

# Keigo Kusumegi

<https://keygoksmg.github.io/> | [kk929@cornell.edu](mailto:kk929@cornell.edu)

**Computational Social Scientist** and PhD Student in Information Science with expertise in network science, machine learning, and data-driven methods to study collective phenomena and innovation. Seeking to apply advanced analytical techniques to projects focused on **the broad impact of AI on scientific discovery and societal impact**

## Education

---

<b>Cornell University</b> (Ithaca, NY), PhD program in Information Science	Aug. 2023 –
Advisor: Yian Yin	
<b>University of Tsukuba</b> (Ibaraki, Japan), MSs in Policy and Planning Sciences	Apl. 2021 – Mar. 2023
<b>University of Tsukuba</b> (Ibaraki, Japan), B.S. in Policy and Planning Sciences	Apr. 2017 – Mar. 2021

## Publications

---

<sup>¶</sup> denotes equal contributions

1. **Kusumegi, K.**, <sup>¶</sup> Yang, X., <sup>¶</sup> Ginsparg, P., De Vaan, M., Stuart, T., Yin, Y. Scientific production in the era of large language models. *Science*, 390, 1240-1243 (2025) [link](#)
2. **Kusumegi, K.**, Yin, Y. The Specialization of interdisciplinary innovators in science and technology. (working paper)
3. **Kusumegi, K.**, Daniel E. A., Yukie S. Dissecting the Gender Divide: Authorship and Acknowledgment in Scientific Publications. (under review) [link](#)
4. **Kusumegi, K.**, Sano, Y. Dataset of identified scholars mentioned in acknowledgement statements. *Scientific Data* 9, 461 (2022). [link](#)
5. **Kusumegi K.**, Sano, Y. 2021. Citations and Gender Diversity in Reciprocal Acknowledgement Networks. Preprint at *arXiv:2104.01729* (2021). [link](#)

## Research Experiences

---

<b>Yin's Lab at Cornell University</b>	Aug. 2023 – Current
<i>Graduate Research Assistant</i>	<i>Ithaca, NY</i>
<ul style="list-style-type: none"><li>• Quantified the influence of Large Language Models on scientific production using large-scale publication metadata analysis applying difference-in-differences frameworks</li><li>• Developed a systematic approach to estimate AI usage in scientific articles and quantify its impact on writing styles and citation behavior</li><li>• Evaluated the application of AI across various scientific fields to acceleration in scientific discovery</li><li>• Explored the complex interaction between individual knowledge expertise and project alignment</li></ul>	

## Mentorships

- Supervised undergrad and master's student research projects, providing weekly guidance on methodology
- Provided technical assistance of coding and text analysis for Humanity PhD students at Summer Graduate Fellowship in Digital Humanities

## Sano Lab at University of Tsukuba

---

<i>Graduate Research</i>	Apr. 2021 – Jun. 2023
<i>Ibaraki, Japan</i>	
<ul style="list-style-type: none"><li>• Uncovered hidden academic collaboration networks by parsing contribution acknowledgments in scientific articles, providing a deeper understanding of team dynamics</li></ul>	

## Research Internships and Industrial Experiences

---

### G-RIPS

Jun. 2021 – Aug. 2021

*Research internship*

*Sendai, Japan*

- Designed and implemented an optimization algorithm of wireless base station placement using topology and graph theory

### OPTiM Corporation

Sep. 2020 – Oct. 2020

*Research internship*

*Tokyo, Japan*

- Developed CNN-based multi-object detection model for industrial quality control

### MARK Creative

Jan. 2019 – Jul. 2023

*Software engineer*

*Ibaraki, Japan*

- Designed and implemented graph database and its API for data analysis and system optimizations

## Awards

---

- |   |      |
|---|------|
| • Funai Overseas Scholarship (\$120K)                             | 2023 |
| • QUAD Fellowship (\$50K)   | 2023 |
| • University of Tsukuba President's Commendation                  | 2023 |
| • Tsukuba Graduate Study Scholarship                              | 2023 |
| • Nakagawa Ikueikai Scholarship                                   | 2021 |
| • Best Poster Award   | 2021 |
| • Scholarship from the Japan Society for The Promotion of Science | 2018 |

## Technical skills

---

Python • Polars • FastAPI • SQL • BigQuery • S3 • PyTorch • Neo4j • LATEX • Docker • Git • Bash