John Snow Labs
Accelerating AI for Healthcare

The year is 1854; the streets of London's West End have become a place of death and disease due to a sudden outbreak of Cholera that claimed hundreds of lives in a matter of weeks. With the discovery of the connection between diseases and microbes being decades away, a physician named John Snow set out to end the outbreak with a unique approach. Snow collected information from the families of the deceased and the ailing and plotted it on a map of the area. This integration of data led him to the source of the outbreak—a water pump on Broadwick Street. After convincing the local council with his studies on the disease pattern, the source pump was shut down, and the outbreak was stopped. Today, John Snow Labs—named after the famed Victorian-era physician—assists healthcare organizations to collect and prepare data for analysis, process natural language, and deploy artificial intelligence (AI) platforms to boost operations.

Healthcare has always been a complex sector to innovate for given the bigger regulatory, technology and talent challenges. Most new technologies that disrupt the business world don't apply out-of-the-box in a healthcare context and require substantial investment and domain expertise to apply. John Snow Labs is a team of specialists built to address this: a third of the company has a Ph.D. or M.D. degree and 80 percent has at least a master's degree. The company counts Intel, Johnson & Johnson, and Roche Pharmaceuticals among its customers.

John Snow Labs offers three primary solutions for healthcare organizations that are planning to incorporate data science into their operations—an AI platform, a frictionless data market, and the Spark Natural Language Processing (NLP) library.

The AI platform is designed to accelerate data science teams in building and deploying AI models. The prime challenge for industry executives today is turning the machine learning (ML) and AI models developed by their data scientists into systems that can be utilized reliably on a production scale. In healthcare, this challenge goes a step further due to the increased rate of model decay caused by the changing patients, hospital clinical guidelines, and local workflows. With John Snow Labs, organizations can deploy and manage hundreds of models—all under the strict security and compliance requirements of healthcare.

For organizations on the lookout for best practices and design patterns on building machine learning pipelines and AI models, look no further than John Snow Labs. The AI platform offers out-of-the-box, enterprise-grade components that cover everything from visualization and collaboration tools for data scientists, to versioning and continuous integration of trained models. The platform is cloud agnostic with existing production deployments on AWS, Azure, and on-premise.

Next up is John Snow Labs' NLP Library for Apache Spark. This is an open source text processing library that's built on top of Apache Spark ML and packages state-of-the-art natural language understanding in easy to use Python and Scala libraries. The library raises the bar on accuracy, performance, and scale in open source NLP. Public benchmarks have shown it to be by far the fastest NLP library on a single machine. It's also the only open source NLP library that can scale to use a cluster. It supports training custom models—essential to support specific domains or languages—and is the only NLP library with optimized builds for deep learning on the latest Intel and Nvidia chips.

Beyond offering commercial support for the library, John Snow Labs licenses Spark NLP for Healthcare. This is an extension of the open source library that productizes a set of healthcare-specific deep learning NLP models. A team of data scientists regularly tracks new research papers to ensure this library stays state-of-the-art—that is, productizing the best performing peer-reviewed papers published to date. Spark NLP for Healthcare was the first to market with the most recent deep learning models for biomedical named entity recognition, negation detection, and entity resolution. It remains the only collection of production-grade, trainable, and scalable implementation of the best performing models for these NLP tasks published in the past two years.

Lastly, the Frictionless Data Market enables data scientists to speed up their progress by leveraging a catalog of over 1,800 datasets on terminologies, measures, providers, guidelines, and 15 other categories. Every data set is curated by a clinical domain expert: the catalog is the result of more than 50 man-years of detailed work by teams of doctors, pharmacists,

public health and medical billing experts. Every dataset undergoes three levels of quality review before its release, to ensure the delivery of clean data to data scientists. Datasets are normalized into one unified type system and come with detailed metadata. The data is kept always up to date—and is easily accessible from 26 analytics platforms.

> **"**
>
> ## John Snow Labs enables healthcare organizations to deploy state-of-the-art artificial intelligence (AI) platforms, models and data in production today
>
> **"**

Customer success is a core value for the company, and each time a customer comes on onboard, John Snow Labs takes a consultative approach to understand their exact needs. This has proven essential since many organizations are taking their first steps into AI, so gaining lessons learned from their first projects is a priority. Consulting and support are provided for all subscriptions directly by the R&D teams—so a data scientist with a question about optimizing healthcare NLP models, for example, can speak directly with the person who coded them. John Snow Labs is a virtual company whose people are in 17 countries and 14 time zones, so someone is almost always nearby.

Highlighting the impact of John Snow Labs' solutions is their success story with healthcare giant Kaiser Permanente. John Snow Labs helped the healthcare services company build an AI platform and assisted their enterprise architecture team in integrating data from medical facilities and developing a model that forecasts patient flow in a hospital. Further, in order to predict who will be admitted from the emergency department, John Snow Labs deployed custom deep learning
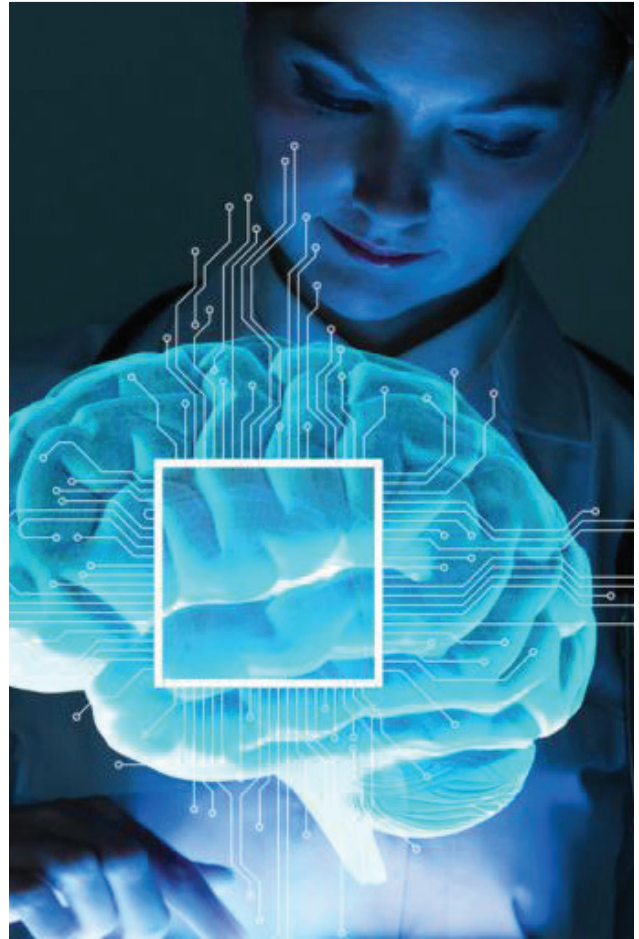
NLP models that mined patient records. This work was presented as a joint case study at the Strata Data Conference earlier this year.

Going forward, John Snow Labs will continue focusing on putting the latest academic research, novel innovations and AI best practices into their customers' hands. Recent releases include a complete data integration, quality & monitoring component; the first scalable model serving and management solution for high-compliance industries; and advanced OCR & data extraction capabilities from long free-text documents. All in all, the company firmly believes that AI is an emerging space for the healthcare and life science industries but one with tremendous scope. John Snow Labs is all set to break new grounds and help organizations realize this potential as fast as possible. **CR**