



# Analytical Information Markup Language (AnIML) A New XML Standard for Analytical Chemistry



Burkhard Schaefer<sup>2</sup>; Mark Bean<sup>1</sup>; Maren Fiege<sup>3</sup>; Gary Kramer<sup>4</sup>

<sup>1</sup>GlaxoSmithKline, Collegeville, PA; <sup>2</sup>BSSN Software, Mainz, Germany; <sup>3</sup>Waters GmbH, Frechen, Germany; <sup>4</sup>National 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. Limit of one membership per company. Standards must be periodically re-approved to remain active.

See adjacent poster on  
AnIML Mass Spectrometry

## AnIML is Fast



## AnIML is Extensible

## AnIML is Flexible

- 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 Remembers the Past

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



## 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 (.atdd)
  - **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 (.animl (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 twos)

JCAMP-DX uses text numbers: "14563.09"

- Scientific precision, human readable

### AnIML uses base64binary (7-bit text encoding of binary)

```
<EncodedValueSet startOffset="0" endOffset="2846">
```

```
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
ABnZmY4AAAAAAAAAYDgAAAAAMZ0Z0AAAAADNzLW4AAAAAACDkAAAMpnpOAAAAAAE4  
AAAAAGdm5jCAAAAAAZAAAGdmNjOAAAAAMpkrOgAAAAADIG6AAAAAAATDoAAAAAzcmNjOgAAA  
ACamWk6AAAAAJqZ6ToAAAAAZ2aG0gAAAAAzM5I6AAAAAM3MkjoAAAAAAzcyNOgAAAAADNzJAGAAA  
AAJqZkToAAAAAmpmPOgAAAAADNzIw6AAAAAJqZgDoAAAAAZ.....
```

```
.....AAAAADNfEQ8AAAAAJ05QzWAAAAAIBCPAAAAADN7E18AAAAADOTQzWAAAAAM9NDPAAAAAZ  
00M8AAAAADQzWAAAAAM5NDPAAAAACaZUI8AAAAAZ6YkPAAAAADN8CM8AAAAAM1MjWAAA  
AAZawhPAAAAADNfB88AAAAAM1sHTwAAAAAMxMcPAAAAADNfBo8AAAAAGeGGTWAAAAAAM1MaPA  
AAAAAZEXs8AAAAAGfmGjwAAAAAMzMPAAAAACaWRk8AAAAAJozGdWAAAAAMWXPAAAAAAZsxg  
8AAAAAJpZGDwAAAAAMrhXpAAAAABrhJ8AAAAAGemEjwAAAAAMrKpAAAAADN7Aw8AAAAAJpZ  
CzwAAAAAZyLpAAAAACb2Qs8AAAAAJf5CjwAAAAAZyYjPAAAAAAQAg8AAAAAM3MBjWAAAAAMl  
KFPAAAAACaWQm8AAAAACgADwAAAAAZc90wAAAAAAQp7AAAAAM2M9TsAAAAAZ6BxWAAA  
ABnZu07AAAAAJZ6zsAAAAA
```

```
</EncodedValueSet>
```

- 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 T<sub>0</sub> in LC (sample introduction - injection)
  - Spectra have parent data point references to T<sub>0</sub>
  - Detectors have trace offsets to align peaks
- Spectral and chromatogram data are separate techniques (chromatograms derive from spectra or reduce to point detectors); axes are different.



### Current Status

- Schema, Naming and Design Rules - **locked**
- Technique Definitions
  - Chromatography - **finished**
  - MS, UV-Vis, Point Detectors, Indexing - **draft**
  - NMR - **ongoing**
- Documentation (official ASTM standard) – **outsourced, funded**
  - **Complete by end-of-year.**
- Official ASTM E13 Standard – **ballot within 12-18 months**

## Zoo Keepers



## AnIML – A New Home for Old Standards

- **JCAMP** - <http://www.jcamp-dx.org/>
- **ANDI** - <http://www.astm.org/Standards/E1947.htm>
- **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.psidev.info/>

Burkhard Schaefer, BSSN Software

Maren 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, IUPAC

Alex Mutin, Shimadzu

Yuri Kalambet, Ampersand.

Bill Browett, GSK

David Martinsen, ACS

James Duckworth, Thermo

Randall Julian, Indigo Biosystems

Mark Mullins, Agilent

Robert J. Lancashire, UWIMONA

Richard Larsen, Jasco

NIST

• Gary Kramer (E.13.15 Chair)

• Peter Linstrom

NIST Interns

• Alex Rühl

• Alex Roth

• Anh Dao Nguyen

• Aykut Arslan

• Dennis Backhaus

• Dominik Pötz

• Frank Masur

• Kordian Placzek

• Martin Peschke

• Ronny Jopp

• Patrick Gleichmann

### Questions ?



<http://www.animl.org>

[Mark.F.Bean@gsk.com](mailto:Mark.F.Bean@gsk.com)

[Gary.Kramer@nist.gov](mailto:Gary.Kramer@nist.gov)