Benjamin Flores
12 March 2026

The Modern Scientific Data Stack

Modern science is producing more data than ever before. Artificial intelligence, machine learning, and advanced analytics promise to transform discovery and manufacturing across pharmaceuticals, biotechnology, and geochemistry.

Yet many laboratories still rely on fragile and manual methods for collecting instrument data.

Scientists often encounter environments where data must be exported manually, parsed through custom scripts, or transferred between systems using unreliable integrations.

These challenges reveal a fundamental architectural gap in laboratory data systems.

The Modern Scientific Data Stack

To support modern scientific workflows, laboratories require layered architecture for reliable data flow. This architecture introduces a new foundational layer in laboratory computing: the **Scientific Data Capture Layer**.
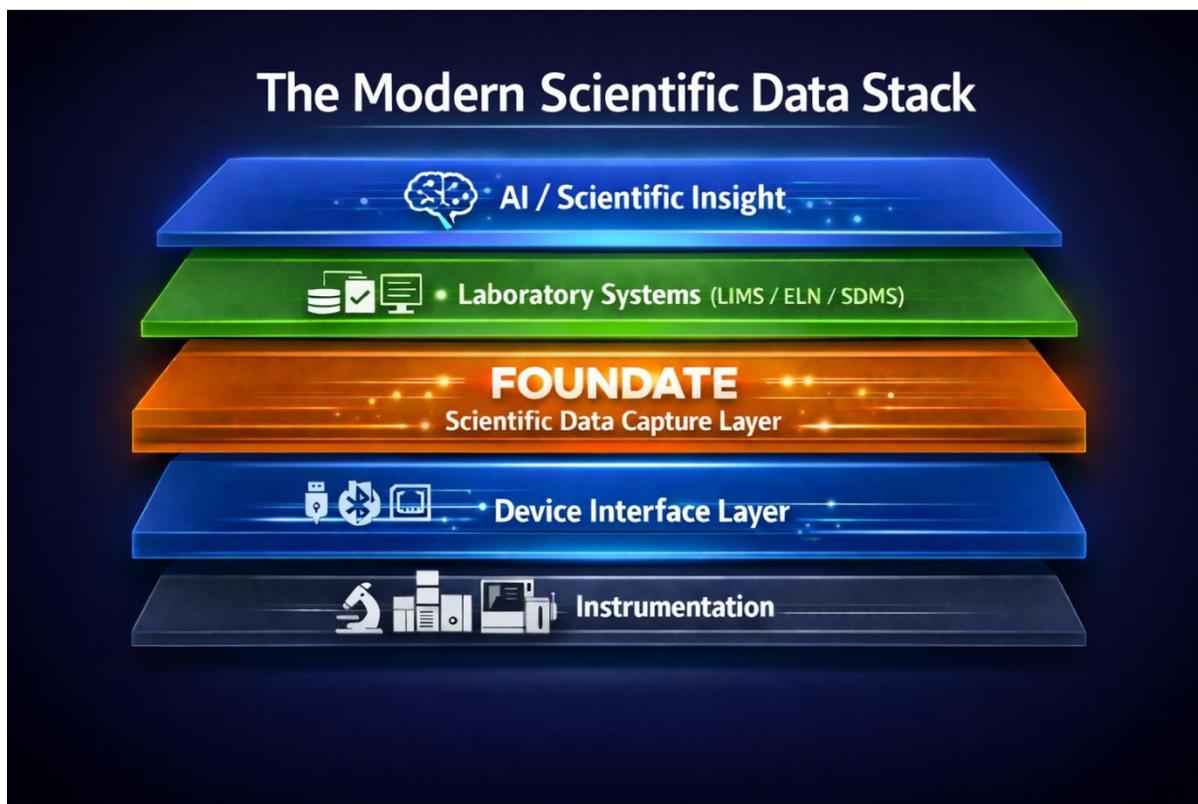
Instrumentation

Physical devices that generate raw scientific measurements.

Device Interface Layer

Protocols and communication channels used to connect instruments, including serial, USB, Bluetooth, Ethernet, and vendor APIs.

Scientific Data Capture Layer

A dedicated infrastructure layer responsible for capturing, normalizing, validating, and securely routing scientific data.

Laboratory Systems

Systems that manage and organize laboratory information, including LIMS, ELN, and SDMS platforms.

Scientific Insight Layer

Advanced analytics, machine learning, and AI models used to generate scientific understanding and operational intelligence.

The Missing Layer

Historically, laboratory environments lacked a standardized **Scientific Data Capture Layer**.

Instead, laboratories relied on custom scripts, manual exports, and custom integrations.

This approach creates fragile data pipelines that are difficult to maintain and scale.

As laboratories move toward automation and AI-driven workflows, reliable data capture infrastructure becomes essential.

A Foundation for Scientific Data Infrastructure

The **Scientific Data Capture Layer** serves as the foundation for modern scientific data systems.

By separating device communication from laboratory systems, organizations can create reliable, scalable, and secure data pipelines.

This architecture enables laboratories to move beyond manual data collection and toward a future of fully automated scientific data infrastructure.