Towards Direct Intent Manipulation: Drag-Based Research Ideation, Evaluation and Evolution

Zijian Ding
ding@umd.edu
University of Maryland, College Park

Fenghai Li
max7@illinois.edu
University of Illinois Urbana-Champaign

Haofei Yu
haofeiy2@illinois.edu
University of Illinois Urbana-Champaign

Joel Chan
joelchan@umd.edu
University of Maryland, College Park

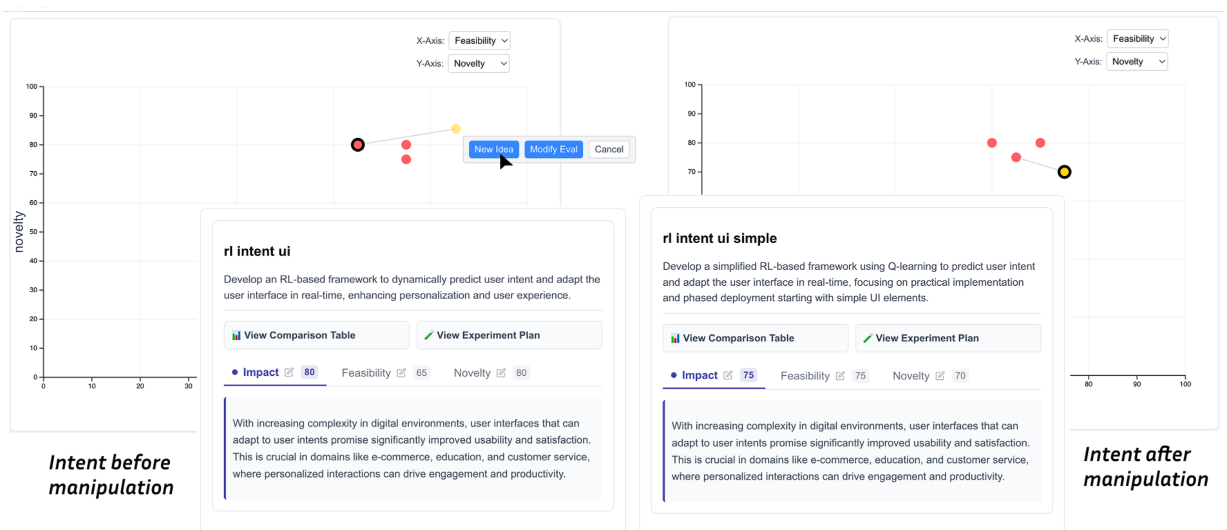

Figure 1: The user can directly manipulate a research intent by dragging and dropping it along the feasibility dimension, generating a more feasible iteration.

ABSTRACT

We present Direct Intent Manipulation (DIM), a GenAI-powered user-interface paradigm that treats nascent research intents as first-class, draggable objects within an ever-updating spatial canvas. Users can directly drag intents along quantitative axes such as novelty and feasibility. Each drag or merge will rewrite the chosen intent and re-scores every item in the landscape, giving immediate visual feedback and preserving context. By extending the classical virtues of direct manipulation—continuous representation and incremental operation—to the abstract level of research ideation, DIM empowers scientists to fluidly generate and evolve hypotheses without wrestling with opaque prompts or rigid forms.

UIST '25 Adjunct, September 28–October 1, 2025, Busan, Korea

ACM ISBN 978-1-4503-8655-5/21/10.
https://doi.org/10.1145/3474349.3480203

CCS CONCEPTS

• **Human-centered computing → Interaction design theory, concepts and paradigms.**

KEYWORDS

Direct Manipulation, Intent-based User Interface, Research Ideation

ACM Reference Format:
Zijian Ding, Fenghai Li, Haofei Yu, and Joel Chan. 2025. Towards Direct Intent Manipulation: Drag-Based Research Ideation, Evaluation and Evolution. In *The Adjunct Publication of the 37th Annual ACM Symposium on User Interface Software and Technology (UIST '25 Adjunct), September 28–October 1, 2025, Busan, Korea.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3474349.3480203

1 INTRODUCTION

Since 2024, there has been a marked increase in research investigating the use of AI to advance scientific discovery [6, 7, 11]. Scientific inquiry necessitates the systematic execution of a wide range of tasks—including literature review, experimental design, and empirical investigation—many of which are increasingly being facilitated by reasoning-capable models augmented through reinforcement learning [6, 11, 17].

Nevertheless, the capacity to evaluate and iteratively refine appropriate research directions—particularly in selecting contextually relevant, feasible, and scientifically meaningful research ideas—remains a significant limitation of current autonomous systems. Human expertise plays a critical role in this process, especially in exercising the nuanced judgment required to discern which research pathways warrant further pursuit and which should be discarded. For instance, Yang et al. [22] report that GPT-4's evaluations of research ideas exhibit a correlation with assessments from doctoral students in the range of 30% to 50%, typically differing by one point on a five-point Likert scale. Such discrepancies are particularly salient in evaluative dimensions such as conceptual clarity and methodological validity [10].

To advance human–AI alignment in research ideation and evaluation, we build on recent work in intent-based user interfaces [2]. We conceptualize brief, ambiguous research ideas as examples of **intent** (e.g., "I want to know more about X," or "I want to solve Y") and introduce **Direct Intent Manipulation (DIM)**—an interaction paradigm that enables direct manipulation of research intents within a spatial visualization.

2 BACKGROUND

2.1 From Execution-oriented Interfaces to Intent-based User Interfaces

With the advent of LLMs, users increasingly convey goals in natural language while autonomous agents execute low-level GUI operations. This shifts interaction needs from fine-grained direct control over task execution details to the need for *Intent-based User Interfaces* that "bridges user intents with task execution" [2]. This paradigm is exemplified by IntentPrism [21], a browser plug-in that surfaces a live "intent tree" so users can capture, refine, and visualise LLM-recognised intents during web foraging, and by CoLadder [23], a hierarchical prompt editor that lets programmers externalise their intent into modular blocks. Complementary agentic systems show how LLMs can translate free-form prompts into concrete actions—such as LAVE for video-editing via script-like text [20]. Collectively, these projects position intent as a first-class UI element and open a design space in which AI both recognises and actuates user goals.

2.2 Direct Manipulation at the Intent Level

One commonly adopted method for interacting with user intent is through prompt-based chat interfaces, exemplified by systems such as AI Co-Scientist [6]. While prompting affords a flexible means of conveying straightforward and well-defined intentions, it proves inadequate when users attempt to articulate ambiguous intents that require precision. For example, a prompt like "Please make the research idea a little more novel" is underspecified—the phrase "a little more novel" lacks quantifiable clarity, leaving its interpretation open-ended and context-dependent.

Direct manipulation could be a powerful approach for expressing verbally ambiguous intent precisely. Ben Shneiderman introduced the term direct manipulation in 1982 to characterize user interfaces that provide continuous visual feedback, support reversible actions, and enable rapid, incremental interaction [16]. This interaction paradigm played a key role in the transition from command-line interfaces to graphical user interfaces (GUIs).

Direct manipulation builds upon humans' evolved spatial reasoning abilities [8]. Even when dealing with abstract concepts, individuals can adapt to their spatial and embodied representations. Gärdenfors models concepts as regions in a geometric space, arguing that spatial reasoning bridges the gap between discrete symbols and subsymbolic data [5]. Semantic-interaction systems such as ForceSPIRE [4] map drag gestures to latent-model updates, enabling analysts to "teach" their algorithms in situ, while Drag-and-Track extends this paradigm to machine-learning parameter spaces [13].

However, many GenAI interfaces prioritise inspection and exploration over direct manipulation. Luminate, for example, lets users observe AI-generated content along multiple dimensions but restricts revisions to prompt-based commands [19]. Similarly, the AI Policy Projector visualises policy landscapes alongside a tabular authoring panel, but does not offer direct manipulation of policies within these spatial representations [9].

Some recent work has explored direct manipulation of intents in creative writing contexts. For example, PatchView [1] enables world-builders to rearrange generative "dust" patches to navigate and shape narrative spaces. We call this emerging interaction paradigm "Direct Intent Manipulation", and develop it further by extending it to the research ideation context.

3 SYSTEM DESIGN

Our system instantiates Direct Intent Manipulation by projecting AI-evaluated research intents onto a spatial canvas where users use direct manipulation to refine and merge them, preserving the immediacy of direct manipulation while operating on abstract intent.

3.1 Design of Direct Intent Manipulation

In this context, iteration allows users to refine, reframe, or evolve their research intents — represented as nascent research ideas — based on both human and AI evaluations [14]. Existing systems that support iterative development of research intents often employ node-link diagram interfaces to help users visualize and update hypotheses [3] and ideation trajectories [14].

Moving beyond iteration, merging provides a mechanism for users to evolve one intent into another by integrating multiple scientific or design intents [15, 18]. Prior tools for intent merging have primarily utilized modular, faceted grid interfaces [15, 18] or template-based approaches [12].

Building on this body of work, our prototype supports three core types of direct manipulation of research intents via drag-and-drop interactions within the spatial canvas:

(1) Revise a research intent along dimensions such as novelty and feasibility,
(2) Revise the evaluation of an intent, and
(3) Merge two intents to synthesize new intents.

3.2 Example Usage Scenario

Upon receiving the user's input query, "rl for user intent interface," the system generates three initial research intents and corresponding evaluation of its potential impact, feasibility, and novelty. Then the user selects the intent titled "Context-Aware RL UI." and aims to enhance its feasibility. By dragging this intent towards more feasible direction, the system generates a refined intent: "Simplified Context-Aware RL UI." The user then expresses interest in both "Probabilistic RL UI" and the newly generated "Simplified Context-Aware RL UI." To investigate a potential synergy between these two intents, the user performs a drag-and-drop operation to merge them. This interaction prompts the system to generate a synthesized hybrid intent "Hybrid Context-Aware UI."

REFERENCES

[1] John Joon Young Chung and Max Kreminski. 2024. Patchview: LLM-powered Worldbuilding with Generative Dust and Magnet Visualization. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*. ACM, Pittsburgh PA USA, 1–19. https://doi.org/10.1145/3654777.3676352

[2] Zijian Ding. 2024. Towards Intent-based User Interfaces: Charting the Design Space of Intent-AI Interactions Across Task Types. http://arxiv.org/abs/2404.18196 arXiv:2404.18196 [cs].

[3] Zijian Ding, Michelle Brachman, Joel Chan, and Werner Geyer. 2025. "The Diagram is like Guardrails": Structuring GenAI-assisted Hypotheses Exploration with an Interactive Shared Representation. (2025).

[4] Alex Endert, Patrick Fiaux, and Chris North. 2012. Semantic Interaction for Visual Text Analytics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 473–482. https://doi.org/10.1145/2207676.2207741

[5] Peter Gardenfors. 2004. *Conceptual spaces: The geometry of thought.* MIT press.

[6] Juraj Gottweis, Wei-Hung Weng, Alexander Daryin, Tao Tu, Anil Palepu, Petar Sirkovic, Artiom Myaskovsky, Felix Weissenberger, Keran Rong, Ryutaro Tanno, Khaled Saab, Dan Popovici, Jacob Blum, Fan Zhang, Katherine Chou, Avinatan Hassidim, Burak Gokturk, Amin Vahdat, Pushmeet Kohli, Yossi Matias, Andrew Carroll, Kavita Kulkarni, Nenad Tomasev, Vikram Dhillon, Eeshit Dhaval Vaishnav, Byron Lee, Tiago R D Costa, José R Penadés, Gary Peltz, Yunhan Xu, Annalisa Pawlosky, Alan Karthikesalingam, and Vivek Natarajan. [n.d.]. Towards an AI co-scientist. ([n. d.]).

[7] Mourad Gridach, Jay Nanavati, Khaldoun Zine El Abidine, Lenon Mendes, and Christina Mack. 2025. Agentic AI for Scientific Discovery: A Survey of Progress, Challenges, and Future Directions. https://doi.org/10.48550/arXiv.2503.08979 arXiv:2503.08979 [cs].

[8] Edwin L. Hutchins, James D. Hollan, and Donald A. Norman. 1985. Direct Manipulation Interfaces. *Human–Computer Interaction* 1, 4 (Dec. 1985), 311–338. https://doi.org/10.1207/s15327051hci0104_2 Publisher: Taylor & Francis _eprint: https://doi.org/10.1207/s15327051hci0104_2.

[9] Michelle S. Lam, Fred Hohman, Dominik Moritz, Jeffrey P. Bigham, Kenneth Holstein, and Mary Beth Kery. 2024. AI Policy Projector: Grounding LLM Policy Design in Iterative Mapmaking. https://doi.org/10.48550/arXiv.2409.18203 arXiv:2409.18203 [cs].

[10] Ruochen Li, Teerth Patel, Qingyun Wang, and Xinya Du. 2024. MLR-Copilot: Autonomous Machine Learning Research based on Large Language Models Agents. https://doi.org/10.48550/arXiv.2408.14033 arXiv:2408.14033 [cs].

[11] Chris Lu, Cong Lu, Robert Tjarko Lange, Jakob Foerster, Jeff Clune, and David Ha. 2024. The AI Scientist: Towards Fully Automated Open-Ended Scientific Discovery. http://arxiv.org/abs/2408.06292 arXiv:2408.06292 [cs].

[12] Stephen MacNeil, Zijian Ding, Kexin Quan, Thomas j Parashos, Yajie Sun, and Steven P. Dow. 2021. Framing Creative Work: Helping Novices Frame Better Problems through Interactive Scaffolding. In *Creativity and Cognition (C&C '21)*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/3450741.3465261 00000.

[13] Daniel Orban, Daniel F. Keefe, Ayan Biswas, James Ahrens, and David Rogers. 2019. Drag and Track: A Direct Manipulation Interface for Contextualizing Data Instances within a Continuous Parameter Space. *IEEE Transactions on Visualization and Computer Graphics* 25, 1 (Jan. 2019), 256–266. https://doi.org/10.1109/TVCG.2018.2865051

[14] Kevin Pu, K. J. Kevin Feng, Tovi Grossman, Tom Hope, Bhavana Dalvi Mishra, Matt Latzke, Jonathan Bragg, Joseph Chee Chang, and Pao Siangliulue. 2024. IdeaSynth: Iterative Research Idea Development Through Evolving and Composing Idea Facets with Literature-Grounded Feedback. https://doi.org/10.48550/arXiv.2410.04025 arXiv:2410.04025 [cs].

[15] Marissa Radensky, Simra Shahid, Raymond Fok, Pao Siangliulue, Tom Hope, and Daniel S. Weld. 2025. Scideator: Human-LLM Scientific Idea Generation Grounded in Research-Paper Facet Recombination. https://doi.org/10.48550/arXiv.2409.14634 arXiv:2409.14634 [cs].

[16] Ben Shneiderman. 1982. The future of interactive systems and the emergence of direct manipulation. *Behaviour & Information Technology* 1, 3 (1982), 237–256.

[17] Chenglei Si, Diyi Yang, and Tatsunori Hashimoto. 2024. Can LLMs Generate Novel Research Ideas? A Large-Scale Human Study with 100+ NLP Researchers. http://arxiv.org/abs/2409.04109 arXiv:2409.04109 [cs].

[18] Arvind Srinivasan and Joel Chan. 2024. Improving Selection of Analogical Inspirations through Chunking and Recombination. In *Proceedings of the 16th Conference on Creativity & Cognition (C&C '24)*. Association for Computing Machinery, New York, NY, USA, 374–397. https://doi.org/10.1145/3635636.3656207

[19] Sangho Suh, Meng Chen, Bryan Min, Toby Jia-Jun Li, and Haijun Xia. 2024. Luminate: Structured Generation and Exploration of Design Space with Large Language Models for Human-AI Co-Creation. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–26. https://doi.org/10.1145/3613904.3642400 arXiv:2310.12953 [cs].

[20] Bryan Wang, Yuliang Li, Zhaoyang Lv, Haijun Xia, Yan Xu, and Raj Sodhi. 2024. LAVE: LLM-Powered Agent Assistance and Language Augmentation for Video Editing. In *Proceedings of the 29th International Conference on Intelligent User Interfaces*. ACM, Greenville SC USA, 699–714. https://doi.org/10.1145/3640543.3645143

[21] Zehuan Wang, Jiaqi Xiao, Jingwei Sun, and Can Liu. 2025. IntentPrism: Human-AI Intent Manifestation for Web Information Foraging. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–11. https://doi.org/10.1145/3706599.3719744

[22] Zonglin Yang, Xinya Du, Junxian Li, Jie Zheng, Soujanya Poria, and Erik Cambria. 2024. Large Language Models for Automated Open-domain Scientific Hypotheses Discovery. In *Findings of the Association for Computational Linguistics: ACL 2024*, Lun-Wei Ku, Andre Martins, and Vivek Srikumar (Eds.). Association for Computational Linguistics, Bangkok, Thailand, 13545–13565. https://doi.org/10.18653/v1/2024.findings-acl.804

[23] Ryan Yen, Jiawen Stefanie Zhu, Sangho Suh, Haijun Xia, and Jian Zhao. 2024. CoLadder: Manipulating Code Generation via Multi-Level Blocks. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology (UIST '24)*. Association for Computing Machinery, New York, NY, USA, 1–20. https://doi.org/10.1145/3654777.3676357