

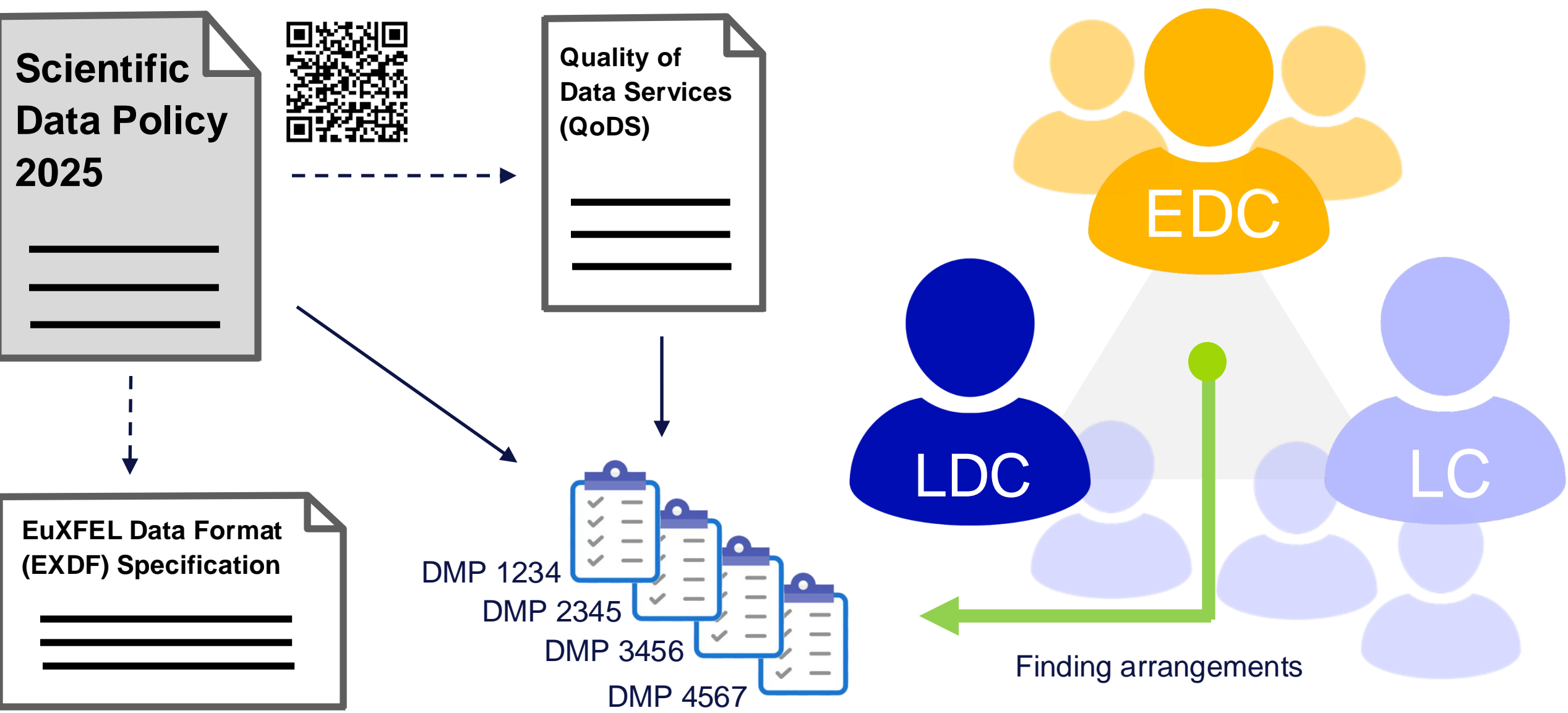
# Data Management Plans at European XFEL: first experience and outlook



Fabio Dall'Antonia, Luca Gelisio, Dennis Goeries, Steffen Hauf, Luis Maia, Janusz Malka, Philipp Schmidt and Krzysztof Wrona, on behalf of the data policy and data reduction working group. European X-ray Free Electron Laser Facility GmbH, Schenefeld, Germany.

See also:

- F. Dall'Antonia et al. *The new Scientific Data Policy of the European XFEL*. **Poster #**
- L. Maia et al. *Proposal Lifecycle Services at the European XFEL*. **Poster #**
- J. Malka et al. *Data Management Infrastructure for European XFEL*. **Poster #**



DMPs formulate agreements between the experiment team, represented by the experiment data contact (EDC) and the facility staff, represented by the local data contact (LDC) and local (instrument) contact (LC).

## Features and purposes

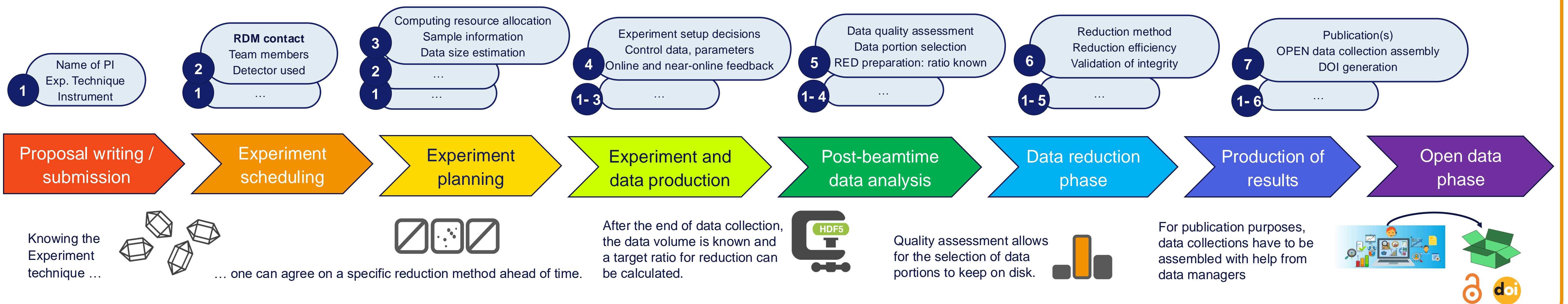
Data Management Plans (DMPs)<sup>[1]</sup> are a key instrument for the application of FAIR data practice<sup>[2]</sup> as anchored in the new Scientific Data Policy of the European XFEL and aim at:

- Optimising the beamtime planning, operation and data management through structured communication;
- Realising **specific** as well as **dynamic** workflows for each proposal by alignment of EuXFEL services to the individual needs of experiments;
- Providing documentation and reference for purposes such as future proposal submissions, experiment planning and (open) data-reuse.

## DMPs are guiding the proposal data lifecycle

DMPs will be integrated to the proposal workflow, incrementally adding information. Data present in proposal services like the user portal will automatically be injected. Updates will be possible any time.

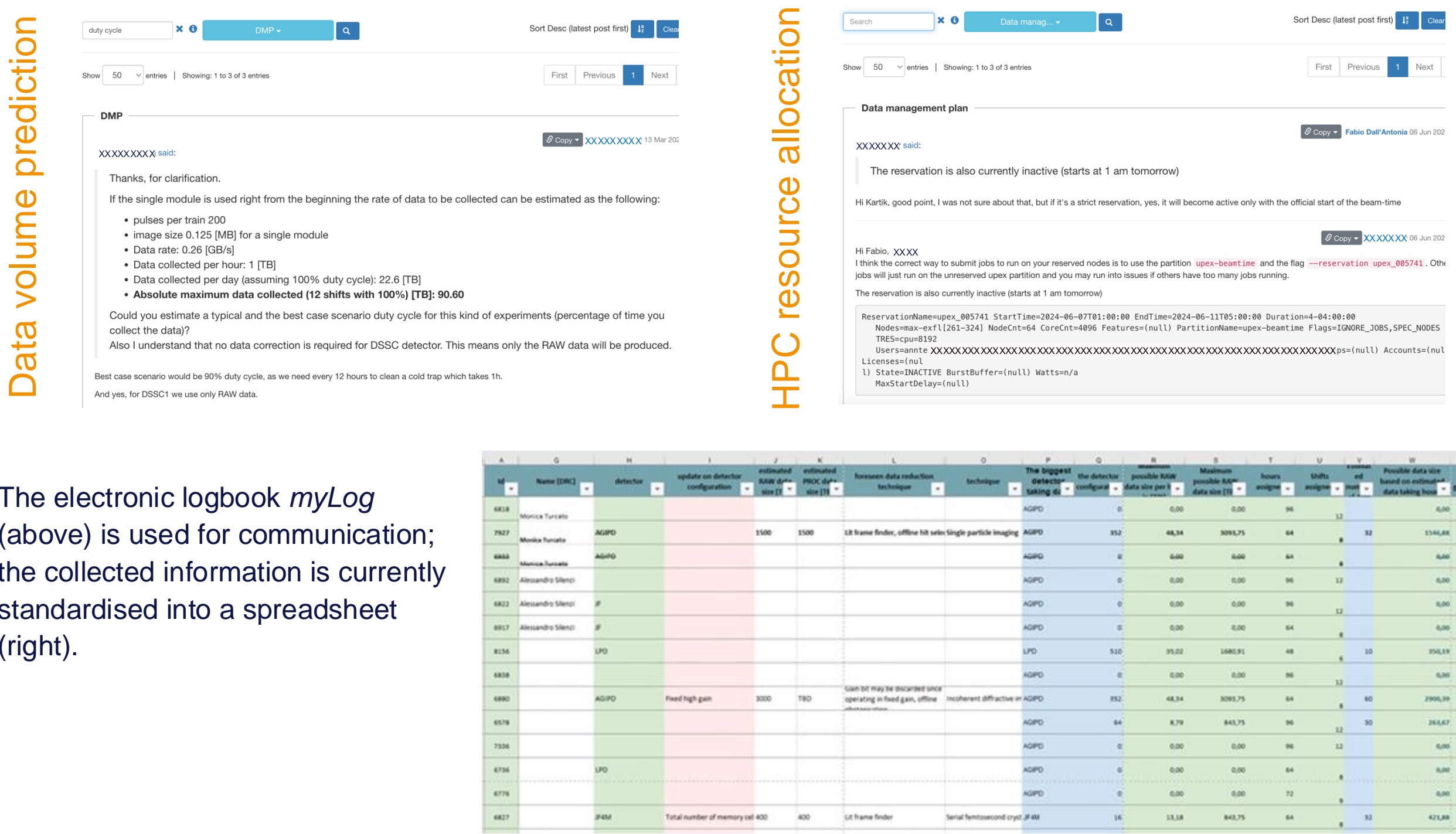
Inspired by:  
DOI 10.5281/zenodo.5639428 [1]



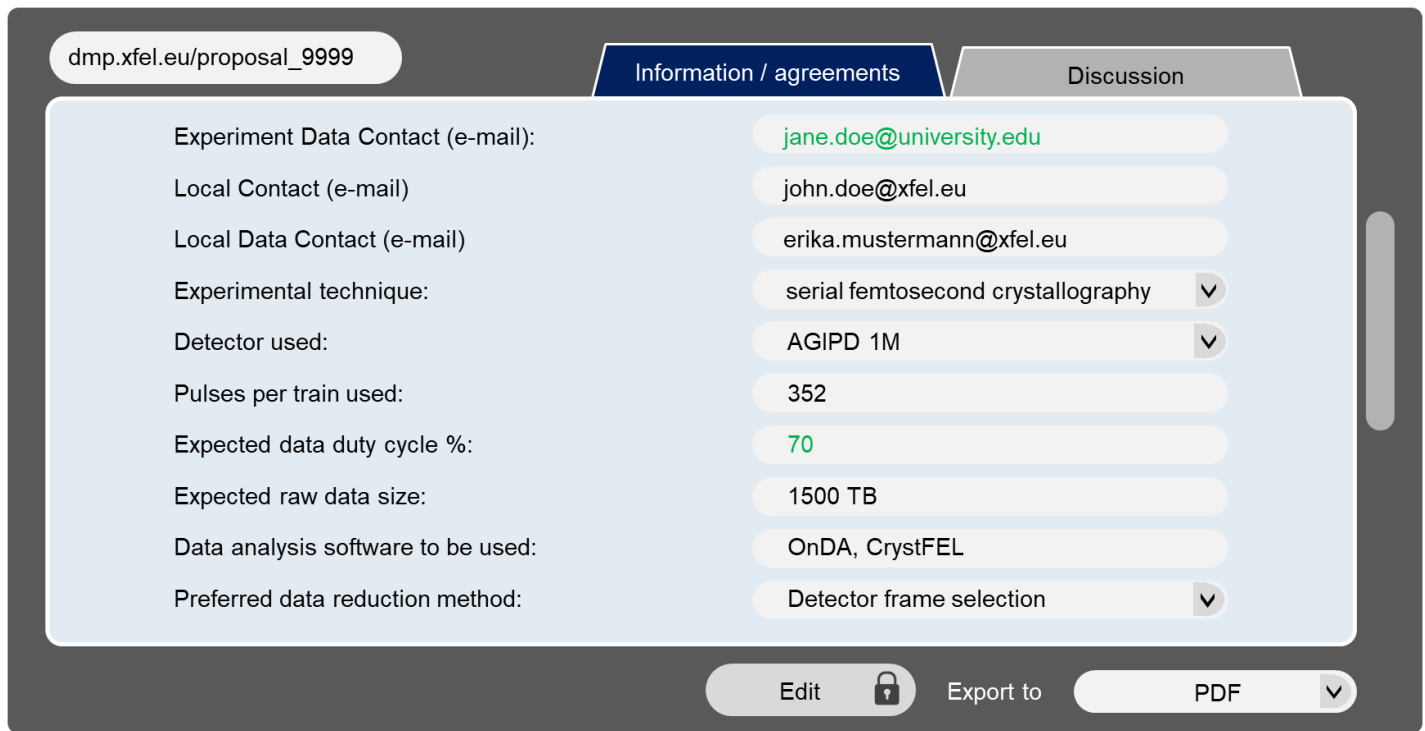
## Implementation roadmap

- 2024, Apr-June**  
First pilot phase with DMPs for six selected proposals. Main aim: exploring the scope of requirements
- 2024, Aug-Oct**  
Addressing data retention, reduction and archival for proposals beyond the end of embargo: Pilot for post-beamtime DMP arrangements in the future
- 2025, Feb-June**  
Second pilot phase with DMPs for all proposals of the run. Main aim: testing the upscaling of user support resources
- 2026+**  
DMPs are mandatory for all proposals by virtue of the new Scientific Data Policy

## DMPs in practice: standard procedures...



## The future



- DMP as a service
- Full integration of existing systems
- Automated workflows
- Dedicated interface

## ... and particular challenges

- Challenge: Large Data Volume**  
Selected pulse-trains prior to acquisition, reducing raw data by 70% (final dataset: 0.63 PB).
- Challenge: New Instrument Camera**  
Adapted an existing solution for online preview using expertise from another instrument.
- Challenge: Non-standard Software**  
Deployed custom software on Maxwell to enable non-standard tool usage
- Challenge: Non-standard Data Format**  
Converted time-tagged binary data to supported formats post experiment.

