Documentation Traps in Classroom Settings

Sujit Kumar Kamaraj
University of Utah
sujit.kamaraj@utah.edu

Vineet Pandey
University of Utah
vineet.pandey@utah.edu

Abstract

Group projects are increasingly shaped by AI tools that let students revise and reframe each other's work *invisibly*. Existing tools–shared text documents, feedback via comment threads, self-reported AI disclosures–offer limited structured accounts of how group submissions evolve. One solution to this problem is building better provenance tracking tools that record students' work explicitly for instructors to see. However, such efforts fail to track contributions that happened outside the tracking environment. Yet these are often the moments where a student's understanding–or lack of it–is formed, making their invisibility a pedagogical problem. These untracked contributions include silent agreement to a teammate's revision, judgment exercised in declining an AI suggestion, and negotiation that resolves which group member's work is used: moments that shape a group's work as much as any visible edit. We identify five such instances where provenance tracking tools might fail, distinguishing what such tools can document from other *ways of knowing*. We pose open questions about whether such tools should track more, restructure collaboration to be legible, or whether instructors need entirely different ways to assess group contribution, such as conversations with the group.

CCS Concepts

• **Human-centered computing → Collaborative and social computing**.

Keywords

Contribution tracking, Group projects, Human-AI collaboration

1 "Nobody Did This": A Vignette

Maya, Jordan, and Sev are assigned a group project in an introductory HCI class: identify five ideas for solving a design challenge and create paper prototypes for each. Sev starts a shared document and drafts five initial ideas. Overnight, Jordan takes her outline, feeds it to an AI assistant, and asks it to "make the ideas stronger". The AI restructures the five ideas around efficient computation instead of Sev's original framing of user experience. Jordan likes the new framing and pastes it back into the document. The document's comment thread, where the group usually flags changes to each other, stays empty: there was no feature prompting Jordan to note that the paragraph came from an AI conversation. The next morning, Sev assumes Jordan developed the text himself and says nothing. Maya reads the new ideas, doesn't know the framing shifted, and spends an hour sketching paper prototypes around the new ideas. When they meet, Jordan cannot explain why the efficient computation framing works: he asked the AI to "make the ideas stronger" without a strong sense of what he wanted it to change. The group decides to proceed with the new framing since the deadline is here...

Two weeks later they submit a polished, coherent set of paper prototypes. Their instructor, Professor Ikhile, reviews the prototypes with no visibility into the process: five prototypes, one framing, and no sign that it began somewhere else. She has no way to ask "whose idea was the efficient-computation framing?" because as far as the document is concerned, there was never a fork in the road. She grades the group as a whole, and privately wonders whether Maya, the strongest designer, carried the project, and whether Sev's shy nature allowed her to express her ideas.

2 Why do we need to track contributions in a group project when students use AI?

Maya, Jordan, and Sev's submission never records the process through which their ideas evolved: from Sev, through Jordan's prompt, through AI's reframing, into the version the whole team built on. The lack of tracking is where the group's shared understanding of "what we're framing and why" detached from its origin. Without a structured way to represent how the submission evolved, Maya spent an hour sketching paper prototypes on Jordan's AI-generated ideas with no knowledge of Sev's initial ideas. Sev, Jordan, and Maya's submission is not an unusual case: it's the result of how group work in classrooms is structured today.

Group projects and assignments are increasingly asynchronous and distributed. Students contribute ideas at different times, build on each other's work, and hand off partially-formed concepts across sessions and team members. Yet the tools available to them–shared documents, chat threads, comment threads in software applications–offer no structured way to represent how a submission evolved or what earlier versions were considered. The lack of support might force students into continuing other's work in a superficial manner: they can see what a teammate wrote, but not why, what was tried before, and which directions were abandoned. The result is misinterpretation, redundant exploration, and work that might fail to develop beyond their initial articulation. The indiscriminate use of AI by students makes tracking their contributions on group projects both urgent and extremely difficult. As AI tools not designed for pedagogical purposes become more capable and see widespread use by students, instructors are increasingly unable to distinguish students' intellectual growth from mindless delegation to AI that leads to an illusion of learning [4, 7]

3 How can we track contributions in a group project when students use AI?

One common response to the challenge of invisible AI contributions is asking students to declare what AI contributed (e.g. when submitting an assignment) and subsequently conducting in-person exams to test their learning. HCI research consistently shows that self-reports are notoriously unreliable: people misremember, underreport, and rationalize after the fact [5, 9]. Other approaches include copy-pasting prompt histories (but it takes additional manual student effort), screenshotting interaction logs (they are difficult to parse/analyze at scale), or tracking low-level behavioral data like keystrokes (but it raises serious privacy concerns). When students' actual contributions become visible, instructors can reliably and effortlessly observe–rather than guess–the level of AI use by students–from no AI use to full AI use without human oversight [6].

Better provenance tracking tools can solve some of these challenges. First, effective tracking would need to ensure students' use of AI is tracked more transparently without disrupting or changing students' workflow. Second, tracking tools need to make AI contributions directly observable to instructors, enabling systematic assessments that are more accurate and require less instructor effort than current hacky idiosyncratic approaches. In the learning sciences and HCI, process tracing methods–such as think-aloud protocols, writing logs–make cognitive processes visible during a task [2, 3]. Making the learning process visible has intrinsic educational value: Schön's concept of reflection-on-action holds that examining one's own process is itself a mechanism for learning [8]; Bandura's social learning theory mentions that observing others' processes is a powerful source of learning [1]. However, such a tool tracks documented provenance (what happened, in what order), and not witnessed contribution (whether a choice reflected judgment, negotiation, or restraint that others could see and hold someone to).

4 What contributions does provenance tracking fail to track?

Provenance tracking tools would track contributions made directly to a document or submission but might fail to track *indirect contributions.* We identify five places such tools fail to track contribution in group work.

1) Off-platform deliberation: Group work in practice happens across multiple spaces and provenance tracking tools might only track some of them. Two or more students can discuss an AI generated idea verbally, over a call, in another shared document, or in a text thread outside the tracked environment, then one of them makes the change in the tracked document, and only that final edit gets recorded. The documented provenance shows one contributor and one iteration from a prior version, with no trace that the idea was jointly discussed before it became a new version.

2) Internalized inspiration: Some contributions can occur internally and might not be tracked as a contribution. For example, a student might take inspiration from something the AI mentioned, think about it by themselves, and make changes based on their internal iteration. Provenance tracking tools will mark such an interaction as the student's contribution without tracking the original inspiration AI provided.

3) Invisible judgment: When a student evaluates an AI suggestion and declines it–recognizing it's wrong or weak–that judgment is an intellectual contribution, but it leaves no artifact and therefore no contribution is tracked. Provenance tracking tools track what was kept and not the reasoning that filtered out what wasn't. A student doing careful curation looks identical, in the documented provenance, to a student who never questioned anything. Invisible judgment has a pedagogical cost as well: declining a flawed suggestion is often the clearest evidence of a student's developing judgment; but such a moment leaves no trace for an instructor to recognize or reinforce.

4) Negotiated synthesis: Group work might involve asymmetric contribution leading to a single new contribution being tracked. When two students' positions merge into one idea, the documented provenance can show that a merge happened but not the social process behind it. Provenance tracking tools cannot track who conceded, whose framing dominated, and whether the synthesis was actually one person's idea with some minor ideas borrowed from others.

5) Non-artifact contributions: Group work involves multiple contributions that do not change the artifact. For example, group work requires resolving conflicts, keeping the group on schedule, assigning responsibilities, and deciding on the path followed by the group. All of these contributions shape the final work substantially but never get recorded since they do not produce versions of the assignment itself. A documented provenance built around artifact evolution structurally cannot represent contributions that are not artifact-producing.

Each of these five gaps names a moment where a real contribution existed but was never witnessed, in the sense the workshop uses the term: produced under conditions where others could recognize it as an act of judgment, respond to it, and hold someone answerable for it. These gaps are not missing entries tools fail to log; they are instances of accountability that documentation was never designed to track. We identify these gaps in relation to an educational context, however, we believe that all five gaps can occur in multiple other contexts. For example, teams working in an office might discuss details outside of workspaces leading to *off-platform deliberation*, and/or *negotiated synthesis*.

5 Open Questions for the Workshop

Based on the gaps identified, we would like to present three open questions that we believe will enrich discussions during the workshop.

Provenance tracking tools' failure modes are not obviously fixable by adding more granular tracking or tracking more environments. A student's decision to reject an AI suggestion is a cognitive event, not an artifact event; there may be no version-history structure, however fine-grained, that tracks it without asking the student to narrate their own thinking, which reintroduces the self-report problem such tools exist to avoid. If such tools are the wrong solution for tracking such contributions, **how can we avoid the self-report problem when social situations are reintroduced?** One way forward could be to ask students to anchor their decisions with concrete evidence (e.g., from the data or prior literature). Such solutions might create additional work for students and instructors.

One response to provenance tracking tools' blind spots is tracking more surfaces, and logging more channels. A different response is procedural: structure group work so that negotiation and synthesis happen on the record by design (e.g., requiring disagreements to be resolved in a tracked space rather than a side conversation). This trades some of the messiness of real collaboration for legibility, which might raise its own cost. We'd like the workshop's help answering: **should accountability infrastructures try to**

match how groups actually work, or push that work toward forms that can be witnessed? In an educational setting, forcing negotiation onto the record may double as a learning scaffold, since externalizing one's reasoning to a teammate is itself a mechanism by which learning happens.

Instructors' dilemma isn't only that they lack information; it's that they are structurally outside the situation in which the contribution was formed. We're interested in whether accountability in educational settings needs a different forum altogether: **how can instructors understand a groups' contributions without relying on self-reports?** An instructor's task is to assess learning along with output quality, which means a classroom accountability infrastructure needs to surface what a student contributed, and whether that contribution reflects students' learning. Peer assessment within the group might reconstruct more of the missing social context (who conceded, whose framing dominated) than any log could, but it might introduce its own incentives to misrepresent or smooth over conflict.

6 Conclusion

Taken together, these five gaps suggest that better provenance tracking–however complete–cannot on its own restore what group work has lost: a shared, situated understanding of who did what, and why. Our vignette is deliberately mundane–no one in it acts maliciously, and no single decision seems consequential in isolation–because that is how contribution dissolution happens in practice. We see the workshop as an opportunity to think collectively about whether accountability in AI-mediated group work is a tracking problem, a design problem, or a problem that requires rethinking the forums in which student contributions are assessed altogether.

References

[1] Albert Bandura and Richard H Walters. 1977. *Social learning theory*. Vol. 1. Prentice-hall Englewood Cliffs, NJ.
[2] K Anders Ericsson. 2017. Protocol analysis. *A companion to cognitive science* (2017), 425–432.
[3] Linda Flower and John R Hayes. 1981. A cognitive process theory of writing. *College Composition & Communication* 32, 4 (1981), 365–387.
[4] Hao-Ping Lee, Advait Sarkar, Lev Tankelevitch, Ian Drosos, Sean Rintel, Richard Banks, and Nicholas Wilson. 2025. The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In *Proceedings of the 2025 CHI conference on human factors in computing systems*. 1–22.
[5] Richard E Nisbett and Timothy D Wilson. 1977. Telling more than we can know: Verbal reports on mental processes. *Psychological review* 84, 3 (1977), 231.
[6] Mike Perkins, Leon Furze, Jasper Roe, and Jason MacVaugh. 2024. The Artificial Intelligence Assessment Scale (AIAS): A framework for ethical integration of generative AI in educational assessment. *Journal of University Teaching and Learning Practice* 21, 6 (2024), 49–66.
[7] James Prather, Paul Denny, Juho Leinonen, Brett A Becker, Ibrahim Albluwi, Michelle Craig, Hieke Keuning, Natalie Kiesler, Tobias Kohn, Andrew Luxton-Reilly, et al. 2023. The robots are here: Navigating the generative ai revolution in computing education. In *Proceedings of the 2023 working group reports on innovation and technology in computer science education*. 108–159.
[8] Donald A Schön. 2017. *The reflective practitioner: How professionals think in action*. Routledge.
[9] Lucille Alice Suchman. 1987. *Plans and situated actions: The problem of human-machine communication*. Cambridge university press.

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009