An Analytical Information Markup Language (AnIML) Standard for Analytical Chemistry

Burkhard Schaefer1; Mark Bean1; Maren Fiege3; Gary Kramer4
1 GlaxoSmithKline, Collegeville, PA; 2BSSN Software, Mainz, Germany; 3Waters GmbH, Frechen, Germany; 4National Institute of Standards and Technology, Gaithersburg, MD

Origin of (AnIML) Species

AnIML = working together

Role of the ASTM E13.15 Subcommittee
• AnIML created by the ASTM E13.15 Subcommittee on Analytical Data Standards.
• ASTM facilitates creation and maintenance of standards worldwide (and was the original home of the ASMS society).
• ASTM is self-supporting: standard documentation is sold, ensuring longevity.
• Balloting is limited to ASTM members and is consensual.

AnIML House Rules

AnIML Standard
• Schema (.xsd)
  • AnIML Core Schema – used to validate and create AnIML data files
  • AnIML Technique Schema – used to validate and create AnIML Technique Definition Documents (below)
  • XML Documents (.xsd)
• AnIML Technique-Definition Documents
  • One per technique, freely extensible
  • Written by technique domain experts.
  • Constrained by the AnIML Technique Schema
• Documentation (PDFs to be made available by ASTM)
• AnIML Naming and Design Rules
  • Elements and terms used by AnIML standard.
  • Official explanatory documentation
• AnIML Data Files (.anml (but still xml))
  • Syntax (tag elements, attributes, relationships, and data types) constrained by AnIML Core Schema
  • Semantics (terminology) constrained by AnIML Technique-Definition Documents

Binary (AnIMLS came two by two)
JCAMP-DX uses text numbers: "14563.09"
• Scientific precision, human readable

AnIML is Fast
• Summary reports or peak lists
• Complex hyphenated techniques
• 96-well plate LC-UV-MS with ELSD
• Post-processing appended to vendor-neutral data
• E-mail, archive, or database results with metadata

AnIML is Extensible
• Preserves data integrity for regulated environments
• Supports Tabular Data (name-value pairs)
• Tagged values are also permitted

AnIML Remembers the Past
AnIML leans on JCAMP experience and replaces ANDI. AnIML uses IUPAC or other official terminologies wherever possible.

AnIML is Flexible
• Preserves data integrity for regulated environments
• Supports Tabular Data (name-value pairs)
• Tagged values are also permitted

Audit Trails and Digital Signatures
Microtitre Plates and Sample Information

AnIML Hybrids

Combining Families of Techniques
• Sample Alteration (no detection, no data)
  • Sample Separation - Chromatography Chemical Reactions
• Detection (spectra and chromatograms)
  • UV, IR, MS, NMR, ELSD, CLND, FID, etc.
• Post-Processing (consume old data and produce new)
  • Some change the axis units (transformation)
  • Some retain the axis units (smoothing, baseline subtraction)
  • Some produce tables (peak finding)
  • Some link multiple analyses (quantitation = AnIML 2.0)

Example: LC separation with UV, ELS, MS detection and post-processing (baseline detection, peak finding, smoothing, spectral summation)
• Timelines count from T0 in LC (sample introduction - injection)
• Spectra have parent data point references to T0
• Detection have trace offsets to align peaks
• Spectral and chromatogram data are separate techniques (chromatograms derive from spectra or reduce to point (transformation)
• axes are different

Zoo Keepers

AnIML – A New Home for Old Standards

• JCAMF - http://www.jcamp-dx.org/
• NetCDF - http://www.unidata.ucar.edu/software/netcdf/
• GAML - http://www.gaml.org/
• SpectroML - http://jla.sagepub.com/content/6/6/76.full.pdf
• HUPO - http://www.posdev.info/

Burkhard Schaefer, BSSN Software
Mark Fiege, Waters
Mark Bean, GSK
Dale O’Neill, Agilent
Jamie McQuay, Scimatic Software
Stuart Chalk, UNF
Todd Miller, Waters
Michael Leiber, Agilent
Adam Patkin, Perkin Elmer
Joe Koury, ASTM
Tony Davies, IFAC
Alex Mulh, Shimadzu
Yui Kalambat, Amersham
Bill Browett, GSK
David Martinsen, ACS
James Duckworth, Thermo
Randall Julian, Indigo Biosystems
Mark Mullins, Agilent
Robert J. Lancashire, UWIMONA
Richard Larsen, Jasco

Burkhard Schaefer, BSSN Software
Mark F. Bean, GSK
Gary Kramer, ASTM

NIST
• Gary Kramer (E.13.15 Chair)
• Peter Linskorn

NIST Interns
• Alex Ruhl
• Alex Roth
• Anh Dao Nguyen
• Aykul Avlan
• Dennis Buchan
• Dominik Piltz
• Frank Mazur
• Kordian Placzek
• Marcin Pehlka
• Renny Jopp
• Patrick Schaefer

Questions?
http://www.animl.org
Mark F. Bean@gsk.com
Gary.Kramer@nist.gov

ASMS 2012, Vancouver, BC