

AnIML From 30,000 Ft

What is AnIML? What can I do with it?

Burkhard A. Schaefer
BSSN Software



bssn software

Overview

- Motivation
- Introduction to the AnIML data standard
- AnIML Use Cases
 - Archiving
 - Generic Software
 - Integration
- Summary



Introduction to the AnIML Data Standard

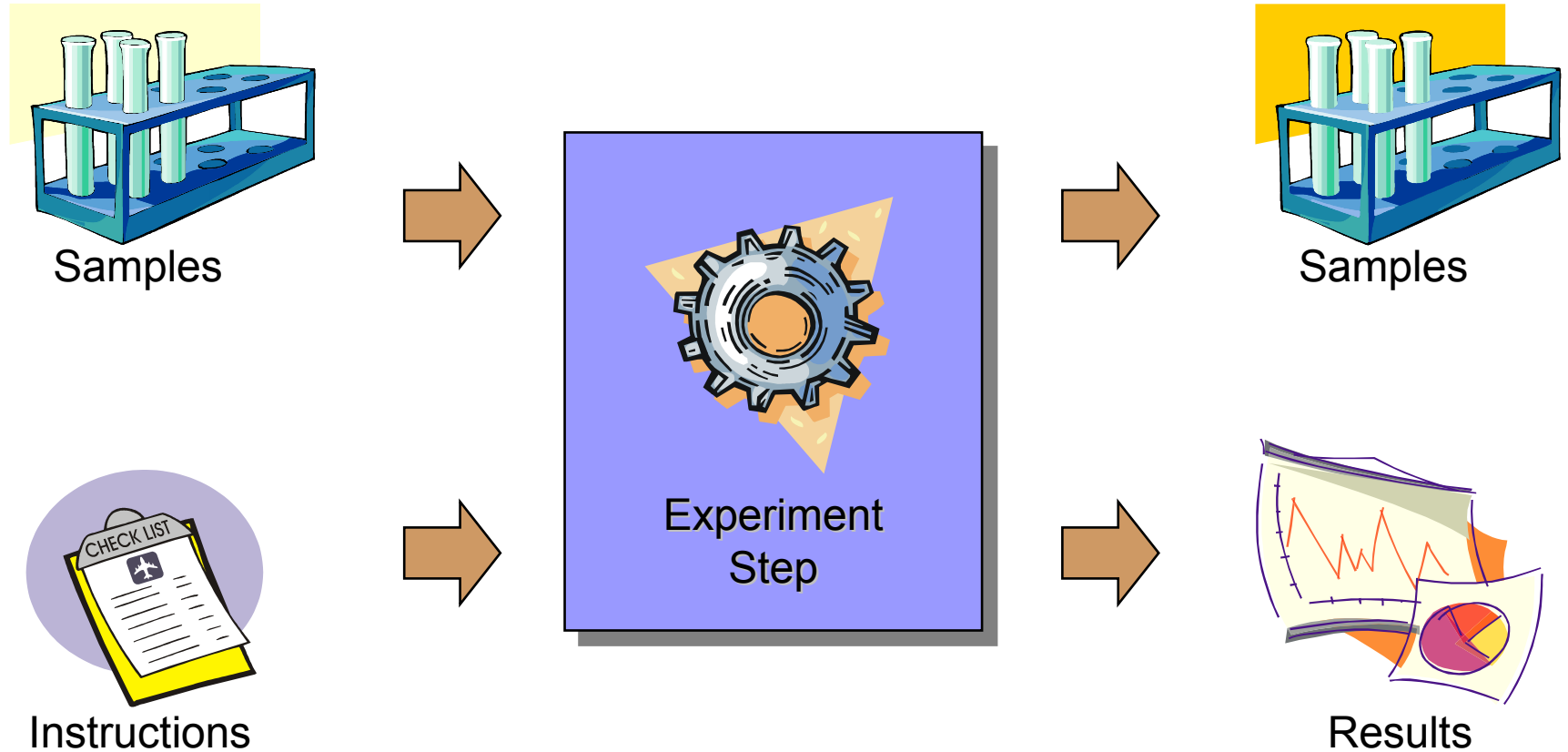
Outline of AnIML

- Analytical Information Markup Language
- Upcoming ASTM XML standard for analytical data, developed by a consortium of industry, academia, vendors, and government bodies
- **Core** provides generic data container to represent “arbitrary” analytical data
- **Technique Definitions** explain how to use the Core for specific techniques

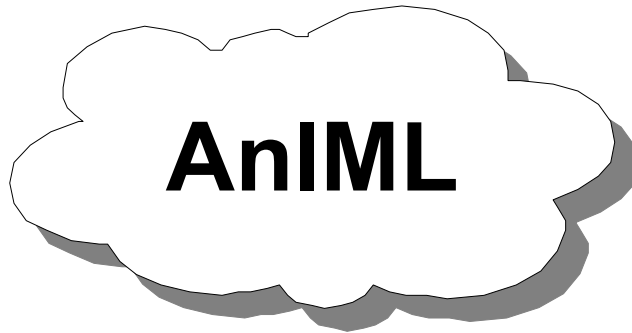
AnIML Design Goals

- Ability to capture data from multiple analytical techniques, possibly combined
- Capturing of sample and process data
- Room for extension by vendors and end users
- Audit trails, digital signatures, and validation for regulatory compliance

AnIML Experiment Model



Possible Roles of AnIML



- Interactive viewing
- Instrument integration
 - Commodity techniques
 - One-offs and custom hardware
- Data management
- Data analysis
- Long-term archiving
- Result delivery
- Publishing, “open access”
- Regulatory submissions



Optimizing Long-Term Data Archival

bssn software

Long-Term Data Archival

- Many challenges to manage:
 - Repository and physical storage
 - Data format
 - Conversion tools
 - Needed as long as a particular source format is used
 - Software for data access
 - Needs to be deployed – and upgraded “forever”
 - Validation
 - Needed initially and with every upgrade
 - ...

Data Formats

- Every allowed data format constitutes a liability
- Software to read the format needs to be preserved as long as the format is current
- Long tail impacts total cost of ownership
- Why not PDF?
 - Stores text and image output, not the underlying data
 - Can not be post-processed

AnIML as an Archival Format

- Open specification
- Captures all required data across techniques:
 - Sample, method and workflow description
 - Instrument and operator
 - Results
 - Identifiers, tracking information (batch, lot, time, ...) to relate information back to original context
- Human readable, even without software tools
- Verbose

Software for Data Access

- “Less is more”
- One tool for any instrument and technique
 - Proprietary software not required
 - Less upgrades & associated re-validation
- Lightweight tool with restricted set of features
 - Low validation effort
- Combination of AnIML and unified archiving tools can reduce total cost of ownership

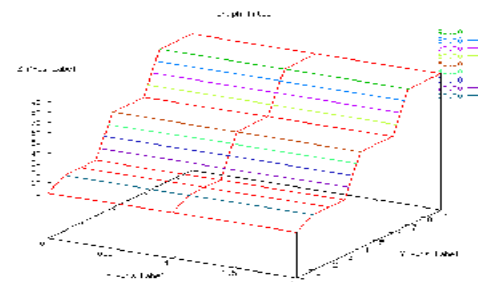


Leveraging Generic Software Applications

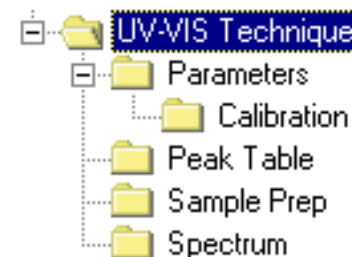
bssn software

AnIML and Generic Software

- Many things can be done without understanding the underlying technique:
 - Displaying, reporting
 - Searching and querying
 - Validation and compliance
 - Digital signatures and audit trails
 - Long-term archival

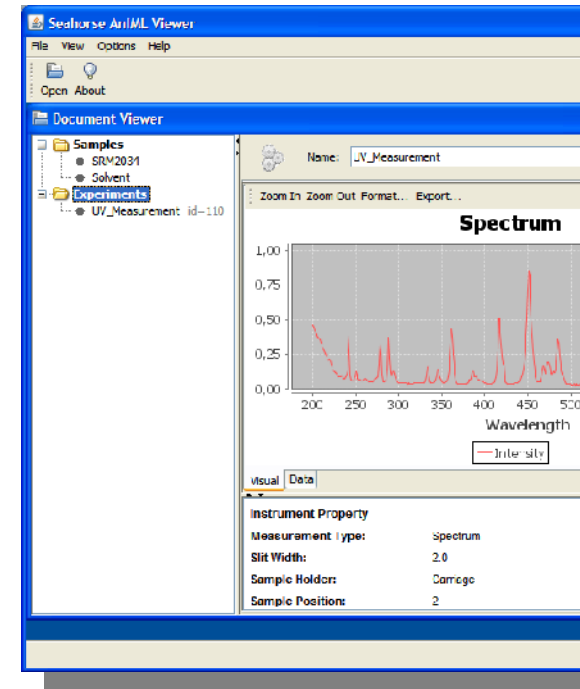


Temperature = 5 °C



Benefits of using generic software components

- Proprietary instrument software is not required to view the data
- Only a compact set of software tools need to be deployed to end user PCs
- Viewing archived data requires only a single tool
- Software can be reused as new techniques are adopted



Seahorse Scientific Workbench

The screenshot displays the Seahorse Scientific Workbench software interface, which is used for managing laboratory workflows and data analysis. The main window is titled "Seahorse" and features a menu bar (File, Edit, Window, Help) and a toolbar with icons for Open, Repository, Save, Import, Techniques, Report, and About.

The central "Experiment Workflow" window shows a process flow starting with "Raw Sample" (ID: raw) and "Solvent" (ID: solvent). These inputs feed into a "Dilute" step (ID: prep1), which produces a "Diluted Sample" (ID: diltec). This sample then goes through a "Filter" step (ID: prep2). The output of the filter is used for a "Measurement" (ID: measure) and a "UV_Measurement" (ID: 110).

On the left, the "Samples" and "Experiments" panels provide a hierarchical view of the data. The "UV-Vis" window (ID: 110) displays instrument properties:

Instrument Property	
Measurement Type:	Spectrum
Slit Width:	2.0 nm
Sample Holder:	Carnage
Sample Position:	2
Scan Speed:	266.75 nm/s
Integration Period:	1.0 s
Spectral Bandwidth Minimum:	0.1 nm
Spectral Bandwidth Maximum:	3.0 nm

The "Wavelength Range" is set from 200.0 to 700.0 nm. The "Measurement Description" includes:

Measurement Description	
Identifier:	SRM 2334 Holmium Oxide Solution
Sample Path Length:	10.0 mm
Experiment Duration:	66.6 s
Project:	AnML
Number of Scans:	1

The "data.ms.animl" window shows a "Total Ion Current Chromatogram" (TIC) for "TIC: SSTD020NT 5259 DFTPPNT". The x-axis is "Time [min]" (3.5 to 5.0) and the y-axis is "Abundance" (0 to 3,000,000). The "UV Measurement" window shows a "Spectrum" plot of "Intensity [A]" vs "Wavelength [nm]" (175 to 500 nm), with peaks labeled at 361, 416, and 461 nm.

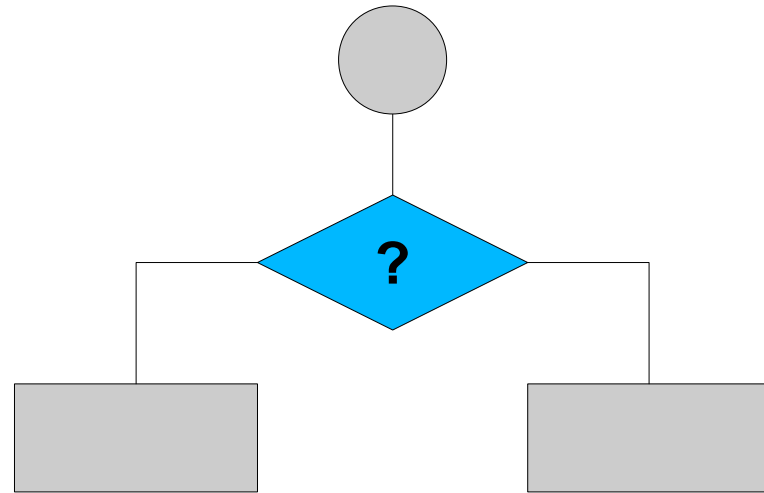
The "Mass Spectrum #409" window shows a "Spectrum" plot of "Abundance" vs "Mass" (20 to 160), with a peak at 152.1334.

Seahorse Scientific Workbench

The screenshot displays the Seahorse Scientific Workbench interface. The 'Microplate Viewer' window shows a data table with columns 1-8 and rows A-P. A heatmap window shows a grid of colored cells (red, purple, blue) corresponding to the data. An 'Image' window shows an X-ray of a knee joint labeled 'LEFT'. A gene diagram window shows the 'Abca12' gene structure with exons and introns, and a list of genes: Ognk1, Kcrlb2, I11r1, Idh1, Abca12, Fnl, Pax3, Bcl2, Cfh, Ccl247, Spina1.

	1	2	3	4	5	6	7	8
A	0.214286	2.948276	5.918719	3.369458	5.231527	2.150246	2.497537	3.812808
B	8.504926	8.807882	9.679803	9.26601	10.05665	8.150246	8.320197	8.541872
C	10.14532	10.293103	10.278325	10.514778	10.810345	9.509852	10.618227	10.027094
D	10.640394	10.396552	10.344828	10.544335	10.396552	9.731527	14.179803	10.049261
E	10.788177	13.293103	10.610837	10.736453	10.522167	10.293103	10.529557	10.522167
F	10.655172	10.374384	10.315271	10.344828	10.13793	10.13793	10.13793	10.13793
G	10.152709	10.758621	10.610837	10.721675	10.721675	10.721675	10.721675	10.721675
H	10.492611	10.455665	10.463054	10.507389	10.278325	10.278325	10.278325	10.278325
I	0.837438	3.280788	9.546798	0.022167	10.22660	10.22660	10.22660	10.22660
J	7.293103	7.972906	11.24630	7.972906	7.972906	7.972906	7.972906	7.972906
K	8.3867	9.103448	10.78078	9.103448	9.103448	9.103448	9.103448	9.103448
L	8.889163	9.576355	10.21182	9.576355	9.576355	9.576355	9.576355	9.576355
M	9.783251	10.093596	10.44827	10.093596	10.093596	10.093596	10.093596	10.093596
N	9.598522	10.049261	10.10098	10.049261	10.049261	10.049261	10.049261	10.049261
O	10.130542	10.359606	10.37438	10.359606	10.359606	10.359606	10.359606	10.359606
P	10.566502	10.05665	10.24876	10.05665	10.05665	10.05665	10.05665	10.05665

Abca12
ATP-binding cassette, sub-family A (ABC1), member 12
protein coding gene



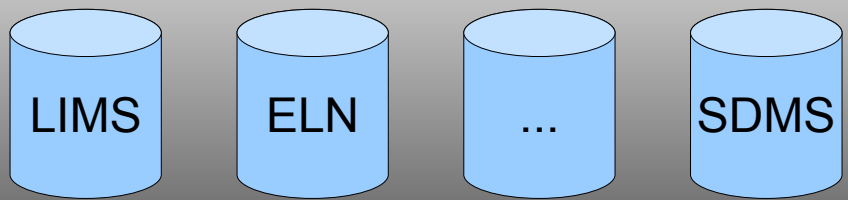
The Vision of Total Integration

Today's Instrument Integration

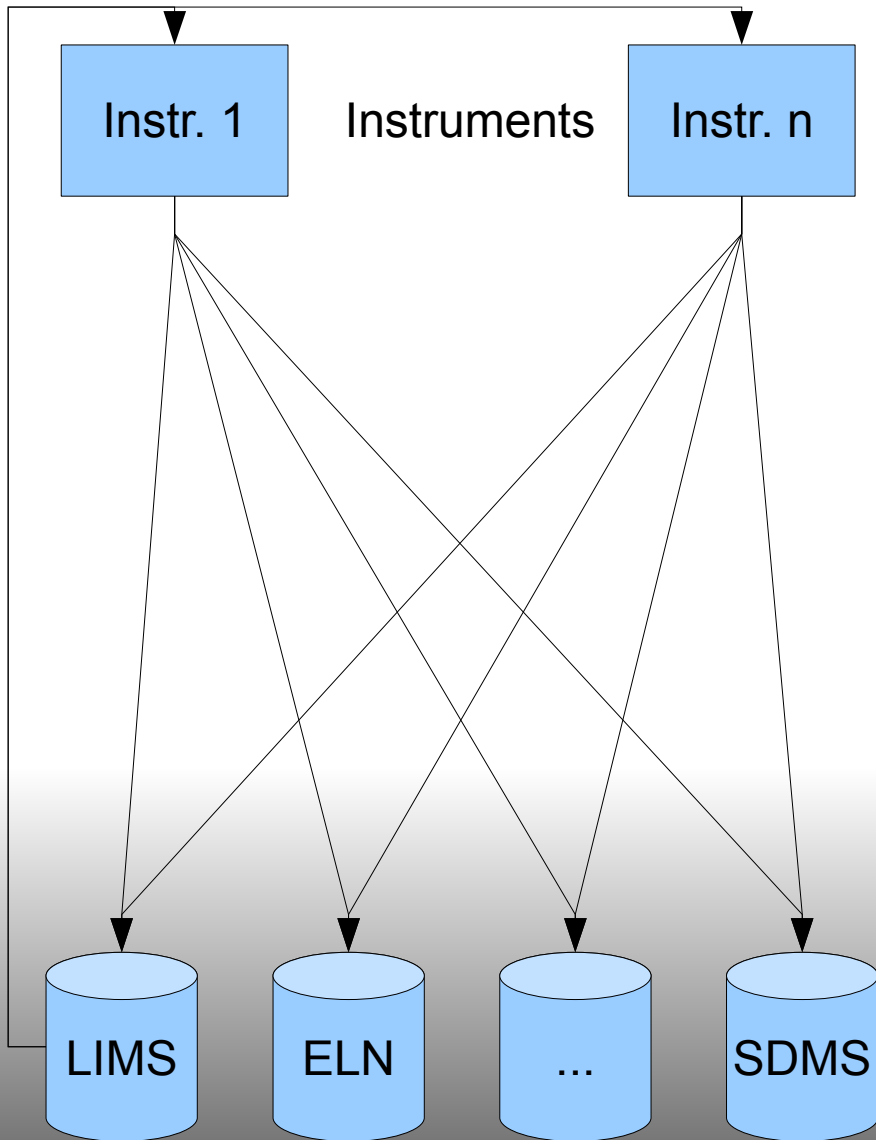


- Lab informatics landscape consists of many data producers and data consumers

Integration ?

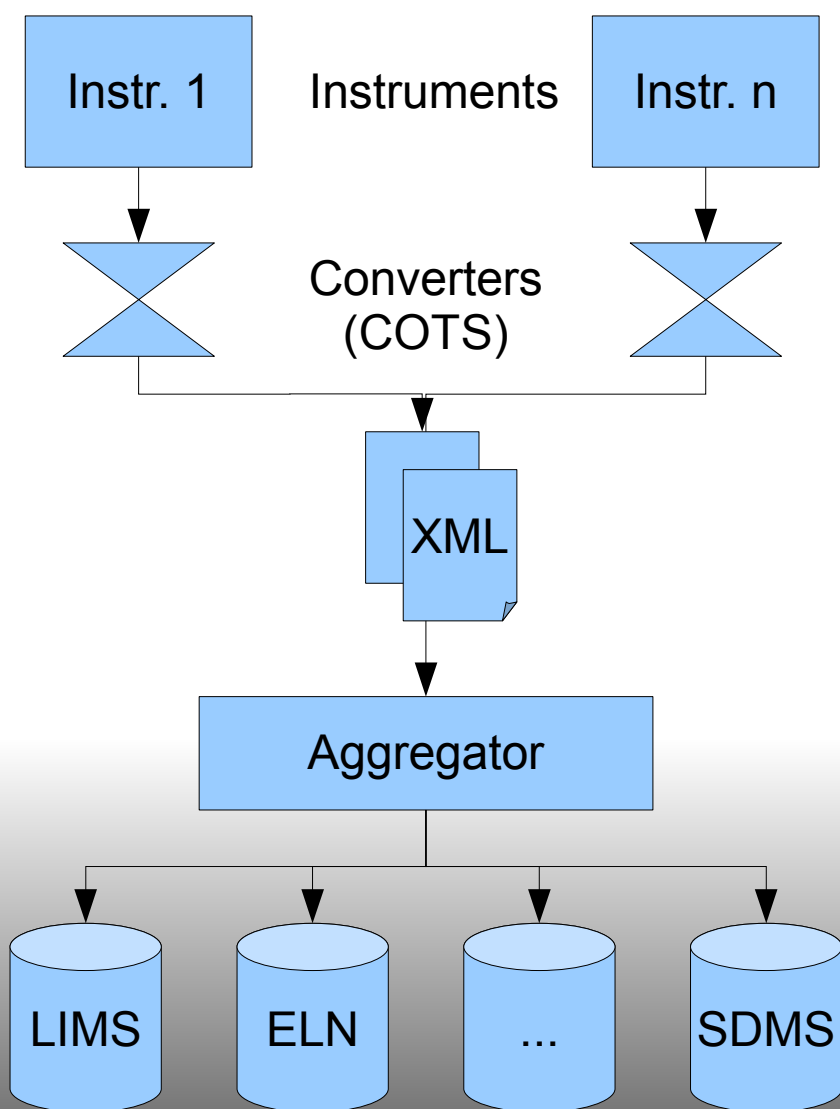


Today's Instrument Integration



- Lab informatics landscape consists of many data producers and data consumers
- Linking instruments and data systems often requires a 1-to-1 integration approach

Standards-Based Integration



- Instrument data is converted into AnIML
- Standard data is aggregated
- Target-specific fields are propagated to data systems
- Low n+1 cost

Summary

- Adopting an open, standards-based approach to data management
 - Enables use of data in new ways
 - Delivers long-term business value
- A combination of AnIML and generic software tools can help us get closer to this goal

AnIML From 30,000 Ft

What is AnIML? What can I do with it?

Live AnIML Showcase
Booth #5155



bssn software