Designing Mixed-Initiative Collaboration Between Humans and Increasingly Autonomous AI Systems

Kevin Pu
Department of Computer Science
University of Toronto
Toronto, Ontario, Canada
jpu@dgp.toronto.edu

Abstract

The recent development of increasingly autonomous AI systems is transforming human-AI interaction from reactive tool use to collaborative, mixed-initiative workflows. My thesis posits that effective collaboration with such systems requires interaction models that enable users to specify, refine, and generalize their intent and feedback at the appropriate level of task abstraction, keeping AI initiative steerable as systems grow more capable and distributed. My research develops this thesis across progressively higher abstraction levels: specifying and generalizing user intent in web automation, designing proactive programming agents and grounding their timing in the user's cognitive state, and preserving user agency in open-ended research ideation through structured, manipulable representations. My ongoing projects address the culminating challenge of steering multi-agent and long-running AI systems, generalizing lightweight user feedback and accumulated context into persistent, inspectable representations that scale human oversight. Together, these efforts contribute to a research agenda toward scalable, cognitively-aligned, and mixed-initiative AI systems that support human creativity, oversight, and evolving task intentions.

Keywords

Mixed-Initiative Design; Proactive AI; Human-AI Interaction

ACM Reference Format:
Kevin Pu. 2026. Designing Mixed-Initiative Collaboration Between Humans and Increasingly Autonomous AI Systems. In *The 39th Annual ACM Symposium on User Interface Software and Technology (UIST Adjunct '26), November 02–05, 2026, Detroit, MI, USA.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3830397.3841711

1 INTRODUCTION

Recent advances in large language models (LLMs) have reshaped interactions with intelligent systems across various tasks, from automating programming scripts to collaborating on creative ideation. These systems are no longer limited to reactive or narrowly scoped tasks. With enhanced contextual understanding, reasoning ability, and initiative, modern AI increasingly acts as a proactive collaborator, capable of making suggestions, taking actions, and adapting dynamically to evolving user goals [4–6].

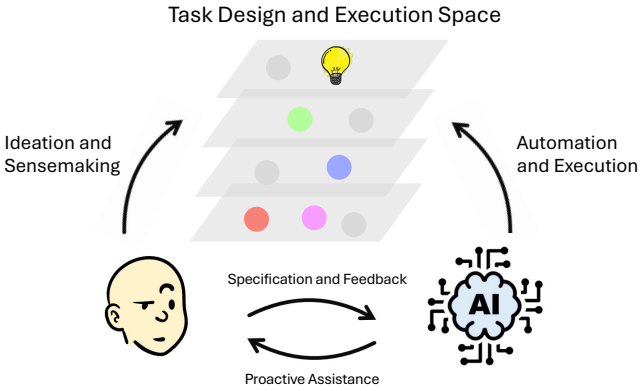


Figure 1: My research focuses on designing mixed-initiative interactions for human-AI collaboration, where automation and proactive assistance serve to enhance and assist human productivity and creativity across different task abstractions.

These developments open exciting opportunities for enhancing human productivity and creativity, but they also raise new questions around interaction design: How should initiative be shared between human and AI agents? How can systems remain transparent, steerable, and responsive across different task contexts and abstraction levels? And how can users maintain control and oversight when working with increasingly autonomous and distributed AI systems?

My research explores these questions through the lens of human-centered, mixed-initiative interaction. My thesis is that effective collaboration with increasingly autonomous AI requires interaction models that enable users to specify, refine, and generalize their intent and feedback at the appropriate level of task abstraction—keeping AI initiative steerable as systems grow more capable and distributed. To develop this thesis, my work has deliberately progressed up levels of task abstraction: from specifying semantic intent in structured web automation, to designing and timing proactive assistance in programming, to co-creating in open-ended research ideation, and now to steering multi-agent systems. Across these stages, I investigate how to structure AI initiative, surface appropriate feedback mechanisms, and balance control between users and systems.

As illustrated in Figure 1, my research aims to build generalizable frameworks for mixed-initiative collaboration that adapt to user needs, support flexible delegation, and foster meaningful co-agency between humans and increasingly capable AI systems. My ongoing

This work is licensed under a Creative Commons Attribution 4.0 International License.
UIST Adjunct '26, Detroit, MI, USA
© 2026 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2855-6/26/11
https://doi.org/10.1145/3830397.3841711

work addresses the culminating challenge of this agenda: steering multi-agent, long-running AI systems where per-action oversight and single-session user context no longer scale. Ultimately, my research seeks to contribute to an emerging body of work in HCI that designs human-centered, mixed-initiative AI systems capable of flexibly adapting to human cognitive workflows and dynamically evolving task intentions.

2 PAST CONTRIBUTIONS

2.1 Semantic Intent Specification And Generalizable Automation

The core challenge in utilizing an automated system is effectively specifying the user's intent such that it accurately translates to system execution across task scenarios and edge cases, especially as user intents evolve in response to execution results. My initial research addressed this challenge by enhancing traditional programming-by-demonstration (PBD) web automation with (1) interaction techniques to express semantic intent and (2) intent-to-execution mapping that creates generalizable automation.

SemanticOn [4] enables users to specify rich, content-based conditions for web automation—beyond the structural or syntactic criteria of prior systems—by combining natural language prompts with direct content highlighting, which the system maps to executable semantic filters using NLP and CV models. Users iteratively refine these conditions through continuous interaction as tasks evolve or errors arise. *DiLogics* [6] extends this approach to tasks requiring diverse, conditional logic, such as filling web forms from heterogeneous spreadsheet data: it semantically segments input data into structured steps, elicits UI demonstrations for each, and generalizes the automation to novel inputs via semantic matching, requesting human input only when segmentation is ambiguous. In both evaluations, participants successfully created automations that handled complex, previously unseen conditions and valued the systems' support for iterative intent refinement.

These projects established a foundational pattern for my thesis: users can express high-level intent that systems generalize beyond literal demonstrations, while lightweight, iterative refinement keeps automation aligned with evolving goals. This insight of generalizing user input to reduce specification effort while preserving control recurs throughout my subsequent work.

2.2 Designing and Evaluating Proactive Programming Assistance

Building on the foundational work of specifying and generalizing user intent for program automation, my subsequent project, Assistance or Disruption [5], explores a critical next step: proactively assisting users during open-environment programming. Early AI code-assistance tools, such as GitHub Copilot, generally operate in a reactive manner, where users explicitly invoke assistance as needed. However, observational studies highlight inefficiencies and cognitive burdens associated with user-initiated help-seeking [2, 8].

To explore the potential for proactive, agent-initiated assistance, we evaluated three design probes representing increasing levels of AI autonomy and visibility: (1) *Reactive* (assistance provided strictly upon explicit user request), (2) *Proactive with Direct Execution* (the system automatically executes certain code modifications without direct user approval), and (3) *Proactive with Agent Presence and Scopes of Control* (the AI visually presents proposed actions within the editor, allowing users to quickly confirm, modify, or dismiss suggestions at both line-level and global scope). A within-subjects user study with 18 participants revealed nuanced trade-offs: while proactive conditions significantly increased users' perceived efficiency and ease of task completion, excessive or mistimed proactivity introduced disruptions, negatively impacting user trust and task flow.

This study contributes empirical insights and actionable design principles for proactive AI integration, and it surfaced a key open problem for my thesis: *when* an AI agent takes initiative matters as much as *what* it does, motivating my subsequent work on principled intervention timing.

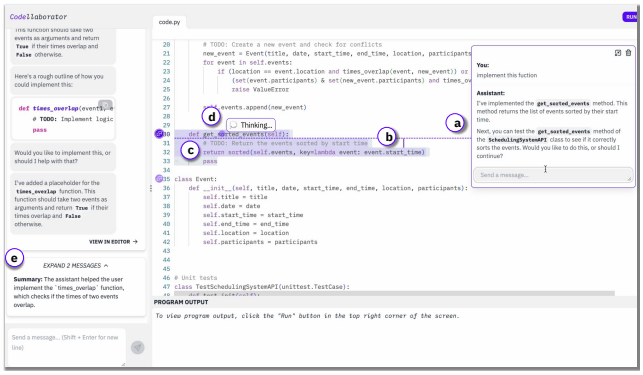


Figure 2: The Design Probe UI. The AI adds code with its caret ⓑ in the editor and replies to the user in the breakout to discuss the next steps. The purple highlight ⓒ indicates the provenance of the added code. The proactive AI agent's cursor displays "Thinking..." ⓓ to indicate its working process of generating a response.

2.3 Timing Proactive Assistance via User Cognitive State

Directly addressing the timing problem identified above, *ProMemAssist* [7] investigates how proactive agent assistance can be timed by explicitly modeling the user's mental state, specifically their working memory (WM) load. Traditional proactive systems rely on predefined heuristics or task-based triggers, neglecting users' dynamic cognitive availability and thereby causing mistimed interruptions. ProMemAssist instead leverages multi-modal sensor input from smart glasses to construct a real-time WM model grounded in cognitive theory, which informs a timing predictor that balances the assistance's potential value against its cognitive interruption cost. A user study with 12 participants completing cognitively demanding tasks showed that ProMemAssist delivered more selective assistance and received higher engagement compared to an LLM baseline, with fewer perceived interruptions.

This project demonstrates that grounding proactive behavior in models of the user's cognitive state makes AI initiative principled rather than heuristic, which becomes a building block for the adaptive, user-aware initiative my thesis envisions.

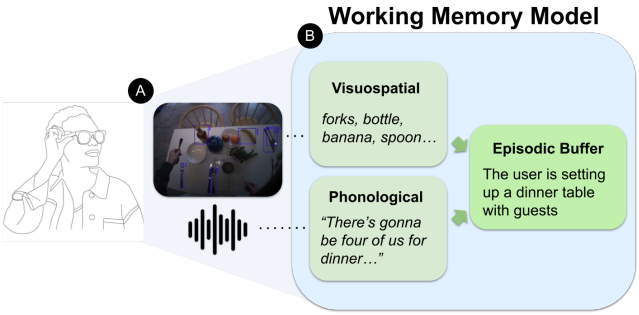


Figure 3: ProMemAssist's working memory (WM) model. ⓐ Multi-modal sensor streams from smart glasses capture the user's visual and auditory environment in real time. ⓑ Grounded in cognitive theories of WM, perceived information is encoded as visuospatial and phonological memory items (e.g., detected objects, overheard speech) and bound into an episodic buffer representing the user's current task context (e.g., setting up a dinner table with guests). This WM model informs a timing predictor that balances an assistance's value against its interruption cost.

2.4 Mixed-Initiative Collaboration in Open-Ended Research Ideation

Extending beyond programming into more abstract, creative cognitive tasks, my work, *IdeaSynth* [3], investigates how AI can effectively support researchers during the early, formative stages of idea generation and refinement. Unlike structured tasks with clear success criteria, scientific ideation is inherently exploratory, non-linear, and iterative, and traditional AI systems often overemphasize generative capabilities, potentially undermining user agency or causing idea fixation [1, 9].

To address these challenges, *IdeaSynth* provides a structured yet flexible interface that explicitly foregrounds the co-evolutionary nature of ideation. It decomposes research ideas into core facets, such as problem statements, methods, and evaluation strategies, and provides literature-grounded AI-generated suggestions that users can flexibly explore, integrate, or modify. A study involving 20 researchers demonstrated that participants effectively used the system to iteratively develop and refine research ideas, reporting enhanced creativity and productivity while maintaining a sense of authorship and agency.

By demonstrating effective collaboration patterns in high-abstraction, cognitively demanding tasks, *IdeaSynth* shows that structuring ideas as explicit, manipulable representations lets users steer AI contributions in open-ended work. This insight directly informs my ongoing research on steerable representations for autonomous and multi-agent systems.

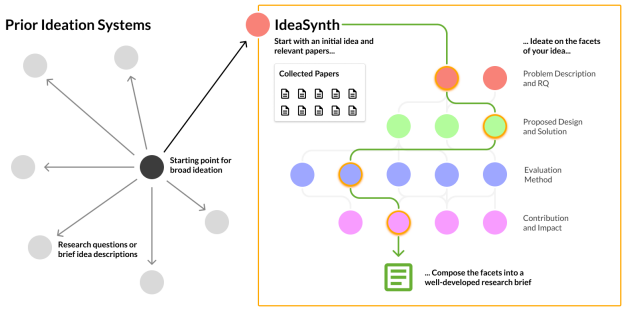


Figure 4: While prior systems focused on helping users explore diverse ideas broadly, IdeaSynth is an ideation tool that helps users more deeply develop an initial research idea into a concrete and specific research brief. Intuitively, researchers often dissect projects into different facets to further develop them; IdeaSynth allows users to structure their idea as faceted idea nodes, explore and evolve different variations, and, finally, compose them into a complete research brief.

3 CURRENT AND ONGOING WORK

My past work shows that users can specify generalizable intent, that AI initiative must be carefully designed and timed, and that explicit, manipulable representations preserve user agency in open-ended tasks. My ongoing research brings these threads together to address the culminating challenge of my thesis: as AI systems become multi-agent and long-running, overseeing every AI system's action no longer scales, yet high-level guidance alone sacrifices control. My two current projects investigate a shared approach: generalizing lightweight user feedback and accumulated context into persistent, inspectable representations that amplify human steering without proportionally increasing oversight effort.

3.1 Generalizing User Feedback to Steer Multi-Agent Systems

As multi-agent systems tackle increasingly complex tasks, users must oversee distributed workflows of agents with different roles and local contexts. Existing steering approaches face a trade-off between feedback granularity and generalizability: users either provide granular feedback by monitoring extensive agent execution traces, or rely on high-level guidance (e.g., global system prompts) that is too coarse to govern situated agent decisions. In a formative study with 8 multi-agent practitioners, I found that steering demands labor-intensive inspection of intermediate outputs, that feedback fragments across parallel tasks and must be repeatedly restated, and that users act like *managers of distributed work*, preferring selective intervention at consequential moments over continuous supervision.

Grounded in these findings and organizational theories of information processing, my ongoing work introduces a multi-agent steering paradigm built on three mechanisms: the system detects when task progress is under-specified or misaligned and proactively solicits targeted user feedback at these decision-relevant moments; it extracts that feedback into reusable, inspectable guidelines with

scope and provenance; and it propagates guidelines across parallel, downstream, and future tasks handled by different agents. A within-subjects study (N=8) against an ablated baseline showed that this approach shifted steering from reactive correction to proactive guidance, nearly tripling checkpoint approval rates, halving task revisions, and significantly improving composite ratings of control, transparency, and effort. This project directly extends my thesis arc: it applies the intent-generalization pattern from my PBD work to the multi-agent setting, and grounds *when* to solicit feedback in detected task state, echoing the principled intervention timing of my proactive assistance research.

3.2 Externalizing User Context as a Steerable Representation

Complementing task-level steering, my second ongoing project examines how systems can make a user's *accumulated context* equally steerable, including their background, goals, and evolving understanding across long-horizon interactions. LLM chat tools increasingly personalize responses, but existing approaches store user context as flat, opaque memories: users cannot see how their context is organized, when it shapes a response, or how to revise it when it becomes outdated or misaligned. I am developing a chat system that reframes personalization as an interactive, user-facing layer rather than a hidden system capability. The system organizes conversations into semantic topics and, for each topic, externalizes an inspectable and editable model of the user's background and evolving understanding; it surfaces personalization *provenance* directly in the conversation via inline markers linking responses to the prior chats that informed them; and it supports low-friction steering through lightweight response labeling and direct manipulation of the structure, while proactively suggesting exploration directions grounded in the user's knowledge state.

In a 7-day deployment study (N=8) as participants' primary chat tool, they developed stable mental models of what the system knew about them, reported that transparent personalization increased their control and trust, and sustained the labeling mechanism because it doubled as a personal reading aid, with 95% of user model updates driven by these lightweight signals rather than manual editing.

Where my earlier work helped users specify intent for a task, this project treats the user's evolving intent and knowledge *itself* as an explicit, manipulable artifact, extending steerability from individual tasks to the longitudinal human-AI relationship.

3.3 Dissertation Trajectory and Feedback Sought

I am a fifth-year PhD candidate expecting to graduate in 2027. My dissertation will synthesize these projects into a framework for mixed-initiative interaction across task abstraction levels, contributing interaction techniques for specifying and generalizing intent, empirical design principles for proactive AI initiative, and steering paradigms for increasingly autonomous AI systems. In the remainder of my PhD, I plan to deepen the multi-agent steering and accumulated personalized context directions. At the symposium, I would particularly value the community's feedback on (1) how to unify user-state models (e.g., cognitive load) and task-state models (e.g., detected misalignment) into a single account of when AI systems should take initiative, and (2) how to synthesize my past works into a unified, generalizable interaction framework that remains relevant as underlying AI capabilities improve.

4 CONCLUSION

As AI systems grow in autonomy, complexity, and generality, the challenge of aligning their behavior with human needs and workflows becomes increasingly critical. My research contributes to this space by developing interaction models that keep AI initiative steerable across levels of task abstraction, whether through semantic intent specification, cognitively-aligned proactive support, structured co-creation, or generalizing user feedback and context into persistent, inspectable representations. Across diverse task domains—from automation to ideation to multi-agent orchestration—my work emphasizes user agency, principled AI initiative, and feedback mechanisms grounded in real-world needs. Looking ahead, I will work toward a unifying framework for when and how AI systems should take initiative, drawing on both user-state and task-state models. My goal is to design human-AI systems that are not only more capable, but also more collaborative, allowing users to remain productive, creative, and in control as AI capabilities continue to evolve.

References

[1] Barrett R Anderson, Jash Hemant Shah, and Max Kreminski. 2024. Homogenization effects of large language models on human creative ideation. In *Proceedings of the 16th Conference on Creativity & Cognition*. 413–425.

[2] Ranim Khojah, Mazen Mohamad, Philipp Leitner, and Francisco Gomes de Oliveira Neto. 2024. Beyond Code Generation: An Observational Study of ChatGPT Usage in Software Engineering Practice. *Proc. ACM Softw. Eng.* 1 (2024), 1819–1840. https://api.semanticscholar.org/CorpusID:269302970

[3] Kevin Pu, K. J. Kevin Feng, Tovi Grossman, Tom Hope, Bhavana Dalvi Mishra, Matt Latzke, Jonathan Bragg, Joseph Chee Chang, and Pao Siangliulue. 2025. IdeaSynth: Iterative Research Idea Development Through Evolving and Composing Idea Facets with Literature-Grounded Feedback. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 145, 31 pages. doi:10.1145/3706598.3714057

[4] Kevin Pu, Rainey Fu, Rui Dong, Xinyu Wang, Yan Chen, and Tovi Grossman. 2022. Semanticon: Specifying content-based semantic conditions for web automation programs. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*. 1–16.

[5] Kevin Pu, Daniel Lazaro, Ian Arawjo, Haijun Xia, Ziang Xiao, Tovi Grossman, and Yan Chen. 2025. Assistance or Disruption? Exploring and Evaluating the Design and Trade-offs of Proactive AI Programming Support. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 152, 21 pages. doi:10.1145/3706598.3713357

[6] Kevin Pu, Jim Yang, Angel Yuan, Minyi Ma, Rui Dong, Xinyu Wang, Yan Chen, and Tovi Grossman. 2023. DiLogics: Creating Web Automation Programs with Diverse Logics. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. 1–15.

[7] Kevin Pu, Ting Zhang, Naveen Sendhilnathan, Sebastian Freitag, Raj Sodhi, and Tanya R. Jonker. 2025. ProMemAssist: Exploring Timely Proactive Assistance Through Working Memory Modeling in Multi-Modal Wearable Devices. In *Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology (UIST '25)*. Association for Computing Machinery, New York, NY, USA, Article 56, 19 pages. doi:10.1145/3746059.3747770

[8] Priyan Vaithilingam, Tianyi Zhang, and Elena L Glassman. 2022. Expectation vs. experience: Evaluating the usability of code generation tools powered by large language models. In *Chi conference on human factors in computing systems extended abstracts*. 1–7.

[9] Samangi Wadinambiarachchi, Ryan M Kelly, Saumya Pareek, Qiushi Zhou, and Eduardo Velloso. 2024. The Effects of Generative AI on Design Fixation and Divergent Thinking. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–18.