

Ayush Manish Agrawal

Phone: +1(929)-372-5500 | **LinkedIn:** <https://www.linkedin.com/in/ayushmagrawal> |

Email: ayush.agrawal7661@gmail.com | New York, NY

SUMMARY

Applied AI Engineer at Microsoft, specializing in production multi-agent LLM systems: end-to-end agent framework design, tool-use affordances, and ablation-driven evaluation. Shipped ARLI at enterprise Azure scale; currently building various knowledge-graph-based applications that autonomously resolve incidents across the org. Research foundation in deep learning interpretability and federated learning (NeurIPS 2020 workshops). Co-author, Grounding Enterprise Text-to-SQL in Tribal Knowledge — accepted at Microsoft MLADS 2026 (poster/talk), with camera-ready journal publication pending in MSJAR 2026. Deepening post-training, fine-tuning, and RL fundamentals in PyTorch, and building self-improving loop techniques for autonomous task development.

WORK EXPERIENCE

Applied AI Engineer

Microsoft | October 2022 - Present

- Own the high-level architecture, integration, and orchestration for CIS incident automation, enabling scalable triage, categorization, and mitigation workflows that reduce manual incident handling effort by up to 60%, improve SLA compliance, and support faster resolution of high-severity incidents
- Architected and built CIS Copilot (ARLI), an enterprise-scale multi-agent system enabling planners, SREs, and leadership to query real-time buildout status, incident impact, and post-analysis insights across CIS platform data using natural language
- Designed domain-specialized LLM tooling for Copilot, including schema-aware indexing, few-shot retrieval, fuzzy entity matching, and business-rule driven query enforcement, allowing agents to reason like subject-matter experts while maintaining accuracy and explainability under ambiguous inputs
- Implemented autonomous NL2SQL and real-time visualization workflows, enabling agents to generate SQL, execute queries, and dynamically render Mermaid-based operational visualizations, reducing reliance on manually maintained dashboards and improving operational decision velocity.
- Delivered ML-driven ETA prediction systems for GPU buildouts, achieving 79% accuracy within ± 7 days and outperforming SLA-based estimates in 83% of cases, with incident-aware rolling updates enabling sub-hour latency for live milestone refreshes
- Collaborating cross-team on new features for Microsoft Scout, Microsoft's always-on enterprise AI agent, to expand its autonomous, agentic capabilities

Software Engineer

University of Nebraska – Board of Regents | April 2020 – October 2022

- Built a Python library for automating project versioning and tagging during incoming merge requests optimizing our team's DevOps lifecycle by 30%
- Architected and delivered a full-stack issue-tracking application (.NET Core, React), integrating hardened input validation, whitelisting, and ReCaptcha to enhance platform security and drive a 55% increase in user issue reporting.

RESEARCH EXPERIENCE

Research Engineer

Manifold Computing | September 2020 – September 2022

- Designed and developed an innovative metric (RWC) for interpreting and visualizing the learning dynamics within Deep Neural Networks, utilizing established Computer Vision architectures as baselines
- Implemented and conducted a series of experiments employing the RWC in PyTorch to analyze and evaluate deep learning models
- Authored and submitted a paper in the NewInML Workshop at NeurIPS 2020, securing acceptance for a contributed talk

Research Engineer

OpenMined | July 2020 – December 2020

- Investigated various ML problems to gain insights into the complexity of systems from the perspective of data privacy
- Reproduced several Federated Learning (FL) algorithms using PyTorch under varying system parameters
- Published a proposal introducing a novel metric for assessing the complexity of Federated Learning (FL) systems, securing acceptance for a contributed talk at the workshop

PUBLICATIONS

- [Accepted] Agrawal, A.M. (2026), Forsbach, R, Gao, H, Triaui M, Duvurru, A. Grounding Enterprise Text-to-SQL in Tribal Knowledge. In: Microsoft Systems Journal of Applied Research (MSJAR), Introduces table rules - retrievable per-table constraint files authored by domain experts - as a complement to schema and few-shot retrieval in NL2SQL agents..
- Agrawal, A.M., Tendle, A., Sikka, H., Singh, S., Kayid, A. (2021). Investigating Learning in Deep Neural Networks Using Layer-Wise Weight Change. In: Arai, K. (eds) Intelligent Computing. Lecture Notes in Networks and Systems, vol 284. Springer, Cham. https://doi.org/10.1007/978-3-030-80126-7_48
- Ayush Manish Agrawal, Atharva A. Tendle, Harshvardhan D. Sikka, Sahib Singh, Amr Kayid. "Investigating Learning in Deep Neural Networks using Layer-Wise Weight Change." In Advances in Neural Information Processing Systems (NIPS) NewInML Workshop 2020.
- Mulay, Ajinkya, Ayush Manish Agrawal, and Tushar Semwal. "FedPerf: A Practitioners' Guide to Performance of Federated Learning Algorithms." In Advances in Neural Information Processing Systems (NIPS) Pre-register Workshop 2020

EDUCATION

University of Nebraska - Lincoln

Bachelor of Science, Computer Science

Minors: Mathematics, Business Administration

TECHNICAL SKILLS

Languages: Python, SQL, TypeScript, KQL

ML & Frameworks: Autogen, Langchain, AgentFramework, PyTorch, Hugging Face Transformers

Cloud Platforms: Azure, Google Cloud Platform