Gaurav Sett

Contact | gauravsett@icloud.com | linkedin.com/in/gauravsett | gauravsett.com

Education

RAND   **Washington, DC**
Incoming PhD Candidate in Policy Analysis — Pardee RAND Graduate School   *2024-09—*
- Technology and Security Policy Scholar.

Georgia Tech   **Atlanta, GA**
MS in Computer Science — College of Computing   *2023-01—2023-08*
- Machine Learning specialization; 4.0 GPA.

BS in Computer Science — College of Computing   *2019-08—2022-12*
- Artificial Intelligence and Human-Computer Interaction threads; 3.9 GPA, Honors Program.

Employment

RAND   **Washington, DC**
Technology and Security Policy Fellow   *2023-11—*
- Conducting research on standards and regulation for AI systems.
- Supporting requests from policymakers for technical expertise on AI governance.

Washington Post   **Washington, DC**
Data Science Intern   *2022-05—2022-08*
- Deployed transformer language model to analyze breaking news for journalists and editors.

Federal Reserve Board   **Washington, DC**
Economics Research Intern   *2021-05—2021-08*
- Built application regularly used by policymakers to analyze quarterly earnings calls.

Georgia Tech Research Institute   **Atlanta, GA**
Undergraduate Research Intern   *2020-05—2020-12*
- Gathered data and analyzed political discourse about COVID-19 on social media platforms.

Research

RAND

Publications
- Smith & **Sett** (2024). Considerations for Dual-Use Foundation Models with Widely Available Weights. *RAND Expert Insights.*
- Elmgren, **Sett**, & Smith (2024). Considerations on AI Model Red-Teaming and Standards. *RAND Expert Insights.*

Presentations
- **Sett** (2024). Responsible Capability Scaling. *Frontier Model Evaluation Science Day*, Washington, District of Columbia.

Georgia Tech

Publications
- Conklin, Bae, **Sett**, Hoffman, & Biddle (2023). Report on the Conference on Ethical and Responsible Design in the National AI Institute: A Summary of Challenges. *Forthcoming.*

Presentations
- Conklin & **Sett** (2023). AI Safety, Governance, and Alignment Tutorial. *2023 IEEE International Symposium on Ethics in Engineering, Science, and Technology (ETHICS)*, West Lafayette, Indiana.

Leadership

AI Safety Initiative at Georgia Tech
Founder, Executive Director *2022-08—2023-10*
- Founded group hosting events, seminars, bootcamps, and research projects in AI alignment and governance.
- Managed programs for over 100 students, leading team operations, and supporting career transitions.

Supervised Program for Alignment Research
Co-Founder, Program Manager *2023-07—2023-09*
- Launched program facilitating AI safety research projects for 15 advisors and 80 students across universities.
- Recruited participants, managed applications, and organized 18 projects in alignment, engineering, and policy.

Teaching

Georgia Tech
Graduate Teaching Assistant — Introduction to Cognitive Science *Spring 2023*
Graduate Teaching Assistant — Introduction to Cognitive Science *Summer 2023*

Skills

Machine Learning | Python | R | SQL | PyTorch | TensorFlow | HuggingFace | SciKit-Learn | Pandas | NumPy | Matplotlib

Software Engineering | Java | C | JavaScript | HTML/CSS | Git | APIs | AWS | React | Django | Agile