

CELMATIX

# Next-generation Women's Health Company Relies on a Streaming Platform to Help Them Transform Reproductive Health Care



## Website

<https://celmatix.com/>

## Industry

Bio-tech

## Challenge

Needed to better support physicians with relevant information and extract data from Electronic Medical Records, which are known for being complicated, messy and highly heterogeneous.

## Solution

Implemented Confluent Platform to stream Electronic Medical Record data securely and in real time to help doctors and patients make data-informed decisions faster.

## Results

- Harmonized and managed disparate data.
- Achieved market differentiation with the ability to scale and expand.
- Enables access to vast and disparate data in real time.

Celmatix is changing the face of women's health. Through the development of digital tools and genetic insights focused on fertility, they are disrupting how women approach their lifelong reproductive health journey by empowering them and their physicians with more personalized information. To address the needs of the one in eight women in America who face challenges with fertility, Celmatix applies big data and genomics to reveal what a woman's personal metrics and DNA may reveal about her fertility. The firm takes a new and unique look at fertility, leveraging data collected from Electronic Medical Records (EMRs) as well as genetic data collected from partners through their Personalized Reproductive Medicine (PReM) Initiative.

Celmatix is empowering and transforming the way women and their physicians make informed and proactive reproductive health decisions by leveraging real-time genomics data and applying technologies such as big data analytics, machine learning, A/I and whole-genome DNA sequencing. Celmatix's research is helping clinics optimize patient outcomes and improve the patient experience.

"No one in our space is doing anything similar to this – no one else is applying internet-scale technologies," said Andrew FigPope, Lead Data Architect at Celmatix. "For us to scale the way we want to and to future-proof our business, I believe Apache Kafka and Confluent are invaluable investments."

## CHALLENGE

In order to leverage big data analytics and whole genome DNA sequencing to enable the innovation of future women's health solutions, Celmatix needed to address a variety of data-related challenges:

1. EMR data is notoriously complicated, messy and highly heterogeneous, resulting in one clinic often having a different way of collecting data than another. Transforming and organizing data into standard data elements from different EMR systems is a major challenge.
2. Celmatix found themselves asking how to best connect to each of their partnering clinics in real-time to provide the best experience possible for both the physician and the patient.

When dealing with sensitive patient information, security is always top of mind, so being able to pull and anonymize data while adhering to necessary security measures is vital.

Confluent simplifies our processes and reduces complexity, allowing us to be more effective.

— Andrew FigPope, Lead Data Architect at Celmatix

## SOLUTION

The need for improved data extraction and demand for data in real-time led Celmatix to adopt a streaming platform. After establishing their vision and need for providing real-time insights to physicians, they implemented Confluent Platform to leverage the stability and scalability of Apache Kafka, and to take advantage of Confluent Platform's added features such as Schema Registry and the Kafka Connect API, as well as Confluent's support team.

"Confluent is directly impacting our data engineering teams in how we collect, manage and analyze data, but that's only the beginning. The data itself is cascading outwards to other crucial elements of our organization, such as our in-house research teams," said FigPope. Celmatix's internal teams now have access to consistent data resources allowing them to develop other inventions to improve reproductive health. "Those teams formerly were presented with a mass

of spreadsheets with inconsistent, de-identified parsed data. If you think about Confluent Platform as creating ripples from a pond, it is primarily impacting our data engineering and DevOps teams, but the rippling outward effects ultimately positively impact physicians and patients who use the derivative information to inform their treatment decisions.”

Celmatix has a team dedicated to developing a formal data model that handles electronic medical record data. To support that team, Celmatix is utilizing Schema Registry to catalogue the data and provide context to particular systems. Schema Registry is helping Celmatix normalize the heterogeneous data types received from the various clinics and databases into common formats.

Celmatix’s internal teams now have access to consistent data resources allowing them to develop other inventions to improve reproductive health.

They also use the Kafka Connect API to pull a subset of that data from EMRs to then anonymize, clean, standardize and transform the data into measurable insights for physicians. That process delivers data to a relational database containing those standardized versions of the data along with the ability to monitor the streams as they flow for analytics.

Celmatix uses Confluent to stream the data from the clinics, so they can run it through their proprietary models to produce relevant reports.

To note, Celmatix considered Hadoop as an alternative solution, but FigPope said it would have required them to build a framework to manage the transformers in a converged manner. Instead, Celmatix can use the framework Confluent had already built. “Kafka is simply doing what we were going to have to build anyway, but in a better way and in a way the Celmatix team respects,” said FigPope.

## RESULTS

With Kafka and Confluent Platform, Celmatix is able to conduct their business in a way that wouldn't be possible otherwise. "Confluent simplifies our processes and reduces complexity, allowing us to be more effective," said FigPope.

Results include:

- **Achieved market differentiation with the ability to scale and expand.** "Confluent allows us to manage our data and products systematically, in an expandable and reasonable way. We start with our organized streaming platform, so adding a new clinic is not a big job," said FigPope. "There is no other way we could do what we do and scale how we need to without Kafka. Our market differentiation is fabulous—no one else in our industry is doing this and we are maintaining a huge lead."
- **Harmonized and managed disparate data.** "Confluent Platform helps us manage the heterogeneity of EMR data by allowing us to build modular stream builders so we don't have to write new code for every clinic we add," said FigPope. "When we write unique code, Confluent Platform allows it to be standardized, harmonized and managed in a separate place from the tooling infrastructure."
- **Enabled access to vast and disparate data in real time.** "Physicians need rapid and accurate insights when considering fertility treatments for their patients to ensure timely care. When the models used for each patient are needed promptly, we use Confluent to stream the data from the clinics on-demand," said FigPope. "It's as easy as running the request through the model and yielding back a report."

For us to scale the way we want to and to future-proof our business, I believe Apache Kafka and Confluent are invaluable investments.

—

Andrew FigPope, Lead Data Architect at Celmatix