

SOLUTION BRIEF

Healthcare and Life Sciences
Data Analytics Solutions



Montefiore Creates Data Analytics Platform to Advance Patient Care

Addressing value-based healthcare with Intel® Xeon® processors and Franz AllegroGraph*

“You have to have the appropriate information on the person sitting in front of you to make sure that you’re serving them correctly. If that’s not saving lives, I’m not sure what is.”

Andrew Racine,
System Senior Vice President
and Chief Medical Officer,
Montefiore

Challenge

Located in the Bronx, Montefiore Health System serves one of the most ethnically and socioeconomically diverse populations in the US. The complex includes, but is not limited to, Montefiore Medical Center, Albert Einstein College of Medicine, and a research facility. Unlike a pay-per-service model, as an accountable care organization Montefiore delivers value based on patients' long-term health—during their hospital or clinic visit and after they return to the community.

Like all healthcare organizations, Montefiore faces complex challenges—from government pressures to reduce costs and stringent regulatory guidelines to diverse patient populations and disruptive technologies. A focus and investment in precision medicine has brought the nation's and industry's attention to expanding the breadth of patient data in order to personalize treatment for individuals and historically underrepresented groups. Further, understanding patients requires information on a complex array of factors, some of which may not even be known during a clinical interaction, such as the home and work environment, nutrition, and genetics. Says Dr. Andrew D. Racine, system senior vice president and chief medical officer at Montefiore, “The challenge where you’ve got hundreds of thousands of patients impacting the institution at any given point is to have the appropriate information about each one of those patients at the fingertips of the therapist who’s interacting with them at the time of that interaction.”

The industry has long collected data on patients—it is not uncommon for hospitals to gather more than 100 data points per patient per day. Data ranges from unstructured free text information to images and waveforms to data from sensors and monitoring devices. Access to accurate data is vital for assessing risk from intubation to drug interactions. But often this data cannot be analyzed quickly, nor can hospital data be easily combined with external data sources such as those from pharmaceutical companies and researchers. Many medical facilities are not equipped to support the significant processing and memory requirements of big data analytics. Healthcare organizations that do integrate advanced technology often find these solutions costly and labor intensive to maintain. And HIPAA and other patient privacy regulations can make accessing and sharing data difficult.

Solution

To optimize healthcare based on advanced data analytics and make sure clinicians have the right information available in time to impact patient outcomes, Montefiore has deployed Semantic Data Lake*, a solution that brings together varied and vast amounts of raw data for deeper analysis to flag patients who are at risk or help clinicians identify optimal treatment plans. Drawing on extensive experience in patient care and medical research, this innovative solution enables relevant data to directly inform and impact patient care.

Montefiore

The platform integrates both structured and unstructured data ranging from basic science, clinician records, and population demographics to community, environmental, behavioral, and wellness research data. By assessing a holistic and realistic profile of patients—along with relevant science, clinical population histories, drug information, and medical imaging—Semantic Data Lake has the capability to improve care, identify at-risk patients, and personalize medicine, while reducing error and inefficiency.

Because the solution operates in near-real time, caregivers have access to timely, pertinent, actionable insight at the point of care. All of this is enabled by an automated system based on clinically relevant algorithms that parse massive data sets at lightning speed. This allows clinicians and medical researchers without technology expertise to concentrate on their discovery process. Semantic Data Lake provides insight based on evolving data. This insight may pinpoint effective treatments, identify risks, alert caregivers to critical care situations where timely interventions may be necessary or preventive, and highlight similarities and differences from similar cases.

The solution is groundbreaking. Says Dr. Parsa Mirhaji, director of Clinical Research Informatics at Montefiore and Einstein, “The ability to conduct real-time analysis over new combinations of data such as patient information, genetic data, medical device data, clinical trials, drug information, and public health data will fuel discoveries, significantly improve efficiencies, and personalize care.”



The Semantic Data Lake helps Montefiore identify and address critical issues for intervention at the point of care

Technical details

Semantic Data Lake is built on Intel® architecture with Cloudera's Hadoop* distribution and Franz's AllegroGraph, a high-performance semantic graph database enabling analytics based on industry standards. The data lake includes data from Montefiore's own institutions, as well as from sources such as the PharmGKB databank (which correlates genetic variations and drug responses), the National Institute of Health's Unified Medical Language System (UMLS), and the Online Mendelian Inheritance in Man, a continuously updated catalog of the human genome and genetic disorders.

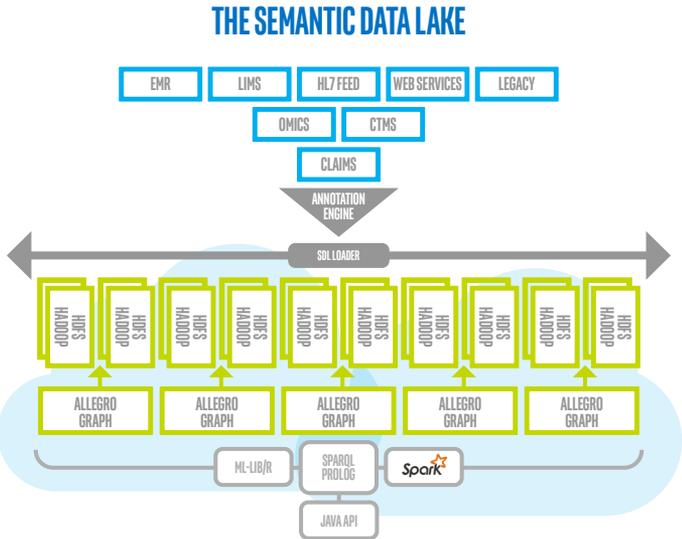
Semantic Data Lake automatically injects metadata into the data layer to create ties between different components of the data. This makes the data “discoverable” via an innovative interrogation mechanism. On top of the basic data framework, Montefiore injects knowledge bases and ontologies—basically the semantics that characterize the data and gives it meaning and a specific context.

The solution provides a breadth of data management technologies including:

- Incorporation of a semantic data lake in which myriad information types such as billing codes, patient events, medical procedures, and more are centralized and codified using semantic relations, industry standards, and domain-specific knowledge bases.
- Deployment of predictive analytics at a scale to leverage this data to anticipate and account for various patient outcomes in timeframes in which treatment can be administered to affect care.
- Use of machine learning algorithms to integrate the results of previous outcomes that significantly impact the analysis and effects of future patient objectives.
- Incorporation of an ontological pipeline that rapidly integrates new data sources (e.g., wearables, home sensing technology, and IoT) and requirements into existing models, and validates the clinical process for highly targeted patient subsets.
- Inclusion of disposable data marts akin to sandboxes that quickly provision analysis-ready, project-specific environments in which data scientists can manipulate data and analytics results, while simultaneously allowing others to leverage that same data for their own purposes without duplication of data or redundant pipelines.

Running on high-performance Intel® Xeon® E5-2690 v3 processors, the solution delivers the performance, reliability, and scalability to handle massive data processing in near-real time. Security is built into the Intel® processors at the hardware level, providing enhancements such as hardware-accelerated encryption and role-based access control. Intel's work on optimization of Cloudera and open source Apache Hadoop provides an agile foundation for memory-intensive computation. The Intel and Cloudera platform operates as a distributed framework, with the ability to easily scale out as demand grows. AllegroGraph allows Montefiore to extract sophisticated decision insights and predictive analytics from highly complex, distributed data in a way that is not possible using conventional databases.

Montefiore worked closely with Intel throughout the development process. Intel donated the hardware and contributed expertise on Cloudera and Hadoop, data integration, networking infrastructure, analytics, and security. This allowed Montefiore to focus on its understanding of healthcare priorities and interactions and realize its vision faster without the technical hurdles.



Semantic Data Lake architecture enables open-ended queries with integrated responses from multiple data sources

Solution Ingredients

- Montefiore's Semantic Data Lake
- Intel Xeon E5-2690 v3 processors
- Franz distributed AllegroGraph



Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software, or service activation. Performance varies depending on system configuration. No computer system can be absolutely secure. Check with your system manufacturer or retailer to learn more at intel.com.

Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase. For more complete information about performance and benchmark results, visit intel.com/performance.

Cost reduction scenarios described are intended as examples of how a given Intel-based product, in the specified circumstances and configurations, may affect future costs and provide cost savings. Circumstances will vary. Intel does not guarantee any costs or cost reduction.

1. Montefiore data, 2017.

Intel, the Intel logo, and Xeon are trademarks of Intel Corporation or its subsidiaries in the U.S. and/or other countries.

*Other names and brands may be claimed as the property of others.

Impact

For its first pilot of the Semantic Data Lake, Montefiore and Einstein chose to address a pressing issue for accountable care: identifying high-risk patients in need of critical, time-sensitive intervention. Says Dr. Mirhaji, "We wanted to use predictive analytics to flag any patient hospitalized at Montefiore Health System locations who is at risk of death or in need of intubation within the following 48 hours, which is the window of opportunity to complete an effective intervention."

The combination of automated predictive and preventive approaches has proven effective in identifying patients at highest risk, and providing consistent clinical decision support to relevant practitioners, just in time and throughout the entire Health System. Accurate prediction of prolonged ventilation detects patients with more than 70% likelihood of an event, 48 hours in advance of a fatal episode or respiratory failure in the hospital.¹

Montefiore and Einstein drew on historical data from its own patient population and that of the Mayo Clinic to define a predictive algorithm that would generate a score identifying the level of patient risk. The system also helps advise clinicians on the most effective next steps. Explains Dr. Mirhaji, "It creates risk scores based on the patient's likelihood of a major event. Then there's another engine that kicks in based on those risk scores and other factors to determine what we can do for that particular patient to avoid the crisis. It can send a personalized checklist of proposed interventions to the practitioner in charge of that case."

Intubation identification is being used successfully at Montefiore, but it represents only one of nearly unlimited possibilities Semantic Data Lake offers for improving patient care and revolutionizing care models and medical research.

Racine summarizes the impact: "Semantic Data Lake gives us the ability to search the information that's available to us in a much more efficient way in order to find patterns in the data and to use those patterns for clinical purposes. This is now being instituted throughout our critical care systems in all of the hospitals within the medical center."

Where to get more information

For more information on Intel® data center solutions, visit intel.com/analytics.

To explore the Montefiore Health System, visit Montefiore.org.

To learn about Franz semantic graph technologies, visit franz.com.