

## SUMMARY

I'm a Data Scientist working in the healthcare space, mainly around clinical NLP and structured medical data. I enjoy taking messy, complex clinical information and turning it into something clear, usable, and meaningful. Most of my work revolves around building data pipelines and ML systems using Python, SQL, Databricks, and cloud tools. A big part of my interest is clinical coding systems like SNOMED CT and ICD, and how we can make medical data talk the same language across systems. For me, the most exciting part is working on problems where data actually connects to real human impact.

## EXPERIENCE

**Models Hub Team Lead**

Feb 2026 – Present

### John Snow Labs

- Collaborated with cross-functional teams to translate healthcare data into production-ready solutions.
- Built and maintained scalable data pipelines using Python, SQL, and Databricks for processing large-scale structured and unstructured healthcare data.
- Developed machine learning / NLP workflows for extracting insights from clinical datasets.

**Data Scientist**

Nov 2025 – Feb 2026

### John Snow Labs

- Worked on clinical coding and interoperability challenges (SNOMED CT, ICD mapping, structured medical data representation).
- Improved data quality and consistency for downstream analytics and ML use cases.

**Data Scientist — Python | Web Scraping | Automation | ML | DL**

Mar 2020 – Sep 2025

### Self-Employed

- Worked on a variety of freelance software projects, focusing on practical data-driven and automation-based solutions.
- Developed Python-based tools and scripts for data processing, automation, and web scraping tailored to client needs.
- Designed and implemented lightweight applications and backend logic to support data workflows and business processes.
- Built and maintained data pipelines and integrations across different systems depending on project requirements.
- Translated loosely defined client ideas into structured technical solutions, from prototype to working implementation.

**Data Scientist**

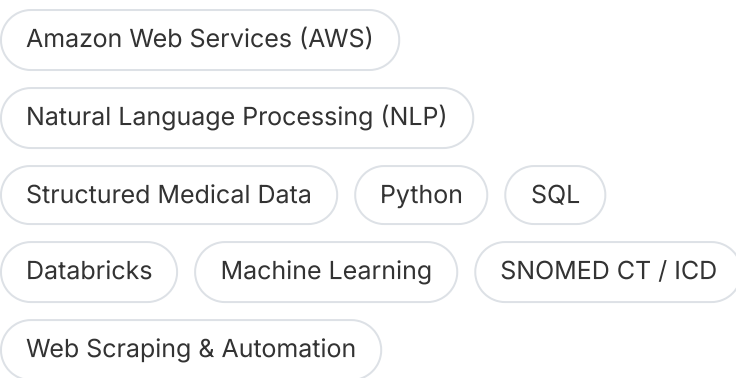
Dec 2018 – Mar 2020

### yDataLabs

- Built end-to-end data workflows using Python and SQL, with a strong emphasis on collecting and structuring data from various web sources.
- Developed and maintained scraping systems to gather large volumes of data, then cleaned and organized it for analysis.
- Applied machine learning techniques on collected datasets to explore patterns, trends, and predictive signals.

- Built feature sets and prepared training data to support ML experiments and iterative model improvements.

## SKILLS



## CERTIFICATIONS

AWS Partner: Technical Accredited — Training Badge

## WRITING

[Clinical De-Identification at Scale: Pipeline Design and Speed–Accuracy Trade-offs Across Infrastructures](#) — John Snow Labs / Medium

[From Doctor Notes to Structured Data: The Hidden Pipeline of Healthcare Systems](#) — Medium