Gaining actionable insights from research and development data

Change and flexibility in data collection and analysis at Virent's renewables research and development facility

Key benefits

- Standardized data capture and ensured data consistency and integrity, while reducing the time spent repeating experiments and searching for information.

- Centralized and comprehensive data storage platform helped to integrate three different departments and dramatically improved the ability to quickly search for and access historic results.

- Flexible and extensible platform provided the means to independently configure new system functionality on Virent's timeline as research progressed and analytical methods evolved – rather than needing to involve a third party for custom functionality.

Data management challenges, including data for patent portfolio

Like most startup companies, Virent's initial data storage and management was done on an as-needed basis through simple spreadsheets. This homegrown solution became unwieldly as the company expanded and evolved over time. Disparate file storage locations, incomplete data capture, inconsistent calculation methodologies, and non-centralized complex customized Excel macros made data mining difficult for a company whose most valuable asset is its data and the patents derived therefrom. "The Excel based data management systems had so many shortcomings, which became even more apparent after implementing LIMS," said Virent researcher, Matt Van Straten.

"Originally, most researchers were very apprehensive to break from the comfortable Excel-based system and pressed to maintain some form of its functionality. Very quickly after implementing LIMS, almost all of those requests faded away." Virent to date has an extensive patent portfolio, and protecting and preserving the data that is the foundation of those patents is paramount.

To preserve and maintain the value generated through the daily research, it was clear that a strategic approach to data management was needed. Virent decided that implementation of a LIMS with an Electronic Laboratory Notebook (ELN) would provide a centralized, standardized way to capture data and reduce the cost of repeating experiments and searching for data.

Thermo Fisher
SCIENTIFIC

"The implementation of [the] LIMS has helped organize data in a uniform, easy-to-interpret manner. It allows new users to search through the results from old, unfamiliar experiments without questioning validity or accuracy."

– Chris Holland, Researcher

Selecting Thermo Scientific Core LIMS and Core ELN software

In order to be an effective solution to meet Virent's data needs, the system needed to be:

Flexible

Virent's lab-scale pilot plants are all unique. Putting structure in place to standardize data collection and manipulation without constraining the design of research campaigns was critical.

Integrated

The LIMS needed to accommodate the unique data handling requirements of the Analytical, Operations, and Research groups. Integrating these three departments, along with an ELN, onto one platform was key to unlocking the full value of a LIMS for Virent.

Configurable/extensible

Cutting edge research changes rapidly and the ability to independently modify existing and configure new entity types in line with evolving requirements was essential.

Thermo Scientific™ Core LIMS™ and Thermo Scientific™ Core ELN™ software met Virent's requirements by providing capabilities to structure data collection and management in an environment that demands flexibility. The Virent configurations highlight the flexible, extensible nature of Thermo Fisher™ Platform for Science™ software and demonstrate how it supports complex requirements to service a variety of customer applications.

About Virent

Virent, Inc. ("Virent") has developed novel technology for the generation of renewable, direct replacement fuels and chemicals using a wide range of biomass derived feedstocks. Its patented technology features catalytic chemistry to convert plant-based materials into a full range of products identical to those made from petroleum, including gasoline, diesel, jet fuel, and chemicals for plastics and fibers. The products are drop-in replacements that enable full utilization of existing logistics infrastructure without blending limitations. Founded in 2002, Virent has repeatedly demonstrated its status as a global renewables leader, most recently through production of the world's first 100% plant based polyester shirts using its BioFormPX® product.

Virent is located in Madison, WI, and has a state of the art research and development facility, including:

20+ lab-scale pilot plants
Automated for 24/7 operation

2 demo-scale plants
Scaled to 100x lab-scale

Catalyst development laboratory
Production of catalysts customized/tuned for specific processing step

Metal fabrication shop
Provides prototype development capabilities and supports fast plant configuration changes

Analytical lab
Rapid analysis turnaround for catalysts, feedstocks, intermediate streams, and finished products.

fig.1. Virent's laboratory facilities

Implementation challenges & opportunities

Implementation can be an overwhelming process, requiring a significant investment of time and resources to accurately define and vet requirements. Team members struggled to fit implementation work in around their normal daily responsibilities. Understanding workflows, both as they currently existed and how they could be improved, also required team discussion. "Implementation of a new system can be a great opportunity for procedural improvements," said Dana Hatch, Analytical Laboratory Manager. For Virent, there were three key risks addressed during implementation:

Change management

Users were comfortable with their existing spreadsheets and paper log sheets, so it was important to put them at ease about using the new system. Getting them comfortable started early on by identifying department representatives to be system super users. Those super users were part of the implementation team, providing input into how the system was configured. They also served as resources for their coworkers during training and after implementation.

Clear messaging about both the benefits and the challenges that the LIMS would bring was equally important to help set expectations. Some end users saw the time it took to complete their daily tasks increase while others realized the benefits of the downstream efficiencies that the system provided. The objective of the LIMS implementation at Virent was to improve data reproducibility, consistency, and searchability, as well as the breadth of data collected. It was important to give the end users, whose tasks would take longer, insight into the broader project perspective and goals so they could understand why they were being asked to change.

Scope

After defining the initial requirements, super users continued to identify areas that could be integrated into the LIMS. Research also continued to evolve, with new assays and innovative processing steps being defined. It was important to create a parking lot for items for future functionality. This meant integration of areas such as the biomass analysis, batch reactors, catalyst assays, and the larger scale demo plants were put on hold initially to focus on the workflows that comprised the majority of the day to day work.

"Setting a limited scope early in the LIMS configuration process was a benefit to project timing," said Hatch. The system was not optimized at go live, and there were still some activities whose data was not captured in the system, but reigning in the scope allowed the team to implement the system more quickly and gain user experience on a daily basis. This approach also allowed the team to better define and prioritize the updates to include in future expansion projects.

Historical data

Integration of historical data into the LIMS was not practical. The substantial amount of data from years of research was valuable, but it was not in a format conducive to back-populating the LIMS database. However, side-by-side comparisons of pre- and post-LIMS data was a requirement.

Tableau™, a third party data visualization software program, was the chosen solution for downstream analysis and visualization. Tableau provided a means to perform complex analysis and visualization of LIMS data while also allowing for the comparison of key metrics from historical Excel data.

"The connection between LIMS and Tableau has significantly reduced the need for overly complicated spreadsheets with calculations that reference data from questionable or sometimes no longer existent data sources. It also allows users to easily compare and contrast the results from experiments that were performed yesterday with experiments from years ago. This helps maintain the relevance of old work and prevents unnecessary duplication of old experiments which saves both time and money."

– Chris Holland, Researcher

Supporting processing equipment modifications through configuration

From a configuration perspective, one of the challenges that required an innovative solution was the integration of the pilot plant model. Lab-scale processing equipment is highly configurable, almost like a set of tinker toys, and new processing lines or reactors can be modified rapidly. Data management strategies need to adjust just as quickly to allow for rapid process configuration, testing, and assessment. This required that the sample data be supplemented with data describing the individual unit operations as well as the overall design of the process configuration.

In Virent's case, a system administrator can quickly build a process configuration in LIMS, Figure 2. Instances of that configuration then define the data the technician should collect during the experiment (e.g., temperatures and pressures) and triggers the analytical tests that should be performed on the samples produced in the experiment, as well as the calculations that are needed on the resulting data.
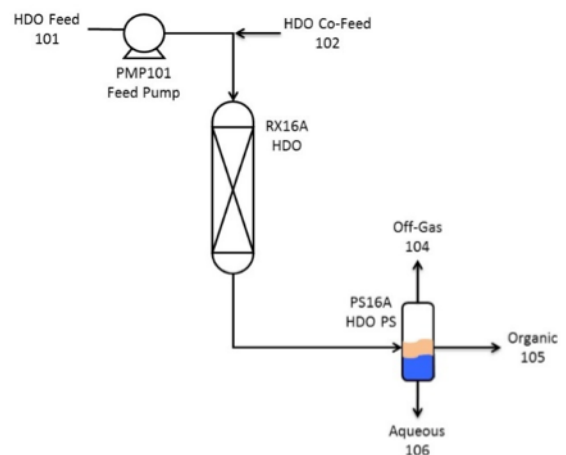

fig.2. Example of a pilot plant configuration

Meeting needs and exceeding expectations

Virent went live with Core LIMS and ELN software in August 2012 with v3.9, and upgraded to v4.1 a year later. With four years of being in the system, it is hard to imagine life without it. Operators, Analysts, and Researchers have one centralized system for data entry and retrieval. Workflows and calculations have been standardized, and there is high level of data integrity. "The implementation of LIMS has helped organize data in a uniform, easy-to-interpret manner. It allows new users to search through the results from old, unfamiliar experiments without questioning validity or accuracy," said Holland.

Pre- versus post-LIMS data analysis

With the research data centralized, the Virent team has experienced improvements in the availability of the data and the ease with which it is entered and retrieved. When researchers want to analyze their data in a new way, they simply query the LIMS database and do an analysis within Tableau.

Revisiting data that predates LIMS is much less convenient. First, it requires time to track down where the data was stored, and then additional time to understand how it was worked up. It is not uncommon for the pre-LIMS data to be incomplete or unable to be analyzed in a new way. If the comparison desired is critical, it may require a duplicate experiment to be run, this time taking advantage of the LIMS data collection and storage.

"When I receive a request to pull together some old data, the first thing I think about is whether the data is coming from before or after we implemented LIMS, because if it is in LIMS it'll be quick and easy, but if it was before LIMS, then I know it's going to take some effort to get it worked up."
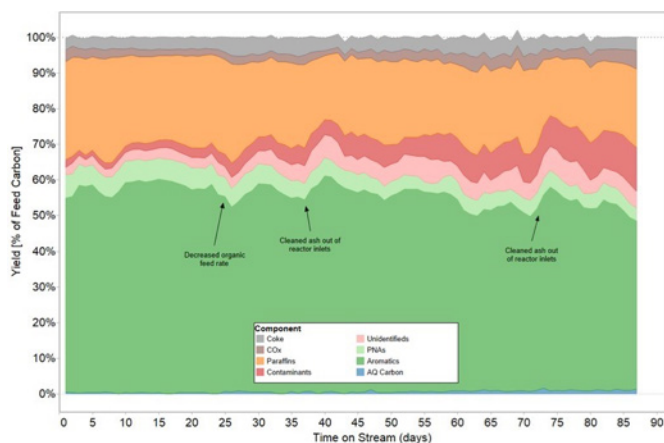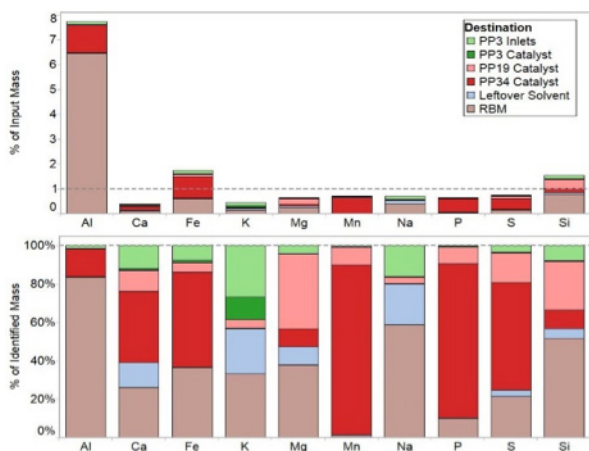
– Matt Van Straten, Researcher

fig.3. Visualization of LIMS data using Tableau

Adding capabilities over time via configuration, without outside assistance

Virent's LIMS Administrator was able to continue expanding the scope of the system using only the out of the box administrative capabilities. Since going live, over 20 new or modified assays were added including relational functionalities between assays. Additionally, entity types were configured to accommodate new requirements around different reactor types, non-standard feedstocks, and larger-scale operations. Being able to extend the functionality of the system on Virent's timeline, rather than relying on a third party, has been a particular advantage of the system.

Providing insights for process and resource optimization

In addition to meeting the specified user requirements, the LIMS has provided additional insights above and beyond the original expectations. For example, a system query can show the analytical sample load per analyst, Figure 4. With four years' worth of historical data, those trends can be used to inform resource planning. Turnaround time, experimental success rates, budgeting, and annual emissions reports are just a few additional areas where the LIMS has exceeded expectations.
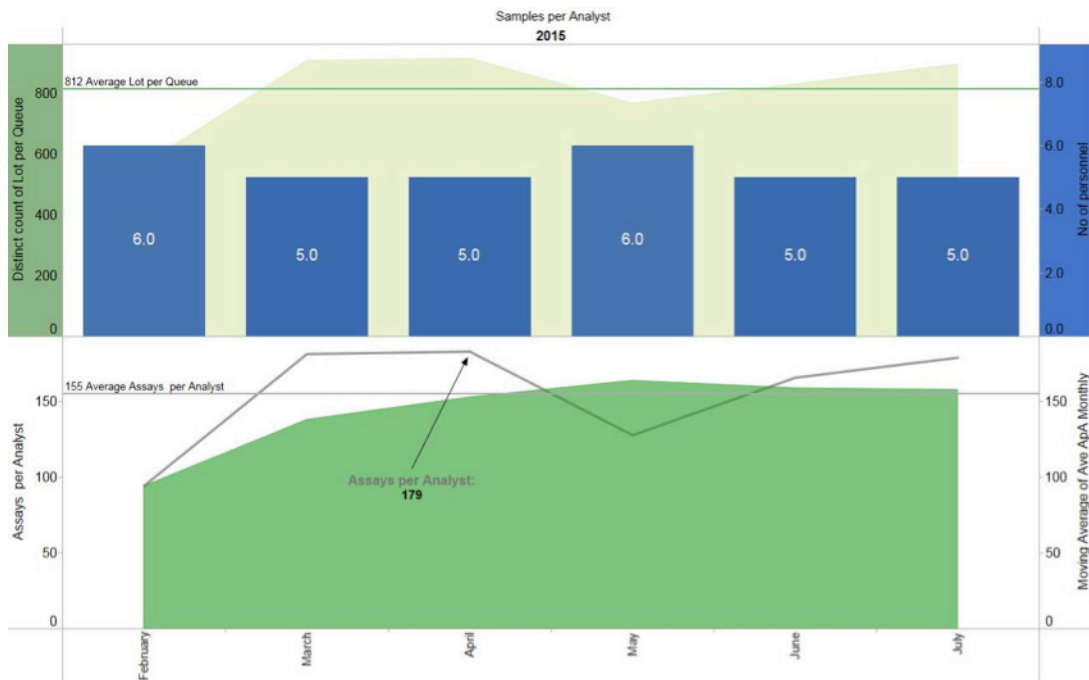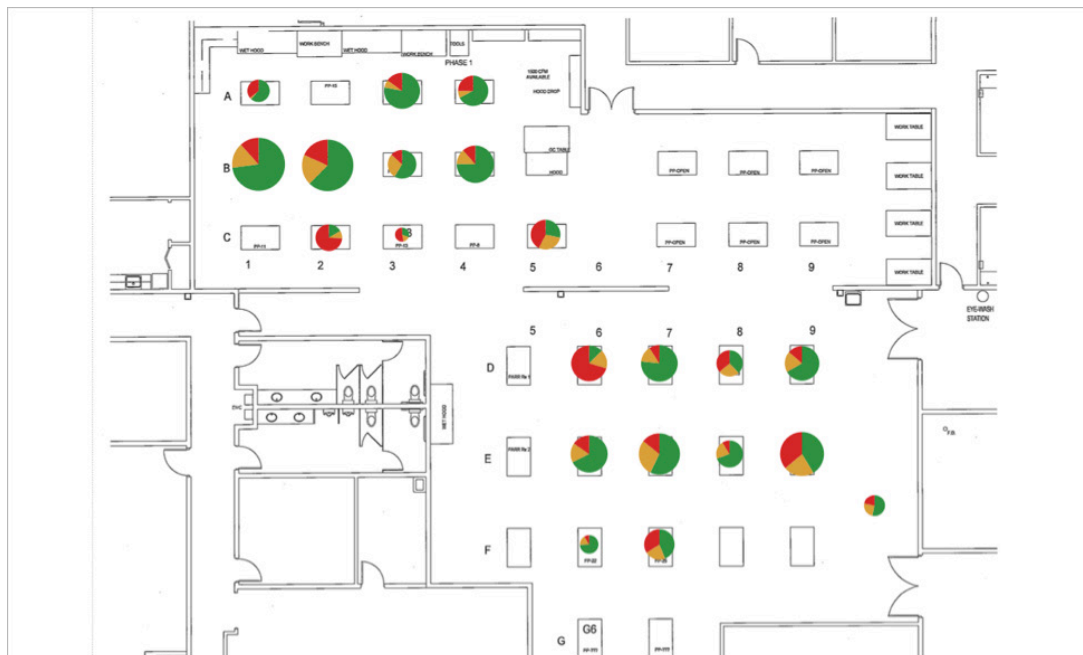

fig.4. Analytical sample load data

fig.5. Using Tableau to map some of Virent's pilot plant data stored in Core LIMS

Simplifying and standardizing lab notebook data management

The ELN has replaced traditional paper lab notebooks for the collection of research and method development data resulting in multiple benefits. The electronic notebooks eliminate the chance of lost information from a misplaced paper notebook or bad penmanship. The centralized location of the notebooks allows for collaboration on experiments and ease in data retrieval from a coworker's notebook. Links can be easily generated to LIMS entities for rapid background information on a sample used in the ELN. Virent's Analytical Lab currently uses the ELN to monitor all instrument maintenance and repairs so this information is now in one location and the query tool is helpful in finding information on past repairs.

Final thoughts

Virent's research and development efforts in renewable fuels and chemicals continue to advance at a rapid pace. Data management requirements are complex and evolve according to Virent's unique workflows and analysis needs. Where traditional LIMS would have become an obstacle that impeded progress, Core LIMS and ELN software has empowered Virent to adapt to changing requirements and enabled on-demand configuration modifications due to the flexible and extensible nature of the system. The centralized, standardized data management solution provided by the LIMS and ELN has given Virent a system that meets their needs today, while positioning them to be successful as their innovative process requirements evolve in the future.

Find out more at **thermofisher.com/informatics**

Thermo Fisher
SCIENTIFIC