



Enabling the Breakthroughs of Tomorrow: The Data Lakehouse for Clinical Trials



Nathan Johnson

VP, Digital Innovation
eClinical Solutions

njohnson@eclinicalsol.com

Meet the Speaker

Nathan Johnson

Vice President, Digital Innovation
eClinical Solutions

- 20 years experience in clinical research as a SAS programmer and innovator.
- expertise in statistical analysis and reporting, standards development, and data management.
- passionate about reshaping clinical trials through digital transformation, intelligent technology, and increased automation.
- degree in biostatistics from Case Western Reserve University.

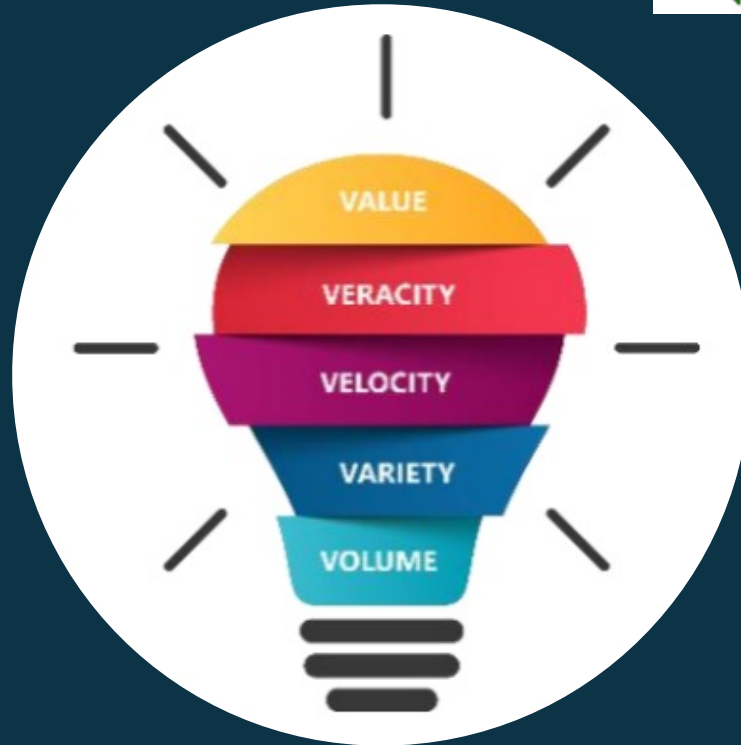


Clinical Data Problem



Clinical Data Problem

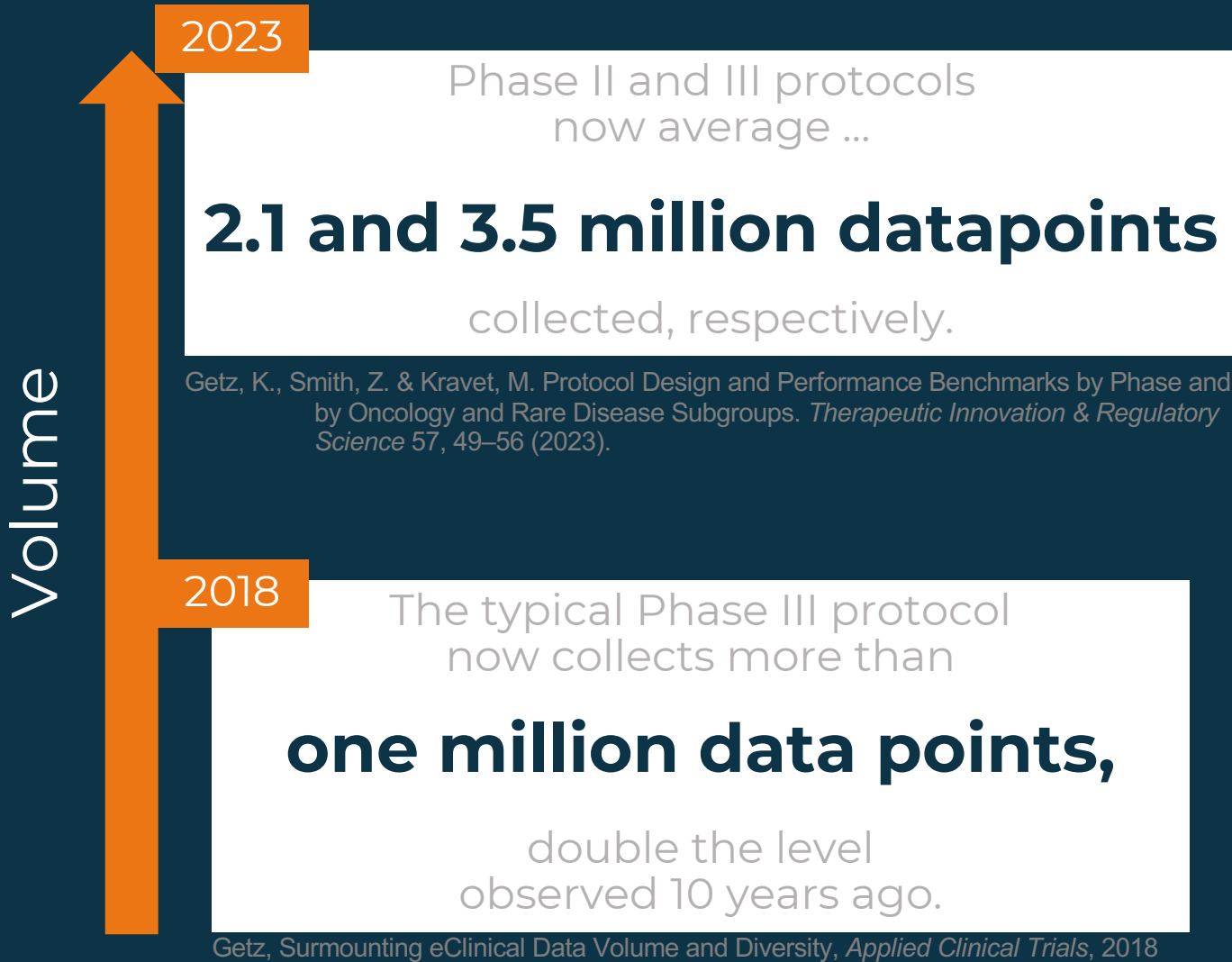
The 5Vs of Clinical Data



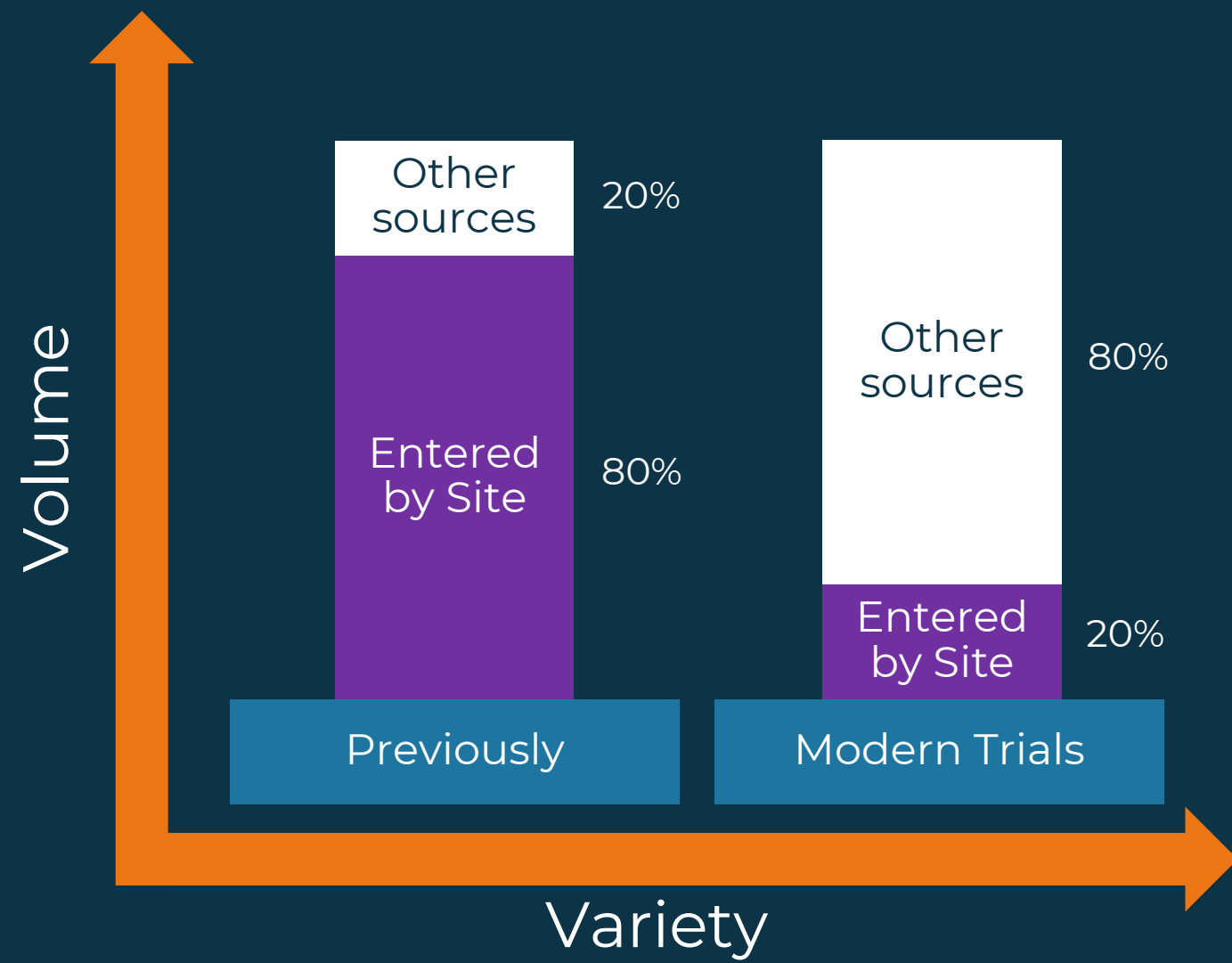
Reference: <https://scdm.org/the-5vs-of-clinical-data/>



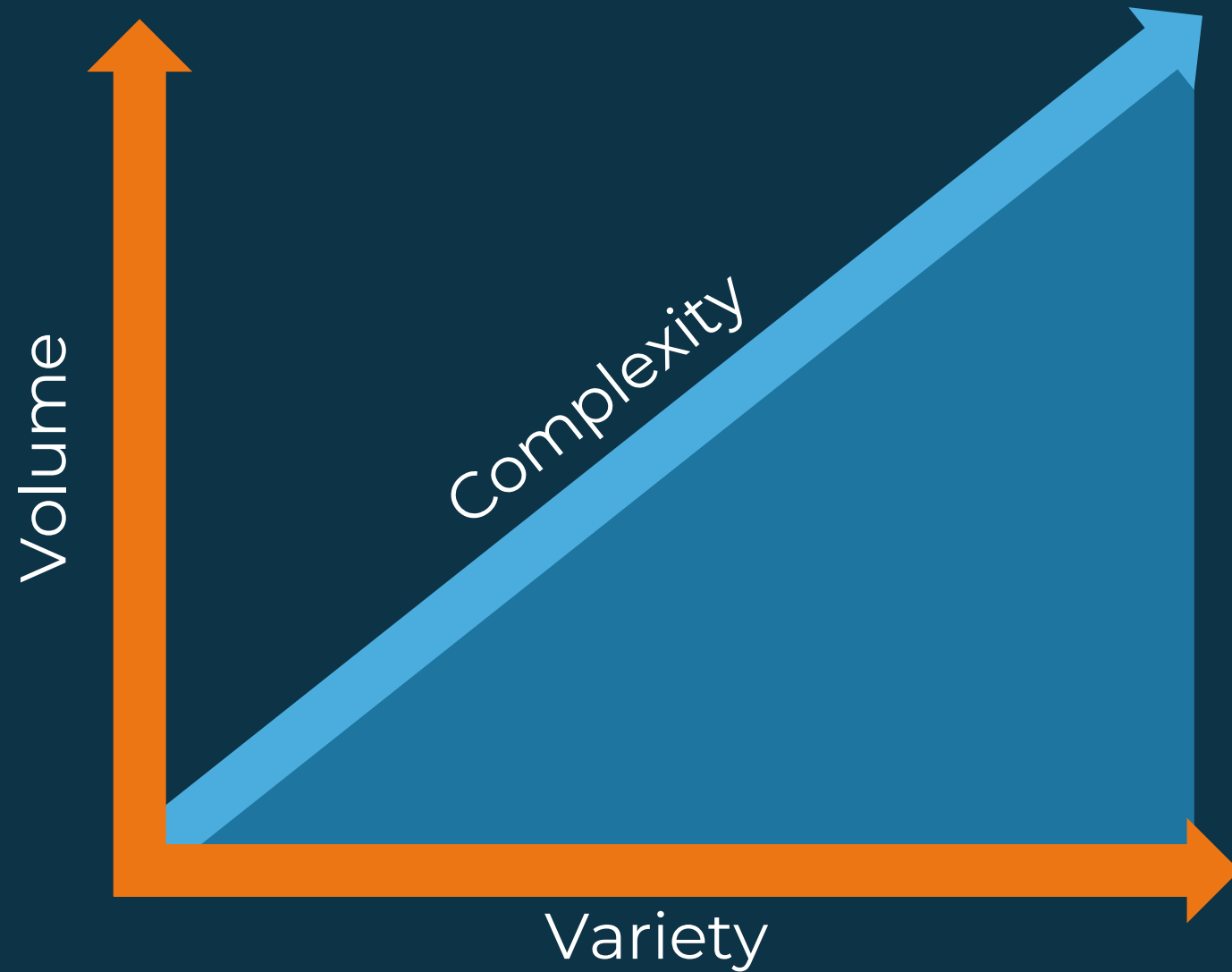
Clinical Data Problem



Clinical Data Problem



Clinical Data Problem



The Reality of Modern Trial Data: Changes in Perspective

EDC



Data Manager

Image Credit: <https://monstertruck.fandom.com/>



The Reality of Modern Trial Data: Changes in Perspective

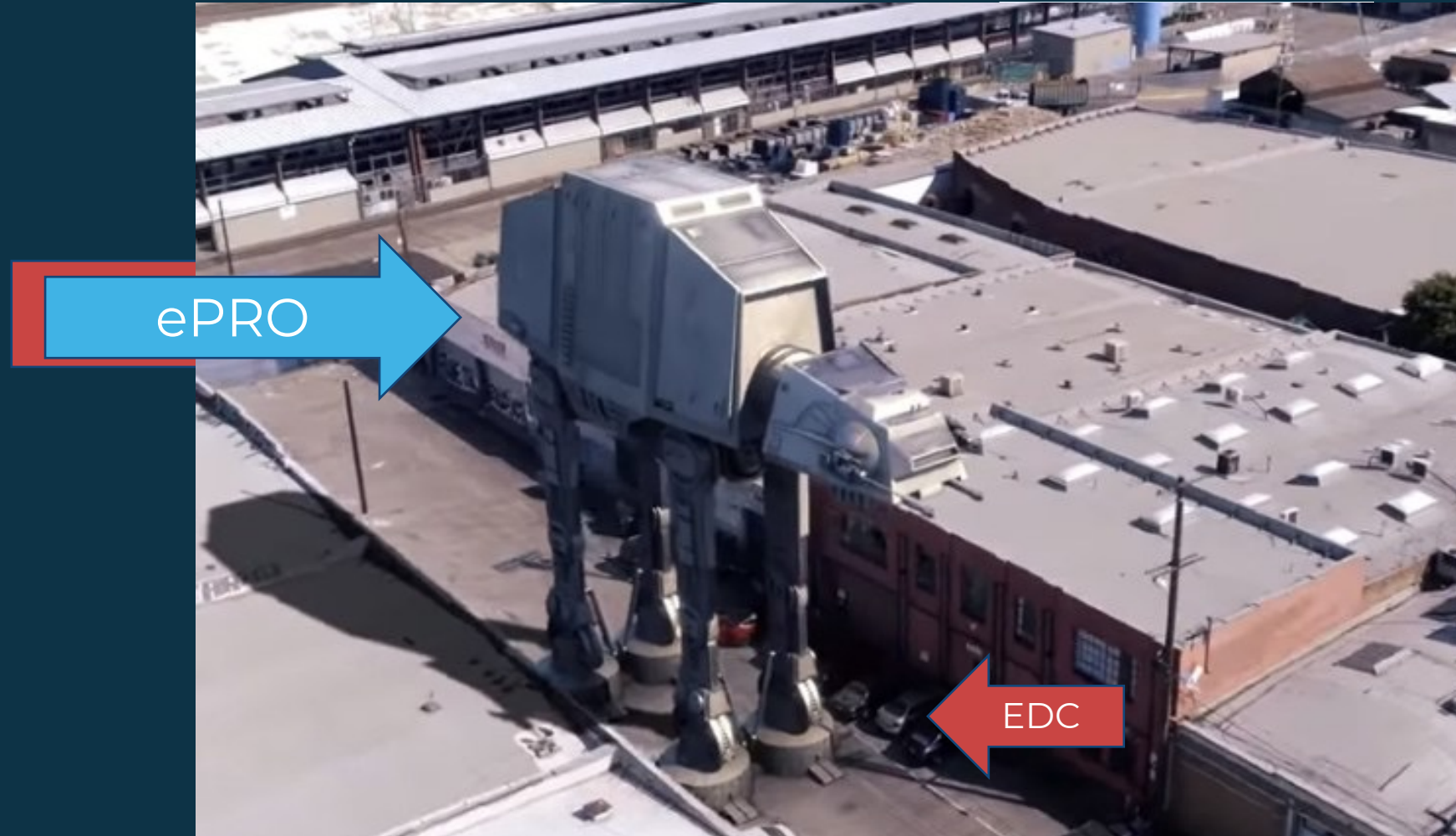


Image Credit: Corridor Crew corridordigital.com



The Reality of Modern Trial Data: Changes in Perspective



Image Credit: Corridor Crew corridordigital.com



Challenges in Handling Data

Data Security and Privacy



Timeliness and Real-time Monitoring



Data Collection and Management



Data Quality Control



Data Standardization and Interoperability



Regulatory Compliance



Data Integration and Analysis



Data Sharing and Collaboration



Modern Data Architectures



Data Warehouses



Integration from disparate sources

Standardization into consistent format

Data mining and statistical analysis

Business intelligence and decision-making

Data Warehouse Challenges



Complexity

Cost

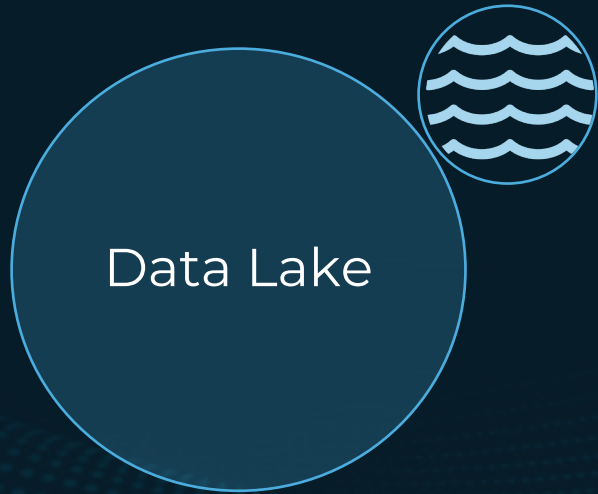
Data Latency

Scalability

Data Silos

Data Quality

Data Lakes



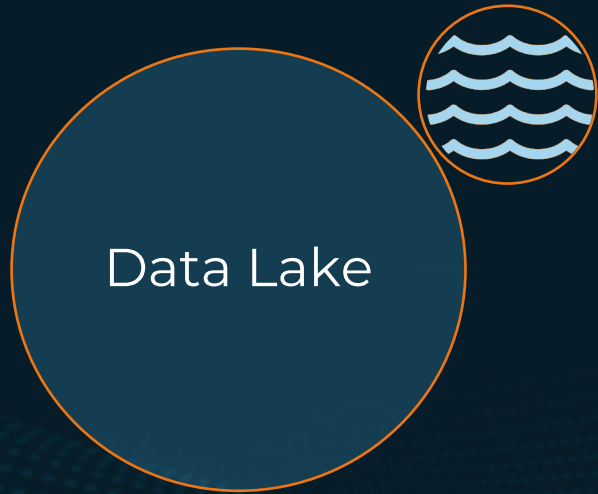
Data ingestion from various sources

Directly-accessible storage in native format

Data processing

Advanced analytics

Data Lake Challenges



Data Quality and Consistency

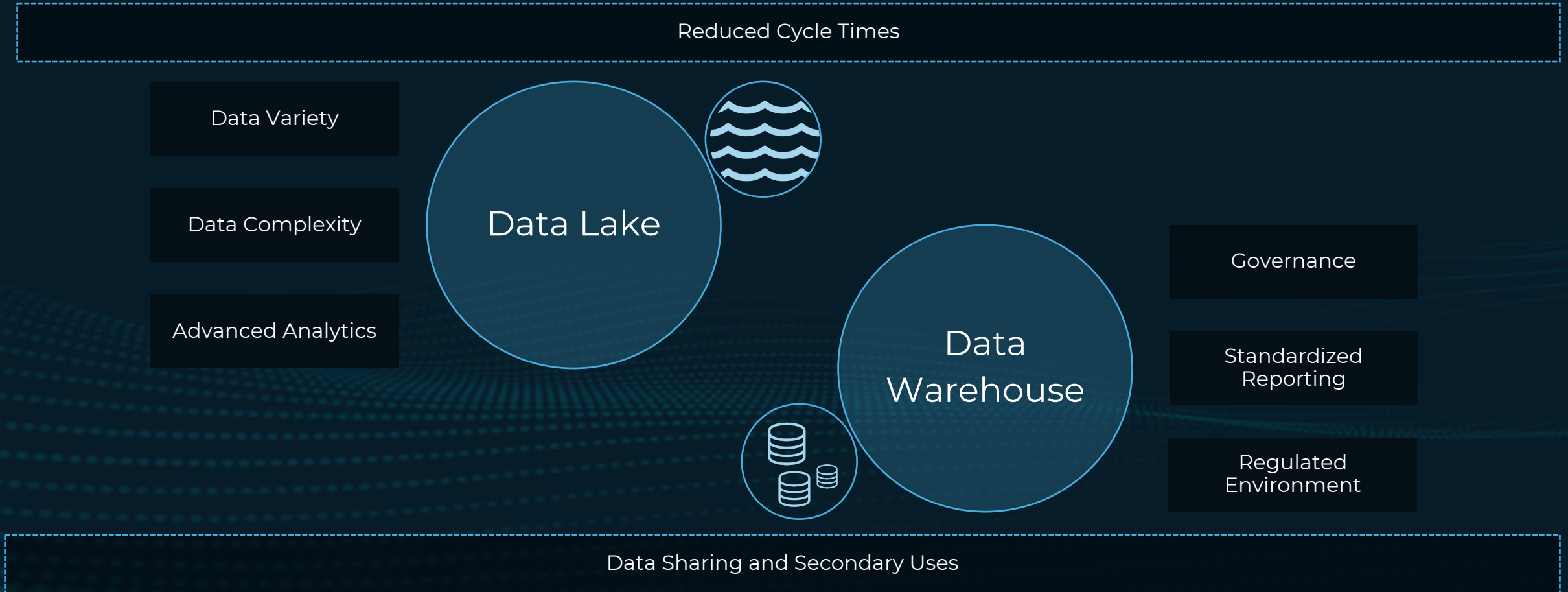
Data Governance and Security

Data Discovery and Accessibility

Processing and Analysis Complexity

Cost and Scalability

Architectural Pressure



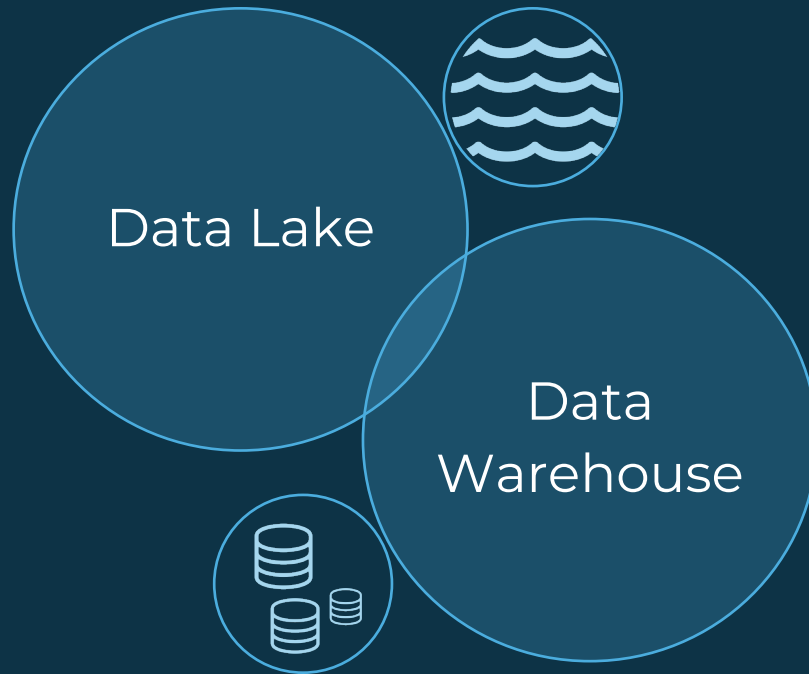
The Digital Blueprint



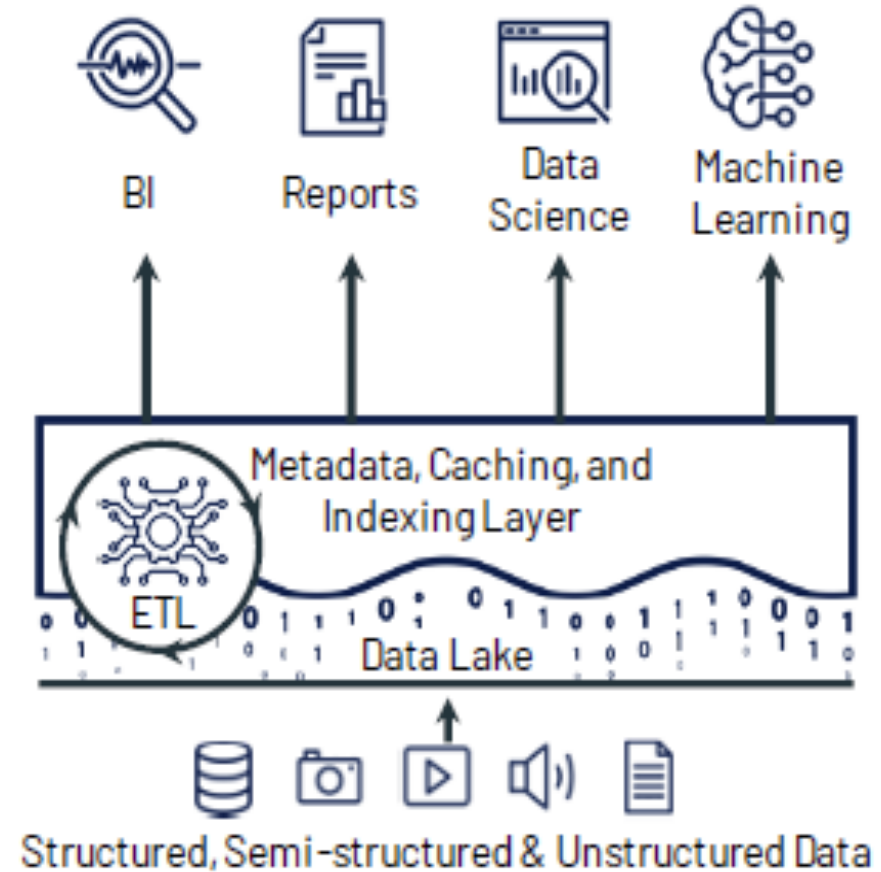
Lakehouse for Clinical Trials



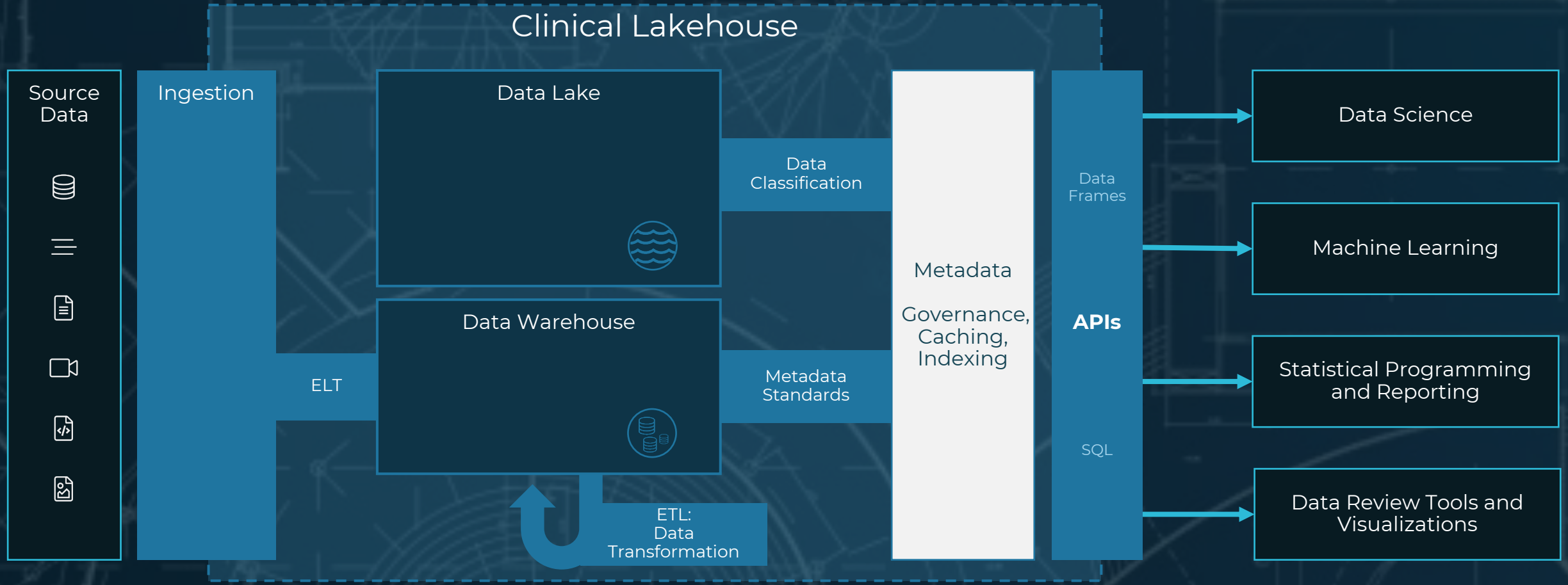
Lakehouse Architecture Explained



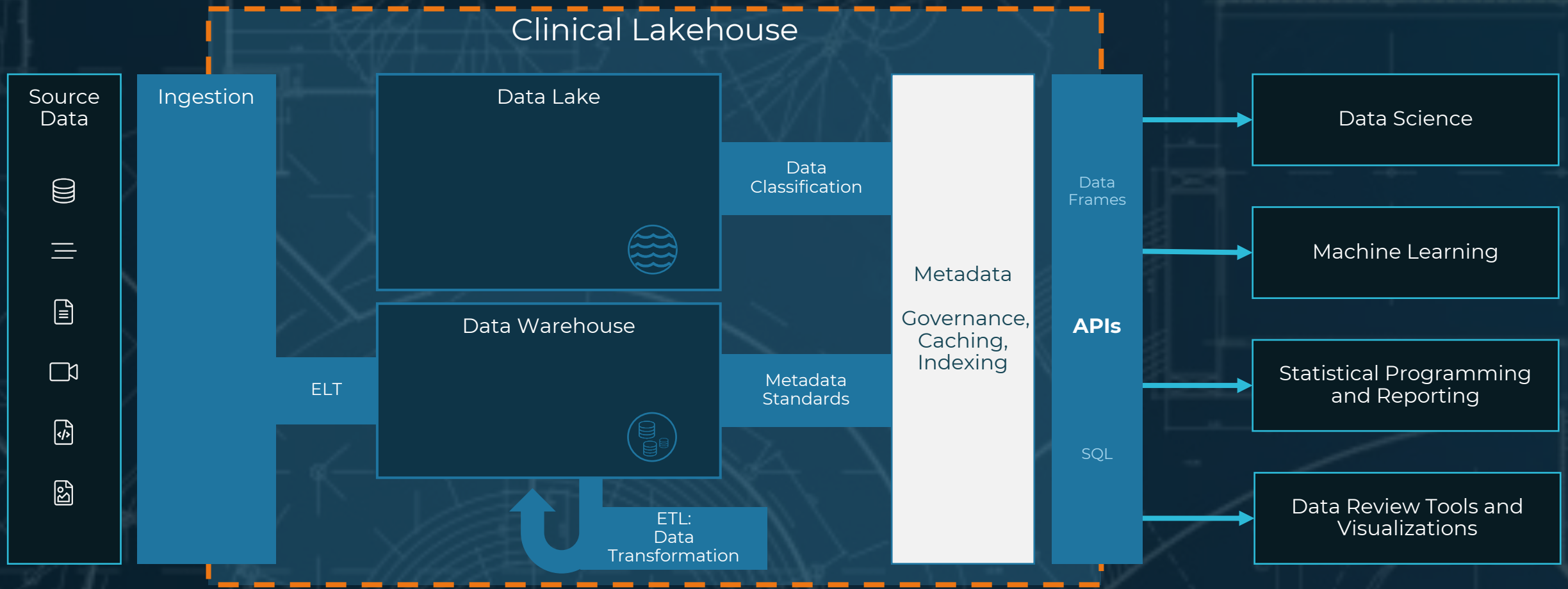
Michael Armbrust, Ali Ghodsi, Reynold Xin, Matei Zaharia
11th Annual Conference on Innovative Data Systems Research
(CIDR '21), January 11–15, 2021, Online.



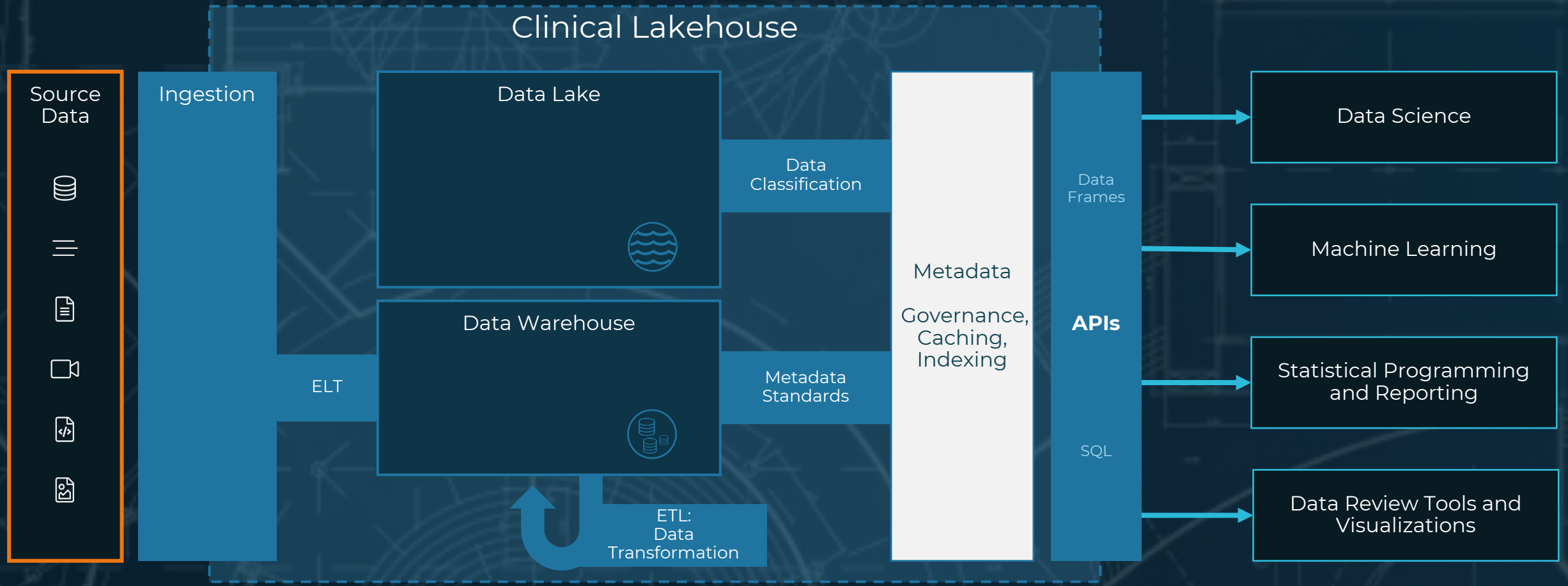
Case Study Lakehouse Adoption



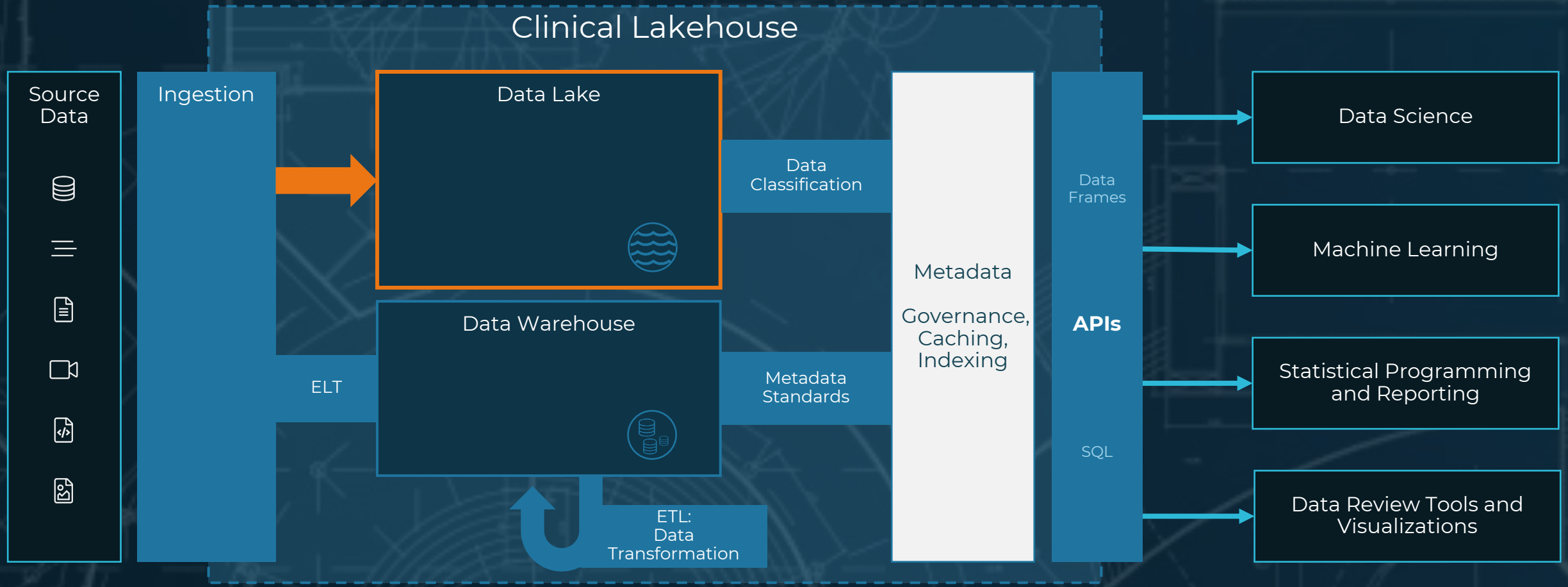
Case Study Lakehouse Adoption



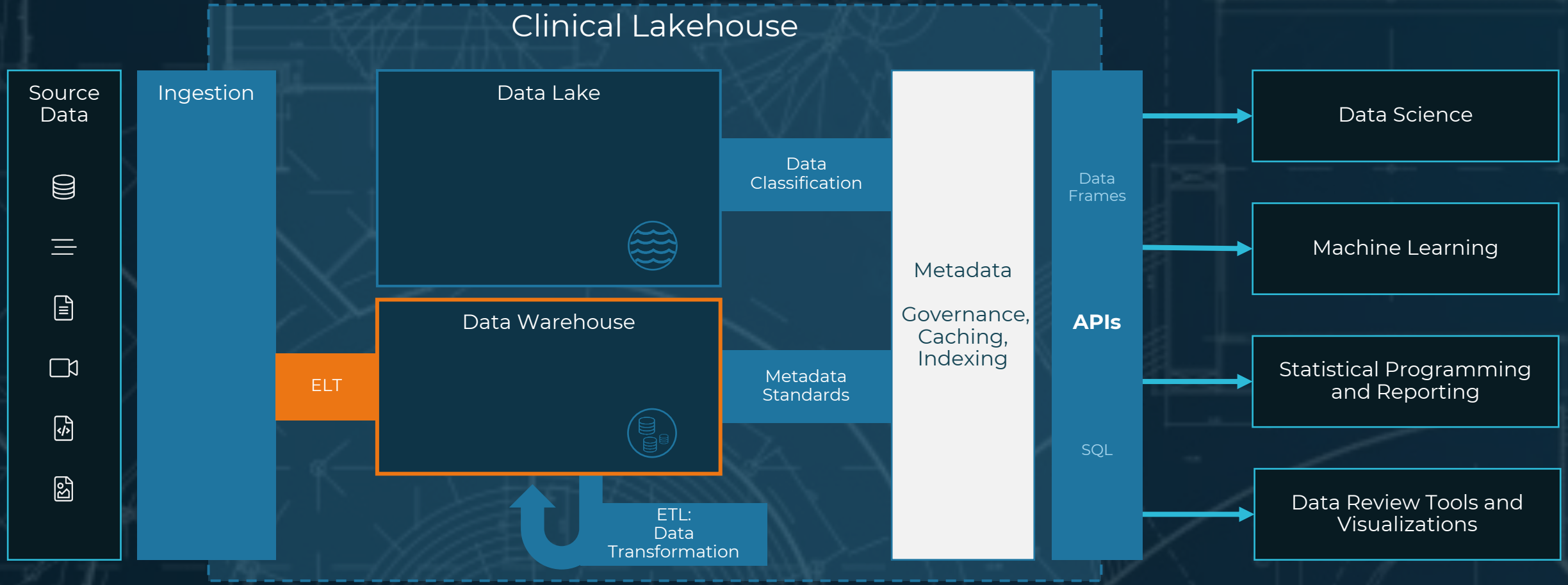
Case Study Lakehouse Adoption



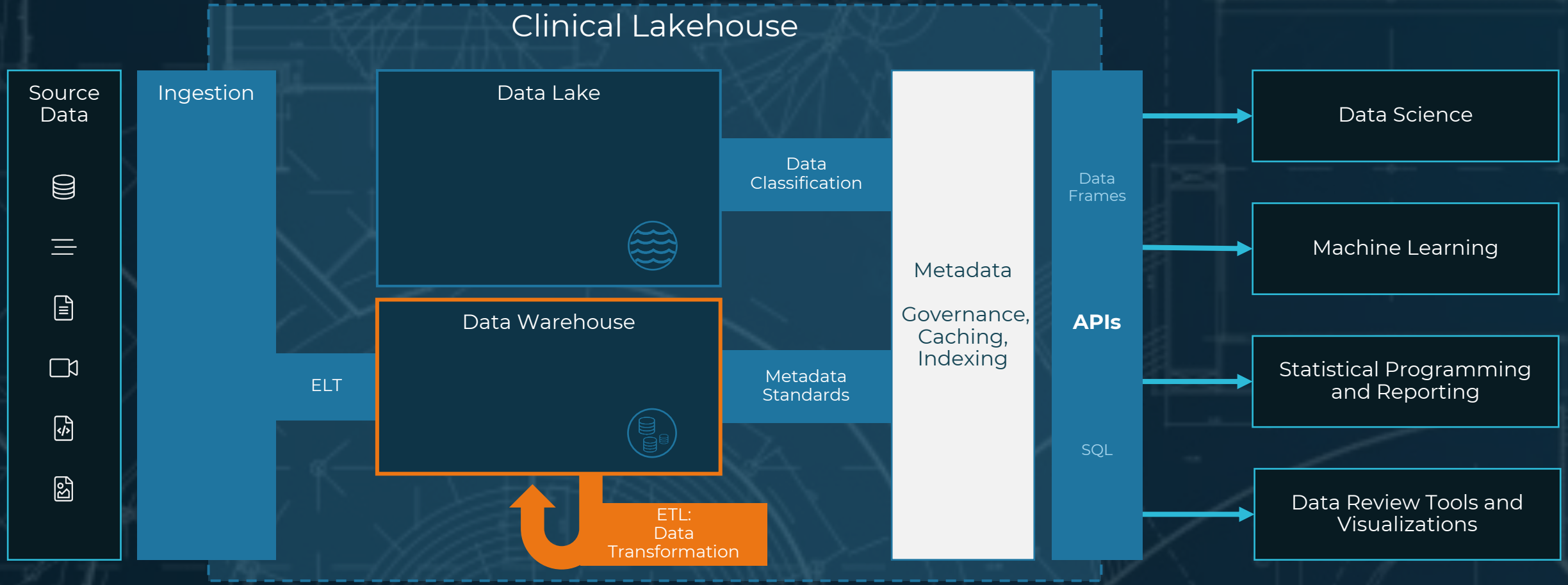
Case Study Lakehouse Adoption



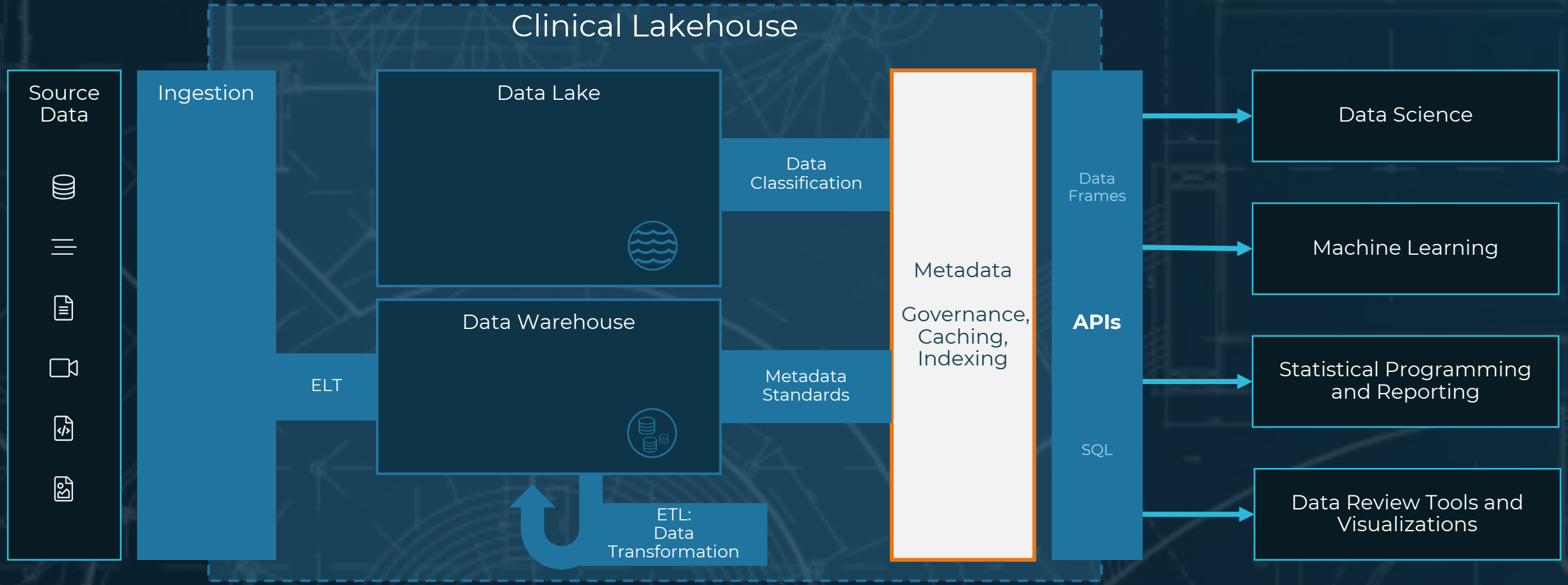
Case Study Lakehouse Adoption



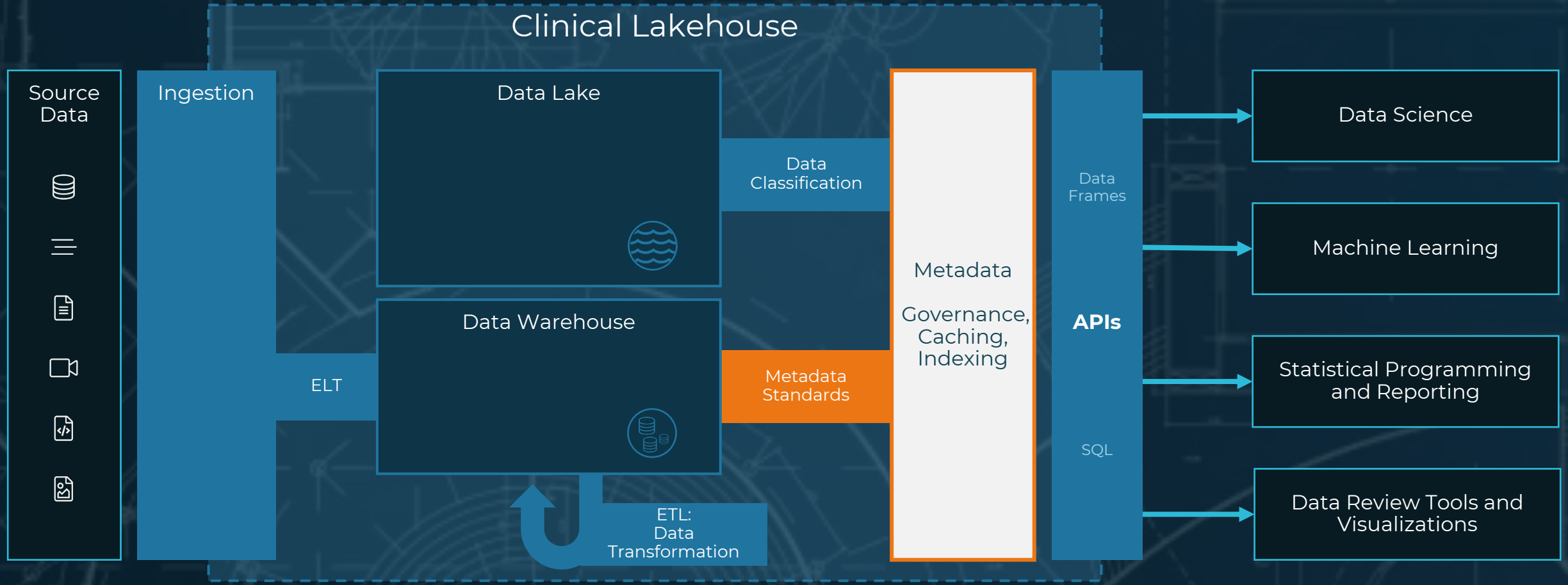
Case Study Lakehouse Adoption



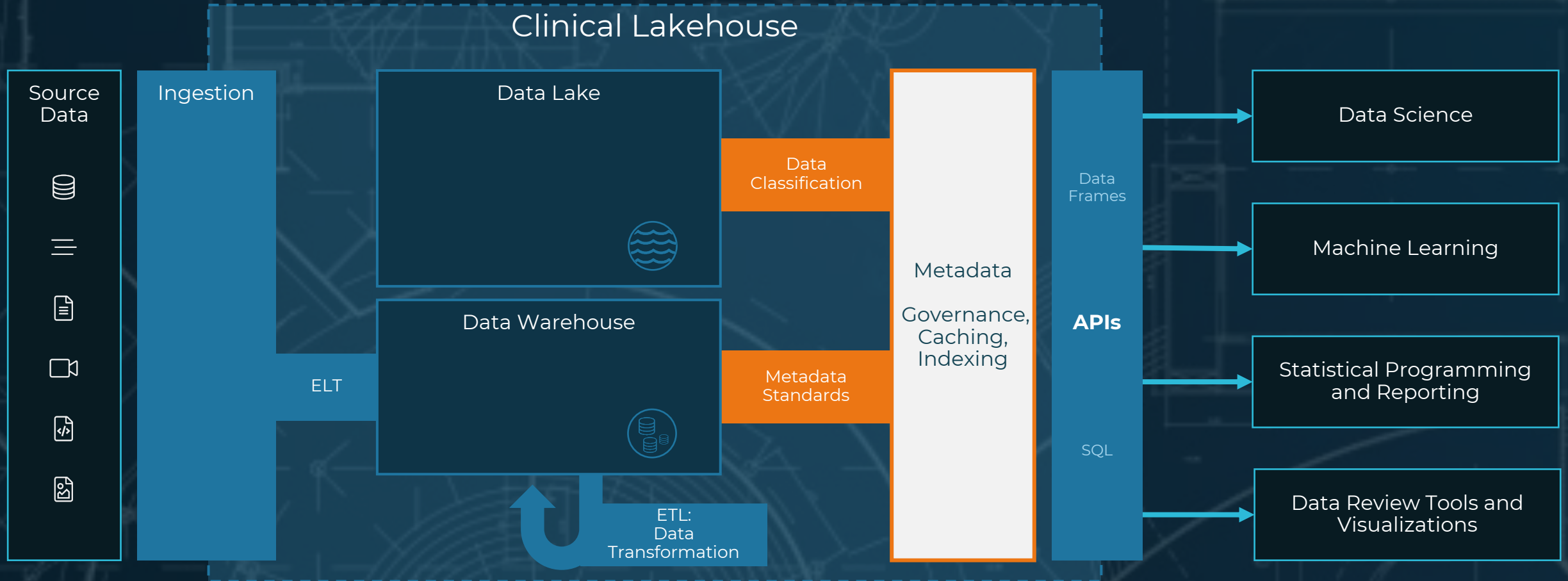
Case Study Lakehouse Adoption



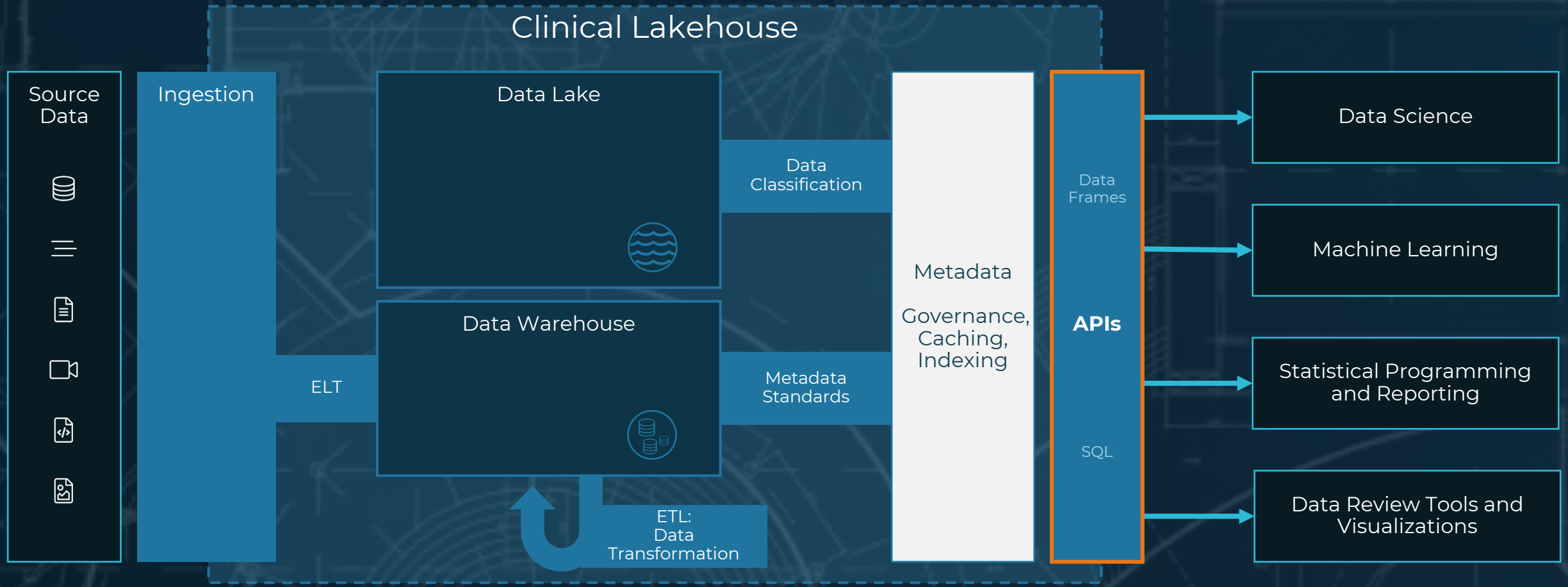
Case Study Lakehouse Adoption



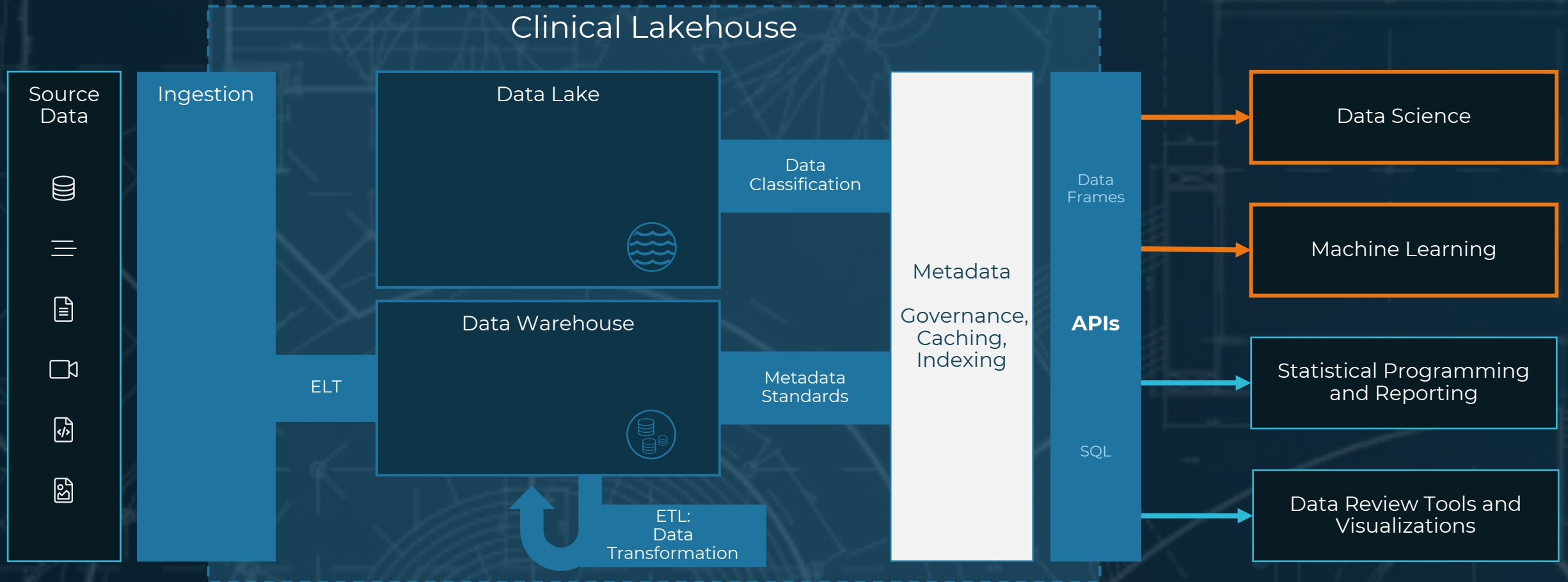
Case Study Lakehouse Adoption



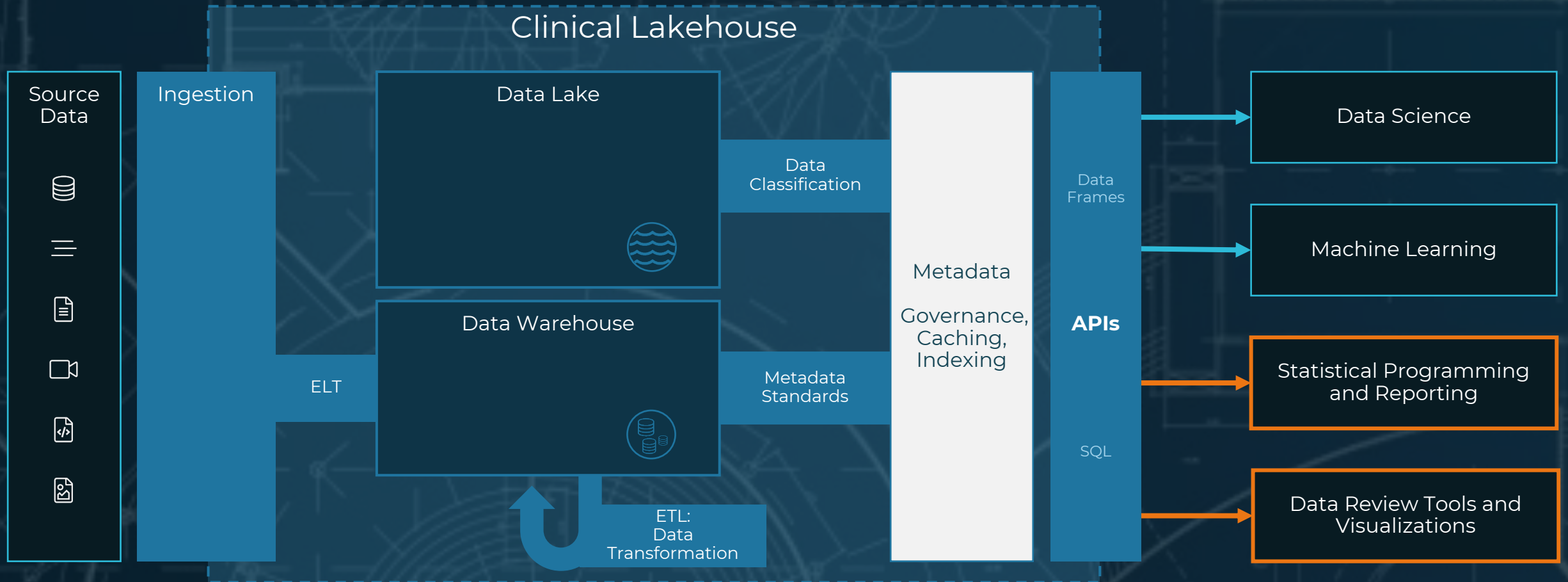
Case Study Lakehouse Adoption



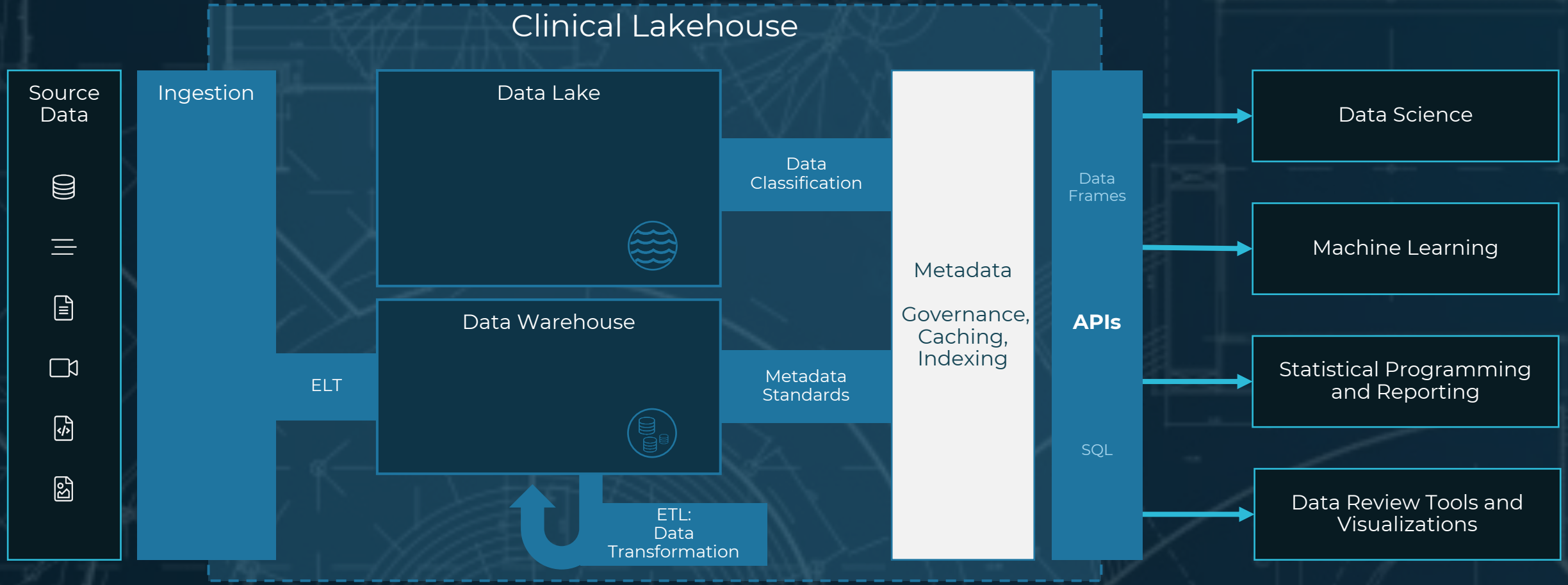
Case Study Lakehouse Adoption



Case Study Lakehouse Adoption

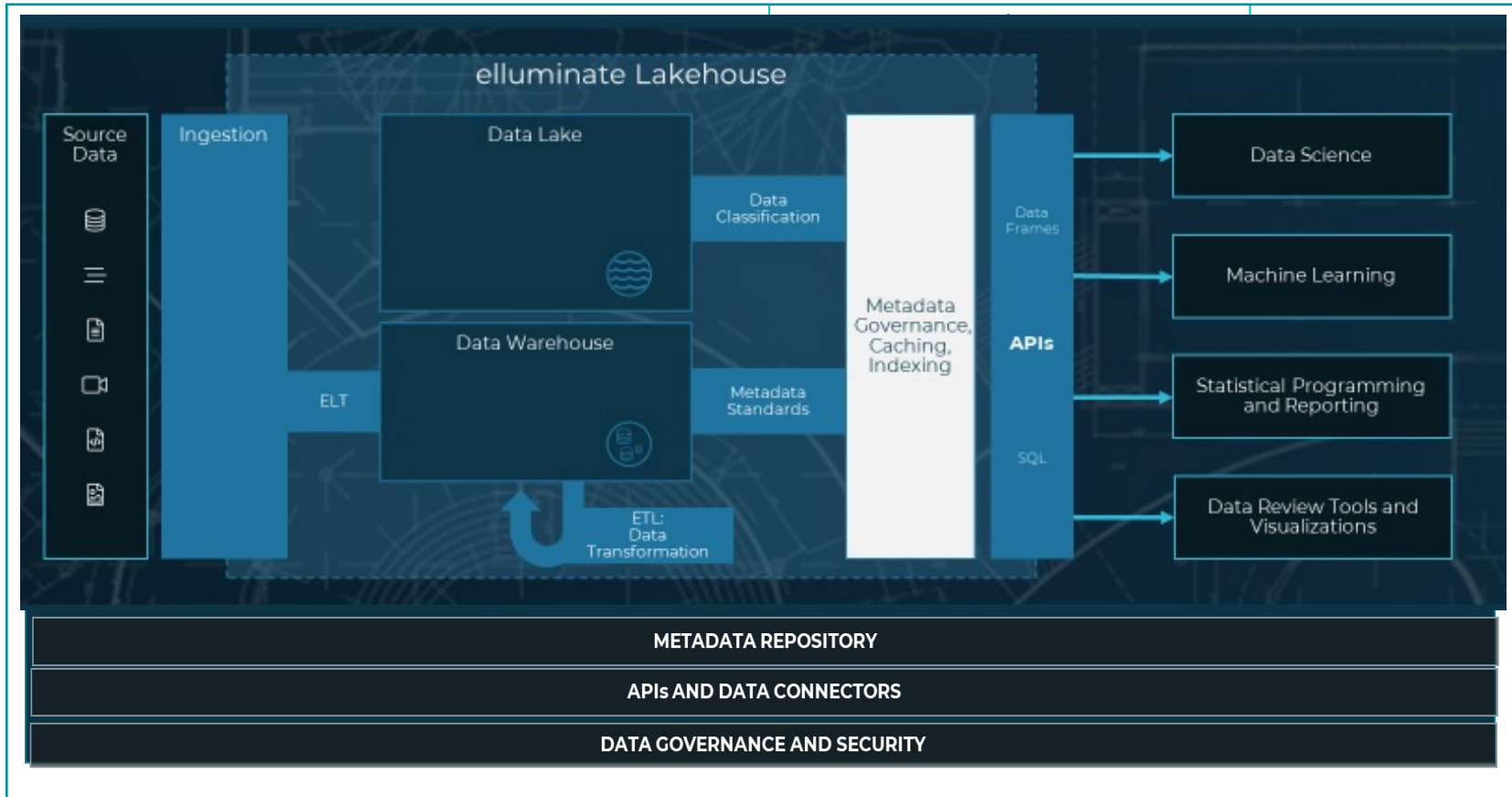


Case Study Lakehouse Adoption



So, why is the Lakehouse critical now?

Clinical Lakehouse architecture better supports modern clinical trials while ensuring a solid foundation for future digitization



Seamless user experience across systems and capabilities

Data access and sharing, no matter where the data lives

Interoperable ecosystem of tools and capabilities

AI augmentation to automate processes and decisions

High compute for advanced analytic workloads

Build your own models where the data lives



Walter Reed National
Military Medical Center

Thank you!

eclinicalsol.com
[877.355.8668 \(877-ELLUMN8\)](tel:8773558668)
info@eclinicalsol.com

[@eClinical_Sol](#)
www.linkedin.com/company/eclinical-solutions/

