

LIMS, AI & Automation

The future of chemical data analysis and management: Toward unified, standardised, and web-based analytical workflows

Richard Lee, Director of Core Technologies, ACD/Labs

Rising expectations for analytical data

Across chemical and pharmaceutical research organisations, expectations placed on analytical data have increased dramatically. Analytical measurements are no longer viewed solely as confirmatory outputs at the end of an experiment; instead, they are expected to inform decisions in real time, support parallel research efforts, and serve as reusable assets for future discovery. Higher experimental throughput, increasing regulatory scrutiny, and the growing adoption of data-driven approaches such as artificial intelligence (AI) and machine learning (ML) have all intensified the demand for analytical data that is reliable, accessible, and context-rich.

Despite this growing reliance, many laboratories continue to struggle with fundamental challenges in how analytical data is captured, processed, stored, and reused. The gap between what organisations need from their data and what their current systems can realistically provide is widening. Addressing this gap requires rethinking not only tools, but also the underlying architecture and philosophy of analytical data management.

Modern laboratories have evolved into ecosystems of standalone applications that are poorly integrated with one another.

The challenge of analytical data heterogeneity

One of the most persistent obstacles in chemical research environments is analytical data heterogeneity. Modern laboratories employ a wide range of analytical techniques—such as NMR, LC, MS, and hyphenated methods—across instruments from multiple vendors, each generating data in proprietary or semi-proprietary formats. Over time, this has led to ecosystems composed of numerous standalone applications, each optimised for a specific technique or instrument but poorly integrated with one another.

As a result, analytical data and the workflows built around it often become fragmented. Data may reside on local workstations, shared network drives, instrument computers, or isolated databases, frequently without consistent metadata or contextual linkage to the experiments that generated it. This fragmentation has several downstream consequences: data becomes difficult to locate, challenging to compare across experiments, and cumbersome to reuse for new purposes.

The problem is not merely one of inconvenience. When analytical data cannot be reliably found, interpreted, or reprocessed, organisations lose opportunities to extract additional value from past work. In many cases, analyses are repeated unnecessarily, or results are trusted without sufficient transparency into how they were generated.

Desktop-centric tools and their limitations

Compounding the issue of data heterogeneity is the continued reliance on desktop-based analytical software. While these tools are often powerful and familiar to scientists, they impose IT limitations that are increasingly misaligned with how research organisations operate today.

Desktop applications typically restrict data access to specific machines or user accounts, making collaboration more difficult—particularly in geographically distributed teams. They also place a significant burden on IT departments, which must manage software installation, updates, licensing, and compatibility across diverse environments. Scaling such systems to accommodate new instruments, additional users, or changing workflows can be slow and costly.

From a strategic perspective, desktop-centric models also inhibit the development of shared analytical knowledgebases. When data processing and

interpretation occur locally, the outputs are often stored as static files or reports, detached from the raw data and the broader experimental context. This limits their usefulness for downstream analysis, auditing, or automation.

Why standardisation and centralisation matter

To support faster and more confident decision-making, analytical data must be both standardised and centralised. Standardisation ensures that data generated by different instruments and techniques can be interpreted consistently, while centralisation allows data to be accessed, searched, and reused across projects and teams.

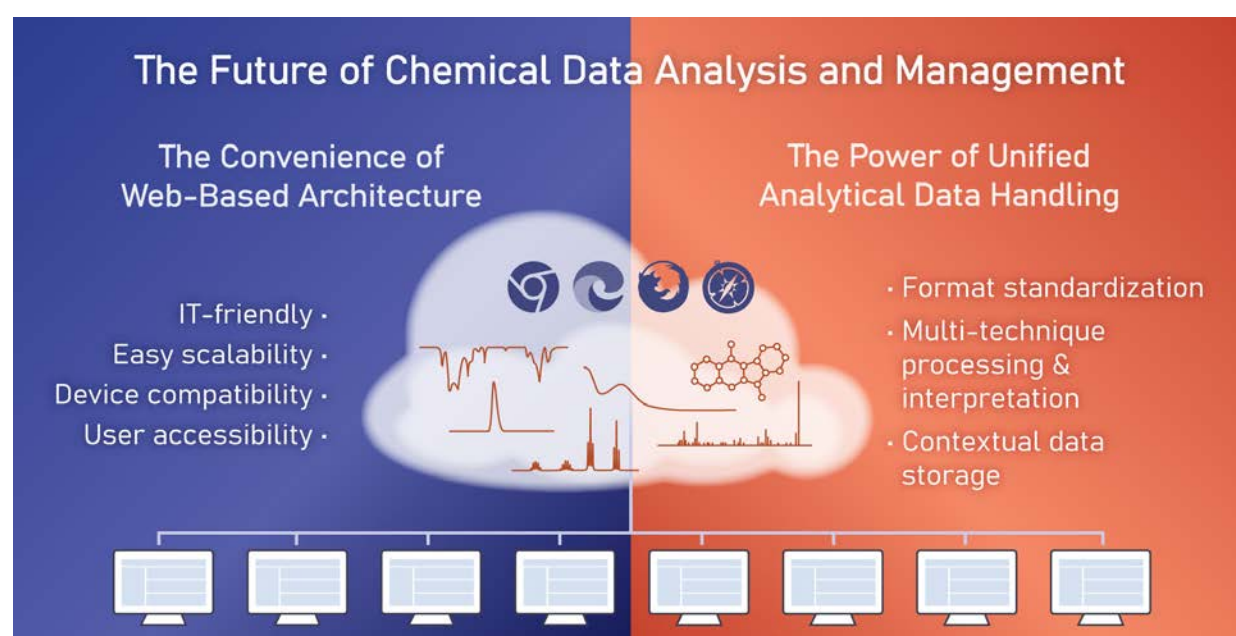
Importantly, standardisation does not imply oversimplification. Chemical research depends on rich, multidimensional data—spectra, chromatograms, structures, annotations, and metadata—that cannot be reduced to simple tables without losing critical information. Effective standardisation preserves this richness while imposing enough consistency to enable automation and interoperability.

Centralised, contextual data storage provides a foundation for more advanced capabilities. When analytical data is stored alongside chemical structures, experimental conditions, and interpretive annotations, it becomes possible to perform chemically intelligent searches, compare results across studies, and reconstruct the full analytical history of a compound or experiment.

Why web-based analysis makes sense for modern labs

If analytical data must be both standardised and centralised to support confident decision-making, then the tools used to work with that data must reflect those same principles. A web-based platform is naturally aligned with standardised, centralised analytical workflows, as it functions as a shared hub rather than a collection of isolated, machine-specific applications.

Unlike desktop software, web-based analysis environments make analytical data immediately accessible across teams and locations, supporting collaboration, reuse, and consistent interpretation. They also simplify deployment, updates, and maintenance, ensuring that standardised workflows and data models are applied uniformly without relying on individual user installations.



For modern laboratories, these characteristics make web-based platforms a practical foundation for standardised and centralised analytical data workflows. They enable advanced analysis and collaboration while addressing the fragmentation, accessibility, and maintenance challenges that have long constrained traditional desktop-centric approaches.

The scientist's perspective: Holistic and accessible analysis

For scientists, the benefits of unified analytical data handling are immediate and tangible. A single environment that supports multiple analytical techniques enables more holistic data analysis, allowing results from different methods to be examined together rather than in isolation. This integrated view supports better interpretation and reduces the risk of overlooking inconsistencies or correlations.

Accessibility is another critical factor. Web-based analytical platforms allow scientists to access data from any location and device, supporting remote and flexible work arrangements without compromising functionality. This is particularly valuable in collaborative research settings, where teams may span multiple sites or organisations.

Equally important is the ability to store analytical data in context. When processed and interpreted data are captured alongside chemical structures, analytical and experimental metadata, they become easier to search, retrieve, and reuse. Over time, this creates a living repository of analytical knowledge that can be easily leveraged for future insights rather than a static archive of files that will rarely be accessed.

The IT perspective: Deployability, scalability, and governance

From an IT standpoint, modern analytical data management solutions must address long-standing challenges related to deployment, maintenance, and governance. Web-based architectures significantly reduce the overhead associated with installing and updating software on individual machines, while centralised user management simplifies access control and compliance.

Scalability is another key consideration. As organisations add new instruments, adopt new techniques, or expand globally, analytical data infrastructure must be able to grow without requiring extensive reconfiguration or redevelopment. Flexible deployment models—whether on-premises, in the cloud, or in hybrid configurations—allow organisations to align analytical data management with broader IT strategies.

Crucially, centralised systems also improve data governance. Consistent data handling practices, audit trails, and permission models help ensure that analytical data remains secure, traceable, and compliant with internal and external requirements.

These systems typically operate on abstracted, tabular data and lack native support for rich analytical objects such as spectra or chromatograms.

Automation: Opportunity and reality

Automation is often presented as a solution to many of the inefficiencies associated with analytical data handling. In principle, automated dataflows can reduce manual effort, improve consistency, and accelerate turnaround times. In practice, however, the heterogeneity of both analytical data and the software tools used to handle it has made automation in chemical research difficult to implement and sustain.

While enterprise automation and workflow platforms have emerged across many industries, their applicability to chemical research is often limited. These systems typically operate on abstracted, tabular data and lack native support for rich analytical objects such as spectra or chromatograms. As a result, they struggle to accommodate the complexity of chemical workflows.

In the absence of suitable platforms, many laboratories rely on custom scripts, macros embedded in legacy applications, or scheduled batch processes. These solutions may work under narrow conditions but are often brittle, difficult to maintain, and dependent on specialised expertise. As workflows evolve, such automation frequently breaks down, forcing teams to revert to manual processes.

Empowering Scientific Automation with Low-Code and No-Code Tools

What chemical research organisations increasingly need is a different approach to automation—one that is built on standardised, centralised analytical data and designed for use by scientifically trained staff. Low-code and no-code platforms tailored to analytical workflows allow scientists to define, modify, and extend dataflows without writing custom software.

By embedding automation capabilities directly within analytical data environments, organisations can ensure that automated processes operate on live, context-rich data. This approach reduces reliance on external developers, shortens turnaround times for workflow changes, and improves resilience as research needs evolve.

Moreover, self-administered automation scales more effectively. As new instruments are introduced or research priorities shift, workflows can be adapted incrementally rather than reengineered from scratch. This agility is particularly important in fast-moving research environments where flexibility is a competitive advantage.

Building a foundation for AI, ML, and advanced analytics

The quality and structure of analytical data ultimately determine the success of downstream initiatives such as business intelligence, AI, and ML. Models trained on incomplete, inconsistent, or poorly contextualised data are unlikely to deliver reliable insights, regardless of their sophistication.

Unified, standardised analytical data from appropriate informatics systems provide the foundation needed to support these advanced applications. By ensuring that data is complete, interpretable, and traceable, organisations can create datasets suitable for predictive modelling, pattern recognition, and knowledge discovery.

Equally important is institutional memorialisation. When analytical data is captured and curated systematically, it becomes a long-term asset that retains value beyond the tenure of individual scientists or projects. This continuity supports both innovation and risk mitigation.

From fragmentation to enablement

The future of chemical data analysis and management lies in moving beyond fragmented, desktop-centric tools toward unified, web-based platforms that treat analytical data as a strategic asset. Standardisation and centralisation are not ends in themselves, but enablers of accessibility, automation, and insight.

By adopting modern analytical data architectures that serve both scientists and IT stakeholders, organisations can improve productivity today while laying the groundwork for tomorrow's data-driven research models. In doing so, they shift analytical data handling from a bottleneck to a leading asset.



Read, Share and Comment on this Article, visit: www.labmate-online.com