LIFE SCIENCES—
Breakthrough Data Access for a Breakthrough Industry

Molecula gives life sciences the world-changing power of data access

MARKET TRENDS: Dedicated to the protection and improvement of human, animal, and natural life, pioneers in life sciences are tirelessly effecting breakthroughs on behalf of humankind. As the industry has exploded with new technology, the double-edged sword of big data has come to the forefront. From the biotech to pharmaceutical, biomedical, life systems, and medical device industries, big data is of paramount importance to gaining a greater understanding of our world and commercializing new products that will impact lives.

The very data that makes machine learning, gene editing, synthesis, sequencing, and automation possible has become so large and unwieldy that it has also become one of the biggest obstacles to overcome.

In life sciences today, data scientists and analysts are playing increasingly important roles as they provide the analytics necessary to make time-sensitive decisions in a competitive, dynamic environment. They need access to massive amounts of high velocity data in order to succeed. It’s not trivial to get vast amounts of complex data into a single place where it is interlinked and prepared for machine processing. Instead of spending their time exploring data and creating new analyses, too many data scientists and analysts are working overtime just to get access to the data.

Some of the greatest challenges facing life sciences initiatives today include:

- RISING GLOBAL COMPETITION
- SILOED AND BROKEN LEGACY SYSTEMS
- LAG IN R&D
- DNA SEQUENCING & TESTING
- ENSURING COMPLIANCE
- INCREASING ACCESS TO AND EFFECTIVENESS OF CLINICAL TRIALS
- COMMERCIALIZATION OF PRODUCTS

DATA AT THE SPEED OF THOUGHT™.
Molecula for Life Sciences—

The sheer amount of data, strict compliance requirements, and ever-growing collection of disparate data silos can be overwhelming. Molecula’s big data solution is different in that it doesn’t apply yet another bandaid to the problem but gets to the root of the Data Death Spiral—the cumulative result of uncoordinated data projects that are too narrowly focused and cause further fragmentation as data gets copied and moved from system to system, rendering the data effectively useless.

Molecula is an enterprise feature store that simplifies, accelerates, and controls big data access to power machine-scale analytics and AI so organizations and companies can achieve their business objectives. With Molecula, life sciences companies are able to securely access and perform computations on any and all data at unprecedented speeds, with incredibly low latency, and with a fraction of the hardware.

Molecula Empowers Companies in the Life Sciences Industry to:

- **PREDICT** outcomes in real-time by leveraging 100% of distributed data
- **DELIGHT** users and drive adoption with zero latency
- **ENABLE** continuous analysis and machine learning with millisecond performance
- **MINIMIZE** security and privacy risks with anonymized-ready data
- **REDUCE** time and cost to get value from AI and machine learning
Use Cases—

Below are just some of the ways the life sciences industry is leveraging data access. What will you do with continuous, real-time analysis of your most important data?

**SHORTEN DRUG APPROVAL TIMELINES SUCH AS THE REAL-TIME ONCOLOGY REVIEW (RTOR) PILOT**

RTOR aims to explore a more efficient review process to ensure that safe and effective treatments are available to patients as early as possible, while maintaining and improving review quality and balancing the review team’s workload through data and analysis standardization, and early iterative engagement.

**DEVELOP THERAPEUTIC INTERVENTIONS**

Pharmacological research to discover new treatments and medicines advances at a much faster pace when bolstered by large, real-time datasets.

**IDENTIFY GENETIC MUTATIONS**

One cancer patient can produce a terabyte of digital data. With the ability to analyze all data across millions of cancer patients, mutations and patterns become visible for the first time.

**PREVENT EPIDEMICS**

Big data access at scale empowers scientists to combine demographic, geographic, and socio-economic data along with disease trends and health strategies to prevent unnecessary illness in populations.

**PREDICT VIRUS EVOLUTION**

Testing against massive databases, DNA sequencing can help in the discovery of new viruses. With sufficient access to the right data, scientists can determine the sequence of the nucleic acid to aid in virus identification that was previously impossible.

**INCREASE STATISTICAL POWER IN CLINICAL TRIALS**

Access to data supplements trials with additional datasets such as next generation sequencing resulting in more balanced results. This minimizes confounders such as unforeseen biases. It also enables researchers to focus on specific subgroups and better randomize samples.

**IMPROVE DRUG TRIAL EXECUTION AND SUCCESS**

The ability to bring in new data and update data on an ongoing basis and instantly analyze it as part of a trial directly impacts the success of a trial. Analyzing unlimited patient attributes and being able to slice and dice data across all trials, past and present, provides for more rigorous evaluation of side effects during trials.
Why Life Sciences Companies Love Molecula—

- Continuous ad-hoc analytics across unlimited study results, genetic, IoT and other relevant data
- Millisecond data updates and access allows instant feedback to make real-time decisions
- High concurrency analytics across massive sets of users provides a flexible data system that allows for acquisition and growth
- 50-90% lower TCO from a reduced data footprint and infrastructure improvements
- Secure data format to share data across ecosystem and reduce compliance risk

How Molecula Will Help You—

Molecula’s feature store is an overlay to conventional big data systems that centralizes an organization’s most important entities and their attributes across various source systems, while maintaining up-to-the-second updates. The Molecula feature store continuously extracts and updates features to provide data engineers, data scientists, and application developers a single access point to graduate from reporting and explaining with human-scale data to predicting and prescribing real-time business outcomes with all data.

FEATURE STORE HIGHLIGHTS:

- **Extracts Features from Data**
  ‘AI Ready’ feature store that continuously extracts and updates features in real time

- **Reduces Footprint**
  Lossless reduction in data footprint, up to 85%, without copying or moving data

- **Offers Single Point of Access**
  Centralized, ultra low latency access to all of your data

- **Supports ML Workloads**
  High concurrency queries for machine-scale analytics and ML

- **Eliminates Pre-Processing**
  Performant Joins at query time, with no pre-aggregation or pre-processing

- **Enables Time-Oriented Filtering**
  Track and filter time at a feature level

- **Delivers Secure, Cell-Level Control**
  Granular access control of feature sharing down to the cell level

- **Easily-Adoptable Overlay Implementation**
  Extension framework enabling seamless integration into existing environments

WHAT ARE FEATURES?

Feature extraction is the process of reducing the dimensionality of data to more efficiently represent its values. A feature, or column, represents a measurable piece of data that can be used for analysis: Name, Age, Sex, Fare, and so on. Features are also sometimes referred to as “variables” or “attributes.” This technique was pioneered by data scientists who needed to prepare data for demanding machine learning and AI workloads.
Customer Highlight—

Elligo, a leading clinical research platform has selected Molecula to enable their accelerated development of new pharmaceutical, biotechnology, medical device, and diagnostic products. As the ultimate healthcare-enabling research organization, Elligo maintains the integrity of the trusted patient and healthcare physician relationship, building local healthcare communities, and leveraging electronic health records to bring clinical research as a healthcare option to more patients.

“As the leader in bringing healthcare to clinical research, Elligo is leveraging the power of data to directly improve the execution of clinical trials. True data readiness creates the ability to instantly provide synthetic data, novel data access and streamlined analytics. Molecula accelerates our ability to provide those capabilities and insights to our customers, and this partnership begins to fundamentally change the way clinical trials are executed.”

- Michael Ibara; Vice President, Data Science

Get Started: Unlock Your Data Opportunity—

To effectively compete in today’s economy, life sciences innovators must leverage their information assets and connect with additional outside data to comprehensively understand and take action on their most important opportunities. Potentially invaluable data sources include: genetic, biometric, clinical trial, IoT, regulation, supplier, competitor, and worldwide population data.

Contact us today to learn more about how Molecula’s feature store will transform your business. Ask us how you can unlock real-time value and actualize use cases never possible before.