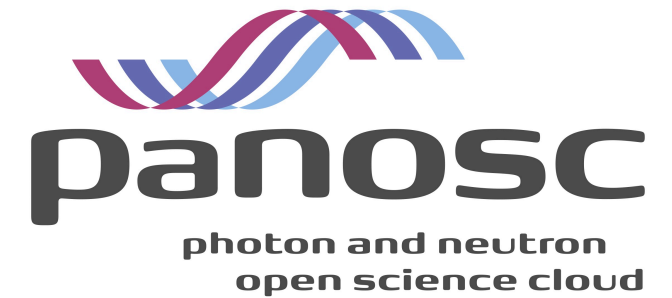


Introduction to **myMdC** the European XFEL metadata catalogue



Luís Maia,
On behalf of European XFEL

11th February 2020



Agenda

Introduction

- Context
- Purpose
- Landscape & Architecture

myMdC application

- Roles and permissions
- Key features
 - ▶ Proposal authorisation, team administration and access to real data
 - ▶ Proposal Run types and Samples
 - ▶ Proposal workflow and notifications
 - ▶ Run management activities (quality assessment, calibration request, describe runs)
 - ▶ DatCite (DOI) management
 - ▶ DAQ integration and
 - ▶ Data Source Groups management
 - ▶ (...)

Governance & Resources

Next steps (& EOSC)

Context (I)

Scientific Data Management

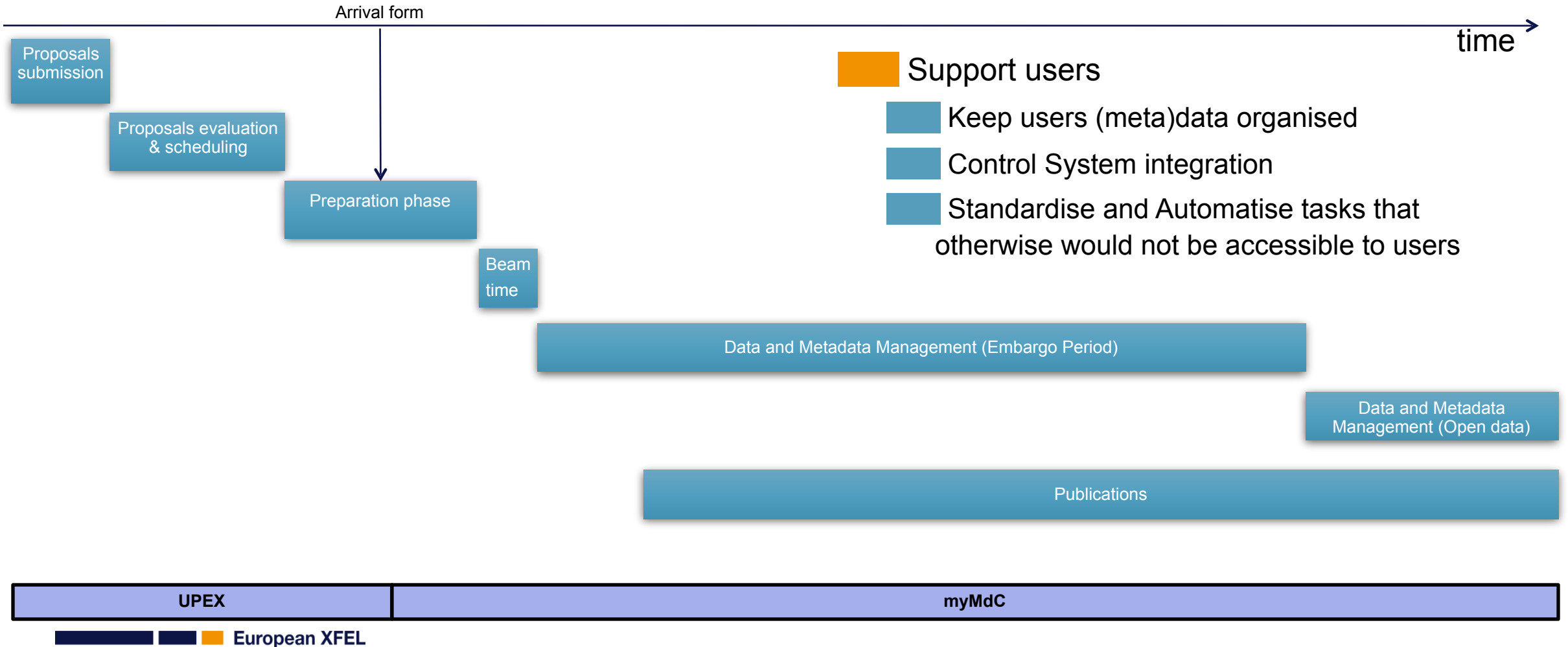
- Organize and manage data in a coherent way
- Aggregate and record large volumes of data
- Provide possibility of inspecting data during experiment
- Enable users to efficiently analyse experiment data when back at home institute
- Prepare for open access era

The screenshot shows a web browser displaying the XFEL website. The address bar shows the URL: https://www.xfel.eu/users/experiment_support/policies/scientific_data_policy/index_eng.html. The page has a navigation menu with items: **Scientific data policy**, Policy for samples and substances, User-supplied instrumentation/equipment, and Electrical work. Below the menu is a breadcrumb trail: Home > Users > Experiment support > Policies > Scientific data policy. The main heading is **Scientific data policy**. There are two document cards: 1. **Scientific data policy**, with the subtitle "As approved by the European XFEL Council on 30 June 2017" and a **DOWNLOAD** button. 2. **Data retention policy** with a **DOWNLOAD** button.

- General principles
 - Open access
 - Long term data curation
 - Embargo period with restricted access






Context (II)

Users support (and Experiments lifecycle)







Context (III)

Data Management Activities

-  Catalog and document the data generated at European XFEL
-  Data migration services
-  Keep track of data location
-  Allow data publishing, exportation and authorisation management
-  Manage permissions and accesses to the data in an uniform manner and guarantee that only authorised people have access to the (meta)data

Storage	Quota	Safety	Lifetime	Comment
dcache.raw	None	Tape Archive	6 months (t.b.c)	Raw data
raw	None	None	2 months (t.b.c)	Fast accessible raw data, lifetime not guaranteed
usr	5TB	Snapshots + Tape Backup	24 months (t.b.c)	user data, results
proc	None	None	6 months (t.b.c)	processed data e.g. calibrated
scratch	None	None	6 months	Temporary data (lifetime not guaranteed)
dcache.cal	None	Tape Archive	10 years	Calibration constants
cal	None	Tape archive	6 months	
user home	20GB	Snapshots + Tape Backup	Lifetime of the account	
archive.raw	None	-	Long-term	Long term means 5 years and XFEL will strive for 10 years
archive.cal	None	-	10 years	

Data Retention Policies control and execution

-  Data Quality Assessment
-  Data Calibration
-  Data Archiving
-  Data Cleanup

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

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

Registry

Users accounts
 Staff organisational group
 Instrument support groups*

LDAP

Users Proposal groups*

GPFS

Repository location
 Universal folders management
 Proposal folders creation
 Runs data migration
 Folders removal*
 Data Audit and reconciliation

Elog

Elog logbook config file generation
 Integration with Elog replacement tool

SLURM

Priority queues reservation for beamtime*

Karabo - DAQ

Universal Run Management
 Run tagging and basic details
 Number of files and size
 Data Source Groups definition

Calibration Pipeline

Trigger Processed data generation*

calCat

Document used CCV on data calibration

Grid-FTP / Globus / Off-shore data exportation

Transfer Agent Off-shore data migration

DataCite / Tind.io

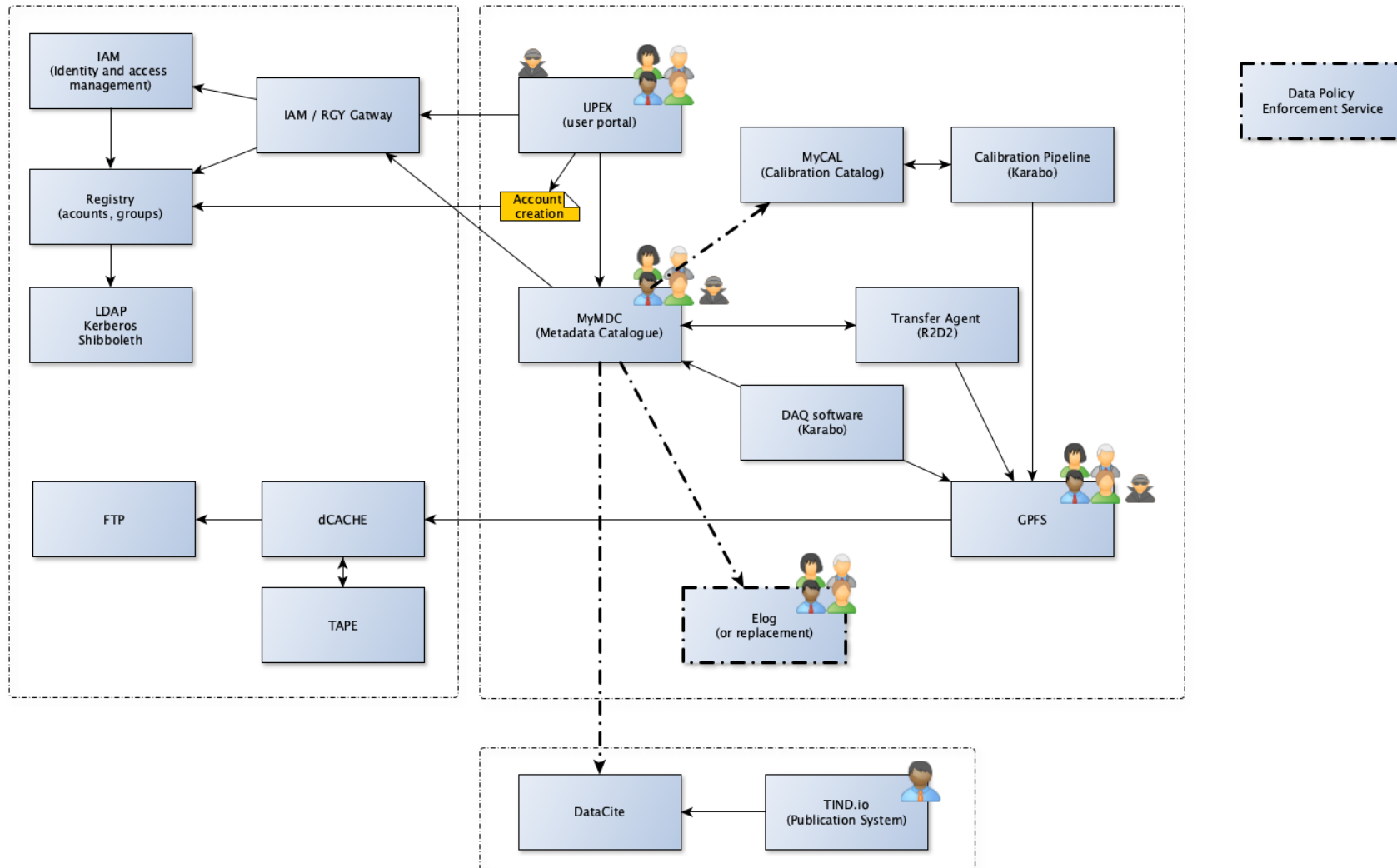
DOI generation and publish*
 Public DOI end point

Data Policy enforcement Service

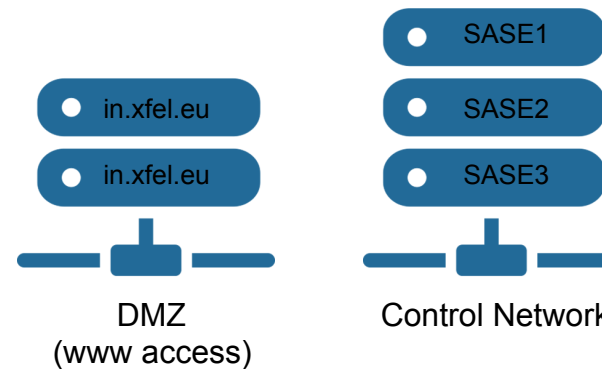
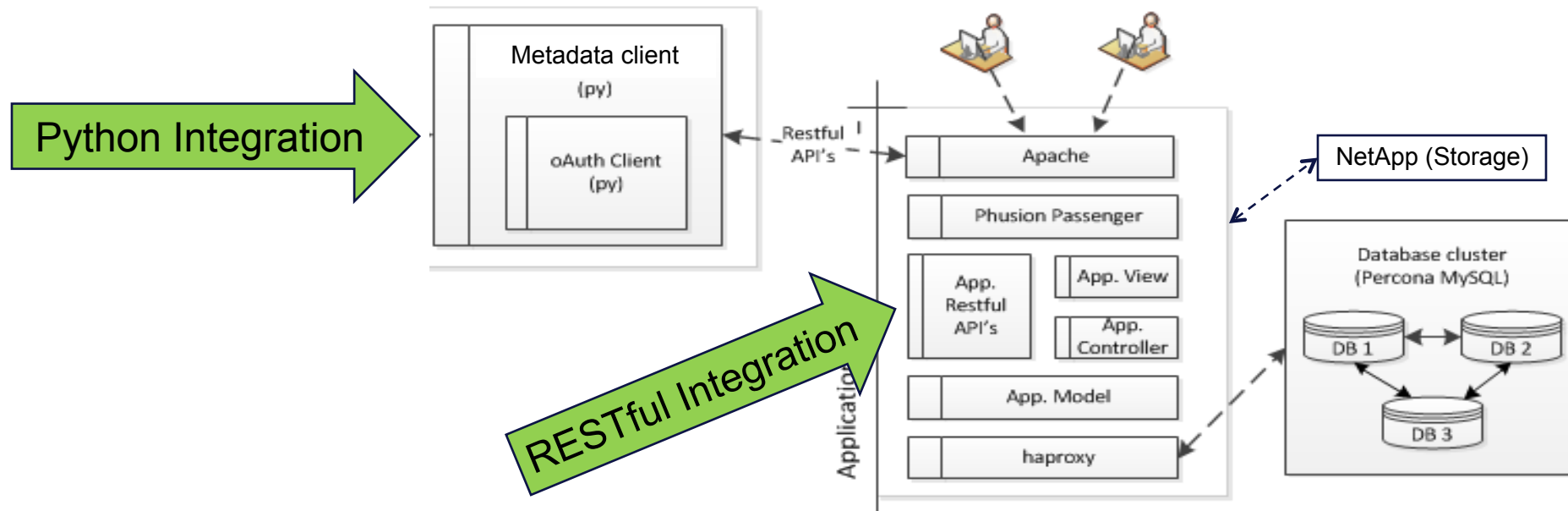
Notifications*

Statistics

Landscape



Architecture



Agenda

Introduction

Context

Purpose

Landscape & Architecture

myMdC application

Roles and permissions

Key features

- ▶ Proposal authorisation, team administration and access to real data
- ▶ Proposal Run types and Samples
- ▶ Proposal workflow and notifications
- ▶ Run management activities (quality assessment, calibration request, describe runs)
- ▶ DatCite (DOI) management
- ▶ DAQ integration and
- ▶ Data Source Groups management
- ▶ (...)

Governance & Resources

Next steps (& EOSC)

Roles and permissions (I)

myMdC roles

Administration

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

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

List of Roles

Dashboard / Roles

List

Filter

Refresh

<input type="checkbox"/>	Id	Name	Flg default
<input type="checkbox"/>	1	admin	0
<input type="checkbox"/>	2	trusted	0
<input type="checkbox"/>	3	user	0
<input type="checkbox"/>	4	read_only	1
<input type="checkbox"/>	6	KARABO_SYS	0
<input type="checkbox"/>	7	UPEX_SYS	0
<input type="checkbox"/>	8	GPFS_SYS	0
<input type="checkbox"/>	9	CAL_SYS	0

8 roles

Roles and permissions (II)

Responsibilities in the context of instrument and proposal

LS: Instrument's Leading Scientist

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

BE: Instrument's Beamline Expert

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

SG: Instrument's 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 Beamline Expert

T_Part: Execution team Participant

Any user in myMdC

T_DM: Execution team Data Manager

Any user in myMdC

Others: Non users or non members

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

Roles and permissions (III)

Proposal permission Grid

What / Who	PI	MP	T_Part	T_DM	LC	BE	LS	SG	Others
Index/search Proposals	Y	Y	Y	Y	Y	Y	Y	Y	N
Show Proposal (General, PI, Runs, Team)	Y	Y	Y	Y	Y	Y	Y	Y	N**
Edit Proposal	Y	Y	N	N ¹	Y	Y ²	Y	N	N
New Proposal	N	N	N	N	N	N	N	N	N
Delete Proposal	N	N	N	N	N	N	N	N	N
History Page	Y	Y	Y	Y	Y	Y	Y	Y	N
Public Proposal DOI page	Y	Y	Y	Y	Y	Y	Y	Y	Y
Proposal Runs List	Y	Y	Y	Y	Y	Y	Y	Y	N**
Proposal Show Run Type	Y	Y	Y	Y	Y	Y	Y	Y	N**
Proposal Edit Run Type	Y	Y	N	N	N	Y ³	Y'	N	N
Proposal Create Run Type	Y	Y	N	N	N	Y ³	Y'	N	N
Proposal Delete Run Type	N	N	N	N	N	N	N	N	N
Proposal Show Sample	Y	Y	Y	Y	Y	Y	Y	Y	N**
Proposal Edit Sample	Y ⁴	Y ⁴	Y ⁴	Y ⁴	N	Y ³	Y'	N	N
Proposal Create Sample	N	N	N	N	N	Y ³		N	N
Proposal delete Sample	N	N	N	N	N	N	N	N	N
Proposal Index Runs	Y	Y	Y	Y	Y	Y	Y	Y	N**
Proposal Runs List Update	Y	Y	N	Y	N	Y ³	Y'	N	N
Proposal Create Runs	N	N	N	N	N	N	N	N	N
Proposal Delete Runs	N	N	N	N	N	N	N	N	N



Grant access to data

Notes:

(*): includes the ability to assess the quality of the run

('): instrument's leading scientist is also always a Beamline expert

(1): data managers can change Preferred Data output field

(2): for proposals in states: R, A, I

(3): for proposals in states: T, N, WLC, DLC, WOFC, DOFC, WONC, DONC, ONNA, R, A, I, F

(4): for proposals in states: N, WLC, DLC, WOFC, DOFC, WONC, DONC, ONNA, R, A, I

(**): Proposal is open data

Agenda

Introduction

Context

Purpose

Landscape & Architecture

myMdC application

Roles and permissions

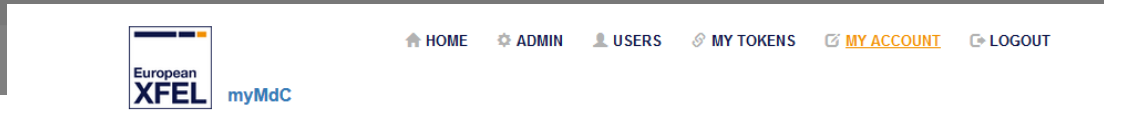
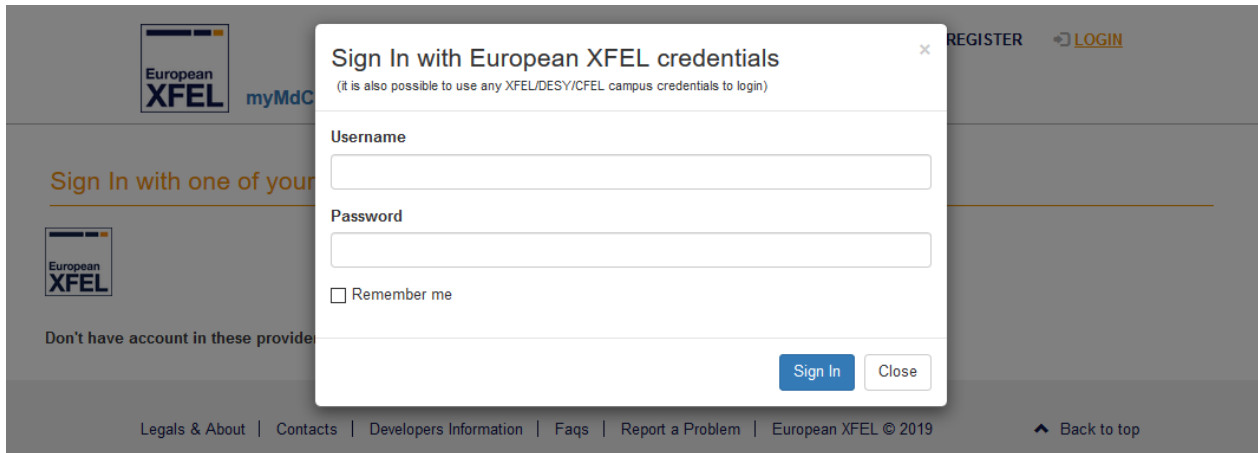
Key features

- ▶ Proposal authorisation, team administration and access to real data
- ▶ Proposal Run types and Samples
- ▶ Proposal workflow and notifications
- ▶ Run management activities (quality assessment, calibration request, describe runs)
- ▶ DatCite (DOI) management
- ▶ DAQ integration and
- ▶ Data Source Groups management
- ▶ (...)

Governance & Resources

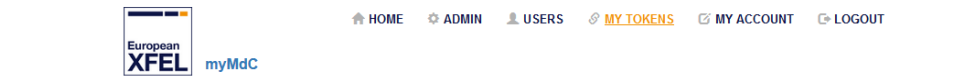
Next steps (& EOSC)

Key Features Authentication



Edit User

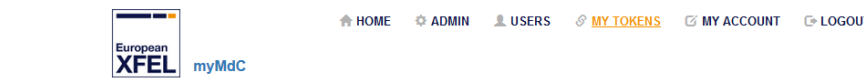
- Users
- Proposals
- Data Source Groups
- Home
- Manage Faqs Categories
- Manage Faqs
- Faqs
- Developers Information



Your applications

Name	Callback URL		
MdC token	urn:ietf:wg:oauth:2.0:oob	<input type="button" value="Edit"/>	<input type="button" value="Destroy"/>

- Users
- Proposals
- Data Source Groups
- Home
- Manage Faqs Categories
- Manage Faqs
- Faqs
- Developers Information
- Legals & About



Application: MdC token

Application Id:

 Secret:

 Scopes:

Callback urls:

Actions

-
-

- Users
- Proposals
- Data Source Groups
- Home
- Manage Faqs Categories
- Manage Faqs
- Faqs
- Developers Information
- Legals & About

Key Features

Application Layout & Responsiveness

The screenshot shows the myMdC application interface. At the top left is the European XFEL myMdC logo. The navigation bar includes links for HOME, ADMIN, USERS, MY TOKENS, MY ACCOUNT, and LOGOUT. A 'Signed in!' notification is visible. The main content area features a 'Welcome to the myMdC' message and a 'Quick Introductory Tour' button. A red arrow points from the 'ADMIN' link to the 'Users' menu item in the dropdown. Another red arrow points from the 'USERS' link to the 'Users' menu item. A third red arrow points from the 'USERS' link to the 'Proposals' menu item. A fourth red arrow points from the 'USERS' link to the 'Data Source Groups' menu item. A fifth red arrow points from the 'USERS' link to the 'Manage Faqs Categories' menu item. A sixth red arrow points from the 'USERS' link to the 'Manage Faqs' menu item. Below these arrows, the text 'Administrators Only!' is written in red. The footer contains links for Legals & About, Contacts, Developers Information, Faqs, Report a Problem, and European XFEL © 2019, along with a 'Back to top' button.

European XFEL myMdC MENU

Welcome to the myMdC

The Data Management portal for European XFEL users

[Quick Introductory Tour](#)


Please Note, that you should use your XFEL/DESY/CFEL campus credentials to login in this system.

- Users
- Proposals
- Data Source Groups
- Manage Faqs Categories
- Manage Faqs
- Faqs
- Developers Information
- Legals & About

Legals & About | Contacts [▲ Back to top](#)
Developers Information | Faqs
Report a Problem
European XFEL © 2019

Key Features

Proposal details



HOME ADMIN USERS MY TOKENS MY ACCOUNT

Proposals

Show 10 entries Search: TSTSA2 Copy Print Export

Number	Title	Instrument	Beamtime status	Start date			
900054	TST SASE2 DAQ tests	TSTSA2	Finished (Beamtime)	2018-09-25 00:00:01 +0200			
900058	TST SASE2 DAQ tests	TSTSA2	Ready	2019-01-21 00:00:01 +0100			

Showing 1 to 2 of 2 entries (filtered from 144 total entries)

First Previous 1 Next Last

[+ New](#)

Proposal no. 900058

[Back](#)
[Edit](#)
[Clone](#)
[+ Add another](#)
[Runs](#)
[Delete](#)
[Beamtime status](#)

[General](#)
[Public Information](#)
[Runs](#)
[Team](#)
[History](#)

General

#id: 99
DOI: 10.22003/XFEL.EU-DATA-900058-00 Internal Only: DOI not registered!

Proposal Number: 900058
Name: p900058
Title: TST SASE2 DAQ tests
URL:
Abstract:
Proposal Folder: p900058
Proposal Path: /gpfs/xfel/exp/TSTSA2/201930/p900058/raw
Repository: XFEL_GPFS_ONLINE_SASE_2
Preferred data output: Raw [Preferred data output](#)

Proposal Last Run: 20
Instrument: Test SASE2
Instrument Cycle: 201930
Principal Investigator: Djelloul Boukhelef
Main Proposer: Nasser Al-Qudami
Local Contact: ---
Start date: 2019-01-21 00:00:01 +0100
End date: 2019-06-30 23:59:58 +0200
No. delivered shifts: ---
Beamtime 1: 2019-01-21 00:00:01 +0100 - 2019-06-30 23:59:58 +0200
Release at: 2022-12-31 23:59:58 +0100
open data?: No
Beamtime status: Ready [Beamtime status](#)

Start Beamtime date: ---
End Beamtime date: ---
Available: Yes
Proposal system: ---
Description:
Created by: Luis Maia
Creation date: 2019-01-23 14:36:37 +0100
Modified by: Luis Maia
Last modification date: 2019-03-13 14:50:53 +0100

All Proposal Run types

[+ New](#)

All Proposal Samples

• [e0001]: Test DAQ

• No Sample

Users

[Proposals](#)

Data Source Groups

Home

Manage Faqs Categories

Manage Faqs

Faqs

Developers Information

Legals & About

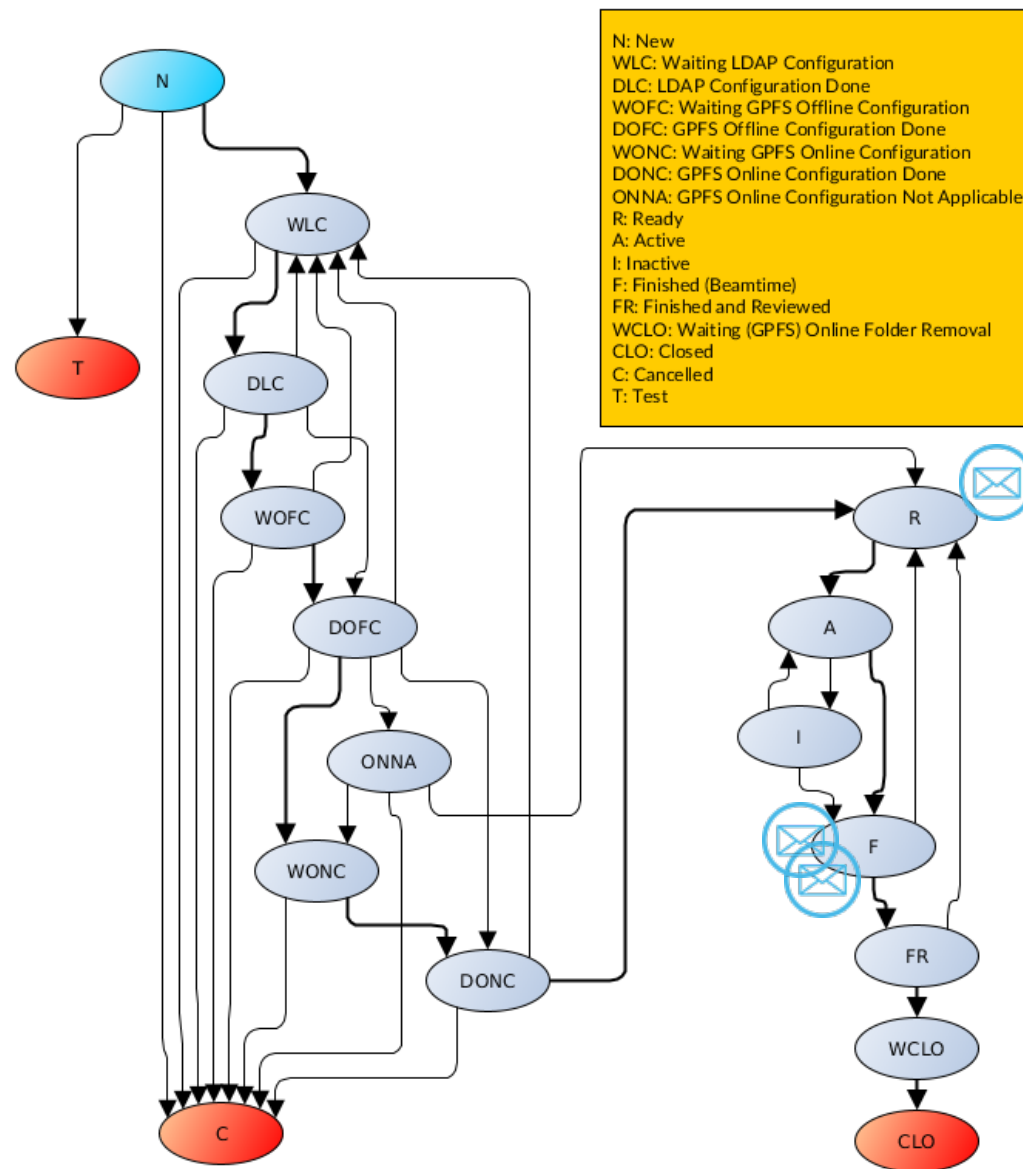
Key Features

Proposal workflow and notifications

no. delivered shifts: ---
 Beamtime 1: 2019-01-21 00:00:01 +0100 - 2019-06-30 23:59:58 +0200
 Release at: 2022-12-31 23:59:58 +0100
 open data?: No
 Beamtime status: Ready
 Start Beamtime date: Ready
 End Beamtime date: Active
 Proposal system: ---
 Description:
 Created by: Luis Maia

no. delivered shifts: ---
 Beamtime 1: 2018-09-25 00:00:01 +0200 - 2018-12-31 23:59:58 +0100
 Release at: 2021-12-31 23:59:59 +0100
 open data?: No
 Beamtime status: Finished (Beamtime)
 Start Beamtime date: Ready
 End Beamtime date: Finished (Beamtime)
 Proposal system: Finished and Reviewed
 Description:
 Created by: Janusz Malka
 Creation date: 2018-09-25 10:47:10 +0200
 Modified by: System

myMdC Proposals' Beamline status workflow



Key Features

Proposal DOIs integration

Proposal no. 900058

Back
Edit
Clone
Add another
Runs
Delete
Beamtime status

[General](#)
[Public Information](#)
[Runs](#)
[Team](#)
[History](#)

Public Information

DOI: 10.22003/XFEL.EU-DATA-900058-00 Internal Only: DOI not registered!
Proposal Number: 900058
Name: p900058
Title: TST SASE2 DAQ tests

Back
Edit
Clone
Add another
Runs
Delete
Beamtime status

Proposal no. 002012

[General](#)
[Public Information](#)
[Runs](#)
[Team](#)
[History](#)

Public Information

Back
Edit

DOI: 10.22003/XFEL.EU-DATA-002012-00
Proposal Number: 002012
Name: p002012
Title: Serial Femtosecond Crystallography at MHz repetition rates
Instrument: SPB/SFX SASE1
Instrument Cycle: 201701
Principal Investigator: Henry Chapman
Main Proposer: Anton Barty
Local Contact: Richard Bean
Beamtime 1: 2017-09-14 20:00:00 +0200 - 2017-09-19 08:00:00 +0200
Beamtime 2: 2018-03-29 20:00:00 +0200 - 2018-04-02 08:00:00 +0200
Release at: 2021-04-02 08:00:00 +0200
open data?: No

Back
Edit
Clone
Add another
Runs
Delete
Beamtime status

Users

Proposals

Data Source Groups

Home

Manage Faqs Categories

Manage Faqs

Users

Proposals

Data Source Groups

Home

Manage Faqs Categories

Manage Faqs

Faqs

Developers Information

Legals & About



[HOME](#) |
 [HANDBOOK](#) |
 [FACTSHEETS](#) |
 [FAQs](#) |
 [RESOURCES](#) |
 [REGISTRATION AGENCIES](#) |
 [NEWS](#) |
 [MEMBERS AREA](#)

The DOI® System

ISO 26324



Resolve a DOI Name

Type or paste a DOI name, e.g., 10.1000/xyz123, into the text box below. (Be sure to enter all of the characters before and after the slash. Do not include extra characters, or sentence punctuation marks.)

SUBMIT

Clicking on a DOI link (try this one: <https://doi.org/10.1109/5.771073>) takes you to one or more current URLs or other services related to a single resource. If the URLs or services change over time, e.g., the resource moves, this same DOI will continue to resolve to the correct resource or service at their new location.

<https://in.xfel.eu/metadata/doi/10.22003/XFEL.EU-DATA-002012-00>

This is the web site of myMdc | Proposal no. 002012



[REGISTER](#)
[LOGIN](#)

Proposal no. 002012

DOI: 10.22003/XFEL.EU-DATA-002012-00
Proposal Number: 002012
Name: p002012
Title: Serial Femtosecond Crystallography at MHz repetition rates
Instrument: SPB/SFX SASE1
Instrument Cycle: 201701
Principal Investigator: Henry Chapman
Main Proposer: Anton Barty
Local Contact: Richard Bean
Beamtime 1: 2017-09-14 20:00:00 +0200 - 2017-09-19 08:00:00 +0200
Beamtime 2: 2018-03-29 20:00:00 +0200 - 2018-04-02 08:00:00 +0200
Release at: 2021-04-02 08:00:00 +0200
open data?: No

Key Features

Proposal runs (data) management

Proposal no. 900054

Run Number (alias)	Run type	Sample Name	Start date	Run status	Data Assessment	Calibration	Run Comment	Edit
0082	Test DAQ	No Sample	2019-01-23 14:20:18 +0100	Closed	Unknown/Unclear	[Calibration]	[Comment]	[Edit]
0081	Test DAQ	No Sample	2019-01-23 12:10:47 +0100	Closed	Run Quality	[Calibration]	[Comment]	[Edit]
0080	Test DAQ	No Sample	2019-01-23 03:13:08 +0100	Closed	Unknown/Unclear	[Calibration]	[Comment]	[Edit]
0079	Test DAQ	No Sample	2019-01-22 14:29:27 +0100	Closed	Unknown/Unclear	[Calibration]	[Comment]	[Edit]
0078	Test DAQ	No Sample	2019-01-22 14:16:09 +0100	Closed	Unknown/Unclear	[Calibration]	[Comment]	[Edit]
0077	Test DAQ	No Sample	2019-01-22 14:07:09 +0100	Closed	Unknown/Unclear	[Calibration]	[Comment]	[Edit]

Proposal no. 900054

Proposal Runs

Run Number (alias)	Run type	Sample Name	Start date	Run status	Data Assessment	Calibration	Run Comment	Edit
0082	Test DAQ	No Sample	2019-01-23 14:20:18 +0100	Closed	Unknown/Unclear	[Calibration]	[Comment]	[Edit]
0081	Test DAQ	No Sample	2019-01-23 12:10:47 +0100	Closed	Run Quality	[Calibration]	[Comment]	[Edit]
0080	Test DAQ	No Sample	2019-01-23 03:13:08 +0100	Closed	Unknown/Unclear	[Calibration]	[Comment]	[Edit]
0079	Test DAQ	No Sample	2019-01-22 14:29:27 +0100	Closed	Unknown/Unclear	[Calibration]	[Comment]	[Edit]
0078	Test DAQ	No Sample	2019-01-22 14:16:09 +0100	Closed	Unknown/Unclear	[Calibration]	[Comment]	[Edit]
0077	Test DAQ	No Sample	2019-01-22 14:07:09 +0100	Closed	Unknown/Unclear	[Calibration]	[Comment]	[Edit]

Run Description

This is the run 0079 description

Proposal no. 900054 > Edit Run no. 79

Run Number: 79

Run Alias: []

Run type: Test DAQ

Sample: No Sample

Run Status: Closed

Run Quality: Unclear (migrate data to Maxwell)

Description: This is the run 0079 description

Audit fields

Created by: KARABO Migrator

Created by: 2019-01-22 14:29:27 +0100

Modified by: Luis Maia

Last modification date: 2019-04-02 04:15:18 +0200

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 Manager &	
Peter Zalden	Data Manager &	
Nasser Al-Qudami	Data Manager &	
Janusz Malka	Data Manager &	
Juergen Hannappel	Participant	

+ add member

Audit fields

Created by Janusz Malka

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

Key Features

Integration with data acquisition

European XFEL myMdc

HOME ADMIN USERS MY TOKENS MY ACCOUNT LOGOUT

Proposal no. 900054

Back Edit Clone Add another Runs Delete Beamtime status

General Public Information **Runs** Team

Proposal Runs

Run Number (alias)	Run type	Sample Name	Start date	Run status	Data Assessment	Run Comment	Edit
0057	Test DAQ	No Sample	2019-01-07 16:52:45 +0100	Closed	Run Quality		
0056	Test DAQ	No Sample	2019-01-07 16:48:05 +0100	Closed			
0055	Test DAQ	No Sample	2018-12-20 12:43:50 +0100	Closed			

Run Quality dropdown menu:

- Good (migrate data to Maxwell)
- Unclear (migrate data to Maxwell)
- Not interesting (data won't be migrated to Maxwell)

SPB_RUN_CONTROL

SPB Run Controller

Proposal number: 2042 900015 Retrieve proposal

source	type	behavior	r	Data source	Data aggregator
<input type="checkbox"/>	IRU_...			0 @EB_COND	SPB_DAQ_DATA/DA/1
<input type="checkbox"/>	SPB_...			1 @IRU_MOTO...	SPB_DAQ_DATA/DA/1
<input type="checkbox"/>	SPB_...			2 @SPB_EHC...	SPB_DAQ_DATA/DA/1
<input checked="" type="checkbox"/>	AGIP...			3 @SPB_IRU_I...	SPB_DAQ_DATA/DA/1
<input type="checkbox"/>	SPB_...			4 @SPB_AGIP	SPB_DAQ_DATA/DA/2

Push to DAQ

```

[09:32:00]: Apply configuration successfully.
[09:32:04]: Monitoring incoming data.
[09:40:16]: Ignore incoming data.
[10:28:29]: Monitoring incoming data.
[10:48:49]: Ignore incoming data.
[10:48:53]: Initiate 'configure' action
[10:48:54]: Apply configuration successfully.
[10:49:13]: Monitoring incoming data.
[10:51:00]: Start data recording (run-number:437)
[10:51:01]: Data recording in progress (run-number: 437)
[10:52:53]: Initiate 'tune' action
[10:52:54]: Run closed successfully
[10:53:34]: Start new run ...
[10:53:35]: Data recording in progress (run-number: 438)
    
```

Proposal: RC DA1 DA2

Data source: RC DA1 DA2

Topology: RC DA1 DA2

Assign: RC DA1 DA2

Configure: RC DA1 DA2

Monitor: RC DA1 DA2

Ignore data Apply configuration

Run Type: Dark runs General

Sample: No sample 3-deoxy-D-manno-2-octulosonate-8-phospho

Train Id: 59141103 Previous run 438

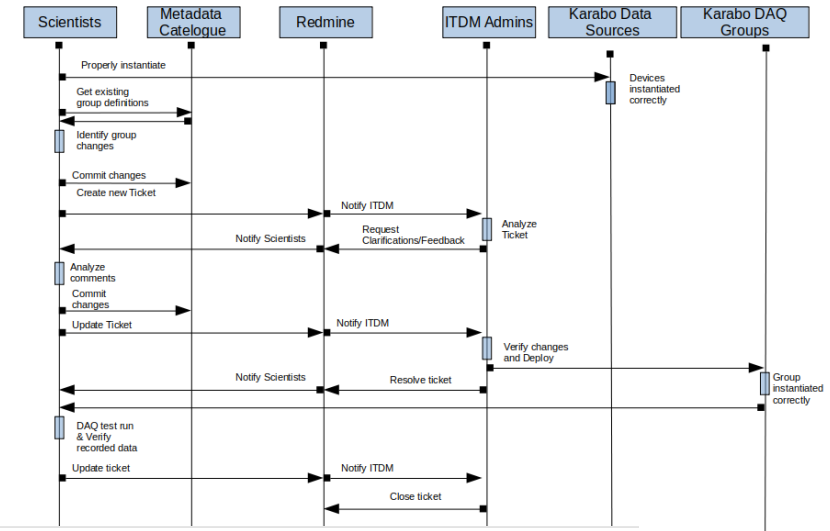
Monitor data Start run Stop run

Key Features

Data Source Groups management

The screenshot shows the myMdc web interface. On the left, a 'Documents' sidebar lists files like 'fxe_active_data_source_groups.xls' and 'hed_active_data_source_groups.xls'. The main area displays a list of documents with details like modification dates and sizes. On the right, the 'FXE Run Controller' interface is visible, showing a 'Proposals' table with columns for 'source', 'Data source', and 'Data aggregator'. Below the table are controls for 'Push to DAQ', 'Run Type', and 'Sample'. A status bar at the bottom shows 'Run closed successfully' and a progress indicator for 'RC'.

source	Data source	Data aggregator
0	@FXE_RR_SYS_DIGI	FXE_DAQ_DATA/...
1	@FXE_X2TIMER	FXE_DAQ_DATA/...
2	@FXE ENERGY	FXE_DAQ_DATA/...
3	@FXE_XAD_JF1M_1	FXE_DAQ_DATA/J...



Deployed data source groups details [Expand all](#)

ITLAB

ITLAB_DAQ_DATA/DM/PARTHA2_GROUP deployed at: 2019-11-13 15:18:59 +0100

Source	Monitored
ITLAB_DAQ_DATA/DM/RUN_CONFIGURATOR	1
ITLAB_DAQ_DATA/DM/SRC1:DAQOUTPUT	1
ITLAB_DATA_DATA/DM/SRC2	0

ITLAB_DAQ_DATA/DM/30OCT_34_GROUP deployed at: 2019-11-08 10:14:53 +0100

Source	Monitored
TEST1	1

- XFEL_TESTS_TOPIC:** You don't have permission to access this Instrument data.
- SPB:** You don't have permission to access this Instrument data.
- FXE:** You don't have permission to access this Instrument data.
- SA1:** You don't have permission to access this Instrument data.
- SA2:** You don't have permission to access this Instrument data.

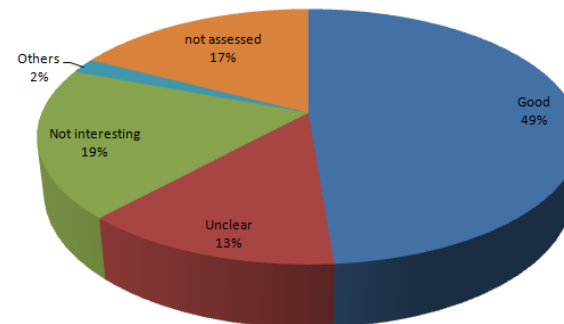
Key Features Statistics

	Mar. 2019	Jan. 2020
Users	1296	1886
Groups	346	662

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

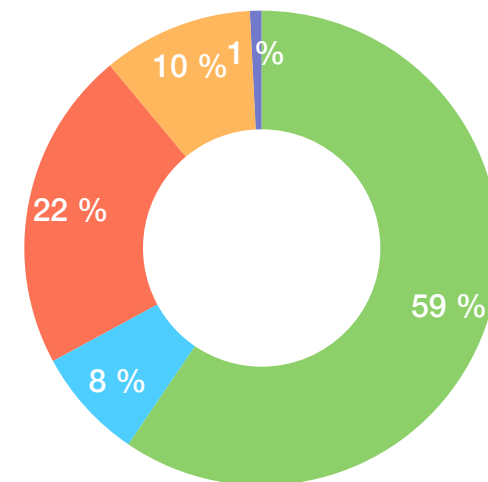
	Mar. 2019	Jan. 2020
Proposals	144	287
Proposals with published DOI	3	4
Samples in all proposals	540	1238
Run Types in all proposals	488	1136
Users in all proposals	2833	6039
Runs in all proposals	29736	64596
Runs migrated to Maxwell	18304	43341
Runs Calibrated (via myMdC)	463	8018

Runs Assessment



Runs Assessment (Jan 2020)

RUNS ASSESSMENT (JAN. 2020)	
Good	38.428
Unclear	4.913
Not interesting	14.120
Not assessed	6.620
Others	515
TOTAL RUNS	64.596



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

Agenda

Introduction

Context

Purpose

Landscape & Architecture

myMdC application

Roles and permissions

Key features

- ▶ Proposal authorisation, team administration and access to real data
- ▶ Proposal Run types and Samples
- ▶ Proposal workflow and notifications
- ▶ Run management activities (quality assessment, calibration request, describe runs)
- ▶ DatCite (DOI) management
- ▶ DAQ integration and
- ▶ Data Source Groups management
- ▶ (...)

Governance & Resources

Next steps (& EOSC)

Governance & Resources

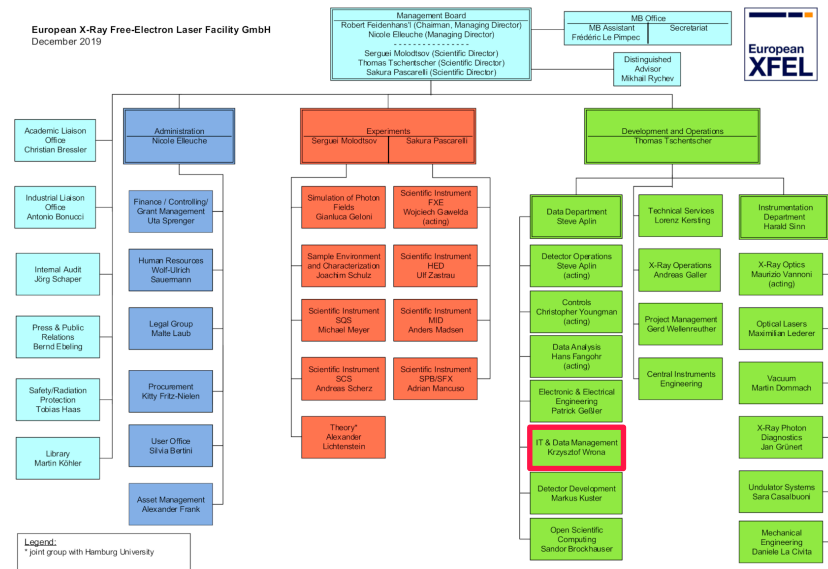
- Product Owner
 - Data Department (Operational/Support groups)

- Architecture Owner
 - Web DevOps

- Team Members
 - 3 FTEs(<50%)
 - (...)

Stakeholders

- Data Management experts
- Detector experts
- Instrumentation groups
- Instruments experts
- Users
- Administration



Agenda

Introduction

- Context

- Purpose

- Landscape & Architecture

myMdC application

- Roles and permissions

- Key features

- ▶ Proposal authorisation, team administration and access to real data
- ▶ Proposal Run types and Samples
- ▶ Proposal workflow and notifications
- ▶ Run management activities (quality assessment, calibration request, describe runs)
- ▶ DatCite (DOI) management
- ▶ DAQ integration and
- ▶ Data Source Groups management
- ▶ (...)

Governance & Resources

Next steps (& EOSC)

Next Steps

Areas of improvement

- Proposals search functionalities and information usability
- Metadata Client and Metadata Tools python libraries
- Data Audit and Reconciliation
- Data Processing pipelines integration
- Data Source Groups definition and integration within Karabo Control System
- Repositories, data migration and data calibration services integration with generic “Transfer Agent” distributed task Executer
- Data reduction R&D project (Event lists / multiple data assessment stages / etc.)
- Integration with the future electronic logbook – definition of interfaces
- Data Policy Enforcement Service
- (...)

Areas of improvement - EOSC related

- Datacite membership
- Automatised DOI generation by Local Contact and Leading Scientists
- Job Scheduler
- Data migration from XFEL/DESY campus to NCBJ (via Transfer Agent)
- data.xfel.eu
- Globus “integration”
- OAI-PMH integration
- Integration with cloud initiatives within PaNOSC (e.g. OpenAIRE, B2FIND, etc.)
- (...)

Links

myMdC: Metadata Catalog Web Server

<https://in.xfel.eu/metadata/> (website, intro, FAQs)

https://in.xfel.eu/gitlab/ITDM/metadata_catalog (code)

<https://in.xfel.eu/redmine/projects/metadata-catalogue> (questions, suggestions, issue tracking & project planning)

Email it-support@xfel.eu (for urgent support)

Metadata_client: Metadata Client python RESTful APIs client

https://in.xfel.eu/gitlab/ITDM/metadata_client (code)

<https://in.xfel.eu/redmine/projects/metadata-catalogue> (questions, suggestions, issue tracking & project planning)

Email it-support@xfel.eu (for urgent support)

Thank you!

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

Data Retention Policy – under review

Storage	Quota	Safety	Lifetime	Comment
dcache.raw	None	Tape Archive	6 months (t.b.c)	Raw data
raw	None	None	2 months (t.b.c)	Fast accessible raw data, lifetime not guaranteed
usr	5TB	Snapshots + Tape Backup	24 months (t.b.c)	user data, results
proc	None	None	6 months (t.b.c)	processed data e.g. calibrated
scratch	None	None	6 months	Temporary data (lifetime not guaranteed)
dcache.cal	None	Tape Archive	10 years	Calibration constants
cal	None	Tape archive	6 months	
user home	20GB	Snapshots + Tape Backup	Lifetime of the account	
archive.raw	None	-	Long-term	Long term means 5 years and XFEL will strive for 10 years
archive.cal	None	-	10 years	

myMdc – DB model

