Qeios

Peer Review

# Review of: "An AI-Human Collaborative Review of AI-Human Collaborative Reviews"

**Iliano Fasolino**[1]

1. Department of Computer Science, University of Milan, Italy

The reflexive move is the interesting part of this paper. A narrative review of human–AI scholarly writing that discloses its own tool chain—Ai2 Paperfinder for the seed query, ResearchRabbit for expansion, NotebookLM for extraction, with the prompts printed—and that refuses to let those tools author the prose, is a useful object to study. The Abstract's line that "AI tools were not used for content authoring" sets a clear boundary. Pathak (JFB5FZ) treated the piece mainly as a timely overview. Arsentev (6HHPN7) already documented the missing corpus counts, the "consensus" built on a handful of sources, the "Null" journal field on Lee [11], the mid-sentence "In conclusion," fragment, the gap between endorsing PaperCard and not meeting its version/date fields, the health-scope mismatch, the title-vs-corpus mismatch, and—in his §2.5—the overbroad hallucination-freedom claim plus the unreported manual-vs-NotebookLM disagreement count. The author has not replied. Those asks remain unmet. I agree with Arsentev's §2.5 punch and will not re-own it; what I add is how that missing count lands under the title's own promise and under the paper's Computer Science register.

Materials and Methods asserts that the workflow is "free from the risk of producing fictitious or hallucinatory research publications," on three grounds: Semantic Scholar backing for Paperfinder, DOI checks in ResearchRabbit, and NotebookLM reading only uploaded PDFs. As Arsentev noted, those three controls speak to one failure mode—citing a paper that does not exist—not to NotebookLM (or the human summarizer) misstating what a real PDF argued. The text says every upload was manually read and checked against the tool. That is the right control. The result of the control is never given. Even zero disagreements on a dozen extractions—if that is the corpus size—would be evidence; an unreported check is not. For a manuscript whose title presents itself as a demonstration of collaborative review

 doi.org/10.32388/6TUEPO 

practice, that missing count is the finding the demonstration needed, and its absence is the finding the reader gets instead.

The same circularity shows up in the confidence language. The Discussion opens by pairing “human creativity” with “the infallible accuracy of a software tool.” A few pages earlier, the paper itself walks through fabricated references, paper mills, red teaming, and context-induced bias. Grammar and formatting can be mechanically consistent; calling that infallible accuracy is the wrong word in a Computer Science submission that spends a section on hallucination. Soften it to what was meant, and keep the caution and the speed claim in the same register.

I also worry about how much of the Results is theme-labeling rather than appraisal. “In the corpus of studies reviewed, there was a consensus…” starts the Benefits subsection, then four citations follow, one of which the author correctly disqualifies as AI-as-sole-author rather than collaboration. “Unanimous ‘no’” on AI co-authorship rests on Lee [11] alone in that subsection, while ICMJE/COPE-style policy is only gestured at later via the STM material. Arsentev already asked for counts and a study table; I will add that without design labels (empirical study vs commentary vs editorial) and without any reported exclusion, the word consensus is doing editorial work the search cannot support. A single Paperfinder prompt that itself says “generalised pre-trained” rather than “generative,” with no date limit and no second database, is a thin net. The author admits the thinness in the closing limitations paragraph. That admission should pull the Abstract back from “excellent exemplars … improving the flow and quality of the output,” which this study never measured.

PaperCard dates/versions, the health frame, the title’s second half, and the broken “In conclusion,” line remain unmet after Arsentev; author replies still 0. I will not re-brief them.

I would keep the disclosed tool chain and the refusal to let AI draft the prose. I would replace the hallucination-freedom sentence with a narrow claim about indexed citations plus a reported manual-vs-NotebookLM disagreement count (Arsentev’s ask), cut “infallible” and “consensus” / “unanimous” where the citations cannot bear them, and publish a one-page included-study table with retrieval counts and access dates. None of that needs a new literature pass; it needs the author to report what the methods paragraph says was already done.

## Declarations

**Potential competing interests:** No potential competing interests to declare.