

Requirements and Goals of the Analytical Information Markup Language (AnIML)

Burkhard A. Schäfer

Introduction

Dealing with electronic data is one of the most daunting aspects of today's laboratory environment. The devices used for analytical measurements come from many different sources and rarely produce their data in the same format. Ensuring proper integration, exchange between systems, and long-term readability for analytical data is a challenge.

These issues will be addressed by the Analytical Information Markup Language (AnIML). AnIML will provide a generic data container^[1] that can hold experimental data from any analytical technique, while at the same time permitting the formal delineation of technique-specific constraints for using this data container, i.e., how the data for specific measurement techniques should be captured represented in the data file. This will allow collection of relevant analytical data in a single place, even for complex workflows using multiple, different analytical techniques.

By supporting embedded digital signatures and audit trails, the authenticity of measurement data can be verified at any time, without requiring access to the application that originally created the data. This is important for regulatory reasons, for intellectual property protection, and for long-term data archiving.

AnIML will provide a number of features to facilitate long-term storage. An AnIML data file is completely self-contained, and all AnIML files are in text-based, XML format, making them human-readable.

As AnIML will provide data in a generic container, it is possible to reuse software components to handle experimental data from many techniques. This can save valuable development resources for equipment and software manufacturers. It also reduces the number of applications that need to be introduced, maintained, and validated in a laboratory and facilitates the creation of generic data viewers.

What has been happening?

AnIML is a joint effort of the ASTM E13.15 Subcommittee on Analytical Data and the IUPAC CPEP Subcommittee on Electronic Data Standards (SEDS). It is being created by a group of individuals from industry, academia, and government. Our industrial partners come from both vendor and end-user communities, ensuring a more balanced view.

During the committee meetings, it has proven extremely helpful to create a framework of requirements and rules for the new data standard to help define, as well as focus on, the tasks to be accomplished. These requirements have been laid out in a comprehensive document that is now being made publicly available.

The AnIML Requirements Document is intended to provide anyone interested in AnIML with information on the development goals for the standard, while acting as a guideline for committee development as well as for company and developer deployment of the new standard. Publication of the document is intended to solicit feedback and spark fruitful discussions with the goal of making AnIML a widely accepted and used data standard.

Where does it go from here?

AnIML is on the way to being standardized using established ASTM processes. The committee is working to prepare the necessary documents for balloting. The E13.15 committee is working closely together with IUPAC CPEP SEDS to create and finalize the data dictionaries and structures for the initially supported analytical techniques. These include UV/Vis, infrared, NMR, and mass spectrometry, together with liquid and gas chromatography. More techniques will follow. More information on current proceedings and plans can be found in a recent article^[2].

The AnIML developers are happy to work with other parties and organizations to facilitate the creation of data dictionaries for additional analytical methods and application domains and are also actively soliciting feedback on AnIML. If you have a particular use case in mind where you think AnIML could be helpful, please let them know.

References

[1] B. A. Schäfer, D. Poetz, G.. W. Kramer: Documenting laboratory workflows using the Analytical Information Markup Language, Journal of the Association for Laboratory Automation, Volume 9, 2004, Issue 6, Pages 375-381.

[2] Tony Davies: Herding AnIMLs, Chemistry International, Volume 29, 2007, No. 6, Pages 21-23. Available on the Web at http://www.iupac.org/publications/ci/2007/2906/pp1_animls.html.