

# Electron microscopy facility data management: NexusLIMS



*Nexus***LIMS**

Gretchen Greene, June Lau\*, Marcus Newrock, Ray Plante, Joshua Taillon, Ryan White  
Office of Data and Informatics

# HOUSEKEEPING

## NIST Product Disclaimer:

Certain commercial equipment, instruments, or materials are identified in this presentation to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.

## QR codes throughout this presentation:



Follow our work with your phone!

<https://github.com/usnistgov/NexusLIMS>

# Microscopy facility overview



643 Titan



642 JEOL 3010



642 Philips 400



642 Titan



647 Leo 1525



647 Quanta 400



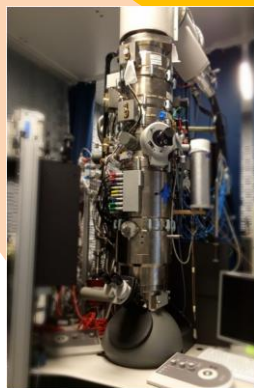
647 Gemini 300



642 JEOL 7100



643 Quanta 200



642 Titan (ETEM)



681 Helios



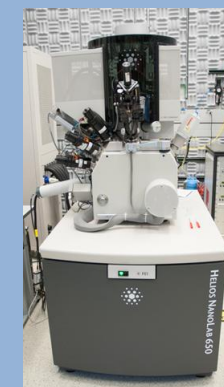
681 Hitachi S5500



630 JEOL 7800



680 Helios (FIB 1)



680 Helios (FIB 2)



680 Titan

# We need a LIMS

(Laboratory Information Management System)

## Design Philosophy

1. Don't reinvent the wheel – use what's available.
2. Future-proof through modularity – technologies change!
3. Don't try to change established workflows and habits – adapt to them instead.



# 30 seconds digression

# Why do we love Wikipedia?

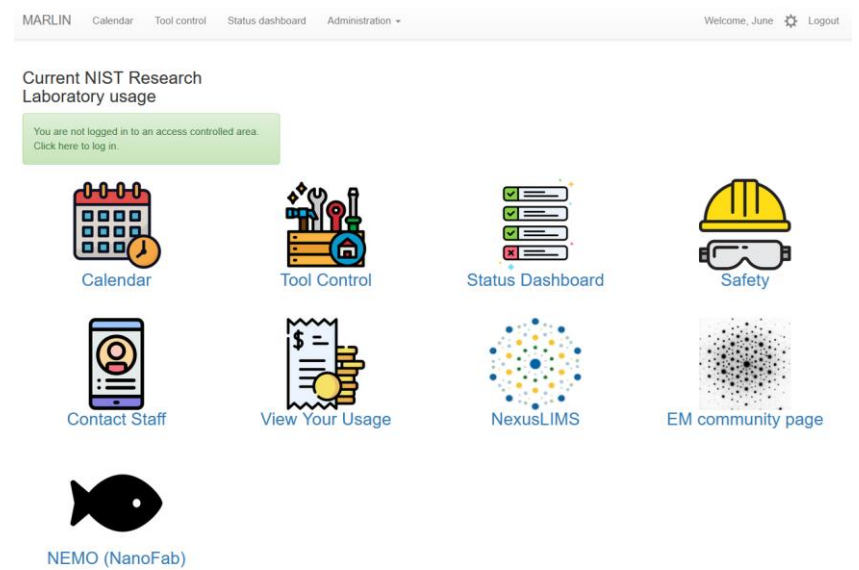
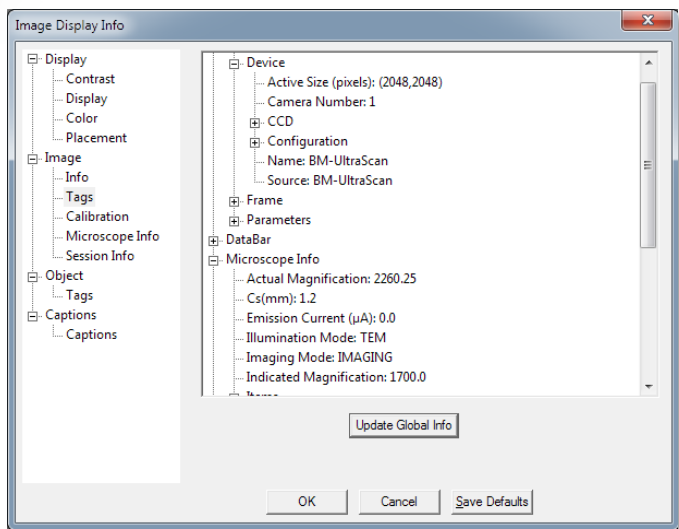
The screenshot shows the Wikipedia article for Jennifer Doudna. The page layout includes a top navigation bar with 'Not logged in', 'Talk', 'Contributions', 'Create account', and 'Log in'. Below this is a search bar and tabs for 'Article' and 'Talk'. The main content area features the article title 'Jennifer Doudna', a sub-header 'From Wikipedia, the free encyclopedia', and several paragraphs of text. A 'Contents' table of contents is visible below the text. On the right side, there is a portrait of Jennifer Doudna and a biographical infobox. The left sidebar contains various navigation and utility links.

Callouts from the image:

- Intuitive headline**: Points to the article title 'Jennifer Doudna'.
- Short synopsis right away**: Points to the sub-header 'From Wikipedia, the free encyclopedia'.
- Well-organized content with a predictable flow**: Points to the 'Contents' table of contents.
- Immediate visual connection to the subject**: Points to the portrait of Jennifer Doudna.
- Changes are tracked**: Points to the 'Edit' button in the top navigation bar.
- People can collaborate and contribute**: Points to the 'View history' button in the top navigation bar.

First knowledge stop for humans

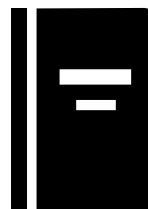
# Reusing our digital crumbs



Microscopy/spectroscopy metadata

electronic lab notebook (ELN)

MARLIN resource management



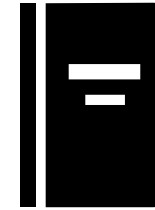
# Sorting metadata into conceptual groups



- Purpose of the experiment
- Sample ID/composition
- User
- Reservation date, time
- Instrument used



- Microscope info
- Imaging conditions
- Detectors in use
- Calibration & reference images
- Datasets



- Sample info
- Setup photos
- Reference and support documents
- Analysis: calculations, graphs, tables

Experimental  
Motivation

Sample Info

Acquisition  
Activity

Data

Analysis

# The NexusLIMS schema

EM schema v1.0  
[doi:10.18434/M32245](https://doi.org/10.18434/M32245)



<i>ID</i>	<i>Property</i>	<i>Obligation</i>
1	Identifier	M
2	SchemaVersion	M
3	LandingPage	M
4	Name	M
5	Owner	M
6	Manufacturer	M
7	Model	R
8	Description	R
9	InstrumentType	R
10	MeasuredVariable	R
11	Date	R
12	RelatedIdentifier	R
13	AlternateIdentifier	R

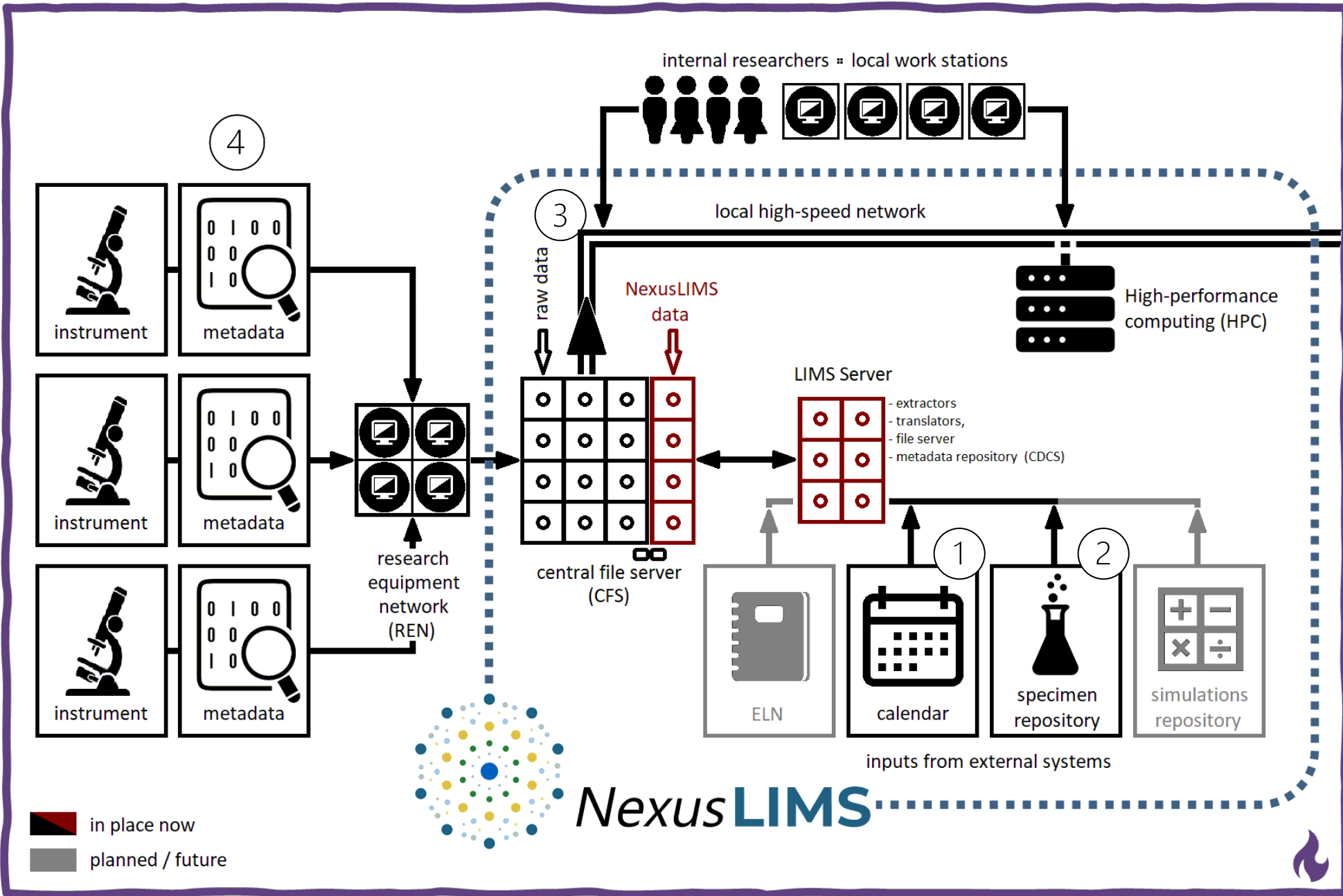
Table 1: Overview of PIDINST properties

<https://doi.org/10.15497/RDA00070>

## Experiment

1. Title
2. Experiment ID
3. Summary
  - a. Experimenter
  - b. Instruments used in the experiment (name, PID - RDA PIDINST v1.0)
  - c. Start
  - d. End
  - e. Motivation
4. Project (ID, link-out)
5. Experiment Ecosystem: microscope, detectors, and sample holders (name, PID - to RDA PIDINST v1.0)
6. Sample (name, PID, link-out)
7. Acquisition activity
  - a. Setup
  - b. Notes
  - c. Dataset(s)

# NexusLIMS overview



# NEMO // MARLIN



<https://github.com/usnistgov/NEMO>

1

tool scheduling



2

specimen info

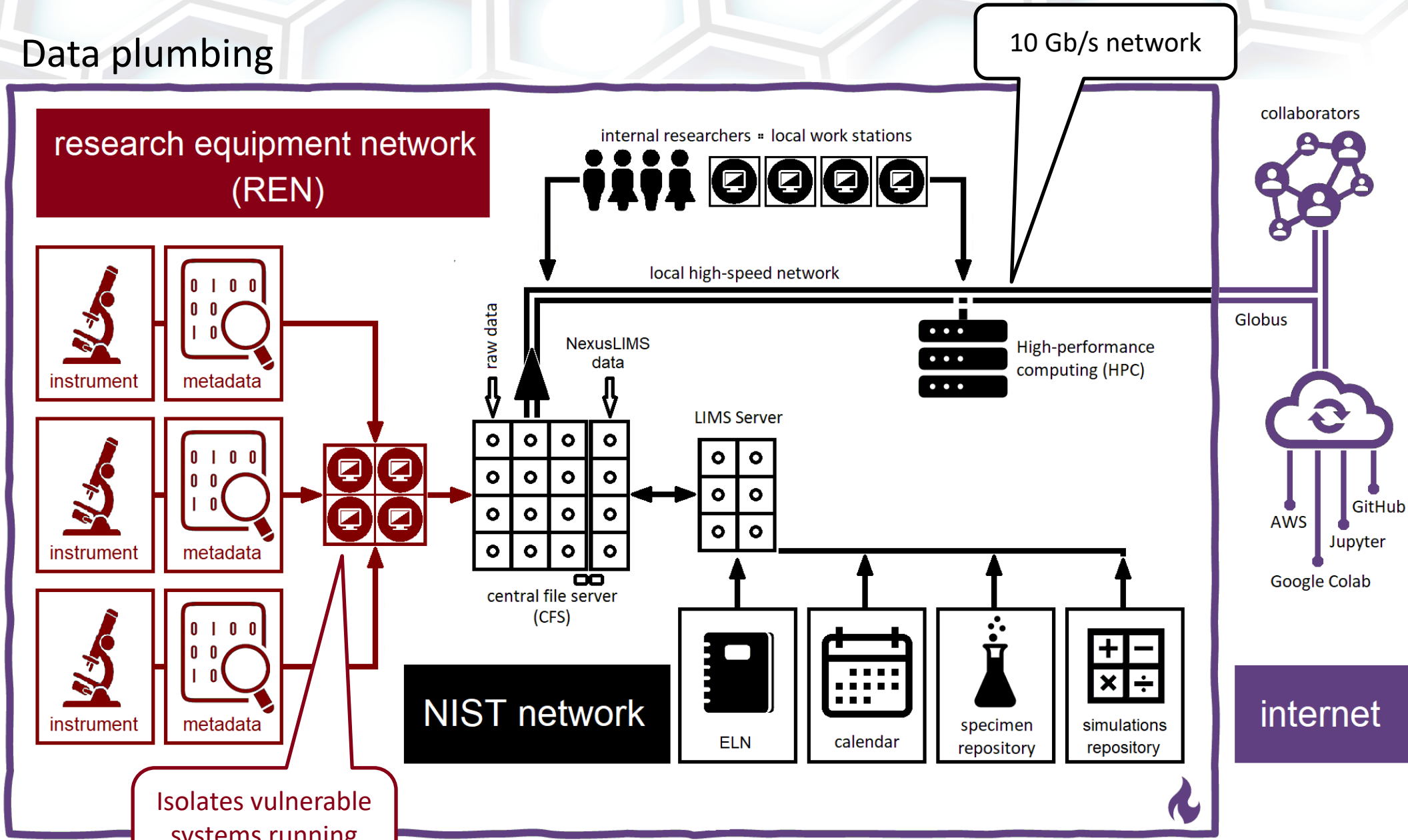


<https://marlin.nist.gov/>

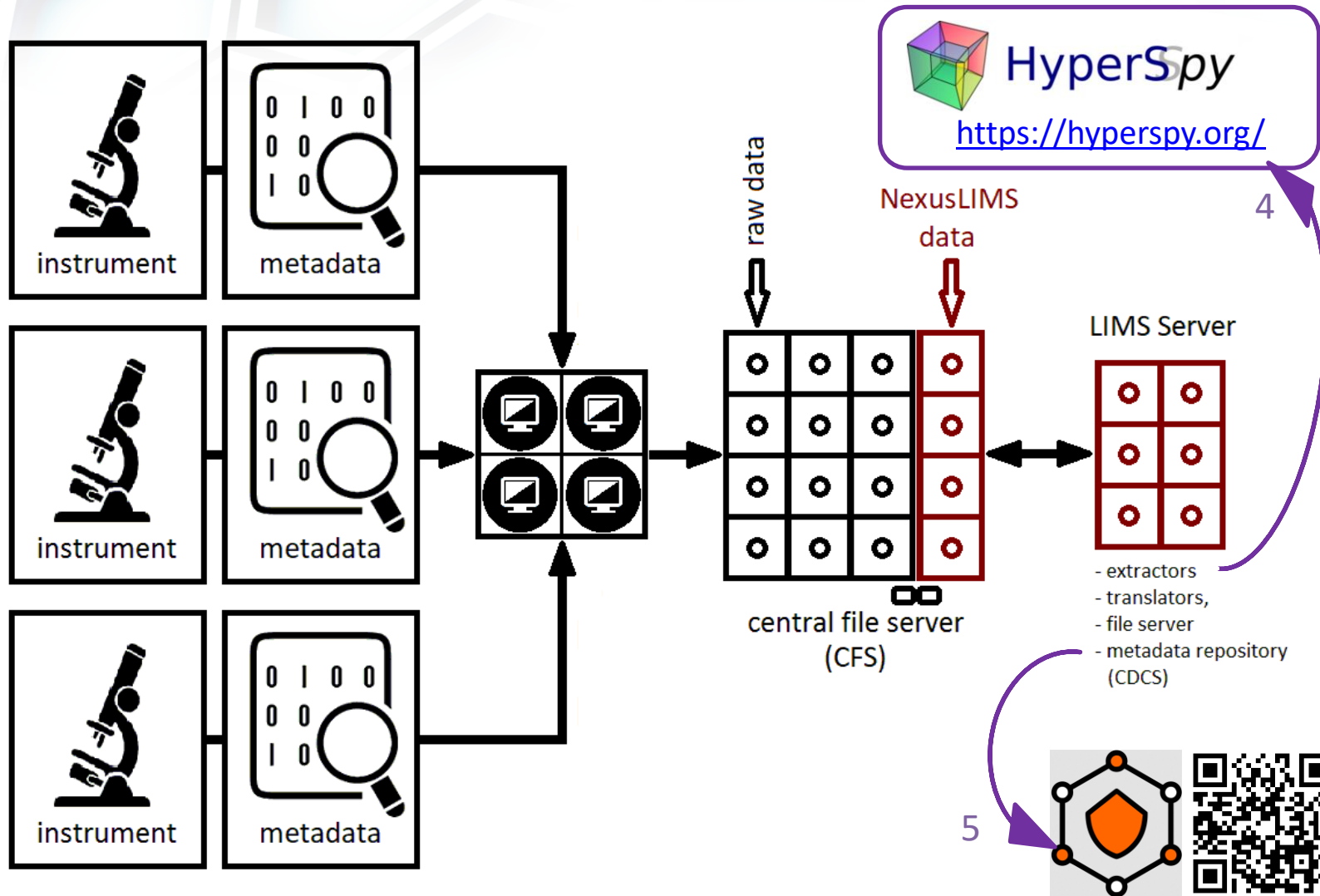
metadata collection at reservation time

# 3

## Data plumbing



# The heart of NexusLIMS



## 4 metadata extraction

1. Microscope data with metadata tags are gathered at CFS location.
2. LIMS server checks MARLIN for event end-time and searches CFS for file changes.
3. Python extractor libraries go to work if changes are detected.
4. Extracted metadata (from experiment, calendar, specimen) is aggregated and written to the CFS in a parallel directory structure.
5. Translation services on the LIMS server populates metadata content into NexusLIMS research records.



<https://github.com/usnistgov/NexusLIMS>

### Experiment activity 7

Activity contents: STEM Imaging, EELS Spectrum Imaging

6 data files 6 dm3

EELS Spectrum Image

```

<meta name="Stage Position
Y">133.41000000000003</meta>
</dataset>
</acquisitionActivity>
<acquisitionActivity seqno="7">
<startTime>2021-04-13T16:22:25.534366</startTime>
<sampleID>11cececc-0bc7-4512-
b8af-0aaae4e7058b</sampleID>
<setup>
<param name="Acquisition Device">DigiScan</param>
<param name="Cs (mm)">1.0</param>
<param name="Data Dimensions">(512, 512)</param>
<param name="Data Type">STEM_Imaging</param>
<param name="Detector" warning="true">GIF
CCD</param>
<param name="GMS Version">2.32.888.0</param>
<param name="Illumination Mode">STEM
NANOPROBE</param>
<param name="Imaging Mode">DIFFRACTION</param>
<param name="Instrument ID">FEI-Titan-
STEM-630901</param>
<param name="Microscope">Titan80-300_D3094</param>
<param name="Name">FEI Technai Remote</param>
<param name="Operation Mode">SCANNING</param>
<param name="Operator" warning="true">
AndrewHerzing</param>
<param name="Specimen" warning="true">STEM</param>
<param name="Stage Position Z">-138.108</param>
<param name="Stage Position
î±">-12.596247942833097</param>
<param name="Stage Position
î²">9.514995512178555</param>
<param name="Voltage">300000.0</param>
</setup>
<dataset type="Image" role="Experimental">

```

6 nm

### Experiment activity 7

Activity contents: STEM Imaging, EELS Spectrum Imaging

Search: <input type="text"> <span><< 1 2 >></span>

Meta D/L

al	≡	↓	↓
al	≡	↓	↓
al	≡	↓	↓
al	≡	↓	↓
al	≡	↓	↓

Meta D/L

ntal	≡	↓	↓
ntal	≡	↓	↓
ntal	≡	↓	↓
ntal	≡	↓	↓
ntal	≡	↓	↓

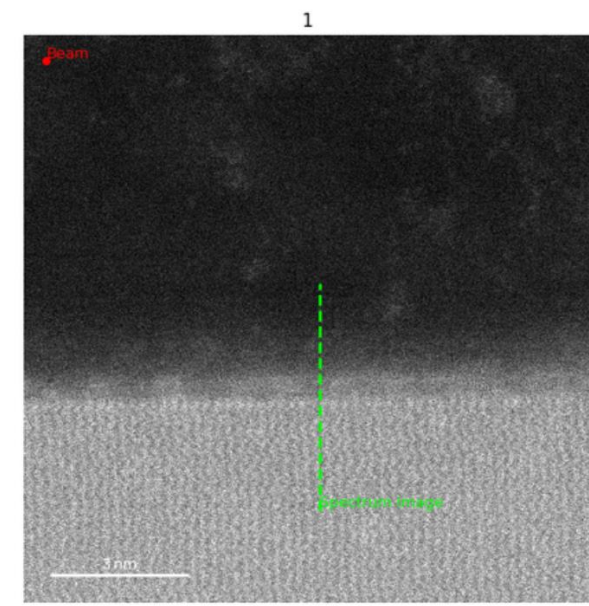
← Back to previous   Download files   **</> Download XML**   Edit this record

activities 94 dm3

April 13, 2021

BED

ay now  
cimen PID



Dataset 21 of 94   Activity 6 of 19




NexusLIMS: A Laboratory Information Management System for Shared-Use Electron Microscopy Facilities  
[10.1017/S1431927621000222](https://doi.org/10.1017/S1431927621000222)

# Persistent identifiers for research records

- For centuries, scientists have maintained lab notebooks with volumes, sections, and numbered pages
  - Allows for easy future reference to specific experiments
- Using a handle server (<https://hdl.net>) provides us a digital alternative!
- Unique identifiers created for every record in the system:
- In the future, "20.500.13060/#####" will always point to the URL for this specific record
- Identifiers are created automatically for every experiment record in the system (user doesn't have to do anything)
- Allows for cross-referencing records or citation of specific experiments

## Organic RT - Radiation damage study

JEOL JEM3010 53 data files in 6 activities  49 dm3 4 tif

Hyeokmin Choe (hnc24) - April 07, 2022

Motivation: Experimental procedure knowledge transfer lab (Eric, Euclid)

Persistent ID: [20.500.13060/18d103b5-9934-4523-aa4b-72a9a1bc4fd8](https://hdl.handle.net/20.500.13060/18d103b5-9934-4523-aa4b-72a9a1bc4fd8) 

Same technology as when you use a DOI to access a scientific paper!

# Human accessible // Machine readable

**NexusLIMS** Browse and Search Records Sharepoint Calendar Tutorial Help Log In / Sign Up

Martensite x Titan x Search Tools

Found 6 Results:

- Martensite Analysis** FEI Titan TEM 95 data files in 9 activities 95 dm3  
 Katz, Michael B. (Fed) - March 11, 2020  
 Motivation: Figure out the martensite orientation relationships with James.
- Martensite Fractions in Steels** FEI Titan TEM 11 data files in 1 activity 11 dm3  
 Katz, Michael B. (Fed) - January 02, 2020  
 Motivation: Figure out what phases are present to confirm the success of the heat treatments.
- Katz x6029** FEI Titan TEM 27 data files in 2 activities 27 dm3  
 Katz, Michael B. (Fed) - November 13, 2018  
 Motivation: Trying to find us some martensite!
- Katz x6029** FEI Titan TEM 0 data files in 0 activities 0  
 Katz, Michael B. (Fed) - May 07, 2019  
 Motivation: Looking for the things between the martensites: precipitates, other phases, etc.

NIST Powered by: CDCS

Human query

Data (.xml)	Query Used
281	
80	"query": [{"\$or": [{"nx.Experiment.acquisitionActivity.setup
8	"query": [{"\$or": [{"nx.Experiment.title": "FEELS"}], [{"nx.E
2	"query": [{"\$or": [{"nx.Experiment.title": "EEDS"}], [{"nx.E
0	"query": [{"\$or": [{"nx.Experiment.title": "CLU"}], [{"nx.Exp
98	"query": [{"\$or": [{"\$or": [{"nx.Experiment.summary
98	"query": [{"\$or": [{"\$or": [{"\$or": [{"nx.Experiment

**Microscopy Data**

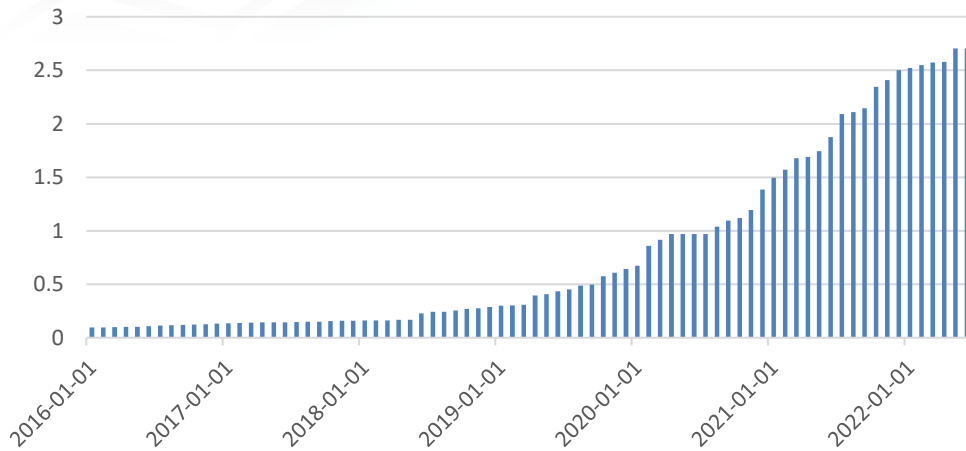
Sheet1 Queries;Data (.xml);Query Used: No Query,281 Magnification Available,80,"query":{"\$or": [{"nx:Experiment.acquisitionActivity.setup.param.@name": "Indicated Magnification"}, {"nx:Experiment.acquisitionActivity.setup.param.@name.#text": "Indicated Magnification"}]}} Spectroscopy;XEELS Da...

Computer query

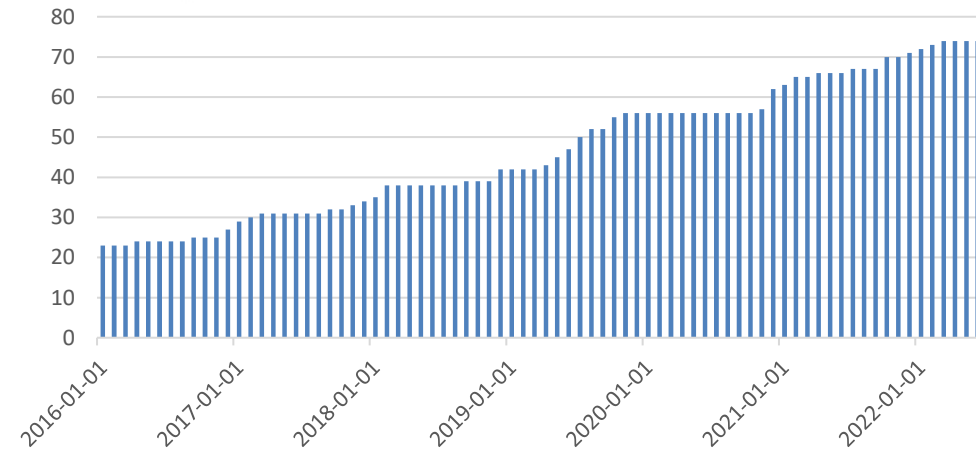
```
<entry>
  <id>FEITitanSTEMEvents(249)</ id>
  <title type="text">An example experiment</ title>
  <updated>2020-09-26T14:25:40-04:00</ updated>
  <link rel="http://schemas.microsoft.com/ado/2007/08/dataservices/related/UserName"
        type="application/atom+xml;type=feed" title="UserName"
        href="FEITitanSTEMEvents(249)/UserName" >
    <content type="application/xml" >
      <d:Name>Taillon, Joshua A. (Fed)</ d:Name>
      <d:WorkEmail>joshua.taillon@nist.gov</ d:WorkEmail>
      <d:Department >641</ d:Department >
      <d:FirstName >Joshua</ d:FirstName >
      <d:LastName>Taillon</ d:LastName >
      <d:Office>02/2607</ d:Office >
    </ content >
  </ link >
  <content type="application/xml" >
    <m:properties >
      <d:ContentTypeID >0x010200B0B03B4BA3198A45ABDA7DB6A835DE33</ d:ContentTypeID >
      <d:TitleOfExperiment >An example experiment</ d:TitleOfExperiment >
      <d:StartTime m:type="Edm.DateTime" >2020-09-14T15:00:00</ d:StartTime >
      <d:EndTime m:type="Edm.DateTime" >2020-09-14T19:00:00</ d:EndTime >
      <d:CategoryValue >User Session</ d:CategoryValue >
      <d:ExperimentPurpose >To test the powers of NexusLIMS</ d:ExperimentPurpose >
      <d:SampleDetails >dummy-sample-1</ d:SampleDetails >
      <d:ProjectID>NexusLIMS</ d:ProjectID >
    </ m:properties >
  </ content >
</ entry >
```

# How is NexusLIMS doing?

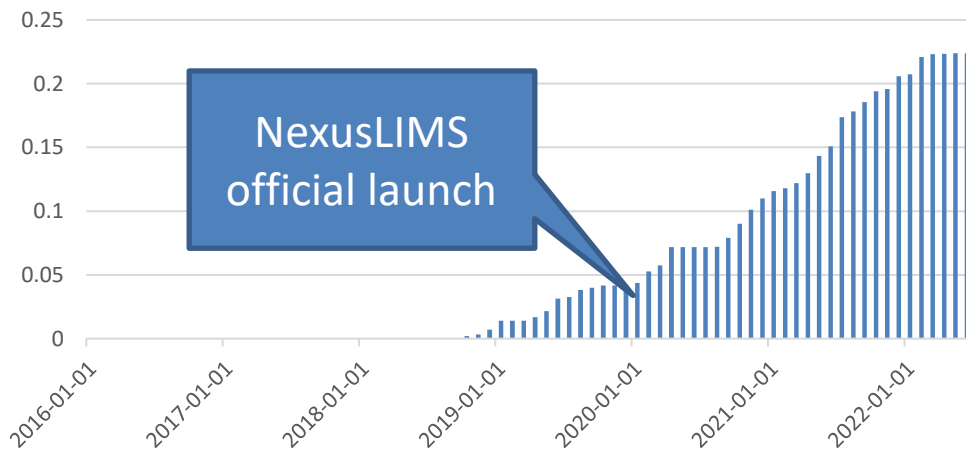
Bytes (CFS) over time (TB)



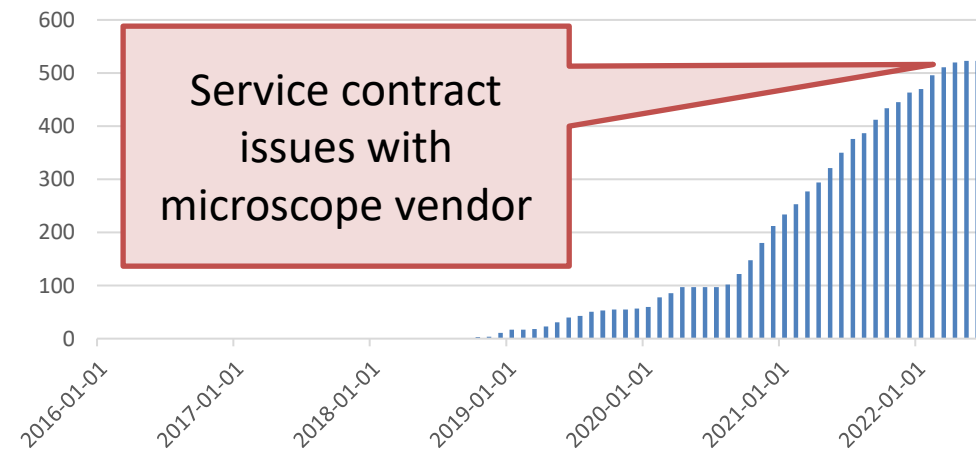
Users with files



Bytes (Nexus) over time (TB)



Nexus records

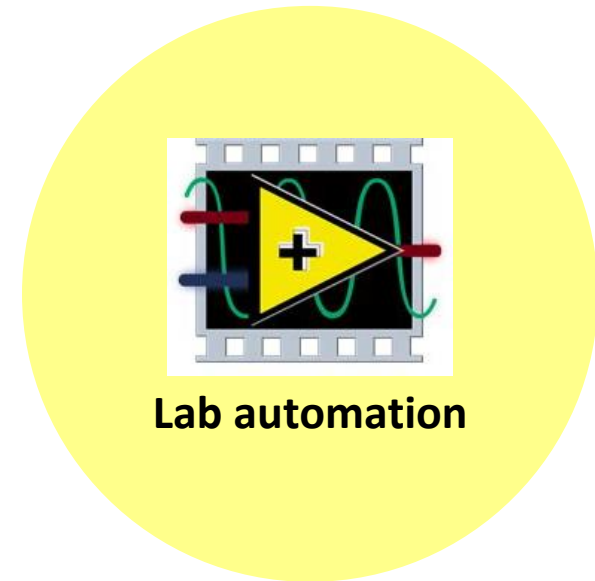


# Future services



**Microsoft Graph**

Microsoft 365    Windows 10    Enterprise Mobility + Security



# Much work ahead – community can help!

1. Standard vocabulary or dictionaries
2. Units standardization
3. More extractors for proprietary formats
4. Persistent Identifiers (PID)
5. Schema updates
6. More modules:
  - Simulation
  - Analysis
  - ELN
  - ....

Imagine better ways

[june.lau@nist.gov](mailto:june.lau@nist.gov)



xkcd #1838