

Problem Statement:

“With a Next-Generation LIMS like Semaphores Labbit™, the high amount of money and effort ultimately wasted when a LIMS Solution never gets deployed or fails adoption may become a problem of the past, as it was a real and painful lesson learned for too many in the past.”

If you were unfortunately involved in one of these debacles, we feel for you! We have some examples of LIMS projects with legacy vendors, one lasting 3-years with \$3 million dollars wasted, and another lasting 5-years with \$5 million dollars wasted. There are several factors in play here and they can become a perfect storm. It can often start before the vendor selection process. We created a “Top Ten List of Reasons” why LIMS implementations had major problems or failed based on our experiences in the industry:

1. We all agreed we needed a LIMS, but everyone in our lab had a different idea of what a LIMS is supposed to do.
2. Our requirements were pulled from a spreadsheet somewhere, and we did not create any business process maps or laboratory workflows of the areas that our LIMS were going to serve.
3. We didn't capture improved business practices like naming conventions, optimized and harmonized workflows, or create business SOPs to follow.
4. We assumed the modules of our LIMS solution would handle all these details for us.
5. We tried to take on too much scope off the hop because we couldn't incrementally deploy process capabilities because the LIMS required us to adopt everything all at once.
6. We needed to customize many of our workflows because the out of the box solution couldn't support our way of working
7. We were going to deploy a LIMS designed for static processes in early research.
8. We wanted to use our LIMS for everything we do in our business.
9. We didn't want our LIMS to be scientifically aware.
10. We didn't focus properly on our real needs and provide value for those needs so we could prioritize them.

What is LIMS and Why it Holds Utmost Importance

Laboratory Information Management Systems (LIMS), are software solutions used in labs to manage samples, workflows, and data, improving efficiency, and maintaining quality. They automate many tasks, enforce standard operating procedures, organize data, and generally store and manage information from laboratory processes, making them essential for modern labs that need to demonstrate quality, comply with regulations, or deliver a high level of customer service. In our opinion they can include the management of the following needed laboratory operation processes:

1. Request
2. Samples

3. Test
4. Analysis
5. Reporting

These issues get compounded by the software partner selling something that the company is either not ready for or where the software is not a good fit for their business model. We often see that requirements have not been collected, the present and future business needs of the customer are not thoroughly understood, and all of it has been not well understood by the vendor, in turn leading them to propose an inappropriate solution.

Underlying all these challenges is lacking a shared understanding of the personas and work streams of the business. A critical component of Business Process Mapping (BPM) is tying the workflows to specific personas or roles. (We go into detail below on BPM) When attempting to select a solution for a lab group, it is critical to visualize and thus understand how the different personas in the lab are going to interact with the solution/system. Are they sample centric, test centric, experiment centric, request centric, analysis centric or reporting centric? Are they technicians, lab scientists, managers, or lab operations? How do all these roles interplay and hand off work? Most often it's a mix of some or even all the above. Getting these details from the start is so important and will determine who and how these people will interact with the software.

It's our opinion that a modern LIMS solution should be built from the ground up to consider the fundamentals like, FAIR data, processes, and instruments/automation as well as the capturing and intercalation of the scientific business process mapping into a functional graphical user interface (GUI). This is a game changer as it solves the "we don't need to process map what we do" problem and forces configuration based on BPM. Now you have vetted real requirements and user stories based on the real documented and in-use laboratory workflows. We have observed many times where 40% of the requirements gathered were incorrect in projects over the years! Reasons for this included:

1. Poor requirement gathering techniques (i.e., text only, no process maps)
2. Requirements changed over the long implementation process due to process change or evolution
3. Missed stakeholders and or end users

Semaphore Solutions created a solution, Labbit LIMS, that addresses these issues directly in the software itself. This was all done on purpose because before the team at Semaphore Solution built Labbit, they spent years supporting and configuring other LIMS systems. They got to see things like solution bloat, the inability to handle certain data types and the poor configurability issues, and even the need to customize. They set out to build a solution based on the best available standards and technologies to let laboratories work with their informatics, rather than despite them. Technologies like RDF Knowledge Graphs, Ontologies, BPMN, Digital Provenance, and taking the FAIR principles as requirements rather than an

afterthought or add-on module. This intent to build a next generation LIMS platform is already proving its worth in the life sciences industry.

A **knowledge** graph (Figures 1 and 2) uniquely encodes the relationships between data entities explicitly into the database rather than implicitly like in other database management systems. This results in a knowledge graph-structured data model or topology. This graph structure is able to represent and operate on data in a human AND machine-readable form as well as combine entities of various types in one network. Knowledge graphs are often used to store interlinked objects, events, situations or abstract concepts or entities-- while also encoding the free-form relationships underlying these entities. This is where the term semantic comes from. (Wikipedia. Timón-Reina et al do a wonderful job of exploring why graph databases are highly applicable to the biomedical domain. (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8130509/>))

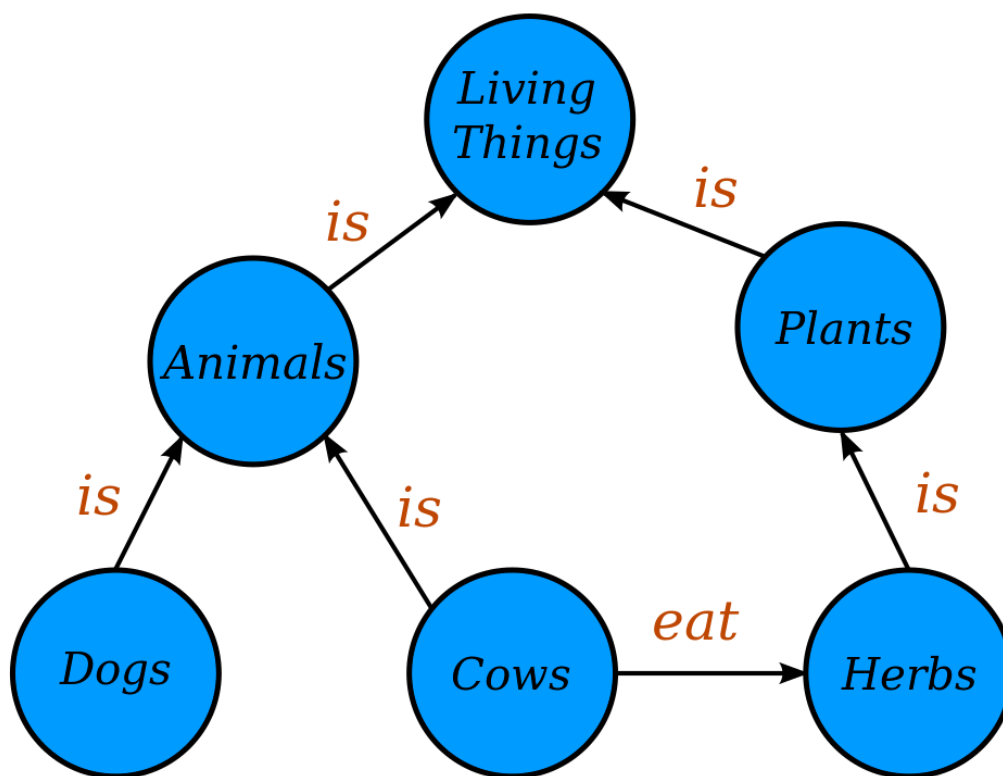


Figure 1-Simple Knowledge Graph Representation Example

In a drug and or therapy environment knowledge graph, the primary entities connected by relationships are molecules, drugs, targets, diseases, variants, biological functions, pathways, locations and more. The relationships have multiple attributes, including relationship type, direction, effect, context and source. The causality of the relationships is represented through direction. Causal relationships frequently carry information about the direction of effect (activation and inhibition) that can be leveraged in powerful analytics. Relationships are annotated with the full experimental context (e.g., tissues or organisms). Entities also

have attributes; for example, they are mapped to public identifiers and synonyms to support data integration.

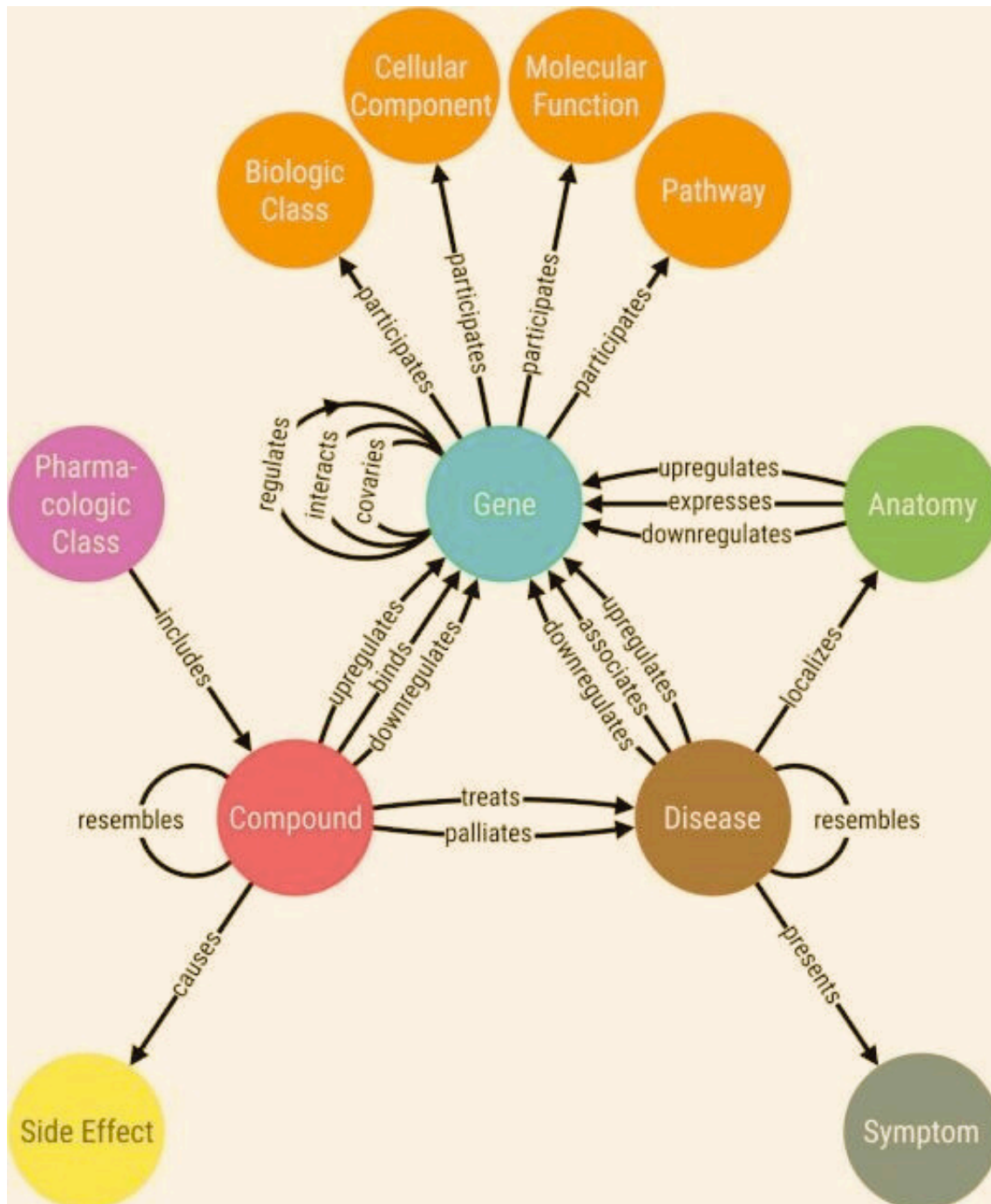


Figure 2- More Complex Drug/Therapy Discovery Example
(<https://www.ncbi.nlm.nih.gov/corecgi/tileshop/tileshop.fcgi?p=PMC3&id=127563&s=103&r=1&c=1>)

In informatics, an **ontology** makes up a representation, formal/standard naming, and definitions of the categories, properties, and relations between the concepts, data, or entities that pertain to a particular domain. More simply, an ontology is a way of capturing and managing the properties and relationships, by defining a set of terms and relational expressions that represent the entities in that domain. (Wikipedia). There are two critical things to know about Ontologies. The first is that once created an ontology can map onto any

other published ontology, thus removing the need to translate data between domains. And the second is the amount of work being put into domain-specific ontologies by their respective practitioners. There are currently over 1135 different ontologies mentioned on Bioportal alone, and that's not even referencing ontologies outside of the biomedical domain! (<https://bioportal.bioontology.org/>). Before you start data modeling it's wise to peruse for an ontology that has already done that work for you!

“The capabilities that Labbit can provide are fundamental to two main goals in industrialized science today: reproducible science and FAIR data and processes.”

The ability to succeed in laboratory informatics depends on how you can handle complexity and capture the details. The ability to drive future efficiencies will all be based on the data quality and level of associated contextualization (metadata) captured. You simply can't go back and impute data with made up metadata! Although this is being debated on a consistent basis.

What is BPM and Why it is Critical

Business Process Mapping (BPM) is a discipline of capturing, in these cases, the scientific business process that occur in a particular laboratory environment. (See Figure 3) Business Process Model and Notation (BPMN) is a standard used in creating graphical BPMs in a Business Process Model. Standards are essential in driving consistency and reproducibility.

So why produce BPMs and Business Process Models captured in a laboratory? There are multiple reasons or answers which are very pragmatic and straightforward. They include:

1. It provides an element of truth and detail of what's occurring in the laboratory.
2. It reduces the time needed to gather requirements and ensures they are accurate and concise.
3. It dramatically improves training and reduces the learning curve for new hires or new trainees.
4. It substantially aids in process optimization and harmonization by revealing reasons for change.
5. As living documents, periodic reviews of the BPMs can ensure consistency and any need for improvement or change.
6. It provides management a detailed understanding of laboratory processes from a high to low level.
7. It leads to a dramatic increase in overall lab operations and efficiency over time.
8. End to end BPM diagrams linked to process metadata at high- and low-level steps can provide insight into why metadata is not reaching the final repository correctly for utilization in data mining.
9. BPM diagrams themselves provide the metadata of the process which operated on any given sample, significantly enhancing audit trails with fulsome context and reproducibility.

Who and Where

LIMS systems are not ideal for very dynamic laboratory processes, they are used for static and repeatable processes. Case in point, Research processes are not conducive to LIMS systems except where there are processes that are highly repeatable like assay screening and other types of testing centers. Development is a heavy use area for LIMS systems because the processes are more static, and the testing can be managed with a competent LIMS system. Development LIMS also pass/fail results on approved product specifications. Once you're in the development stage, and beyond, regulation and validation of software becomes critical. The good news is next generation LIMS that natively utilize BPMN, like Semaphore's Labbit, has been built with validation and CFR 21 Part 11 compliance in mind. See (<https://labbit.com/resources/validation-deconstructed-navigating-ig-og-cq-and-pq>) One noted challenge is to keep user requirements at a high enough level so that they change less frequently than the functional requirements and map them carefully to the higher-level ones. If functional requirements are handled mostly by the vendor software, then customers can rely on vendor documentation and procedures for validation.

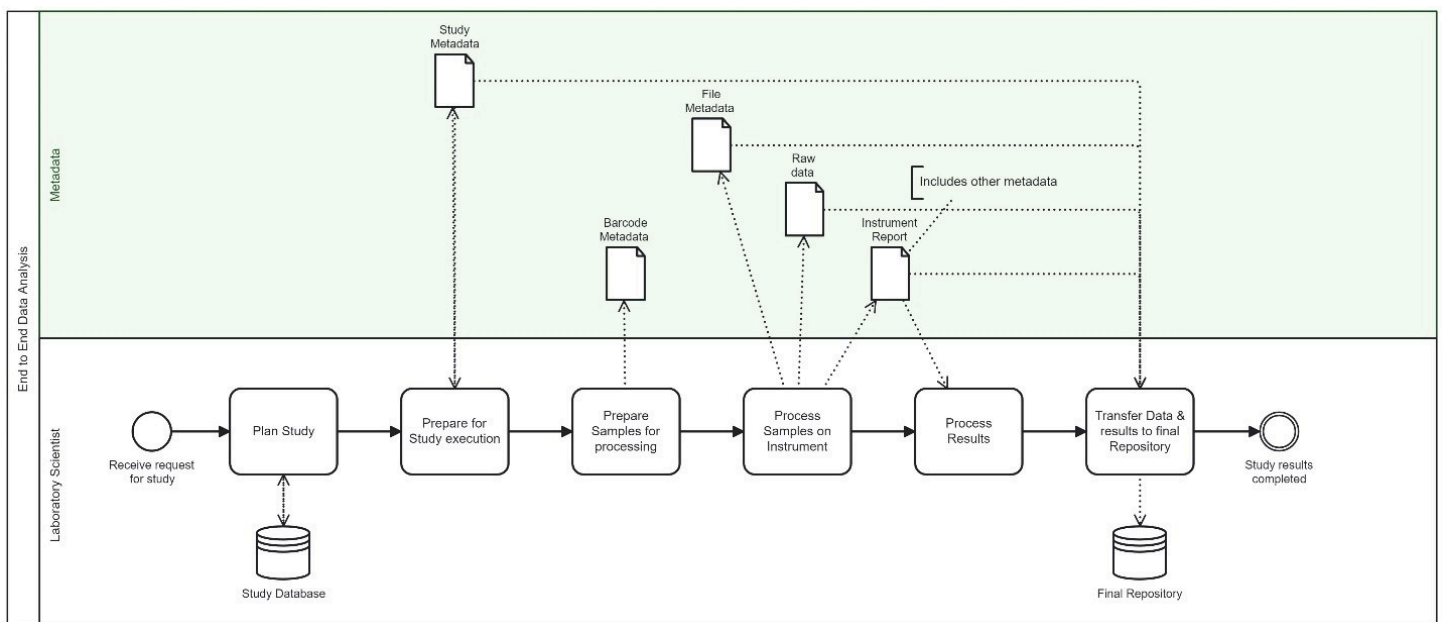


Figure 3- Example Simple Business Process Map (BPM)

How

In today's externalized R&D environment hosting scientific solutions on premise is risky. You want to take full advantage of cloud native capabilities and leverage the scalability, the seamless access (with security), and the reduction in cost for efficient support and maintenance. This is especially true where collaboration and analysis of data across sites,

domestically or globally, is needed. Performance is one of the key adoption parameters and having it hosted in the cloud will usually provide better performance.

If you had major LIMS failures, or your current LIMS is costing you too much and the technical debt is building, it may be time to look at a next-generation LIMS that has taken the proper technological steps to guard your processes and data so that you get a high ROI and much lower TCO than with previously attempted or deployed solutions.

When 20/15 Visioneers encounters a scientific or laboratory informatics solution that was built with expertise and innovative forward thinking we write about it. We are experts in this field and have worked in the industry for over 30 years and prepared clients and salvaged many client projects. If you found this Industry perspective useful let us know.