CiteRead: Integrating Localized Citation Contexts into Scientific Paper Reading

Napol Rachatasumrit*
Carnegie Mellon University
Pittsburgh, PA, USA
napol@cmu.edu

Jonathan Bragg
Allen Institute for AI
Seattle, WA, USA
jbragg@allenai.org

Amy X. Zhang†
University of Washington
Seattle, WA, USA
axz@cs.uw.edu

Daniel S. Weld
Allen Institute for AI & University of Washington
Seattle, WA, USA
danw@allenai.org

ABSTRACT

When reading a scholarly paper, scientists oftentimes wish to understand how follow-on work has built on or engages with what they are reading. While a paper itself can only discuss prior work, some scientific search engines can provide a list of all subsequent citing papers; unfortunately, they are undifferentiated and disconnected from the contents of the original reference paper. In this work, we introduce a novel paper reading experience that integrates relevant information about follow-on work directly into a paper, allowing readers to learn about newer papers and see how a paper is discussed by its citing papers in the context of the reference paper. We built a tool, called CiteRead, that implements the following three contributions: 1) automated techniques for selecting important citing papers, building on results from a formative study we conducted, 2) an automated process for localizing commentary provided by citing papers to a place in the reference paper, and 3) an interactive experience that allows readers to seamlessly alternate between the reference paper and information from citing papers (e.g., citation sentences), placed in the margins. Based on a user study with 12 scientists, we found that in comparison to having just a list of citing papers and their citation sentences, the use of CiteRead while reading allows for better comprehension and retention of information about follow-on work.

CCS CONCEPTS

• **Human-centered computing** → **Interactive systems and tools**.

KEYWORDS

interactive documents, reading interfaces, scientific papers, citations, citances, annotations

*Work done during an internship at AI2
†Work done while employed by AI2

IUI '22, March 22–25, 2022, Helsinki, Finland

ACM ISBN 978-1-4503-9144-3/22/03.
https://doi.org/10.1145/3490099.3511162

ACM Reference Format:
Napol Rachatasumrit, Jonathan Bragg, Amy X. Zhang, and Daniel S. Weld. 2022. CiteRead: Integrating Localized Citation Contexts into Scientific Paper Reading. In *27th International Conference on Intelligent User Interfaces (IUI '22), March 22–25, 2022, Helsinki, Finland*. ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3490099.3511162

1 INTRODUCTION

Scientific progress involves a cumulative endeavor across many scientists that are building on and discussing each other's work. Through the use of citations, authors of research papers situate their novel contributions in the context of previously published scientific literature. Citations can also be used to trace *forward* in time to understand how a paper relates to work that came afterward. Citing papers can sometimes describe newer work that directly improves on or uses the work described in a reference paper. Citing papers also provide commentary in the form of citation sentences, or "citances" [19], from other scientists about how the work in a reference paper was received by the field or where it is situated in light of contemporary work. This information is valuable when a scientist is trying to understand the state of the art in a field or has encountered a relevant paper while conducting a literature review.

While readers of a paper can easily make use of citations provided by the paper's authors to determine prior work, few tools exist for understanding and exploring the follow-on work that came after a reference paper was published. One set of tools is provided by scientific search engines, such as Google Scholar, that list all the papers that cite a reference paper. However, these lists can be long and are undifferentiated, with little information about how the citing paper makes use of or refers to the reference paper, requiring the reader to dig into each citing paper for details. Other scientific search engines, such as Semantic Scholar and Scite,[1] go further to extract and present the relevant citance from the citing paper and also categorize citances according to how they discuss the reference paper. For instance, Semantic Scholar groups citances by intent (citing as background, method, or result) [5], while Scite characterizes citances as supporting, contrasting, or just mentioning the reference paper's findings. However, the presentation of these citances remains a standalone list, disconnected from the experience of reading the reference paper. Thus, someone reading a passage

[1] https://scite.ai

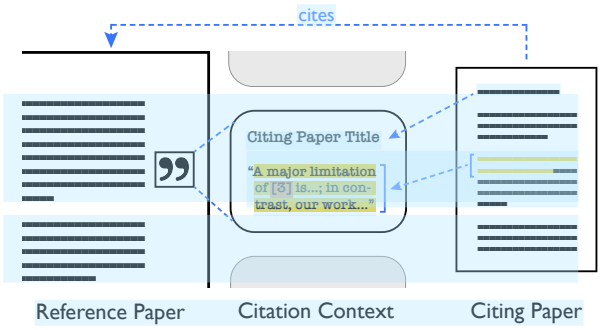


Figure 1: CiteRead provides an augmented reading experience for a reference paper by 1) finding subsequently published citing papers, 2) extracting the citation context—the text that is surrounding citations to the reference paper, 3) selecting a subset of these citation contexts deemed to be interesting, and 4) localizing the citation context in a relevant part of the reference paper, where an interface affordance allows quick perusal as a marginal note.

in a paper—for instance, the presentation of a particular finding—might easily miss a crucial citing paper that has direct relevance to that passage. Even if readers choose to peruse a paper's list of citing papers and their citances, it may be difficult to understand the significance of a citance without the relevant local context from within the reference paper.

In this work, we present a novel paper reading experience, called CiteRead, that integrates important information from citing papers directly alongside relevant text in the reference paper (Fig. 1). We take inspiration from social annotation systems [26], such as Google Docs and Hypothes.is, that enable readers of a document to leave commentary "in the margins" of the document by anchoring their comment to a specific location. This allows readers to seamlessly alternate between reading the document and reading contextually relevant commentary about the document. Much like comments on a document, citing papers also provide commentary from the scientific community on a reference paper through the use of citances and other context around a citation within the citing paper. However, unlike in social annotation systems, authors of citing papers typically do not provide localization information when making a citation. In addition, not all citing papers make sense as an inline annotation within a reference paper, as some citing papers only have an indirect connection to the paper or broadly discuss the paper as a whole.

To determine how to address these issues, we first conducted a formative study with seven scientists to find out what kinds of citations would benefit from localization in a reference paper and what information from citing papers would be useful to readers. We found that participants were interested in localizing citances that discuss topics like results comparison and limitations of the reference paper but were not interested in more generic citances. From these findings, we built CiteRead, which includes an automated pipeline that first selects good candidate citing papers for localization from the full list of citing papers. The pipeline then finds an appropriate location in the reference paper to place the citation context of the citing paper. As some kinds of citation contexts can be localized to a particular number or sentence, while others only discuss the reference paper at the level of an entire section, our localization approach also accounts for how fine-grained to make each placement. Finally, CiteRead takes the output of the pipeline and populates a novel reading interface that integrates citation contexts into the margins of a paper's PDF file. While reading, users can quickly see where in a paper there is relevant commentary from a citing paper and click in to read useful context about the citing paper, including the citance itself, additional surrounding text, and summary information about the citing paper as a whole, such as the title and abstract.

Through a user study with 12 scientists, we evaluated CiteRead's usability and its ability to help readers with comprehension of a paper and its follow-on work. We found that in comparison to reading a paper with a companion list of citances from citing papers, the use of CiteRead for reading allowed for better comprehension and retention of information about follow-on work. From qualitative feedback, we found that participants benefited from both the localization, which they used for finding and re-finding relevant citations, as well as the interactive interface, which reduced the need for context switching.

In summary, we make the following contributions:

- We conducted a formative study, which revealed what types of information from citing papers would be helpful while reading a reference paper, and where to display that information.
- Based on this study and additional pilots, we define the problem of augmenting a PDF reader with localized and contextualized citation contexts to help scholars better understand a reference paper in the context of follow-on work.
- We developed a method to select citation contexts to display in the PDF reader, which fall into the categories of interest identified in the formative study. Our method identified citations of interest to user study participants, using a simple linear combination of features that were derived from formative study insights (e.g., penalize similarity to reference paper abstract) and leveraging state-of-the-art NLP embeddings.
- We also developed a method for *localizing* citation contexts in the reference paper according to user preferences elicited from the formative study. Our approach is the first to localize at the section level (formative study participants frequently preferred it, but prior NLP-focused approaches have localized to a finer granularity of up to five sentences).
- We implemented CiteRead, a scientific paper reading interface that enhances one's reading experience by annotating citation contexts from follow-on work directly in the paper. These affordances allow users to seamlessly alternate between the reference paper and citation contexts.
- We conducted a quantitative and qualitative evaluation of CiteRead, which showed that CiteRead helps users better understand follow-on work and validates the benefits of localization and seamless integration of relevant information.

2 RELATED WORK

2.1 Citations for Exploring and Understanding Research

Citations provide an important way for authors of research papers to characterize how they are contributing to a body of knowledge within a field. As a result, much research has analyzed and applied citation information towards supporting researchers and the research process.

Some researchers have analyzed citation sentences, or "citances," within citing papers [19] as well as the broader context around a citation. Analyzing these citation contexts can help researchers determine what other researchers think about the paper in question and learn about considerations that might not be discussed in the reference paper (e.g., limitations of the approach used). In their analysis, Elkiss et al. [10] found that while many citations overlap to some extent with the abstracts of the papers, some citations focus on aspects of these papers different from the abstracts. In fact, the set of all citances of a paper contain 20% more concepts compared to the abstract of the reference paper [9]. Other researchers note that citances state facts more concisely [21].

In addition, citation contexts as well as citation network information can be used to analyze papers towards specific applications beneficial to researchers, including paper summarization, paper recommendation, automated review paper generation, and finding emerging trends.

While researchers have explored applications of citation information to support determining whether to read a paper or getting the gist of a paper without reading it, prior work has not yet focused on using citation information to enhance the reading experience once a person has decided to read a paper. In this work, we focused on how citations can enhance the reading experience of those who have already opened and are scrolling through the pages of a paper.

2.2 Automated Techniques for Selecting and Localizing Citation Contexts

A body of NLP work has studied the problem of selecting important citations [14, 16, 20, 23]. A small amount of NLP work has studied how to identify spans of text in the reference paper corresponding to the citation sentences in citing papers. In NLP, the goal of this work has been to summarize the reference paper via the citation contexts from citing papers; different from our work, localization is used in service of summarization. For example, Cohan & Goharian [8] evaluate on the TAC 2014 Biomedical Summarization benchmark[2], and show that identifying spans in the reference paper improves citation context-based summarization. The CL-SciSumm shared task [4] has run for several years, and also focuses on localization in service of summarization by citing papers. Unlike our work, this line of work views localization as a subtask for automatic summarization; in contrast, our localization approach seeks to optimize usefulness for immediate human consumption in a novel paper reading interface. Thus, for example, our approach seeks to match visually salient attributes, or localize to the section level, as frequently preferred by formative study participants; in contrast, the NLP approaches match text in the reference paper only up to a finer granularity of five sentences [4].

2.3 Reading Interfaces for Scientific Papers and Digital Documents

A range of interfaces have been developed to aid readers of scientific papers and other types of digital documents. Many systems have explored various augmentation techniques to improve active reading experiences, such as embedding word-scale visualizations [12], layering dynamic content on static images [17], and generating on-demand visualizations [1]. Several scientific paper reading interfaces incorporate social annotation by readers, including Nota Bene [26] and Fermat's Library.[3] In addition, ScholarPhi [15] automatically augments scientific papers with term definitions and equation diagrams, and offers navigation aides within a paper. Different from user-generated annotations or automated definitions and navigation aides, our work seeks to repurpose author-written paper text as a new form of paper annotations. This approach is complementary to these earlier efforts, and can also be seen as a bootstrapping method for social annotation, which could help to attract users to a platform prior to the existence of user-generated annotations.

3 FORMATIVE STUDY

To understand how citations and citances might be beneficial to researchers, we conducted a small formative study. Seven researchers (four doctoral students and three professional researchers) were recruited through mailing lists and Slack channels to participate in participatory design sessions. Three participants identified as NLP researchers, and five participants identified as HCI researchers.

Before the session, we assigned one of two papers for participants to read. The papers were "Generalized Fisheye Views" [11] and "Scientific Article Summarization Using Citation-Context and Article's Discourse Structure" [7]. We chose these two papers as they cover disciplines our participants had expertise in (HCI and NLP, respectively) and are well-cited. Assignment was counterbalanced such that some participants had to read a paper from their specific discipline, while others had to read a paper outside their discipline.

The studies were conducted remotely via Zoom, a video conferencing platform, and Figma, a collaborative design tool. Each session began with a semi-structured interview about the participant's background and reading experiences, with a focus on how they utilized citations and citances. Participants then took part in a 40-minute participatory design session with the experimenter. We provided the paper that they came to the session having already read, together with some selected citing papers, and participants were asked to select citations or citances that they found useful and sketch how they wanted to present them on the reference paper. After the session, they were asked to organize their designs into groups and rate them according to the benefits. We asked participants to think aloud and share their screen during the sessions, and we recorded the audio, video, and screen share.

[2] http://www.nist.gov/tac/2014/BiomedSumm/

[3] https://fermatslibrary.com

3.1 Findings

All participants mentioned that they usually use search engines (e.g., Google Scholar and Semantic Scholar) to look for papers that cite the reference paper. Their stated aim was mostly to find broadly similar but more recent papers (instead of looking for a particular aspect or answer in the papers). Participants usually relied on the title, citation count, venue, and year to filter the list of citing papers. They then used the abstract of the citing paper (and sometimes its introduction) to gauge its relevance. Only one of them reported scanning the citances provided in Semantic Scholar as a way to select relevant papers.

Participants also gave feedback on what kinds of citances they found interesting versus uninteresting:

- Participants expressed that citances were useful to understand how other researchers frame the reference paper. They further used this information to verify that they correctly understood the main point of the reference paper. Citances that discuss result comparisons were noted to be particularly useful.
- Citances that discuss the limitations of the reference paper were also highly appreciated. One participant noted: "Authors [of the reference paper] usually put an emphasis on the novelty but not the other aspects of the work, like limitations ..." Participants also mentioned that they found it useful when a citing paper claims to have a solution to the issue.
- Generic citations were not judged to be particularly interesting or useful. If there was a high information overlap between the citance and the reference paper's abstract, then it was likely to be unhelpful.

Finally, participants gave insights into what information they needed to better understand a citation, including information from the reference paper and from the citing paper:

- Participants considered or looked for specific locations in the reference paper that would be the best place to present a given citance, but when asked, they said a precise location was unnecessary and maybe even a distraction—aligning at the section level often seemed sufficient.
- Several types of information about a citing paper were identified as useful for contextualizing the citance and deciding whether it was important. Particpants wanted an abstract or TLDR summary [3] to understand the gist of the citing paper. They wanted to know when each citing paper was published, with the most recent work first. Publication venue and citation count were also deemed useful as a signal of the credibility of the citing paper. Interestingly, the names of citing authors were not judged useful—participants felt they were unlikely to recognize most names.
- Sometimes the citing *sentence* did not contain the requisite information that the participant needed to understand the citation. Participants declared an interest in the opportunity to see the whole paragraph (or maybe section) of the citing paper, in order to understand it. Sometimes, related figures or tables were also needed.

4 OUR APPROACH

Guided by insights from our formative study, we developed CiteRead, a scientific paper reading interface that enhances the reading experience by annotating citation contexts from follow-on work directly on the paper, thereby allowing users to seamlessly alternate between the reference paper and citation contexts. CiteRead consists of three main components: an automated method for selecting citation contexts, a method for localizing selected citation contexts in the reference paper, and a web-based client.

4.1 Selection Method

Inspired by features proposed by prior work for identifying important citations in scientific papers [16, 20, 23], we selected a similar feature set to use in our selection method. Unlike prior work, we calculated a score of each citance using a linear combination of features and hand-tuned parameters instead of training a binary classifier. While a learned classifier might achieve higher precision, the need for a large training dataset was a barrier. Unfortunately, existing datasets used by prior work were not applicable for us due to differences in objectives. For example, the dataset used by Valenzuela et al. [23] did not consider result comparison as an important citation, but our formative study had shown that researchers found those highly useful. When it came to choice of features, however, we relied heavily on previous machine learning models (e.g., SciBert [2], SPECTER [6], and Cohan's citation intent classifier [5]).

In the end, our selection approach used the following features (we annotate whether the feature is a positive or negative signal):

- **(F1) Total number of direct citations**: The total number of times the citing paper cited the reference paper. (positive)
- **(F2) Small co-citations**: A Boolean feature, set to one if there are fewer than three other citations in the citance. (positive)
- **(F3) Citation appears in table or caption**: A Boolean feature, set to one if at least one citation appears in a table or a caption. (positive)
- **(F4) Sentence length**: An ordinal feature based on the word count in the citance. (positive)
- **(F5) Similarity between papers**: Cosine similarity between the SPECTER embedding of the citing paper and the reference paper. (positive)
- **(F6) Similarity to abstract content**: The highest cosine similarity between the SciBert embedding of the citance and sentences in the reference paper's abstract. (negative)
- **(F7) Age**: An ordinal feature based on the age of citing paper. (negative)
- **(F8) Contains cue words**: A Boolean feature, set to one if the citance contains a signaling keyword, such as 'however' or 'whereas.' (positive)
- **(F9) In generic sections**: A Boolean feature, set to one if the citance is from (or subsection of) generic sections, such as 'Introduction' or 'Background.' (negative)
- **(F10) In important sections**: A Boolean feature, set to one if the citance is from (or subsection of) important sections, such as 'Method' or 'Results.' (positive)

- **(F11) Intent classification**: An ordinal feature based on the citance's intent classification [5], with values set according to observed usefulness of the intent. (positive)

Note that this set of features is informed by our formative study. For example, (F6) reflects the observation that citances that overlap with the reference abstract are uninteresting to researchers. CiteRead computes these features, calculates the weighted sum, and selects the highest 15 citation contexts to present in the interface (we heuristically limited the number of citation contexts to 15 from the observations during pilots).

4.2 Localization Method

Similarly to the selection method, we adopted an approach that leverages heuristics and existing pretrained machine learning models instead of training a new model due to mismatch between available datasets and our new task. We ran each selected citance through our localization method to localize them on the reference paper. Fig. 2 illustrates the procedure. First, if the citance contains a number, we tried to search if the target paper also contains the same number, and localized the citance there if it exists. If not, we tried to find the most relevant section for citances using a combination of category classification and similarities between key terms and section titles. After we found the most relevant section, we searched for the best sentence in the section using a combination of similarities between the citance's key terms and each word in the sentence and a similarity between the citance and the sentence. If none of the sentences in the section passed the threshold, we localized the citance near the section title instead.

4.3 Interface

CiteRead's user interface was implemented as an extension of the ScholarPhi reader [15] using the React framework. Based on the insights from our formative study, we designed CiteRead to emphasize: (1) reducing context switching, (2) providing sufficient details about the citation context, and (3) minimizing intrusiveness. We envision that CiteRead can help researchers in both linear reading experiences and information foraging scenarios. CiteRead's interface consists of three main components:

Annotation. CiteRead annotates selected citances from our selection and localization method in the margin of the reference paper in the reading interface (see Fig. 3a). The icon inside the annotation represents the type of associated citation. For example, a citation that is comparing its own results to the results from the reference paper will be shown as a chart icon. The citance of the citation will be shown as a tooltip when a reader hovers over an annotation, and the dedicated citation sidebar will be opened when a reader clicks on an annotation. When there are multiple citations of the same type located in close proximity, CiteRead aggregates them into a single stacked annotation.

Sidebar. Each selected citation appears as a card in the sidebar grouped by the page in the reference paper it is located in. The citation card design was inspired by our formative study and includes the information that participants deemed useful. A reader can click a citation card to scroll the paper to the annotation, and similarly click an annotation to scroll the sidebar to the associated citation card. To make the card as compact as possible, lengthy text in the card is truncated, but a reader can hover over the truncated text to see the full text in a tooltip. The citation number (or other citation format) is highlighted to help readers quickly identify where in the citance is it talking about the reference paper. If a reader needs more context to understand a citation, they can click the title of a card to navigate to a detailed card.

Detailed Card. The detailed card provides more context to a reader by including sentences or paragraphs near the citance (i.e., sentences with 50 words around the citance). The detailed card also includes a TLDR [3], or an auto-generated short summary, of the citing paper to help readers understand the overview context. In addition to the same information shown in a citation card in the sidebar, a detailed card includes a list of authors and other related metrics, such as reference count and citation velocity. When a reader wants to explore the citing paper further, they can use links in the card to arrive at Semantic Scholar's paper detail page or the paper's PDF.

4.4 Additional Implementation Details

We used VILA [22] to extract text from source PDFs. We used a pretrained SPECTER model [6] to compute embeddings through the Hugging Face Sentence Transformer interface. We used TextRank [18], a graph-based ranking model, to retrieve key phrases. CiteRead was implemented within the React framework as an extension of the open-source ScholarPhi reader [15].

5 USER STUDY

We conducted a lab evaluation of CiteRead to explore the potential utility of integrating localized citation contexts into the reading experience. More specifically, we considered the use case of a scientist who is reading a paper and is interested in learning about relevant follow-on papers, including how those citing papers relate to and provide commentary about the paper they are reading. We conducted a within-subjects experiment where we compared using CiteRead for reading a paper with citation contexts in the margins to a baseline where readers have access to the paper and a companion list of citing papers, representing the standard way readers find out about follow-on work today. We also examined the overall usability of CiteRead while participants were using the tool during the experiment. Finally, after the study, we asked participants to discuss how they could see themselves using the tool in their work. We obtained Internal Review Board (IRB) approval prior to conducting our study.

5.1 Study Design

5.1.1 Participants. In order to reduce potential variation across participants due to differing research topics and expertise, we focused on researchers who are current doctoral graduate students and working in the field of Human-Computer Interaction (HCI). We recruited a total of 12 graduate students as participants, through university and industry-affiliated mailing lists and Slack channels. All participants had experience with reading and writing research papers in HCI as well as understanding of technical topics and terminology in the research papers used for this study. Participants were compensated $30 (USD) via PayPal. Study sessions were one hour long and held remotely over Zoom.

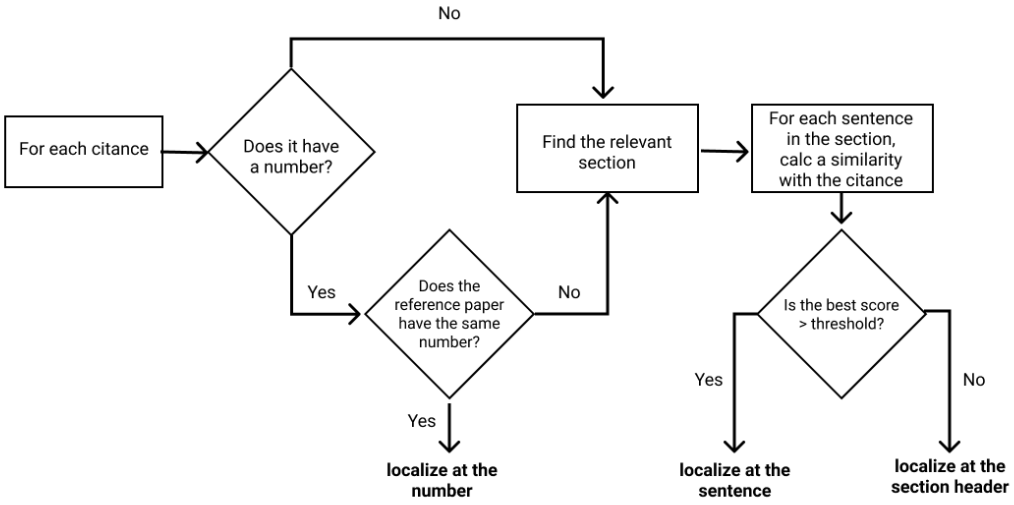


Figure 2: Flow chart illustrating CiteRead's localization procedure.

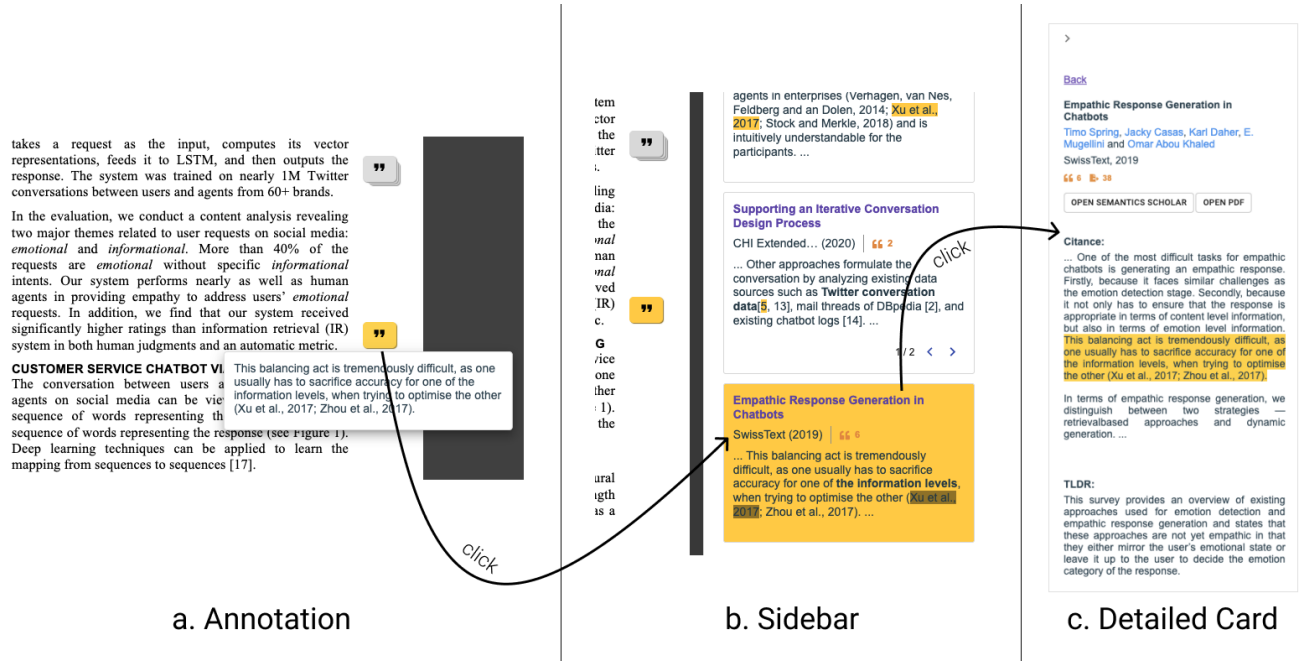


Figure 3: The CiteRead interface. (a) Annotated quote marks (") in the reference paper margins correspond to localized citation contexts, which on hover provide the citation sentence (citance); (b) the sidebar shows a corresponding card with additional citation context details, such as the citing paper title (hovering on the annotated quote highlights the card and clicking scrolls to the card); and (c) clicking on the card opens additional context in the sidebar, including an automatically generated paper TLDR.

5.1.2 Experiment Conditions. As our study was within-subjects, each participant completed both of the following two conditions, with order of the conditions counterbalanced:

- CiteRead **Condition**: Participants were presented with the CiteRead reader interface (Fig. 3) loaded with a pre-selected research paper in PDF format. We ran all of its citing papers through the selection and localization steps of the CiteRead

pipeline so that the interface contained citation contexts of a subset of all the citing papers in the margins.

- **Baseline Condition**: Participants were presented with a pre-selected research paper in PDF format in the web browser's default PDF viewer. As shown in Figure 4, participants also received a companion interface containing a list of citing papers that mimics how one would explore follow-on work today. This list view, modeled after the citations section of Semantic Scholar's paper details page, has a card for each paper, containing its title, list of authors, conference name, year, abstract, citation count, and a list of citances categorized by intent. At the top of the interface, there is a search bar with a dropdown menu to filter results by their intent categorization, as well as a dropdown menu to sort results by relevance, citation count, and recency. Much like CiteRead, participants can click to view a citing paper's Semantic Scholar page and PDF file. To have a fairer comparison to CiteRead, as the full list of citing papers can be long, we reduced the list of citing papers to only the ones that passed the selection step in the CiteRead pipeline. Thus, the number of citation contexts that a reader has access to is equivalent in both conditions.

5.1.3 Procedure. Within each condition, we had each participant go through the following procedure:

(1) We began with a quick tutorial of the respective interface. Participants were then given a few minutes to explore and ask any questions about the interface.
(2) Participants were then given 15 minutes to read their assigned paper and explore any of the presented citing papers and their citation contexts. For the last 10 minutes of this time, participants also had access to three questions that asked about how follow-on work related to the assigned paper. We did not provide participants with any multiple-choice answer options at this point.
(3) After the 15-minute period, participants had 5 minutes to answer the three questions, this time with multiple-choice options available. Participants had to attempt to answer the questions at least once without the use of the interface or any of the reading materials. After the first pass, participants were then allowed to use the interface.
(4) Finally, participants completed a NASA-TLX questionnaire [13] on task load.

After the conclusion of both conditions, participants filled out a System Usability Scale questionnaire (SUS) regarding the CiteRead condition and also participated in a semi-structured interview discussing their experience in the study with both conditions and how they could see themselves using a tool like CiteRead. We recorded participants' screen share during the task, and we also automatically logged participant interactions within the tools.

We chose to split up the 15 minutes that participants had to read the materials into 5 minutes without access to questions and 10 minutes with access to questions in order to strike a balance between linear reading behavior and skimming in search of answers. From our formative study, we expect that both of these reading strategies would be used in real settings, where sometimes readers are carefully reading in a linear fashion, while other times, they are in search of information within a particular section and skipping around.

Table 1: Number of citances that were localized at different levels of granularity for the reading materials.

	Header	Sentence	Number
Paper 1	6	9	0
Paper 2	4	10	1

5.1.4 Reading Materials. As all our participants were HCI researchers, we chose two papers for participants to read that are in the field of HCI. Participants first read "A New Chatbot for Customer Service on Social Media" [25] (Paper 1), a paper published in ACM CHI 2017 with 371 citations, followed by "Serendipity: Finger Gesture Recognition using an Off-the-Shelf Smartwatch" [24] (Paper 2), a paper published in ACM CHI 2016 with 131 citations. We selected these two papers according to the following criteria: (1) published in ACM CHI, the top conference in HCI, within the last decade, (2) have over 100 citations, and (3) are relatively short papers, due to time constraints for our study. Each paper also covers a different subfield in HCI (conversational agents and wearable interfaces, respectively), allowing us to examine the applicability of CiteRead on two papers with differing structure, and also reducing the potential learning effect from participants reading one paper followed by the other. Table 1 shows the number of citances that were localized at different levels of granularity for each paper.

5.1.5 Test Questions. We wrote questions according to insights gained from our formative study. Specifically, we created three questions for each paper that asked about (1) limitations of the approach in the reference paper according to follow-on work, (2) other approaches tried by follow-on work, and (3) comparisons of the results of follow-on work to the results in the reference paper. These three questions were the most-frequently mentioned questions that participants in our formative study said they had when exploring the follow-on work of a paper.

Each question had six answer options, with the ability to select all options that apply. We wrote the answer options such that each question had three correct answer options, with the remaining three incorrect options serving as distractors. The three correct answer options were written according to the following categories:

- **Local**: This answer can be determined from reading the paper without examining any of the citation contexts.
- **Obvious**: This answer can be determined by reading the citance of one of the citing papers shown to the participant without any further investigation into the citing paper.
- **Hidden**: This answer can only be determined by investigating more of the context surrounding the citance of a citing paper shown to the participant, such as the adjacent paragraphs or the abstracts of the citing paper.

We chose to vary where the participants had to look to determine the answer in this way, in order to assess how well our interface allows a user to go from the reference paper to a localized citance to additional context around the citance in search of the answer.

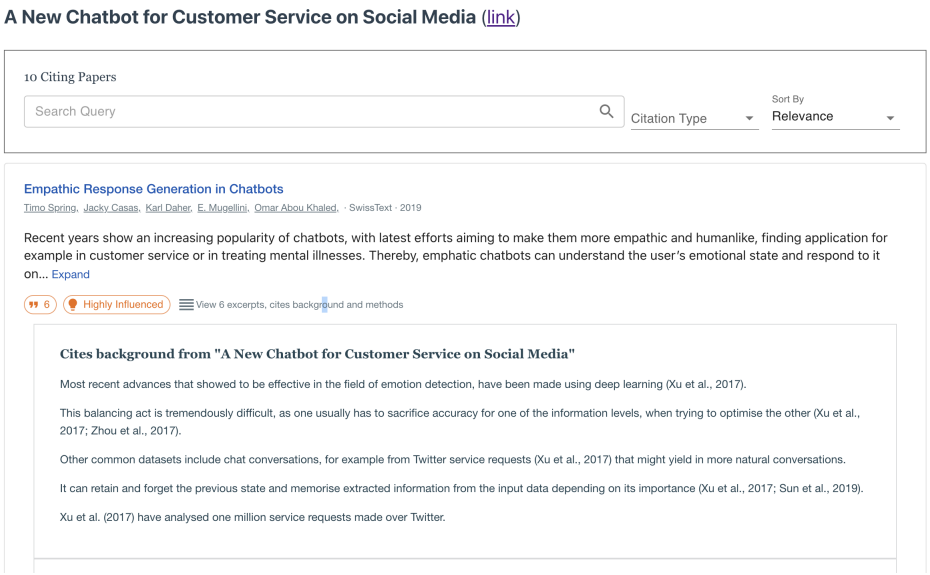


Figure 4: The baseline interface, modeled on the Semantic Scholar citation browsing interface. The interface provides search, faceted browsing of citation types, and sorting by relevance and other criteria. Citations are shown along with additional supplementary information such as citation contexts, intents, and whether they were "highly influenced" by the reference paper.

We include the full set of questions and multiple-choice answer options for each paper in Appendix A.

5.2 Quantitative Experiment Results

Quantitative results from the lab experiment were strongly positive. As shown in Figure 5a, participants using CiteRead scored significantly higher on the comprehension test (M=62.5, SD=24.2) compared to Baseline (M=43.3, SD=13.7), based on a two-tailed paired t-test ($p = 0.029$). There were 3 out of 12 participants who achieved a higher score with the baseline interface, but these participants also performed worse on average compared to the participants who achieved a better score with CiteRead (38% vs 64%). Additionally, participants reported lower cognitive load for all NASA-TLX criteria, as shown in Table 2 (2.05 points lower on average, range of 1–20), and this difference was significant for the Performance workload dimension (10.00 vs. 12.85, $p = 0.039$; lower value indicates increased sense of success).

Participants rated the system high in terms of usability, with average and median SUS scores of 78.54 and 80.0, respectively. Figure 6 illustrates the distribution of the participants' answers, which was positive for all questions. Generally, all participants successfully used CiteRead to explore the selected citations and navigate around the interface with ease. We observed a minor disruption in the system where the sidebar did not scroll to the associated card when participants clicked on the same annotation twice in a row, but most participants quickly noticed the issue and continued smoothly with the session.

More detailed analysis of test results by category (Figure 5b) shows that participants using CiteRead achieved higher test scores not just overall, but also on each sub-category of questions. Unexpectedly, participants using CiteRead even scored higher on the Local questions (78% vs 63%), whose answers can be found in the reference paper text itself without the use of citation contexts. Unlike the overall accuracy comparison, sub-category accuracy differences were not statistically significant.

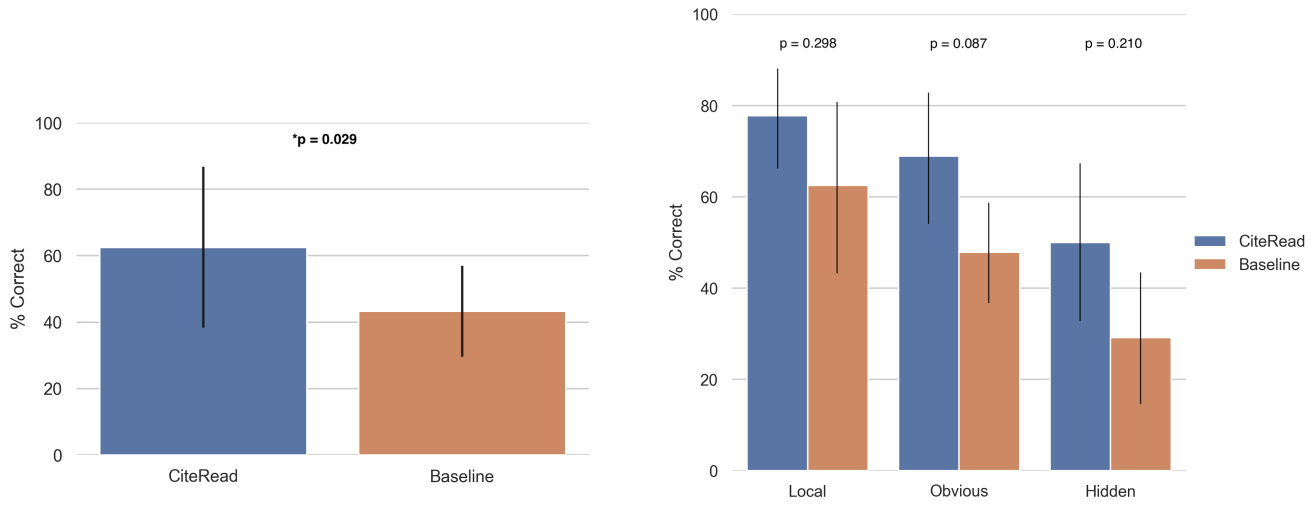
Table 2: Mean cognitive load (NASA-TLX) scores by interface, with standard deviation shown as uncertainty. Lower values indicate lower cognitive load (better). The range of possible values is 1–20. The asterisk indicates 0.05-significance.

	CiteRead	Baseline	p
Mental	14.52 ± 4.12	16.43 ± 3.25	0.244
Temporal	13.33 ± 4.77	15.24 ± 4.28	0.305
Performance	10.00 ± 3.76	12.86 ± 3.55	*0.039
Effort	12.86 ± 4.64	14.05 ± 4.79	0.558
Frustration	7.38 ± 4.30	9.76 ± 5.64	0.193

Overall performance aggregated across both conditions did not meaningfully differ between the two papers; average test scores were 53.37 ± 20.26 and 52.52 ± 23.81 for Papers 1 and 2, respectively ($p = 0.921$). We note, however, that participants reported higher cognitive load for Paper 2 for all NASA-TLX criteria, and this difference is significant for the Temporal criterion (5.67 vs 4.33, $p = 0.047$), indicating that participants felt increased time pressure.

5.3 Qualitative Findings

From the responses to the semi-structured interview questions asked after the completion of the experiment, we collected quotes



(a) Overall Accuracy **(b) Accuracy by Question Type**

Figure 5: Mean test accuracy for two interfaces, with standard deviation shown using uncertainty bars. (a) Overall, participants using CiteRead scored significantly higher than baseline ($p = 0.029$). (b) Participants using CiteRead also scored higher on each individual category of questions, though the differences were not significant at the 0.05 level.

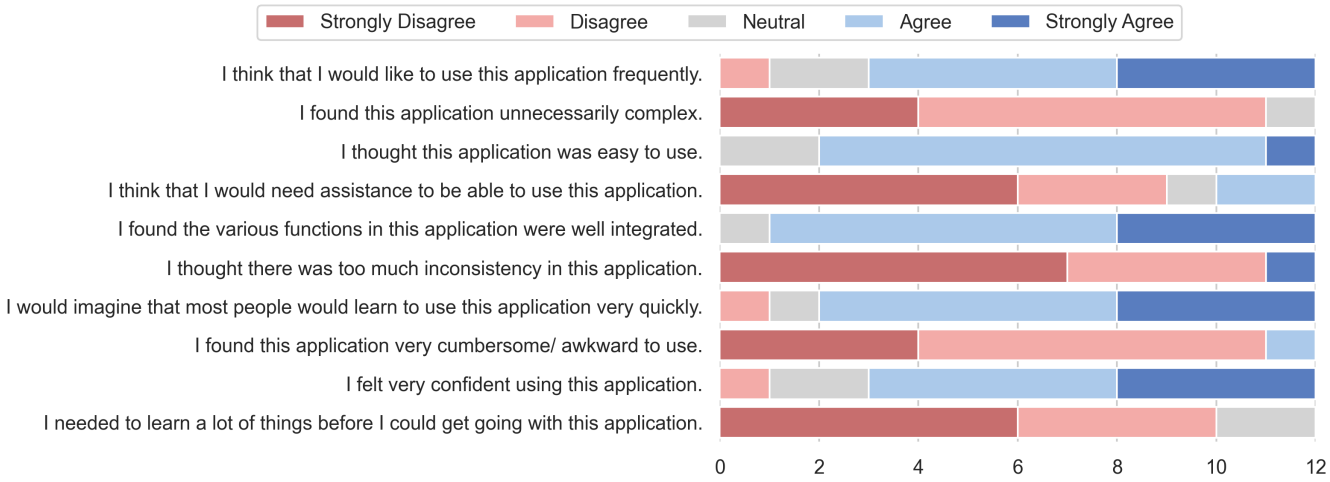


Figure 6: Distribution of System Usability Scale for CiteRead, showing that participants overwhelmingly agreed with favorable statements and disagreed with unfavorable ones.

that discussed either the localization of citation contexts in CiteRead or the interactive experience provided by the CiteRead interface, and we summarize participants' responses below.

5.3.1 Localization. Overall, participants described benefiting from the localization of citation contexts in CiteRead in a number of ways. One way that localization helped was to make it easier to find relevant citation contexts. For instance, Participant 12 (P12) would go to a particular section that seemed relevant to answer a question and then look for annotations in that area:

> "Having the location-based annotation, which I found fairly accurate, is really helpful for answering specific questions. Like, for different types of sensors, I figured in Related Work they were probably going to talk about past work and types of sensors, and there was a citation cluster around there, I thought that might be the place where the information about other types of sensors might be." –P12

In this way, for participants seeking answers to a question about follow-on work—for instance, results that build on the results in the reference paper—the section headings of the reference paper

can act as an index into the relevant citation contexts. We also observed some participants searching for particular keywords in the reference paper and then looking at the annotations surrounding the keyword matches. On the other hand, a few participants expressed their confusion when the localization was not accurate or a connection between a citance and the corresponding location in the reference paper was not obvious:

> "Why is this placed here? …I don't know why I'm seeing what I'm seeing." –P2

In this case, participants could suffer from the additional cognitive load associated with trying to think of plausible connections. Other participants described how localization helped them re-find content by providing them with a spatial mental map that they could use to more easily remember the location:

> "It was easier when I had to go back and find the information. I can think about where it was, like spatially in the paper. It's very analogous to how you would mark up a piece of paper. Even when you see it once...you might not process it the first time, but later you might be reminded of that one thing you sort of read before, and it was in this corner of the page." –P4

5.3.2 Interactive Interface. When it came to the interface, participants expressed that CiteRead allowed them to stay within the context of the reference paper throughout the task, while in the baseline condition, they had to context switch frequently due to navigating between multiple tabs. For instance, the following participants said:

> "It was easier not having to jump between multiple tabs [in the CiteRead condition]. Having them all in one page is really easy to navigate." –P1

> "It could be cognitively very disruptive [in the baseline condition]. When I jumped back to the paper, I have to think again about what I am looking for." –P10

Many participants also liked how the CiteRead interface allowed for progressive disclosure, where they could dive deeper if they wanted to get more information. By having different presentations including the tooltip on hover on the PDF annotation, the sidebar, the more detailed card, and all the way to the full citing paper in a new window, this ensured that participants would only see as much information as they wanted in that moment and not more. One participant said:

> "I really liked how concise [the interface] was. Like how you can see almost all of them with some information, then click it once to access extra information." –P2

Other participants such as P6 also mentioned that the ability to access additional context around a citance as well as the paper abstract and TLDR summary helped them be more confident in their understanding of what the citance was about. Some participants suggested additional information that could be included in the citation context interface, such as the section where the citance came from. However, many participants mentioned that the stacked annotations were quite confusing especially when they were in close proximity but not highly related:

> "Stacked UI could be confusing. It is not clear why things are anchored in the same location." –P3

6 DISCUSSION

Our study results overall provide evidence of the positive benefit of localizing citation contexts of citing papers directly into the reading experience of a reference paper. Not only does CiteRead make it easier for readers to find relevant follow-on work, the use of annotations in the margins of a reader application allows for seamless exploration of a reference paper along with follow-on work. From both the user study as well as the formative study, we found that participants expressed a desire to answer the kinds of questions posed in the user study around how follow-on work discusses or builds on a reference paper. For instance, one participant in the user study said:

> "I definitely had those moments of wondering about these types of questions, but going through the list in Google Scholar with only abstracts does not help with the specific questions." –P9

Unfortunately, typical methods for answering these questions today are extremely cumbersome, involving manually scanning through a long list of potentially hundreds of papers, finding the citation contexts, and then interpreting them in light of the parts of the reference paper they discuss. Even our baseline condition in the user study is a step up from typical methods as we used our selection method to reduce the list to a more relevant and manageable set. For instance, when asked how long it would take to answer the questions posed in our user study using existing tools, one participant said:

> "It could be very, very long, especially if it is a domain that I am not familiar with. I might spend 3–5 days to do a lit review to understand these aspects." –P1

Another participant expressed that the task of answering questions about follow-on work is so difficult without appropriate tools that it may not even make sense to try:

> "They might not ask these types of questions because they do not have the tools to answer...The kind of questions I can ask will also change depending on the tools accessible to me." –P8

Thus, while we note that participants in our study found even the CiteRead condition to be somewhat mentally taxing and were not completely successful in answering the questions, we are bringing the effort associated with a nearly impossible, potentially multi-day task down to an approachable level. Through our pipeline of selection, localization, and integration into the reading experience, we can significantly reduce the manual effort required for researchers to tackle these important questions.

6.1 Design Considerations and Implications

6.1.1 Selection versus Scale. We devised a selection process that significantly whittled down the number of citing papers to provide as annotations within a reference paper through artificially limiting to 15. This had the benefit of leading to a reading experience that was not overwhelming due to a high volume of annotations in the margins. While our sidebar interface provides the capability for

additional scaling through the stacking of citation context cards, with the ability to then paginate through a stack, a large stack of citation contexts may be difficult to explore effectively. However, while a paper may have many hundreds or thousands of citations, it is likely that most of them would not be useful as a localized annotation within a reading experience, as many cite the reference paper as background information or discuss the paper as a whole. For instance, in testing our selection method, we found that most citances that cite a reference paper simply provide a high level summary of the paper, which our formative study confirmed as unhelpful. Such citations are less directly relevant to a reader of a reference paper compared to a citation that refutes a key finding or uses an alternative method to achieve better results. Thus, between being more selective versus incorporating more annotations, we argue that it would be more important to be selective and achieve high precision when elevating a citing paper to an annotation in the margins for a reader. Additionally, a sorting mechanism would become more critical as the number of selected citances grows, to help readers sift through them in a more efficient way. One participant suggested that:

> "Some kind of cue for you to know which one to focus on or some metrics that these annotations are more important would be really helpful" –P5

As a paper ages, it accumulates more and more citations, particularly if it is regarded as a foundational piece of work. It is possible that these kinds of papers might benefit from a different approach than the one we have outlined. For instance, it might be useful to go further up the citation chain in order to allow readers to more quickly find more contemporary relevant literature. This is because the latest work may not be directly comparing against an older paper (though it might cite it as background) but instead comparing against other contemporary papers that represent the state of the art.

6.1.2 Localization. Some of the participants in our user study discussed how sometimes it was not immediately clear why an annotation was placed where it was or what it was anchored to. We also found that at times we disagreed on or were unsure of where a particular citation context should be localized, as the reference paper touched upon the relevant point in multiple places. In addition, our localization method allows for the granularity of the localization to vary, where sometimes an annotation is linked to a section as a whole as opposed to a particular sentence. In CiteRead's current interface, each annotation occupies a singular point in the margins, and annotations are not differentiated from each other. An alternative approach could be to more clearly indicate what an annotation is anchored to. For instance, much like in Google Docs, the system could highlight the relevant anchor sentence or section header upon hovering over an annotation. To address the issue of multiple potential places for an annotation, annotations could potentially reside in more than one location. However, one drawback with both of these approaches is the risk that they could detract from the reading experience by introducing more clutter into the page.

6.1.3 Improving the Reading Experience. By placing citation contexts alongside the reference paper, we saw how readers were able to use the reference paper as an index or mental map in order to quickly locate relevant citations contexts. One interesting additional observation that we made was that the presence of annotations in the margins of the reference paper helped to also structure how participants read so that they focused on different parts of the paper itself. We note that users in the CiteRead condition performed better even on finding local answers that were directly in the paper and did not require reading any of the citation contexts. It turned out that the answers for those local options could be found in close proximity to the annotations, though we did not purposefully intend to collect answers from those places. Our hypothesis is that there is a reverse effect of the localization where, in addition to helping users understand citances better, the annotations also attract users' attention to the parts of the reference paper discussed by other papers, which might signify the importance of that portion of the paper. This could have potential applications for interfaces or techniques to support skimming papers.

7 LIMITATIONS AND SOCIETAL IMPACTS

Our evaluation was limited in several ways. First, our controlled lab experiment evaluated the benefits of CiteRead for a specific slice of scientists (HCI researchers) on a small set of papers. While our results are positive, more studies are needed to determine the generality of our findings. More extensive evaluation on a large range of papers is also required to determine the general effectiveness of our automated selection and localization techniques. Our end-to-end system evaluation provided support for its effectiveness in terms of strong quantitative and qualitative results corroborating the value of these features, but evaluations of each individual technique on many papers would provide additional evidence. We also note the need for longitudinal studies to better understand how tools like CiteRead are used outside of laboratory settings. In this work, we used existing extracted data from Semantic Scholar for selection and localization, which were pre-processed before study sessions. However, in the real world, it is impossible to exhaustively pre-process all publications, and readers are likely to encounter documents that have not been processed. Moreover, this processing includes parsing a document to get both layout and content information, the quality of which may depend on the format of the accessed paper (e.g., images, PDFs, and TeX files). A potential alternative to pre-processing is to process the document in real-time, but further investigation into its feasibility would be necessary.

We also note several potential risks associated with our approach. As with any discovery tool that redirects user attention, we must be cognizant of possible harms this may produce. For example, drawing attention to a subset of citing papers risks unfairly promoting some follow-on work over others. Adding annotations to a paper may divert attention away from the paper. CiteRead's choice of localization and citation context to show may also misrepresent the citing paper author's initial intent. We believe CiteRead's benefits merit these risks, and note several mitigating factors. While CiteRead does select a subset of follow-on work to supply, we believe that lowering the barrier to finding follow-on work while reading will result overall in increased discovery. With respect to diverting attention away from the reference paper, we are encouraged by the lower cognitive load scores reported in the user study,

as well as the higher comprehension scores for information located in the reference paper. Lastly, we believe the additional discovery benefit to the authors of the citing paper in general outweighs potential misrepresentation risk.

8 CONCLUSION AND FUTURE WORK

In this work, we present a system, CiteRead, that integrates information from follow-on work directly in the scientific paper reading experience. Through a formative study, we discovered what types of information from citing papers scientists are interested in consuming while reading. Based on these findings, we developed novel techniques to select and localize citation contexts in ways that support these discovered information needs. CiteRead provides a seamless interface for alternating between reading the paper and commentary from follow-on work. Our quantitative and qualitative evaluation of CiteRead demonstrates the benefits of this approach for understanding follow-on work, while reducing cognitive load. Finally, we synthesize a set of design considerations and implications for future tools that integrate commentary into the reading experience, through automated selection and localization.

We intend ultimately to deploy our system on a broad range of papers. Towards this end, we plan to improve the robustness of our citation context selection and localization approaches by collecting a new annotated dataset and tuning our models using supervised learning. We also plan to conduct field studies with a wider range of scientists, and investigate the potential engagement benefits of CiteRead. Finally, we are also excited to study how to integrate user commentary with automatically-localized author commentary.

While we have tested our approach in the scientific literature domain, we are excited about the possibilities of localized discussion to augment reading experiences across a range of domains. For example, news articles are often accompanied by discussion threads. However, these discussion threads are not localized to relevant sections in the article; instead, they are typically located below the article, in a sidebar (e.g., New York Times), or on a separate web page altogether (e.g., Hacker News[4]). We envision the application of integrated reading experiences like CiteRead to enable seamless switching between articles and automatically-localized commentary for a broad range of domains.

ACKNOWLEDGMENTS

We thank Marti Hearst, Andrew Head, Dongyeop Kang, Arman Cohan, Kyle Lo, Matt Latzke, Shannon Shen, Tal August, and Raymond Fok for helpful early discussions, as well as the anonymous reviewers for useful feedback on the manuscript. We also thank the researchers who participated in our user studies and assisted with piloting the system.

REFERENCES

[1] Sriram Karthik Badam, Zhicheng Liu, and Niklas Elmqvist. 2018. Elastic documents: Coupling text and tables through contextual visualizations for enhanced document reading. *IEEE transactions on visualization and computer graphics* 25, 1 (2018), 661–671.

[2] Iz Beltagy, Kyle Lo, and Arman Cohan. 2019. SciBERT: A Pretrained Language Model for Scientific Text. In *In Proceedings of the Conference on Empirical Methods in Natural Language Processing*.

[3] Isabel Cachola, Kyle Lo, Arman Cohan, and Daniel S. Weld. 2020. TLDR: Extreme Summarization of Scientific Documents. In *Findings of the Association for Computational Linguistics: EMNLP 2020*.

[4] Muthu Kumar Chandrasekaran, Guy Feigenblat, Eduard H. Hovy, Abhilasha Ravichander, Michal Shmueli-Scheuer, and Anita de Waard. 2020. Overview and Insights from the Shared Tasks at Scholarly Document Processing 2020: CL-SciSumm, LaySumm and LongSumm. In *Proceedings of the First Workshop on Scholarly Document Processing*.

[5] Arman Cohan, Waleed Ammar, Madeleine van Zuylen, and Field Cady. 2019. Structural Scaffolds for Citation Intent Classification in Scientific Publications. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics*.

[6] Arman Cohan, Sergey Feldman, Iz Beltagy, Doug Downey, and Daniel S. Weld. 2020. SPECTER: Document-level Representation Learning using Citation-informed Transformers. arXiv:2004.07180 [cs.CL]

[7] Arman Cohan and Nazli Goharian. 2015. Scientific Article Summarization Using Citation-Context and Article's Discourse Structure. In *In Proceedings of the Conference on Empirical Methods in Natural Language Processing*.

[8] Arman Cohan and Nazli Goharian. 2017. Contextualizing citations for scientific summarization using word embeddings and domain knowledge. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1133–1136.

[9] Anna Divoli, Preslav Nakov, and Marti A. Hearst. 2012. Do Peers See More in a Paper Than Its Authors? *Advances in Bioinformatics* 2012 (2012).

[10] Aaron Elkiss, Siwei Shen, Anthony Fader, Günes Erkan, David J. States, and Dragomir R. Radev. 2008. Blind men and elephants: What do citation summaries tell us about a research article? *Journal of the Association for Information Science and Technology* 59 (2008), 51–62.

[11] George W. Furnas. 1986. Generalized fisheye views. In *Proceedings of the 1986 CHI Conference on Human Factors in Computing Systems*.

[12] Pascal Goffin, Tanja Blascheck, Petra Isenberg, and Wesley Willett. 2020. Interaction techniques for visual exploration using embedded word-scale visualizations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.

[13] S. G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in psychology* 52 (1988), 139–183.

[14] Saeed-Ul Hassan, Anam Akram, and Peter Haddawy. 2017. Identifying Important Citations Using Contextual Information from Full Text. *2017 ACM/IEEE Joint Conference on Digital Libraries (JCDL)* (2017), 1–8.

[15] Andrew Head, Kyle Lo, Dongyeop Kang, Raymond Fok, Sam Skjonsberg, Daniel S. Weld, and Marti A. Hearst. 2021. Augmenting Scientific Papers with Just-in-Time, Position-Sensitive Definitions of Terms and Symbols. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*.

[16] Petr Knoth, Phil Gooch, and Kris Jack. 2017. What Others Say About This Work? Scalable Extraction of Citation Contexts from Research Papers. In *International Conference on Theory and Practice of Digital Libraries*.

[17] Damien Masson, Sylvain Malacria, Edward Lank, and Géry Casiez. 2020. Chameleon: Bringing Interactivity to Static Digital Documents. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.

[18] Rada Mihalcea and Paul Tarau. 2004. TextRank: Bringing Order into Text. In *In Proceedings of the Conference on Empirical Methods in Natural Language Processing*.

[19] Preslav Nakov, Ariel S. Schwartz, and Marti A. Hearst. 2004. Citances: Citation Sentences for Semantic Analysis of Bioscience Text. In *SIGIR'04 workshop on Search and Discovery in Bioinformatics*.

[20] David Pride and Petr Knoth. 2017. Incidental or Influential? - Challenges in Automatically Detecting Citation Importance Using Publication Full Texts. In *International Conference on Theory and Practice of Digital Libraries*.

[21] Ariel S. Schwartz and Marti A. Hearst. 2006. Summarizing Key Concepts using Citation Sentences. In *Proceedings of the Workshop on Linking Natural Language Processing and Biology: Towards Deeper Biological Literature Analysis*.

[22] Zejiang Shen, Kyle Lo, Lucy Lu Wang, Bailey Kuehl, Daniel S. Weld, and Doug Downey. 2022. VILA: Improving Structured Content Extraction from Scientific PDFs Using Visual Layout Groups. arXiv:2106.00676 [cs.CL]

[23] Marco Valenzuela, Vu A. Ha, and Oren Etzioni. 2015. Identifying Meaningful Citations. In *AAAI Workshop: Scholarly Big Data*.

[24] Hongyi Wen, Julian Ramos Rojas, and Anind K. Dey. 2016. Serendipity: Finger Gesture Recognition using an Off-the-Shelf Smartwatch. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*.

[25] Anbang Xu, Zhe Liu, Yufan Guo, Vibha Sinha, and Rama Akkiraju. 2017. A New Chatbot for Customer Service on Social Media. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*.

[26] Sacha Zyto, David Karger, Mark Ackerman, and Sanjoy Mahajan. 2012. Successful classroom deployment of a social document annotation system. In *Proceedings of the 2012 CHI Conference on Human Factors in Computing Systems*. 1883–1892.

[4] https://news.ycombinator.com/

A TEST QUESTIONS

A.1 Paper 1: "A New Chatbot for Customer Service on Social Media"

Q1: How do alternative approaches compare against the approach used in this paper? Select one or more statements that were made, either in this paper or in a citing paper.

- GRU models outperform the current work's (baseline) LSTM model with respect to BLEU score
- Convolutional Neural Network yields similar result to the current work's LSTM model for BLEU score
- Bidirectional Long Short-term Memory encoder with attention-based architecture gets better results compared to plain LSTM encoder-decoder used in this paper
- Human responses still outperform generated responses on appropriateness in this paper, but the responses generated by a tone-aware chatbot are perceived as appropriate as the responses by human agents
- There was no statistically significant difference between this paper's approach and human agents on empathy for emotional requests
- Chatbots based on GRU models have shown better evaluation results than human's responses on attentiveness.

Q2: In contrast to the dataset used in this paper to train chatbots, what have other papers tried to use as datasets instead? Select one or more answers.

- Twitter conversations
- Mail threads of DBpedia
- Extracted data from mobile apps
- Ubuntu dialogue corpus
- The NPS Chat corpus

Q3: In contrast to the factors used in this paper for human evaluation, what factors do other papers use to do human evaluation of chatbots? Select one or more answers.

- Humor
- Helpfulness
- Flexibility
- Attentiveness
- Perceived humanness

A.2 Paper 2: "Serendipity: Finger Gesture Recognition using an Off-the-Shelf Smartwatch"

Q1: In contrast to the sensors used in this paper, what other types of sensors have been used in gesture recognition systems? Select one or more answers, and specify whether the the sensors are in this paper or in a citing paper.

- PPG
- Sonar Sensors
- EKG
- EMG
- Electrical Impedance Tomography

Q2: What are the limitations of this paper's approach? Select one or more statements that were made, either in this paper or in a citing paper.

- While the average accuracy is high (87%), it might be impractical as a commercial solution to require users to provide a lot of instances upfront to train the model.
- This paper only explored the feasibility of using motion sensors when the user is not moving.
- The gesture set is fixed with limited gestural vocabulary and not easily modifiable.
- Leveraging around-device position information from acoustic processing techniques could induce noisy feedback that interferes with gesture detection.
- Accelerometers can only be used for coarse hand gestures. Training data is extremely difficult to collect.

Q3: Which of these statements have been made about evaluation of gesture recognition? Select one or more statements that were made, either in this paper or in a citing paper.

- Gesture sets vary extensively between systems, so it is challenging to compare their results.
- It is very common to use multiple different gesture sets in evaluation.
- From the perspective of new system design and evaluation, it is difficult to evaluate what does the new system add in terms of class of gesture.
- Besides the set of five gestures evaluated in the current work, some works also examine a bending gesture.
- False negative is more important than false positive in gesture recognition evaluation