



## Uncovering Breakthrough Insights that Help Save Lives

“We were looking for a company that was as curious about the data as we were. With Cloudera, we established a relationship of discovering what was possible.”

— Aaron Black, Chief Data Officer, Inova Translational Medicine Institute

Inova is a not-for-profit integrated health system that serves more than two million people each year.

### Challenge

Which treatment is most effective for each patient? The Inova Translational Medicine Institute works to answer this question by driving innovation in precision medicine. But to achieve its goal, Inova faced two significant challenges: bringing together massive volumes of genomic and patient data for advanced analysis, and enabling faster exploration of that data.

“With our previous data warehouse, it could take weeks and months to pull data together for researchers,” said Aaron Black, chief data officer of the Inova Translational Medicine Institute. “Additionally, our scale was getting so big that we couldn’t continue on that path.”

### Solution

Inova has generated petabytes of genomic and patient data, and needed to provide a way to process that data into a single data infrastructure. After processing and optimizing this data, Inova provided its researchers with fast access to terabytes of genomic and patient data in a single data set using a **Cloudera analytic database**. With access to a wider range of data and the ability to more easily explore the data, researchers can more quickly test new theories and uncover new patterns that may not have been apparent before.

“Before, researchers spent 80 percent of their time on data wrangling, and only a little sliver of their time on the analytics,” said Black. “We’re now in the process of reversing that.”

### Implementation

In its search for a modern data platform, Inova sought a collaborative approach. “We looked for a company that was as curious about the data as we were,” said Black. “With Cloudera, we established a relationship of discovering what was possible.”

To gain executive buy-in, Black’s team demonstrated the expected return on investment through a Proof of Concept. “While we ultimately implemented Cloudera on-premise, we chose Cloudera on **Amazon Web Services** for our Proof of Concept because it was easy to build the cluster without spending a lot of upfront capital,” said Black. “Once we made our decision and built the on-premise cluster, it only took a few weeks to bring the entire dataset down to the cluster on-premise.”



## Key Highlights

### Industry

- Healthcare

### Location

- Headquarters: Falls Church, Virginia, USA

### Business Applications Supported

- Clinical research
- Genomic research

### Impact

- Reduces query times from hours and days to seconds and minutes
- Increases research productivity with a greater percentage of time spent on research versus data integration
- Provides insights to deliver precision medicine and improve patient health

### Data Sources

- External vendors
- Electronic health records
- Lab data
- Patient surveys

### Solution

- Modern Data Platform: Cloudera Enterprise
- Workloads: Analytic Database
- Components: Apache Impala (incubating)

### Big Data Scale

- Petabytes and growing

## Results

Inova researchers can now answer questions magnitudes faster than they could previously. It's a major step in improving healthcare as new medical discoveries can dramatically change treatment plans, and, ultimately, patient outcomes.

“Our goal was to match the speed at which researchers think,” said Black. “What Cloudera has done is made that possible. Now we’re moving towards getting answers in minutes and seconds and can find correlations that we couldn’t before. Ultimately, we can put the data together in novel ways to understand the evolution of diseases so that we can help keep our patients well.”

## About Cloudera

Cloudera delivers the modern platform for machine learning and advanced analytics built on the latest open source technologies. The world's leading organizations trust Cloudera to help solve their most challenging business problems by efficiently capturing, storing, processing and analyzing vast amounts of data. Learn more at [cloudera.com](http://cloudera.com).

---

[cloudera.com](http://cloudera.com)

1-888-789-1488 or 1-650-362-0488

Cloudera, Inc. 1001 Page Mill Road, Palo Alto, CA 94304, USA

© 2017 Cloudera, Inc. All rights reserved. Cloudera and the Cloudera logo are trademarks or registered trademarks of Cloudera Inc. in the USA and other countries. All other trademarks are the property of their respective companies. Information is subject to change without notice.