The Semantic Reader Project: Augmenting Scholarly Documents through AI-Powered Interactive Reading Interfaces*

Kyle Lo[α] Joseph Chee Chang[α] Andrew Head[ψ] Jonathan Bragg[α] Amy X. Zhang[ω] Cassidy Trier[α]
Chloe Anastasiades[α] Tal August[α] Russell Authur[α] Danielle Bragg[ρ] Erin Bransom[α]
Isabel Cachola[ι] Stefan Candra[α] Yoganand Chandrasekhar[α] Yen-Sung Chen[α]
Evie Yu-Yen Cheng[α] Yvonne Chou[α] Doug Downey[α] Rob Evans[α] Raymond Fok[ω]
Fangzhou Hu[α] Regan Huff[α] Dongyeop Kang[υ] Tae Soo Kim[κ] Rodney Kinney[α]
Aniket Kittur[χ] Hyeonsu Kang[χ] Egor Klevak[α] Bailey Kuehl[α] Michael Langan[α]
Matt Latzke[α] Jaron Lochner[α] Kelsey MacMillan[α] Eric Marsh[α] Tyler Murray[α]
Aakanksha Naik[α] Ngoc-Uyen Nguyen[α] Srishti Palani[σ] Soya Park[τ] Caroline Paulic[α]
Napol Rachatasumrit[χ] Smita Rao[α] Paul Sayre[α] Zejiang Shen[τ] Pao Siangliulue[α]
Luca Soldaini[α] Huy Tran[α] Madeleine van Zuylen[α] Lucy Lu Wang[ω]
Christopher Wilhelm[α] Caroline Wu[α] Jiangjiang Yang[α] Angele Zamarron[α]

Marti A. Hearst[β] Daniel S. Weld[α]

[α]Allen Institute for AI [ω]University of Washington [β]University of California, Berkeley
[ψ]University of Pennsylvania [χ]Carnegie Mellon University [τ]Massachusetts Institute of Technology [κ]KAIST
[ι]Johns Hopkins University [υ]University of Minnesota [σ]University of California, San Diego [ρ]Microsoft Research

{kylel,josephc,jbragg,cassidyt,danw}@allenai.org,head@seas.penn.edu,axz@cs.uw.edu,hearst@berkeley.edu

ABSTRACT

Scholarly publications are key to the transfer of knowledge from scholars to others. However, research papers are information-dense, and as the volume of the scientific literature grows, the need for new technology to support the reading process grows. In contrast to the process of *finding* papers, which has been transformed by Internet technology, the experience of *reading* research papers has changed little in decades. The PDF format for sharing papers is widely used due to its portability, but it has significant downsides including: static content, poor accessibility for low-vision readers, and difficulty reading on mobile devices. This paper explores the question "Can recent advances in AI and HCI power intelligent, interactive, and accessible reading interfaces—even for legacy PDFs?" We describe the Semantic Reader Project, a collaborative effort across multiple institutions to explore automatic creation of dynamic reading interfaces for research papers. Through this project, we've developed ten research prototype interfaces and conducted usability studies with 300+ participants and real-world users showing improved reading experiences for scholars. We've also released a production research paper reader that will incorporate novel features as they mature. We structure this paper around challenges scholars and the public face when reading research papers—discovery, efficiency, comprehension, synthesis, and accessibility—and present an overview of our progress and remaining open challenges.

1 INTRODUCTION

The exponential growth of scientific publication [7, 8] and increasing interdisciplinary nature of scientific progress [71, 94] makes it increasingly hard for scholars to keep up with the latest developments. Academic search engines, such as Google Scholar and Semantic Scholar help scholars discover research papers. Automated summarization for research papers [11] helps scholars triage between research papers. But when it comes to actually *reading* research papers, the process, based on a static PDF format, has remained largely unchanged for many decades. This is a problem because digesting technical research papers is difficult [2, 5].

In contrast, interactive and personalized documents have seen significant adoption in domains outside of academic research. For example, news websites such as the *New York Times* often present interactive articles with explorable visualizations that allow readers to understand complex data in a personalized way. E-readers, such as the Kindle, provide in-situ context to help readers better comprehend complex documents, showing inline term definitions and tracking occurrence of characters in a long novel. While prior work has focused on authoring support tools [21, 22, 54] that can reduce effort in creating interactive scientific documents [34, 38], they have not seen widespread adoption due to a lack of incentive structure [27]. Furthermore, millions of research papers are locked in the rigid and static PDF format, whose low-level syntax makes it extremely difficult for systems to access semantic content, augment interactivity, or even provide basic reading functionality for assistive tools like screen readers [6].

Fortunately, recent work on layout-aware document parsing [39, 86, 104] and large language models [4, 10, 83] show promise for accessing the content of PDF documents, and building systems that can better understand their semantics. This raises an exciting challenge: *Can we create intelligent, interactive, and accessible reading interfaces for research papers, even atop existing PDFs?*

*[Click to open in the Semantic Reader] [Download the version with alt-text]

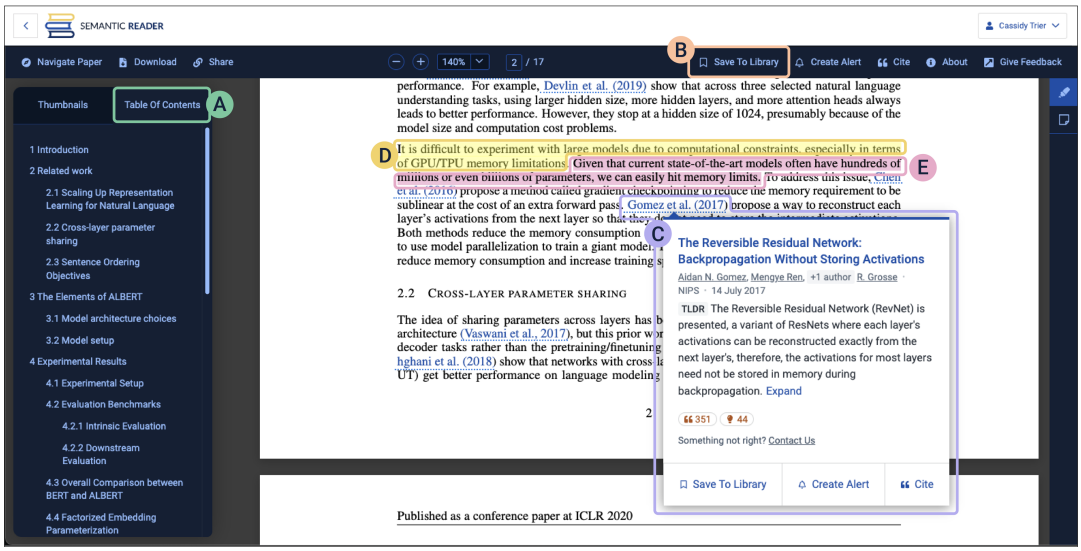

Figure 1: The Semantic Reader Project consists of research, product, and open science resources. The Semantic Reader product[1] is a free interactive interface for research papers. It supports standard reading features (e.g., (A) table of contents), integration with Semantic Scholar (e.g., (B) save to library), useful augmentations atop the existing PDF (e.g., (C) in-situ Paper Cards when clicking inline citations), and integration with third-party features (e.g. (D) Hypothes.is[5] for user highlights). We continues to integrate research features into this product as they mature (e.g., (E) Scim automated highlights §3.1).

To explore this question, we present the **Semantic Reader Project**, a broad collaborative effort across multiple non-profit, industry, and academic institutions to create interactive, intelligent reading interfaces for research papers. This project consists of three pillars: research, product, and open science resources. On the research front, the Semantic Reader Project combines AI and HCI research to design novel, AI-powered interactive reading interfaces that address a variety of user challenges faced by today's scholars. We developed research prototypes and conducted usability studies that clarify their benefits. On the product front, we are developing the Semantic Reader (Figure 1),[1] a freely available reading interface that integrates features from research prototypes as they mature.[2] Finally, we are developing and releasing open science resources that drive both the research and the product. These resources together open-source software,[3,4] AI models [11, 20, 42, 86], and open datasets [50, 61] to support continued work in this area.

In this paper, we focus on summarizing our efforts under the *research pillar* of the Semantic Reader Project. We structure our discussion around five broad challenges faced by readers of research papers:

- **Discovery:** Following paper citations is one of the main strategies that scholars employ to discover additional relevant papers, but keeping track of the large numbers of citations can be overwhelming. In §2, we explore ways to visually augment research papers to help readers prioritize their paper exploration during literature reviews.
- **Efficiency:** The exponential growth of publication makes it difficult for scholars to keep up-to-date with the literature—scholars need to skim and read many papers while making sure they capture enough details in each. In §3, we explore how support for non-linear reading can help readers consume research papers more efficiently.
- **Comprehension:** Research papers can be dense and contain terms that are unfamiliar either because the author newly introduces them or assumes readers have prerequisite domain knowledge. In §4, we explore how providing in-situ definitions and summaries can benefit readers especially when reading outside of their domains.
- **Synthesis:** The sensemaking [84] process of synthesizing knowledge scattered across multiple papers is effortful but important. It allows scholars to make connections between prior work and identify opportunities for future research. In §5, we explore how to help readers collect information from and make sense of many papers to gain better understanding of broad research topics.
- **Accessibility:** Static PDFs are an ill-suited format for many reading interfaces. For example, PDFs are notoriously incompatible with screen readers, and represent a significant barrier for blind and low vision readers [6]. Furthermore, an increasing number of scholars access content on mobile devices, on which PDFs of papers are difficult to read. In §6, we explore methods for converting legacy papers to more accessible representations.

Specifically, we present ten research prototypes developed in the Semantic Reader Project—CiteSee [15], CiteRead [82], Scim [28], Ocean [75], ScholarPhi [33], Paper Plain [1], Papeo [92], Threddy [43],

[1]Semantic Reader: https://www.semanticscholar.org/product/semantic-reader
[2]Available for over 369K papers as of February 2023.
[3]For UI development: https://github.com/allenai/pdf-component-library
[4]For processing PDFs: https://github.com/allenai/papermage
[5]Hypothes.is: https://web.hypothes.ishttps://web.hypothes.is

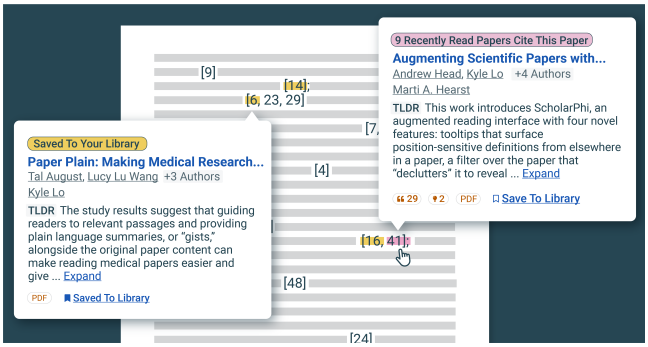

Figure 2: CiteSee [15] highlights citations to familiar papers (e.g., recently read or saved in their libraries) as well as unfamiliar papers to help readers avoid overlooking important citations when conducting literature reviews. Clicking on *Expand* surfaces additional context, such as citing sentences from recently read papers.

Relatedly [73], and SciA11y [91, 97]—and explain how they address these reading challenges. We conclude by discussing ongoing research opportunities in both AI and HCI for developing the future of scholarly reading interfaces. We provide pointers to our production reading interface and associated open resources to invite the broader research community to join our effort.

2 UNLOCKING CITATIONS FOR DISCOVERY

Scholars use many methods to discover relevant research papers to read, including search engines, word of mouth, and browsing familiar venues. However, once they find one research paper, it's especially common for scholars to use its references and citations to further expand their knowledge of a research area. This behavior, sometimes referred to as *forward/backward chaining* or *footnote chasing*, is ubiquitous and has been observed across many scholarly disciplines [74]. Supporting this, one popular feature in the Semantic Reader[1] is in-situ Paper Cards that pop up when readers click on an inline citation, dramatically reducing the interaction cost caused by jumping back-and-forth between inline citations and their corresponding references at the end of a research paper (Figure 1). Despite this affordance, during literature reviews, readers may still be overwhelmed trying to make sense of the tens to hundreds of inline citations in each paper [15, 24, 77]. Conversely, when reading a given paper, a reader cannot see relevant follow-on research papers that cited the current paper. Here we discuss how interactive reading interfaces can help scholars more effectively explore citations to important relevant work in both directions with two systems called CiteSee [15] and CiteRead [82].

2.1 Augmenting Citations with CiteSee

While most prior work on supporting research paper discovery has focused on developing bespoke interfaces of recommender systems or visualizations based on paper contents [78, 89], the citation graph [16, 30, 32, 40, 63, 80, 103], or a combination of the two [20, 96], research paper discovery via inline citations in a reading interface is important but under-explored. One study estimates that reading and exploring inline citations accounts for around one in five research paper discoveries during active research [49]. However, while all inline citations are relevant to the current research paper, it is likely that some are more relevant to the current reader than others. For example, a reader reading papers about *aspect extraction of online product reviews* to learn more about *natural language processing techniques* would be less interested in citations to research papers around *e-commerce and marketing*. In addition, citations to the same research papers often have different surface forms across papers (i.e., reference numbers), making it all the more difficult for readers to keep track of all the inline citations they should explore or have already explored during literature reviews.

To address this, CiteSee provides a personalized research paper reading experience by automatically identifying and resolving inline citations in PDFs to research paper entities in our academic graph [50], and visually augmenting inline citations based on their connections to the current reader. First, CiteSee leverages a reader's reading behavior and history as a way to capture their short-term and fluid interests during literature reviews. Using this signal, CiteSee scores and highlights inline citations to help the reader triage them and discover prior work that are likely relevant to their literature review topics (Figure 2). Second, CiteSee leverages research papers saved in the reader's Semantic Scholar paper library and the reader's publication record [50] to understand their longer-term research interests. Using this signal, CiteSee changes the colors of the inline citations to familiar papers so that the reader can both better contextualize the current paper and keep track of citations to papers they have already explored. In addition, CiteSee also helps readers better make sense of the cited papers by showing how they connect to a reader's previous activities; for example, showing which library folders they were saved under or the citing sentences from a familiar research paper (Figure 2). Based on lab and field studies, CiteSee showed promise that providing visual augmentation and personalized context around inline citations in an interactive reading environment can allow readers to more effectively discover relevant prior work and keep track of their exploration during real-world literature review tasks.

2.2 Exploring Future Work with CiteRead

While augmenting inline citations helps readers to triage them, many relevant research papers are not cited in a research paper in the first place, for example, because they were published afterwards. CiteRead is a novel reading interface that helps readers discover how follow-on work has built on or engaged with the research paper [82]. Much like social document annotation systems [109], CiteRead annotates text in the paper with margin notes containing relevant commentary from citing papers [70], thereby helping the reader to become aware of the citing paper and its connection. In order to produce these annotations automatically, CiteRead first filters citing research papers for ones that are most relevant to the reader using a trained model atop a number of features representing citational discourse and textual similarity, i.e. from scientific paper embeddings [20]. CiteRead then localizes citing papers to particular spans of text in the paper being read, and extracts relevant information from the citing paper. Figure 3 shows a research paper annotated with this information from citing papers. Localization is

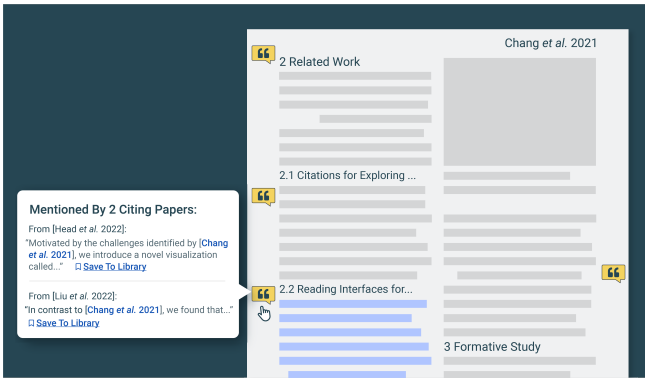

Figure 3: CiteRead [82] finds subsequently published citing research papers, extracts the citation context, and localizes it to relevant parts of the current research paper as margin notes. This allows readers to become aware of important follow on work and explore them in-situ.

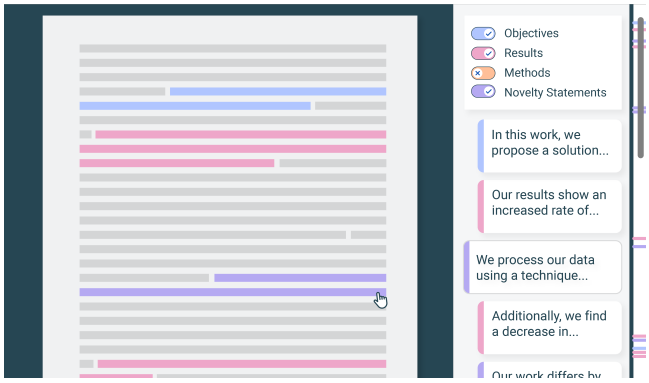

Figure 4: The Scim [28] interface guides reader attention using color highlights corresponding to discourse facets. A sidebar allows users to toggle facets on/off. Clicking a color-coded snippet scrolls the reader to the relevant passage.

a technical challenge because while inline citations reference cited papers, they do not typically reference specific locations in the cited paper; CiteRead determines location by looking for overlapping spans of text (e.g., a number in common in the citing paper and the cited paper) or localizes to the relevant section when this overlap is unavailable. With CiteRead, a reader can directly examine follow-on work while keeping the citation contexts of both the current paper and the citing paper. In a lab study, CiteRead helped readers better understand a research paper and its follow-on work compared to providing readers with a separate interface for faceted browsing of follow-on work.

3 NAVIGATION AND EFFICIENT READING

An important part of reading a paper is knowing what and where to read. Scholars often read papers non-linearly; they might return to a previously-read passage to recall some information, or jump forward to a different section of the paper (or to another paper) to satisfy an information need before jumping back. While jumping can help scholars orient their reading to sections of interest, it can also be a distraction by causing readers to constantly switch contexts. Non-linear navigation can be especially burdensome when the reader is interested in a particular *type* of information (e.g., skimming a paper for the main results), but doesn't know precisely where to find it within the paper. In this section we discuss two systems, Scim [28] and Ocean [75], which demonstrate different approaches to helping readers navigate efficiently through a paper toward high-value, relevant information.

3.1 Guided Reading with Scim

Scholarly reading can be considered a type of sensemaking represented as a continuous interplay between two processes: *information foraging* in which readers identify relevant paper content, and *comprehension* in which readers attempt to integrate the new information into their working model of the paper and with relevant prior knowledge [79, 84]. Distinguishing between relevant and irrelevant content could help facilitate efficient reading. Paper abstracts offer one such separation, in essence an author-crafted determination of relevant content. However, static paper abstracts can leave readers to desire additional details that then require them to skim the paper itself.

Scim [28] addresses this problem via an augmented reading interface designed to guide readers' attention using automatically-created in-situ faceted highlights (Figure 4). Though prior work has explored highlighting as a visual cue for guiding reader attention [18, 102, 105], the efficacy for reading of scholarly text is less well-understood. Scim investigated the following design goals for intelligent highlights in scholarly reading: highlights should be (1) evenly-distributed throughout a paper, (2) have just the right density (too few highlights will present the guise of an inept tool, and too many will slow a reader down), and (3) highlight several key categories of information in the paper. Because readers often skim for common types of information, Scim uses a pretrained language model [99] to classify salient sentences within papers into one of four information facets: research objectives, novel aspects of the research, methodology, and results, coupled with heuristics that ensure an even distribution of highlights. Usability studies of Scim have shown these highlights can reduce the time it takes readers to find specific information within a paper. Readers found Scim particularly useful when skimming text-dense papers, or for papers that fell outside their area of expertise. Moreover, readers learned to use both Scim's inline highlights and a sidebar summary of highlights to augment their existing reading strategies.

3.2 Low-Vision Navigation Support and Reader-Sourced Hyperlinks with Ocean

The task of navigating between sections and retrieving content can be particularly challenging for blind and low-vision readers due to limitations in auditory information access or small viewports under high magnification [90]. Even when related content is linked, a small viewport can make navigation difficult and necessitate scrolling [75]. Most existing tools such as for auditory skimming [45] do not address such challenges associated with low-vision and magnification.

Ocean [75] minimizes scrolling requirements for low-vision readers by providing bi-directional, viewport-preserving hyperlinks that enable navigating to and from associated content without disrupting the viewport. Based on reported findings from interviews with low-vision readers, Ocean also allows for easily revisiting portions of the paper with tabbed reading. Since papers do not always provide hyperlinks and automated link creation is imperfect, Ocean includes an authoring interface that allows readers to create and share paper links during reading. An exploratory field deployment study with mixed-ability groups of low-vision and sighted readers revealed that readers found value in creating and consuming these links, and that reader-created links can increase trust.

4 IN-SITU EXPLANATIONS FOR BETTER COMPREHENSION

Could an augmented reading application help readers understand a paper by reducing the cognitive load associated with reading a paper? In this section, we discuss several ways in which interactive reading aids can help a reader understand a paper with less work through three systems: ScholarPhi [33], PaperPlain [1] and Papeo [92]. In particular, papers can be augmented with definitions of terms and symbols, provide plain-language summaries of paper passages, and connect readers with alternative forms of expression (for instance, video clips of research talks) that offer more approachable explanations of the paper's content.

4.1 Defining Terms and Symbols with ScholarPhi

Understanding a paper requires understanding its vocabulary. However, this is by no means an easy task—a typical paper may contain dozens of acronyms, symbols, and invented terms. And often, these terms appear without accompanying definitions [69]. How can we design interactive aids that present definitions of terms when and where readers most need them? ScholarPhi [33] takes as its basis the term gloss—an extension to a reading interface that shows a reader an explanation of a phrase when they click it. Glosses appeared in early research interfaces for reading hypertext [107] and

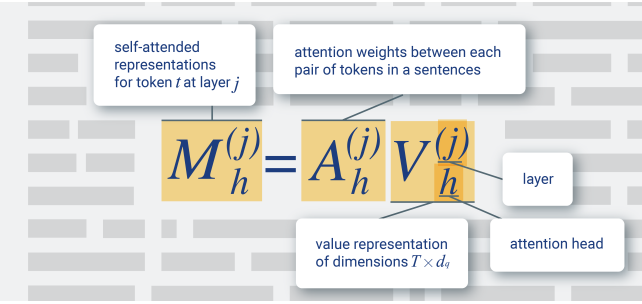

Figure 5: ScholarPhi [33] shows definitions of terms and symbols in pop-up tooltips. When a reader selects a formula, all known definitions of symbols are shown simultaneously. To let readers select nested symbols (e.g., "h" in "$V_h^{(j)}$"), ScholarPhi supports "drill-down" subsymbol selection.

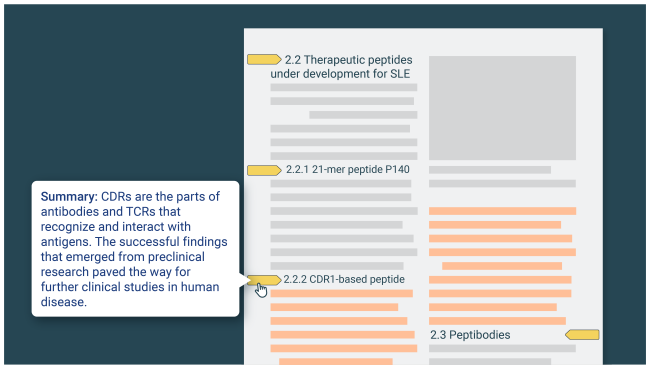

Figure 6: Paper Plain [1] provides in-situ plain language summaries of passages called "gists" to help readers who are overwhelmed by complex textual passages. Readers access gists by clicking a flag next to a section header. These gists are generated by large language models.

have since become part of widely-used reading interfaces including Wikipedia and Kindle.

That said, familiar gloss designs do not work well for scientific papers, where glosses run the risk of distracting readers, terms have multiple meanings, and phrases (specifically math symbols) are difficult to unambiguously select. The ScholarPhi design addresses these challenges. First, it aims to reduce distraction by showing definitions with high economy: glosses show multiple definitions and and in-context usages within a compact tooltip. Second, it provides position-sensitive definitions, revealing definitions that appears most recently prior to the selected usages of terms. Terms and definitions are automatically identified using a pretrained language model [42]. Finally, it provides easier access to definitions of mathematical symbols. Readers can access definitions of both a symbol and the subsymbols it is made of through a multi-click, "drill-down" selection mechanism. Furthermore, when a reader selects a formula, they can see definitions for all symbols at once, automatically placed adjacent to the symbols in the formula's margins (see Figure 5).

In a usability study, the above interactions reduced the time it took readers to find answers to questions involving the understanding of terminology. All readers reported they would use the definition tooltips and formula diagrams often or always if available in their PDF reader tools.

4.2 Simplifying Complex Passages with Paper Plain

Helping a reader understand individual terms and phrases only addresses part of the problem. Papers often contain passages so dense and complex that individual definitions are not enough to help someone read the passages, especially if they are a novice or non-expert in a field [9]. Can we make complex texts more approachable by incorporating plain language summaries in the margins of the text? With Paper Plain [1], when a reader encounters a section they find difficult to read, they can access a plain language summary of that section by clicking a button adjacent to the section

header (see Figure 6). These summaries are generated by prompting a large language model with section text [10].

Furthermore, Paper Plain helps guide readers using these summaries as an "index" into the text. A sidebar containing questions a reader may have about the text (e.g., *What did the paper find?* or *What were the limitations?*) provides links into answering passages identified using a question-answering system [106] alongside their associated plain language summaries. These features were designed to help readers understand the "gist" of passages that contain unfamiliar vocabulary, providing support beyond that of individual term definitions. Drawing inspiration from prior interactive reading affordances for term definitions [41], in-situ question answering [17, 108], and guiding reading [26], Paper Plain seeks to bring these features together into a holistic system capable of supporting reading of a paper by a non-expert readership. In a usability study, readers made more frequent use of passage summaries than definition tooltips when both were available, suggesting the potential value of plain language summaries as allowing readers to bypass definitions of individual terms when acquiring a broad understanding of a paper.

4.3 Fusing Papers and Videos with Papeo

Sometimes, the best explanation of an idea is non-textual. Videos can enhance understanding [65] while also requiring less mental load [66], and various tools have been designed to facilitate searching and browsing for explanations in informational videos such as lectures [47, 52, 57, 76] and tutorials [46, 48, 93]. Similarly, for research papers, an algorithm might be better explained through an animation, a user interface might be better showcased through an screen recording, compared to the proses of a paper [37]. Instead of consuming the two formats independently, could interactive reading interfaces offer readers access to these alternative, more powerful descriptive forms as they read? For this, Papeo [92] was developed as a tool that supplements papers with more engaging, concise, dynamic presentations of information by linking excerpts of talk videos to corresponding paper passages. To grant authors more control over how their work is presented, we developed an AI-supported authoring interface for linking paper passages and videos efficiently: candidate passages are linked to excerpts of videos as suggestions using a pretrained language model [100], and an author interactively confirms or refines them.

Unlike text-skimming with Scim (§3.1) and Paper Plain (§4.2), video-skimming in Papeo combines multiple modalities to explain complex information. For example, instead of reading a long text description of an interactive system, readers could see the system's behavior in a screen recording video with the author's commentary, and switch to corresponding passages to see implementation details or design motivations if desired. Our early-stage evaluations of Papeo suggest that readers can use these interactions to fluidly transition between watching video and reading text, using video to quickly understand, and then selectively descending into the text when they desire a detailed understanding of the paper.

5 SCAFFOLDING SYNTHESIS WITH RELATED WORK SECTIONS

Scientific breakthroughs often rely upon scholars synthesizing multiple published works into broad overviews to identify gaps in the current literature [81]. For this, scholars periodically compile survey articles to help other scholars gain a comprehensive overview of important research topics. For example, some fields have dedicated outlet for such articles (e.g., the *Psychological Bulletin* [5]). However, survey articles require significant time and effort to synthesize, and can quickly become outdated with the exponential growth of scientific publication [7].

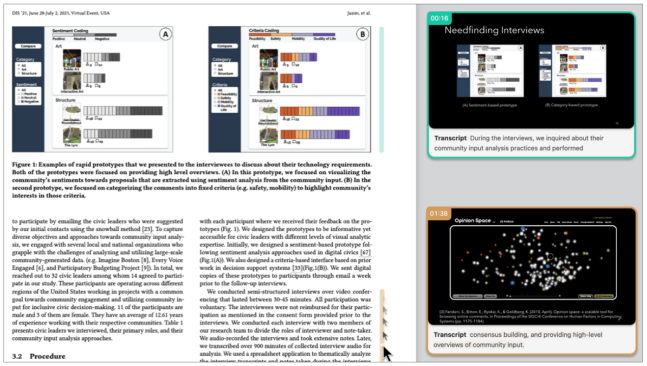

Figure 7: Papeo [92] enables authors to map segments of talk videos to relevant passages in the paper, allowing readers to fluidly switch between the two formats. Color-coded bars show the mapping between the two formats, and allow readers to scrub through video segments for quick previews.

Instead, scholars in fast-paced disciplines often rely on the related work section when they need to better understand the broader background when reading a paper. While related work sections also summarize multiple prior works, unlike comprehensive survey articles, they typically provide partial views of the larger research topic most relevant to a single paper. There is an opportunity to build better tooling for scholars to consume and synthesize multiple related work sections across many papers to gain richer and more comprehensive overviews of fast-paced domains. The Threddy [43] and Related [73] projects explored this opportunity using two different approaches: clipping and organizing research threads mentioned across papers [43], and directly exploring and reading related work sections extracted across many papers [73].

5.1 Clipping and Synthesizing across Papers with Threddy

Clipping and note-taking is one common approach to supporting synthesis across multiple documents. Prior work has pointed to the importance of tightly integrating clipping and synthesis support in the reading process, and how incurring significant context-switching costs can be detrimental to sensemaking [51, 79, 84]. Therefore, recent work has developed tools aimed at reducing the cognitive and interaction costs of clipping [12, 60] and structuring [13, 53, 58, 59, 88] to support everyday online researchers [13], programmers [58], and students [88]. However, designing clipping and synthesis support tools for research papers is relatively under-explored and introduces exciting new research opportunities. For example, additional organizational structures for literature reviews

(*e.g.,* threads of prior work instead of tables [13, 58]), and research paper discovery (*e.g.,* based on inline citations in clipped text).

For this, Threddy [43] is a thread-focused clipping tool integrated into scholars' paper reading process to support literature review and discovery. Using Threddy, readers can select and save sentences into a sidebar from the related work sections of a paper. The system maintains rich context for each clip, including its provenance and inline citations. This allows readers to navigate back to the clipped paper and cited papers afterward. In the sidebar, readers can further organize clips collected across papers into a hierarchy of threads to form their view of the research landscape. The content of the sidebar is preserved across papers that were read over time, and provides valuable context for subsequent reading based on the emerging threads of research the reader have curated. Finally, readers can further expand their coverage by exploring paper recommendations for each thread, based on the referenced papers in the corresponding clips. A lab study showed that Threddy was able to lower the interaction costs of saving clips while maintaining context, allowed participants to curate research threads without breaking reading flows, and discover interesting new papers to further grow their understanding of the research fields.

5.2 Reading and Exploring Related Work Sections across Papers with Relatedly

In contrast to Threddy, which aims to improve readers' existing literature review process through enhanced in-situ clipping and synthesis [43], the Relatedly system introduced a novel workflow that allows readers to explore many related work sections across papers in an interactive search and reading interface to quickly gain a comprehensive overview of rich research topics [73]. While prior work have explored providing overview structure of multiple documents based on citations [16, 80], semantic similarity [36, 85], or human computation [14, 31, 62], they could still lead to complex structures that are hard to interpret [35] or require significant crowdsourcing efforts. Relatedly sidesteps these issues by *reusing* existing related work paragraphs in published papers which already cite sets of related references with descriptions connecting them [73]. As an example, consider a scholar trying to better understand the space of *online misinformation*. With *online misinformation* as the query term, Relatedly shows the reader a list of paragraphs that describe and cite multiple relevant prior work. Using a pretrained language model for summarization [56], Relatedly generates short and descriptive titles for each paragraph, and uses a diversity-based ranking algorithm so that the reader can quickly see and explore paragraphs describing different research threads, such as *Fact Checking Datasets*, *Social Media and Misinformation*, and *Fake News Detection Techniques*.

One challenge here is that paragraphs of the same threads often cite overlapping prior work, making them hard to explore and read while keeping track of which papers were new versus already explored. For this, Relatedly provides reading and cross-referencing support by keeping track of paragraphs and references explored by the readers. This allows Relatedly to help readers prioritize their reading for both breadth and depth. Specifically, Relatedly dynamically re-ranks paragraphs and highlights sentences to spotlight unexplored and dissimilar references for breadth, but also allow readers to explore clusters of paragraphs that cited similar references for depth. A usability study comparing Relatedly to a strong document-centric baseline showed that Relatedly led to participants writing summaries that were rated significantly more coherent, insightful, and detailed after 20 minutes of literature review.

6 DYNAMIC DOCUMENTS FOR IMPROVED ACCESSIBILITY

A range of disabilities cause people to read scientific documents using a wide variety of devices and reading tools. For example, blind and low vision readers may use assistive reading technology such as screen readers, screen magnification, or text-to-speech to read documents [90]. Furthermore, people without disabilities face situational impairments, such as the inability to view a screen while driving or may have a preference for consuming content on a small, mobile device.

Many of these reading tools, such as screen readers, do not function properly on document formats designed for print such as PDF unless the document has been manually post-processed to add information about reading order, content type, etc., which is rarely performed on scientific documents [6, 98]. Further, certain content elements such as figures require the addition of alternative text in order to be read aloud at all (figure captions typically assume the reader can see the figure and do not provide the same semantic content as alt text). High magnification reduces the viewport (the amount of visible content) and can dramatically increase the amount of scrolling and panning required, especially for multi-columnar formats that are commonly used by scientific documents. Visual scanning for information may be impacted or unavailable in these settings, making it more difficult to find and navigate between content in the document [75].

One way to render legacy PDF content more accessibly is to parse and convert it into a more flexible format, such as XML or HTML, which can then be formatted for mobile devices and augmented for reading by screen readers. The SciA11y system[6] demonstrates this approach, automatically converting 12M academic PDFs to HTML [97]; a user study with blind and low vision participants demonstrated strong user appreciation of the output, though some errors remain (e.g., failing in certain cases to distinguish footnotes from body text, difficulty parsing math equations) [98]. When available, alt text can be automatically categorized into semantic content types, enabling new reading experiences that allow skipping or prioritizing certain types [19]. Other approaches provide complementary benefits, such as interfaces tailored for low-vision readers (§ 3.2), as well as the range of reading support systems outlined above.

7 DISCUSSION AND FUTURE WORK

There are additional directions to explore to better support scholarly activities through the Semantic Reader Project.

Towards a full-featured reading experience. One question is how to integrate the different kinds of functionality across these projects into one coherent user interface, especially as we migrate research features into the production interface. Another question is how

[6] A demo of a subsequent version is available at https://papertohtml.org/

to develop support for the oftentimes social and collaborative nature of scholarly reading. Scholars frequently leverage their social networks and other social signals for paper discovery [44], work in groups to conduct literature review triage and synthesis, or engage in reading group discussions to aid comprehension. Existing augmentations within the Semantic Reader product could imbue social information, such as providing signals from one's co-author network (e.g., in CiteSee §2.1) or aggregate navigation traces (e.g., in Scim §3.1). The publicly-available Semantic Reader tool could also scaffold the creation of novel crowd- or community-sourced content, such as author- or reader-provided explanations, commentary, or verification of paper content. Finally is the question of how we can allow the scholarly community to step in where current AI systems fall short, such as by fixing improperly-extracted content or incorrect generated text which are especially problematic for interfaces such as SciA11y (§6).

Advancing AI for scholarly documents. The Semantic Reader Project presents an opportunity for further AI research in scholarly document processing, especially when paired with human-centered research grounded in user-validated systems and scenarios. The bar for deploying AI models to support real-world reading is high; we often found during iterative design and usability studies that even slight errors in these models can have detrimental effects on the readers. Until recently, interface design could require months of development of bespoke AI models which creates a barrier for quickly iterate different system designs. Recent advancements in scaling large language models (LLMs) has altered this landscape by enabling researchers to experiment with a wide range of new NLP capabilities at relatively low cost [10]. This has the potential of significantly lowering the cost of human-centered AI design by incorporating user feedback in earlier stages of system development to create AI systems that work in symphony with the users beyond pure automation [87]. For example, when developing Paper Plain (§4.2), LLMs enabled us to quickly test different granularities and complexity-levels of plain language summaries with participants, eschewing the need for expensive changes to data requirements and model retraining. In the near-term, we will revisit interface designs relying on bespoke AI models to evaluate whether LLMs can close the gap between research prototype and ready-for-production (e.g., more accurate definition identification for ScholarPhi §4.1). Longer-term, we will explore whether LLMs can power new interactions (e.g., user-provided natural language queries while reading [23, 95]). While recent work has shown that these models can occasionally make critical errors or generate factually incorrect text when processing scientific text [72], we remain cautiously optimistic about developing ways to address their limitations [25, 55].

Ethics of augmented papers. Finally, all the new interfaces for reading that we propose pose a number of important ethical considerations that will require further research and discussion. One aspect that arises with any system for elevating certain papers or certain content over others is bias. For instance, using signals such as citation counts faces the risk of a "rich get richer" bias, which can reflect other kinds of documented biases [3, 64, 101]. As a result, systems such as CiteSee (§2.1) or Relatedly (5.2) should carefully consider additional signals of relevance such as semantic similarity to surface newer and overlooked papers. Another tension that we have encountered is the potential discrepancy between author desires and reader desires for how a work is presented and how much control to provide authors. For instance, in our work on Papeo (§4.3), we found that authors desired control over placement of their talk video snippets, even as they found automated mapping support to be helpful. In other cases, authors might not have the requisite expertise (e.g., they may not have a good sense of reader needs or what non-experts are confused by) or may have the wrong incentives. Future work should consider author perspectives on these augmented experiences. A related issue is around systems for more efficient reading or synthesis, which may encourage readers to take shortcuts that lead to incorrect understanding, sloppy research, or even outright plagiarism. Instead of simply seeking to increase reading throughput uniformly, our systems should enable *triage*, so that readers can dedicate time for thoughtful and careful reading when the content is important. For instance, our systems could design pathways that, while they may be more efficient, do not obfuscate the full context (e.g., Scim §3.1), and that encourage good practices such as verification and provenance tracing. A final consideration is around what is ethical reuse of a paper's contents to support reader experiences outside of that paper and its licensing implications. For instance, CiteRead (§2.2) extracts paper citances and places them in the cited paper, and Relatedly (§5.2) extracts related work sections from different papers for users to explore. Recent trends in *open science and datasets* [29, 61, 67, 68] point to a promising future where we could continue to explore different ways to *remix and reuse* scholarly content across context so that future scientists can take fuller advantage of prior research.

8 CONCLUSION

This paper describes the Semantic Reader Project, which currently consists of ten research prototypes focusing on supporting scientists around Discovery [15, 82], Efficiency [28, 75], Comprehension [1, 33, 92], Synthesis [43, 73], and Accessibility [91, 97] when reading research papers. Validating our approach of augmenting existing PDFs of research papers, we have seen tremendous adoption of the freely-available Semantic Reader product[1] which has grown to 10k weekly users.[7] While we focused on augmenting PDF documents to support common scholar reading practices, all of our reading interfaces are built with web technologies—allowing these novel interactions to extend to future publication formats which can be rendered in web browsers. We plan to continue experimenting with novel AI-powered intelligent reading interfaces, as well as migrating successful interactive features into the product. Finally, we offer a collection of freely-available resources to the larger research community, including datasets of open-access research papers [61], APIs for accessing the academic citation graph [50], machine learning models for processing and understanding research papers [11, 20, 42, 86],[4] and open-source software for rendering and augmenting PDF documents for developing reading interfaces.[3] We hope by providing these resources we can enable and encourage the broader research community to work on exciting novel intelligent reading interfaces for research papers with us.

[7] As of late February, 2023

ACKNOWLEDGMENTS

This project is supported in part by NSF Grant OIA-2033558, NSF Grant CNS-2213656. NSF RAPID Award 2040196, and ONR Grant N00014-21-1-2707.

REFERENCES

[1] Tal August, Lucy Lu Wang, Jonathan Bragg, Marti A. Hearst, Andrew Head, and Kyle Lo. 2023. Paper Plain: Making Medical Research Papers Approachable to Healthcare Consumers with Natural Language Processing. *ACM Transactions on Computer-Human Interaction* (2023). To appear.

[2] Charles Bazerman. 1985. Physicists reading physics: Schema-Laden Purposes and Purpose-Laden Schema. *Written Communication* 2, 1 (Jan. 1985), 3–23.

[3] Joeran Beel and Bela Gipp. 2009. Google Scholar's ranking algorithm: The impact of citation counts (An empirical study). *2009 Third International Conference on Research Challenges in Information Science* (2009), 439–446.

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

[5] Daryl J Bem. 1995. Writing a review article for Psychological Bulletin. *Psychological Bulletin* 118, 2 (1995), 172.

[6] Jeffrey P. Bigham, Erin L. Brady, Cole Gleason, Anhong Guo, and David A. Shamma. 2016. An Uninteresting Tour Through Why Our Research Papers Aren't Accessible. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI EA '16)*. Association for Computing Machinery, New York, NY, USA, 621–631. https://doi.org/10.1145/2851581.2892588

[7] Lutz Bornmann, Ruediger Mutz, and Robin Haunschild. 2020. Growth rates of modern science: a latent piecewise growth curve approach to model publication numbers from established and new literature databases. *Humanities and Social Sciences Communications* 8 (2020), 1–15.

[8] Jeffrey Brainard. 2020. Scientists are drowning in COVID-19 papers. Can new tools keep them afloat. *Science* 13, 10 (2020), 1126.

[9] Mary Anne Britt, Tobias Richter, and Jean-François Rouet. 2014. Scientific Literacy: The Role of Goal-Directed Reading and Evaluation in Understanding Scientific Information. *Educational Psychologist* 49 (2014), 104 – 122.

[10] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel Ziegler, Jeffrey Wu, Clemens Winter, Chris Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language Models are Few-Shot Learners. In *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin (Eds.), Vol. 33. Curran Associates, Inc., 1877–1901. https://proceedings.neurips.cc/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf

[11] Isabel Cachola, Kyle Lo, Arman Cohan, and Daniel S. Weld. 2020. TLDR: Extreme Summarization of Scientific Documents. In *Findings of EMNLP*.

[12] Joseph Chee Chang, Nathan Hahn, and Aniket Kittur. 2016. Supporting Mobile Sensemaking Through Intentionally Uncertain Highlighting. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. Association for Computing Machinery, New York, NY, USA, 61–68. https://doi.org/10.1145/2984511.2984538

[13] Joseph Chee Chang, Nathan Hahn, and Aniket Kittur. 2020. Mesh: Scaffolding Comparison Tables for Online Decision Making. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '20)*. Association for Computing Machinery, New York, NY, USA, 391–405. https://doi.org/10.1145/3379337.3415865

[14] Joseph Chee Chang, Aniket Kittur, and Nathan Hahn. 2016. Alloy: Clustering with crowds and computation. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 3180–3191.

[15] Joseph Chee Chang, Amy X Zhang, Jonathan Bragg, Andrew Head, Kyle Lo, Doug Downey, and Daniel S Weld. 2023. CiteSee: Augmenting Citations in Scientific Papers with Persistent and Personalized Historical Context. *ArXiv* abs/2022.99999 (2023).

[16] Duen Horng Chau, Aniket Kittur, Jason I Hong, and Christos Faloutsos. 2011. Apolo: making sense of large network data by combining rich user interaction and machine learning. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 167–176.

[17] Vinay K. Chaudhri, Britte Haugan Cheng, Adam Overholtzer, Jeremy Roschelle, Aaron Spaulding, Peter E. Clark, Mark T. Greaves, and David Gunning. 2013. Inquire Biology: A Textbook that Answers Questions. *AI Mag.* 34 (2013), 55–72.

[18] Ed H. Chi, Lichan Hong, Michelle Gumbrecht, and Stuart K. Card. 2005. ScentHighlights: highlighting conceptually-related sentences during reading. In *Proceedings of the 10th International Conference on Intelligent User Interfaces*. Association for Computing Machinery, San Diego, CA, USA, 272.

[19] Sanjana Shivani Chintalapati, Jonathan Bragg, and Lucy Lu Wang. 2022. A Dataset of Alt Texts from HCI Publications: Analyses and Uses Towards Producing More Descriptive Alt Texts of Data Visualizations in Scientific Papers. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility* (Athens, Greece) *(ASSETS '22)*. Association for Computing Machinery, New York, NY, USA, Article 30, 12 pages. https://doi.org/10.1145/3517428.3544796

[20] Arman Cohan, Sergey Feldman, Iz Beltagy, Doug Downey, and Daniel S Weld. 2020. Specter: Document-level representation learning using citation-informed transformers. *arXiv preprint arXiv:2004.07180* (2020).

[21] Matthew Conlen and Jeffrey Heer. 2022. Fidyll: A Compiler for Cross-Format Data Stories & Explorable Explanations. *ArXiv* abs/2205.09858 (2022).

[22] Matthew Conlen, Megan Vo, Alan Tan, and Jeffrey Heer. 2021. Idyll studio: A structured editor for authoring interactive & data-driven articles. In *The 34th Annual ACM Symposium on User Interface Software and Technology*. 1–12.

[23] Pradeep Dasigi, Kyle Lo, Iz Beltagy, Arman Cohan, Noah A. Smith, and Matt Gardner. 2021. A Dataset of Information-Seeking Questions and Answers Anchored in Research Papers. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics, Online, 4599–4610. https://doi.org/10.18653/v1/2021.naacl-main.365

[24] Andrew S. Denney and Richard Allan Tewksbury. 2013. How to Write a Literature Review. *Journal of Criminal Justice Education* 24 (2013), 218 – 234.

[25] Graham Dove, Kim Halskov, Jodi Forlizzi, and John Zimmerman. 2017. UX design innovation: Challenges for working with machine learning as a design material. In *Proceedings of the 2017 chi conference on human factors in computing systems*. 278–288.

[26] Kristina Dzara and Ariel S Frey-Vogel. 2019. Medical Education Journal Club for the Millennial Resident: An Interactive, No-Prep Approach. *Academic pediatrics* (2019).

[27] Editorial Team. 2021. Distill Hiatus. *Distill* 6, 7 (July 2021). https://doi.org/10.23915/distill.00031

[28] Raymond Fok, Hita Kambhamettu, Luca Soldaini, Jonathan Bragg, Kyle Lo, Andrew Head, Marti A. Hearst, and Daniel S. Weld. 2023. Scim: Intelligent Skimming Support for Scientific Papers. In *28th Annual Conference on Intelligent User Interfaces (IUI '23)*.

[29] Paul Ginsparg. 2011. ArXiv at 20. *Nature* 476, 7359 (2011), 145–147.

[30] Marco Gori and Augusto Pucci. 2006. Research Paper Recommender Systems: A Random-Walk Based Approach. *2006 IEEE/WIC/ACM International Conference on Web Intelligence (WI 2006 Main Conference Proceedings)(WI'06)* (2006), 778–781.

[31] Nathan Hahn, Joseph Chang, Ji Eun Kim, and Aniket Kittur. 2016. The Knowledge Accelerator: Big picture thinking in small pieces. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 2258–2270.

[32] Jiangen He, Q. Ping, Wen Lou, and Chaomei Chen. 2019. PaperPoles: Facilitating adaptive visual exploration of scientific publications by citation links. *Journal of the Association for Information Science and Technology* 70 (2019).

[33] 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* (2021).

[34] Andrew Head, Amber Xie, and Marti A Hearst. 2022. Math Augmentation: How Authors Enhance the Readability of Formulas using Novel Visual Design Practices. In *CHI Conference on Human Factors in Computing Systems*. 1–18.

[35] Marti A Hearst. 1999. The use of categories and clusters for organizing retrieval results. *Natural language information retrieval* (1999), 333–374.

[36] Marti A Hearst. 2006. Clustering versus faceted categories for information exploration. *Commun. ACM* 49, 4 (2006), 59–61.

[37] Tim Niclas Höffler and Detlev Leutner. 2007. Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction* 17 (2007), 722–738.

[38] Fred Hohman, Matthew Conlen, Jeffrey Heer, and Duen Horng Chau. 2020. Communicating with Interactive Articles. *Distill*. https://doi.org/10.23915/distill.00028

[39] Yupan Huang, Tengchao Lv, Lei Cui, Yutong Lu, and Furu Wei. 2022. LayoutLMv3: Pre-training for Document AI with Unified Text and Image Masking. *Proceedings of the 30th ACM International Conference on Multimedia* (2022).

[40] Zan Huang, Wingyan Chung, Thian-Huat Ong, and Hsinchun Chen. 2002. A graph-based recommender system for digital library. In *JCDL '02*.

[41] Abhinav Jain, Nitin Gupta, Shashank Mujumdar, Sameep Mehta, and Rishi Madhok. 2018. Content Driven Enrichment of Formal Text using Concept Definitions and Applications. *Proceedings of the 29th on Hypertext and Social Media* (2018).

[42] Dongyeop Kang, Andrew Head, Risham Sidhu, Kyle Lo, Daniel Weld, and Marti A. Hearst. 2020. Document-Level Definition Detection in Scholarly Documents: Existing Models, Error Analyses, and Future Directions. In *Proceedings of the First Workshop on Scholarly Document Processing*. Association for Computational Linguistics, Online, 196–206. https://doi.org/10.18653/v1/2020.sdp-1.22

[43] Hyeonsu B Kang, Joseph Chee Chang, Yongsung Kim, and Aniket Kittur. 2022. Threddy: An Interactive System for Personalized Thread-based Exploration and Organization of Scientific Literature. *arXiv preprint arXiv:2208.03455* (2022).

[44] Hyeonsu B Kang, Rafal Kocielnik, Andrew Head, Jiangjiang Yang, Matt Latzke, Aniket Kittur, Daniel S Weld, Doug Downey, and Jonathan Bragg. 2022. From Who You Know to What You Read: Augmenting Scientific Recommendations with Implicit Social Networks. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 302, 23 pages. https://doi.org/10.1145/3491102.3517470

[45] Taslim Arefin Khan, Dongwook Yoon, and Joanna McGrenere. 2020. Designing an Eyes-Reduced Document Skimming App for Situational Impairments. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (2020).

[46] Kandarp Khandwala and Philip J. Guo. 2018. Codemotion: expanding the design space of learner interactions with computer programming tutorial videos. *Proceedings of the Fifth Annual ACM Conference on Learning at Scale* (2018).

[47] Juho Kim, Philip J. Guo, Carrie J. Cai, Shang-Wen Li, Krzysztof Z Gajos, and Rob Miller. 2014. Data-driven interaction techniques for improving navigation of educational videos. *Proceedings of the 27th annual ACM symposium on User interface software and technology* (2014).

[48] Juho Kim, Phu Tran Nguyen, Sarah A. Weir, Philip J. Guo, Rob Miller, and Krzysztof Z Gajos. 2014. Crowdsourcing step-by-step information extraction to enhance existing how-to videos. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2014).

[49] Donald W. King, Carol Tenopir, Songphan Choemprayong, and Lei Wu. 2009. Scholarly journal information-seeking and reading patterns of faculty at five US universities. *Learned Publishing* 22 (2009).

[50] Rodney Kinney, Chloe Anastasiades, Russell Authur, Iz Beltagy, Jonathan Bragg, Alexandra Buraczynski, Isabel Cachola, Stefan Candra, Yoganand Chandrasekhar, Arman Cohan, et al. 2023. The Semantic Scholar Open Data Platform. *arXiv preprint arXiv:2301.10140* (2023).

[51] Aniket Kittur, Andrew M. Peters, Abdigani Diriye, Trupti Telang, and Michael R. Bove. 2013. Costs and Benefits of Structured Information Foraging. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. Association for Computing Machinery, New York, NY, USA, 2989–2998. https://doi.org/10.1145/2470654.2481415

[52] Rebecca Krosnick. 2015. VideoDoc : combining videos and lecture notes for a better learning experience.

[53] Andrew Kuznetsov, Joseph Chee Chang, Nathan Hahn, Napol Rachatasumrit, Bradley Breneisen, Julina Coupland, and Aniket Kittur. 2022. Fuse: In-Situ Sensemaking Support in the Browser. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology* (Bend, OR, USA) *(UIST '22)*. Association for Computing Machinery, New York, NY, USA, Article 34, 15 pages. https://doi.org/10.1145/3526113.3545693

[54] Shahid Latif, Zhengzhong Zhou, Yoon Kim, Fabian Beck, and Nam Wook Kim. 2021. Kori: Interactive Synthesis of Text and Charts in Data Documents. *IEEE Transactions on Visualization and Computer Graphics* PP (2021), 1–1.

[55] Min Kyung Lee, Sara B. Kiesler, Jodi Forlizzi, Siddhartha S. Srinivasa, and Paul E. Rybski. 2010. Gracefully mitigating breakdowns in robotic services. *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (2010), 203–210.

[56] Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. 2020. BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, Online, 7871–7880. https://doi.org/10.18653/v1/2020.acl-main.703

[57] Ching Liu, Juho Kim, and Hao-Chuan Wang. 2018. ConceptScape: Collaborative Concept Mapping for Video Learning. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (2018).

[58] Michael Xieyang Liu, Jane Hsieh, Nathan Hahn, Angelina Zhou, Emily Deng, Shaun Burley, Cynthia Taylor, Aniket Kittur, and Brad A Myers. 2019. Unakite: Scaffolding developers' decision-making using the web. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*. 67–80.

[59] Michael Xieyang Liu, Aniket Kittur, and Brad A. Myers. 2022. Crystalline: Lowering the Cost for Developers to Collect and Organize Information for Decision Making. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 68, 16 pages. https://doi.org/10.1145/3491102.3501968

[60] Michael Xieyang Liu, Andrew Kuznetsov, Yongsung Kim, Joseph Chee Chang, Aniket Kittur, and Brad A. Myers. 2022. Wigglite: Low-Cost Information Collection and Triage. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology* (Bend, OR, USA) *(UIST '22)*. Association for Computing Machinery, New York, NY, USA, Article 32, 16 pages. https://doi.org/10.1145/3526113.3545661

[61] Kyle Lo, Lucy Lu Wang, Mark Neumann, Rodney Kinney, and Daniel Weld. 2020. S2ORC: The Semantic Scholar Open Research Corpus. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, Online, 4969–4983. https://doi.org/10.18653/v1/2020.acl-main.447

[62] Kurt Luther, Nathan Hahn, Steven P Dow, and Aniket Kittur. 2015. Crowdlines: Supporting synthesis of diverse information sources through crowdsourced outlines. In *Third AAAI Conference on Human Computation and Crowdsourcing*.

[63] Jock D. Mackinlay, Ramana Rao, and Stuart K. Card. 1995. An organic user interface for searching citation links. In *CHI '95*.

[64] Daniel Maliniak, Ryan Powers, and Barbara F. Walter. 2013. The Gender Citation Gap in International Relations. *International Organization* 67 (2013), 889 – 922.

[65] Richard E. Mayer and Roxana Moreno. 1998. A Cognitive Theory of Multimedia Learning: Implications for Design Principles. In *CHI 1998*.

[66] Richard E. Mayer and Roxana Moreno. 1998. A Split-Attention Effect in Multimedia Learning: Evidence for Dual Processing Systems in Working Memory. *Journal of Educational Psychology* 90 (1998), 312–320.

[67] Erin C McKiernan, Philip E Bourne, C Titus Brown, Stuart Buck, Amye Kenall, Jennifer Lin, Damon McDougall, Brian A Nosek, Karthik Ram, Courtney K Soderberg, et al. 2016. How open science helps researchers succeed. *elife* 5 (2016).

[68] Gerry McKiernan. 2000. arXiv.org: the Los Alamos National Laboratory e-print server. *International Journal on Grey Literature* (2000).

[69] Sonia K Murthy, Kyle Lo, Daniel King, Chandra Bhagavatula, Bailey Kuehl, Sophie Johnson, Jonathan Borchardt, Daniel S Weld, Tom Hope, and Doug Downey. 2022. ACCoRD: A Multi-Document Approach to Generating Diverse Descriptions of Scientific Concepts. *arXiv preprint arXiv:2205.06982* (2022).

[70] Preslav I Nakov, Ariel S Schwartz, Marti Hearst, et al. 2004. Citances: Citation sentences for semantic analysis of bioscience text. In *Proceedings of the SIGIR*, Vol. 4. Citeseer, 81–88.

[71] Keisuke Okamura. 2019. Interdisciplinarity revisited: evidence for research impact and dynamism. *Palgrave Communications* 5, 1 (2019), 1–9.

[72] Yulia Otmakhova, Karin Verspoor, Timothy Baldwin, and Jey Han Lau. 2022. The patient is more dead than alive: exploring the current state of the multi-document summarisation of the biomedical literature. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, Dublin, Ireland, 5098–5111. https://doi.org/10.18653/v1/2022.acl-long.350

[73] Srishti Palani, Aakanksha Naik, Doug Downey, Amy X. Zhang, Jonathan Bragg, and Joseph Chee Chang. 2023. Relatedly: Scaffolding Literature Reviews with Existing Related Work Sections. 22 (2023).

[74] Carole L Palmer, Lauren C Teffeau, and Carrie M Pirmann. 2009. Scholarly information practices in the online environment. *Report commissioned by OCLC Research. Published online at: www. oclc. org/programs/publications/reports/2009-02. pdf* (2009).

[75] Soya Park, Jonathan Bragg, Michael Chang, Kevin Larson, and Danielle Bragg. 2022. Exploring Team-Sourced Hyperlinks to Address Navigation Challenges for Low-Vision Readers of Scientific Papers. In *Proceedings of the 25th ACM Conference On Computer-Supported Cooperative Work And Social Computing*.

[76] Amy Pavel, Colorado Reed, Björn Hartmann, and Maneesh Agrawala. 2014. Video digests: a browsable, skimmable format for informational lecture videos. *Proceedings of the 27th annual ACM symposium on User interface software and technology* (2014).

[77] Silvio Peroni, Alexander Dutton, Tanya Gray, and David M. Shotton. 2015. Setting our bibliographic references free: towards open citation data. *J. Documentation* 71 (2015), 253–277.

[78] Simon Philip, Peter Bamidele Shola, and Abari Ovye John. 2014. Application of Content-Based Approach in Research Paper Recommendation System for a Digital Library. *International Journal of Advanced Computer Science and Applications* 5 (2014).

[79] Peter Pirolli and Stuart Card. 1999. Information Foraging. *Psychological Review* 106, 4 (1999), 643–675. https://doi.org/10.1037/0033-295x.106.4.643

[80] Antoine Ponsard, Francisco Escalona, and Tamara Munzner. 2016. PaperQuest: A Visualization Tool to Support Literature Review. *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (2016).

[81] Jason Portenoy, Marissa Radensky, Jevin D West, Eric Horvitz, Daniel S Weld, and Tom Hope. 2022. Bursting scientific filter bubbles: Boosting innovation via novel author discovery. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–13.

[82] Napol Rachatasumrit, Jonathan Bragg, Amy X. Zhang, and Daniel S. Weld. 2022. CiteRead: Integrating Localized Citation Contexts into Scientific Paper Reading. *27th International Conference on Intelligent User Interfaces* (2022).

[83] Colin Raffel, Noam M. Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. 2019. Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. *ArXiv* abs/1910.10683 (2019).

[84] Daniel M. Russell, Mark J. Stefik, Peter Pirolli, and Stuart K. Card. 1993. The Cost Structure of Sensemaking. In *Proceedings of the INTERACT '93 and CHI '93*

Conference on Human Factors in Computing Systems (Amsterdam, The Netherlands) *(CHI '93)*. Association for Computing Machinery, New York, NY, USA, 269–276. https://doi.org/10.1145/169059.169209

[85] Dafna Shahaf, Carlos Guestrin, and Eric Horvitz. 2012. Metro maps of science. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. 1122–1130.

[86] 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. *Transactions of the Association for Computational Linguistics* 10 (2022), 376–392. https://doi.org/10.1162/tacl_a_00466

[87] Ben Shneiderman. 2022. *Human-centered AI.* Oxford University Press.

[88] Hariharan Subramonyam, Colleen Seifert, Priti Shah, and Eytan Adar. 2020. TexSketch: Active Diagramming through Pen-and-Ink Annotations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376155

[89] Kazunari Sugiyama and Min-Yen Kan. 2010. Scholarly paper recommendation via user's recent research interests. In *JCDL '10*.

[90] Sarit Felicia Anais Szpiro, Shafeka Hashash, Yuhang Zhao, and Shiri Azenkot. 2016. How People with Low Vision Access Computing Devices: Understanding Challenges and Opportunities. *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility* (2016).

[91] Semantic Reader team. Unpublished demo application. Paper to HTML. https://papertohtml.org

[92] Semantic Reader team. Unpublished demo application; in submission.. Papeo. https://papeo.app

[93] Anh Tuan Truong, Peggy Chi, D. Salesin, Irfan Essa, and Maneesh Agrawala. 2021. Automatic Generation of Two-Level Hierarchical Tutorials from Instructional Makeup Videos. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (2021).

[94] Richard Van Noorden et al. 2015. Interdisciplinary research by the numbers. *Nature* 525, 7569 (2015), 306–307.

[95] David Wadden, Shanchuan Lin, Kyle Lo, Lucy Lu Wang, Madeleine van Zuylen, Arman Cohan, and Hannaneh Hajishirzi. 2020. Fact or Fiction: Verifying Scientific Claims. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics, Online, 7534–7550. https://doi.org/10.18653/v1/2020.emnlp-main.609

[96] Chong Wang and David M. Blei. 2011. Collaborative topic modeling for recommending scientific articles. In *KDD*.

[97] Lucy Lu Wang, Isabel Cachola, Jonathan Bragg, Evie Yu-Yen Cheng, Chelsea Haupt, Matt Latzke, Bailey Kuehl, Madeleine N van Zuylen, Linda Wagner, and Daniel Weld. 2021. SciA11y: Converting Scientific Papers to Accessible HTML. In *Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility* (Virtual Event, USA) *(ASSETS '21)*. Association for Computing Machinery, New York, NY, USA, Article 85, 4 pages. https://doi.org/10.1145/3441852.3476545

[98] Lucy Lu Wang, Isabel Cachola, Jonathan Bragg, Evie Yu-Yen Cheng, Chelsea Hess Haupt, Matt Latzke, Bailey Kuehl, Madeleine van Zuylen, Linda Wagner, and Daniel S. Weld. 2021. Improving the Accessibility of Scientific Documents: Current State, User Needs, and a System Solution to Enhance Scientific PDF Accessibility for Blind and Low Vision Users. *arXiv: 2105.00076 [cs.DL]* (2021). arXiv:2105.00076 [cs.DL]

[99] Wenhui Wang, Hangbo Bao, Shaohan Huang, Li Dong, and Furu Wei. 2021. MiniLMv2: Multi-Head Self-Attention Relation Distillation for Compressing Pretrained Transformers. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*. Association for Computational Linguistics, Online, 2140–2151. https://doi.org/10.18653/v1/2021.findings-acl.188

[100] Wenhui Wang, Furu Wei, Li Dong, Hangbo Bao, Nan Yang, and Ming Zhou. 2020. MiniLM: Deep Self-Attention Distillation for Task-Agnostic Compression of Pre-Trained Transformers. In *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin (Eds.), Vol. 33. Curran Associates, Inc., 5776–5788. https://proceedings.neurips.cc/paper_files/paper/2020/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

[101] Samuel F Way, Allison C Morgan, Daniel B Larremore, and Aaron Clauset. 2019. Productivity, prominence, and the effects of academic environment. *Proceedings of the National Academy of Sciences* 116, 22 (2019), 10729–10733.

[102] Alan J. Wecker, Joel Lanir, Osnat Mokryn, Einat Minkov, and Tsvi Kuflik. 2014. Semantize: visualizing the sentiment of individual document. In *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces*. Association for Computing Machinery, Como, Italy, 385–386.

[103] Feng Xia, Haifeng Liu, Ivan Lee, and Longbing Cao. 2016. Scientific Article Recommendation: Exploiting Common Author Relations and Historical Preferences. *IEEE Transactions on Big Data* 2 (2016), 101–G112.

[104] Yiheng Xu, Minghao Li, Lei Cui, Shaohan Huang, Furu Wei, and Ming Zhou. 2019. LayoutLM: Pre-training of Text and Layout for Document Image Understanding. *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (2019).

[105] Qian Yang, Gerard de Melo, Yong Cheng, and Sen Wang. 2017. HiText: Text Reading with Dynamic Salience Marking. In *Proceedings of the 26th International Conference on World Wide Web Companion*. Association for Computing Machinery, Perth, Australia, 311–319.

[106] Wonjin Yoon, Jinhyuk Lee, Donghyeon Kim, Minbyul Jeong, and Jaewoo Kang. 2019. Pre-trained Language Model for Biomedical Question Answering. In *PKDD/ECML Workshops*.

[107] Polle T. Zellweger, Bay-Wei Chang, and Jock D. Mackinlay. 1998. Fluid Links for Informed and Incremental Link Transitions. In *Proceedings of the Conference on Hypertext and Hypermedia*. ACM, 50–57.

[108] Tian Zhao and Kyusong Lee. 2020. Talk to Papers: Bringing Neural Question Answering to Academic Search. *ArXiv* abs/2004.02002 (2020).

[109] Sacha Zyto, David Karger, Mark Ackerman, and Sanjoy Mahajan. 2012. Successful classroom deployment of a social document annotation system. In *Proceedings of the sigchi conference on human factors in computing systems*. 1883–1892.