

# Yuxin Hou

Entrepreneur · AI Product Development · Computational Social Science

houyx15@gmail.com | +1 (609) 212-6931 | +86 135-2179-8217 | WeChat: houyx15

[github.com/houyx15](https://github.com/houyx15) | [houyx15.github.io](https://houyx15.github.io)

88 College Road West, Princeton, NJ 08544

## Education

|                                                                          |                      |
|--------------------------------------------------------------------------|----------------------|
| Princeton University (current)   Visiting Scholar                        | 2026–2027            |
| Peking University (current)   PhD, Computational Sociology; GPA 3.88/4.0 | 2024–2028 (expected) |
| Tsinghua University   M.A., Education; GPA 3.95/4.0                      | 2021–2024            |
| Tsinghua University   B.E., Mechanical Engineering; GPA 3.51/4.0         | 2015–2021            |

## Projects

### Higgz Technology | Co-founder and CTO 2016–2022

- **Core algorithm design:** design the company’s first semi-automated reasoning algorithm for an AI tutor, translating teaching needs into algorithms and product features.
- **Product development:** own zero-to-one AI education product design and development, delivering two generations of student and teacher products.
- **Technical team management:** build and lead a 20+ person team across product, design, engineering and QA; coordinate cross-functional execution and delivery.
- **Data feedback loop:** establish an education research team connecting learning analytics, teaching experiments and product iteration; translate learning feedback into product improvements.

### Mind Imprint | Independent product designer and full-stack developer 2026

- **Zero-to-one product design and development:** independently define requirements, design interactions and implement the full stack for student research, writing and teacher assessment; deliver the platform, website and source code.
- **Architecture and implementation of four core agents:** define agent responsibilities, task orchestration and tool use; combine context isolation, persistent project state and on-demand retrieval for cross-stage collaboration. Connect guidance and assessment through structured outputs, evidence validation and student feedback.

[Official Website](#) | [Product](#) | [Source code: GitHub](#)

## Research Experience

### Large-scale text analysis and pretrained model fine-tuning

- **Cross-societal public opinion structures:** develop text cleaning, LLM-assisted annotation, BERT pretraining and task fine-tuning for user-level topic and stance measurement. Compare opinion structures using correlation matrices, spectral analysis, bootstrap intervals and shuffled baselines.
- **Gender differences in social-media participation:** process 146.8 million posts from 225,000 accounts; apply significance tests and robustness checks across balanced samples, outlier treatment, time windows and lexicons to assess gendered consumption and expression.
- **Network structure and opinion polarization:** use network analysis and simulation to test how network structure, perception bias and conformity shape polarization, linking individual behavior to population-level opinions.

### LLM alignment bias and internal representations

- **LLM-based social survey simulation:** evaluate bias in LLM-generated surveys through marginal and joint distributions; implement empirical copula and resampling baselines, and use internal representation analysis to investigate bias sources.
- **Gender bias in large language models:** combine output-level bias measurement with internal representation analysis to investigate gender encoding and its association with generation bias.

## Publications

---

Yang, F., **Hou, Y.**, Pang, X. (2025). Imputing International Trust Survey Data Using Large Language Models. Quarterly Journal of International Politics, 10(4), 1–31. Online 2026. In Chinese.

**Hou, Y.**, Huang, J. (2025). Natural language processing for social science research: A comprehensive review. Chinese Journal of Sociology, 11(1), 121–157.

Yan, K., **Hou, Y.** (2023). Policy Diffusion of the Strategy of Reinvigorating China through Talent Development: Quantitative Research Based on Policy Literature from 2002 to 2021. Journal of National Academy of Education Administration, 7, 45–57, 77. In Chinese.

Yan, K., **Hou, Y.** (2022). The Construction of a Powerful Nation of Education in China: Policy Diffusion Process and Features. Tsinghua Journal of Education, 43(5), 79–92. In Chinese.

Yan, K., Jiang, Y., Wu, H., Zhang, Y., **Hou, Y.** (2022). Promoting the Discipline-Major Dual Governance System Reform: Rational Analysis and Realization Path. Higher Education Review, 10(2), 242–256. In Chinese.

## Working Papers

---

**Hou, Y.**, He, Y., Ni, C., Yu, J., Huang, J., Mei, W. Comparing Opinion Structures Across Societies: Different in Surveys, Similar on Social Media.

Yu, J., Ngai, R., Chen, T., **Hou, Y.** Embedding Gender Norms. (Authors in alphabetical order.)

**Hou, Y.**, Liu, A. Gendered Participation in Information Domains: Weibo Accounts' Circulation and Expression of Public Affairs and Celebrity Culture.

Huang, J., Xie, Y., **Hou, Y.** Your friends are more polarized than you: scale-free network biases opinion perception and boosts online opinion polarization.

Xie, Y., Liang, L., Yang, Y.M., Huang, J., Xie, Y., **Hou, Y.**, Ma, Z., Wang, W., Sun, Y., Zhang, C. The Race between AI Superpowers: Social Factors Favoring China over the United States.

## Additional Experience

---

|                                                                       |                  |
|-----------------------------------------------------------------------|------------------|
| <b>UNICEF China Education Office</b>   Project intern                 | 2023             |
| <b>Shiguang Poetry</b>   Data collection / product planning volunteer | 2022–2024        |
| <b>PEER</b>   Education program design and technical support          | 2020, 2021, 2023 |

## Technical Skills

---

- **Agent engineering:** Multi-agent orchestration, skills and tool calls, context isolation, persistent project state and long-term memory; structured outputs, evidence validation and human-in-the-loop interactions.
- **Pretrained models:** Language model pretraining and fine-tuning, including BERT; LLM-assisted annotation, text classification and stance detection; LLM bias evaluation and internal representation analysis.
- **Software engineering:** Python, Go, JavaScript / TypeScript and SQL; full-stack web development, Electron desktop apps, FastAPI services and PostgreSQL data modeling.
- **Product design:** Zero-to-one requirements research, interaction design, prototyping and delivery; AI education products, learning workflows, process assessment and student–teacher experiences.
- **Data analysis:** Large-scale text cleaning and social-media analysis; correlation and spectral analysis, bootstrap, significance testing, resampling and robustness checks; network analysis, simulation and empirical copulas.