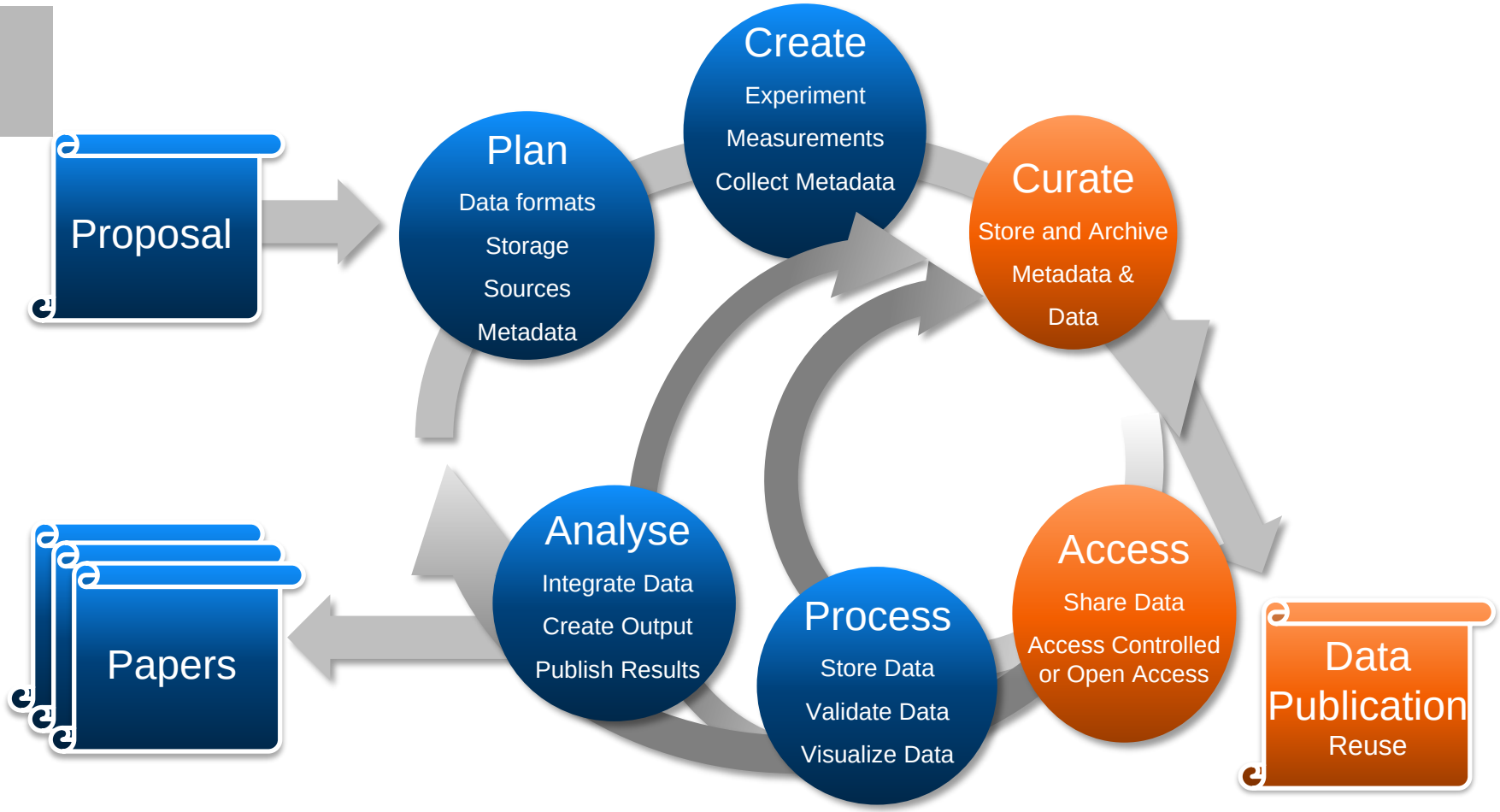
Carlo Minotti (PSI), Stephan Egli (PSI)

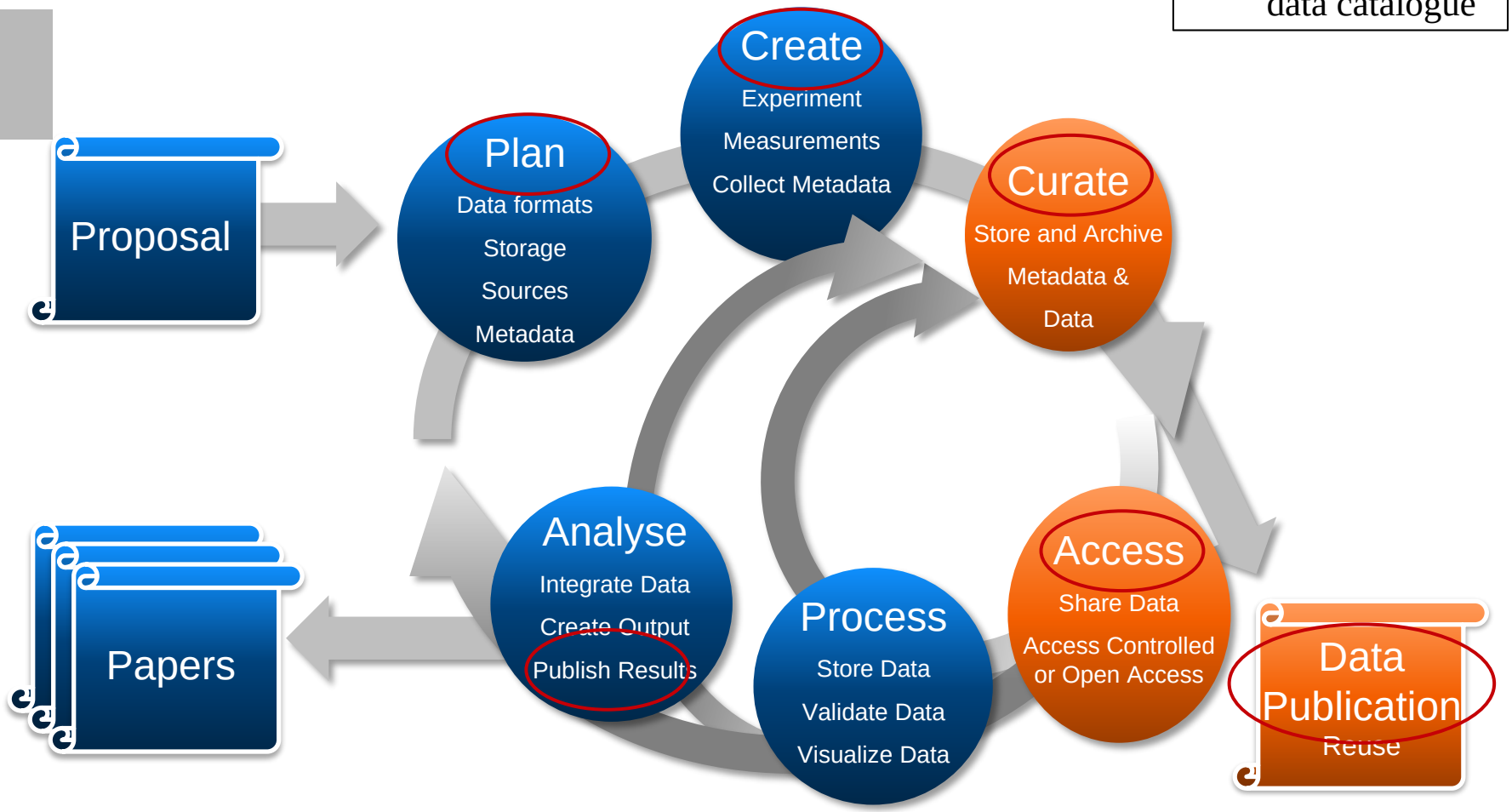
SciCat as a Tool for Users and Data Managers

DESY User Meeting,
Satellite Meeting "Science driven data management solutions" - 24/01/22

A typical user journey



Where data catalogue is used



Where does SciCat help the Scientists?

- **Organize** the scientific data into datasets
- Annotate the Datasets with **administrative** and **flexible scientific metadata**
- Make the data **searchable/discoverable**
- Provides the infrastructure for **publishing** the data, DOI generation
- Can be used as frontend for **longterm** storage (Archive) solutions of mass data (PB regime)
- Supports both **open access** and **embargoed** data

Discover data via WebUI

User specific data

Archive Interface

Facet search

Help About 0 sistomcat

Search Clear

Text Search

Location

Group

p18788 | 2300

p18762 | 10

p18761 | 49

p18748 | 147

p18675 | 18

My Data | Public Data | All | Archivable | Retrievable | Work In Progress | System Error | User Error

Add to Cart

Items per page: 25 1 – 25 of 111272

Name	Source Folder	Size	Start Time	Type	Proposal ID	Group	Data Status
029_estaillades1_q01_fw085_ss	...1_fw085_ss	1 TB	2020-12-23 Wed 00:05	derived		p17614	retrievable
020_estaillades1_q01_fw085_us	...1_fw085_us	729 GB	2020-12-23 Wed 00:05	derived		p17614	retrievable
019_estaillades1_q01_fw085_us	...1_fw085_us	376 GB	2020-12-23 Wed 00:05	derived		p17614	retrievable
018_estaillades1_q01_fw085_us	...1_fw085_us	376 GB	2020-12-23 Wed 00:05	derived		p17614	retrievable
031_estaillades1_q01_fw085_ss	...1_fw085_ss	4 TB	2020-12-22 Tue 22:02	derived		p17614	retrievable
20201214_ANAXAM/11_360_	...AM/11_360_	47 GB	2020-12-14 Mon 20:59	raw	unknown	p17896	archivable
20201214_ANAXAM/10_360_	...AM/10_360_	47 GB	2020-12-14 Mon 20:37	raw	unknown	p17896	archivable
09_360/09_360_S13_	...9_360_S13_	47 GB	2020-12-14 Mon 20:09	raw	unknown	p17896	archivable
09_360/09_360_S12_	...9_360_S12_	47 GB	2020-12-14 Mon 20:03	raw	unknown	p17896	archivable
09_360/09_360_S11_	...9_360_S11_	47 GB	2020-12-14 Mon 19:57	raw	unknown	p17896	archivable
09_360/09_360_S10_	...9_360_S10_	47 GB	2020-12-14 Mon 19:52	raw	unknown	p17896	archivable
09_360/09_360_S09_	...9_360_S09_	47 GB	2020-12-14 Mon 19:46	raw	unknown	p17896	archivable
09_360/09_360_S08_	...9_360_S08_	47 GB	2020-12-14 Mon 19:40	raw	unknown	p17896	archivable
09_360/09_360_S07_	...9_360_S07_	47 GB	2020-12-14 Mon 19:35	raw	unknown	p17896	archivable
09_360/09_360_S06_	...9_360_S06_	47 GB	2020-12-14 Mon 19:29	raw	unknown	p17896	archivable

Editing of Metadata

Help

You are editing Published Data record.

Title *

Real-Time Imaging Reveals Distinct Pore-Scale Dynamics During Transient and Equilibrium Subsurface Multiphase Flow

Creators *

Catherine Spurin | Tom Bultreys | Maja Rücker | Gaetano Garfi | Christian M. Schlepütz | Vladimir Novak | Steffen Berg | Martin J. Blunt | Samuel Krevor

Publisher *

PSI

Resource Type *

derived

Description *

This published data collection contains five datasets obtained by X-ray tomographic microscopy of a carbonate rock sample 5 mm in diameter and 20 mm in length. Both brine and nitrogen gas are injected into the sample at a total flow rate of 0.1 ml/min (the brine made up 85% of this total flow rate). Data were collected and processed at the TOMCAT beamline X02DA of the Swiss Light Source. The first three datasets contain the scanned volume reconstruction during unsteady-state dynamics, while last two datasets contain the same scanned volume during steady state dynamics.

Abstract *

In the related publication to these data sets, we explore the flow dynamics for two-phase flow in a porous medium (a bioclastic carbonate rock). We use state-of-the-art synchrotron X-ray tomography to capture the fluid dynamics within the pore space, with a scan time of 1 second and a temporal resolution (scan repetition rate) of 2 s. The rock sample was initially saturated with brine (DI water doped with 15%wt. KI) before brine and nitrogen gas were injected simultaneously. As the gas establishes a path through the pore space, the flow dynamics are transient. Eventually, an equilibrium is established, where the gas saturation oscillates about a constant mean value; this is referred to as steady state. There are 5 data sets, 3 of which capture the unsteady state dynamics, and 2 of which capture the steady state dynamics. The images were captured with a voxel size of 2.75 µm3. In these data sets we observe that the pore scale dynamics evolve as the macroscopic flow transitions from unsteady state to steady state. We observe that the saturation of the gas plateaus out before the differential pressure across the core. This suggests that gas phase is more mobile during unsteady state.

download link

https://doi2.psi.ch/datasets/das/work/p17/p17614/Data10/disk1/

related publications

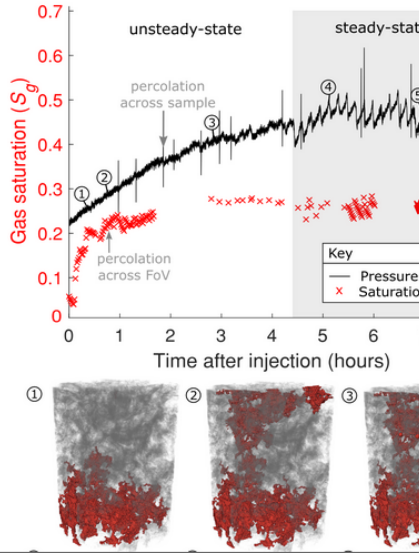
C. Spurin, T. Bultreys, M. Rücker, G. Garfi, C. M. Schlepütz, V. Novak, S. Berg, M. J. Blunt, and S. Krevor. Real-Time Imaging Reveals Distinct Pore-Scale Dynamics During Transient and Equilibrium Subsurface Multiphase Flow. Water Resour. Res. 56, 433 (2020). https://doi.org/10.1029/2020WR028287

Update Published Data | Cancel

Drop a file here

or

Browse



Published Data = List of Datasets + Metadata + DOI

Real-Time Imaging Reveals Distinct Pore-Scale Dynamics During Transient and Equilibrium Subsurface Multiphase Flow

Catherine Spurin, Tom Bultreys, Maja Rücker, Gaetano Garfi, Christian M. Schlepütz, Vladimir Novak, Steffen Berg, Martin J. Blunt, Samuel Krevor; PSI (2021)

Abstract

In the related publication to these data sets, we explore the flow dynamics for two-phase flow in a porous medium (a bioclastic carbonate rock). We use state-ofthe-art synchrotron X-ray tomography to capture the fluid dynamics within the pore space, with a scan time of 1 second and a temporal resolution (scan repetition rate) of 2 s. The rock sample was initially saturated with brine (DI water doped with 15%wt. KI) before brine and nitrogen gas were injected simultaneously. As the gas establishes a path through the pore space, the flow dynamics are transient. Eventually, an equilibrium is established, where the gas saturation oscillates about a constant mean value; this is referred to as steady state. There are 5 data sets, 3 of which capture the unsteady state dynamics, and 2 of which capture the steady state dynamics. The images were captured with a voxel size of 2.75 µm3. In these data sets we observe that the pore scale dynamics evolve as the macroscopic flow transitions from unsteady state to steady state. We observe that the saturation of the gas plateaus out before the differential pressure across the core. This suggests that gas phase is more mobile during unsteady state.

Publication details

DOI https://doi.org/10.16907/46a4d882-4dec-4097-8289-8f6311a4aa36

Resource Type derived

Related Publications C. Spurin, T. Bultreys, M. Rücker, G. Garfi, C. M. Schlepütz, V. Novak, S. Berg, M. J. Blunt, and S. Krevor. Real-Time Imaging Reveals Distinct Pore-Scale Dynamics During Transient and Equilibrium Subsurface Multiphase Flow. Water Resour. Res. 56, 433 (2020). https://doi.org/10.1029/2020WR028287

Datasets

Description This published data collection contains five datasets obtained by X-ray tomographic microscopy of a carbonate rock sample 5 mm in diameter and 20 mm in length. Both brine and nitrogen gas are injected into the sample at a total flow rate of 0.1 ml/min (the brine made up 85% of this total flow rate). Data were collected and processed at the TOMCAT beamline X02DA of the Swiss Light Source. The first three datasets contain the scanned volume reconstruction during unsteady-state dynamics, while last two datasets contain the same scanned volume during steady state dynamics.

20.500.11935/64af1e80-c539-4a90-a051-b7db5e6e714d

20.500.11935/e151f4d6-198a-47e7-ac63-0b258ef36ed3

20.500.11935/441fdcd9-fa0c-491c-b102-d114cc841609

20.500.11935/b9782901-be3b-40fe-91d0-3e0a784337c4

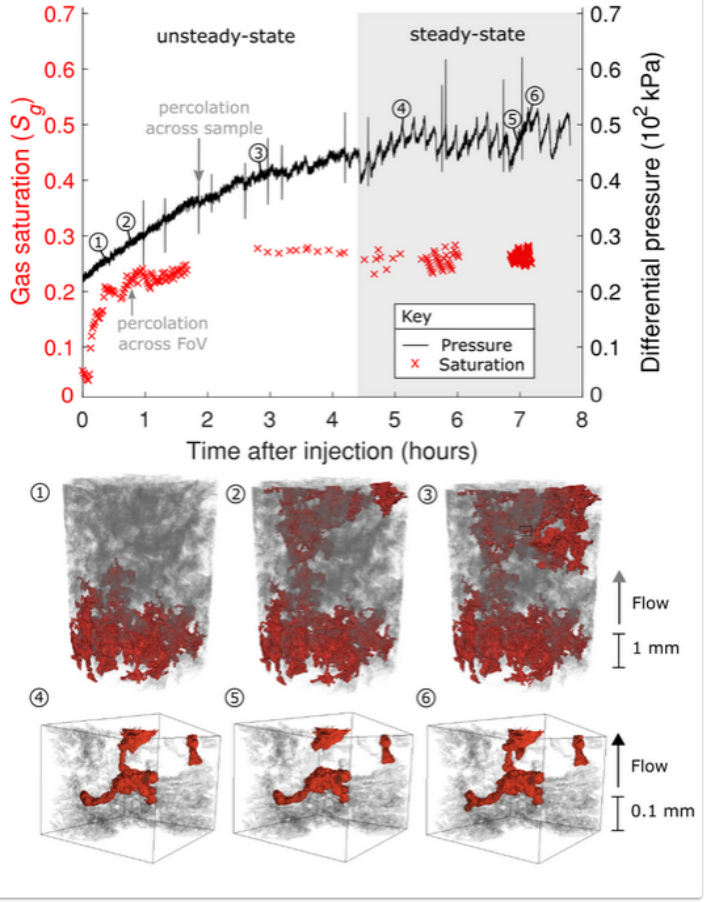
20.500.11935/5899a0eb-7e3b-451f-b01e-17ddfc0d0938

Actions

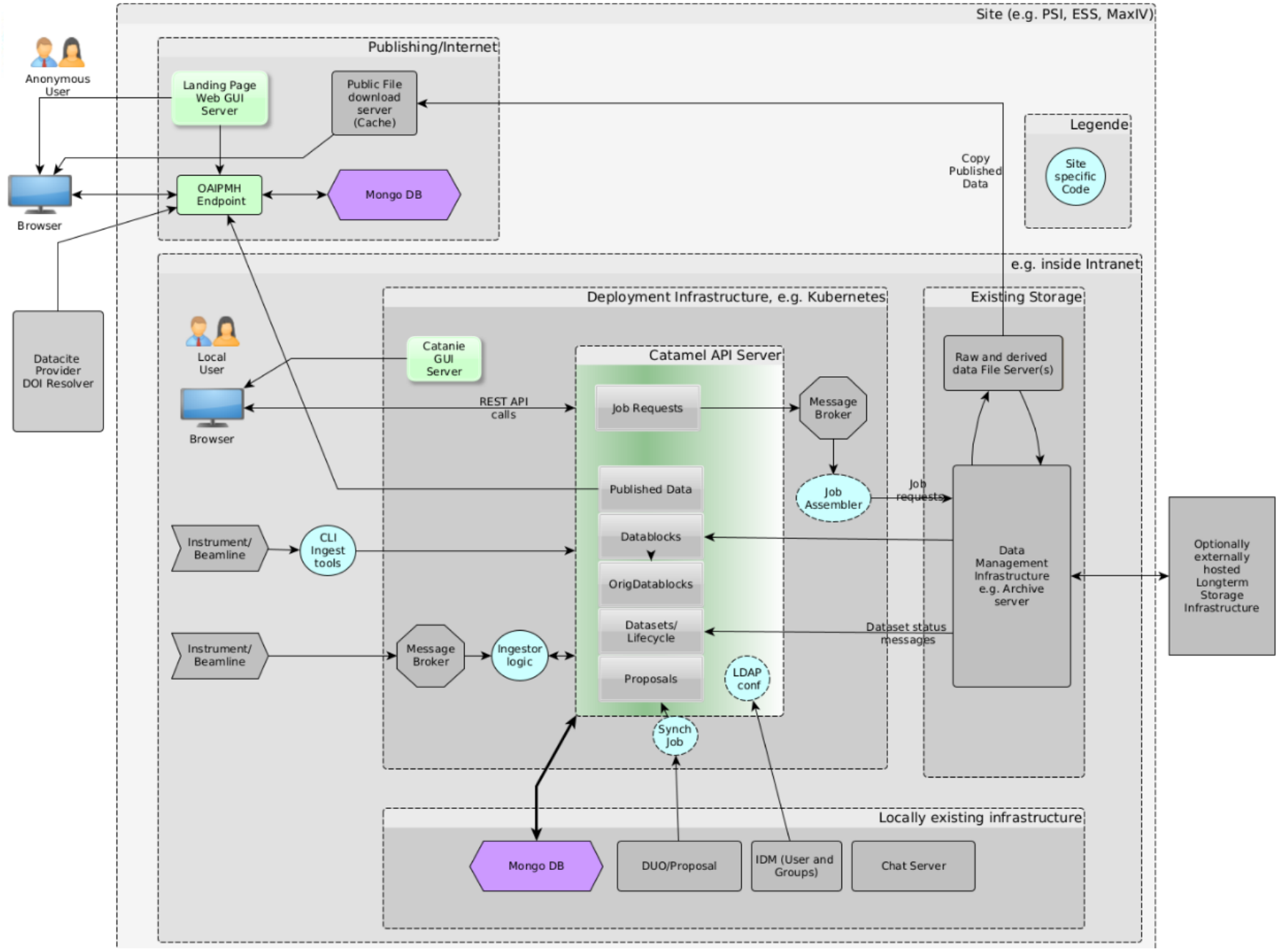
To access the data associated with this DOI click below and follow the instructions

Access Data

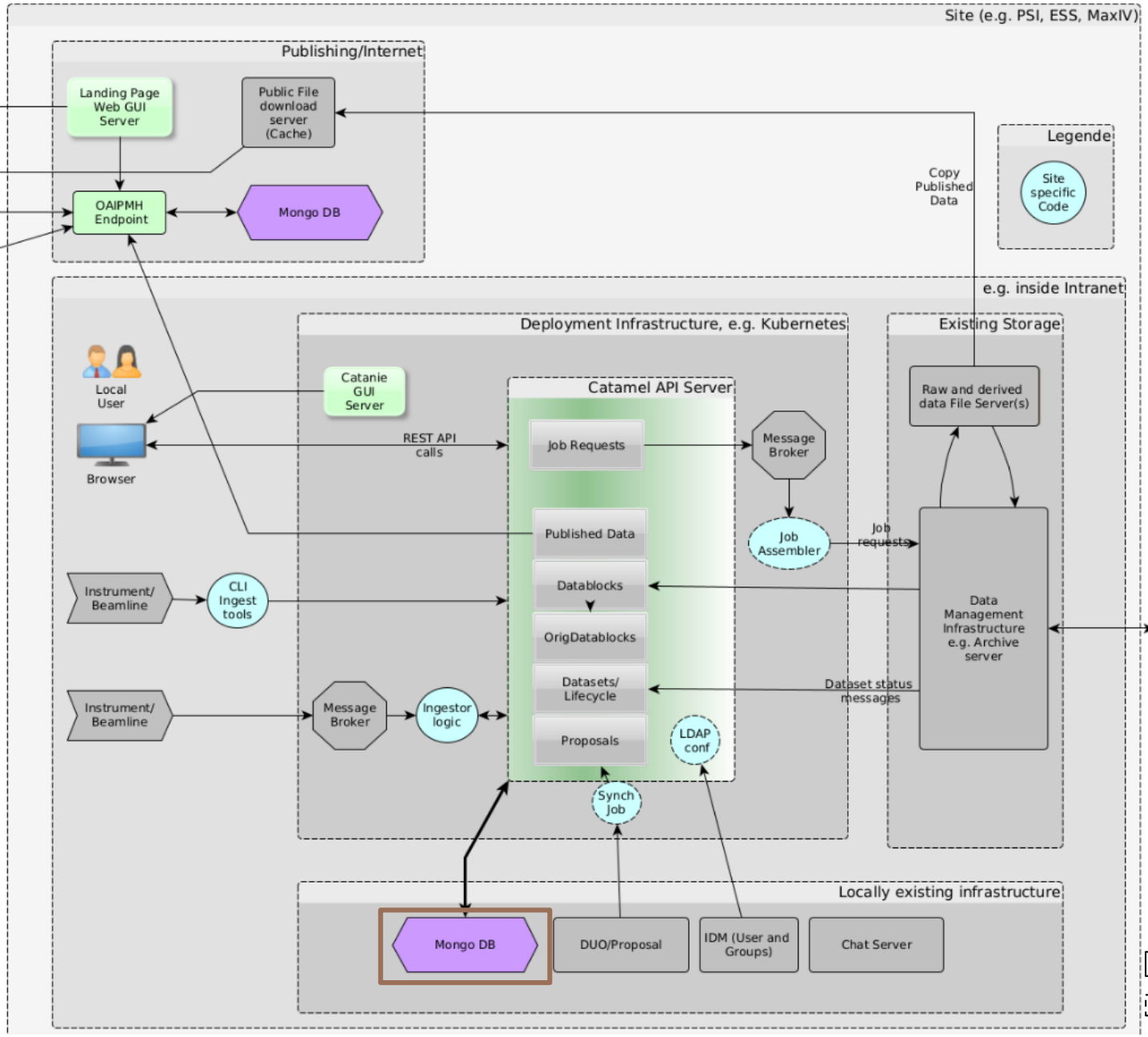
This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License



Data Catalog Architecture and Integration

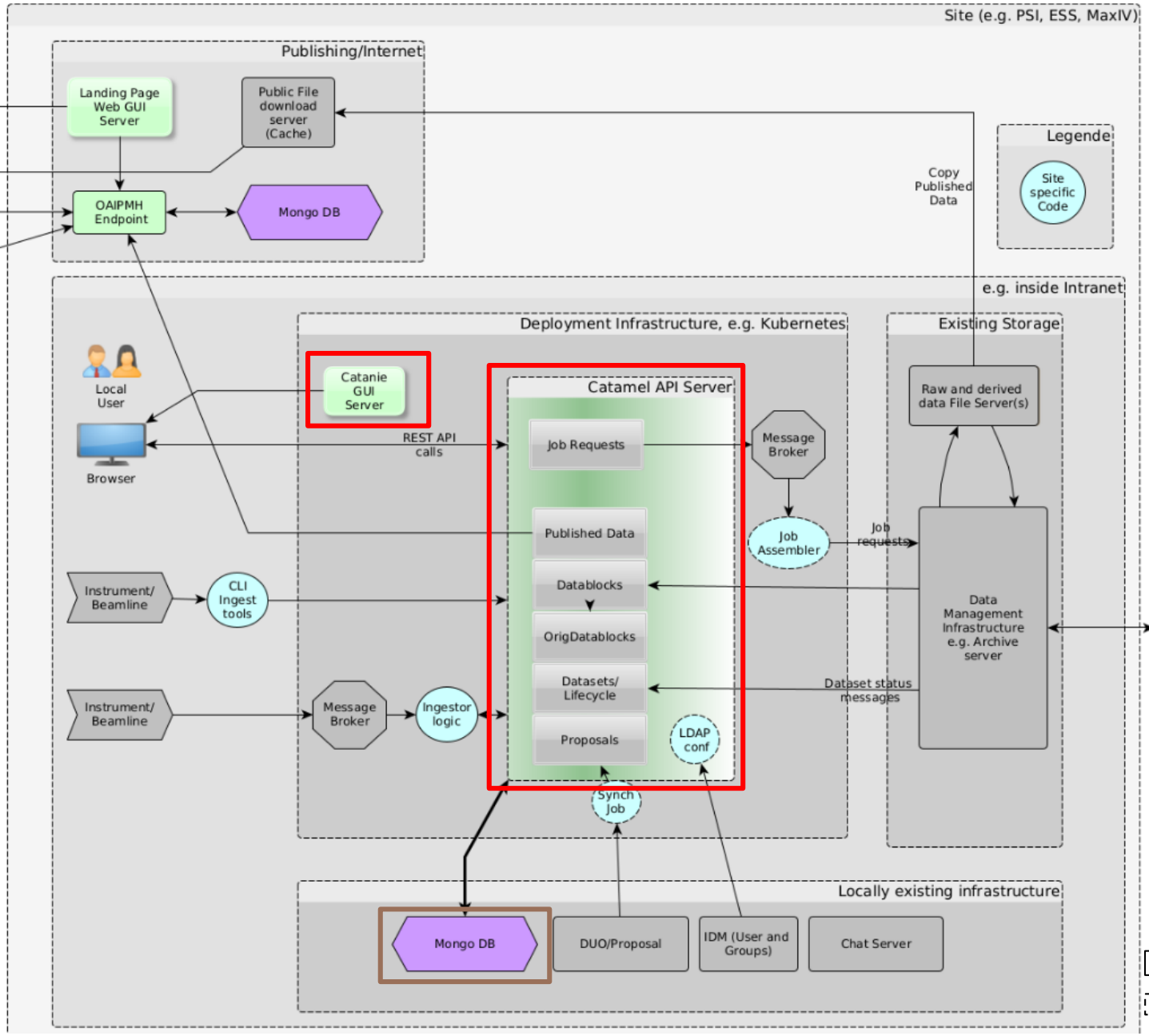


Data Catalog Architecture and Integration



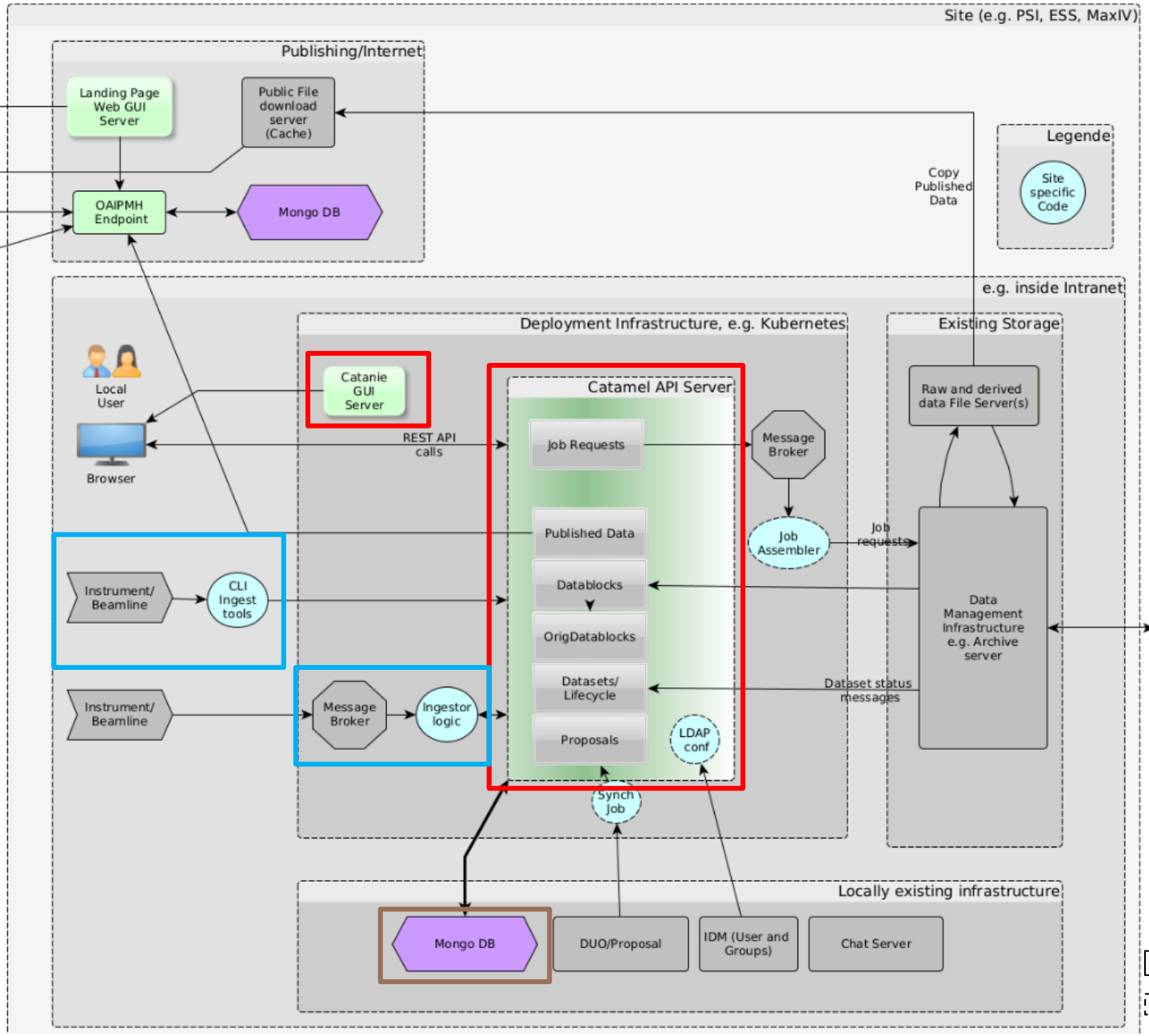
1. Setup Mongo DB

Data Catalog Architecture and Integration



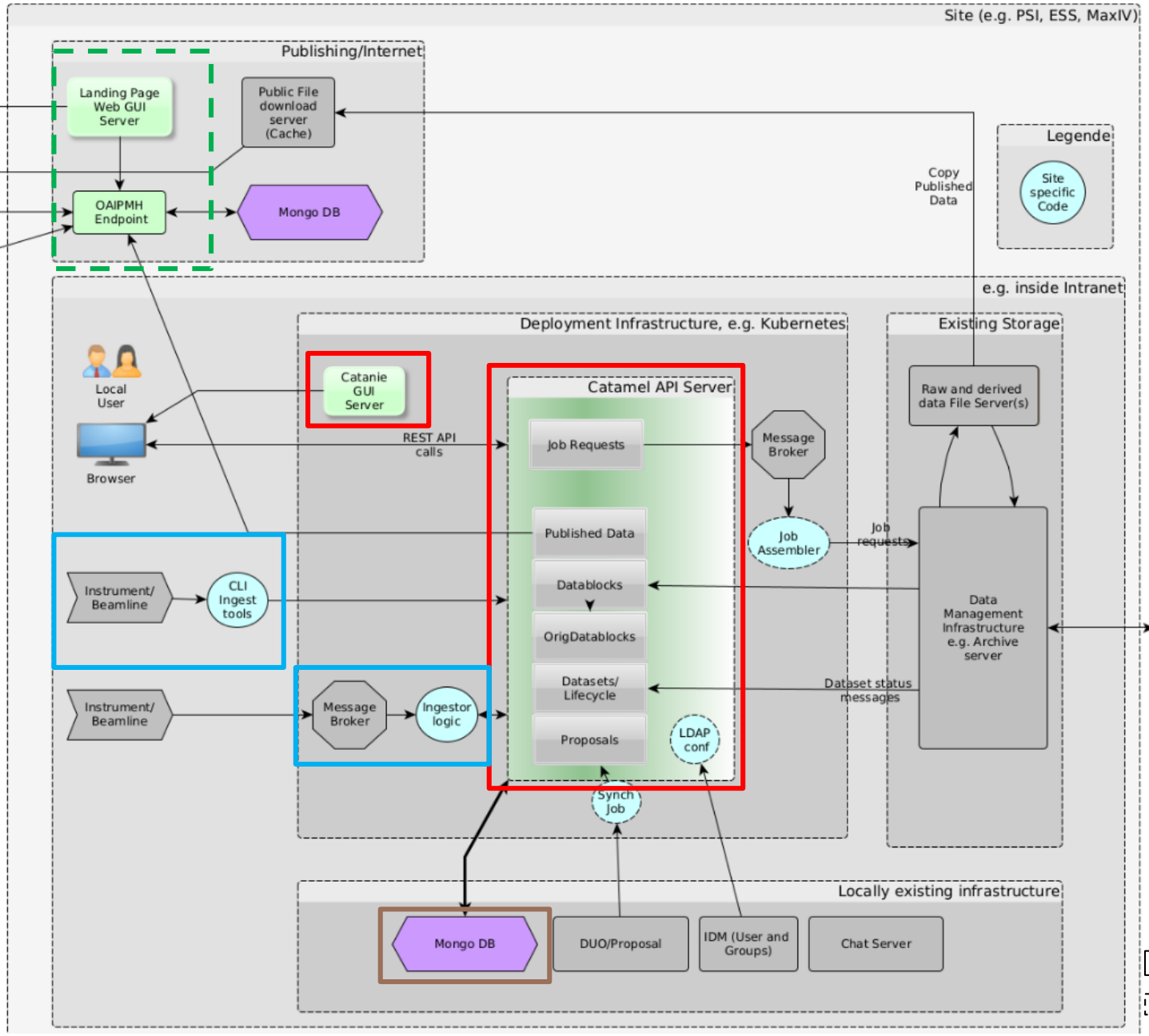
1. Setup Mongo DB
2. Setup SciCat

Data Catalog Architecture and Integration



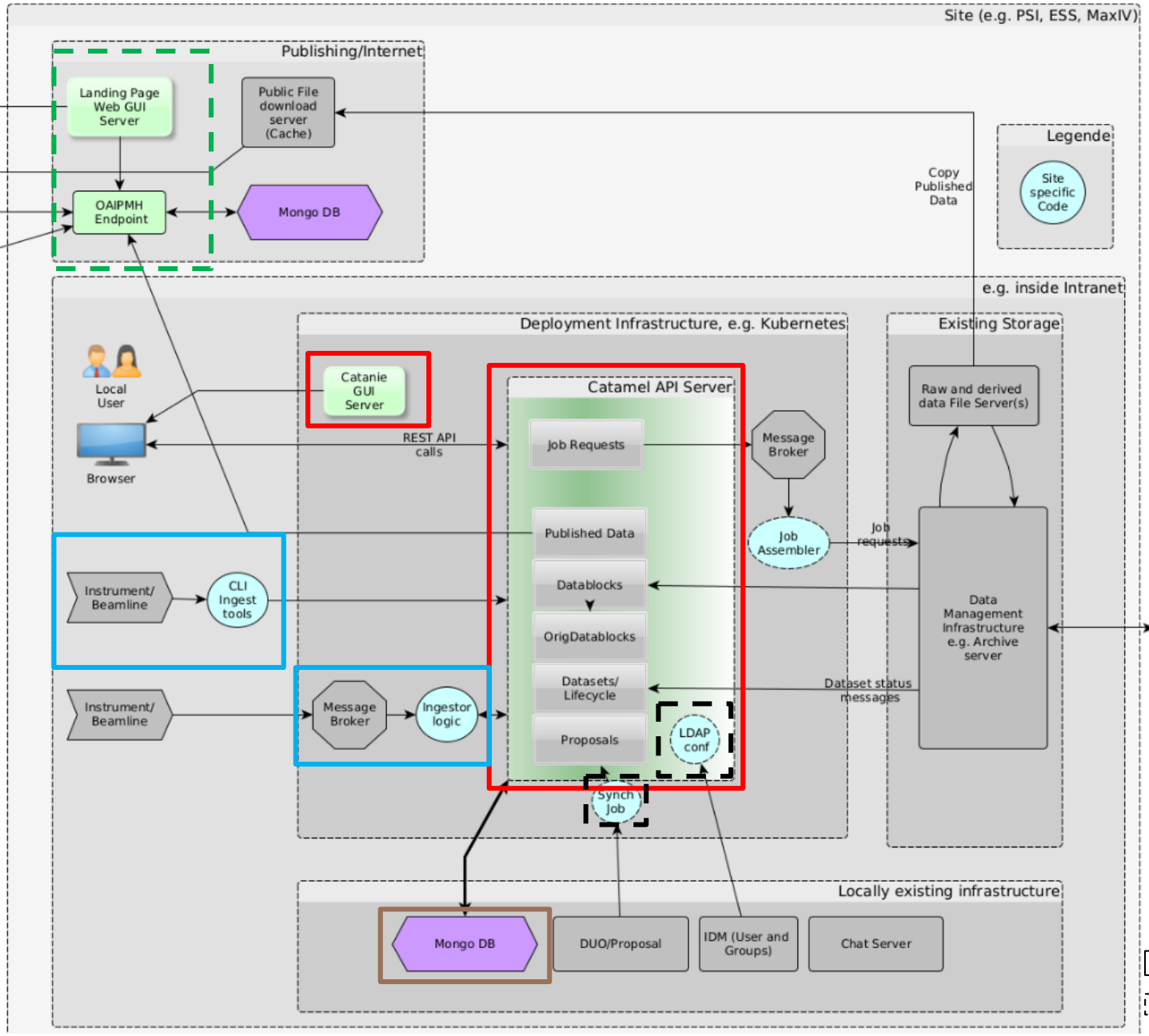
1. Setup Mongo DB
2. Setup SciCat
3. Define ingest process

Data Catalog Architecture and Integration



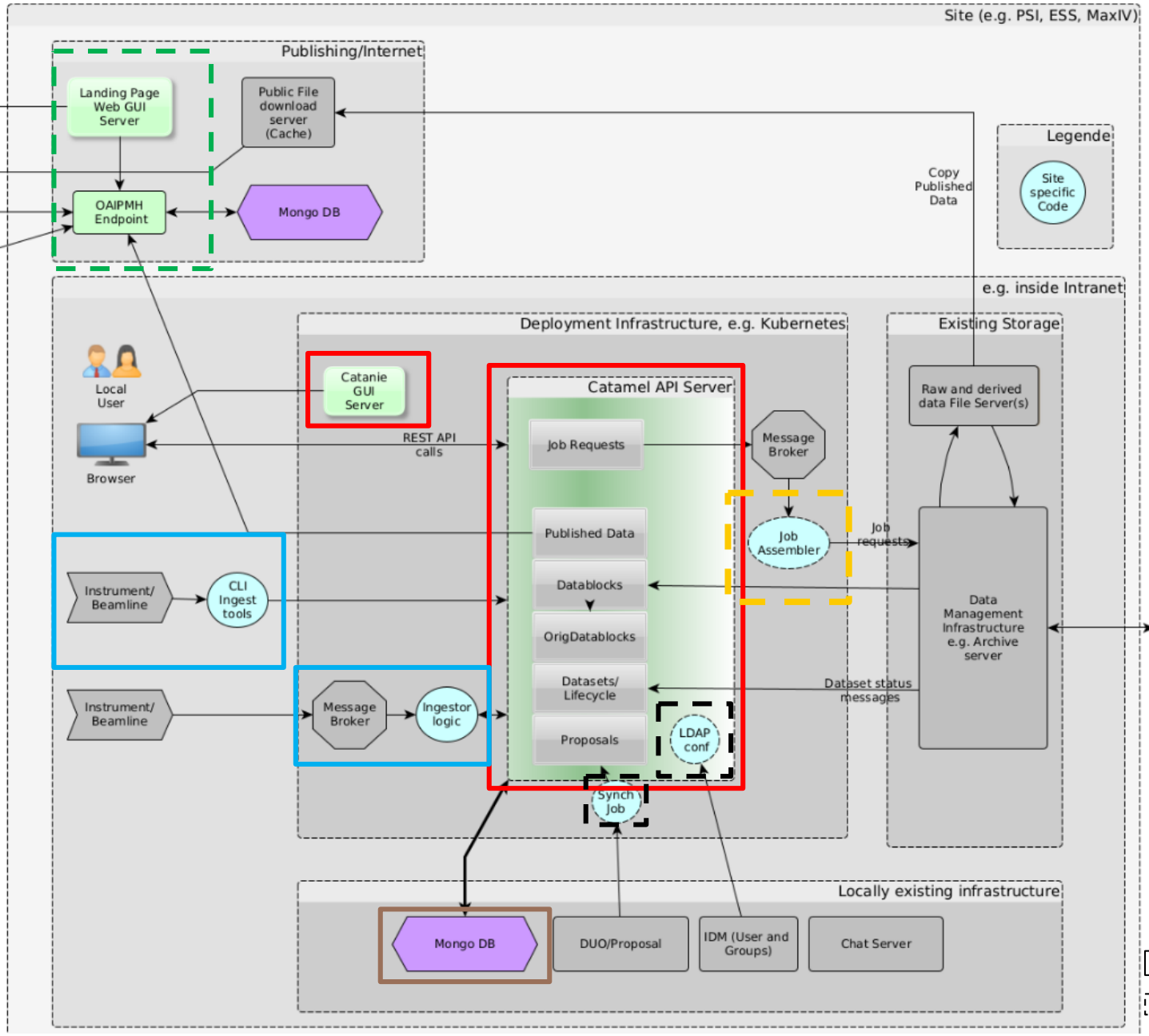
1. Setup Mongo DB
2. Setup SciCat
3. Define ingest process
4. Setup additional SciCat packages for published data

Data Catalog Architecture and Integration



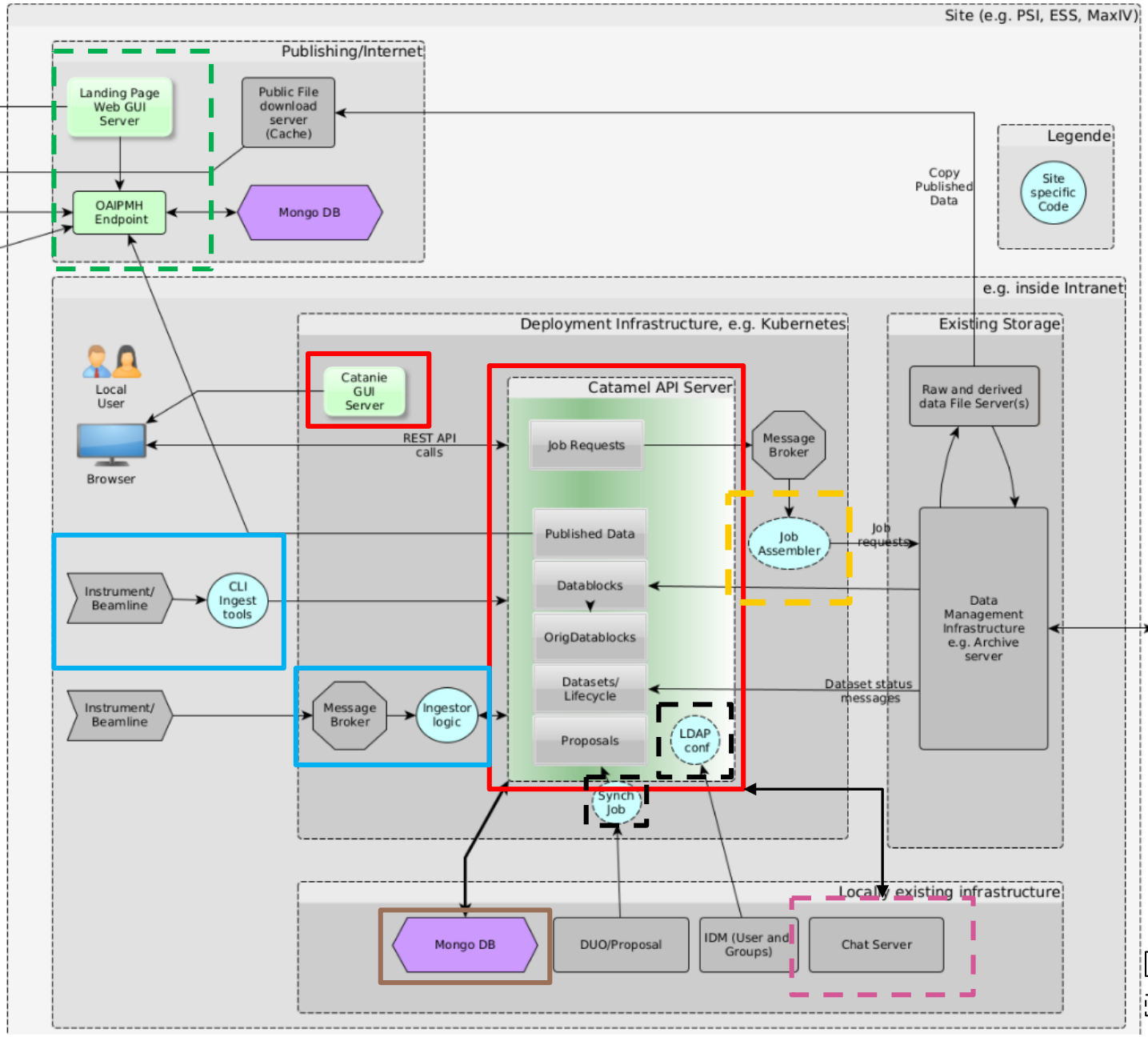
1. Setup Mongo DB
2. Setup SciCat
3. Define ingest process
4. Setup additional SciCat packages for published data
5. Connect to IDM/ proposal system

Data Catalog Architecture and Integration



1. Setup Mongo DB
2. Setup SciCat
3. Define ingest process
4. Setup additional SciCat packages for published data
5. Connect to IDM/ proposal system
6. Connect to long term storage

Data Catalog Architecture and Integration



1. Setup Mongo DB
2. Setup SciCat
3. Define ingest process
4. Setup additional SciCat packages for published data
5. Connect to IDM/ proposal system
6. Connect to long term storage
7. Interface to ELN or analysis workflow

One-click Reference Installation

README.md

SciCat

Files for running SciCat with docker-compose.

Steps

1. Clone the repository

```
git clone https://github.com/SciCatProject/scicatlive.git
```

2. Run with the following command inside the directory

```
docker-compose up -d
```

3. SciCat will now be available on http://localhost. The Loopback API explorer of catamel is available at http://localhost/explorer/, the one for the search-api at http://localhost/panosc-explorer/.

Add Your Local Configuration

1. Add your local configuration to config.local.js
2. Uncomment the `volumes:` line and the line containing `config.local.js` in the catamel service section in docker-compose.yaml (if commented)
3. Restart the docker containers

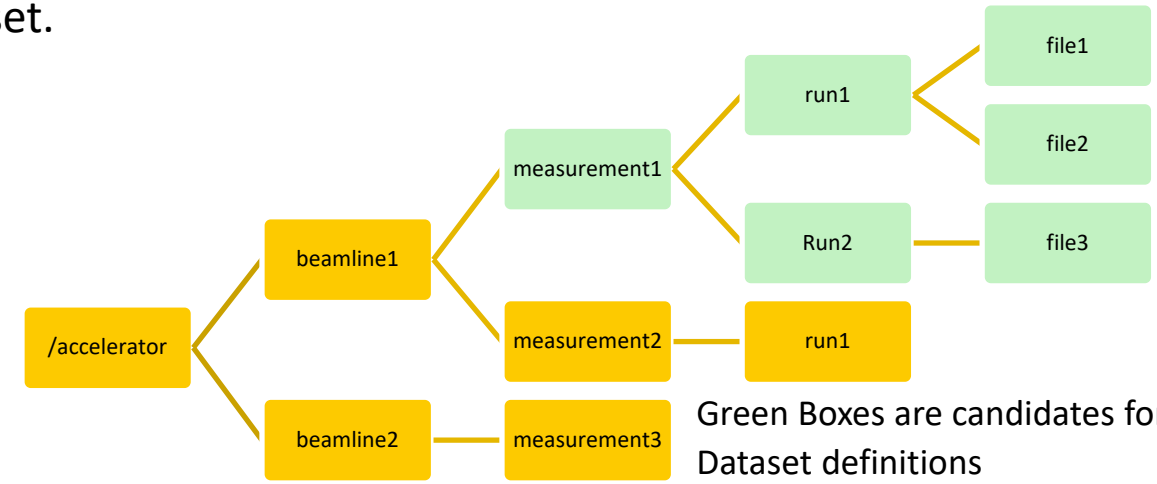
Add LDAP Authentication

1. Add your LDAP configuration to providers.json
2. Uncomment the `volumes:` line and the line containing `providers.json` in the catamel service section in docker-compose.yaml
3. Restart the docker containers

https://github.com/SciCatProject/scicatlive

Metadata ingestion: 1. start e.g. from existing folder structure to define Datasets

- Datasets are the smallest unit for archiving, retrieving and publication
- Create them by defining a list of files, e.g. for raw data list all the files that logically belong to a measurement/data taking run, or any other criteria. For example: define all the files in the same directory (e.g. measurement1) as part of one dataset.



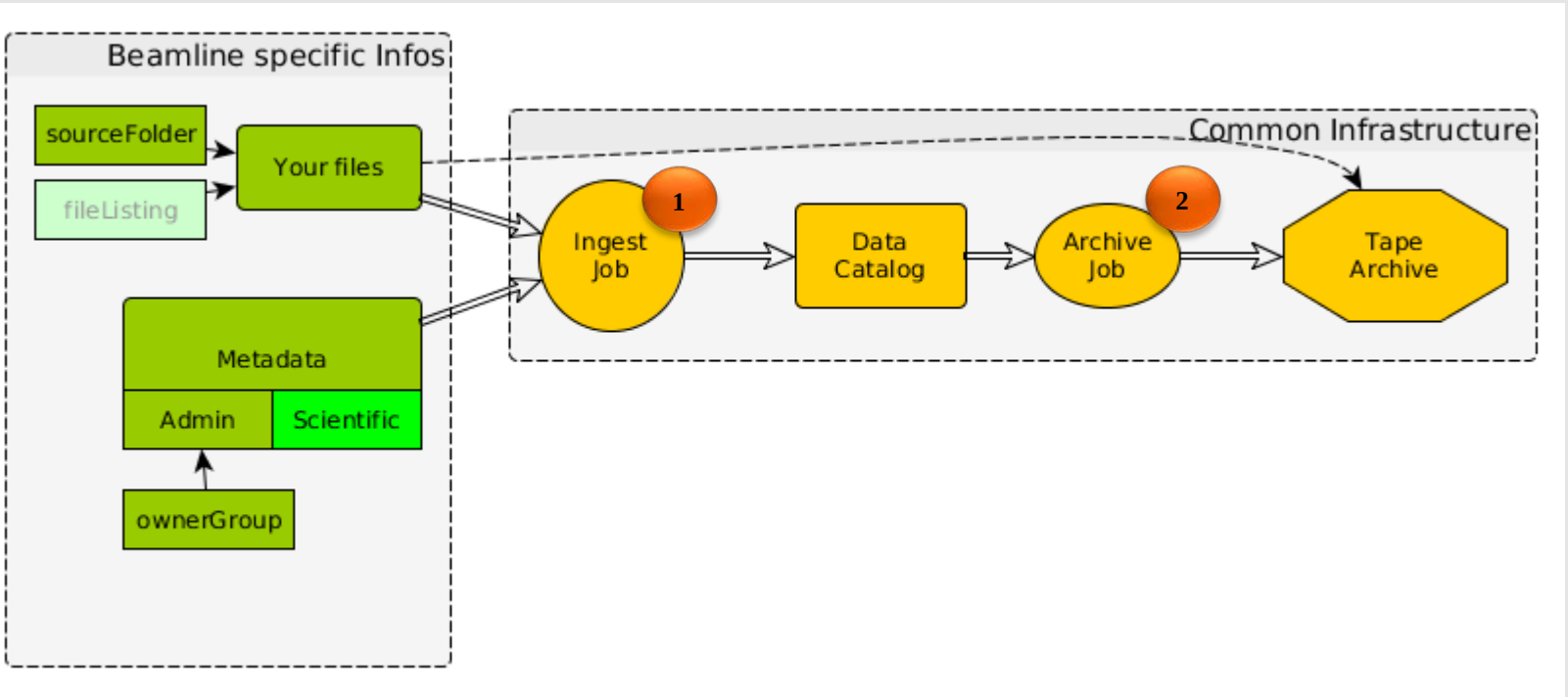
- In addition to “raw” Datasets you can create “derived” datasets containing the results of your analysis derived from the raw data. This ingest step is usually done by the user pursuing the analysis

2. Define Scientific Metadata

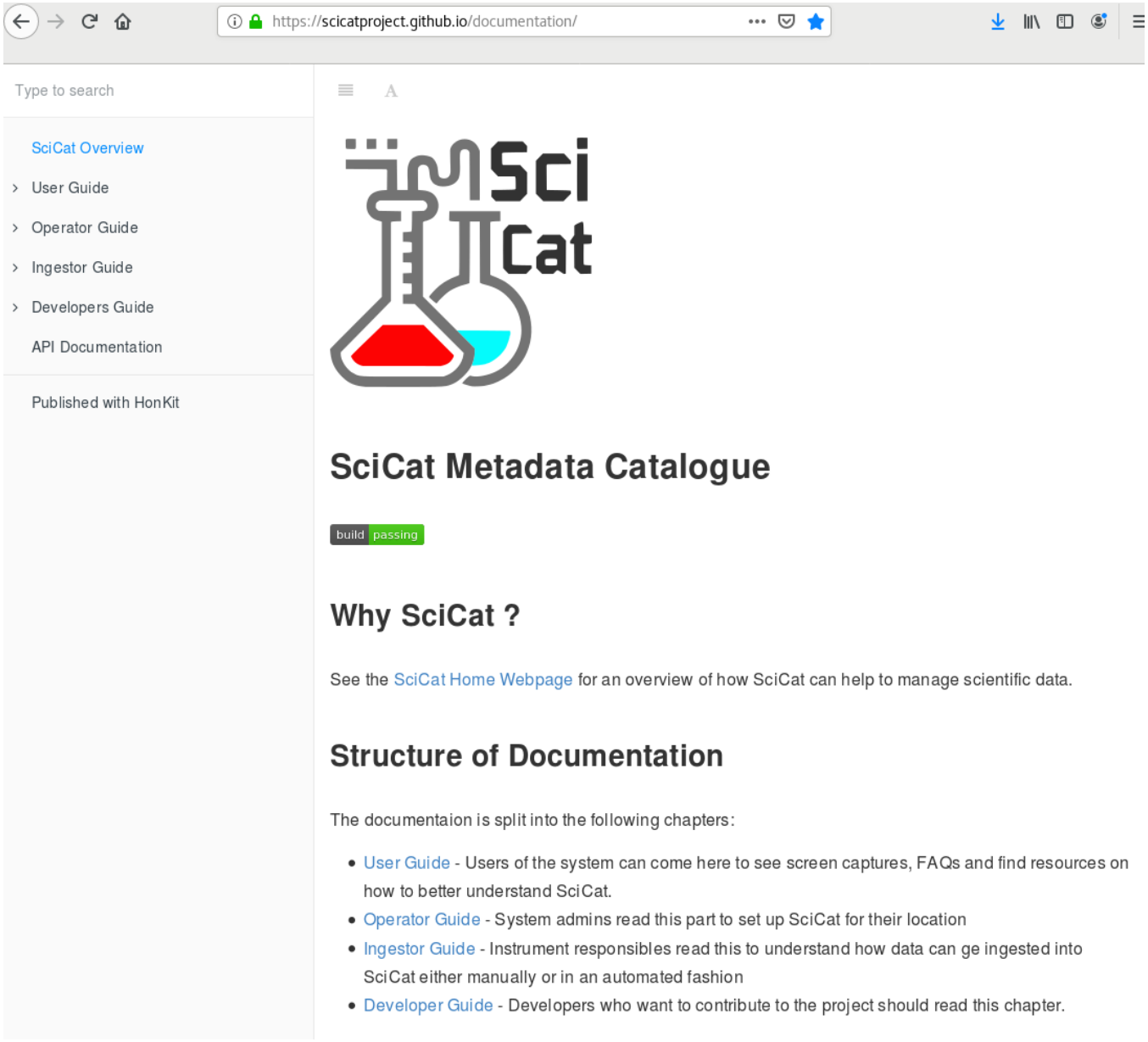
- The definition of scientific meta data is fully flexible.
- Ideally following a standard if it exists, e.g. NeXus based HDF5 files, extracted from instrument.
- Example:

```
“scientificMetadata": {
    "beamlineParameters": {
        ”monostripe": "Ru/C",
        ”ring_current": {
            "value": 0.402246,
            "units": "A"
        },
        ”beam_energy": {
            "value": 22595,
            "units": "eV"
        }
    },
    "detectorParameters": {
        ”objective": 20,
        ”scintillator": "LAG 20um",
        ”exposure_time": {
            "value": 0.4,
            "units": "s"
        }
    }…
}
```

3. Adding data is a simple (two step) procedure



Full documentation for users and operators

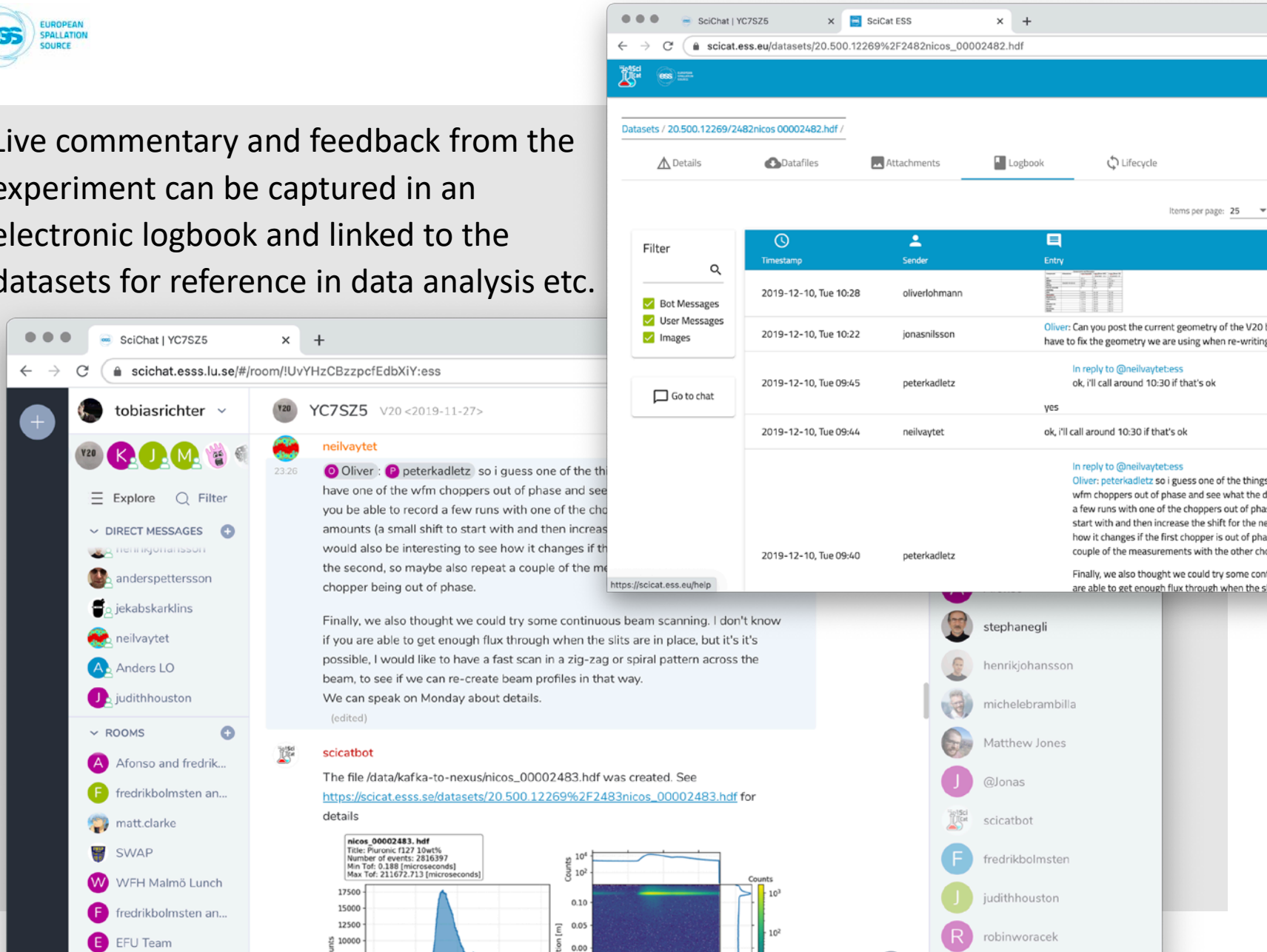


Electronic Lab Notebook (ELN) integration

- That is a big topic which can easily fill a separate talk
- There are many existing ELN solutions already, often very specialized and tailored for a specific use case or scientific domain.
- Strategy of SciCat to meet the challenges:
  - try to be a «good citizen» in the landscape of existing applications: this means be designed with **easy interfacing** in mind from the start.
  - API centric approach, e.g.: Python SDK
  - Message driven interfacing, loose coupling of components
  - Support for attachments
  - Flexible metadata structure, therefore open to accept new ad-hoc informations from ELNs
- Example of such integrations:
  - SciChat: integrate SciCat API with an existing Matrix/element based chat system
  - SciLog: This is in early state,WIP at PSI, optimized for beamline usage
  - Jupyter Notebooks

Link with SciChat Electronic Logbook

Live commentary and feedback from the experiment can be captured in an electronic logbook and linked to the datasets for reference in data analysis etc.



Sites using or planning to use SciCat and contact information

- European Spallation Source (Sweden/Denmark)
- Paul Scherrer Institut (Switzerland)
- MaxIV Laboratory (Sweden)
- Advanced Light Source - micro-CT (USA)
- Bundesamt für Materialwirtschaft (Germany)
- Rosalind Franklin Institute (United Kingdom)
- Shanghai Facility SSRF (China)
- Beijing High Energy Photon Source (China)
- Deutsches Elektronen-Synchrotron - PETRA (Germany)
- SOLEIL (France)

- Several of the sites contribute actively to the SciCat codebase

- **Fortnightly developers'/operators' meetings + internal chat channel**
- **Please get in contact for questions about SciCat onboarding at this email address: carlo.minotti@psi.ch or max.novelli@ess.eu**