tetrascience | The Tetra Scientific Data and AI Cloud:

A science-led approach

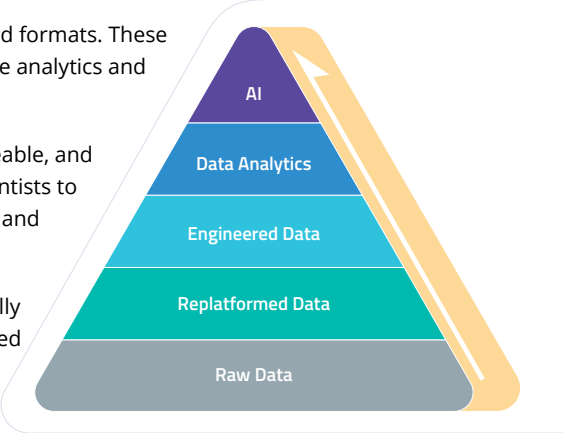
FACT SHEET

With an average of twelve years and $2.3 billion to develop a new therapeutic, life sciences organizations are seeking a new paradigm for more efficient and cost effective approaches. The rise in computational power, democratization of cloud computing, and the emergence of advanced neural networks has led the industry to expect AI and ML to radically accelerate and improve scientific outcomes, reducing development time and costs while increasing competitiveness.

However, AI/ML algorithms need large, well-engineered datasets, which are challenging to produce due to the complexity and variety of scientific data. Typically, these datasets are siloed, static, locked in proprietary formats, and lack the structure and scale required for AI applications, rendering them unsuitable for scientific AI.

The Scientific AI journey

Raw data as generated by scientists and lab instruments is often unusable for scientific AI and therefore its value is very limited. This data must be transformed to drive AI outcomes and generate significant insights for truly transformational breakthroughs. There is an immutable order of operations in the journey of scientific data from raw to AI-enabled.



1. **Raw data** typically resides in vendor-specific silos, existing in proprietary and unstructured formats. These datasets are too small and static for AI and ML, offering little value and utility for predictive analytics and AI, and are costly to maintain.
2. **Data replatforming** is the necessary first step towards making datasets accessible, reuseable, and compliant. Collecting, centralizing, and contextualizing siloed data in the cloud allows scientists to leverage modern cloud infrastructure natively, providing elasticity for storing, processing, and easily retrieving their data.
3. Sophisticated **data engineering** requires deep scientific and technological expertise to fully convert data into liquid, large-scale, harmonized, and compliant datasets with industrialized taxonomies and ontologies. Only these open, vendor-agnostic, and AI-enabled datasets can become the atomic building blocks for scientific AI.
4. Optimizing scientific data for dashboards, visualization tools, and **data analytics** applications requires precisely structured data models. Fragmented models result in significantly suboptimal outcomes.
5. **AI** can address critical scientific questions, but organizations often lack the resources or skills for effective data strategies. Purpose-engineered, liquid data is necessary for groundbreaking AI-driven outcomes.

The solution

The Tetra Scientific Data and AI Cloud™ is a modern cloud data stack that has been purpose built to elevate the value of scientific data throughout the data journey, bridging the gap between raw scientific data and transformational AI outcomes. As the first and only industry-specific data and AI cloud tailored for science, it enables the rapid assembly of scientific data to generate large-scale, liquid, compliant, and AI-ready datasets for biopharma and other science-based organizations. Data is replatformed and engineered, maintaining integrity and traceability. Built on a scientific lakehouse architecture, it natively supports other cloud data platforms and provides a foundation for advanced data analytics and processing solutions. It facilitates multi-modal data consumption from data discovery and analytics to high-performance computing and AI model training.

The Tetra Scientific Data and AI Cloud automatically collects and aggregates scientific data at every stage of the design-make-test-analyze (DMTA) cycle, linking experiment design, measurements, methods, compound creation, and analyzed data for advanced analytics and AI. It integrates with diverse data sources and liberates data from vendor-specific silos by transforming data into an open, vendor-agnostic

JSON format and providing programmatic REST API and SQL query interfaces. It ensures data is FAIR (findable, accessible, interoperable, reusable), supports compliance standards, and uses robust scientific taxonomies and ontologies to contextualize data. Powering the "Scientific AI Factory", it accelerates the generation of AI-based scientific results at scale. Organizations can rapidly prototype AI models and a Data and AI Workspace allows scientists to access their data through software of their choice.

Scientific use cases

To maximize value for customers, scientific data and workflow solutions must prioritize scientific use cases over individual instrument or application integrations and data schemas. By centering on end-to-end scientific workflows, customers achieve tangible and impactful outcomes through streamlined processes and purpose-built data for Scientific AI.

TetraScience defines a "scientific use case" as a comprehensive experimental workflow that delivers specific scientific outcomes (e.g., water content, impurities, etc). This approach enhances lab productivity through data automation, provides new insights through visualizations and analytics, and creates the foundational data needed for AI/ML.

TetraScience has built the largest library of highly repeatable, deployment-ready **scientific use cases** that is continually expanding. In developing these use cases, TetraScience prioritizes the user experience for scientists and data scientists, considering the typical instruments, applications, and data interactions within specific workflows. Data pipelines are designed to purposefully contextualize data with relevant metadata, including rich taxonomies and robust ontologies, to ensure scientific accuracy.This science-led approach empowers customers to extract valuable insights from their extensive scientific data, enabling faster, more informed decision-making.

Software is not enough

To achieve widespread adoption and a high return on investment (ROI) for scientific data and workflow solutions, it's crucial to align IT, scientific, and data science teams. However, organizations often struggle with limited resources, misalignment between departments, a lack of industry-best practices, and experience with the new data solution. This can result in a "project" mindset focused on short-term gains, ultimately failing to produce AI-ready data that accelerates scientific outcomes.

The Tetra Catalysts helps customers **increase the value of their Tetra Scientific Data and AI Cloud deployment by 5-10x.** The Tetra Catalysts offering is delivered by our elite Tetra Sciborgs, composed of Scientific Data Architects (SDA) and Scientific Business Analysts (SBA) with deep expertise in science, data, and technology, and the TetraScience product. Within an engagement, they work as the "connective tissue" between scientific IT and scientists, focusing on tangible business outcomes. Embedded in your organization, the Sciborgs uncover high impact scientific use cases to maximize ROI and demonstrate how to leverage scientific AI in daily work, pushing the boundaries of scientific outcomes. They develop an effective metadata strategy for enhancing data utilization, help optimize user adoption with support for workflow changes, and enable users to become more data conscious.

Unparalleled scientific impact

Customers gain access to a library of highly repeatable, deployment-ready use cases that empower scientists to automate workflows and create curated datasets for advanced analytics and AI/ML. Additionally, customers benefit from our knowledge base, which contains learnings from our extensive experience with technology vendors, including instrumentation and lab informatics software, as well as insight into best practices from customers across the biopharma industry.

With the skill of our Sciborgs, Tetra Catalysts deliver unprecedented improvements to scientific workflows, accelerating time to market. Example deliverables include discovery and design, solution design documentation with current state and future state diagrams, defined use case requirements and success metrics, visualization and dashboard templates, monthly reports, and more.

Customer value

Charles River Laboratories was able to reduce ADME/Tox wet lab experiments by 50%, cutting sampling points by 50% while enhancing IC_{50} accuracy, decreasing manual data curation by 1.5 FTE, and developing an *in silico* model in only 5 months. This is just one example of why TetraScience is a trusted partner of 12 of the top 25 biopharmas.

Corporate Headquarters | 177 Huntington Avenue, Suite 1703, Boston, MA 02115

Atlanta | 8000 Avalon Boulevard Suite 247, Alpharetta, GA 30009

tetrascience

tetrascience.com