## **Nays Point-to-Point Integrations May Fail**

Here are some challenges R&D IT may face when connecting data sources like instruments with targets like ELN, LIMS, or SDMS:



#### **Contending with Complexity**

Instruments, software, and other data sources generate data in a staggering array of formats, sometimes partially or fully proprietary, often poorly documented. Biopharma workflows, too, may have dozens or hundreds of steps.

Do-it-yourself pipeline-building efforts can bog down in early engineering stages, and take a long time to deliver reliable benefits.

### **Fragility Guaranteed**

Making a prototype data pipeline work on the coder's laptop is one thing. Making it flexible, secure, robust, scalable, and protective of data integrity in production is another.

A lot of platform know-how is required to keep pipelines running correctly, and make them maintainable. Lacking this, R&D IT can easily become mired in frequent data recovery and software maintenance fire drills.





#### **Context Uncertain**

Point-to-point connections prioritize using data immediately to serve requirements of a single scientific workflow. So source data may only be partially extracted and preserved, and broader context is easily lost.

For long-term utility, source data needs to be deeply parsed and often benefits from further enrichment. But doing this well, in a flexible, generalized way, demands deep understanding of scientific use-cases and workflows, and may require lots of input from scientists — taking them away from other important work.

#### **Storage Unplanned-For**

Point-to-point connections often include an archival storage step, but their goal is to put data to work right now, not enhance its value long-term.

That becomes problematic for data scientists and others who want to understand what data exists, then find and consume it easily, without needing to do a ton of re-parsing and curation. To do this right requires indexing, a comprehensive and scalable cloud storage architecture, and harmonization (see below).





#### Harmonization Absent: Not FAIR!

Building a point-to-point connection that works means understanding extracted and parsed data. But unless this understanding is preserved, enriched, harmonized, documented, and shared, the value of your work (and perhaps of the data itself) is being thrown away.

To ensure data's maximum value and usability, every point-to-point connection should make data FAIR, and index and store it for easy access. That means adopting or developing an extensible, comprehensive set of schemas for many different kinds of scientific data, and making harmonization a required step in every point-to-point dataflow.

#### Untangle Point-to-Point Connections with TetraScience

The Tetra R&D Data Cloud is a cloud-native platform for managing scientific data through its whole lifecycle. Productized integrations with popular instruments and software run on a flexible, robust, self-monitoring service architecture, speeding time-to-value and reducing operations costs. Ultra-reliable cloud storage for raw and harmonized data is supported out-of-the-box. Deep parsing, data enrichment, and harmonization make Tetra Data FAIR, letting Scientists, Data Scientists, R&D IT, and other data consumers find and curate data quickly and process it with best-of-breed analytics, visualization, machine learning, or other software tools.

To accelerate your digitalization and integration roadmap, request a demo at tetrascience.com

# o d'itetrascience