Strategic Directions International, Inc.

INSTRUMENT BUSINESS OUTLOOK

Strategic Information for the Analytical & Life Science Instrument Industry

Standardization Efforts Multiply

In the era of productivity pressure, paperless labs and drug development partnerships, standards are becoming a greater issue for laboratory-generated data. Data standardization facilitates the transfer, analysis and storage of data by creating a common language that can be universally read and written. Standards can be applied to many aspects of data, including format and representation. Three standardization efforts, ASTM Committee E13.15/BSSN Software, the Pistoia Alliance and the Allotrope Foundation, each in different stages of progress, are working to implement standards for biopharmaceutical research. In this article, ***IBO*** examines these standardization projects, their progress and challenges, and adoption potential.

AnIML (Analytical Information Markup Language) is a data-format standard for analytical chemistry. Based on XML, it is designed to capture and store analytical data in a single format from multiple analytical techniques, including HPLC, MS, IR spectroscopy, UV/Vis spectroscopy and NMR. Developed by ASTM Subcommittee E13.15 Analytical Data, in collaboration with BSSN Software, the standard was completed last year. The standard's final ASTM paperwork is being finished. The standard document will be available within the next few weeks, according to Burkhard Schaefer, cofounder and president of BSSN. Funding for the final documentation effort was provided by Agilent Technologies, Waters, PerkinElmer, GlaxoSmithKline and Amgen.

The work on AnIML has spanned 10 years. Technical feasibility was not the main problem, according to Mr. Schaefer. The more serious impediments were validation and consensus. "The big problem is actually twofold: one is to validate that the concept that you come up with can actually capture the types of experiments that are being done in the field," he explained. "And the second one, which is the much larger problem, is getting companies who usually are fierce competitors to actually agree that this is the way you want to do things."

All stakeholders were involved from early on, including software vendors, instrument companies, pharmaceutical firms, and representatives of government and academia. BSSN Software was formed in 2005 to facilitate the implementation of the standard. "We work with vendors who would like to make their products standard compliant. But we also work with end-users who would like to deploy the standard in their organizations," said Mr. Schaefer.

Adoption of the AnIML standard by instrument companies, electronic lab notebook firms and LIMS companies is being driven by major end-users, according to Mr. Schaefer. When working with an end-user to implement AnIML, BSSN will often approach an instrument or software vendor for help in developing an AnIML convertor for a particular instrument or platform. "That know-how that gets created at that point, when that conversion is done for this single customer, then flows back to the vendor so that they can do that themselves, or they can take the component that we developed together and make that available to other customers who have that same problem," explained Mr. Schaefer. Alternatively, BSSN can offer the converter as a third-party component. Currently, one hundred instrument models can generate AnIML data.

There are now 30 deployments of AnIML, according to Mr. Schaefer. The largest adopter is the pharmaceutical industry, followed by environmental laboratories. These industries' range of instrumentation and regulatory requirements has driven their adoption of AnIML. For example, the FDA's data archival requirements favor a standard format such as AnIML, in contrast to PDFs, because it can store raw data and make it accessible regardless of future software availability. As Mr. Schaefer explained, "We see pushes from regulators for archiving the entire raw-data set and being able to recreate the report based on the raw data. If all I keep is my final report, I cannot do that. We are starting to see the first impact of this change."

The Allotrope Foundation may also further the adoption of AnIML. Formed last year as the result of a proposal to the International Consortium for Innovation and Quality in Pharmaceutical Development, Allotrope addresses the "costly and widespread inefficiencies arising from the lack of data standards in the pharmaceutical industry," according to an Allotrope statement to ***IBO***. Allotrope has 11 members, including Pfizer, GlaxoSmithKline, Abbott, Amgen and other major pharmaceutical firms. Rather than create new data standards, Allotrope intends to develop an open framework for managing analytical data. "The goal is not to 'reinvent the wheel' but to adopt existing or emerging standards,

and facilitate their adoption through software that instantiates them," stated Allotrope. The framework's components will include metadata dictionaries, data standards and class libraries, or reusable code. One of the first standards that Allotrope is evaluating for use as part of the framework is AnIML.

Describing the framework's progress since Allotrope's launch meeting last month, Allotrope told ***IBO*** that it has formed project teams for Proof of Concept (POC) Delivery, Development/Architecture, Document Standards, Test Data Set and Metadata. "We anticipate that by the end of Q1 2014, the first POC applications will be deployed at Allotrope member companies for testing." Last month, Allotrope selected Osthus to develop and implement the framework.

Although Allotrope seeks to address the pharmaceutical industry's requirements for standardized lab data, those needs are applicable to any industry that uses analytical instruments, according to Allotrope. "For such industries, the high-level requirements are to enable the seamless movements of analytical data through its entire life cycle and to enable scientists to: find the data needed in seconds, regardless of who, what, when, where, why or how the data were generated; click and create complex technical reports or large sections of regulatory submissions in seconds; click and share data with colleagues across their organization; click and compare their data with data obtained from their partner organization; and run laboratory systems smart enough to prevent errors before they occur."

Another requirement that standardized lab data can facilitate is the ability to work with multiple data types stored in different proprietary file formats, especially metadata. "Compounding this complexity, critical metadata necessary to fully document the laboratory experiment is captured incompletely or not at all, inaccurately due to manual free-text entry, or is spread across multiple software applications from different software suppliers, often with incompatible interfaces," said Allotrope in its statement.

The pharmaceutical industry's data necessities are especially complex due to regulatory requirements and long development cycles. Data-format standardization can address this complexity as well as lower costs for preparing other documents such as stability reports, certificates of analysis and regulatory submissions. Such complexity and costs affect drug company partnerships and companies' work with CROs and contract manufacturing organizations.

Alltrope has provided software and instrument vendors with periodic updates on its progress. It will now have a wider engagement with them and other stakeholders. Next year, it will introduce a partner program. "We are building an ecosystem that will allow software companies and instrument vendors to build on the framework, which will create additional opportunities for them to develop new products for even broader markets."

An example of a standard recently released is HELM (Hierarchical Editing Language for Macromolecules), a notational standard representing the structure of complex biological molecules. HELM facilitates the registration, storage, analysis and visualization of biomolecules and more easily allows researchers to share data. This month, the Pistoia Alliance released a HELM software toolkit and a macromolecule editor for drawing and notating HELM.

Officially launched in 2009, the Pistoia Alliance is a nonprofit organization dedicated to the identification and development of best practices, technologies and processes for addressing inefficiencies in R&D workflows. Members include pharmaceutical and biotechnology companies, service organizations, academic groups and software vendors, including ACD/Labs, Accelrys and PerkinElmer. Utilizing an open collaboration process, the Pistoia Alliance helped transfer HELM, which was initially developed and used by Pfizer, to an open source standard. Twenty-four organizations were involved with the HELM project, according to Sergio H. Rotstein, PhD, director of Research Business Technology at Pfizer and Domain Lead for the HELM initiative. "The main challenge for a complex initiative like this one is to keep a team comprised of volunteer members from 24 different organizations engaged and focused on a common goal, and making steady progress."

Describing the process of developing HELM as an open standard, he told ***IBO***, "Once we obtained permission from all the relevant Pfizer legal and corporate organizations, a 'Code Externalization' subgroup was created, consisting of a team of volunteer developers from various project member organizations, which worked together on the cleanup of the code, removing all references to Pfizer proprietary systems, refactoring, etc., ultimately depositing it in the GitHub open-source repository for public consumption."

Software vendors were among those involved in the initiative. "Some of the software vendors donated software developer resources to the code externalization initiative. Others participated in our communications initiatives, helping to lay out the strategy for dissemination of the initiative and the technology," explained Dr. Rotstein. "Others focused primarily on adding HELM support to their systems." He noted that ChemAxon, in particular, played a key role.

The vendors will also be important to HELM's adoption by end-users. Accelrys, ChemAxon, NextMove, Scilligence and Biomax are in various stages of releasing systems that support HELM as an exchange standard, according to Dr. Rotstein. "Additionally, there are several companies, institutes and content providers that have expressed keen interest in adopting the standard and/or the underlying open source technology," said Dr. Rotstein. To promote the standard, Pistoia aims to raise awareness via speaking opportunities and poster sessions. "On the technology front, we continue to develop key enhancements to the technology to lower the barriers to adoption by interested organizations," he explained.