

SciCatCon - Diamond Update

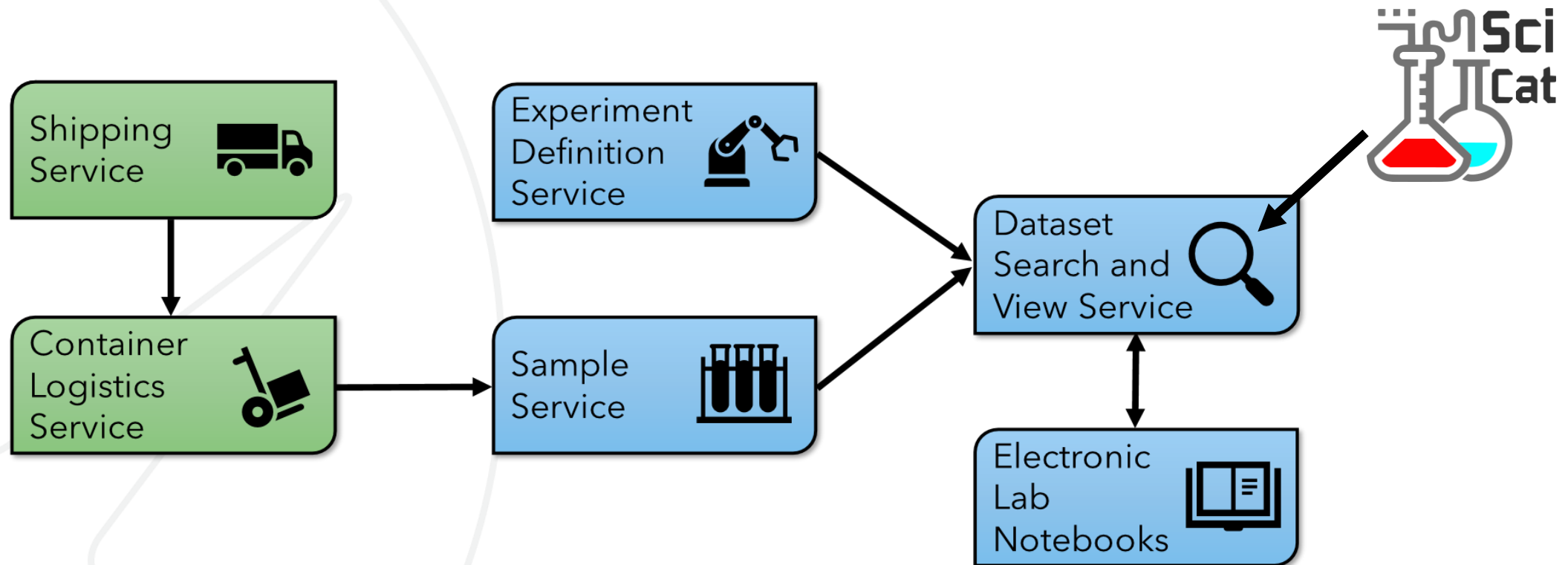
Ian Bush

Universal LIMS Team Lead

Diamond-II Software Infrastructure: Universal LIMS

Universal Laboratory Information Management System (LIMS)

- Universal LIMS to support needs of all groups, except Structural Biology & I19, which will continue with ISPyB/SynchWeb



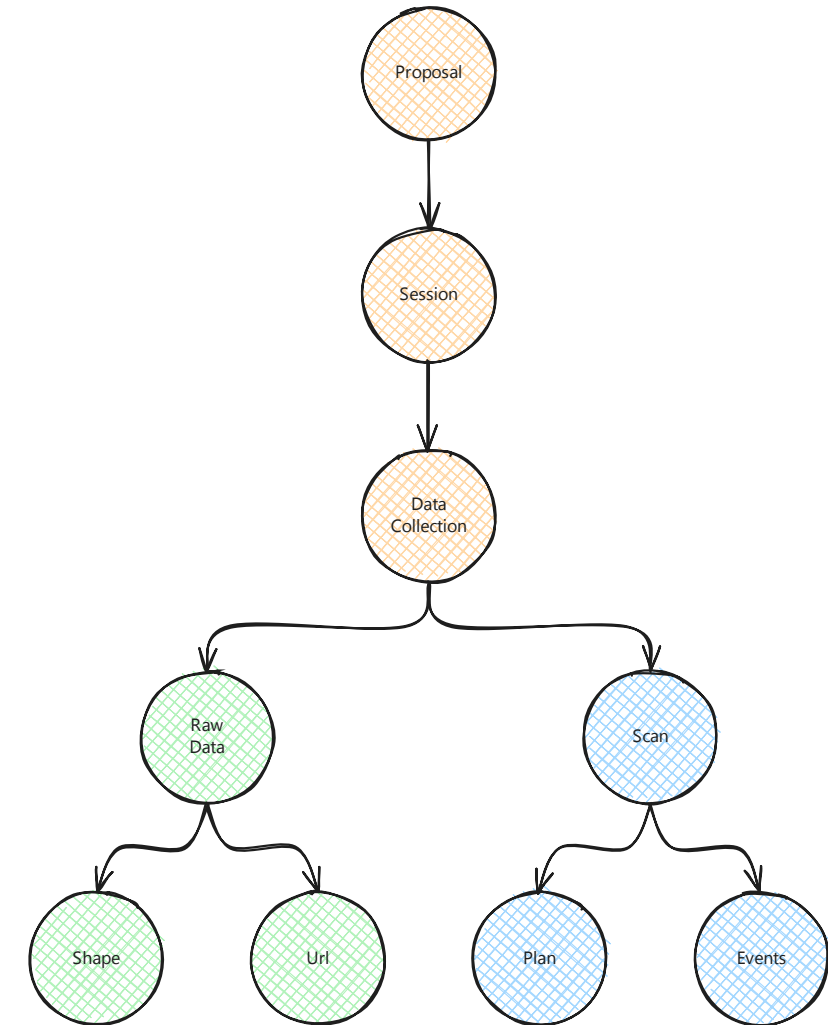
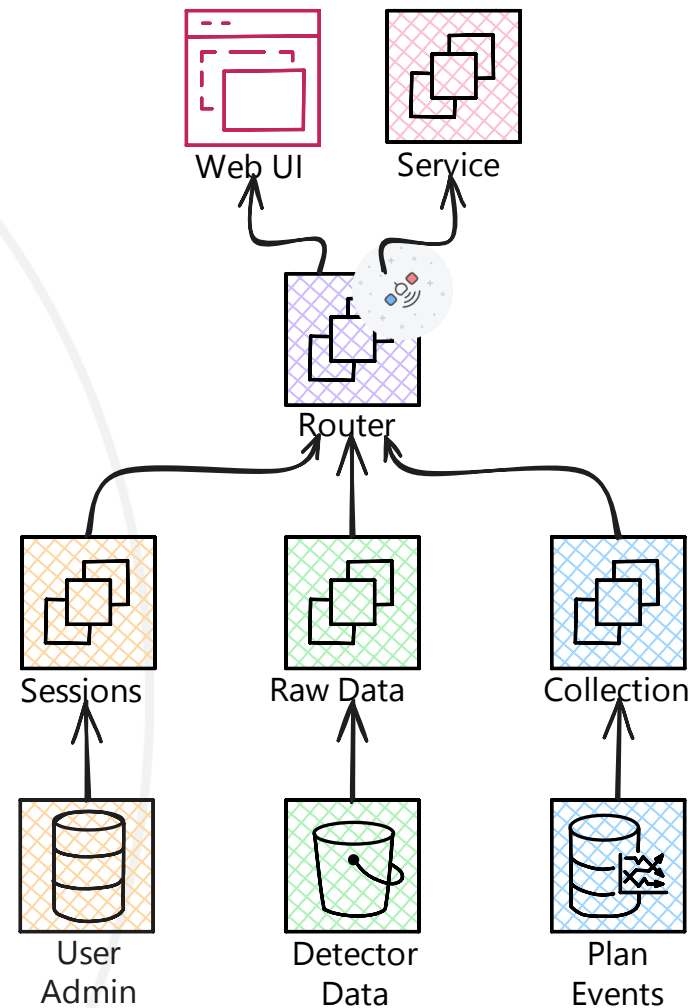
Diamond-II Software Infrastructure: GraphQL API

Adopting Federated GraphQL

- Enables presentation of a single interface to data consumers
- Exposes object relationships via API schema
- Decouples service implementation of domain teams
- Forces conversations at domain team boundaries

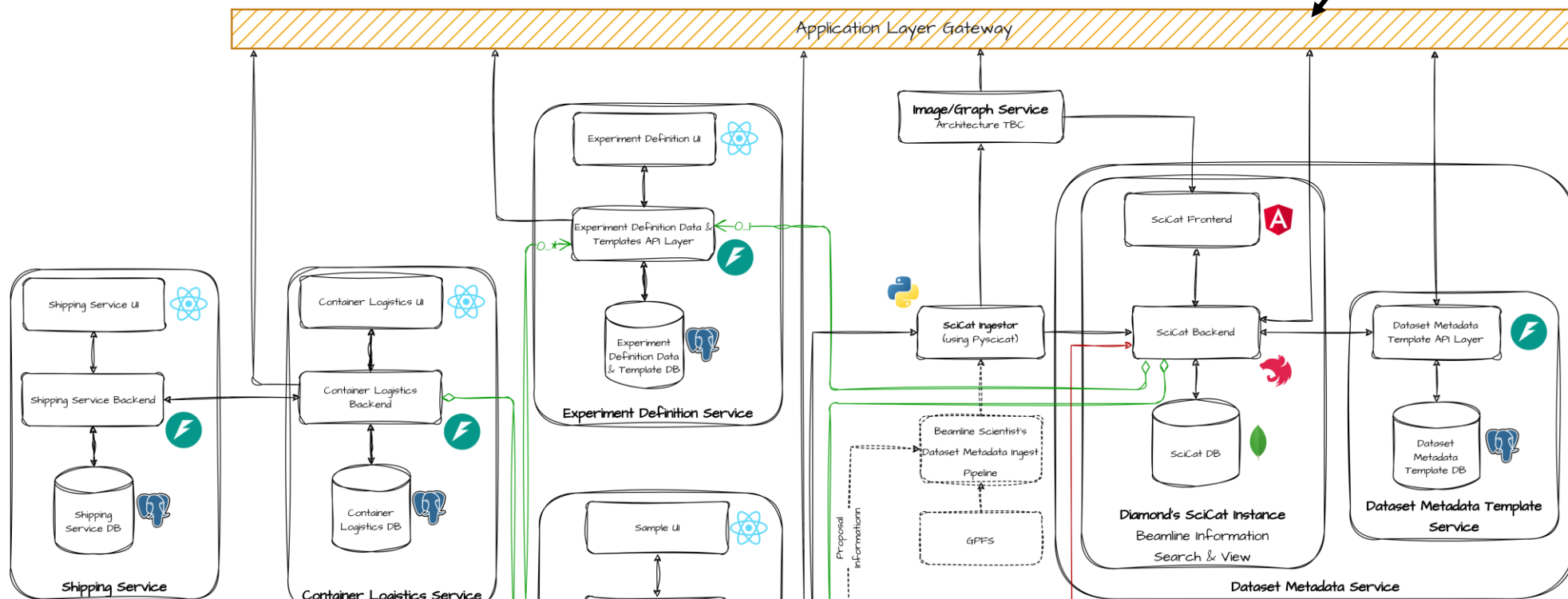
Progress

- Prototype created and evaluated
- Technical Design Steering Group approval



Universal LIMS: Draft Architecture

Common Beamline Architecture
– GraphQL API Gateway



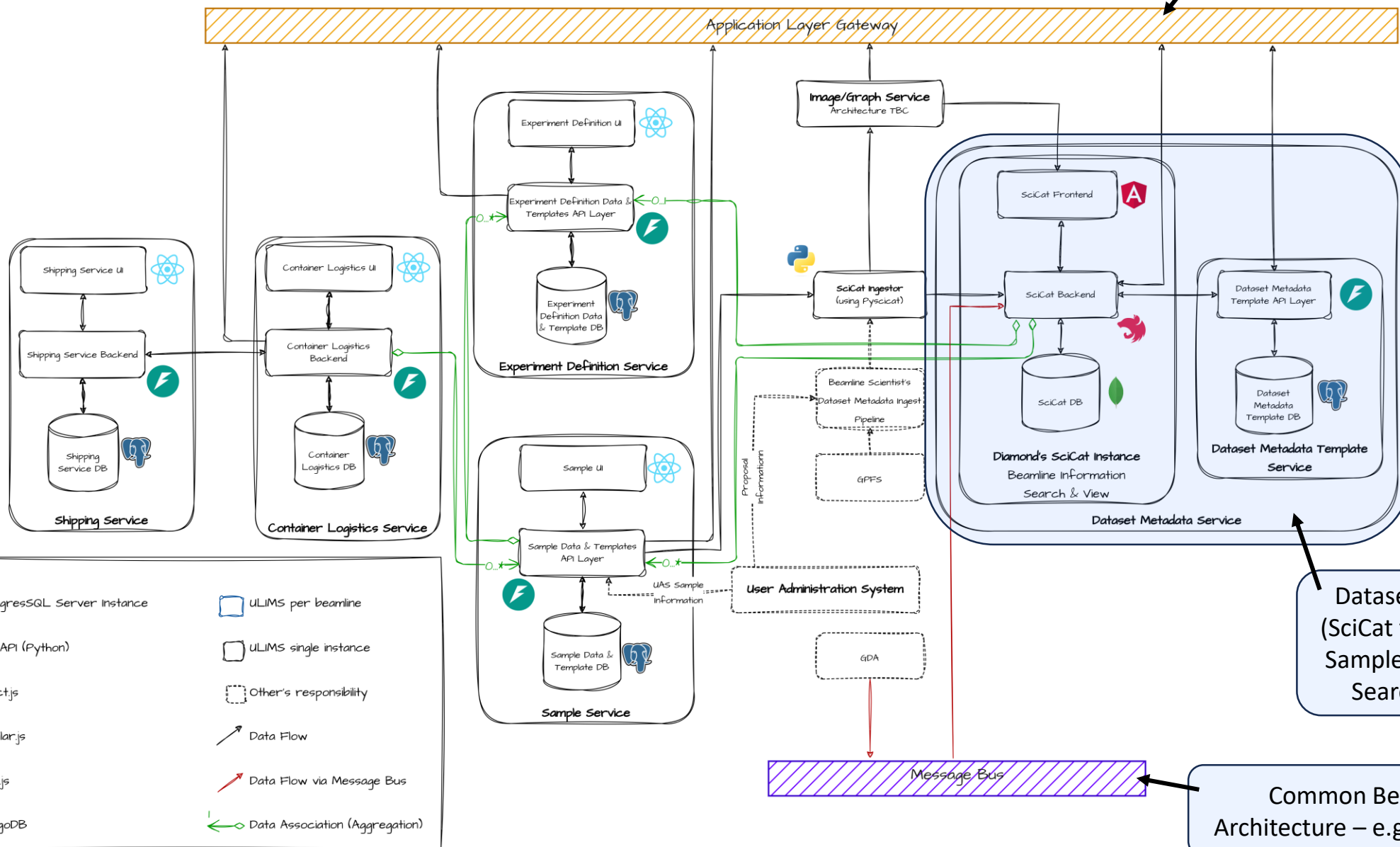
Key

| | | | |
|--|----------------------------|--|--------------------------------|
| | PostgreSQL Server Instance | | ULIMS per beamline |
| | FastAPI (Python) | | ULIMS single instance |
| | React.js | | Other's responsibility |
| | Angular.js | | Data Flow |
| | Nest.js | | Data Flow via Message Bus |
| | MongoDB | | Data Association (Aggregation) |

Common Beamline Architecture
– e.g. RabbitMQ

Universal LIMS: Draft Architecture

Common Beamline Architecture
– GraphQL API Gateway



Dataset Metadata
(SciCat for Dataset &
Sample Information
Search & view)

Common Beamline
Architecture – e.g. RabbitMQ

Key

| | | | |
|--|----------------------------|--|--------------------------------|
| | PostgreSQL Server Instance | | ULIMS per beamline |
| | FastAPI (Python) | | ULIMS single instance |
| | React.js | | Other's responsibility |
| | Angular.js | | Data Flow |
| | Nest.js | | Data Flow via Message Bus |
| | MongoDB | | Data Association (Aggregation) |

Universal LIMS: Templating

- To support a wide variety of experiment types, sample types, dataset types etc. we need a flexible approach
- Templates will provide the ability for users to define the metadata structures needed
- Data is validated against a given template before being stored
- Templates will be immutable and versioned
- Template manager service with a GUI envisioned, so no knowledge of the structure of template format would be needed

```
{
  "title": "B24MetadataSample",
  "type": "object",
  "properties": {
    "Name": {
      "title": "Name",
      "type": "string"
    },
    "Description": {
      "title": "Description",
      "type": "string"
    },
    "HazardGroup": {
      "title": "Hazardgroup",
      "type": "integer"
    },
    "SafetyLevel": {
      "title": "Safetylevel",
      "type": "string"
    },
    "Date": {
      "title": "Date",
      "type": "string"
    }
  }
}
```

A simple B24 Sample Definition example

Universal LIMS: GUI & Search

- GUI – a further template will be used to define, for example, how a particular dataset should be displayed for a beamline in the Data Viewer
- A point-and-click application to build the GUIs will also be developed
- Search – a template can be defined to provide a custom search on the data

```
{
  "type": "VerticalLayout",
  "elements": [
    {
      "type": "Control",
      "scope": "#/properties/FileName"
    },
    {
      "type": "Group",
      "label": "Instrument",
      "elements": [
        {
          "type": "VerticalLayout",
          "elements": [
            {
              "type": "HorizontalLayout",
              "elements": [
                {
                  "type": "Control",
                  "scope": "#/properties/CameraName"
                },
                {
                  "type": "Control",
                  "scope": "#/properties/ObjectiveName"
                }
              ]
            },
            {
              "type": "Control",
              "scope": "#/properties/CameraPixelSize"
            },
            {
              "type": "Control",
              "scope": "#/properties/CameraBinning"
            }
          ]
        }
      ]
    }
  ]
}
```



Scientific Metadata

File Name
test_tomo2_e3C_full.txrm

Instrument

| | |
|----------------------|----------------------|
| Camera Name Pixis | Objective Name 1x |
| Camera Pixel Size | Camera Binning |

*A simple
B24 GUI
Metadata
Definition,
and
associated
GUI*

Diamond's SciCat: Current Status

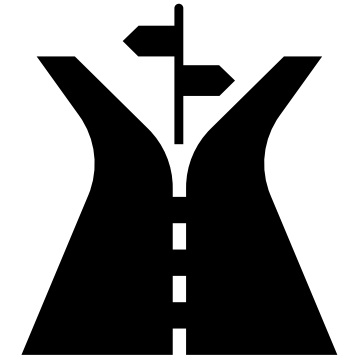
Using SciCat for the Dataset Search and View part of Universal LIMS

- Main Changes:
 - Support for saving data against a template (JSON Schema)
 - Scientific Metadata and Dataset Table in frontend displayed according to JSON forms template
 - Wrapper for PySciCat to make ingesting Diamond data easier (e.g. looking up permissions)
 - Some frontend changes, e.g. to display an image slideshow, show related datasets and formatting of job data
 - <https://mnemosyne.diamond.ac.uk>
- Currently rolled out on one beamline – B24, over 10,000 datasets in production
- Next steps to rollout to I07 (GDA beamline) and EPSIC
- Currently based on an out-of-date SciCat fork – needs updating!
- One small PR to main SciCat repo to fix performance issue with attachments

SciCat at Diamond: Decisions and Challenges

Outstanding Decision on Using SciCat

- Does our intended direction fit with the wider SciCat collaboration with SciCat, e.g. using templates as part of a LIMS system.
- Can we achieve the level of customisation we need, for example for a data collection page?
- Will we be able to integrate a Diamond SciCat instance fully with D-II architecture, particularly the GraphQL gateway? This will need consideration for both serving and ingesting data via the GraphQL gateway.



Thanks for listening – any questions?

Screenshot: SciCat Visit Summary

The screenshot displays the SciCat ESS test web application interface. The browser address bar shows the URL: localhost:4200/datasets?args={\"modeToggle\": \"all\", \"mode\": {}, \"text\": \"\", \"creationTime\": null, \"type\": {}, \"creationLocation\": {}, \"ownerGroup\": {}, \"skip\": 0, \"limi\"}. The application header includes the SciCat logo, ESS logo, and navigation links for Help, About, and Ingestor. The main content area is titled 'Datasets /' and features a search and filter sidebar on the left. The sidebar includes a 'PID Selection' field, a 'Search' section with 'Text Search', 'Location', 'Group', 'Type', 'Keywords', and 'Start Date - End Date' filters, and a '+ Add Condition' button. The main area shows a list of datasets with columns for 'My Data' and 'All Public Data'. The first dataset is 'SIRT_tomo_E3E_area3_idx00000_20240315_0227', which is derived, with a sample ID of '?', created at 15/03/2024 2:49:59, and updated at 15/03/2024 2:50:0. Below it is a 'Hide Related Datasets' button. The second dataset is 'tomo_E3E_area3_idx00000.txrm', which is raw, with a sample ID of 'bi34460_PW-904-G3', created at 2024-03-15T02:28:25.091Z, and updated at 2024-03-15T02:28:26.206Z. Below it is a 'Show Related Datasets' button. The third dataset is 'Patch_tomo_E3E_area3_idx00000_20240315_0227', which is derived, with a sample ID of '?', created at 15/03/2024 2:49:23, and updated at 15/03/2024 2:49:24. Below it is another 'Show Related Datasets' button. The fourth dataset is 'WBP_tomo_E3E_area3_idx00000_20240315_0227', which is derived, with a sample ID of '?', created at 15/03/2024 2:48:47, and updated at 15/03/2024 2:48:49. The interface also includes a 'PID' search field, 'Items per page: 100', '1 - 100 of 7400' items, and 'Sort By' options (ASC, DESC).

A view of a collection of datasets, for example for a single visit. Options to filter/search available in this view.

Screenshot: SciCat Dataset Details View

SciCat DLS

https://memosyne.diamond.ac.uk/datasets/diamond_pid%2Fdb095583-be85-4b36-861f-020295341aef

120%

SciCat diamond

Help About

witb43145

Scientific Metadata

tomo_E3E_area3_idx00000.txrm

File Name

tomo_E3E_area3_idx00000.txrm

Instrument

| | | | |
|-------------------|----------------|-----------------------|-----------------|
| Camera Name | Objective Name | | |
| Pixis | Pixis 1x | | |
| Camera Pixel Size | Camera Binning | Optical Magnification | Zone Plate Name |
| 13 | 1 | 1 | Y |

Images

| | |
|------------------|--------------|
| Number of Images | Image Height |
| 551 | 946 |
| Camera Binning | Image Width |
| 1 | 946 |

09/07/2024

Show Image Metadata

Open Image Slideshow

SSAC #1

1 11 22 33 44 55

A detailed view of the dataset. The GUI templates define the layout under 'Scientific Metadata'.