

How Much Can AI Understand? Toward AI-Assisted Sensemaking of Collaborative Discussion in Groups with Shared History

Soobin Cho
soobin30@uw.edu
Human Centered Design &
Engineering, University of
Washington
Seattle, USA

Mark Zachry
zachry@uw.edu
Human Centered Design &
Engineering, University of
Washington
Seattle, USA

David W. McDonald
dwmc@uw.edu
Human Centered Design &
Engineering, University of
Washington
Seattle, USA

Abstract

AI tools that support collaborative discussion typically treat the discussion as a standalone task, focusing only on its content and setting aside the social context of the group having it. But it is groups with a shared history, with their own norms, hierarchies, and relationships, where the most tangled and complex discussions tend to arise. These discussions cannot be understood apart from that context, and AI that overlooks it risks failing to convey what a discussion means, or even misrepresenting it. Drawing on two studies of how experienced Wikipedia editors read and make sense of discussions, we propose an AI-Assisted Sensemaking Model for Collaborative Discussions, which captures not only a discussion’s arguments but also the norms and participants behind it, along with the context that gives each meaning. In this model, the system supports the early stages of the sensemaking process, and the degree to which it performs interpretive work can range from low to high. We argue that higher interpretive work reduces the burden on users but increases their reliance on the system’s judgment. We then discuss the risks of an insufficiently intelligible system, what it would take to make one more intelligible, and the safeguards it still requires.

1 Motivation

Discussion is essential to how collaborative working groups get work done. It is how members understand each other’s positions, work through disagreement, and reach the decisions needed to move forward. Some discussions are short and resolve quickly, but just as many stretch on and grow complex, which can mean becoming entangled with earlier practices, or having the people involved shift over time. Anyone who has been part of a long-running collaborative working group, whether a project team at a company, a research team at a lab, or a volunteer team, has likely had to make sense of a discussion that grew complex because of a myriad of factors.

CSCW has long studied ways to support discussion in collaborative settings, with a large body of work on tools such as argument visualization [2, 7] and group decision support tools [3, 9]. As AI has advanced, research on autonomous systems that support discussion has grown as well. This includes argument mining [8, 10], which automatically extracts argumentative structure from text, and discussion facilitation bots [4–6], which are conversational agents that take on the role of a discussion facilitator. With the rise of LLMs, practical applications now include agents that generate meeting minutes, while research has investigated how to generate

consensus statements that participants in public deliberation can agree on [12].

However, these AI systems typically focus only on the given discussion content, without considering the social context around it. Research on these systems also tends to assume an ad hoc group formed only for the task at hand. But it is collaborative working groups with their own history where the most complex discussions tend to arise, and these groups’ discussions cannot be separated from their social context. An AI that attempts to assist discussion without taking this social context into consideration risks failing to fully grasp and convey the meaning that contributions within the discussion carry, or even misrepresenting them.

Groups with a shared history usually have explicit or implicit ways of working and communicating, and members differ in their experience, positions within the group’s hierarchy, and relationships with one another, all of which are naturally reflected in how a discussion unfolds. In a project team, for instance, members may know that a given argument belongs to an important stakeholder and that the group tends to defer to it. A new member joining the project would find little help in an AI that summarizes only the arguments themselves, and it would probably take a senior member quietly filling in that context before the newcomer could truly make sense of the discussion.

This raises the question of how sensemaking of discussion unfolds in collaborative working groups with a shared history, and how AI can support it.

2 AI-Assisted Sensemaking Model for Collaborative Discussions

Our understanding of how members of a collaborative working group with a shared history make sense of discussions derives from studies we conducted in the context of Wikipedia. In Wikipedia, contributors work together to evaluate content and build consensus on what should be included in encyclopedic articles. To carry out this collaborative work, editors engage in a large number of discussions, many of which run long and grow complex over time. Wikipedia is governed by an extensive and interlinked system of community-created policies and guidelines [15]. However elaborate these are, they are applied in practice with their own implicit conventions, some of which are informally documented in a space called essays [16], though these essays inevitably fall short of capturing every convention. Registered editors’ identities are reflected in their edit histories and user pages, which convey their Wikipedia experience and background, as well as in their interactions with

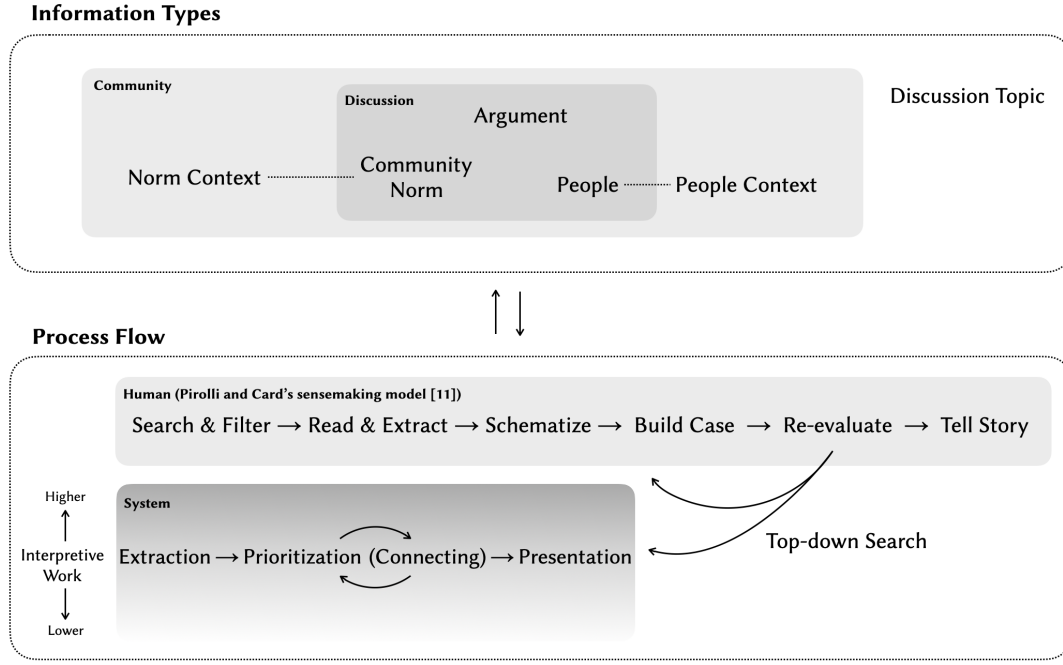


Figure 1: AI-assisted sensemaking model for collaborative discussions. Information types include argument, community norm, norm context, people, people context, and discussion topic. The model retains Pirulli and Card's sensemaking model [11] as the human component and extends it by introducing system support over the early stages of the human sensemaking process. The system performs Extraction, Prioritization, Connecting, and Presentation, operating along a spectrum from lower to higher system interpretive work.

other editors across various pages, which reflect their past relationships and social roles [14].

To understand how experienced Wikipedians read and make sense of complex discussions, we conducted two studies. The first examined their cognitive process as they read discussions unassisted, identifying the information elements they attend to [1]. The second developed a web-based discussion sensemaking prototype with LLM-assisted features and studied how Wikipedians used it, along with the cognitive process behind their use (under review). Based on these two studies, we developed an AI-assisted sensemaking model for collaborative discussions (Figure 1).

Through our study, we found that sensemaking operates over six types of information: argument, community norm, norm context, people, people context, and discussion topic. While argument, community norm, and people are directly expressed within the discussion, norm context and people context lie outside the discussion within the broader community, and discussion topic represents domain-level information that exists outside the community.

We found that participants' sensemaking process can be aligned with Pirulli and Card's sensemaking model [11], so the human part of our model follows their stages: Search & Filter → Read & Extract → Schematize → Build Case → Tell Story. In their model, sensemaking begins with Search and Filter, which identifies and collects relevant information, followed by Read and Extract, which pulls useful evidence from what was collected. This evidence is then Schematized into a structured representation, Build Case develops

an interpretation from that representation, and Tell Story communicates the resulting interpretation. Inquiries or feedback on the interpretation can direct attention back to earlier stages, prompting re-evaluation and further search.

The AI system in our model operates as an overlay on the early stages of the human process. It performs four functions: Extraction, Prioritization, Connecting, and Presentation. Extraction involves identifying and extracting information elements from raw data; Prioritization involves selecting and ordering information based on importance; Connecting establishes relationships between different information elements; and Presentation provides these results as intermediate outputs. Prioritization and Connecting can operate iteratively, repeatedly informing one another as importance is reassessed based on newly formed relationships.

In our model, the degree of interpretive work the system performs can range from low to high. We define interpretive work as how much the system modifies raw data, based on the importance of that information for final interpretation. For every individual information feature, such as a discussion summary, an overview of participant information, or past norm practices, designers need to decide on the degree of its interpretive work. For example, in presenting a discussion summary, one lower-interpretation implementation might preserve much of the original discussion while only lightly condensing it. As the system produces a more selective summary that includes only the points it identifies as important, it performs more interpretive work.

3 Difficulty of Interpreting Discussion Correctly, and What It Means for System Design

Any time an AI system, an LLM or otherwise, is applied to a task, it performs interpretive work, whether or not its designers or users notice. This is not unique to our context or model. However, that work is not always correct. Hallucination is one example of what an incorrect result can look like. In our context of collaborative discussion, being correct means correctly interpreting participants' arguments during discussion sensemaking.

3.1 Interpreting Human

To correctly interpret someone's argument is to accurately understand the meaning behind what they said. This is essentially what Max Weber [13] defines as the goal of interpretive sociology, an attempt to "understand" human behavior through "interpretation." According to Weber, this process aims to arrive at an "intelligible explanation" of that behavior.

Reaching this goal requires keeping human subjectivity in mind. The behavior interpretive sociology seeks to understand is grounded in the actor's own subjectively intended meaning, and this behavior is interrelated with and codetermined by others' behavior. In other words, one's action is shaped by how the actor subjectively understands the meaning of their own action in relation to others. The same subjective logic applies to "associational action," an action oriented toward an established rule or order that applies to the whole group, the actor along with the other actors. The order itself exists objectively, but each actor takes it into account subjectively, and how each actor understands and expects others to follow that order is a subjective judgment.

Irrationality is another feature of social action Weber discusses. He warns that if interpretive analysis only assumes rationality, the researcher will keep running into purposes that can no longer be interpreted as rational means to some further end. Irrational action is diverse, ranging from irrational but understandable behavior to behavior that resists understanding altogether. Weber also notes that even an action starting from the same rational basis can, through incidental factors, take an irrational course.

Underlying both features, subjectivity and irrationality, is emotion. Weber treats emotion as not quite action itself, but closely bound to social action. It shapes how an actor subjectively constructs their action's intended meaning, and can be a source of irrational action.

Weber's account shows just how difficult correctly interpreting someone's behavior is, given human subjectivity and irrationality, and the role emotion plays in both.

3.2 Implications for the System's Interpretive Work

We distinguish between interpretation and interpretive work. Interpretive work is the process and practice through which interpretation happens, while interpretation is the meaning ultimately arrived at through that process. In our model, the system performs interpretive work rather than interpretation, filtering and highlighting the information that matters for the human's eventual interpretation.

Yet even though the system does not interpret participants' arguments directly, its interpretive work still needs to be "intelligible", especially given the complexity of the social setting a discussion like this involves. Whether a system is capable of that is an open question.

3.2.1 The Risks of a System That Falls Short of Intelligibility. Why this capability matters is that the level of interpretive work the system performs, low or high, corresponds inversely to the amount of interpretive work users need to do themselves. As the system's interpretive work increases and data is reshaped around the information it judges to be important, the work users need to do correspondingly decreases. In the discussion summary example above, a more selective summary means users spend less time reading through details to find what matters. But reducing users' interpretive work also means relying more on system-generated judgments. In other words, while more interpretive outputs can offer more useful support for sensemaking, they also require users to trust the system's judgment more.

If a system performing such a high level of interpretive work fails to correctly identify what information matters for making sense of a discussion in a complex social setting, several risks can arise.

One is that it dehumanizes participants, flattening them into generic contributors and stripping away the context that explains why their contribution matters. What's worse, this loss does not happen randomly; it happens in a systematically biased way. Whose contribution is recognized ends up depending less on how important what someone actually said is, and more on how legible their contribution is to the system. This can steer the discussion in the wrong direction and lead to poor decision making.

Another is homogenization, or false consensus. By surfacing only the points it judges important, a system can erase minority opinions and nuance that were actually present in the discussion, points that may have genuinely mattered but are never taken into account. It can also make it look like a cleaner agreement was reached than actually existed, letting the group act on a decision that does not have the buy-in it appears to have, with the underlying disagreement potentially resurfacing as conflict later on.

3.2.2 Designing for Intelligibility. For a system to gain the ability to judge what information matters, it has to be an "intelligible actor" that understands the specific social setting. We argue that the most important step toward this is providing a clear account of what kinds of information matter for understanding that social setting. In our context of collaborative discussion, this means, as our model shows, not focusing only on the arguments themselves, but also making clear that the norms mentioned in the discussion and their norm context matter, as do the participants and their people context. Each of these, norm context and people context, in turn requires its own layered understanding. Norm context, for instance, may need to be understood at three levels, an objective description of the norm, how each member subjectively understands that objective rule, and, since disagreement over the norm likely stems from differences in this subjective understanding, what shared understanding the members eventually reached and enacted through discussion. People context is more complex still, requiring an understanding

of each member's experience, background, role, relationships, and character, among other things.

But does doing this actually make the system a sufficiently intelligible actor? We cannot know for certain. This is why we still need safeguards against the risks of a system that falls short. One safeguard is leaving the final interpretation to humans. This is why, in our model, the system operates only as an overlay on the earlier stages of sensemaking, information search and organization, while leaving Build Case and Tell Story to the human. Another safeguard could be continuing to expose the original data, for transparency and explainability. This could be done by linking any presentation involving interpretive work back to the original data and providing a rationale for the interpretive work performed.

4 Toward a Design Agenda

As AI already takes on a growing role in tools people use for group discussion, such as meeting-minute agents, this raises open questions for design. We see these as falling into three parts, deciding the degree of interpretive work each system feature should perform, making the system itself a more intelligible actor, and building in the safeguards it still requires. We hope this paper offers a starting point for addressing them.

References

- [1] Soobin Cho, Mark Zachry, and David W. McDonald. 2025. Towards Insider Summarization for Mediation Instead of Moderation: Examining Wikipedian Views on Key Elements of Discussion Summaries. *Proc. ACM Hum.-Comput. Interact.* 9, 7, Article CSCW500 (Oct. 2025), 25 pages. doi:10.1145/3757681
- [2] Jeff Conklin and Michael L Begeman. 1988. gIBIS: A hypertext tool for exploratory policy discussion. *ACM Transactions on Information Systems (TOIS)* 6, 4 (1988), 303–331.
- [3] Gerardine DeSanctis and R Brent Gallupe. 1987. A foundation for the study of group decision support systems. *Management science* 33, 5 (1987), 589–609.
- [4] Hyo Jin Do, Ha-Kyung Kong, Pooja Tetali, Karrie Karahalios, and Brian P. Bailey. 2023. Inform, Explain, or Control: Techniques to Adjust End-User Performance Expectations for a Conversational Agent Facilitating Group Chat Discussions. *Proc. ACM Hum.-Comput. Interact.* 7, CSCW2, Article 343 (Oct. 2023), 26 pages. doi:10.1145/3610192
- [5] Soomin Kim, Jinsu Eun, Changhoon Oh, Bongwon Suh, and Joonhwan Lee. 2020. Bot in the Bunch: Facilitating Group Chat Discussion by Improving Efficiency and Participation with a Chatbot (*CHI '20*). Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3313831.3376785
- [6] Soomin Kim, Jinsu Eun, Joseph Seering, and Joonhwan Lee. 2021. Moderator Chatbot for Deliberative Discussion: Effects of Discussion Structure and Discus-sant Facilitation. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 87 (April 2021), 26 pages. doi:10.1145/3449161
- [7] Werner Kunz and Horst WJ Rittel. 1970. Issues as elements of information systems. (1970).
- [8] John Lawrence and Chris Reed. 2020. Argument mining: A survey. *Computational linguistics* 45, 4 (2020), 765–818.
- [9] Jay Nunamaker, Alan Dennis, Joseph Valacich, Doug Vogel, and Joey George. 1991. Electronic Meeting Systems To Support Group Work. *Commun. ACM* 34 (01 1991), 40–61.
- [10] Raquel Mochales Palau and Marie-Francine Moens. 2009. Argumentation mining: the detection, classification and structure of arguments in text. In *Proceedings of the 12th international conference on artificial intelligence and law*. 98–107.
- [11] Peter Pirolli and Stuart Card. 2005. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of international conference on intelligence analysis*, Vol. 5. McLean, VA, USA, 2–4.
- [12] Michael Henry Tessler, Michiel A. Bakker, Daniel Jarrett, Hannah Sheahan, Martin J. Chadwick, Raphael Koster, Georgina Evans, Lucy Campbell-Gillingham, Tatum Collins, David C. Parkes, Matthew Botvinick, and Christopher Summerfield. 2024. AI can help humans find common ground in democratic deliberation. *Science* 386, 6719 (2024), eadq2852. arXiv:https://www.science.org/doi/pdf/10.1126/science.adq2852 doi:10.1126/science.adq2852
- [13] Max Weber. 1981. Some categories of interpretive sociology. *The Sociological Quarterly* 22, 2 (1981), 151–180.
- [14] Howard T. Welsler, Dan Cosley, Gueorgi Kossinets, Austin Lin, Fedor Dokshin, Geri Gay, and Marc Smith. 2011. Finding social roles in Wikipedia. In *Proceedings of the 2011 IConference* (Seattle, Washington, USA) (*iConference '11*). Association for Computing Machinery, New York, NY, USA, 122–129. doi:10.1145/1940761.1940778
- [15] Wikipedia. 2024. Wikipedia:Expectations and norms of the Wikipedia community. Retrieved Feb 09, 2026 from https://en.wikipedia.org/wiki/Wikipedia:Expectations_and_norms_of_the_Wikipedia_community#Main_policies,_guidelines_and_essays
- [16] Wikipedia. 2026. Wikipedia:Essays. Retrieved Feb 10, 2026 from <https://en.wikipedia.org/wiki/Wikipedia:Essays>