T2C2

# NSF DIBBs T2-C2 Project and 4CeeD Tool

Steve Konstanty, Senior Research Programmer
Coordinated Science Laboratory, University of Illinois at Urbana-Champaign

May 2, 2016

A timely and trusted curator and coordinator of scientific data

MRL FREDERICK SEITZ MATERIALS RESEARCH LABORATORY

MNTL micro+nanotechnology lab

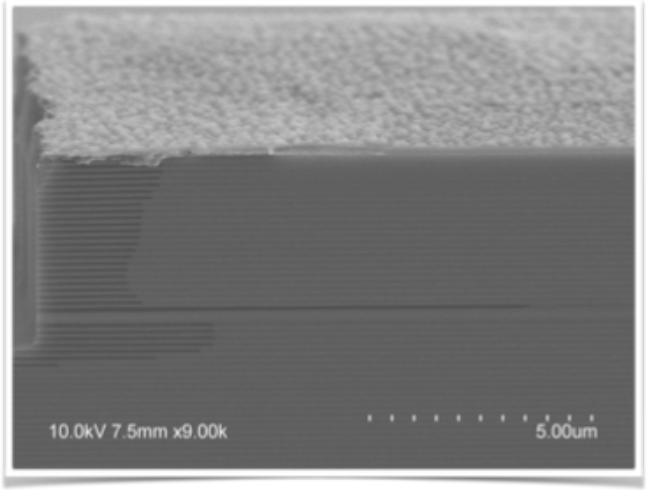
CSL: COORDINATED SCIENCE LAB

# The environment, the team

- An academic institution, first and foremost
  - New researchers every year that need to be trained
    - … and they leave on a regular basis (we hope!)
  - Every experiment every hour is likely to be different
    - … where what was done before could be crucial to what is done next
  - Limiting costs is at a premium
    - … most won't have manufacturers' data tools; images are extracted as **tif** files with loss of metadata
- An interdisciplinary team led by the three IRUs from the College of Engineering
  - Expertise in IT, cybersecurity, material science, device fabrication
  - Working together for the first time and doing very well

# Collected Data

An example of the result from an experiment at MNTL




Result image of 07302013-Oxidation experiment

**Experimental setting:**

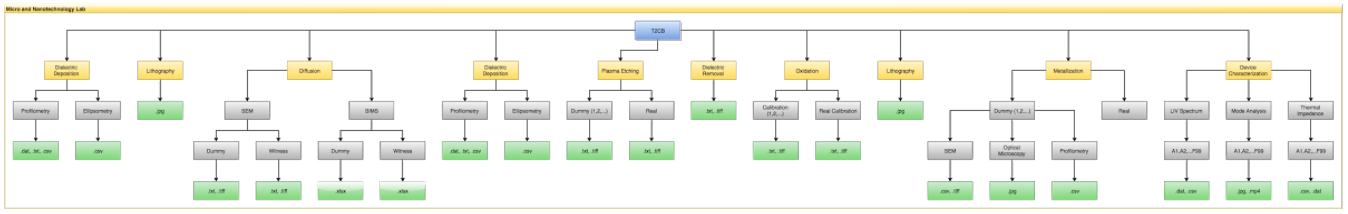
Time 13min
Temp 425 C

(Structured meta data)

**Notes:**

Oxidation depth is about 12um. Oxidation layer composed of Al(0.98)GaAs with thickness of 30 nm. Furnace in 2111 MNT L, 2" diameter quartz.

(Free text)



# Current State of Data Capture in Materials and Semiconductor Domains

- Current situation for experimental data involves manual processes for data capture and storage leading to poor documentation of results

- Data transfer is often done via "sneaker-net" techniques using flash-drives or email
- No data file conversion is available

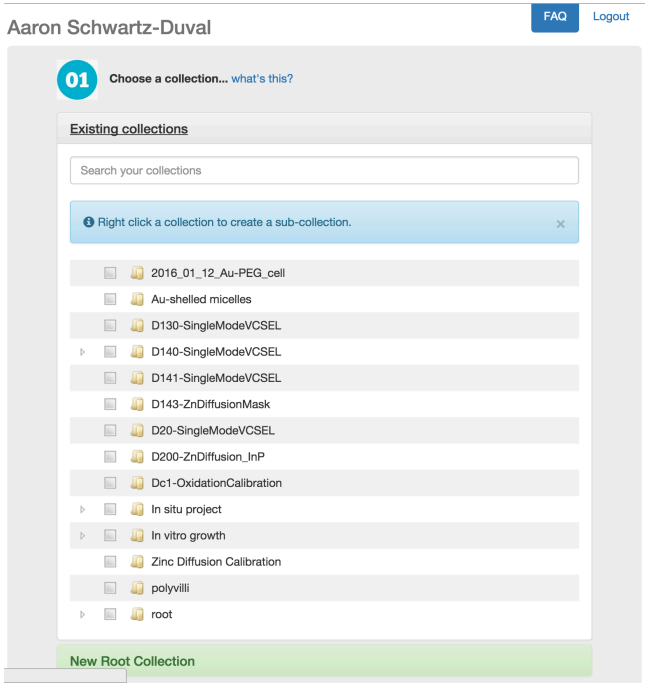
- "Best" results and images are kept, but what is "best" is determined by a narrow, specific scientific objective. "Imperfect" data is often discarded or not available for others to review.

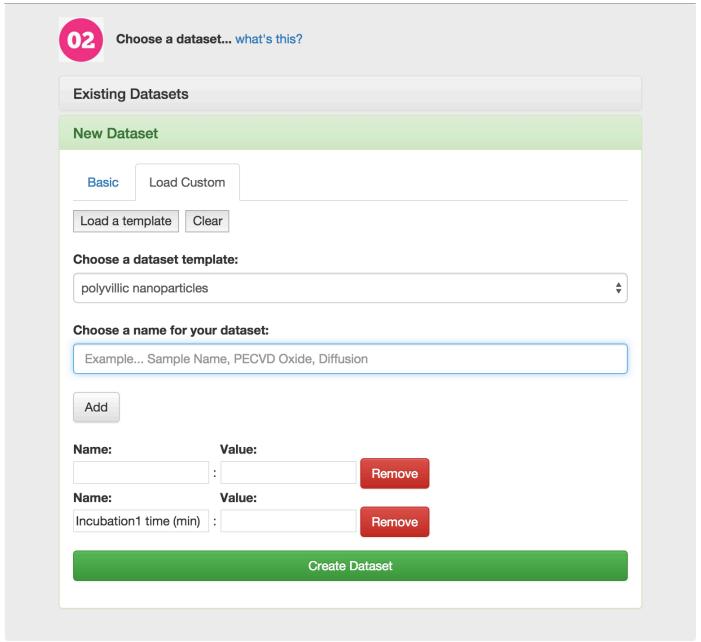
# Proposed Solution

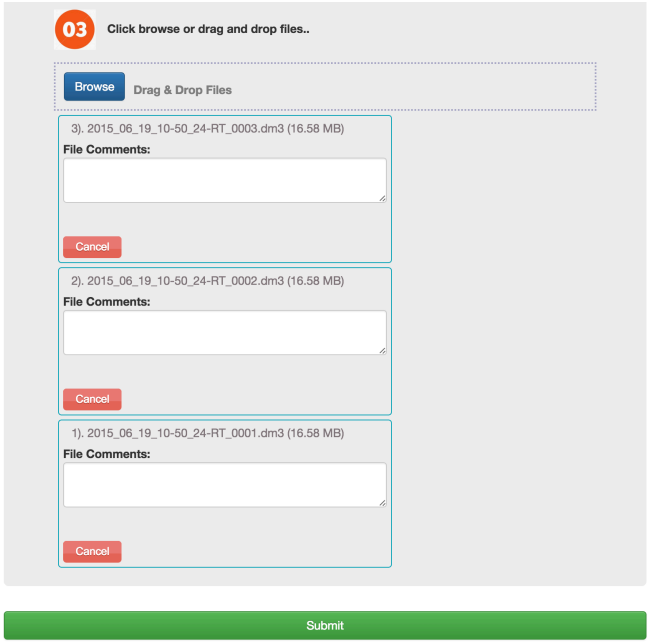
## 4CeeD: Timely and Trusted Capture, Curation, Correlation, Coordination and Distribution

- Timely: Data is collected immediately upon collection
- Trusted: Data is stored so that only those permitted users can see it
- Capture: Data is captured in real-time from microscopes
- Curation: Data is permanently tagged for future usage
- Correlation: Data dependencies are captured
- Coordination: Data of many different types is handled, filtered and classified for easy access and search
- Distribution: Data is disseminated among scientific instruments, clouds and users

# 4CeeD Curator




Create or Select A Collection

Create or Select Dataset

Optional: Choose template and enter metadata

Upload Files

# 071315-ZnDiffusion

Created by Phuong Nguyen
Created on Mar 06, 2016

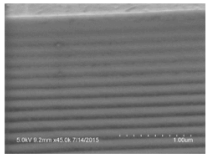


Furnace calibrated for this sample and all further. Tube is 6.24in long and rod 2in shorter than before to avoid previous condensation issue.

Proper disordering found to be 0.72um with zinc presence up to 1um+ deep. Will confirm with SIMS but disordering shows blurred interfaces indicated goal achieved.

+ Add Files | Download All Files | Delete | ★ Follow | Publish | ✔ Create Folder

Files | Metadata | Visualizations | Comments (0)



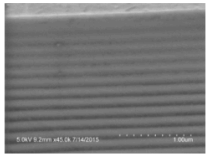
PostDiff.TIF

image/tiff
Mar 06, 2016

0 0 0

Download

★Follow



PostDiff_raw.tif

image/tiff
Mar 06, 2016

0 0 0

Download

★Follow

## Spaces containing the dataset

Select a space + Add

## Collections containing the dataset

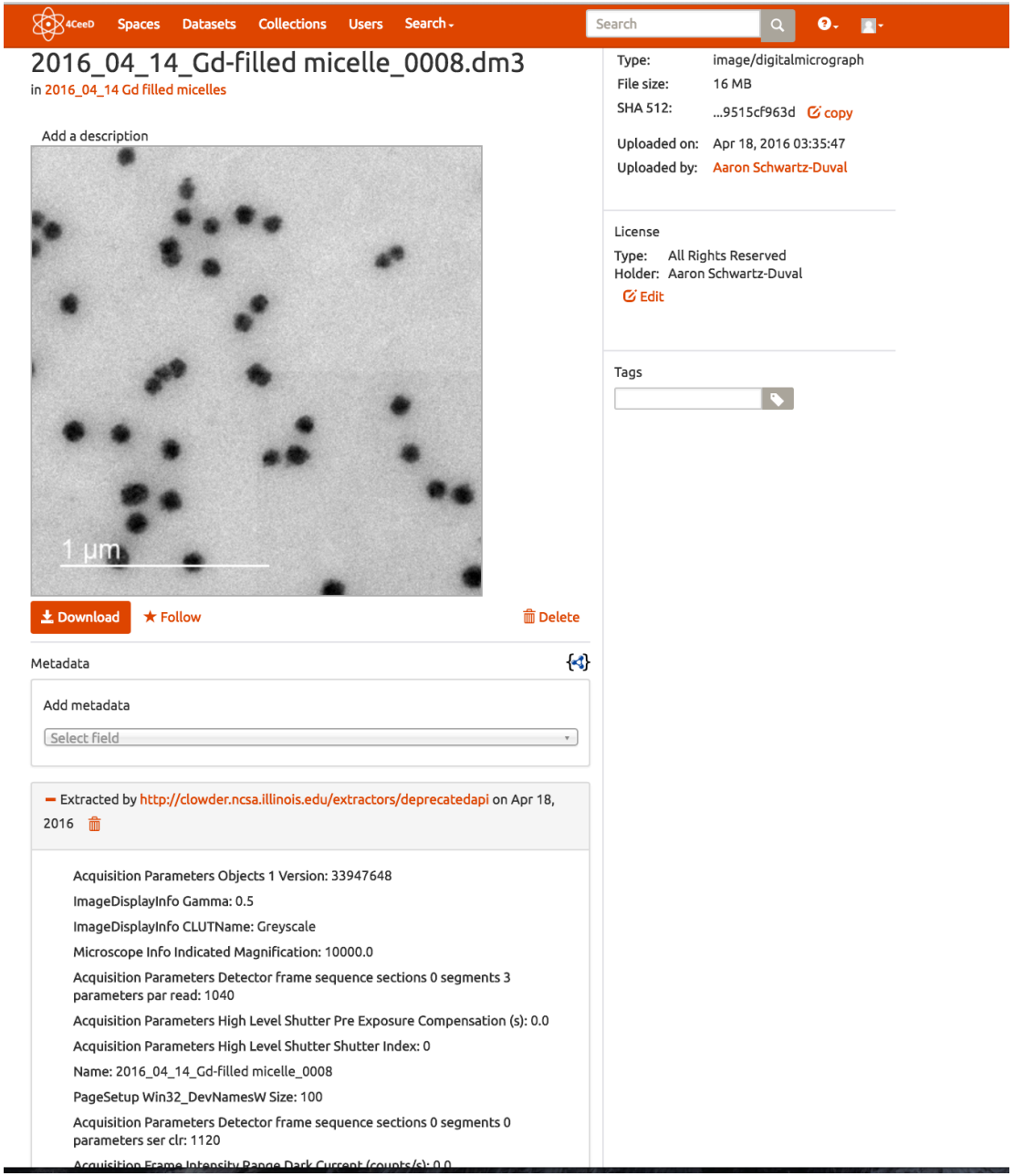
SEM
12 datasets | ✖ Remove

Select a collection + Add

## Tags

Tag

Show tags on files in this dataset


4CeeD
Spaces
Datasets
Collections
Users
Search
Search
2016_04_14_Gd-filled micelle_0008.dm3
in 2016_04_14 Gd filled micelles
Add a description
1 µm
Download
Follow
Delete
Metadata
Add metadata
Select field
Extracted by http://clowder.ncsa.illinois.edu/extractors/deprecatedapi on Apr 18, 2016
Acquisition Parameters Objects 1 Version: 33947648
ImageDisplayInfo Gamma: 0.5
ImageDisplayInfo CLUTName: Greyscale
Microscope Info Indicated Magnification: 10000.0
Acquisition Parameters Detector frame sequence sections 0 segments 3 parameters par read: 1040
Acquisition Parameters High Level Shutter Pre Exposure Compensation (s): 0.0
Acquisition Parameters High Level Shutter Shutter Index: 0
Name: 2016_04_14_Gd-filled micelle_0008
PageSetup Win32_DevNamesW Size: 100
Acquisition Parameters Detector frame sequence sections 0 segments 0 parameters ser clr: 1120
Type: image/digitalmicrograph
File size: 16 MB
SHA 512: ...9515cf963d copy
Uploaded on: Apr 18, 2016 03:35:47
Uploaded by: Aaron Schwartz-Duval
License
Type: All Rights Reserved
Holder: Aaron Schwartz-Duval
Edit
Tags

# Learning from past users "failures"

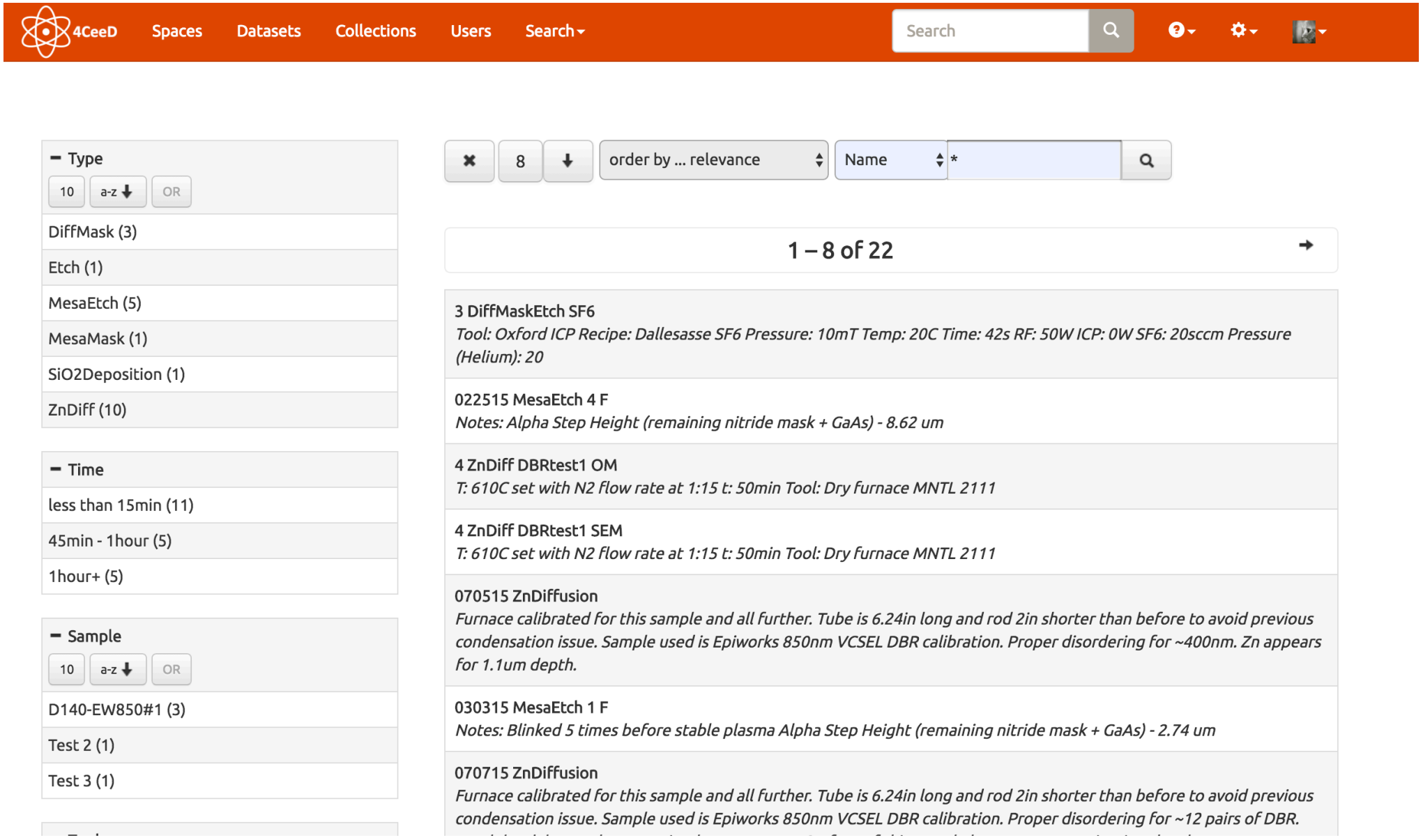
**Discovery**

- *"What is the thickness of Oxidation layer if I set temperature at 500C for 20min?"*
- *"Find all the experiments that produce Oxidation layer below 20nm"*

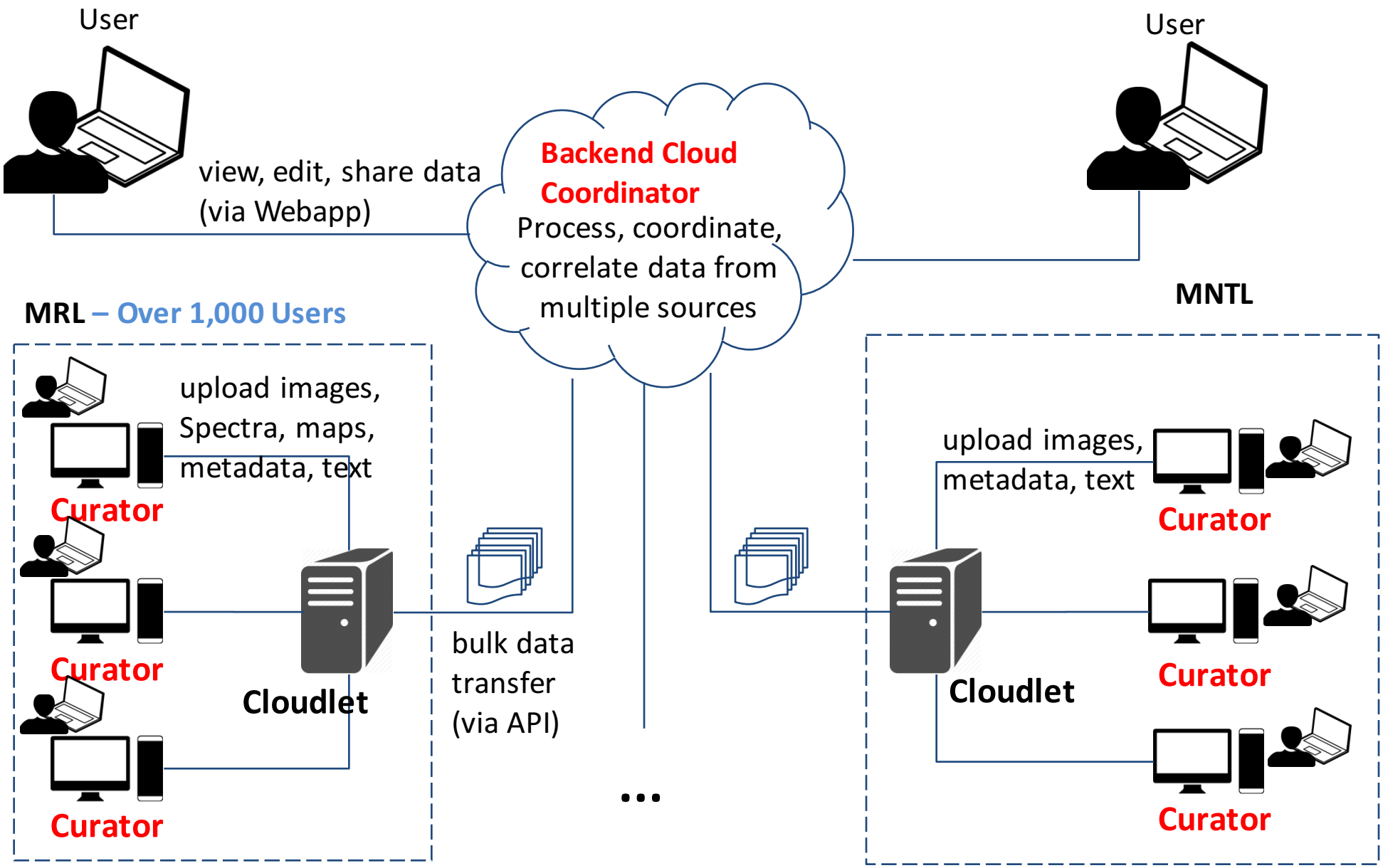
**Error/failure analysis**

- *"What are the common reasons for failed oxidation experiment?"*

# Faceted Search



# Our Approach: 4Ceed Design – Component View



# 4CeeD key features and anticipated benefits

- **What does 4CeeD curator do that nothing else does?**
  - REAL TIME curation of metadata.
  - Metadata Extractors for SEM and proprietary TEM instruments.
  - Clean simple to follow layout and process.
  - Supports complex data models.
  - Uses global and custom templates to maintain accuracy and speed up user defined metadata entry.
  - No data is lost. Support for log, image, and user metadata into single download.

- Time savings (lost time at an expensive machine, lost time for user, better accuracy)
  - Users **can save upwards of 20-30 minutes per lab session**. This can add up to a significant time and money savings over the life of an experiment, with some experiments taking 40-60 hours over 2 months.