

EXPERIENCE

— D&IS Information Solutions and Services, Intelligent Automation Team

RUSH University Medical Center | Chicago, IL

Data Scientist - AI Solutions, December 2025 — Present

- Establish an enterprise AI asset repository and lead the data reconciliation process to govern the AI lifecycle, automate intake and post lifecycle tracking in Power Automate, and support data cataloging efforts.
- Develop high-throughput data integration pipelines in Microsoft Fabric to ingest data from SQL databases and REST APIs executing complex data orchestration, feature engineering, and data mapping to transform data streams for data modeling.
- Redesign data schemas, semantic data layers, and engineered features to track KPIs and statistically track and evaluate operational metrics across 15+ robotic process automations in Power BI.
- Serve as Solutions Architect for a cancer oncology AI initiative, designing data layer diagrams, integration schemas, and system architectures to securely ingest EPIC EHR data into Microsoft Fabric and Azure environments.
- Curate oncology patient cohorts from EPIC databases by harmonizing longitudinal lab metrics with unstructured clinical texts and performing rigorous agent output validation to test and validate an NLP-driven LLM orchestrator extracting tumor markers, clinical staging, and summarization of patient cases for weekly tumor board reviews.

— NCEH/ATSDR Public Health Informatics & Analytics Team

Centers for Disease Control and Prevention | Atlanta, GA

Data Scientist, September 2024 — November 2025

- Managed the development of over ten PowerBI dashboards end-to-end including discovery, deployment, and enhancement stages, ensuring product delivery, visual impact, and strategic insights to stakeholders.
- Performed data validation and quality analysis on the center's publications database with data extraction from two different APIs, redeveloping aspects of the data pipeline in Databricks to improve data quality, and eliminating data duplication by 20%.
- Orchestrated 9 data ingestion pipelines in Azure Data Factory to move and migrate data from SQL databases to dataverse to streamline business processes through automation.
- Deployed a custom optical character recognition (OCR) model achieving a 98% performance accuracy in extracting handwritten text from compliance forms in an automated document submission workflow in Power Automate.
- Developed a retrieval-augmented generation (RAG) chatbot agent as a proof of concept in Copilot Studio to orchestrate knowledge retrieval from a cloud database, validate data features, and automate document submission in a chat interface with real-time user response.

Data Engineer & Visualization Fellow, August 2023 — August 2024

- Engaged with various stakeholders including IT governance, application security, data modernization, human subjects regulation, and scientific clearance to collect and translate business requirements into data specifications for data modeling, dashboard design, and development of analytic tracking tools.
- Designed and developed 6 analytic dashboards within 9 months, achieving dynamic visuals exceeding stakeholder specifications and launching dashboards into production ahead of expected timelines.
- Awarded the director's award for excellence in data transparency through the development and implementation of PowerBI dashboards that improve operational performance and transparency in scientific clearance.

SKILLS

Orchestration & Cloud Tools: Databricks, Fabric, Power Automate, Dataverse.

Data Visualization & BI: PowerBI, Tableau, Matplotlib, Seaborn.

Programming & Data Analysis: Python (Pandas, NumPy, Sci-Kit Learn), SQL, R.

Machine Learning & AI: Copilot Studio, TensorFlow, Keras, Random Forest, SVM, GBoost.

EDUCATION

Jarvis College of Computing and Digital Media at DePaul University (DPU) | Chicago, IL

Master of Science in Data Science concentration in Computational Methods with Distinction | November 2022

CERTIFICATES

Cogito | EPIC // Certificate ID 1628444722 | February 2026

Caboodle Data Model and Clarity Data Model | EPIC // Certificate ID 1664202602 & 1664202603 | March 2026

AI Engineering | IBM // Credential ID 6G6KQ6LFTV09 | September 2024

Machine Learning | Stanford University, DeepLearning.AI // Credential ID 78KSKS5B6CXN | December 2023