Data Curation and Discovery – Making Experiments and Discovery Remotely Accessible and Actionable

Kevin Dressler, Konrad Hilse, Wes Reinhart (2DCC-MIP)

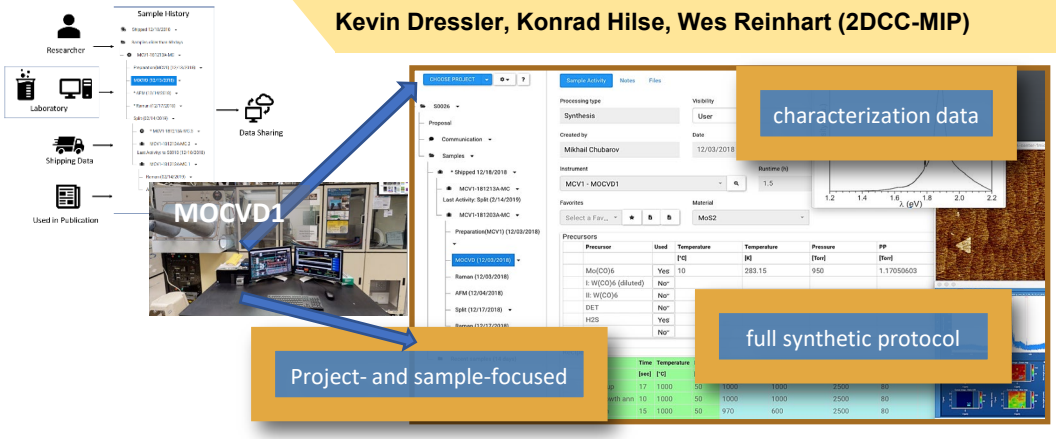
LiST 1.0 (DMR-1539916) is a **traditional database** for data capture and curation. It captures the full history of a sample including growing conditions, characterization, sample quality assessment, sample chain of custody (shipping) and data packages for data used in publications.

- Compatible with staff workflows
- Data automatically uploaded from tools and organized and displayed in the user interface (UI)
- UI enables users to access, search and visualize their data

LiST 1.0: Curation

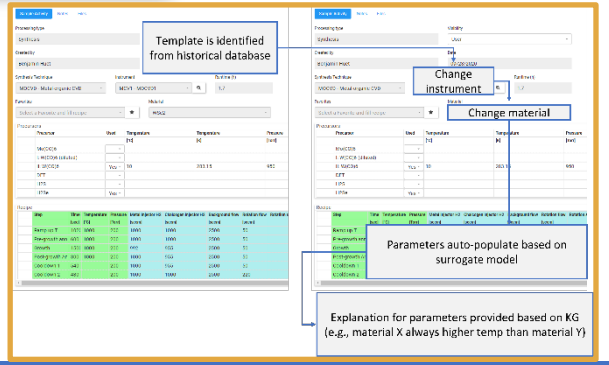
↓

LiST 2.0: Discovery



LiST 2.0 (DMR-2039351) includes a **knowledge graph (KG)** that will contain both a data store and an ontology describing the relationships between samples, tools, characterization data, simulations, etc, in context-aware, graph structured data with machine learning pipelines

- Recommend optimizations of growth conditions
- Accelerate diagnosis of protocols and equipment
- Suggest new experiments



KG discovers relationships → Parameter influence predicted, user fine tunes

LiST 2.0 will provide insights into data previously obscured by complexity, enabling the research community to pursue new modes of discovery