



Allotrope[®] Connect Day 1: Wednesday, 15-May-2024

Breakouts (60 min)

The role of data standards in the era of generative AI

Readouts (45 min)

How are you using data standards today to achieve your digital goals?

Are data standards more important for the **F**, the **A**, the **I**, or **R** in **FAIR**?

What are the challenges within your organizations today for the implementation of data standards?

Are there areas of data standards that are important for your AI/ML initiatives?

Where do you see opportunities for genAI to accelerate the development and implementation of data standards?



How are you using data standards today to achieve your digital goals?

- Not enough, just started, using business terms taxonomy
- Opportunistic, to deliver on specific need
- Local structuring, not driven top down
- Standardized ontologies applied in new application configurations, foundational to FAIR strategy
- Connection of experiments to Measurements are a priority
- Standards are good for access, but often local structuring is preferred by domain experts
- Change management is hard
- Standardization is currently a sliding scale from local practices to full cross-domain standards

Are data standards more important for the F, the A, the I, or R in FAIR?

- Need to ensure accessibility for GxP solutions for many years
 - All four are important, I and R are more challenging
 - Challenge: end user trusting data (completeness and accuracy)
 - Is FAIR data needed equally for the following use cases:
 - Search
 - Primary analytics
 - Modeling and prediction
 - Generative approaches
- Can be done with less stringent FAIR data
- Requires the most FAIR data

What are the challenges within your organizations today for the implementation of data standards?

- Prioritization from senior sponsorship down to SMEs
- Value often seen down the road; lots of change management
- Band aids keep business continuity, seems more appetite for these
- Tipping point, if you can prove ROI or critical project need across domains that requires standardization
- Need teams of experts with experience to be able to rapidly deliver data standardization solutions

Are there areas of data standards are important for your AI/ML initiatives? Where do you see opportunities for genAI to accelerate the development and implementation of data standards?

- LLM used to analyze historical data to determine if it would meet current guidelines. Difficulty reading some graphics and tables, ignored content without notification
- Extract unstructured data to structure it, without standards these are often localized structures that differ across domains, organizations, etc
- genAI to help increase data standard adoption
- How do we apply competency questions to confirm Gen AI has generated something trustworthy? Where can standards with validated content / schemas help?