

# ARPAN KUMAR MAHAPATRA

[mahapata@purdue.edu](mailto:mahapata@purdue.edu) // [arpan.arpan.mohapatra@gmail.com](mailto:arpan.arpan.mohapatra@gmail.com)  
(+1) 765-775-0836 // [GitHub](#) // [LinkedIn](#) // [Website](#)

## EDUCATION

**Purdue University**  
Masters of Science, Computer Science

*January 2024 – December 2025*

**SRM Institute of Science and Technology**  
Bachelors of Science, Computer Science

*May 2017 - June 2021*

## PROFESSIONAL EXPERIENCE

**Eli Lilly & Company | Senior AI Engineer**

*January 2026 – Present*

- Built an org-wide cloud health alerting platform with SLO-driven routing, event ingestion, and incident triage automation, targeting 95%+ production coverage and <3-minute critical alert delivery
- Developed **Python FastAPI backend services** for an enterprise AI knowledge platform, implementing secure GitHub ingestion APIs that automatically parse Markdown, embed content for RAG, and enforce RBAC for self-service knowledge onboarding
- Designed a modular MCP integration framework with OAuth 2.0/PKCE authentication for Stack Overflow and Confluence and built Microsoft Teams Bot Framework backend integrations supporting rich-card responses and continuous knowledge ingestion from engineering discussions.

**Eli Lilly & Company | Tech Intern**

*May 2025 – August 2025*

- Built **Python FastAPI backend services** powering a multi-agent RAG onboarding assistant with REST APIs, tool orchestration, Dockerized backend services, knowledge ingestion, retrieval, and real-time tool execution; reduced onboarding time by 40%, maintained <2-second p95 latency, and achieved an 87% retrieval hit rate.
- Built a metadata-driven knowledge registry with Python-based ingestion pipelines, automated tagging, and retrieval workflows, transforming approximately 3,200 internal technical documents into a searchable AI-ready knowledge layer and reducing manual search effort by 60%.
- Developed CV pipeline for brand-compliance detection using frame-sequence analysis with Python, OpenCV and NumPy; evaluated on 8,400 frames (210 videos, 80/20 split) against a 67% majority-class baseline; achieved 92% accuracy, precision 0.89, recall 0.87, PR-AUC 0.91, with PR-AUC chosen over ROC-AUC to account for class imbalance in violation frequency

**Purdue University | Research Assistant**

*February 2024 – May 2025*

- Developed a deep learning-based annotation tool that improved efficiency of nuclei labelling in histopathology images by 30%
- Implemented semi-supervised learning and clustering methods reducing the need for labelled data by 60%
- Streamlined the annotation process, reducing manual effort by 35% and enhancing nuclei labelling precision, leading to better diagnosis and research outcomes

**Optum Global Solutions | Software Engineer**

*June 2021 - December 2023*

- Implemented automation software for **REST API** testing using Serenity BDD, Java, and Jenkins, handling 20 API's resulting in a 40% reduction in testing cycle time, 95% code coverage, and a 30% decrease in post-release API-related defects
- Developed comprehensive UI testing automation using Java, Cucumber, Cypress integrated with Jenkins, reducing manual testing efforts by 80% and increasing sprint velocity by 25% through faster feedback cycles and improved defect detection
- Delivered a data validation tool using Java, Selenium WebDriver, and Cucumber, achieving a 99% reduction in manual data input, leading to a 70% decrease in data-related defects and a 300% improvement in data processing throughput

## TECHNICAL SKILLS

|                              |                                                                                              |
|------------------------------|----------------------------------------------------------------------------------------------|
| <b>Languages:</b>            | Python, Java, JavaScript, TypeScript, C, C++                                                 |
| <b>AI / Agentic Systems:</b> | LangGraph, LangChain, MCP Server, Multi-Agent Systems, RAG, Vercel AI SDK, NLP               |
| <b>Machine Learning:</b>     | PyTorch, TensorFlow, OpenCV, NumPy, Pandas, CNNs, RNNs, PySpark                              |
| <b>Cloud / Backend:</b>      | AWS, Docker, FastAPI, Node.js, PostgreSQL, MongoDB, MySQL, Databricks                        |
| <b>Web Development:</b>      | React, Next.js, React Native, Tailwind CSS, GraphQL, Redux Toolkit                           |
| <b>Automation / DevOps:</b>  | Jenkins, Postman, Selenium WebDriver, Cucumber, Serenity BDD                                 |
| <b>Miscellaneous:</b>        | 2 <sup>nd</sup> Place in Eli Lilly Hackathon, Charter President of Rotaract Club (2019-2021) |

## PROJECTS

**AI Networking Platform**

*December 2025 – January 2026*

- Developed a full-stack AI networking application with a **Python FastAPI backend**, React Native frontend, MongoDB, and GPT-4o, exposing REST APIs for AI-Powered conversation generation and analytics; validated in a 42-user beta with 78% D7 retention and 61% activation.
- Implemented a streaming audio -> Whisper STT -> GPT-4o extraction pipeline delivering talking points, action items, and tone feedback; achieved 95% extraction accuracy (token-level F1, 300-sample held-out set), <4s p95 end-to-end latency, and ~\$0.006/audio-minute; error analysis surfaced overlapping speech and domain jargon as primary failure modes

## ACADEMIC PAPERS

- **AnnotateAnyCell: Open-Source AI Framework for Efficient Annotation in Digital Pathology** bioRxiv
- **RouteNet-Fermi: Network Modeling With GNN** arXiv:241205649