StreamSets

USE CASES

Predict Readmission

Hospitals and care facilities want to understand admission rates to staff and allocate resources effectively. Detecting patterns in anomalous readmission rates for certain patient segments can lead to actions that improve patient outcomes.

Predict Sepsis and Rapid Response

Oftentimes caregivers only know a patient has sepsis when it is already affecting their health. Monitoring connected devices can detect early warning signs before a patient is effected and trigger action.

Medical Image Repository

Creating a data repository for scanned images and applying deep learning, NLP, and text analysis to the images.

> "StreamSets counts 4 of the top 20 healthcare companies in the U.S amongst its customers."

Genomics and Precision Medicine

Companies are using big data platforms and large corpi of data to map the human genome. These activities can give way to delivering precision medicine based on genetic predispositions.

Real World Evidence for Clinical Trials

Practitioners and patients agree that it takes too long to bring life saving drugs to market. Using more data and implementing simulation and predictive modeling companies can shorten the time to market for crucial drugs.

Clinical Data Lake and Member 360

Creating a single view of customer health that can be utilized by caregivers, insurance providers, and claims adjusters.

StreamSets in Healthcare

Overview

Data is the lifeblood of the healthcare industry, but the world of data is exploding, offering new opportunities and challenges in the form of pervasive intelligence, which is the ability of myriad analytics groups to self-service ready-to-use healthcare data. Machine learning and artificial intelligence, applied by groups throughout an organization, can help in a variety of areas from predicting sepsis rates, driving better accuracy in claims prediction, and delivering precision medicine; the end result being better patient outcomes and more cost-effective service delivery.

Challenges

In the Healthcare industry, things are changing all the time. Bringing in all available data regardless of size, shape, or structure and landing it in a centralized data hub is the goal of many organizations, but how do they achieve that without soaking up resources that distract them from their core focus of healthcare innovation?

Priorities for managing healthcare data include:

Leveraging New Data Sources

Healthcare companies around the world are tasked with digitizing the wide breadth of medical records, leveraging instrumented medical devices for predictive capabilities, and optimizing health care delivery at facilities. New standards like HL7 provide a framework for the exchange, integration, sharing, and retrieval of electronic health information but are there are variations in how different organizations implement it.

Establishing Patient 360

Healthcare companies want to have a full 360 degree picture of a patient's health but often times that information spans systems and is siloed in separate locations.

Enabling a Unified Data Lake

To overcome data silos, as well as issues of growing data volume and variety, a centralized enterprise data lake provides many benefits. However, ingesting data into the lake usually requires custom coding and specialized skills, translating to high costs and lengthy projects that delay getting data to data scientists and analysts.

Protecting Patient Data

As companies move to modern data platforms, the inspection, detection and disposition of personal health information (PHI) is a challenge, especially when data is in motion. HIPAA requirements are increasingly unforgiving and often evolve at a pace that is difficult for companies to adapt in practice. This involves protecting data at origin, in-flight, and at its destination.

StreamSets

USE CASES

Pharma Supply Chain Optimization

Understanding areas by geo and cohort to deliver better stocks and educational material about available pharmaceutical solutions.

Healthcare Plan Fraud

Sometimes patients, providers and facilities mis-report services and charges to healthcare providers. This results in lost capital that is categorized as fraud. By looking at greater trends and detecting anomalies healthcare providers can identify potential fraud before payment is made.

Population Health/Medicaid Services

Federal efforts to create operational efficiency in delivering and paying for medical care. Limited budgets and an abundance of empirical data make these initiatives paramount for government agencies.

ABOUT STREAMSETS

StreamSets transforms how enterprises flow big and fast data from myriad sources into data centers and cloud analytics platforms. Its DataOps platform helps companies build and operate continuous dataflow topologies, combining award-winning open source data movement software with a cloud-native Control Hub. Enterprises use StreamSets to enable cloud analytics, data lakes, Apache Kafka, IoT and cybersecurity.

Founded by Girish Pancha, former chief product officer of Informatica, and Arvind Prabhakar, a former engineering leader at Cloudera. StreamSets is backed by top-tier Silicon Valley venture capital firms, including Battery Ventures, New Enterprise Associates (NEA), and Accel Partners.

For more information, visit **streamsets.com**

Solution

Streamsets helps healthcare companies cost-effectively make new data sources like bedside equipment, patient records and radiology images available to analytics teams and applications, while ensuring data is protected in-motion.

StreamSets Data Collector helps companies design and run batch and streaming pipelines in a fraction of the time using a drag-and-drop environment that minimizes coding and facilitates collaboration. It also detects and handles data drift, which may manifest as added fields or changed data types that can occur without notice when data sources are upgraded.

StreamSets Control Hub allows for management of hundreds of complex data flow topologies, giving you end-to-end visibility into your data movement. This helps modern healthcare organizations develop a DataOps practice and manage the health, delivery, and security of the critical pipelines feeding their analysis and discovery.

StreamSets Data Protector discovers, protects and governs sensitive data in stream. Data Protector can detect PII and HIPAA regulated data based on a large number of standard or custom templates and can apply many obfuscation and routing actions to the data.

Impact

While the problems that healthcare companies face are indeed complex in nature and transformative in scope, the business impact and strategic advantages are paying off for those companies that decide to leverage data in a way that addresses future compliance requirements.

Companies like **Availity** use StreamSets to ingest data into a vast repository which helps lower cost and increase data discovery. Pharma companies like **GlaxoSmithKline** use StreamSets to accelerate the drug development process from 10 years down to 2.

Closing

StreamSets is helping healthcare companies leverage the data they need to realize business value from a growing list of strategic use cases. By providing a platform that intelligently handles the eminent evolution of data systems and practices, monitors and reports on data pipeline operations comprehensive, and removes barriers from leveraging complex data types healthcare companies can build groundbreaking functionality that impacts the health and well-being of our society.

Want to know more about how StreamSets can help your healthcare use case? **Contact a representative today.**