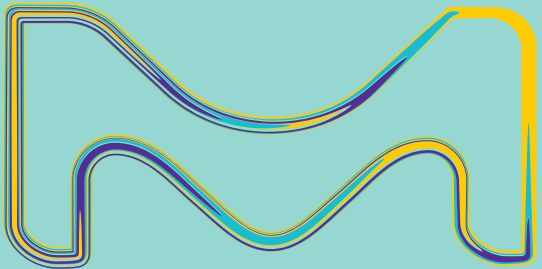


Meeting regulatory requirements in long-term storage and processing of HPLC-MS e-data

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

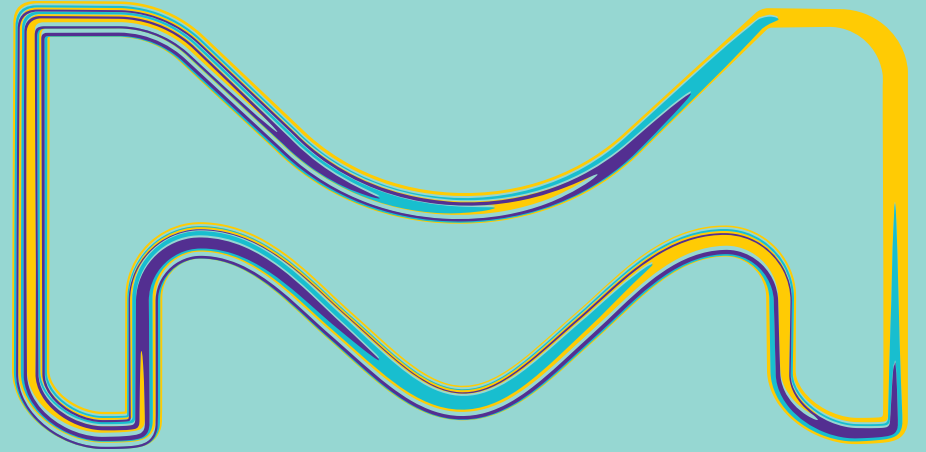


MERCK

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

Data Accessibility

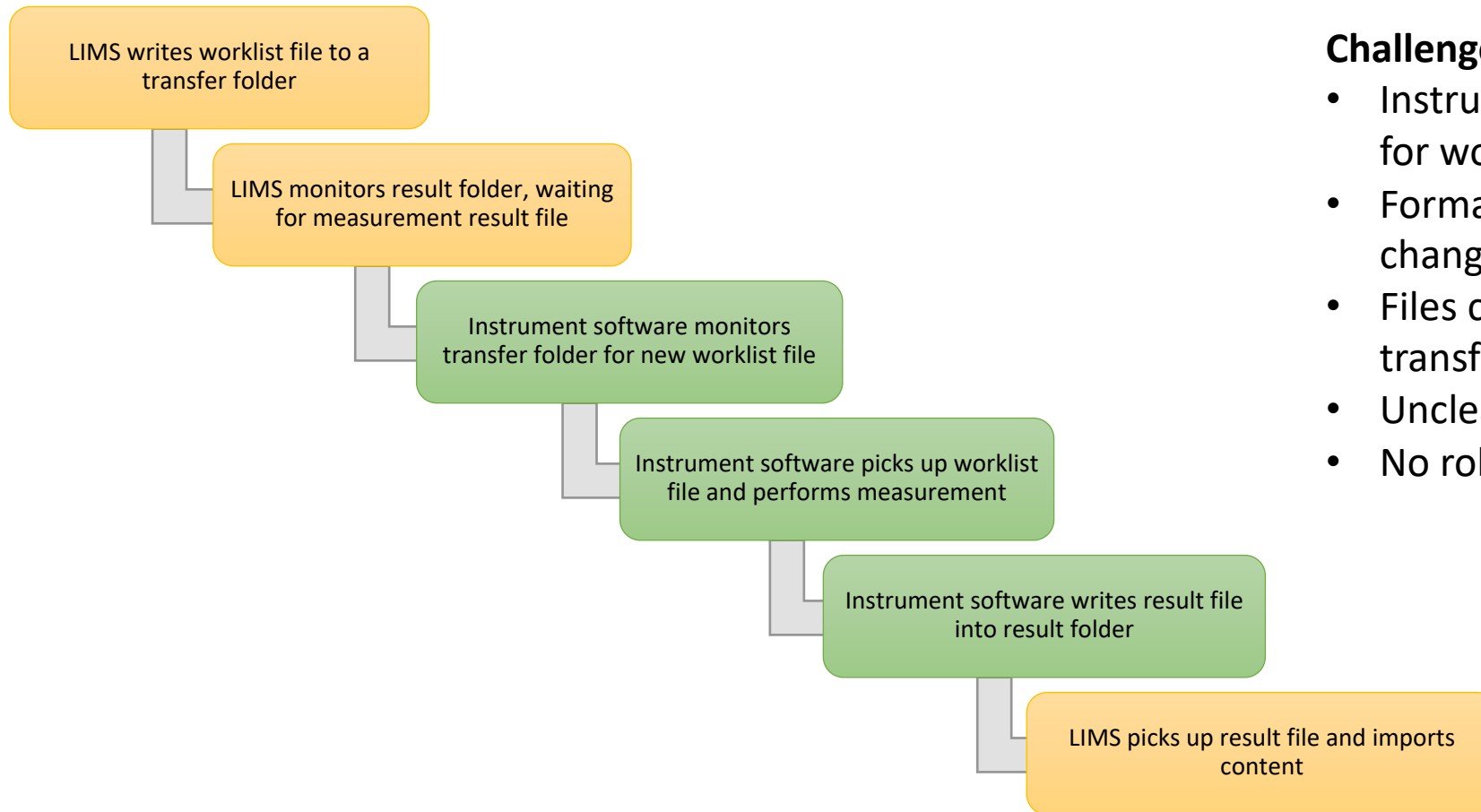
Instrument Integration

Traditional approach

MERCK

BSSN Software

Traditional Approach



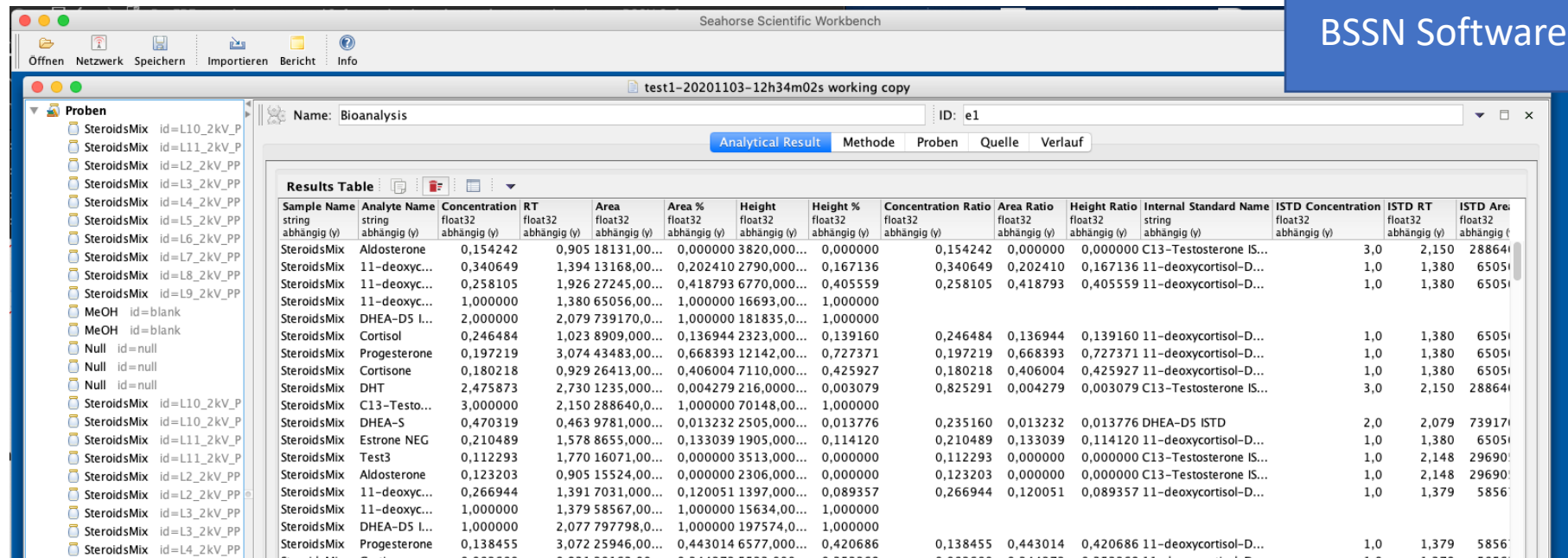
Challenges:

- 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



The screenshot displays the Shimadzu BSSN Software viewer interface. The main window shows a 'Results Table' for a sample named 'test1-20201103-12h34m02s working copy'. The table lists various analytes and their corresponding data points, including concentration, RT, area, and height. The table is organized into columns for Sample Name, Analyte Name, Concentration, RT, Area, Height, and various ratios. The data is presented in a structured format, allowing for easy comparison and analysis of the results.

Sample Name	Analyte Name	Concentration	RT	Area	Area %	Height	Height %	Concentration Ratio	Area Ratio	Height Ratio	Internal Standard Name	ISTD Concentration	ISTD RT	ISTD Area
string abhängig (y)	string abhängig (y)	float32 abhängig (y)	float32 abhängig (y)	float32 abhängig (y)	float32 abhängig (y)	float32 abhängig (y)	float32 abhängig (y)	float32 abhängig (y)	float32 abhängig (y)	float32 abhängig (y)	string abhängig (y)	float32 abhängig (y)	float32 abhängig (y)	float32 abhängig (y)
SteroidsMix	Aldosterone	0,154242	0,905	18131,00...	0,000000	3820,000...	0,000000	0,154242	0,000000	0,000000	C13-Testosterone IS...	3,0	2,150	28864...
SteroidsMix	11-deoxyc...	0,340649	1,394	13168,00...	0,202410	2790,000...	0,167136	0,340649	0,202410	0,167136	11-deoxycortisol-D...	1,0	1,380	6505...
SteroidsMix	11-deoxyc...	0,258105	1,926	27245,00...	0,418793	6770,000...	0,405559	0,258105	0,418793	0,405559	11-deoxycortisol-D...	1,0	1,380	6505...
SteroidsMix	11-deoxyc...	1,000000	1,380	65056,00...	1,000000	16693,00...	1,000000	1,000000	1,000000	1,000000				
SteroidsMix	DHEA-D5 l...	2,000000	2,079	739170,0...	1,000000	181835,0...	1,000000							
SteroidsMix	Cortisol	0,246484	1,023	8909,000...	0,136944	2323,000...	0,139160	0,246484	0,136944	0,139160	11-deoxycortisol-D...	1,0	1,380	6505...
SteroidsMix	Progesterone	0,197219	3,074	43483,00...	0,668393	12142,00...	0,727371	0,197219	0,668393	0,727371	11-deoxycortisol-D...	1,0	1,380	6505...
SteroidsMix	Cortisone	0,180218	0,929	26413,00...	0,406004	7110,000...	0,425927	0,180218	0,406004	0,425927	11-deoxycortisol-D...	1,0	1,380	6505...
SteroidsMix	DHT	2,475873	2,730	1235,000...	0,004279	216,0000...	0,003079	0,825291	0,004279	0,003079	C13-Testosterone IS...	3,0	2,150	28864...
SteroidsMix	C13-Testo...	3,000000	2,150	288640,0...	1,000000	70148,00...	1,000000							
SteroidsMix	DHEA-S	0,470319	0,463	9781,000...	0,013232	2505,000...	0,013776	0,235160	0,013232	0,013776	DHEA-D5 ISTD	2,0	2,079	73917...
SteroidsMix	Estrone NEG	0,210489	1,578	8655,000...	0,133039	1905,000...	0,114120	0,210489	0,133039	0,114120	11-deoxycortisol-D...	1,0	1,380	6505...
SteroidsMix	Test3	0,112293	1,770	16071,00...	0,000000	3513,000...	0,000000	0,112293	0,000000	0,000000	C13-Testosterone IS...	1,0	2,148	29690...
SteroidsMix	Aldosterone	0,123203	0,905	15524,00...	0,000000	2306,000...	0,000000	0,123203	0,000000	0,000000	C13-Testosterone IS...	1,0	2,148	29690...
SteroidsMix	11-deoxyc...	0,266944	1,391	7031,000...	0,120051	1397,000...	0,089357	0,266944	0,120051	0,089357	11-deoxycortisol-D...	1,0	1,379	5856...
SteroidsMix	11-deoxyc...	1,000000	1,379	58567,00...	1,000000	15634,00...	1,000000							
SteroidsMix	DHEA-D5 l...	1,000000	2,077	797798,0...	1,000000	197574,0...	1,000000							
SteroidsMix	Progesterone	0,138455	3,072	25946,00...	0,443014	6577,000...	0,420686	0,138455	0,443014	0,420686	11-deoxycortisol-D...	1,0	1,379	5856...

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

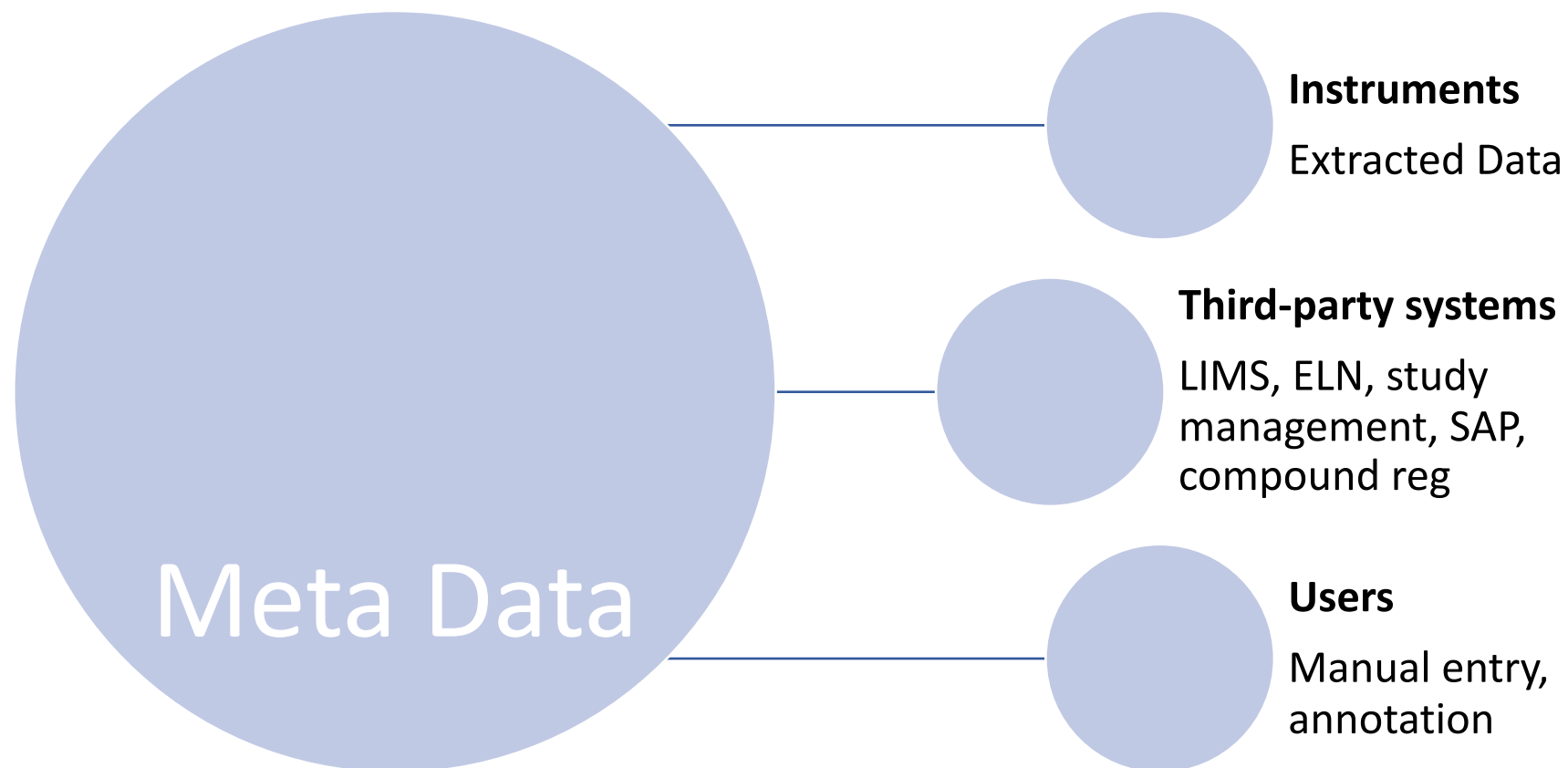
Everything lives in context

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

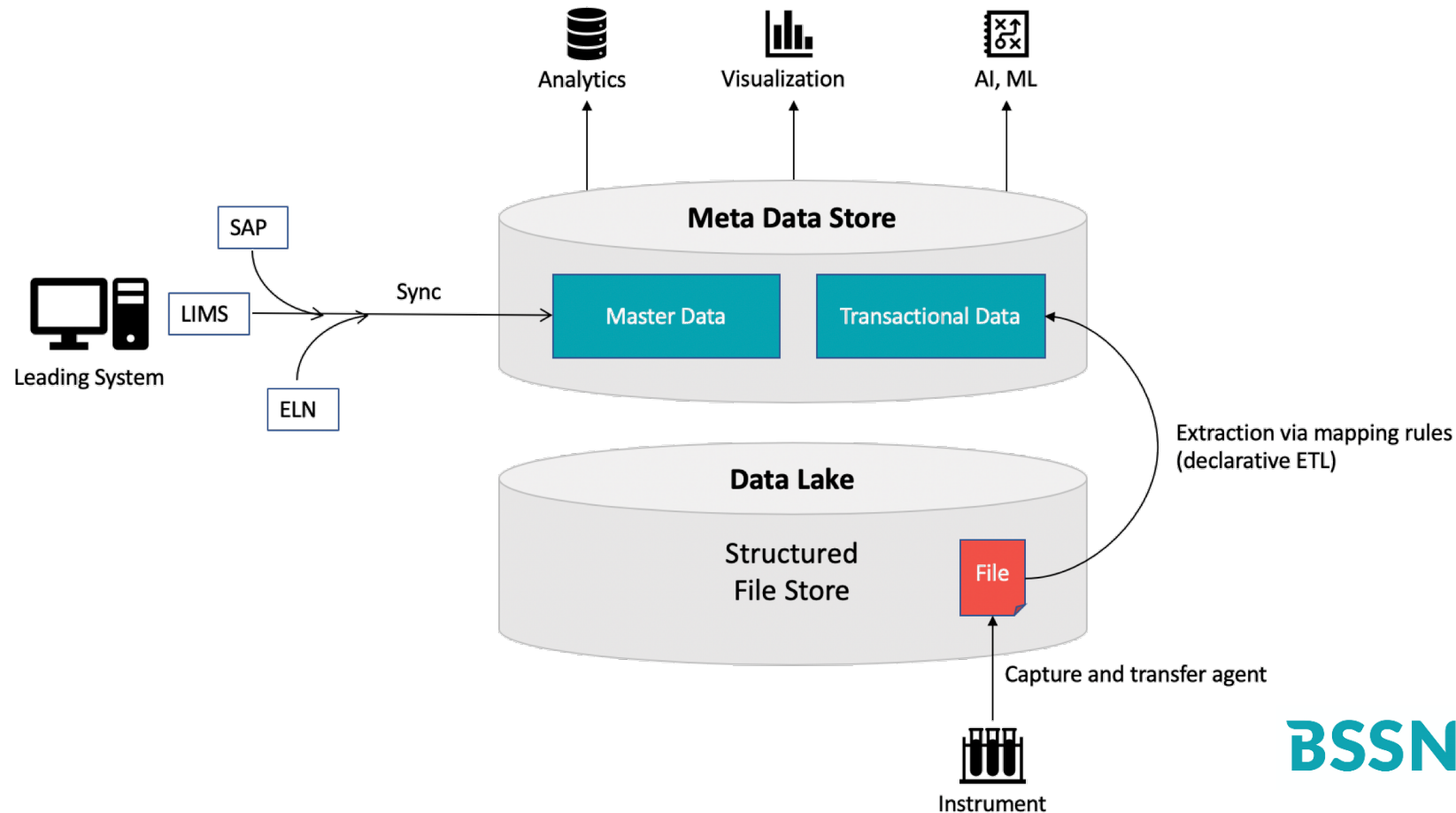
- Clear metadata strategy is required

Meta Data Sources



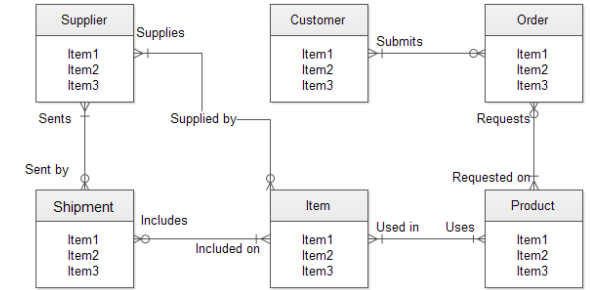
Sea Star Lab Information Hub

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
 - Define your own navigation hierarchy
 - Study > Patient > Sample
 - Product > Production Site > Batch/Lot > Sample > Injection
 - Site > Lab > Instrument > Sample
 - Decoupled from files and folders



CRO Integration

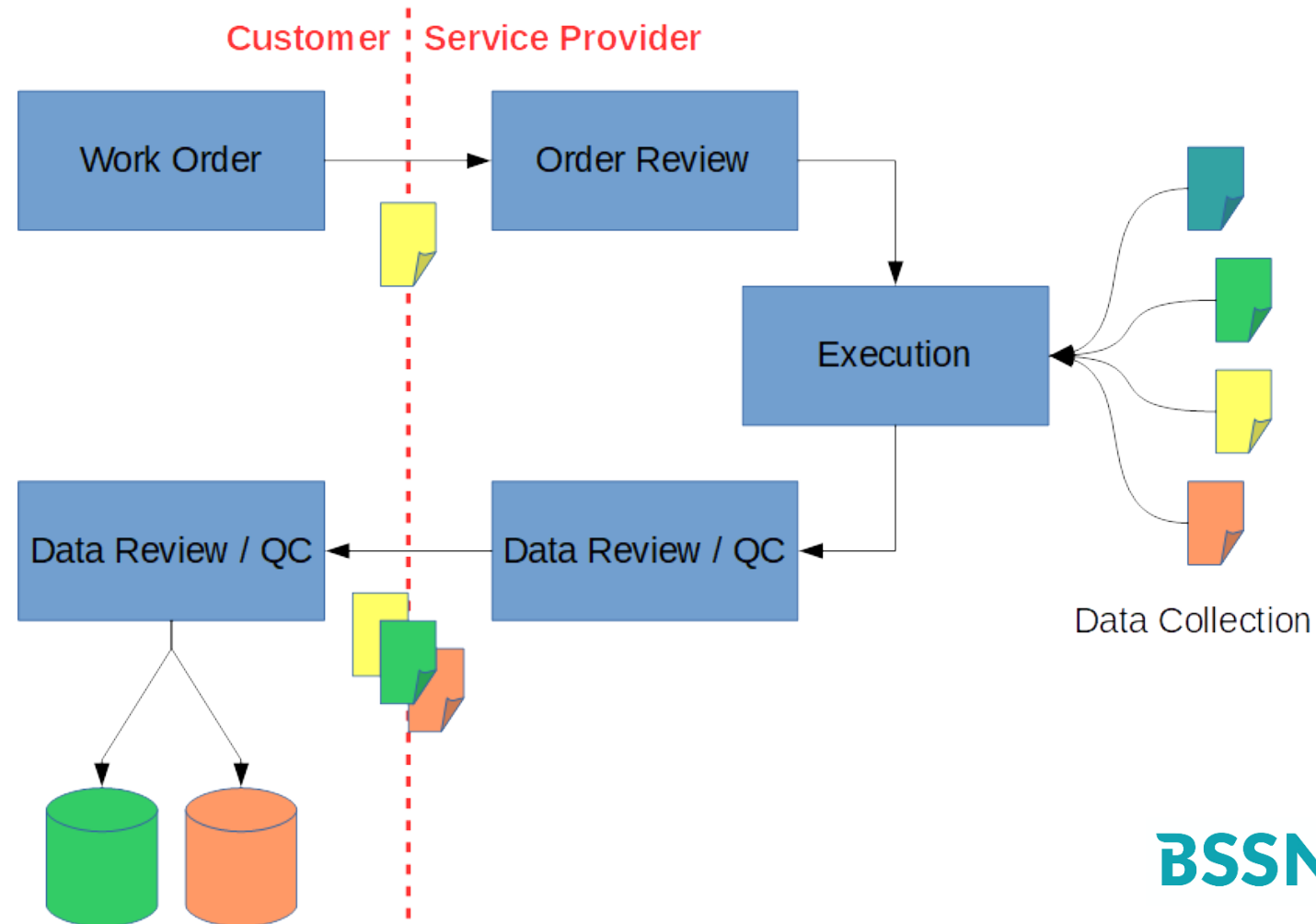
Achieving the same data integrity, completeness and traceability as in-house

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
- AnIML 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:
 - Review of data from previous studies, as requested from archivist
 - Inspection and audit situations



Data Analytics

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

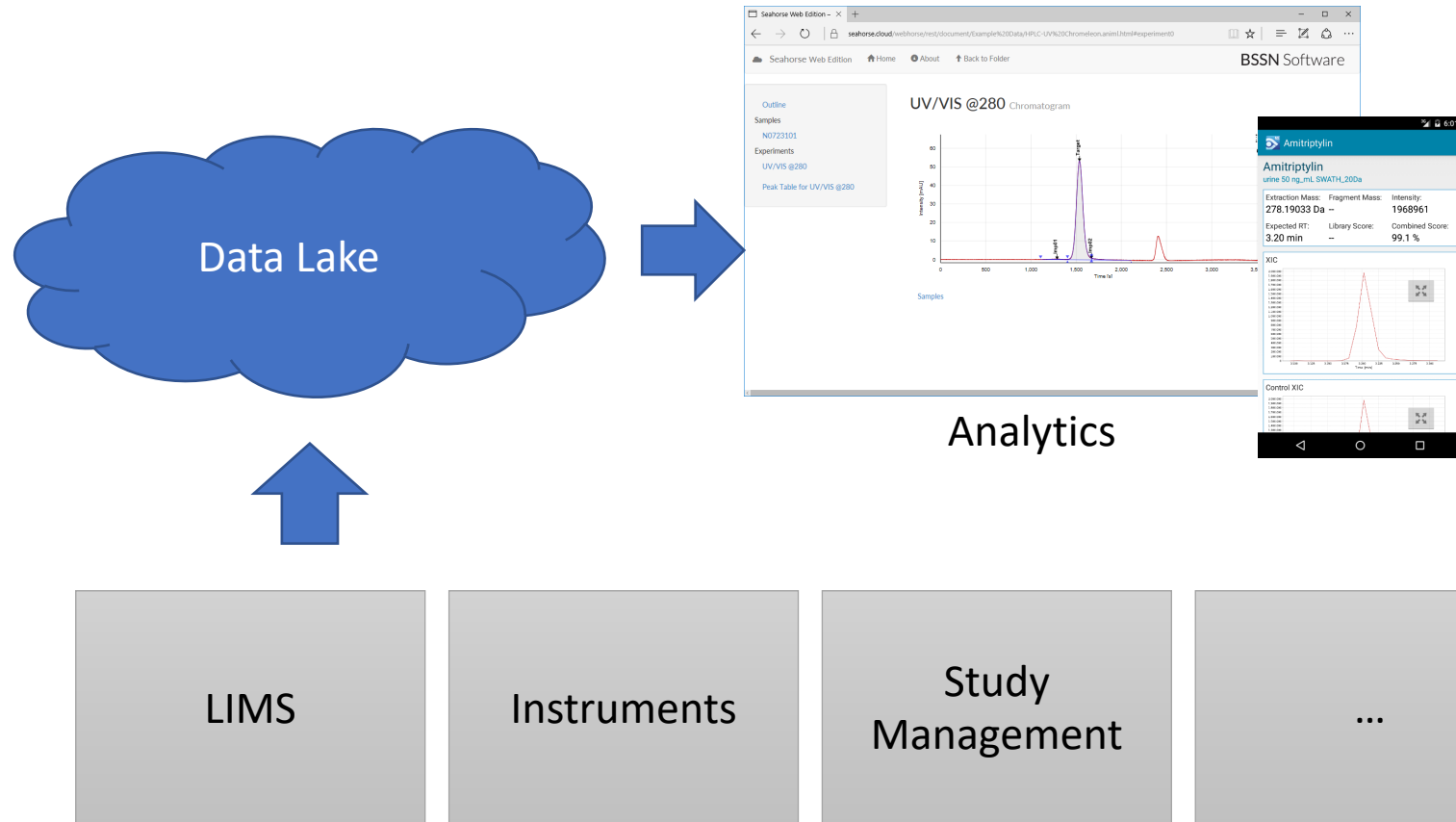
Feeding the Monster

- Analytics tools, machine learning and AI have been around for a long time
- Why haven't we used these technologies?
 - 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"

F R O S T  S U L L I V A N

MERCK



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.

Data
Accessibility

