

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

Kevin Dressler, Konrad Hilde, 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.

The screenshot shows the LiST 1.0 interface with several key features highlighted:

- Sample History:** A sidebar menu with categories: Researcher, Laboratory, Shipping Data, and Used in Publication.
- Data Sharing:** A cloud icon indicating data sharing capabilities.
- Project- and sample-focused:** A central view showing a sample named 'MOCVD1' with associated experimental data.
- Characterization data:** A graph showing a plot of current (I) versus voltage (V) with a peak at approximately 1.6V.
- Full synthetic protocol:** A detailed table of experimental parameters including Preursor, Used, Temperature, Pressure, and PP.

Preursor	Used	Temperature [°C]	Pressure [Bar]	PP [Bar]
Mo(O2)6	Yes	10	283.15	950
W(O2)6 (diluted)	No	-	-	-
W(O2)6	No	-	-	-
DET	No	-	-	-
H2S	Yes	-	-	-
	No	-	-	-

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

The screenshot shows the LiST 2.0 interface with several key features highlighted:

- Template is identified from historical database:** A box highlighting the template selection process.
- Change instrument:** A button for changing the instrument used in the experiment.
- Change material:** A button for changing the material used in the experiment.
- Parameters auto-populate based on surrogate model:** A box highlighting the auto-population of parameters.
- Explanation for parameters provided based on KG (e.g., material X always higher temp than material Y):** A box highlighting the explanation for the parameters.

Preursor	Used	Temperature [°C]	Pressure [Bar]	PP [Bar]
Mo(O2)6	Yes	10	283.15	950
W(O2)6 (diluted)	No	-	-	-
W(O2)6	No	-	-	-
DET	No	-	-	-
H2S	Yes	-	-	-
	No	-	-	-

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