Meeting regulatory requirements in long-term storage and processing of HPLC-MS e-data

Burkhard Schaefer

Director, Head of Core Technologies and Partnering, Merck KGaA
Co-Founder, BSSN Software
SiLA Board Member, ASTM AnIML Task Group Member

BSSN Software from Merck
BSSN Software is now a part of Merck
Key Challenges

How can we take e-data to the next level?
Key Challenges

• Instrument integration
• Data management
• Collaboration with external partners, CROs
• Long-term archival
Instrument Integration

Traditional approach
Traditional Approach

LIMS writes worklist file to a transfer folder

LIMS monitors result folder, waiting for measurement result file

Instrument software monitors transfer folder for new worklist file

Instrument software picks up worklist file and performs measurement

Instrument software writes result file into result folder

LIMS picks up result file and imports content

Challenges:
- Instrument-specific file formats for worklists and result reporting
- Formats are fragile and can change
- Files can be manipulated while in transfer folders
- Unclear if a file is complete
- No robust error reporting
New standards-based approach

• EBF E-Data Working Group Vendor Partnership created a secure XML format for data exchange between LIMS/ELN and instruments

• BSSN Software fully supports this

• Prototype implementation:

![Shimadzu results in BSSN Software viewer](image-url)
Data interoperability

We need the same level of interoperability on the data side
AnIML: Overview

• ASTM XML format for analytical and biological data
• Supports data from multiple analytical techniques, even combined
• Captures of sample and process data, not only results
• Audit trails, digital signatures, and validation for regulatory compliance
• Focus on data accessibility and easy adoption
• Low TCO by design

Frequently presented at EBF in talks, posters, By David Van Bedaf, LabWare, SCIEX, BSSN, ...
Managing Data and Meta Data

Everything lives in context
Everything lives in context

• Instruments have no idea about the context of an experiment
• Context is key to understand the experiment
• Context must be preserved

• Clear metadata strategy is required
Meta Data Sources

Instruments
- Extracted Data

Third-party systems
- LIMS, ELN, study management, SAP, compound reg

Users
- Manual entry, annotation
Sea Star Lab Information Hub
A much more sustainable approach than traditional SDMS
Meta Data

• More than just attributes on files

• Freely definable entities ("lab", "study", "compound", "instrument", "project", "cost center", "target")

• Entities can be navigated and searched

• Dynamic data tree
  o Define your own navigation hierarchy
    ▪ Study > Patient > Sample
    ▪ Product > Production Site > Batch/Lot > Sample > Injection
    ▪ Site > Lab > Instrument > Sample
  o Decoupled from files and folders
CRO Integration

Achieving the same data integrity, completeness and traceability as in-house
CRO Integration

• Distributed workflows across organizational boundaries
• Data collection
• Conversion to AnIML format
• Distributed audit trails
• Cloud-based data transfer
• Data review on partner and sponsor side
• “Inbox” on sponsor side
Cross-Organizational Workflows (CROs)
Standards-Based CRO Integration

• CRO data is (almost) as complete as internal data
• Workflows and data are traceable across organizations
• AnI ML data package includes instrument data
• CRO meta data can become part of internal meta data repository
Digital Archivist

New capabilities for archiving workflows in a digital environment
Digital Archivist Functionality

- Archivist role support
- Moving of data into the archive
- Data access limited to archivist
- Metadata may remain visible to enable navigation & search
- Archivist can place data into “Digital Reading Room” upon request
Digital Reading Room

• Temporary space, holding a collection of documents for review
• Virtual view, no copy, access to audit trail
• Populated by archivist upon request
• Restricted to set of users
• Eyes-only or controllable download
• Read only, notes

• Usage scenarios:
  o Review of data from previous studies, as requested from archivist
  o Inspection and audit situations
Data Analytics

Making our digital assets available for data analysis, machine learning and AI
Feeding the Monster

• Analytics tools, machine learning and AI have been around for a long time
• Why haven’t we used these technologies?
  o Availability of data
• Today, we have the data
• Problem: How to feed tools with this data

“Data !!!”
Data Integration in the Lab of the Future

Instruments, LIMS and other tools publish a holistic picture of the outcome to a global data lake.
F&S Technology Innovation Award 2019 presented to BSSN Software

Category
Global Analytical Instrumentation
Informatics Technology

Award
2019 Global Analytical Instrumentation Informatics Technology Innovation Award

“Frost & Sullivan recognizes that this platform is poised to make phenomenal changes that could disrupt the industry as it breaks through barriers that have restrained laboratory software interoperability and ultimately help users solve the industry’s complex problems through data analysis from all sources”
Summary

Data standards can help us solve e-data challenges through better data accessibility

• Instrument integration
• Data management
• Collaboration with external partners, CROs
• Long-term archival
• Analytics

Community are moving into the right direction. Vendors are ready.