Benchling

Top 6 Questions R&D Organizations Struggle to Answer

The impact of legacy data management on life science

benchling.com

Table of Contents

Top 6 Questions R&D Organizations Struggle to Answer

The impact of legacy data management on life science

The life science industry is shifting to a new model to support the next generation of therapeutic development. R&D has become more specialized and more collaborative, and produces vast amounts of highly complex data. In order to stay competitive, life science companies need data systems that help address a new set of challenges.

Teams are more distributed and specialized than ever before, yet they need to collaborate closely.

For organizations to make informed business decisions, they need to be able to centralize, interlink, and analyze a larger quantity of complex data.

Scientific platforms are rapidly evolving, so processes need to be able to adapt to capture relevant data.

Without addressing these challenges, it's difficult for scientists and R&D leaders to access the information they need to make fully informed decisions. At the level of an individual scientist, this can negatively impact their productivity and the quality of their outputs. For R&D leadership, suboptimal cross-team collaboration, a lack of centralized data, and sluggishness to respond to process changes lead to knowledge gaps that slow down the pipeline.

To better understand the nature of these challenges and identify the core questions that R&D organizations are struggling to answer with their current systems, we spoke with scientists and R&D leaders throughout the industry. In this document, we discuss the six questions that emerged as common themes, as well as suggestions for how you can address them in your own R&D organization.

1 What are the upstream sequences that correspond to a certain physical sample?

Why is this difficult to answer?

- Legacy systems can't connect downstream results to upstream sequence data, creating a disconnect between experimental results and inputs. Scientists have to manually track down sequence data or repeat experiments.

Why is this top-of-mind?

- Maintaining the connection between sequences and samples is crucial to evaluating the performance of the sequence / cell line.
- Downstream team members need access to data from earlier stages of the experiment to inform their work.
- Core teams may receive a sample, but without all of its relevant data and context, the sample becomes meaningless, wasting resources and time.

What's the solution?

- Connect your repository of sequence data to your inventory and process management systems. By integrating these tools, you can prevent gaps in the data and keep the pipeline moving.

The Stats on Legacy Systems

4/10

The satisfaction of R&D teams with the cross-team collaboration enabled by legacy systems.*

**Statistics from a recent poll of Benchling customers before implementing Benchling*

2

For a given candidate, what are all of its related cellular and molecular entities?

Why is this difficult to answer?

- Modern life science R&D involves a wide range of interrelated entities that produce specific types of data. To get the complete picture of a candidate, all of these data need to be captured and interlinked. Legacy software simply wasn't built to model the complexity of biological entities.

Why is this top-of-mind?

- To fully understand a candidate and inform a go/no-go decision, R&D organizations need to understand the full cellular and molecular genealogy of that candidate.
- Compared to traditional small molecule R&D, modern life science R&D is much more complex. The full history of a candidate comprises many different types of data, from plasmids, to cell lines, to oligos.
- Comparing the genealogies of different candidates side-by-side can provide insights for future lead identification.

What's the solution?

- Use tools that are built with biology in mind to capture and interlink the relevant types of data.

The Stats on Legacy Systems

2.2

Hours

The amount of time the average scientist spends each week assembling the genealogy of a single sample.

3

What are all the assays, parameters, and consumables that led to a certain candidate?

Why is this difficult to answer?

- Spreadsheets, emails, and legacy R&D software don't provide an easily accessible experimental history. Data is often lost, inaccurate, or duplicated.

Why is this top-of-mind?

- Data are the foundation of any successful life science R&D program. Without centralizing and analyzing all the data behind a candidate, companies can't get the insights they need to move their pipeline forward.
- Anytime data are siloed is a chance for data to go missing. Connecting assay data with process parameters ensures your team doesn't waste time duplicating experiments.
- Recording experimental results along with all of the relevant context is critical for regulatory approvals.

What's the solution?

- Connect assay data with data on processes and consumables usage. Scientists and decision makers will be able to understand the steps that led to a result, making their processes transparent, repeatable, and error-proof.

The Stats on Legacy Systems

2/10

R&D teams' ease in finding the necessary data related to a specific candidate.

4

What's the real-time status of our projects?

Why is this difficult to answer?

- Life science R&D processes are inherently complex and iterative. Getting a complete picture of project progress is difficult when using systems that weren't built with biology in mind.

Why is this top-of-mind?

- Modern life science R&D involves a large number of disparate teams and processes, so R&D leaders need a way to manage them concurrently.
- Having access to real-time project status means you can more accurately predict downstream timelines and adjust them when necessary.
- Immediate insight into project performance lets you promptly identify and communicate changes in project priorities.

What's the solution?

- Leverage your data management system to unify project tracking with experimental notetaking and results reporting. This will give you a clear line of sight into project progress so you can quickly pivot when necessary.

The Stats on Legacy Systems

4

Hours

The amount of time the average scientist spends each week compiling reports on project progress.

5

What are the bottlenecks in a given process, and how can we iterate?

Why is this difficult to answer?

- It's difficult for organizations to identify bottlenecks and optimize processes without metadata on processes themselves. What's more, rigid software systems make it challenging to iterate on processes once bottlenecks are identified.

Why is this top-of-mind?

- Quickly identifying bottlenecks before they become a larger problem is critical to saving time and resources.
- Immediate insight into what's working and what needs improvement is key to process optimization.
- R&D organizations need to learn as much as possible from each experiment. Tracking less data makes process development more costly.

What's the solution?

- In a flexible digital system, gather data on process performance (ex. process runtime, quality of outputs, reagents used) and results. This will help you quickly identify bottlenecks and ensure that you can rapidly iterate.

The Stats on Legacy Systems

3/10

R&D teams' satisfaction with the data capture and reporting capabilities of legacy systems.

6

How many FTEs and resources do we need to allocate to a new project?

Why is this difficult to answer?

- For novel life science R&D processes with few industry benchmarks, forecasting is largely dependent on the company's own historical data. Unless you're gathering that data in a queryable manner, forecasting can be more of an art than a science.

Why is this top-of-mind?

- Understanding the performance of past projects and the resources they required is key to accurate forecasting.
- Over- or under-investing the right amount of resources in a project upfront can affect the ultimate success of the project.
- Accurate forecasting saves scientists time, reduces unnecessary busywork, and conserves resources.

What's the solution?

- Connect data that describes project performance and time-to-milestone with data that describes reagent usage. With insight both into project performance and reagent usage, you'll be able to innovate more efficiently in the future.

The Stats on Legacy Systems

5/10

R&D teams' confidence in data recorded in legacy systems.

Unlocking the Answers

Life science R&D is developing incredible solutions to global problems. But this groundbreaking work is often being done with outdated tools that weren't designed with the complexity of life science in mind.

If you're asking these questions and the answers seem obscured, you may already know that the tools you're using aren't innovative enough to keep up. The only way life science R&D organizations can get the answers to these six questions quickly, confidently, and reliably is with an end-to-end cloud solution purpose-built for modern life science R&D.

What to look for in a modern life science R&D solution

1. Interconnected molecular biology design and inventory tracking tools that unify sequence data with process data.
2. A solution built with biology in mind that captures the complexity of biological entities.
3. The capability to track, interlink, and surface the numerous types of data that make up the experiment lifecycle.
4. Unified project tracking, experimental note taking, and results reporting that give you a clear view of project progress.
5. Real-time data on process performance, captured in a flexible system. This will allow you to identify process bottlenecks and rapidly iterate.
6. Integrated process performance data and reagent tracking, so you can improve resource and FTE allocation.

Learn how Benchling addresses the core challenges of life science R&D at benchling.com