A Unified Software Platform for Laboratory Informatics

Graham A. McGibbon, Hans de Bie, David Hardy, Ryan Sasaki, Patrick Wheeler, Carol Preisig

Advanced Chemistry Development, Inc. (ACD/Labs), Toronto, Canada.

Overview

- Quality decisions rely on quality information and data
- A software platform for **Unified Laboratory Intelligence (ULI)** should **capture** analytical data and metadata, ideally from a variety of sources and types of instruments, completely and correctly
- **Analysis** with association of chemical structures and reactions is a key aspect of knowledge generation
- **Automation** improves quality and quantity of laboratory output through fewer manual errors, standardized methods, and reduced administration burden on chemists
- **Storage** and **access** with **evaluation** provide the route to leveraging intelligent decision s from a ULI platform
- Software requires effective interaction and **integration interfaces** for communications with human scientists, and also with hardware and instrument systems to offer and maintain consistent quality

Introduction

One aim of R&D organizations is the strategic assimilation of informatics technology for creating knowledge management infrastructure that copes with the variety and volume of scientific workflows and analytical information generated from modern instrumentation.
However, there are simultaneous pressures to:

- Increase value while lowering cost and simplifying IT infrastructure
- Deploy and maintain software
- Satisfy evolving user requirements and preferences

Amalgamation is one key theme for simplification of disparate software technologies such as CDS, ELN, LES, LIMS, SDMS, and data warehouses, which have proliferated throughout organizations to serve various primary purposes (Fig. 1) . These include data archival, intellectual property protection, and workflow tracking. Simple reduction isn't satisfactory, hence technologies that are designed for effectively integrating, automating, and communicating in Laboratory Informatics Environments are of growing interest.

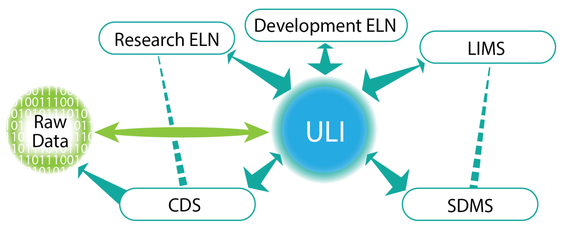

Figure 1—Typical informatics technologies involving analytical data.

Reported here are capabilities in automated workflows involving analytical data with chemical structures. Specifically described is automated homogenization of data from a set of instruments, including NMR structure verification, as one solution. Another widely deployed ACD/Labs solution is for the automation of LC/UV-MS for purity check prior to compound registration.

Methods and the Platform

The ACD/Spectrus Platform v2015 offers web services and automation components as well as thick client interfaces (ACD/Spectrus Processor or ACD/Labs Spectrus Workbooks), and the ACD/Spectrus DB knowledgebase supported by either PostgreSQL or Oracle. The ACD/Spectrus Portal is being developed in Java and designed to be browser independent.

ACD/Labs software technologies, v12 and older, have been migrated to the ACD/Spectrus platform. The platform supports the design of custom forms, dialogs and scripts that enable sample submission projects. User-assisted automation or full automation of data capture, processing, reporting and storage for mass spectrometry screening, structure verification and other workflows are also possible.

ACD/Labs

Advanced Chemistry Development, Inc.
Tel: (416) 368-3435
Fax: (416) 368-5596
Toll Free: 1-800-304-3988
Email: info@acdlabs.com
www.acdlabs.com

Data and Metadata Capture

Key supported objectives for effective data science include:

- Handling a variety of different vendor or generic file formats for chromatographic (LC, GC, LC/UV, GC/MS, LC/MS^n), spectroscopic (UV/vis, IR, 1D or 2D NMR) and other (TGA, DSC) data types.
- Import of metadata

Key objectives for groups and departments include:

- Capabilities to handle sample submission and configure and track samples through a particular workflow
- Freedom to select fit-to-purpose instruments as a result of the vendor-neutral platform
- Capability for automated conversion of proprietary vendor formats into a standardized, normalized, structured format in a homogeneous environment.

Results, Storage and Access

Effective visualization enables evaluation of selected sets of stored information. Human-computer interfaces (HCI), therefore, are vital aspects of a ULI platform, to scientists and organization leaders.

- The data and knowledge gained via interpretations are stored in a digital, standardized, normalized, accessible and easily reusable structured format of the ACD/Spectrus platform database
- Chemical structure (in whole or in part), mixture components and formulation recipes may also be stored with associated data and results
- Template-based reports may be generated and stored within Microsoft Office™ documents directly, available for viewing or interactive editing
- Novel ACD/Spectrus Portal technology for web browser based inspection of data from databased information

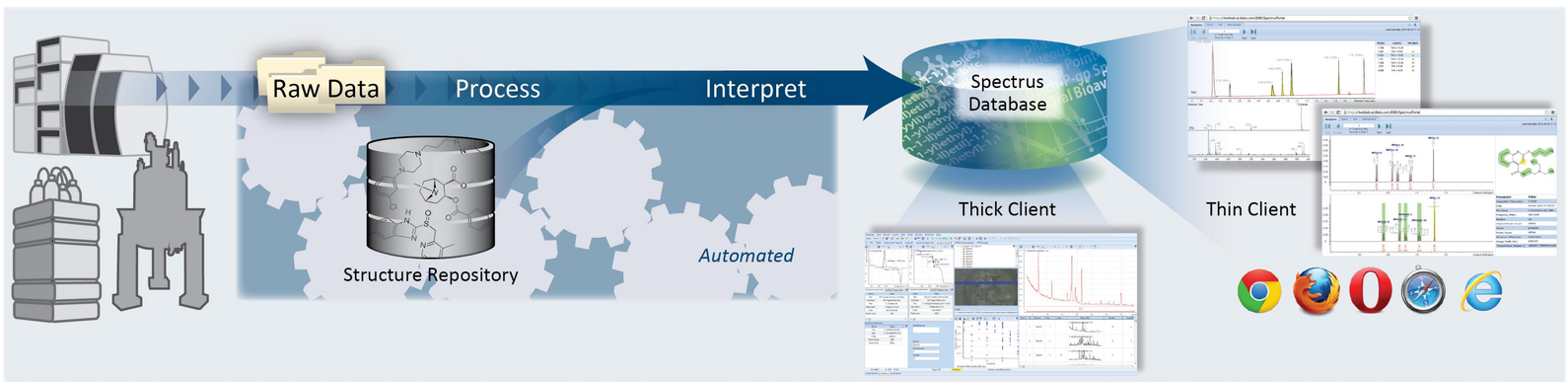

Figure 2—Knowledge management workflow from data capture to access based on a ULI platform.

Analysis, Algorithms and Automation

Consistency of analysis establishes a level of quality in data processing and interpretation, and thus also of extracted or generated results. These are used to make scientific assessments, draw conclusions, make decisions, plan future experiments, and guide business actions. A variety of sophisticated algorithms are available within the Spectrus platform to provide that sought-for consistency. Examples include:

- Targeted and non-targeted screening algorithms for LC/MS [1,2] and for GC/MS workflows have been reported
- John Stafford and co-workers at Lilly reported [3] an approach to impurity analyses using ACD/AutoChrom in chromatographic method development
- Novel clustering approach for spectral data based on similarity [4]
- Custom reporting workflows and templates

Automated Structure Verification (ASV) by NMR [5,6]

The ASV workflow (Fig. 3) begins with fully automated multi-spectrum assignment

- Takes into account chemical shift, multiplicity, integral, and connectivity
- All spectra in project are used to determine appropriate connectivity
- Capabilities exist beyond traditional ^{1}H & ^{1}H-^{13}C HSQC combined verification
- Implemented at Lexicon Pharmaceuticals as part of molecule registry system (see methods).
- Organizations with at least 100 chemists using ASV can recover their investment in under one year.

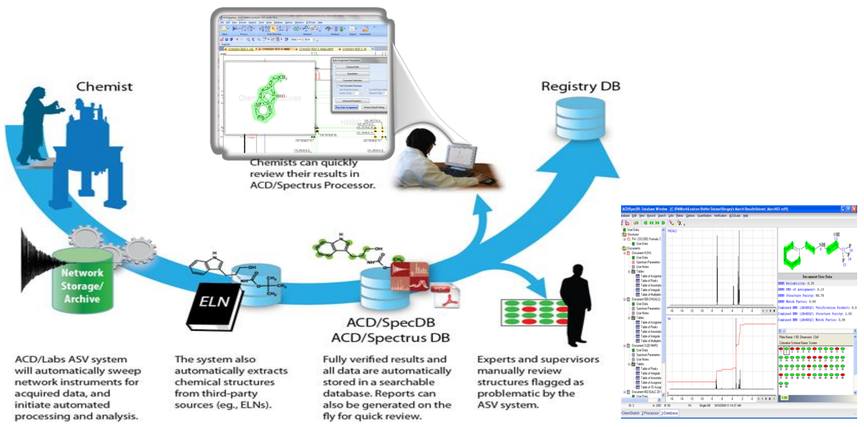

Figure 3—Workflow for NMR Automated Structure Verification.

Conclusions

- ACD/Spectrus provides a software platform on which raw and processed data from different instruments within and across laboratories is consolidated manually or automatically in many workflows
- A platform with an architecture that is amenable to integration and to cloud-based deployment can enable workflows with automated data analysis, interpretation, and knowledge warehousing along with visualization via thick, web, or mobile client interfaces
- ACD/Spectrus solutions deployed at customer sites around the globe feature a wide variety and volume of analytical data in their workflows, and illustrate how ULI technologies address knowledge management needs in chemistry R&D sub-disciplines.

References

1. T. Ito, T. Odake, H. Katoh, Y. Yamaguchi, M. Aoki, J. Nat. Prod. 2011, 74, 983-8. High-throughput profiling of microbial extracts. doi: 10.1021/np100859a.
2. T.R. Croley, K.D.White, J. Wong, J.H. Callahan, S.M. Musser, M. Antler, V. Lashin, G.A. McGibbon, J. Sep. Sci. 2013, 36, 971-9. Combining targeted and nontargeted data analysis for liquid chromatography/ high-resolution mass spectrometric analyses. doi: 10.1002/jssc.201200348.
3. J.D. Stafford, T.D. Maloney, D.P. Myers, J.M. Cintron, B.C. Castle, J Pharm Biomed Anal. 2011 Sep 10;56(2):280-92. A systematic approach to development of liquid chromatographic impurity methods for pharmaceutical analysis. doi: 10.1016/j.jpba.2011.05.028.
4. M. Boruta, SciX 2014. Euclidean Distance Clustering in Spectroscopy, with Applications for Polymorphs, Formulated Products, and Other Areas.
5. S.S. Golotvin, R. Pol, R.R. Sasaki, A. Nikitina, P. Keyes, Magn Reson Chem. 2012, 50, 429-35. Concurrent combined verification: reducing false positives in automated NMR structure verification through the evaluation of multiple challenge control structures. doi: 10.1002/mrc.3818.
6. S.S. Golotvin, E. Vodopianov, R. Pol, B.A. Lefebvre, A.J. Williams, R.D. Rutkowske, T.D. Spitzer, Magn. Reson. Chem. 2007, 45, 803-13. Automated structure verification based on a combination of 1D (1)H NMR and 2D (1)H - (13)C HSQC spectra.