# Hongyi **Du**

Research Assistant · M.S. Student

*Urbana-Champaign, IL, United States*

(+1) 217-778-5314 | hongyid4@illinois.edu | www.hongyidu.ai | Hongyi-Du | Google Scholar


## Education

**University of Illinois at Urbana-Champaign** — *Illinois, United States*

B.S. in Computer Science; M.S. in Computer Science; BS GPA: 3.79/4.00 — *B.S. 2022–2026; M.S. 2026–2027*

- A+/A grades in coursework including graduate Advanced NLP (A+, 100/100), Algorithms, Systems Programming, and Computer Architecture.

**Research Interests.** Multi-agent systems, self-improving agents, and human–AI interaction, with a focus on **organizations as the unit of coordination and learning in agentic AI**: how groups of agents and people coordinate, learn, and stay governable as members change.

## Selected Publications

**RELIC: From Multi-Agent Collaboration to Persistent Organizational Capability**

**Hongyi Du**, Tianyi Zhang, Weijia Zhang, Yi Yang, Haofei Yu, Kunlun Zhu, Tianxiang Dai, Shang Jiang, Zhelun Gao, Jiaxin Pei, Shang Zhu, Jiaxuan You | **arXiv:** 2609.32965 | **Under review** — *Sep. 2026*

**ProtocolBench: Which LLM MultiAgent Protocol to Choose?**

**Hongyi Du**, Jiaqi Su, Jisen Li, Lijie Ding, Yingxuan Yang, Peixuan Han, Xiangru Tang, Kunlun Zhu, Jiaxuan You

**arXiv:** 2510.17149 | **ICML 2026** | Protocol benchmark and learned router — *Oct. 2025*

**RINI: Seeing the Prior Is Not Enough**

**Hongyi Du**, Tianyi Zhang, Heng Wang, Zhelun Gao, Yimei Liu, Ambrose Luo, Annie Hao, Jiayan Ni, Jiawei Han, Jiaxuan You

**arXiv:** 2609.33284 | **Under review** | Research Idea Novelty Inspection — *Sep. 2026*

**Agentic Sociology: Relations and Institutions for Productive Agent Systems**

**Hongyi Du**, Jiaxuan You | **Preprint** — *2026*

**NatureSwarm: Nature-Inspired Coordination Helps Multi-Agent Language-Model Systems Where Central Control Reaches Its Limits**

Kunlun Zhu†, **Hongyi Du**†, Junyu Lu†, Zhenhailong Wang, Cheng Qian, Zhongzhu Zhou, Malavika Venu, Xiangru Tang, Xiaomin Li, Pan Lu, Shang Zhu, Jiaxuan You, James Zou, Heng Ji, Gene Robinson — *2026*

**Project page** | **In preparation for Nature Machine Intelligence** | †Core contributors (co-first authors)

**Ψ-Bench: Evaluating Persona-Sensitive Influencing in Persuasive Dialogues**

Peixuan Han, **Hongyi Du**, Jiayu Liu, Yihang Sun, Yutong Liu, Jiaxuan You

**arXiv:** 2606.02754 | **Under review** — *Jun. 2026*

**ModelingAgent: Bridging LLMs and Mathematical Modeling for Real-World Challenges**

Cheng Qian, **Hongyi Du**, Hongru Wang, Xiusi Chen, Yuji Zhang, Avirup Sil, Chengxiang Zhai, Kathleen McKeown, Heng Ji

**arXiv:** 2505.15068 | **Findings of EMNLP 2025** — *May 2025*

**MultiAgentBench: Evaluating the Collaboration and Competition of LLM Agents**

Kunlun Zhu*, **Hongyi Du***, Zhaochen Hong*, Xiaocheng Yang*, Shuyi Guo*, Zhe Wang*, Zhenhailong Wang, Cheng Qian, Xiangru Tang, Heng Ji, Jiaxuan You | **arXiv:** 2503.01935 | **ACL 2025 Main Conference** | *Co-first authors | 400+ citations — *Mar. 2025*

## Research Experience

**Research Leadership:** Currently lead and mentor nine research interns across four connected projects on agent organizations.

**RELIC: From Multi-Agent Collaboration to Persistent Organizational Capability** — *Champaign, United States*

Lead Researcher – Independent Research Project; advised by Prof. Jiaxuan You; with Shang Zhu (Together AI) — *Jun. 2026 – Present*

- **Content:** Develop RELIC, which turns recurring multi-agent collaboration failures into organization-owned, governed, executable protocols that persist outside member-local memory and continue to govern shared work after member replacement.
- **Protocol Lifecycle:** Design the end-to-end path from visible work friction to evidence-grounded rule proposals, validation and approval, runtime bindings over triggers, role responsibilities, required evidence, and execution consequences, followed by revision or retirement.
- **Environment and Evaluation:** Build an execution-grounded software-production organization (repositories, CI/review/merge, hidden acceptance tests, audit trails); run 360 controlled trials across ten workloads and three models, plus ablations, fresh-member transfer, ProgramBench, and CooperBench.
- **Results and SOTA:** Improve complete-contract delivery from 14.06% to 19.76% (+5.71 pp) over the matched structured control, with gains on all four verified endpoints in every model stratum. Executable transfer reaches 41.2% versus 34.6% for the same rules as text. RELIC achieves 371/469 (79.1%) on valid CooperBench pairs, the best reported peer-structured result, exceeding the strongest released Peer by 20.0 pp and hierarchical Team by 4.7 pp.
- **Human–Agent Organizations:** Designed a secretary agent that turns a person's remarks into organizational objects, such as evidence requests and standing review policies; formative walkthroughs showed that the bottleneck moved from understanding to timing.
- **Open-Source Software:** Released Relic Bridge, bringing Claude Code, Codex, and chat assistants into one Relic organization; used by other researchers.
- **My Role:** First author, team lead, and corresponding author; led the end-to-end work, including problem formulation, system architecture and implementation, benchmark and evaluator construction, experiment execution, analysis, visualization, paper writing, and release preparation.

### ProtocolBench: Which LLM MultiAgent Protocol to Choose?

*Champaign, United States*

Undergraduate Research Assistant at U Lab (UIUC) – supervised by Prof. Jiaxuan You — *May 2025 – Oct. 2025*

- **Content:** Developed ProtocolBench, a protocol-agnostic benchmark for LLM-based multi-agent communication (A2A, ACP, ANP, and Agora) and ProtocolRouter, a scenario-aware selector that chooses one protocol per module.
- **Method:** Built a three-layer evaluation stack (protocol adapters, scenario harness, metrics/logger) and designed representative scenarios (GAIA QA, safety tech, streaming queue, and fail-storm recovery) to jointly evaluate task utility, latency, resilience, and security.
- **My Role:** Acted as team lead and first author, responsible for benchmark design, scenario specification, protocol behavior analysis, and the end-to-end implementation and training of ProtocolRouter.
- **Outcome:** Showed that learned routing achieves 63.3% scenario / 81.7% module selection accuracy and matches or surpasses the best single-protocol baselines, including lower mean latency than ACP in streaming and faster recovery than A2A in fail-storm. Code | arXiv:2510.17149. **Accepted to ICML 2026**.

### RINI: Seeing the Prior Is Not Enough

*Champaign, United States*

Researcher – supervised by Prof. Jiawei Han — *Aug. 2025 – Present*

- **Problem and Evidence:** Show that access to prior work does not ensure correct scientific attribution: among 175 interpretable exposed proposals, 137 recognize the prior's relevance but only 61 correctly attribute the established contribution, while 37 of 71 proposed remaining distinctions are covered by the same prior.
- **Method:** Develop Research Idea Novelty Inspection (RINI), which links exact contribution claims to prior passages, audits ownership and coverage, checks whether the claimed remainder survives the same evidence, and applies at most four validated local replacements while preserving the research question, method, and evaluation plan.
- **Evaluation:** Five human annotators evaluate 1,080 original-revision pairs. On the same 240 originals requiring correction, RINI repairs 72.2%, versus 39.1% for Retrieve-and-Revise and 11.7% for Self-Revision; it improves over same-evidence direct revision by 33.0 pp and introduces zero new contribution errors across all 360 RINI outputs.
- **My Role:** First author, team lead, and corresponding author; lead problem formulation, controlled-exposure studies, system and benchmark design, annotation protocol development, experimental analysis, and paper writing.

### NatureSwarm (formerly EvolvingAgents): Bio-Inspired Design and Evolution for LLM-Based Multi-Agent Systems

*Champaign, United States*

Undergraduate Research Assistant at Blender Lab & U Lab – supervised by Prof. Gene Robinson, Prof. Heng Ji, and Prof. Jiaxuan You; in collaboration with Prof. James Zou and Shang Zhu (Together AI) — *Jun. 2025 – Jun. 2026*

- **Content:** Developed **EvolvingAgents**, a two-pillar bio-inspired LLM multi-agent framework that couples within-task coordination mechanisms with across-task organizational adaptation.
- **Mechanism Design:** Co-designed **stigmergic memory**, **typed reflex signaling**, and an **LLM-driven adaptive immune pathway** to address memory pollution, routine coordination overhead, and non-persistent failure recovery in LLM-based multi-agent systems.
- **Meta-Evolution:** Designed a **Meta-Evolver** that clusters recurrent experience into auditable runtime components with lineage, usage evidence, ratings, promotion, and retirement, enabling persistent adaptation beyond static role or workflow prompts.
- **Nature Environment:** Led the design and implementation of the **Nature Environment**, a long-horizon ecological-civilization testbed with partial observability, resource scarcity, seasonal pressures, hidden recipe discovery, cultural transmission, and civilization-level metrics.
- **Evaluation:** Contributed to experiments on SWE-bench Verified, DSGym, and Nature pilots, showing that bio-inspired mechanisms behave as **conditional affordances**: they help when supplying a missing coordination capability and interfere when adding directive overhead to an already competent policy.
- **My Role:** Core contributor (co-first author); proposed the civilization-style benchmark vision, led Nature scenario and evaluation design, coordinated interfaces across simulation, observation, memory, evolution, and analysis modules, and contributed to the paper's technical narrative and June 2026 manuscript completion.

### ATLAS: Adaptive Technologies for Learning and Reskilling

*Champaign, United States*

Research Lead for Graph-Grounded Personalization and Learner Modeling at UIUC ATLAS Moonshot Project – advised by Prof. ChengXiang Zhai, Prof. Jiaxuan You, Prof. Nancy M. Amato, Prof. Dilek Hakkani-Tür, and Prof. Jiawei Han — *May 2026 – Present*

- **Content:** Lead the graph-grounded personalization and learner modeling direction in the UIUC ATLAS Moonshot project, an Upskill-as-You-Go AI education and reskilling ecosystem connecting learners, training providers, and employers through trustworthy AI assistance.
- **Method:** Design Open Reskill Graph structures that decompose curricula, job requirements, and learning materials into atomic skill nodes, dependency edges, prerequisite relations, evidence links, and adaptive microlearning units.
- **Learner Modeling:** Develop uncertainty-aware learner-state representations that track mastery, missing prerequisites, evidence quality, learning goals, constraints, and preference signals for graph-grounded recommendation and diagnosis.
- **Evaluation:** Lead evaluation design for graph-grounded human-AI learning interactions, including skill-gap diagnosis, learning-gain measurement, recommendation quality, and agentic pedagogy metrics for distinguishing productive AI-pilot behavior from passive autopilot usage.
- **My Role:** Own the personalization and learner-modeling track, including problem formulation, graph schema design, mastery-state modeling, benchmark construction, evaluation protocol design, and alignment with the broader ATLAS ontology and reskilling graph.

### MultiAgentBench: Evaluating the Collaboration and Competition of LLM Agents

*Champaign, United States*

Undergraduate Research Assistant at Blender Lab – supervised by Prof. Jiaxuan You and Prof. Heng Ji — *Aug. 2024 – Feb. 2025*

- **Content:** Developed MultiAgentBench, a benchmark for LLM-based multi-agent systems across coding, games (e.g., Minecraft), and research-style tasks, focusing on planning, memory, and collaborative abilities.
- **Method:** Designed the Werewolf Arena framework to simulate decentralized social interactions and evaluate theory-of-mind reasoning, leadership, cooperation, and emergent distrust among LLM agents.
- **My Role:** Co-first author; conceived and independently implemented the Werewolf scenario and analysis pipeline, and performed detailed behavioral analysis of multi-agent interactions. Discovered the emergent conservative "witch" behavior in the Werewolf environment.
- **Outcome:** Showed that stronger foundation models tend to develop mutual distrust and internal friction in social deduction games; benchmark has been recognized as a standard reference for LLM-based multi-agent collaboration and behavior evaluation. Code | arXiv:2503.01935.