User Portal, Metadata Catalog and Logbook at European XFEL

19th May, 2021

Luís Maia,

On behalf of European XFEL

Agenda

🟧 **<u>Introduction</u>**

    🟦 Context

    🟦 Purpose
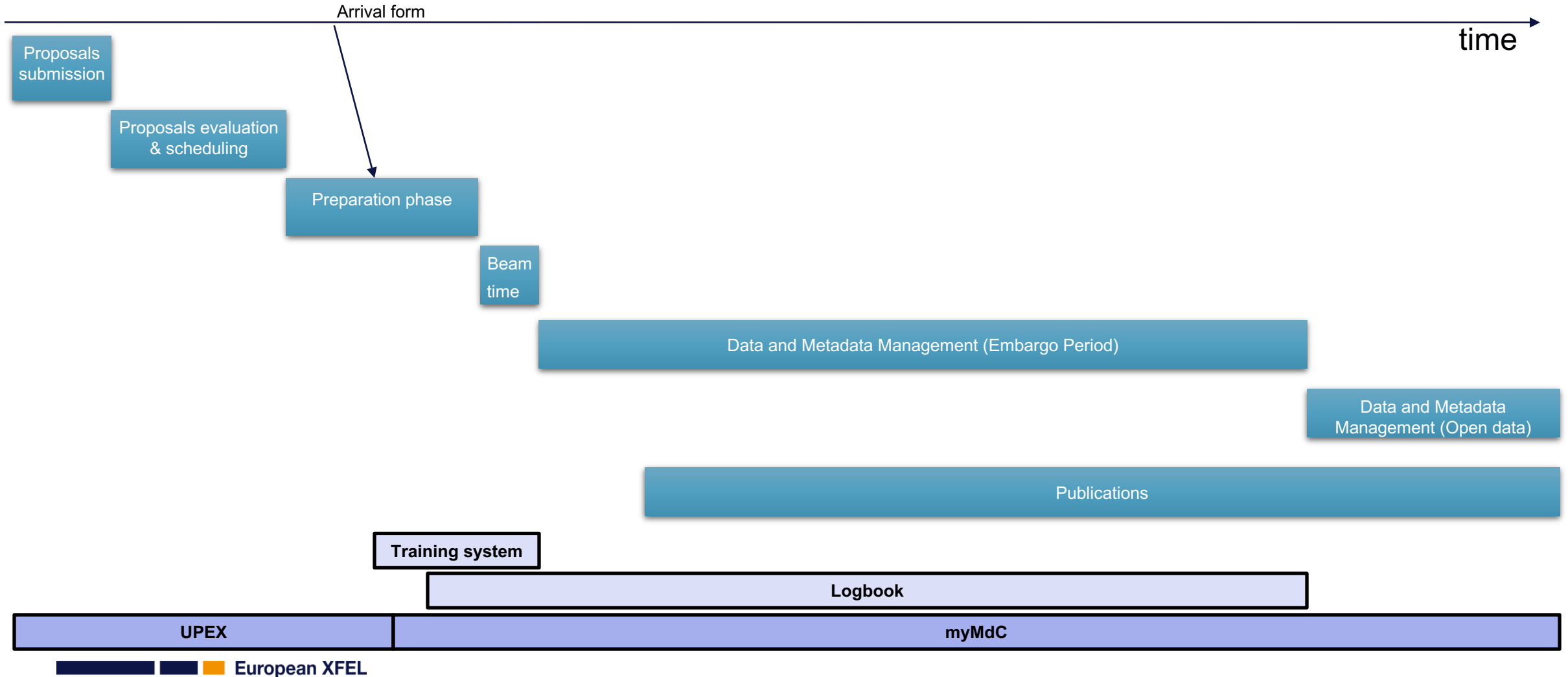
    🟦 Landscape & Architecture

🟧 Governance & Resources

🟧 myMdC application (**optional**)

    🟦 Roles and permissions

    🟦 Key features

European XFEL

Context
Users support (and Experiments lifecycle)

UPEX Purpose (I)

■ First entry point for users

■ Handling Policies & Acceptances

■ Provide a means to orchestrate the facility user public beam time proposal management process

> 7/23 operation mode

■ Organize and manage users personal data and existing accounts in a coherent way

■ IAM
  Personal data

■ Zerberus
  Sanction lists

■ Registry
  Users accounts password

■ MyMdC
  Users and internal proposals
  Proposals Experiment team & Samples sheet

■ Training / Access management system

XFEL.EU public user beam time proposal management process (V. 02 June 2015)

European XFEL
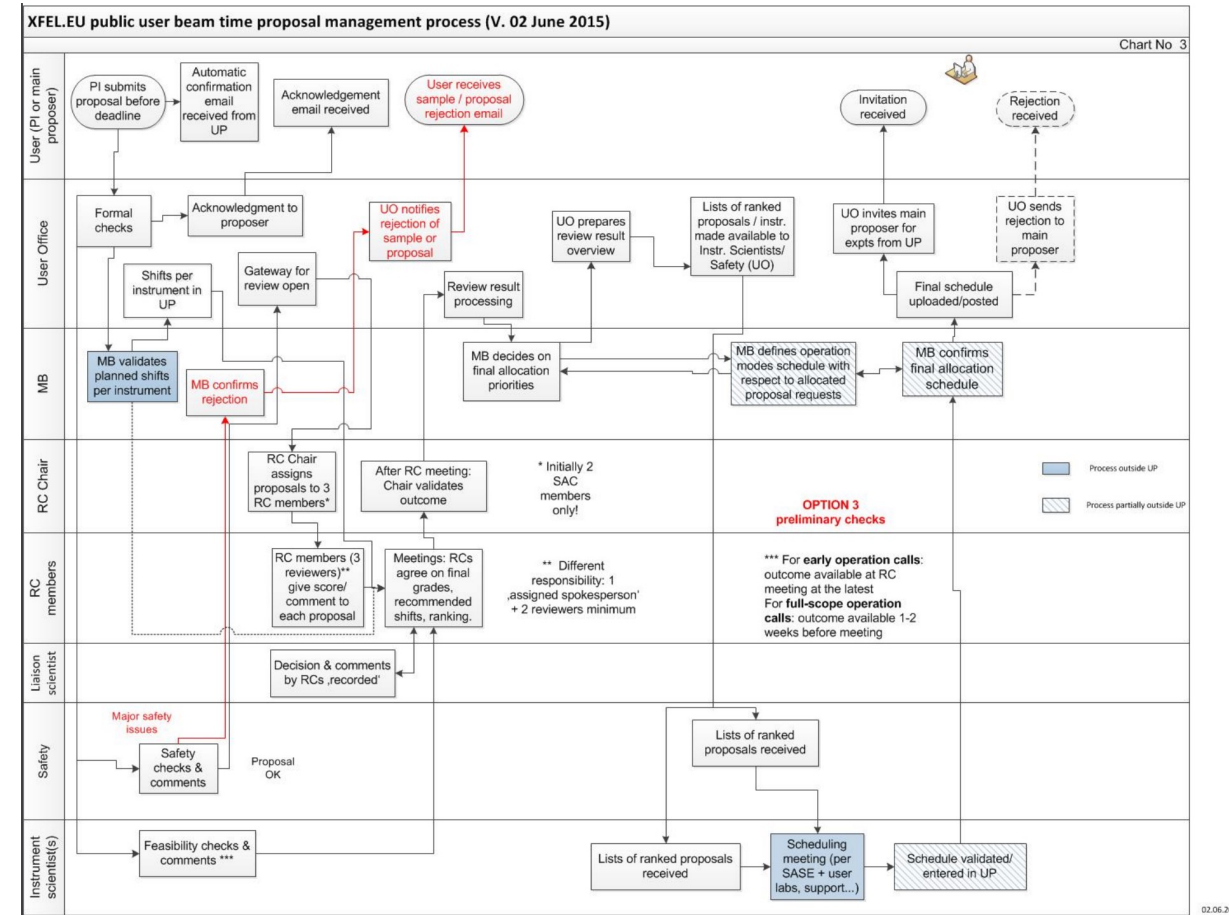
myMdC Purpose (I)

■ Provide a means of storing, retrieving and query raw and run based data in an organised way

■ Organize and manage data and metadata in a coherent way

■ Assist on the execution of the Data Management policies

■ Manage data and metadata authorisation and responsibilities (especially during the Embargo period)

■ Manage and restrict the GLOBAL unique entities used during data taking

■ Proposal number

■ Proposal path

■ Sample

■ Run type

European XFEL

myMdC **Purpose (II)**

Glue together different systems and services

- **UPEX**
  - Users and internal proposals
  - Proposals Experiment team & Samples sheet
  - Check and sync users accepted data policies
  - Trigger proposal team snapshot
  - Internal proposals "fast-track"
  - Users with access to Labs
  - Booked system integration

  > 7/23 operation mode

- **Registry**
  - Users accounts
  - Staff organisational group
  - Instrument support groups*

  > Upgrade Registry 2.0

- **LDAP**
  - Users Proposal groups

  > Upgrade to LDAP Gateway 2.0

- **GPFS**
  - Repository location
  - Universal folders management
  - Proposal folders creation
  - Runs data migration
  - Folders removal*

- **SLURM**
  - Priority queues reservation for beamtime*

- Zoom integration – proposal rooms

European XFEL

- **Data Audit and reconciliation***
  - Run Data location
  - Number of files, size, migration dates

- **Elog**
  - Elog logbook config file generation
  - Integration with Elog replacement tool

- **Karabo - DAQ**
  - Universal Run Management
  - Run tagging and basic details
  - Data Source Groups definition
  - Run Data Source Groups

- **Calibration Pipeline**
  - Trigger Processed data generation
  - Trigger Dark Runs generation and report download

- **calCat**
  - Dark Runs CCV integration
  - Document used CCV on Process data

- **Grid-FTP / Globus / Off-shore data exportation**
  - Repositories
  - Transfer Agent Off-shore data migration

- **DataCite / Tind.io**
  - Automatic DOI generation and publish
  - Public DOI end point

- **RESTful APIs**
  - Metadata Client and Metadata Tools python libraries
  - Document APIs using Swagger

- **Notifications***

- **Statistics**
  - Leading Scientists assessing delivered shifts

Landscape
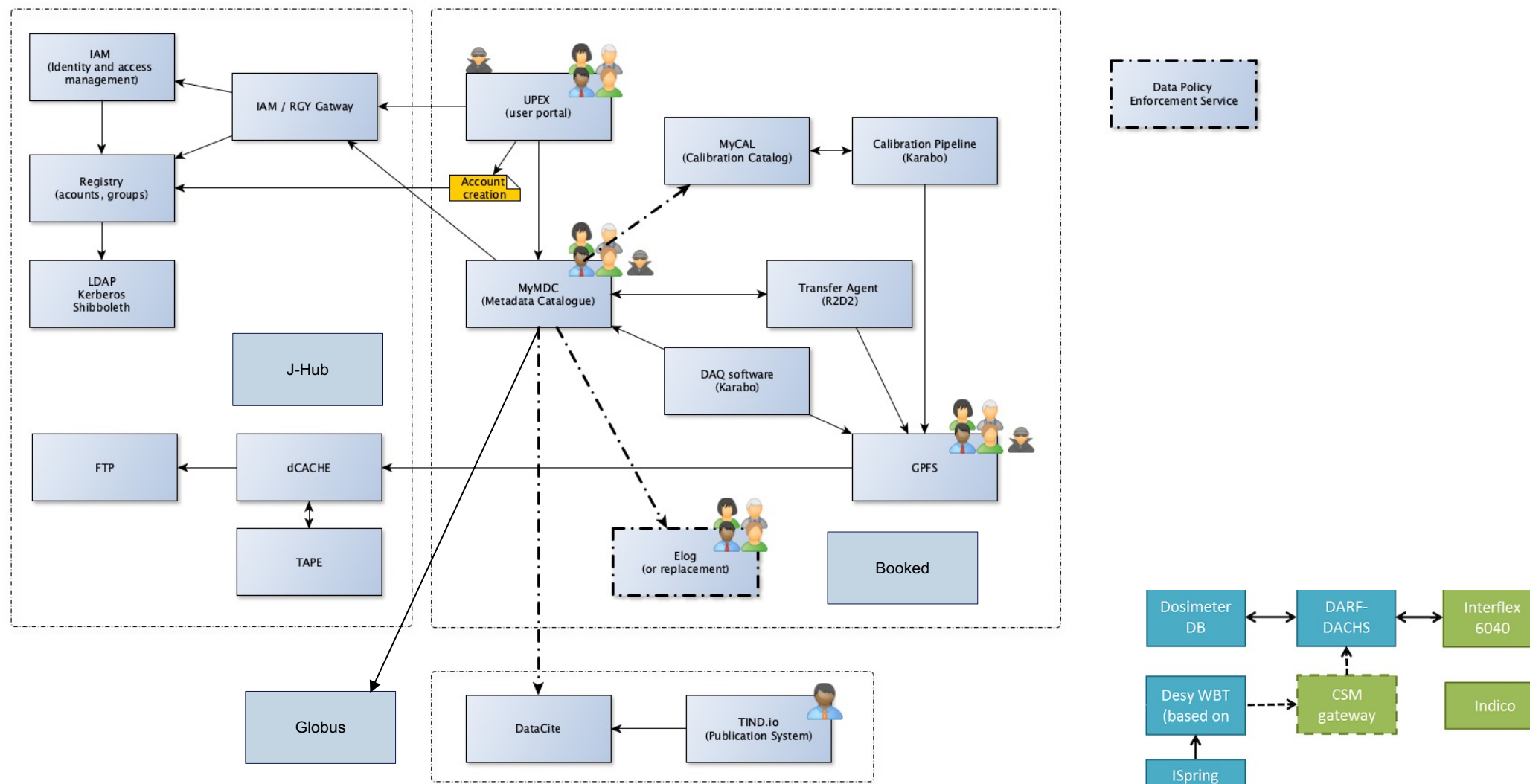
UPEX Architecture (I)

■ Based on DUO instance got from PSI in 2015

in.xfel.eu

DMZ
(www access)

European XFEL
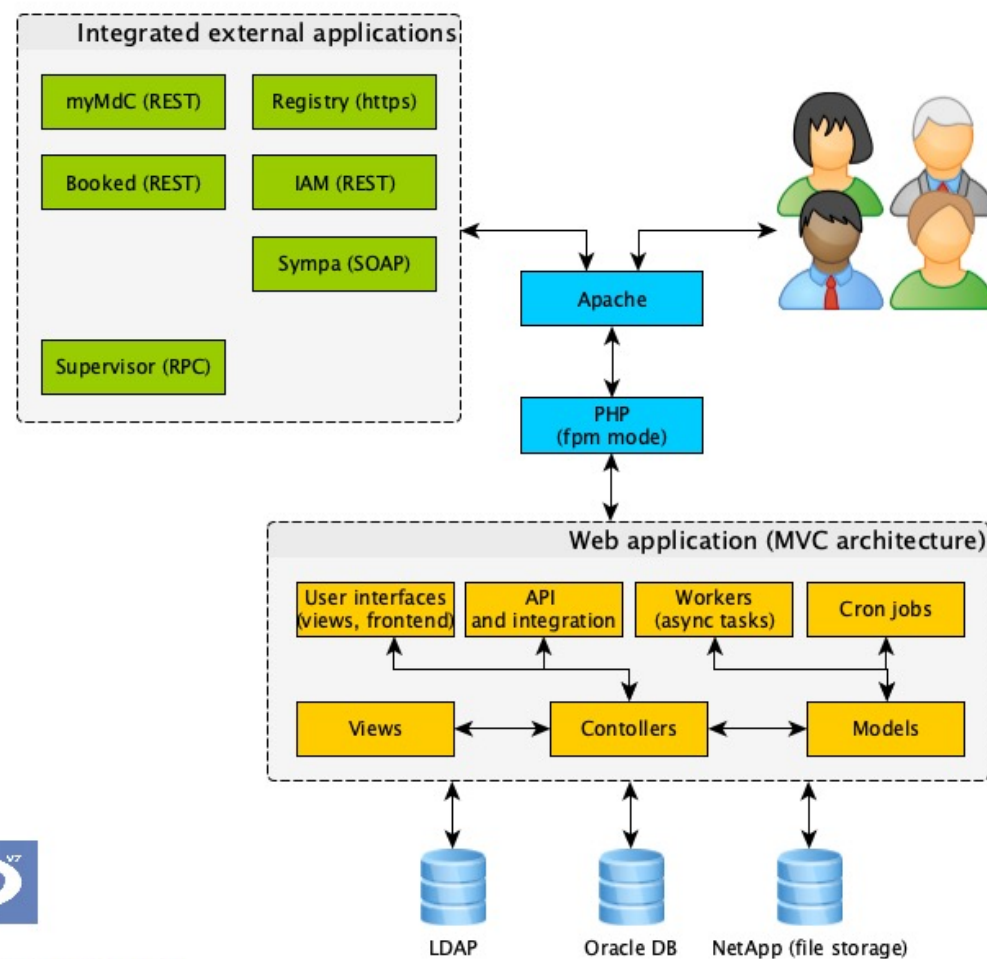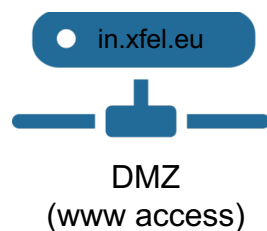
eLog **Architecture (I)**

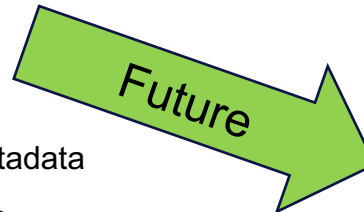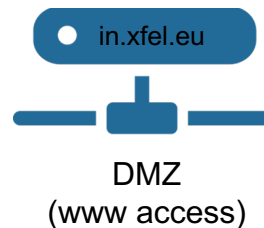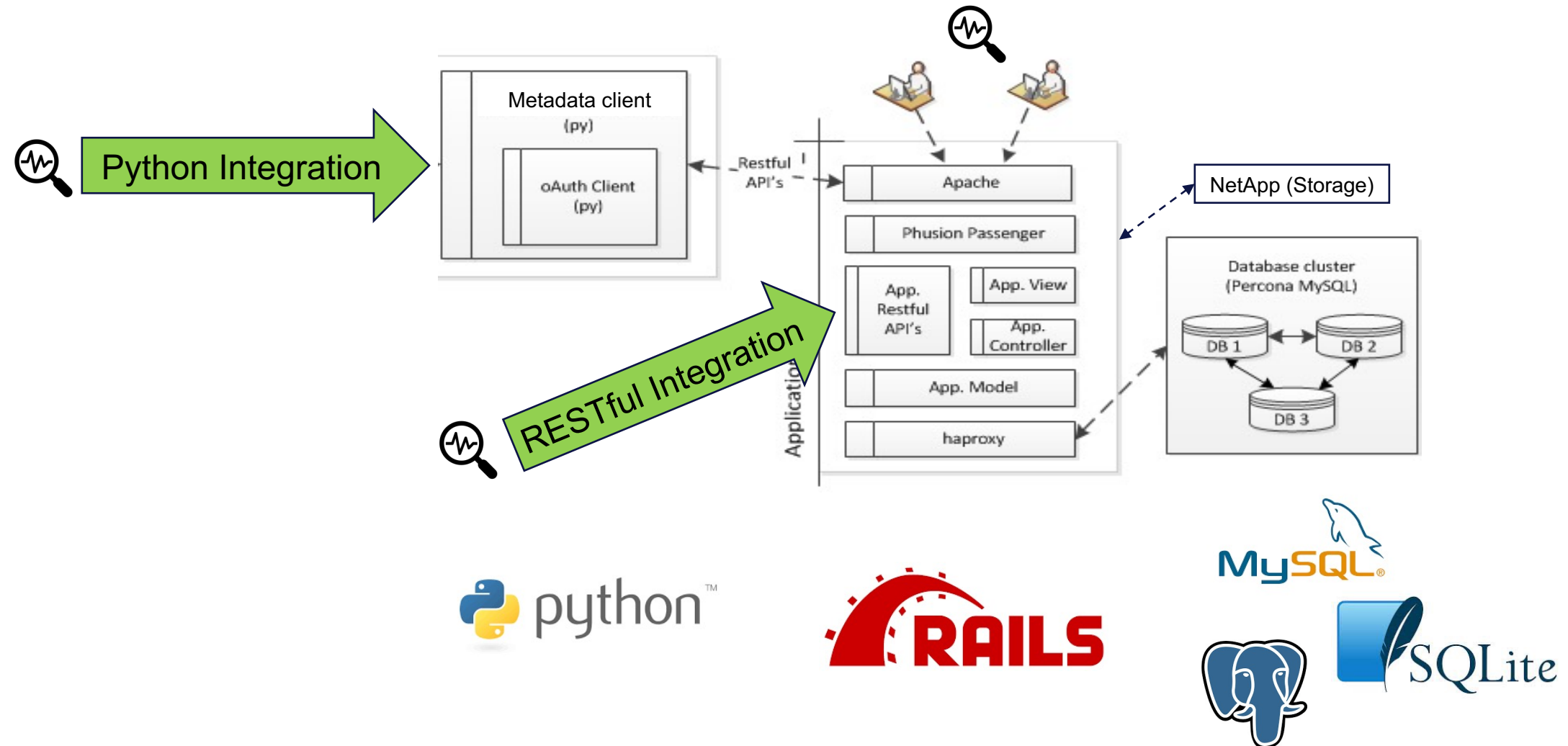- PSI Elog (https://elog.psi.ch/elog/) , adjusted visually (CSS) and configured for XFEL purposes

- There is no real integration with anything
  - Logbooks and Users created by script manually
  - Logbooks permissions managed by instrument experts
  - No integration with Control System and/or parameters metadata

in.xfel.eu

DMZ
(www access)

Future

Based on a chat tool
- Zulip
- Rocket chat

Suggestions

Specific tool
- ESRF logbook
- (...)
- New logbook

European XFEL
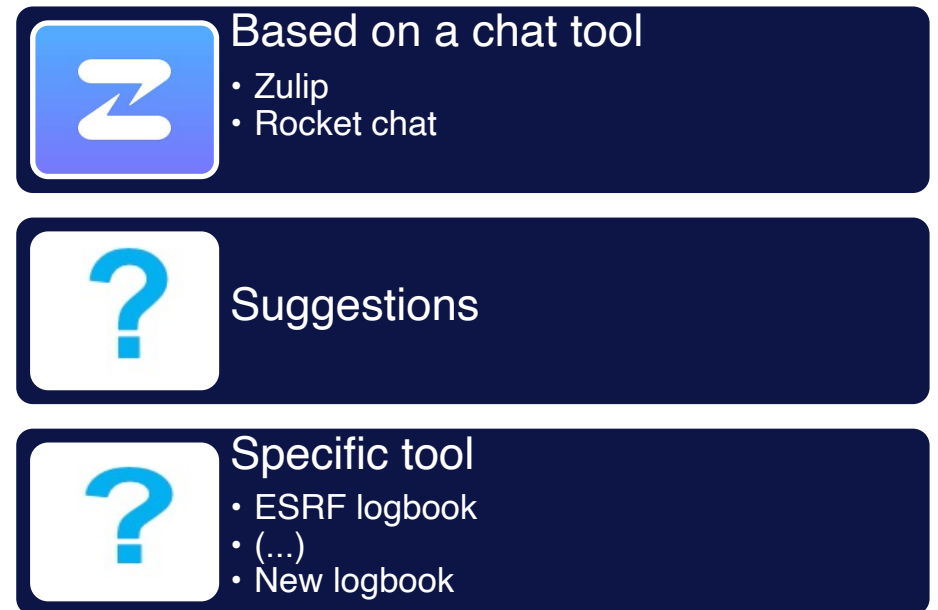
myMdC Architecture (I)

myMdC Architecture (II)

Local → Development → Test → Staging → Production

Xen™

GitLab

Ubuntu / CentOS

Capistrano

PERCONA

influxdb

in.xfel.eu
in.xfel.eu

DMZ
(www access)

SASE1
SASE2
SASE3

Control Network

European XFEL

Agenda

🟧 Introduction

   🟦 Context

   🟦 Purpose

   🟦 Landscape & Architecture

🟧 **<u>Governance & Resources</u>**

🟧 myMdC application (**optional**)

   🟦 Roles and permissions

   🟦 Key features

European XFEL

Governance & Resources

🟧 Product Owner

🟦 Several groups

🟧 Architecture Owner

🟦 Web DevOps

🟧 Team Members

🟦 Luis Maia* (50%)

🟦 Maurizio Manetti (25%)

🟦 Illia Derevianko (50%)

🟦 Bharathi Vanganuru (75% - 2Y)

🟦 Sudhanshu Singh (67% - 3Y)

🟦 Varun Singh (40% - 2Y)

🟧 Stakeholders

🟦 Data Management experts

🟦 Detector experts

🟦 Instrumentation groups

🟦 Instruments experts

🟦 Users

🟦 Administration
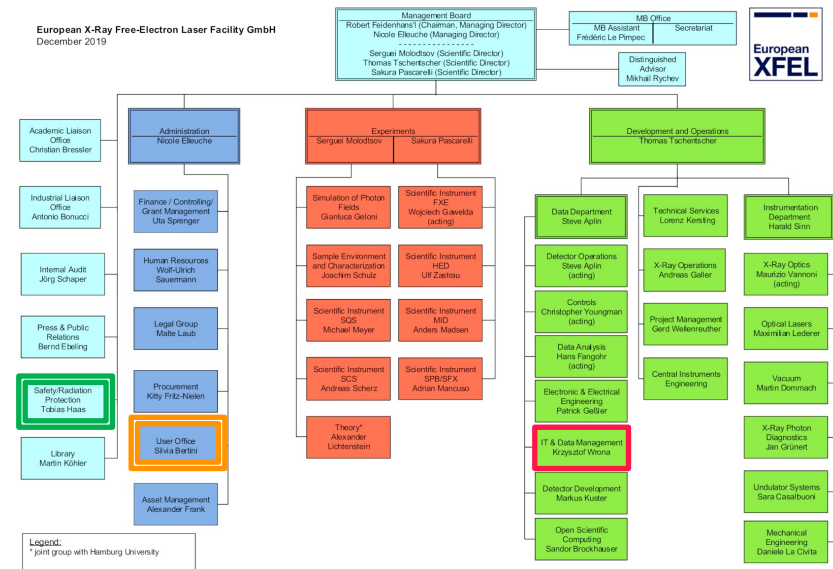
European XFEL

Agenda

🟧 Introduction

🟦 Context

🟦 Purpose

🟦 Landscape & Architecture

🟧 Governance & Resources

🟧 **myMdC application (optional)**

🟦 Roles and permissions

🟦 Key features

European XFEL

Roles and permissions (I)
myMdC roles

🟧 Administration

🟦 **admin** is the role assigned to myMdC administrators and ITDM OCD personnel

🟦 **Trusted** is the role assigned to people that needs to access all system information

🟧 Users & Staff

🟦 **read_only** is the default role for everybody accessing myMdC

🟦 **user** is the role granted to users with UPEX account and the latest Data Policy accepted

🟧 Systems

🟦 Specific roles to be assign to system integration local users

Dashboard / Roles

☰ List   ➦ Export

| Filter | | ⟳ Refresh | ✕ |

	Id ▾	Name	Flg default
☐	1	admin	0
☐	2	trusted	0
☐	3	user	0
☐	4	read_only	1
☐	6	karabo_sys	0
☐	7	upex_sys	0
☐	8	gpfs_sys	0
☐	9	cal_sys	0
☐	10	panosc_user	1
☐	11	monitoring_sys	1

10 roles

European XFEL

Roles and permissions (II)
Responsibilities in the context of instrument and proposal

LS: <u>Instrument's</u> Leading Scientist

- Any user in myMdC that is associated as the leading scientist to the Instrument where the given proposal is taking place

BE: <u>Instrument's</u> Expert

- Any user in myMdC that is associated as an expert user to the Instrument where the given proposal is taking place

SG: <u>Instrument's</u> Support Group

- Any user in myMdC that is a member of the support group associated with the instrument where the given proposal is taking place (e.g. SPBDATA)

PI: Principal Investigator

- Any user in myMdC

MP: Main Proposer

- Any user in myMdC

LC: Local Contact

- Any user in myMdC that is a Instrument Expert

T_Part: Execution team Participant

- Any user in myMdC

T_DM: Execution team Data Manager

- Any user in myMdC

T_UL: Execution team User Lab

- Any user in the execution team

Others: Non users or non members

- Any user in myMdC not having any responsibility on the proposal and not being Administrator

European XFEL

Key Features
Authentication
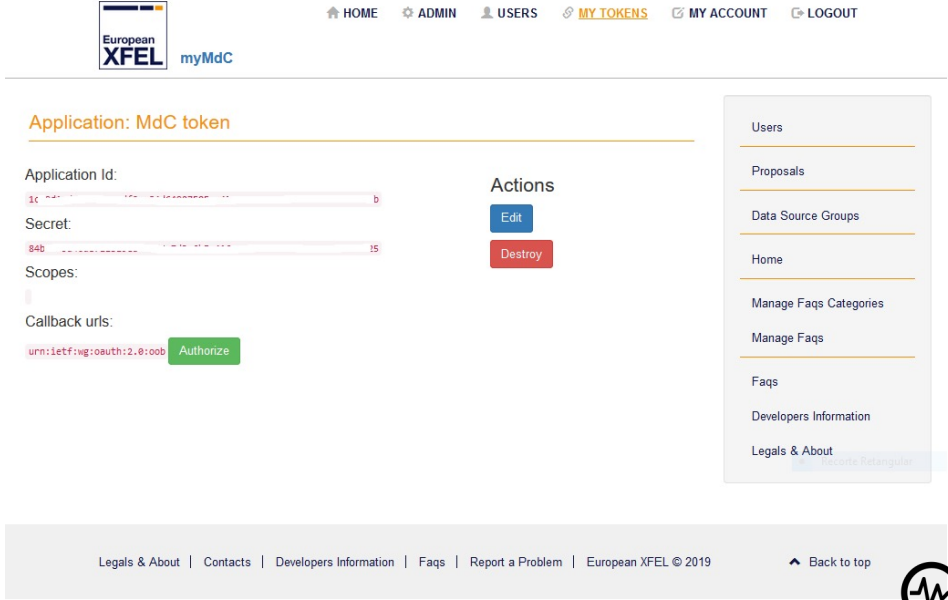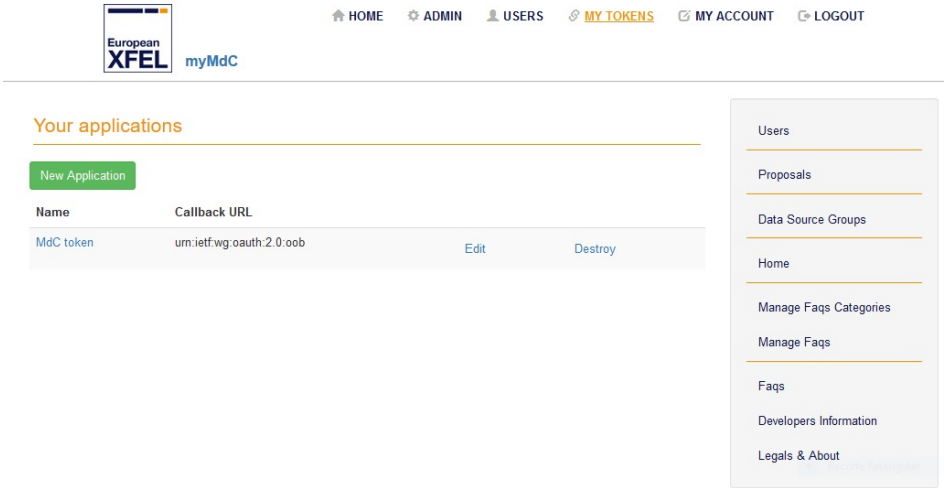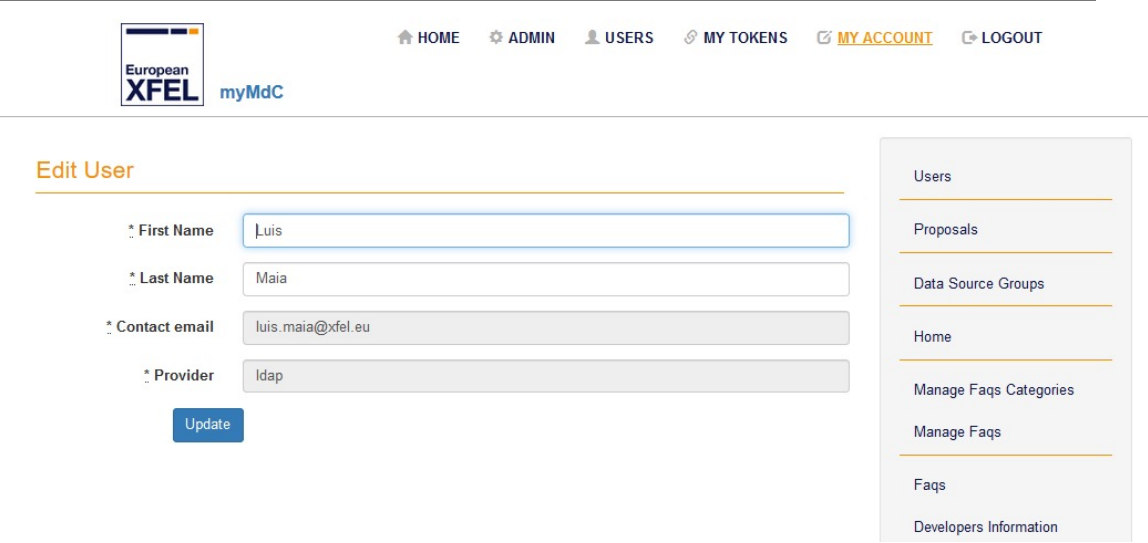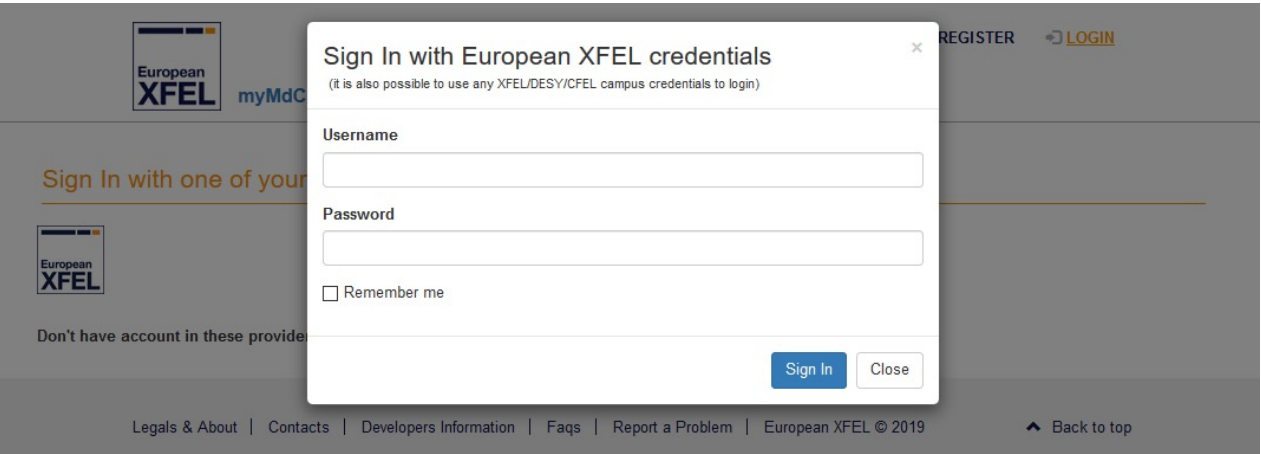
Key Features
Application Layout & Responsiveness

Administrators Only!

Key Features
Proposal details

Key Features
Proposal workflow and notifications

myMdC Proposals' Beamline status workflow

N: New
WLC: Waiting LDAP Configuration
DLC: LDAP Configuration Done
WOFC: Waiting GPFS Offline Configuration
DOFC: GPFS Offline Configuration Done
WONC: Waiting GPFS Online Configuration
DONC: GPFS Online Configuration Done
ONNA: GPFS Online Configuration Not Applicable
R: Ready
A: Active
I: Inactive
F: Finished (Beamtime)
FR: Finished and Reviewed
WCLO: Waiting (GPFS) Online Folder Removal
CLO: Closed
C: Cancelled
T: Test

Key Features
Proposal DOIs integration

Key Features
Proposal runs (data) management

Key Features
Proposal team and History

Proposal no. 900054

Back | Edit | Clone | + Add another | Runs | Delete | Beamtime status ▾

General | Public Information | Runs | Team | History

Experiment Team

Principal Investigator: Djelloul Boukhelef - (boukhele) as Data Manager & Participant
Main Proposer: Nasser Al-Qudami - (alqudami) as Data Manager & Participant
Team members 1: Luis Maia - (maial) as Participant
Team members 2: Jolanta Sztuk-Dambietz - (jsztuk) as Data Manager & Participant
Team members 3: Maurizio Manetti - (manettim) as Participant
Team members 4: Krzysztof Wrona - (wrona) as Data Manager & Participant
Team members 5: Peter Zalden - (zaldenp) as Data Manager & Participant
Team members 6: Janusz Malka - (jmalka) as Data Manager & Participant
Team members 7: Juergen Hannappel - (hannappj) as Participant

Download: CSV | JSON | Elog format

Back | Edit | Clone | + Add another | Runs | Delete | Beamtime status ▾

Team members

* User	Data access	
Luis Maia	Participant	✕
Djelloul Boukhelef	Data Manager &	✕
Jolanta Sztuk-Dambietz	Data Manager &	✕
Maurizio Manetti	Participant	✕
Krzysztof Wrona	Data Ma	
Peter Zalden	Data Ma	
Nasser Al-Qudami	Data Ma	
Janusz Malka	Data Ma	
Juergen Hannappel	Participa	

➕ add member

Audit fields

Created by   Janusz Malka

* User	Data access	* Lab Access	
Luis Maia	Participant	○ Yes ● No	✕
Karen Appel	Data Manager & Participant	○ Yes ● No	✕
Djelloul Boukhelef	Data Manager & Participant	○ Yes ● No	⊗
* User	Data access	* Lab Access	✕

Proposal no. 900054

Back | Edit | Clone | + Add another | Runs | Delete | Beamtime status ▾

General | Public Information | Runs | Team | History

History

2019-02-14 15:27:59 +0100 – ## System ##
| Preferred data output | Processed | ⇒ | Raw |

2019-01-29 18:23:55 +0100 – ## System ##
| Repository | --- | ⇒ | XFEL_GPFS_ONLINE_SASE_2 |

2019-01-29 03:41:47 +0100 – Luis Maia
| Beamtime status | Active | ⇒ | Finished (Beamtime) |

2019-01-23 16:24:10 +0100 – Janusz Malka
| Team changes | --- | ⇒ | Juergen Hannappel (Participant) |

2018-10-19 16:39:06 +0200 – Janusz Malka
| Beamtime status | Ready | ⇒ | Active |

2018-09-27 16:38:46 +0200 – Janusz Malka
| Team changes | --- | ⇒ | Janusz Malka (Data Manager & Participant) |

2018-09-25 17:30:09 +0200 – Luis Maia
| Beamtime status | GPFS Online Configuration Done | ⇒ | Ready |

2018-09-25 17:28:04 +0200 – Luis Maia
| Beamtime status | GPFS Offline Configuration Done | ⇒ | GPFS Online Configuration Done |

European XFEL

Key Features
Proposal Repositories

Key Features
Proposal Calibration

Key Features
Integration with data acquisition

Key Features
RESTful API's

API docs: https://in.xfel.eu/metadata/api-docs/index.html

European XFEL

Key Features
Python client to myMdC RESTful API's


```
exflqr17747:metadata_client maial$ python3 maial_py.py
{'success': True, 'info': 'Got proposal successfully', 'app_info': {}, 'data': {'id': 85,
'number': 900054, 'title': 'TST SASE2 DAQ tests', 'abstract': '', 'proposal_folder': 'p900
054', 'def_proposal_path': '/gpfs/exfel/exp/TSTSA2/201831/p900054/raw', 'doi': '10.22003/X
FEL.EU-DATA-900054-00', 'url': '', 'instrument_id': 110, 'instrument_cycle_id': 31, 'princ
ipal_investigator_id': 51, 'main_proposer_id': 411, 'local_contact_id': None, 'orig_reposi
tory_id': 2, 'last_run': 82, 'begin_at': '2018-09-25T00:00:01.000+02:00', 'end_at': '2018-
12-31T23:59:58.000+01:00', 'release_at': '2021-12-31T23:59:59.000+01:00', 'num_delivered_s
hifts': None, 'flg_public': False, 'flg_proposal_system': None, 'flg_preferred_data_output
': 'R', 'flg_available': True, 'flg_beamtime_status': 'F', 'beamtime_start_at': None, 'bea
mtime_end_at': None, 'description': '', 'beamtimes': [{'id': 73, 'begin_at': '2018-09-25T0
0:00:01.000+02:00', 'end_at': '2018-12-31T23:59:58.000+01:00', 'flg_available': True, 'des
cription': None}], 'team': [{'id': 1674, 'user_id': 51, 'data_access_role': 1}, {'id': 167
5, 'user_id': 170, 'data_access_role': 1}, {'id': 1677, 'user_id': 391, 'data_access_role'
: 1}, {'id': 1678, 'user_id': 399, 'data_access_role': 1}, {'id': 1679, 'user_id': 411, 'd
ata_access_role': 1}, {'id': 1697, 'user_id': 665, 'data_access_role': 1}, {'id': 1673, 'u
ser_id': 1, 'data_access_role': 2}, {'id': 1676, 'user_id': 231, 'data_access_role': 2}, {
'id': 1959, 'user_id': 1131, 'data_access_role': 2}], 'users_ids': [[51, 'boukhele'], [170
, 'jsztuk'], [391, 'wrona'], [399, 'zaldenp'], [411, 'alqudami'], [665, 'jmalka'], [1, 'ma
ial'], [231, 'manettim'], [1131, 'hannappj']]}}
[exflqr17747:metadata_client maial$
[exflqr17747:metadata_client maial$
```

```python
#!/usr/bin/python3
from metadata_client.metadata_client import MetadataClient

# Necessary configuration variables to establish a connection
user_id = '1d_____ab'
user_secret = '84b_____25'
user_email = 'luis.maia@xfel.eu'
#
metadata_web_app_url = 'https://in.xfel.eu/metadata'
token_url = 'https://in.xfel.eu/metadata/oauth/token'
refresh_url = 'https://in.xfel.eu/metadata/oauth/token'
auth_url = 'https://in.xfel.eu/metadata/oauth/authorize'
scope = ''
base_api_url = 'https://in.xfel.eu/metadata/api/'

proposal_number = 900054

# Generate the connection
client_conn = MetadataClient(client_id=user_id,
                             client_secret=user_secret,
                             user_email=user_email,
                             token_url=token_url,
                             refresh_url=refresh_url,
                             auth_url=auth_url,
                             scope=scope,
                             base_api_url=base_api_url)

all_proposal_runs = MetadataClient.get_proposal_info(client_conn, proposal_number)
#all_proposal_runs = MetadataClient.get_proposal_runs(client_conn, proposal_number)
print(all_proposal_runs)
```

Gitlab: https://in.xfel.eu/gitlab/ITDM/metadata_client
PyPi: https://pypi.org/project/metadata-client/

European XFEL

Key Features
Open data (PaNOSC & EOSC)

The main outcomes of other PaNOSC work packages were also presented. In particular, in WP3 – Data Catalogue Services, the activities carried out since the project's start include:

- Development of an Application Programmers Interface (API) for searching for FAIR data;
- **Integration of a search API into the EOSC portal**;
- Use of Nexus/HDF5 standard format for metadata;
- Automation of metadata collection on beamlines;
- **Use of e-logbook to make data FAIRer**;
- Long-term storage (hundreds of Petabytes).
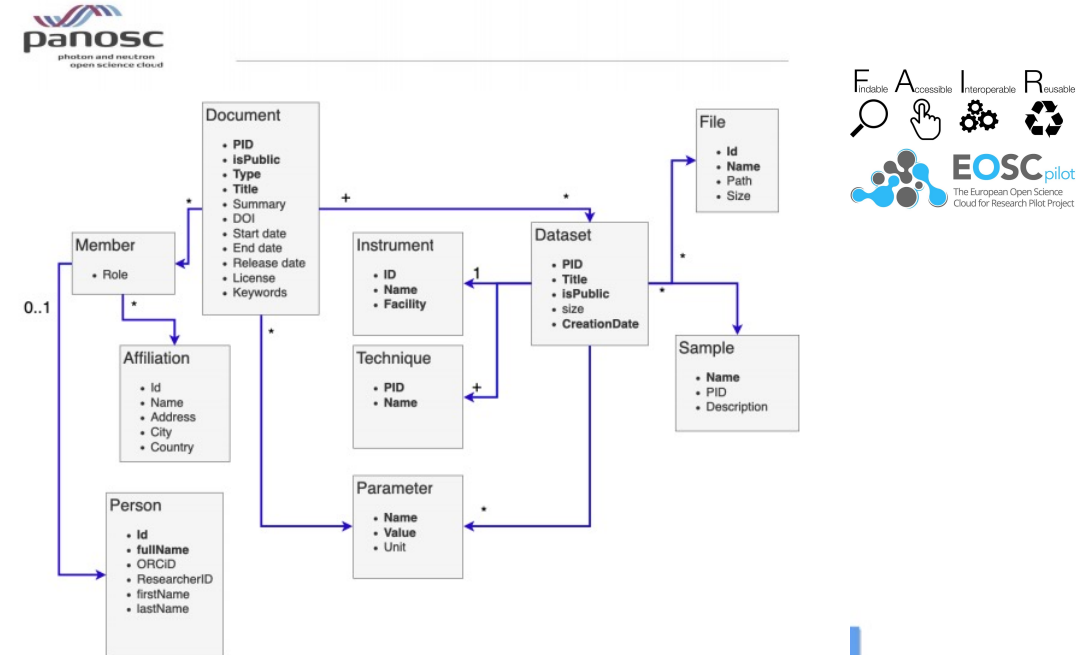
Source: www.panosc.eu/news
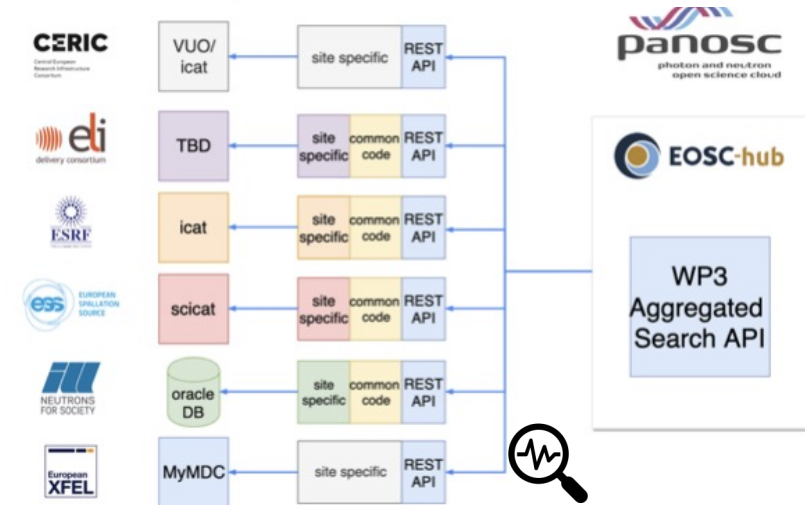

Figure 2: Search API data model

Figure 1: Overview of targeted deployment at partner sites and use of search API.
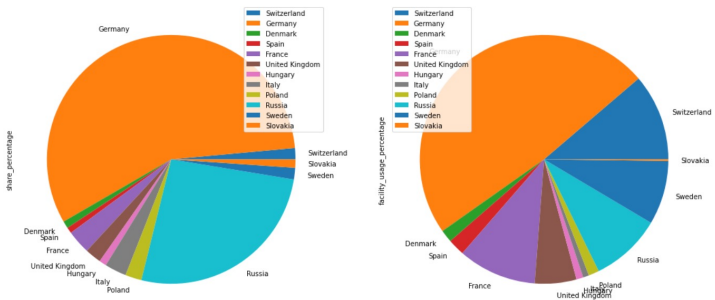
Key Features
Statistics

	Mar. 2019	Jan. 2020	Mar. 2021
Users	1296	1886	2196
Groups	346	662	948

	Mar. 2019	Jan. 2020	Mar. 2021
Instruments	15	20	20
Instrument Cycles	59	115	176
Repositories	11	12	12

	Mar. 2019	Jan. 2020	Mar. 2021
Proposals	144	287	431
Proposals with published DOI	3	4	13
Samples in all proposals	540	1238	1955
Run Types in all proposals	488	1136	1829
Users in all proposals	2833	6039	8886
Runs in all proposals	29736	64596	91139
Runs migrated to Maxwell	18304	43341	65579
Runs Calibrated (via myMdC)	463	8018	12323
Requested Dark Runs			147
Number of Delivered Shifts			

```
European XFEL Shareholders Scientific Usage Statistics
------------------------------------------------------

Scope: Calendar Year 2020 (Facility Run no.5 & Facility Run no.6)
Period: from 2020-01-01 till 2020-12-31
Total number of delivered shifts: 159
Statistics calculated with proposal experiment teams grouped by: Laboratories/Department
s
Final statistics calculation weighted by: Affiliation
```
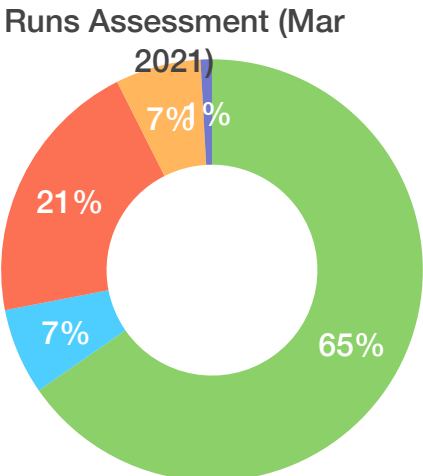

Runs Assessment

Runs Assessment (Jan 2020)

- Good
- Unclear
- Not interesting
- Not assessed
- Others

RUNS ASSESSMENT (MAR. 2021)	
Good	59.545
Unclear	6.034
Not interesting	18.790
Not assessed	5.977
Others	793
TOTAL RUNS	**91.139**
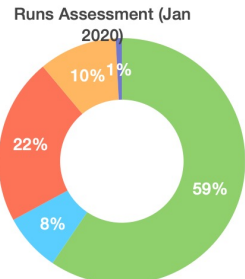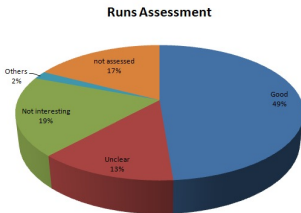
Runs Assessment (Mar 2021)

- Good
- Unclear
- Not interesting
- Not assessed
- Others

European XFEL

Thank you!

Special thank you to all the people that contributed to the project

European XFEL