

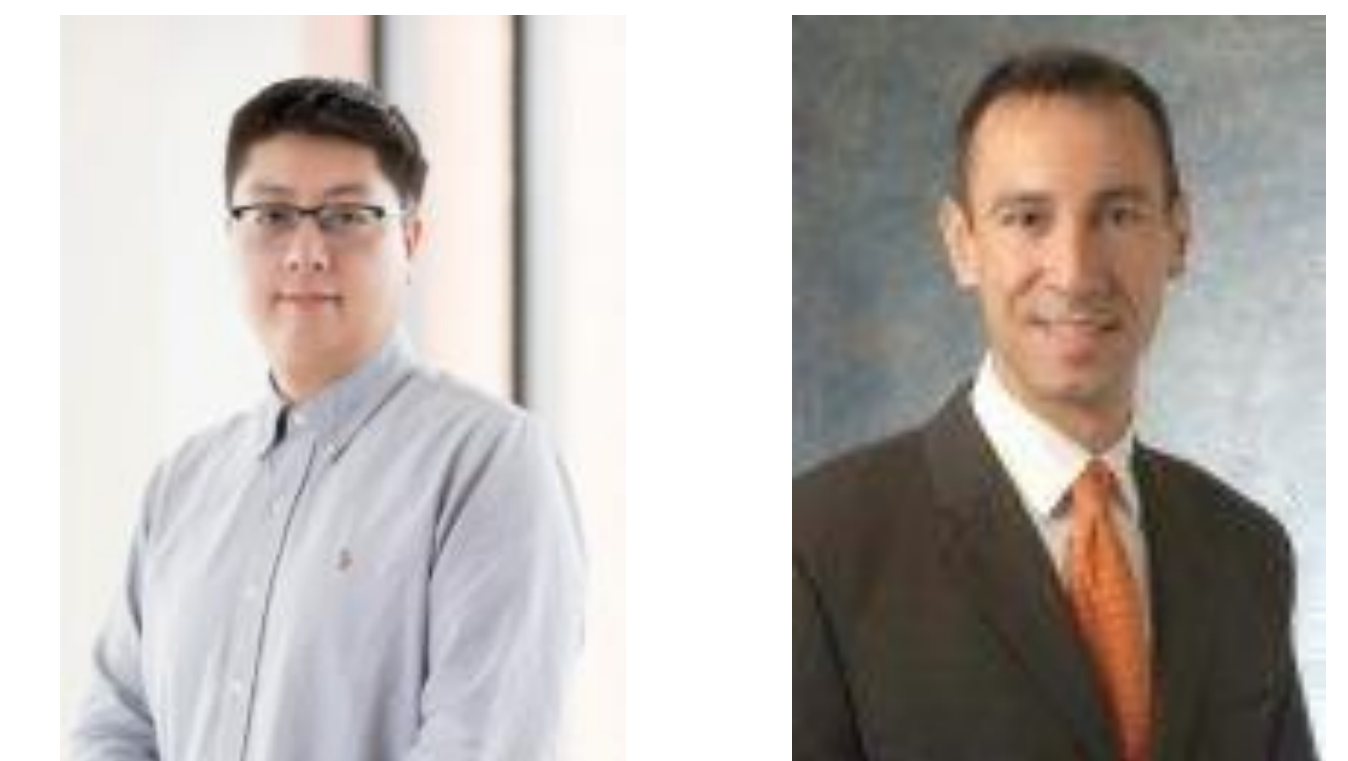
# Elements: Crowdsourced Materials Data Engine for Unpublished XRD Results



PI: Yinghui Wu  
Dpt. of Computer & Data Sciences.

Co-PI: Alp Sehirlioglu  
Dpt. of Materials Sci. and Eng.

OAC-2104007: Project:  
[crux-project.github.io](https://github.com/crux-project)



Case Western Reserve University

## Project Overview

High-value materials data resources such a multiphase X-Ray Diffraction (XRD) remain *unpublished* and never made available for public. In this project, we propose to develop a crowdsourced materials data infrastructure to self-curate and recommend high-value yet unpublished materials (XRD) data.

### Challenges:

- Modeling of raw, isolated XRD data
- Search and access high-value XRD data and the corresponding processing metadata
- Recommend XRD data and resources for scientific workflows

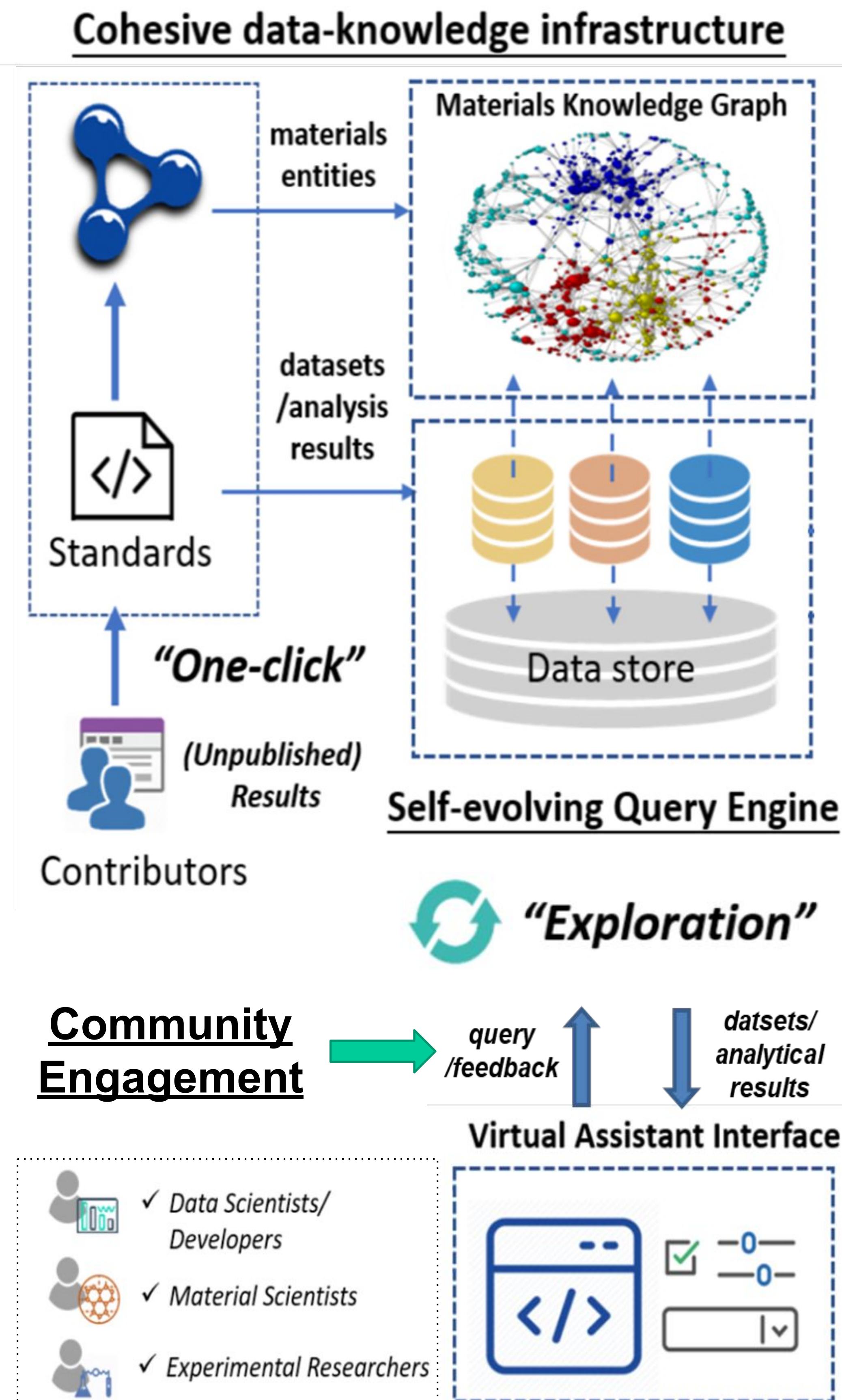
### Proposed Research:

**CRUX:** A **CR**owdsourced data infrastructure, to curate, discover, and recommend **U**npublished **X**RD data and analytical results.

## Research Focus & Scientific Impact

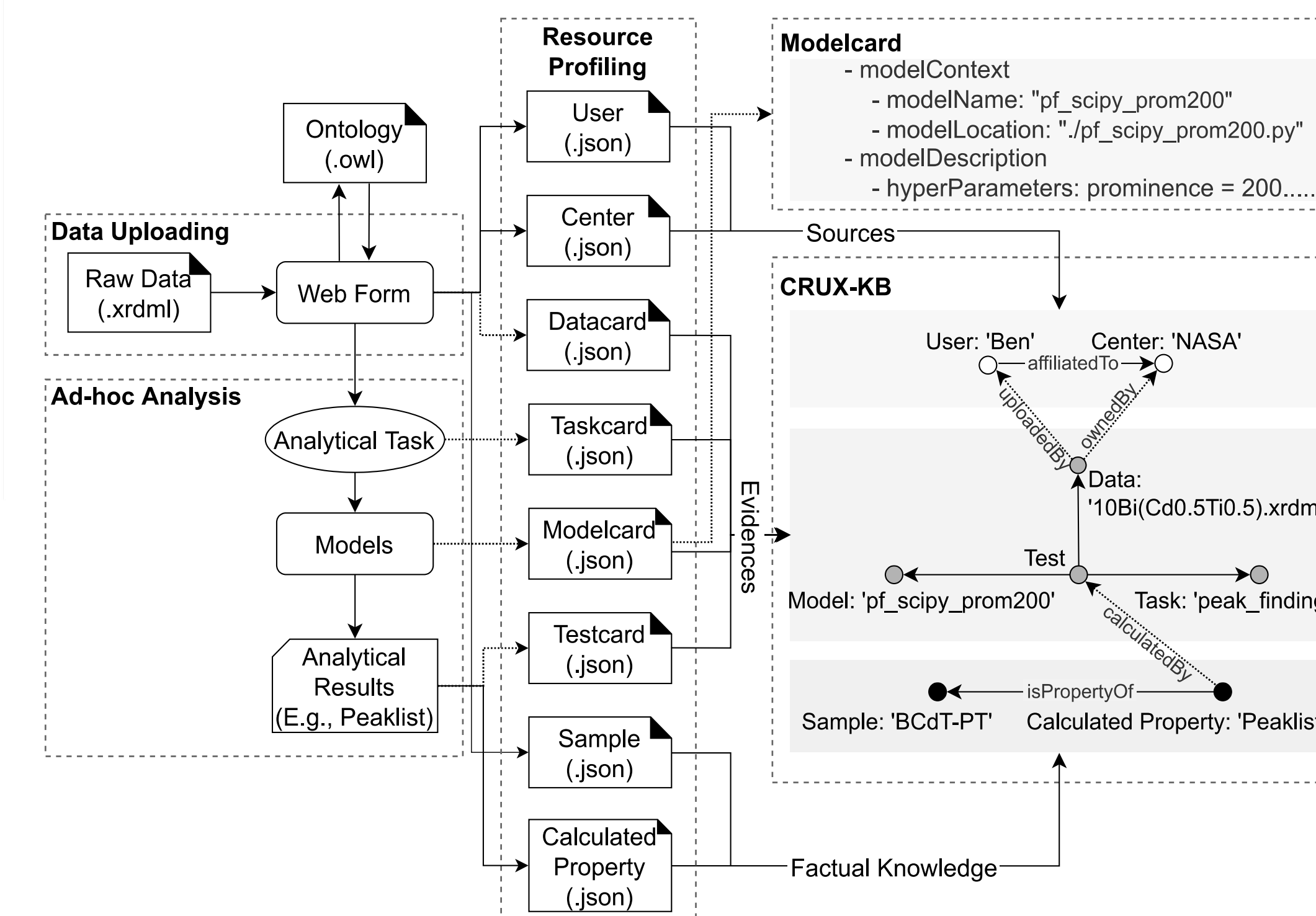
- CRUX-KB:** Three-tier Knowledge Graph Model for coherent data-knowledge representation
- CRUX-IE:** “One-click” data sharing & fusion
- CRUX-Q:** Exploratory query Infrastructure for data search and answer “What”, “Why”, “How” and “Who” questions

## High-level Architecture

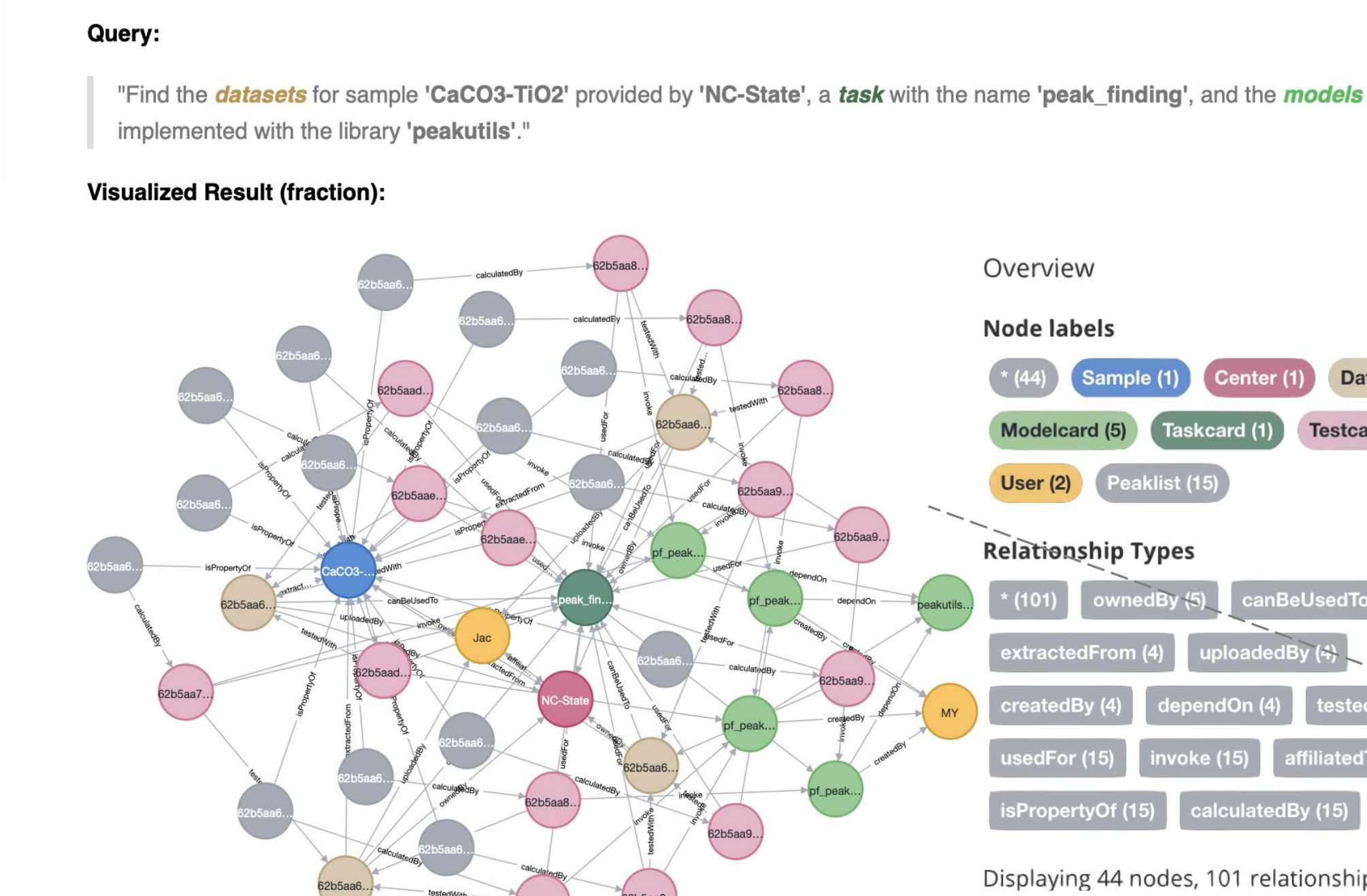


## Technologies

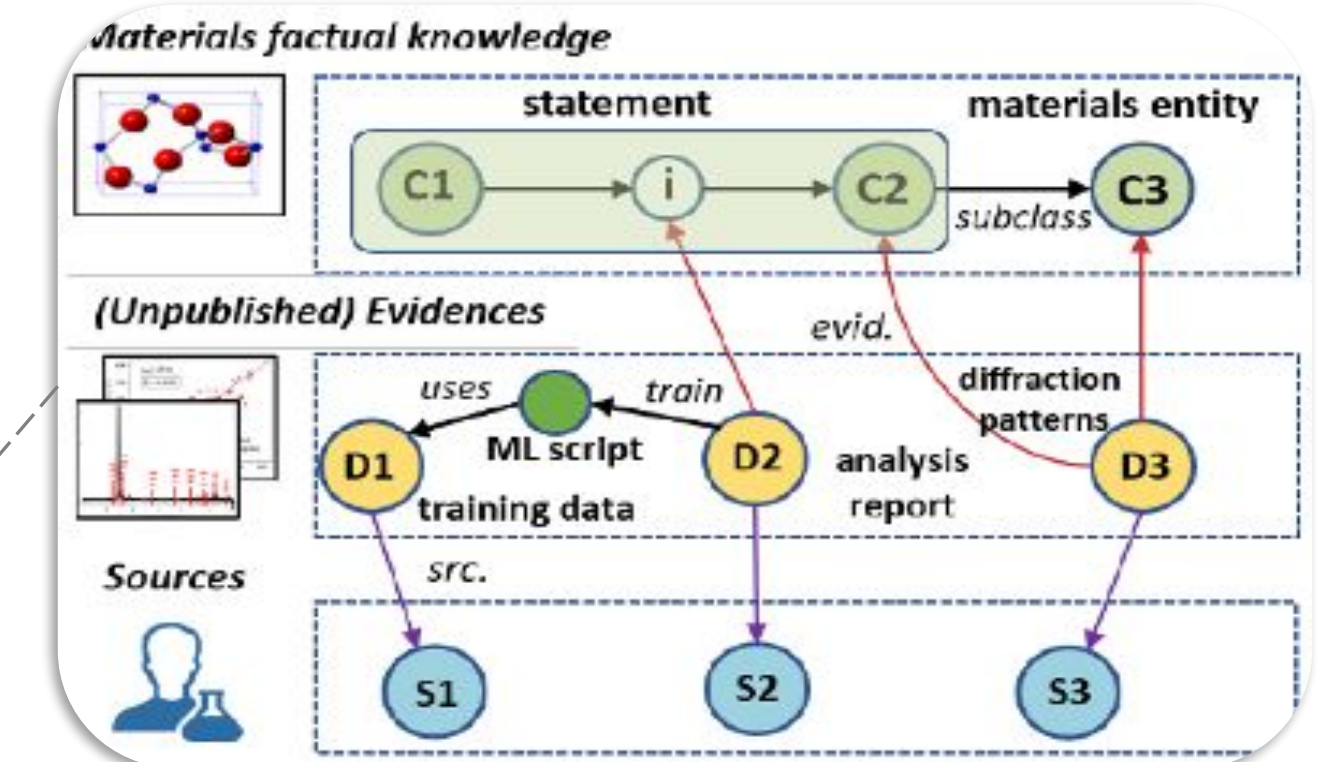
### CRUX Dataflow



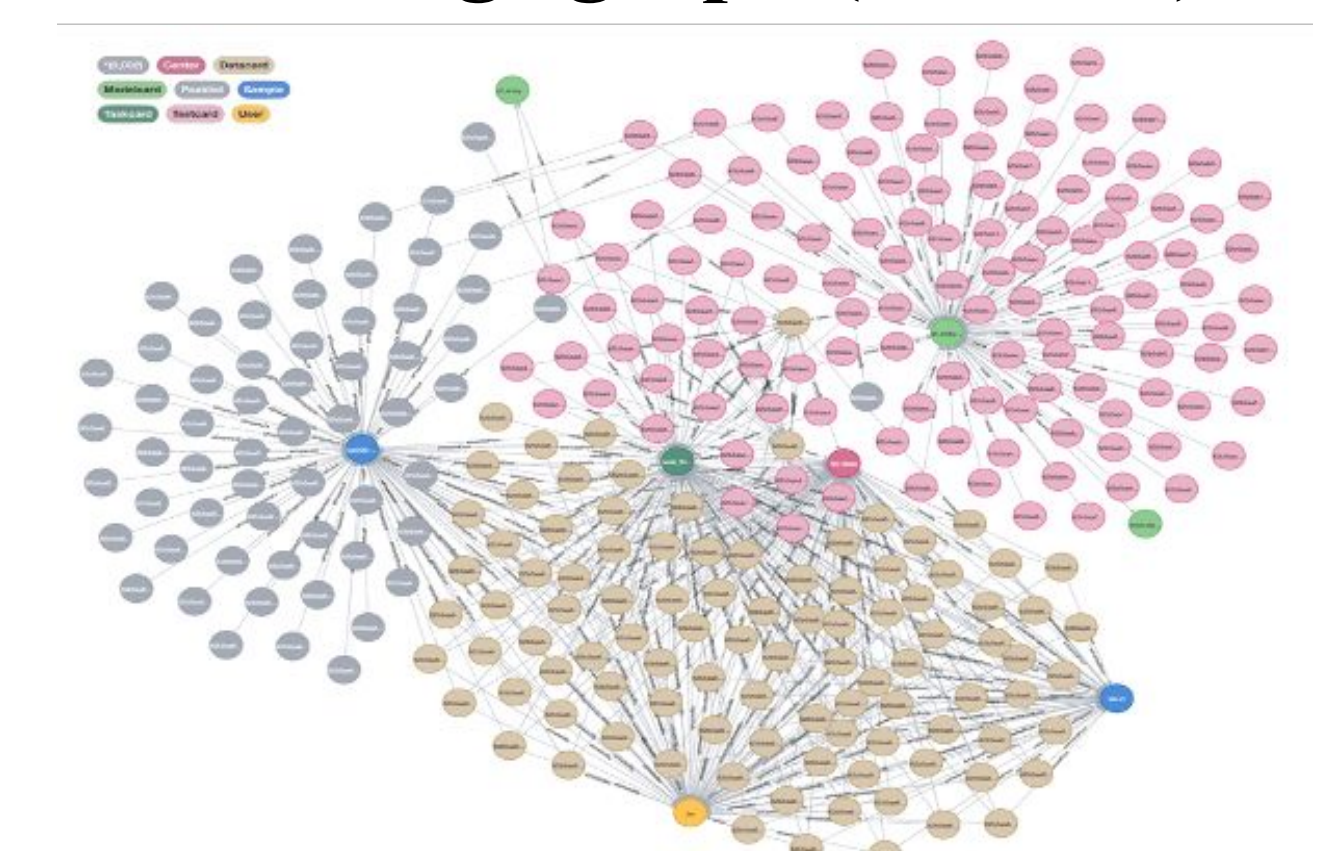
### CRUX Query



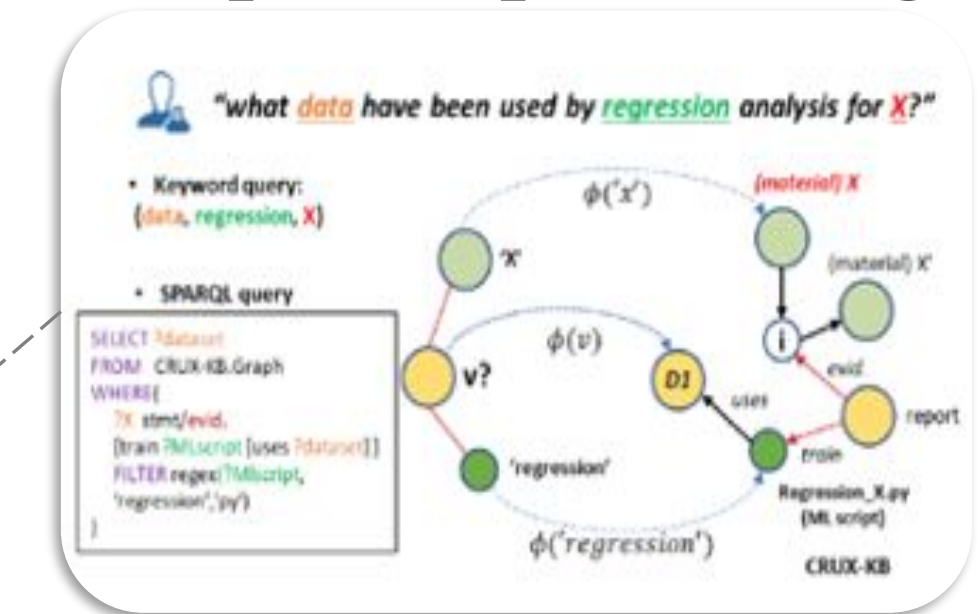
### “Three-tier” Knowledge Graph



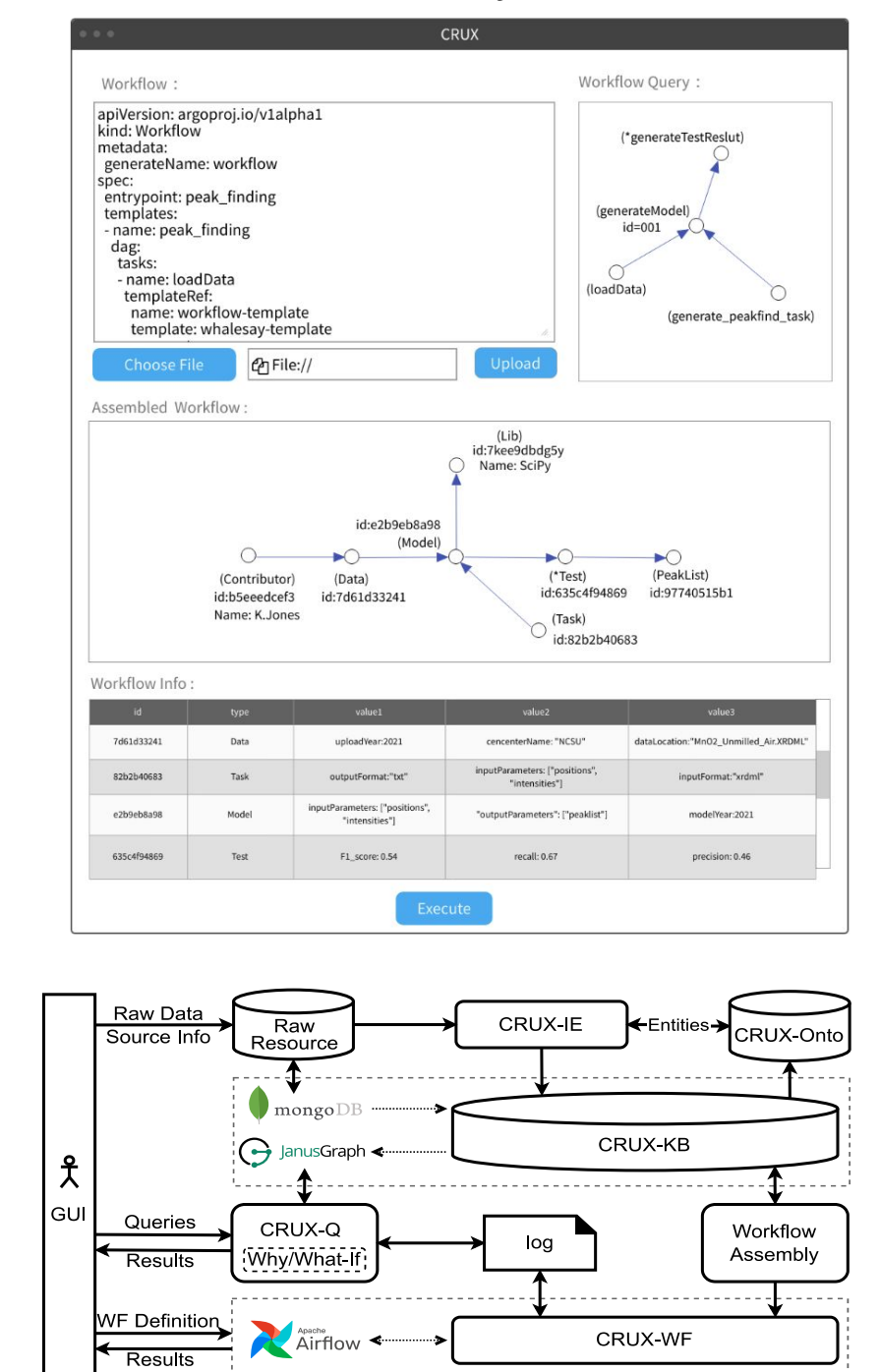
### knowledge graph (fraction)



### Graph Deep Learning



### Demo System



### Data & Query Provenance

## Broader Impact

- Promote collaborative materials research
- Pathways to large ML/AI driven tasks e.g., prediction of crystal structures or solubility limits of different compounds
- Provide a data vault of processing metadata tied to XRD data
- Outreach and promote underrepresented groups
- Curriculum & Tools on “Materials data science”

