From Queries to Conversations: Pragmatic and Interdependent Human–LLM Collaboration in Data Work

VIDYA SETLUR*, Tableau Research, USA

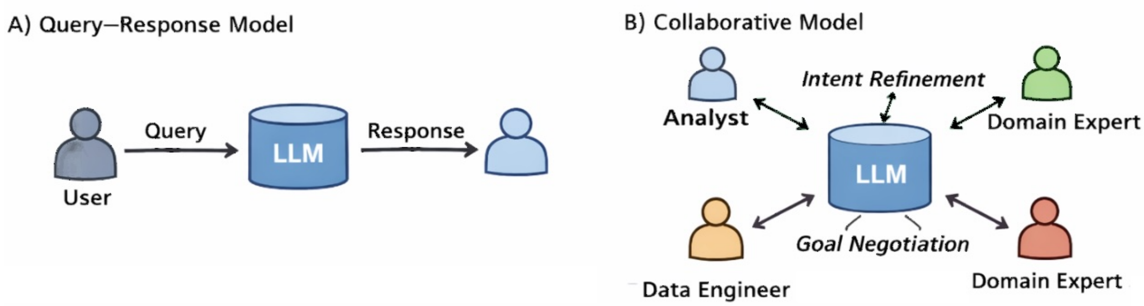

Fig. 1. Traditional LLM use treats interaction as a single-user query-response loop (left). Collaborative data work requires LLMs to mediate meaning, intent, and coordination across multiple stakeholders (right).

Data-intensive workflows are rarely solitary endeavors. Analysts, domain experts, and data engineers collaborate to collect, clean, integrate, annotate, and query data, often negotiating meaning, intent, and responsibility through conversation. As Large Language Models (LLMs) are increasingly embedded in data tools, their role is typically framed as responding to isolated queries or automating discrete steps. We argue that this framing underutilizes LLMs' potential and obscures important challenges in collaborative data work. In this position paper, we propose treating LLMs as *pragmatic collaborators* rather than tools, wherein participants that engage in ongoing conversations, mediate semantic differences, and adapt their behavior based on interdependence among stakeholders. Grounded in Gricean Maxims and Interdependence Theory, we examine how LLMs can support intent refinement, handle vagueness, and maintain discourse coherence across multi-party data workflows. We focus on semantic misalignment and goal misalignment as recurring sources of breakdown, and identify failure cases where violations of pragmatic norms undermine trust and shared outcomes.

Additional Key Words and Phrases: Conversational Data Analysis, Pragmatics, Gricean Maxims, Interdependence Theory, Collaborative Data Work, Human-Centered AI

1 Introduction

Modern data workflows are inherently collaborative. Analysts, domain experts, and data engineers work together to make sense of messy, evolving datasets, each bringing distinct expertise, assumptions, and goals. Much of this work unfolds through conversation: collaborators clarify intent, negotiate terminology, surface implicit assumptions, resolve disagreements, and coordinate decisions over time. These interactions are rarely linear or fully specified at the outset; instead, intent and meaning emerge gradually through back-and-forth dialogue as data and objectives evolve, reflecting long-standing views of language as a form of action in cooperative work [2, 7].

Recent advances in Large Language Models (LLMs) have led to their rapid integration into tools for data cleaning, integration, annotation, and querying [5]. However, most current systems conceptualize LLMs as query–response engines or automation aids, supporting individual users through isolated interactions. As illustrated in Figure 1A, this model assumes well-formed queries, well-interpreted intent, and a single locus of control—assumptions that align with dominant human–LLM interaction paradigms identified in recent taxonomies [1]. In collaborative settings, these

Author's Contact Information: Vidya Setlur, vsetlur@tableau.com, Tableau Research, Palo Alto, CA, USA.

assumptions frequently break down; intent is often underspecified, terminology is contested across roles, and goals shift as collaborators encounter new constraints or insights.

We argue that LLMs must instead be designed as conversational collaborators embedded within multi-stakeholder workflows (Figure 1B). In this framing, LLMs participate in ongoing dialogue, helping collaborators articulate and refine intent, align semantics across perspectives, and coordinate shared outcomes. Rather than simply executing commands or generating answers, LLMs should surface assumptions, prompt clarification when meanings diverge, and adapt their behavior as collaboration unfolds. Designing LLMs in this manner foregrounds pragmatic reasoning [2, 4], semantic alignment, and interdependence as central challenges for human–LLM collaboration in data work [6].

2 Background

Analytical interaction is inherently pragmatic: users rely on context, shared assumptions, and conversational repair to communicate intent, particularly when goals are evolving or underspecified. Prior work has long emphasized that language in cooperative work functions not merely as a medium for information exchange, but as a mechanism for coordinating action and aligning understanding among collaborators [7]. In analytical settings, meaning is often negotiated incrementally through dialogue rather than fully specified upfront.

Gricean Maxims, *Quantity*, *Quality*, *Relation*, and *Manner* provide a useful lens for understanding how meaning is inferred, managed, and repaired in such conversations [2]. These principles have been shown to be particularly relevant in the design of analytical conversational interfaces, where users expect systems to handle ambiguity, refine intent, and respond cooperatively rather than literally [4, 6].

In collaborative data workflows, violations of these maxims are common and consequential. LLMs may provide overly detailed or insufficient explanations (violating *Quantity*), present speculative inferences with unwarranted confidence (violating *Quality*), perform transformations misaligned with collaborators' evolving goals (violating *Relation*), or communicate in ways that obscure assumptions, uncertainty, or rationale (violating *Manner*). While such violations can be disruptive in single-user settings, they are particularly damaging in multi-stakeholder contexts, where breakdowns in shared understanding propagate across roles and undermine collective trust and accountability.

Beyond pragmatics, collaborative data work is shaped by varying degrees of interdependence among participants. Interdependence Theory emphasizes that effective collaboration depends on mutual responsiveness and alignment toward shared outcomes [3]. In data workflows, interdependence varies substantially by task: exploratory querying may tolerate divergence in interpretation, whereas data cleaning and integration often require consensus, and annotation or labeling tasks frequently involve negotiation among stakeholders with differing expertise and priorities [8].

Designing LLM behavior without accounting for these differences in interdependence risks premature autonomy or inappropriate deference—either acting too decisively in contexts that require negotiation, or withholding initiative when coordination support is needed. Recent taxonomies of human–LLM interaction modes underscore the importance of adapting system behavior to task structure and collaborative context, rather than assuming a single, dominant interaction paradigm [1]. To this end, pragmatics and interdependence offer complementary lenses for understanding when and how LLMs should participate in collaborative data work.

3 Challenges in Collaborative Data Work

3.1 Challenge 1: Semantic Misalignment

Semantic misalignment is a recurring challenge in collaborative data work. Collaborators frequently use shared terms that carry different meanings depending on disciplinary background, analytical goals, or domain expertise. For example, an analyst may interpret an "outlier" as a statistical anomaly to be removed or downweighted, while a domain expert may view the same data point as a rare case worthy of further investigation. Similar discrepancies arise around terms such as "important," "large," or "relevant," whose meanings are often context-dependent and implicitly negotiated through interaction [2, 7].

When LLMs are introduced into these workflows, semantic misalignment can be amplified rather than resolved. LLMs trained on broad, general-purpose corpora may default to dominant or statistically prevalent interpretations, privileging one perspective while obscuring others. Prior work on analytical conversational interfaces has shown that failures to surface ambiguity or support semantic clarification can undermine user trust and shared understanding [6]. Without mechanisms to surface alternative interpretations or prompt clarification, LLM-generated outputs can silently reinforce misunderstandings, making semantic disagreements harder to detect and repair. In multi-stakeholder contexts, such failures undermine shared understanding and may lead collaborators to trust or reject system outputs for different—and often unarticulated—reasons [8].

3.2 Challenge 2: Goal Misalignment and Drift

In addition to semantic differences, collaborative data workflows are characterized by evolving and sometimes competing goals. Objectives are rarely fixed at the outset; instead, they shift as collaborators encounter new constraints, insights, or downstream requirements. A workflow that initially frames the task as "clean the data" may later prioritize preserving edge cases, supporting interpretability, or preparing the data for a specific modeling approach. These shifts are often implicit and negotiated through conversation rather than explicitly documented, reflecting the dynamic nature of cooperative analytical work [7].

LLMs that fail to track such goal evolution risk optimizing for outdated or incomplete objectives. For example, an LLM may aggressively remove anomalous values based on an early understanding of the task, even after collaborators have reframed those values as analytically significant. Conversely, excessive deference to earlier instructions can result in missed opportunities to support emerging goals. Empirical studies of human–AI collaboration highlight how such misalignments can lead to breakdowns in trust, coordination, and effectiveness when system behavior diverges from evolving human expectations [8]. In collaborative settings, goal misalignment is particularly problematic because different stakeholders may notice and respond to such drift at different times, leading to fragmented or inconsistent expectations of the system's behavior.

To address these challenges, we propose that LLMs be designed as pragmatic collaborators whose behavior is guided by conversational norms and sensitivity to interdependence among participants. Rather than occupying a fixed role, LLMs should adapt their participation based on the structure of the task, the degree of consensus required, and the evolving state of collaboration. In collaborative data work, LLMs may alternately function as executors of well-specified actions, critics that surface inconsistencies or risks, mediators that expose semantic differences and prompt negotiation, or coordinators that help align goals and sequence activities. The appropriateness of these roles depends on both the level of task interdependence and the degree of alignment among collaborators. Designing LLMs to shift fluidly among these roles allows them to support collaboration without prematurely asserting autonomy or withdrawing support when coordination is needed [1]. Together, these pragmatic considerations highlight that effective human–LLM collaboration in data work is not solely a matter of improving model accuracy or performance. Rather, it requires designing interaction behaviors that support negotiation, repair, and shared understanding across diverse collaborators and evolving tasks.

4 Discussion and Provocation

Reframing LLMs as pragmatic collaborators rather than autonomous tools raises a set of open questions that are central to the goals of the Co-Data workshop. At the heart of these questions is the issue of *agency*: when should an LLM take initiative in a collaborative data workflow, and when should it defer to human judgment? While automation may be desirable for well-specified, low-interdependence tasks, premature initiative in contexts that require negotiation or consensus can undermine collaboration. Conversely, excessive deference can place undue cognitive burden on human collaborators, particularly when coordination or sensemaking support is needed. Determining how and when LLMs should lead, follow, or negotiate remains an open design challenge that depends on task structure, interdependence, and the evolving state of collaboration.

A related provocation concerns how semantic negotiation should be made visible and actionable in user interfaces. Much of the semantic work in collaborative data analysis currently happens implicitly, through conversation and informal repair. When LLMs participate in these workflows, they often obscure semantic decisions by presenting a single, seemingly authoritative interpretation. This raises questions about how interfaces might instead expose competing interpretations, prompt clarification when meanings diverge, or allow collaborators to inspect and negotiate the semantics underlying system actions.

Traditional metrics such as task completion time or accuracy fail to capture whether collaborators share an understanding of goals, trust system behavior, or feel ownership over outcomes. In collaborative data work, success may instead be reflected in the quality of negotiation, the ability to recover from misunderstandings, and the extent to which system behavior supports shared sensemaking. Developing evaluation criteria that account for these relational and process-oriented dimensions is a critical area for future research. Overconfidence in model outputs can discourage critical scrutiny and suppress alternative interpretations. Semantic collapse, where nuanced distinctions are reduced to a single dominant framing, can erase important domain knowledge. Opacity in reasoning obscures how conclusions were reached, making it difficult for collaborators to contest or revise system behavior. Finally, premature autonomy, i.e., acting decisively in contexts that require negotiation can fracture trust and disrupt shared ownership of decisions.

LLMs should surface their assumptions, alternatives, and sources of uncertainty, enabling collaborators to reason about and critique system behavior. Their roles should adapt dynamically based on task interdependence and the degree of alignment among collaborators, shifting between execution, critique, mediation, and coordination as needed. Support for conversational repair, acknowledging misunderstandings, revisiting prior decisions, and updating shared context should be treated as a core capability rather than an edge case. Finally, system reasoning should be made visible and contestable, allowing collaborators to interrogate, challenge, and reshape the contributions of LLMs within the workflow.

To summarize, these provocations underscore that effective human–LLM collaboration in data work is not achieved by making models more autonomous, but by making their participation more accountable, transparent, and responsive to human interactional norms. Grounding LLM behavior in conversational pragmatics and interdependence opens new opportunities for designing human-centered data systems that support collaborative, equitable, and trustworthy workflows beyond query–response paradigms.

References

[1] Jie Gao, Simret Araya Gebreegziabher, Kenny Tsu Wei Choo, Toby Jia-Jun Li, Simon Tangi Perrault, and Thomas W Malone. 2024. A Taxonomy for Human-LLM Interaction Modes: An Initial Exploration. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '24)*. Association for Computing Machinery, New York, NY, USA, Article 24, 11 pages. doi:10.1145/3613905.3650786

[2] H. P. Grice. 1975. Logic and Conversation. In *Syntax and Semantics: Vol. 3: Speech Acts*, Peter Cole and Jerry L. Morgan (Eds.). Academic Press, New York, 41–58. http://www.ucl.ac.uk/ls/studypacks/Grice-Logic.pdf

[3] Harold H. Kelley and John W. Thibaut. 1978. Interpersonal relations: A theory of interdependence. https://api.semanticscholar.org/CorpusID:145761013

[4] Yoonsu Kim, Brandon Chin, Kihoon Son, Seoyoung Kim, and Juho Kim. 2025. Applying the Gricean Maxims to a Human-LLM Interaction Cycle: Design Insights from a Participatory Approach. 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, Article 72, 8 pages. doi:10.1145/3706599.3719759

[5] Marianne Aubin Le Quéré, Hope Schroeder, Casey Randazzo, and Jie Gao. 2025. The State of Large Language Models in HCI Research: Workshop Report. *Interactions* 32, 1 (Jan. 2025), 8–9. doi:10.1145/3705617

[6] Vidya Setlur and Melanie Tory. 2022. How do you Converse with an Analytical Chatbot? Revisiting Gricean Maxims for Designing Analytical Conversational Behavior. 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 29, 17 pages. doi:10.1145/3491102.3501972

[7] Terry Winograd. 1986. A language/action perspective on the design of cooperative work. In *Proceedings of the 1986 ACM Conference on Computer-Supported Cooperative Work* (Austin, Texas) *(CSCW '86)*. Association for Computing Machinery, New York, NY, USA, 203–220. doi:10.1145/637069.637096

[8] Lixiang Yan, Vanessa Echeverria, Gloria Milena Fernandez-Nieto, Yueqiao Jin, Zachari Swiecki, Linxuan Zhao, Dragan Gašević, and Roberto Martinez-Maldonado. 2024. Human-AI Collaboration in Thematic Analysis using ChatGPT: A User Study and Design Recommendations. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '24)*. Association for Computing Machinery, New York, NY, USA, Article 191, 7 pages. doi:10.1145/3613905.3650732