Toward Structured Experimental Layers for Augmented Reading: A Hybrid AI-Assisted Representation of Scientific Logic

Myung Jin (MJ) Kim
Electronics and Telecommunications Research Institute
Daejeon, Republic of Korea
mj@etri.re.kr

Abstract

Scientific papers communicate structured experimental reasoning, including research questions, hypotheses, variables, metrics, and limitations, yet this structure remains embedded in narrative prose. During literature review and peer review, researchers reconstruct this logic manually across papers, while AI-mediated reading tools attempt to infer it through summarization and extraction. As AI becomes increasingly integrated into scholarly workflows, we argue that the experimental structure implicit in scientific writing merits reconsideration as a design concern for augmented reading. We outline a design space for a hybrid, AI-assisted structured layer that complements narrative articles by making elements of experimental logic explicit and inspectable. In this model, AI systems propose structured representations such as hypotheses, study designs, and evaluation metrics, and authors verify or refine them. Such a layer could support cross-paper comparison, clarify reasoning during review, and enable accessibility-oriented transformations while preserving rhetorical nuance and authorial voice. We also discuss risks including oversimplification, superficial compliance, and misplaced trust in automated extraction, and we identify open research questions for evaluating hybrid structured workflows in scientific practice.

CCS Concepts

• **Human-centered computing** → **Interactive systems and tools**; *User interface design*; Visualization; Accessibility systems and tools.

Keywords

augmented reading, structured experimental logic, hybrid AI-human workflows, scholarly communication, accessibility, peer review

1 Introduction

AI-mediated reading is increasingly integrated into scholarly workflows. Researchers use large language models to summarize papers, generate outlines, and answer topical questions, while literature search and citation discovery tools assist in mapping related work. Recent research in HCI has begun to reexamine the reading interface itself. For example, the Semantic Reader Project explores how AI can layer discovery aids, synthesis tools, and accessibility features onto traditional PDFs [6].

At the same time, large bodies of scientific articles, particularly in fields such as HCI, psychology, and applied computer science, continue to be written primarily as narrative prose. Within these narratives, the core structure of experimental reasoning, including research questions, hypotheses, variables, metrics, and limitations, is present but not explicitly encoded in a machine-readable form.

In practice, researchers rarely read papers from beginning to end with equal attention. Instead, they move between sections and extract specific components relevant to their goals. During literature review or peer review, they often reconstruct experimental logic across documents by asking questions such as: What was tested? Under what conditions? Compared to what baseline? Using which metrics?

This reconstructive work is not new. Early augmented reading systems such as Beyond Paper emphasized annotation as a way for readers to externalize understanding while navigating complex documents [9]. Later systems like LiquidText supported spatial rearrangement of excerpts, allowing readers to cluster related passages and build cross-document comparisons [10]. These systems acknowledge that readers actively reorganize material to make sense of it. However, the experimental structure itself remains embedded in prose.

Today, AI systems attempt to approximate this structure through extraction and question answering over structured documents. For example, PDFTriage models document structure to support question answering across long PDFs [8]. Even so, such systems infer structure from prose rather than relying on components that are explicitly authored and encoded. As a result, omissions or ambiguities can be difficult to detect.

We argue that as AI becomes more deeply integrated into scholarly reading workflows, the implicit structure of experimental reasoning deserves renewed attention as a design concern for augmented reading. We explore whether a hybrid, AI-assisted structured layer could complement narrative articles by making elements of experimental logic explicit while preserving flexibility, nuance, and authorial voice.

2 Observations from Experimental Research Practice

This position grows out of routine research practice rather than formal reading theory.

First, literature review frequently involves cross-paper structural comparison. Researchers extract and align independent and dependent variables, participant characteristics, task designs, evaluation metrics, and reported limitations. Although many papers follow formats such as IMRaD, these elements are expressed through varied language and rhetorical styles. Differences in terminology, reporting detail, and writing tone require manual interpretation.

Recent intelligent skimming systems acknowledge this challenge. Scim, for example, highlights sentences corresponding to rhetorical facets such as objectives, novelty, methods, and results, helping readers build a high-level understanding of a paper more efficiently [1]. However, these systems still operate by predicting

structure from prose. The elements they surface are inferred and visually marked, but they are not explicitly authored or validated as structured components. Experimental details therefore remain embedded and may be more or less visible depending on writing style, which reflects diversity of expression rather than error.

Second, peer review often requires reconstructing experimental logic under time constraints. Reviewers assess whether claims are supported by methods and results and whether metrics and limitations are appropriate. When key concepts are distributed across sections, cognitive effort increases. Systems such as ScholarPhi illustrate how localized structural augmentation, for example surfacing definitions of technical terms at the point of use, can reduce this burden without rewriting the document [4]. This suggests that additional structural cues can improve comprehension while preserving the integrity of the original text.

Third, AI tools increasingly mediate this process through summarization. Summaries can omit details or misrepresent emphasis. Recent work on traceable text shows that linking generated summaries back to specific source passages improves reader trust and supports error detection [5]. This line of work underscores a broader point: when structure is generated implicitly, transparency and traceability become essential.

Taken together, these developments suggest that the experimental structure researchers routinely reconstruct might be made more explicit. At the same time, narrative exposition remains essential. It communicates motivation, interpretation, and contextual framing that cannot be reduced to structured fields alone. The question, then, is not whether narrative should be replaced, but how it might be complemented.

3 A Hybrid Structured Layer for Experimental Logic

We outline a design space for a structured layer that complements narrative articles.

Such a layer could encode core elements of experimental reasoning, including research questions or objectives, hypotheses when applicable, independent and dependent variables, study design and conditions, evaluation metrics, sample characteristics, baselines, and stated limitations. These examples are illustrative rather than exhaustive.

We do not propose rigid standardization. Instead, we envision a hybrid workflow in which AI systems generate draft structural representations from a manuscript and authors verify, refine, or reject these suggestions before publication. Community feedback mechanisms could also allow readers to flag inconsistencies or suggest refinements in published work.

Similar to widely adopted tools such as grammar checking or reference extraction, AI assistance could reduce effort while preserving author control. The resulting structured layer would remain inspectable and editable, serving as a complementary representation of experimental logic rather than a substitute for prose argumentation.

4 Implications for Augmented Reading

Making experimental structure explicit may broaden the design space of augmented reading interfaces.

Cross-paper comparison. Structured representations could support systematic alignment of variables, metrics, and study designs across papers, potentially reducing manual reconstruction during literature review. Systems such as CiteRead demonstrate how integrating citation context into the flow of a paper can support cross-document reasoning [7]. A structured experimental layer could extend this approach to methodological comparison.

Clarification of reasoning. Explicit encoding may help expose gaps between claims and supporting evidence during writing or review. Work on AI-resilient text rendering, such as Grammar-Preserving Text Saliency Modulation, shows how interfaces can increase salience while maintaining fidelity to source text [2]. A structured layer could similarly increase the visibility of methodological components without altering narrative content.

Accessibility and modality translation. Structured representations may also enable alternative renderings of scientific content. SciA11y demonstrates how reorganizing PDF papers into semantically structured HTML improves navigation and accessibility for blind and low-vision readers [11]. A machine-readable experimental schema could further support structured audio summaries, navigable method overviews, or adaptive modality transformations that rely less on inference.

These possibilities remain speculative and require empirical validation.

5 Risks and Open Questions

Formalizing experimental structure introduces risks.

Over-standardization may oversimplify nuanced reasoning or privilege certain paradigms, particularly quantitative experimental designs, over qualitative or exploratory approaches. While study protocols may vary creatively, many experimental principles remain consistent. The challenge is to encode structure without constraining legitimate diversity of method.

Structured fields could incentivize superficial compliance, for example authors filling required fields mechanically without improving clarity. AI-generated representations may contain errors, leading to misplaced trust if not carefully verified. Additional author burden may hinder adoption, although community norms such as the adoption of ORCID identifiers show that reporting practices can evolve when perceived benefits outweigh costs.

Key research questions include:

- Does explicit structural encoding measurably improve cross-paper comparison accuracy or efficiency, and under what conditions might it fail?
- How reliably can AI systems extract experimental components, and how should uncertainty be communicated?
- How would hybrid structured workflows influence peer review?
- Can structured layers support accessibility without reducing rich interpretation and contextual nuance to overly simplified schema?

These questions could first be explored through small-scale prototypes and controlled studies before broader adoption is considered.

6 Discussion

As AI systems increasingly mediate scholarly reading workflows, the relationship between narrative text and structured reasoning becomes more apparent. Augmented scholarly interfaces already demonstrate the value of layering additional representations onto traditional articles [6]. We suggest that the experimental logic embedded in scientific writing may benefit from similar treatment.

This proposal is not without precedent. Structured abstracts and reporting guidelines in medicine were introduced to standardize the reporting of key study components and improve clarity and retrieval [3]. These efforts showed that explicitly structuring aims, methods, and results can improve informativeness and usability. However, structured abstracts primarily organize summaries for human readers. Contemporary AI-mediated environments create a different opportunity: to layer machine-operable structural representations alongside full narrative articles so that experimental logic is accessible to both people and computational systems.

Rather than prescribing reform, we offer this as a design provocation for the augmented reading community. The open question is whether hybrid, AI-assisted structural encoding can improve accessibility and comparability while preserving epistemic diversity and rhetorical richness.

References

[1] Raymond Fok, Hita Kambhamettu, Luca Soldaini, Jonathan Bragg, Kyle Lo, Marti Hearst, Andrew Head, and Daniel S Weld. 2023. Scim: Intelligent Skimming Support for Scientific Papers. In *Proceedings of the 28th International Conference on Intelligent User Interfaces (IUI '23)*. Association for Computing Machinery, New York, NY, USA, 476–490. doi:10.1145/3581641.3584034

[2] Ziwei Gu, Ian Arawjo, Kenneth Li, Jonathan K. Kummerfeld, and Elena L. Glassman. 2024. An AI-Resilient Text Rendering Technique for Reading and Skimming Documents. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–22. doi:10.1145/3613904.3642699

[3] James Hartley. 2004. Current Findings from Research on Structured Abstracts. *Journal of the Medical Library Association* 92, 3 (July 2004), 368–371.

[4] 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. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–18. doi:10.1145/3411764.3445648

[5] Hita Kambhamettu, Jamie Flores, and Andrew Head. 2025. Traceable Texts and Their Effects: A Study of Summary-Source Links in AI-Generated Summaries. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25)*. Association for Computing Machinery, New York, NY, USA, 1–7. doi:10.1145/3706599.3719830

[6] Kyle Lo, Joseph Chee Chang, Andrew Head, Jonathan Bragg, Amy X. Zhang, Cassidy Trier, Chloe Anastasiades, Tal August, Russell Authur, Danielle Bragg, Erin Bransom, Isabel Cachola, Stefan Candra, Yoganand Chandrasekhar, Yen-Sung Chen, Evie Yu-Yen Cheng, Yvonne Chou, Doug Downey, Rob Evans, Raymond Fok, Fangzhou Hu, Regan Huff, Dongyeop Kang, Tae Soo Kim, Rodney Kinney, Aniket Kittur, Hyeonsu B. Kang, Egor Klevak, Bailey Kuehl, Michael J. Langan, Matt Latzke, Jaron Lochner, Kelsey MacMillan, Eric Marsh, Tyler Murray, Aakanksha Naik, Ngoc-Uyen Nguyen, Srishti Palani, Soya Park, Caroline Paulic, Napol Rachatasumrit, Smita Rao, Paul Sayre, Zejiang Shen, Pao Siangliulue, Luca Soldaini, Huy Tran, Madeleine van Zuylen, Lucy Lu Wang, Christopher Wilhelm, Caroline Wu, Jiangjiang Yang, Angele Zamarron, Marti A. Hearst, and Daniel S. Weld. 2024. The Semantic Reader Project. *Commun. ACM* 67, 10 (2024), 50–61. doi:10.1145/3659096

[7] Napol Rachatasumrit, Jonathan Bragg, Amy X. Zhang, and Daniel S Weld. 2022. CiteRead: Integrating Localized Citation Contexts into Scientific Paper Reading. In *Proceedings of the 27th International Conference on Intelligent User Interfaces (IUI '22)*. Association for Computing Machinery, New York, NY, USA, 707–719. doi:10.1145/3490099.3511162

[8] Jon Saad-Falcon, Joe Barrow, Alexa Siu, Ani Nenkova, David Seunghyun Yoon, Ryan A. Rossi, and Franck Dernoncourt. 2023. PDFTriage: Question Answering over Long, Structured Documents. arXiv:2309.08872 [cs] doi:10.48550/arXiv.2309.08872

[9] Bill N. Schilit, Gene Golovchinsky, and Morgan N. Price. 1998. Beyond Paper: Supporting Active Reading with Free Form Digital Ink Annotations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '98)*. ACM Press/Addison-Wesley Publishing Co., USA, 249–256. doi:10.1145/274644.274680

[10] Craig S. Tashman and W. Keith Edwards. 2011. LiquidText: A Flexible, Multitouch Environment to Support Active Reading. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. Association for Computing Machinery, New York, NY, USA, 3285–3294. doi:10.1145/1978942.1979430

[11] 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 (ASSETS '21)*. Association for Computing Machinery, New York, NY, USA, 1–4. doi:10.1145/3441852.3476545