PAUL SCHERRER INSTITUT

PSI

SLS 2.0 Controls and Science IT SubProject Update

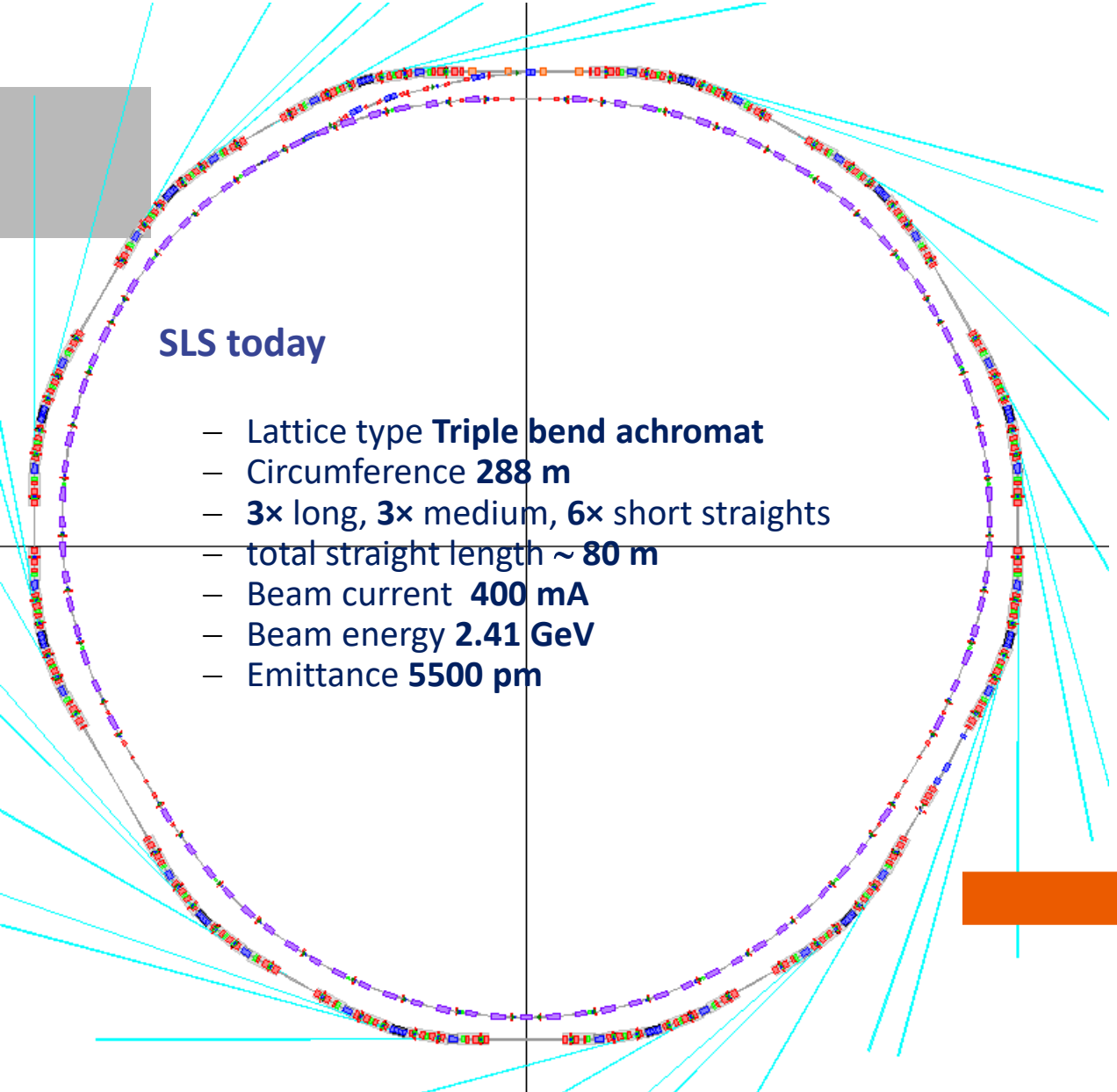
Disclaimer

- All the real work is being done by others
- For technical details please ask colleagues on the beamlines, in controls and science IT.
- I’m a manager and I don’t do the fun stuff anymore.

PSI Overview

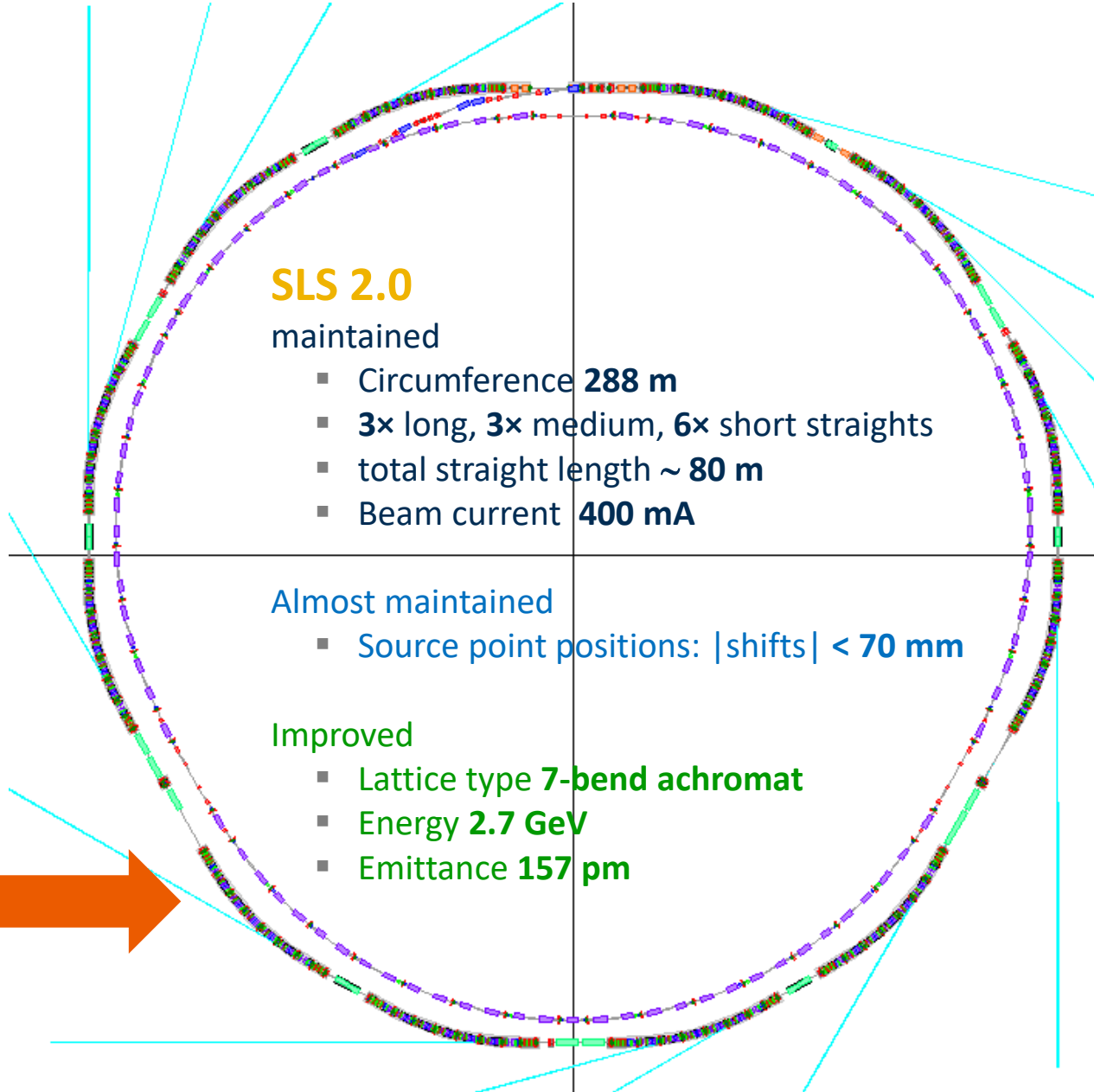
PAUL SCHERRER INSTITUT

SLS → SLS 2.0



SLS today

- Lattice type **Triple bend achromat**
- Circumference **288 m**
- **3×** long, **3×** medium, **6×** short straights
- total straight length ~ **80 m**
- Beam current **400 mA**
- Beam energy **2.41 GeV**
- Emittance **5500 pm**



SLS 2.0

maintained

- Circumference **288 m**
- **3×** long, **3×** medium, **6×** short straights
- total straight length ~ **80 m**
- Beam current **400 mA**

Almost maintained

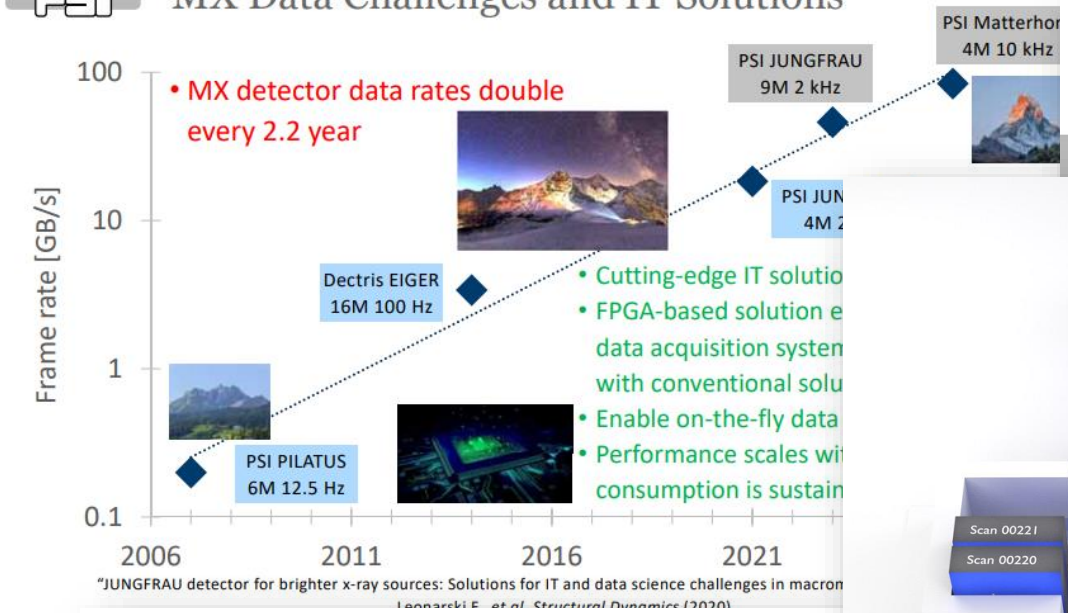
- Source point positions: |shifts| **< 70 mm**

Improved

- Lattice type **7-bend achromat**
- Energy **2.7 GeV**
- Emittance **157 pm**

The Opportunities and Challenges.

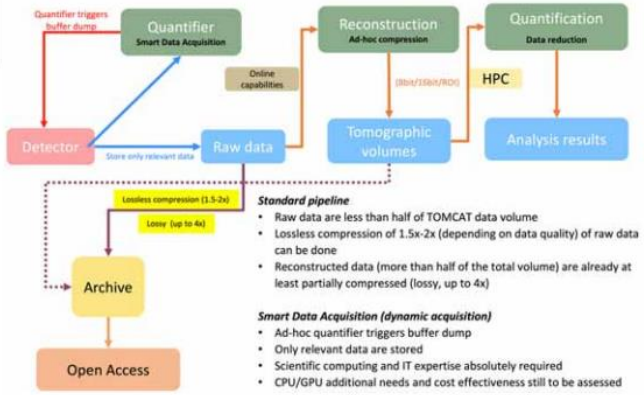
MX Data Challenges and IT Solutions



"JUNGFRAU detector for brighter x-ray sources: Solutions for IT and data science challenges in macrom

Leonarski F., et al., Structural Dynamics (2020)

Towards smart data acquisition



Standard pipeline

- Raw data are less than half of TOMCAT data volume
- Lossless compression of 1.5x-2x (depending on data quality) of raw data can be done
- Reconstructed data (more than half of the total volume) are already at least partially compressed (lossy, up to 4x)

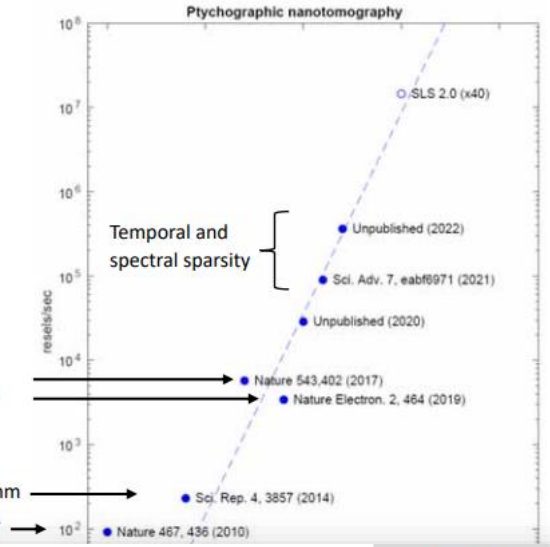
Smart Data Acquisition (dynamic acquisition)

- Ad-hoc quantifier triggers buffer dump
- Only relevant data are stored
- Scientific computing and IT expertise absolutely required
- CPU/GPU additional needs and cost effectiveness still to be assessed

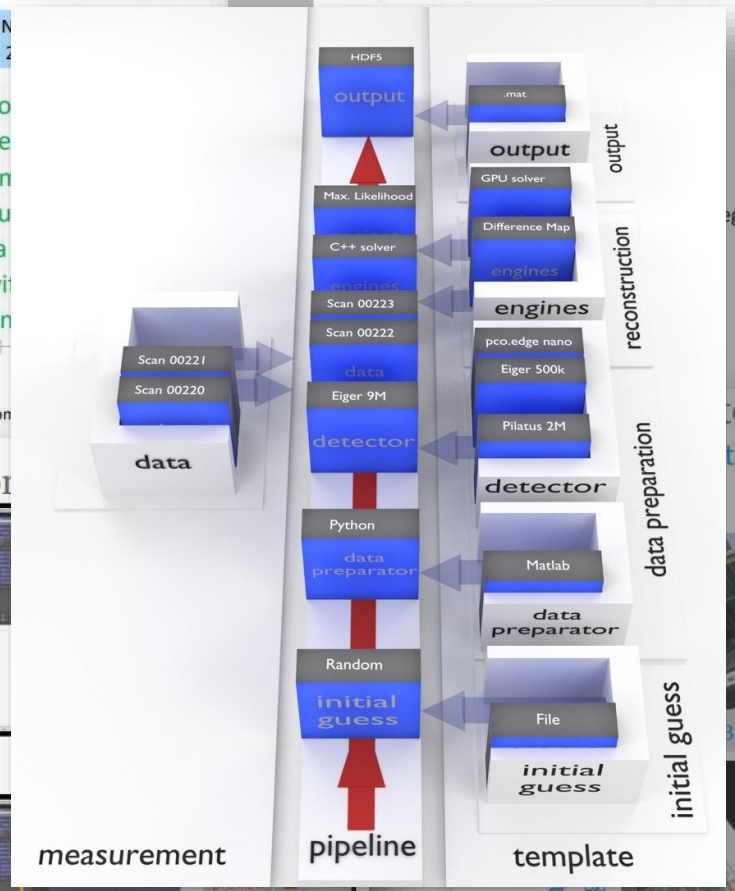
Buurlage et al. (2019), https://doi.org/10.1038/s41598-019-54647-4

Data challenges

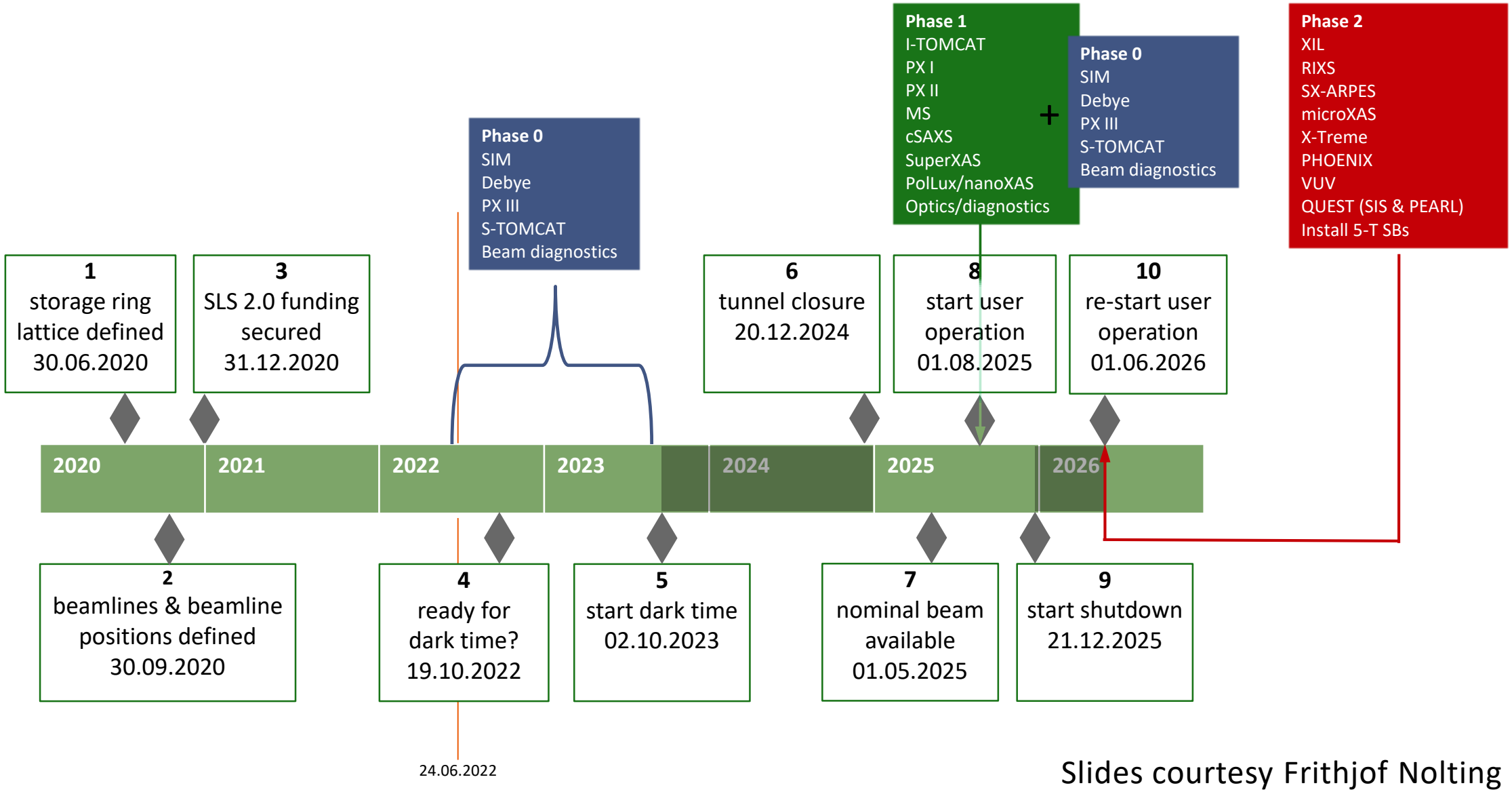
Ptychographic nanotomography is outrunning



egrated circuit at 15 nm
First laminography
Demonstration of 16 nm
First PXCT



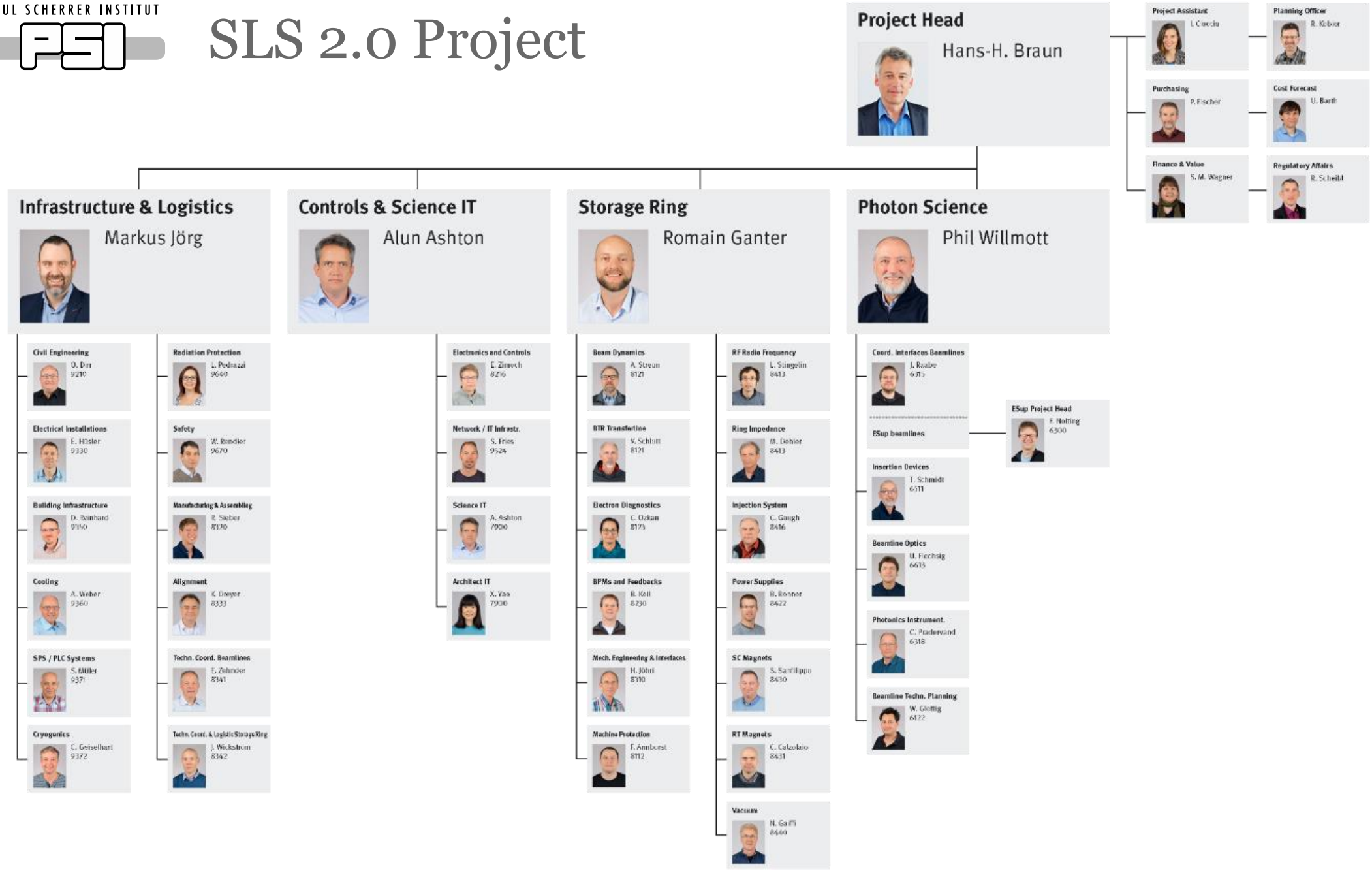
SLS 2.0 Timeline



Slides courtesy Frithjof Nolting

PAUL SCHERRER INSTITUT

SLS 2.0 Project



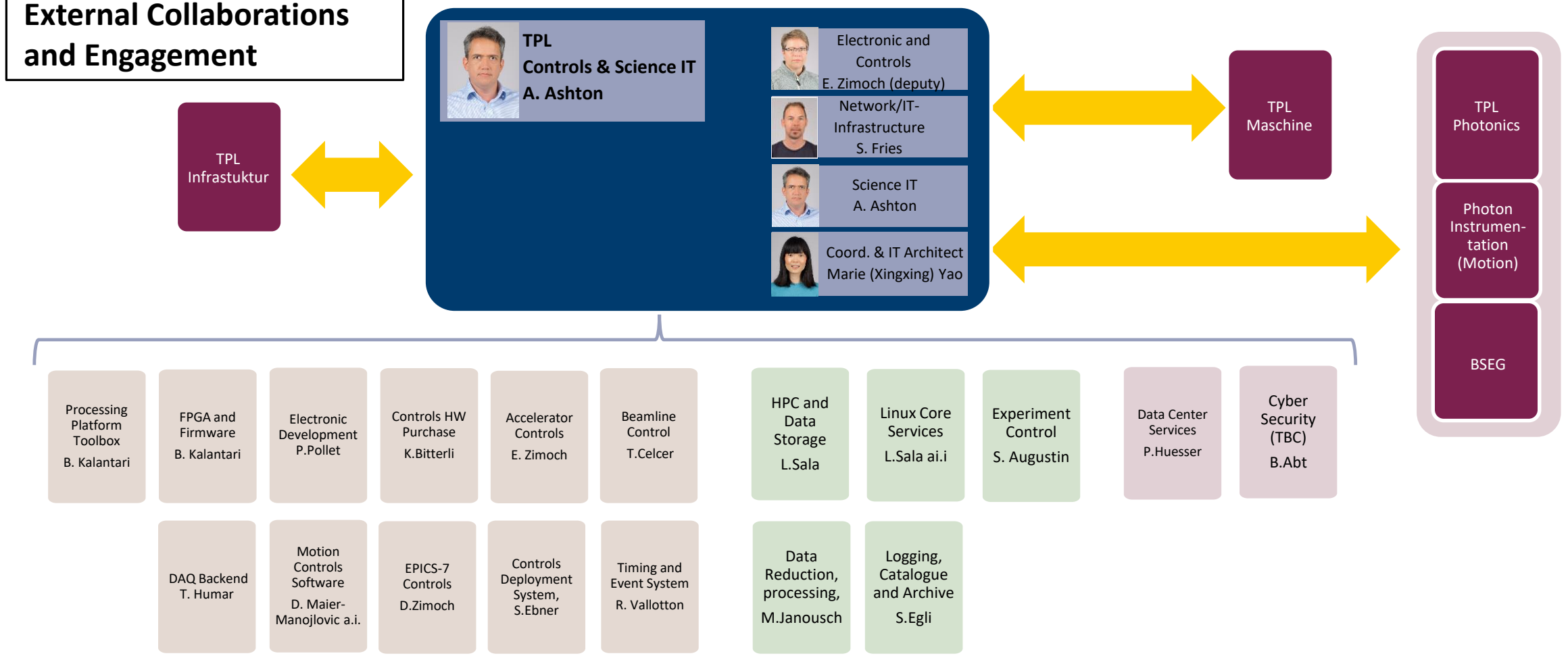
Controls and Science IT Sub-project for SLS 2.0

- A key driver for the establishment of CaSIT was to clearly define a hierarchy and relationship between software and infrastructure groups who are organisationally separated within PSI divisional structures
- Goals for the project include:
  - exploit new technical advances,
  - challenge and revisit existing approaches,
  - refine concepts where needed,
  - to facilitate improved development cycles and software quality
  - strengthened exchange between CaSIT group and Beamline groups covering the whole experiment/measurement lifecycle.
  - to coordinate interactions and exploit the various computer science and IT related activities throughout PSI, Switzerland and the wider synchrotron community.

Schnittstellen (Interfaces) für das Teilprojekt Controls & Science IT: from December 2021

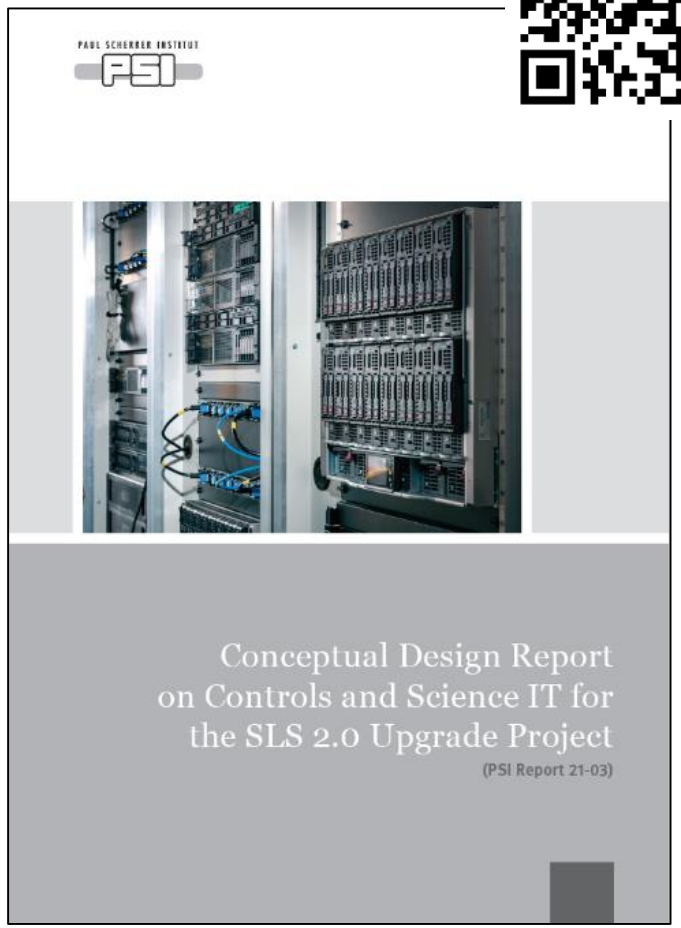
aCaSIT: advisory group for Controls and Science IT
(HBraun, TSchilcher, SBaymani, RPeterhans, AAshton, MarieYao, OBunk)

External Collaborations and Engagement



CaSIT Conceptual Design Report

- 21 authors, 82 pages
- Virtual review on 19th and 20th of May 2021
  - 16 talks
  - 6 reviewers from ESRF, APS, BESSY, Diamond, and SKA
- Variations in the level of detail in the report highlight the level of maturity of different services
- Published and available open access (Nov 2021): https://www.dora.lib4ri.ch/psi/islandora/object/psi%3A39514



Details in the CDR include

- Accelerator and Beamlines
  - EPICS 7
  - Compact PCI-Serial Toolbox
  - Integration of Network and Serial Devices
  - Timing and Event Systems
  - Motion Control
  - EPICS Gateways
  - Legacy Systems
  - GUI and Operator Level Applications
  - Standard Beamline DAQ
  - Networking

Revolution and Evolution

- Beamlines
  - Beamline Experiment Control
  - Data and Metadata Pipeline
  - Data Reduction, Reconstruction and Analysis
  - Data Storage and Computing
  - Experiment Information Management and Data Curation

Available IT resources and software services can easily facilitate or restrict:

Capabilities:

- What the beamlines can do
- How quickly they can develop, share or adopt new techniques

Operations:

- How well you can operate your beamlines
- Steer your experiment
- How well the beamlines can be supported

The experiment:

- How well you can track your experiment
- How much data you can handle
- How quickly you can analyze
- Publish and share your results and data

- Beamlines
  - Beamline Experiment Control
  - Data and Metadata Pipeline
  - Data Reduction, Reconstruction and Analysis
  - Data Storage and Computing
  - Experiment Information Management and Data Curation

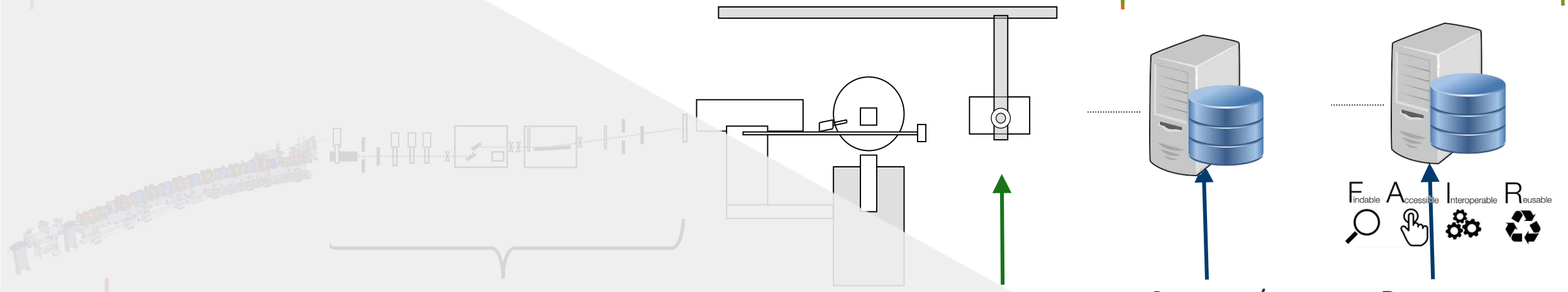
Revolution?

Beamline IT Environment Support @ SLS

@SLS: PSD Beamlines Groups
(predominantly responsible for experiment/analysis/processing software)

GFA-AEK/Electronics and Controls

SCD/Science IT



Findable Accessible Interoperable Reusable

Source

Beamline/Beam Conditioning

Detector

Storage/
Processing

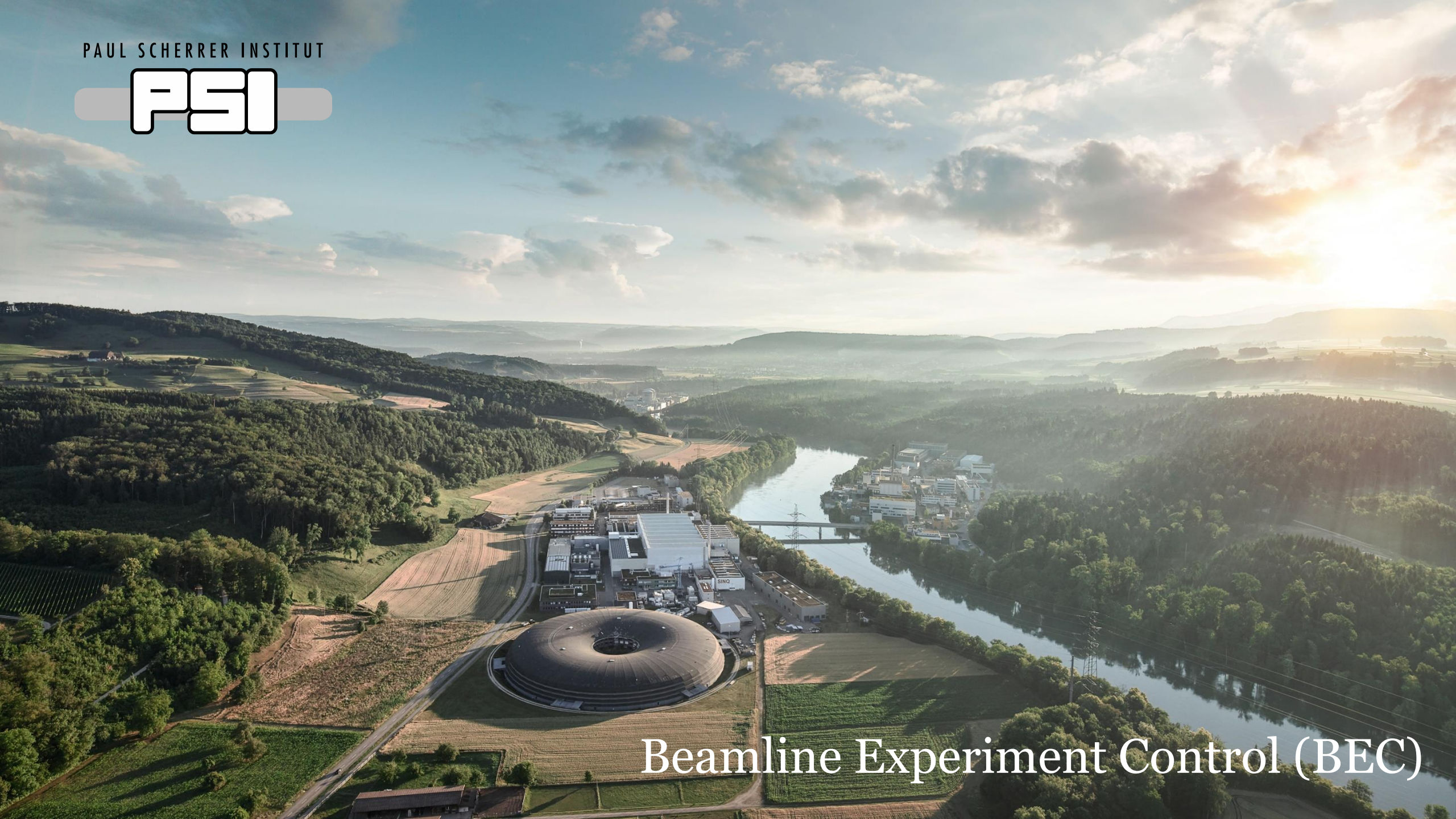
Remote
Resources

LOG-AIT/e.g. Networking

Focus for NoBUGS

- Accelerator and Beamlines
  - EPICS 7
  - Compact PCI-Serial Toolbox
  - Integration of Network and Serial Devices
  - Timing and Event Systems
  - Motion Control
  - EPICS Gateways
  - Legacy Systems
  - GUI and Operator Level Applications
  - Standard Beamline DAQ
  - Networking

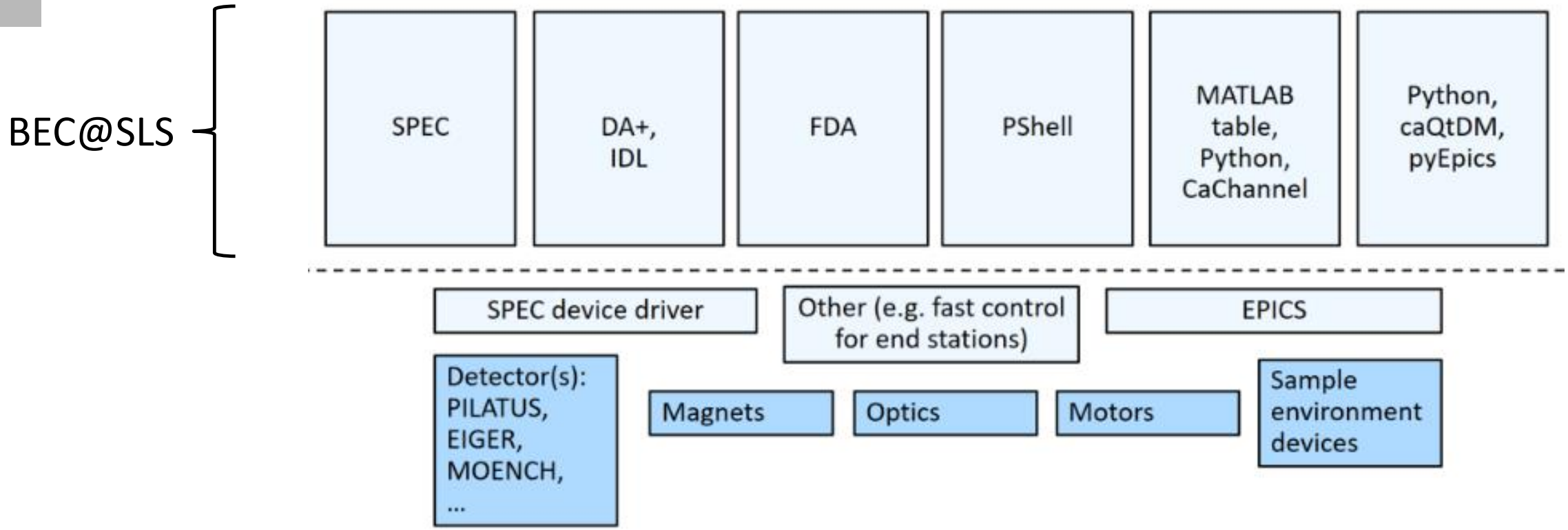
- Beamlines
  - Beamline Experiment Control
  - Data and Metadata Pipeline
  - Data Reduction, Reconstruction and Analysis
  - Data Storage and Computing
  - Experiment Information Management and Data Curation



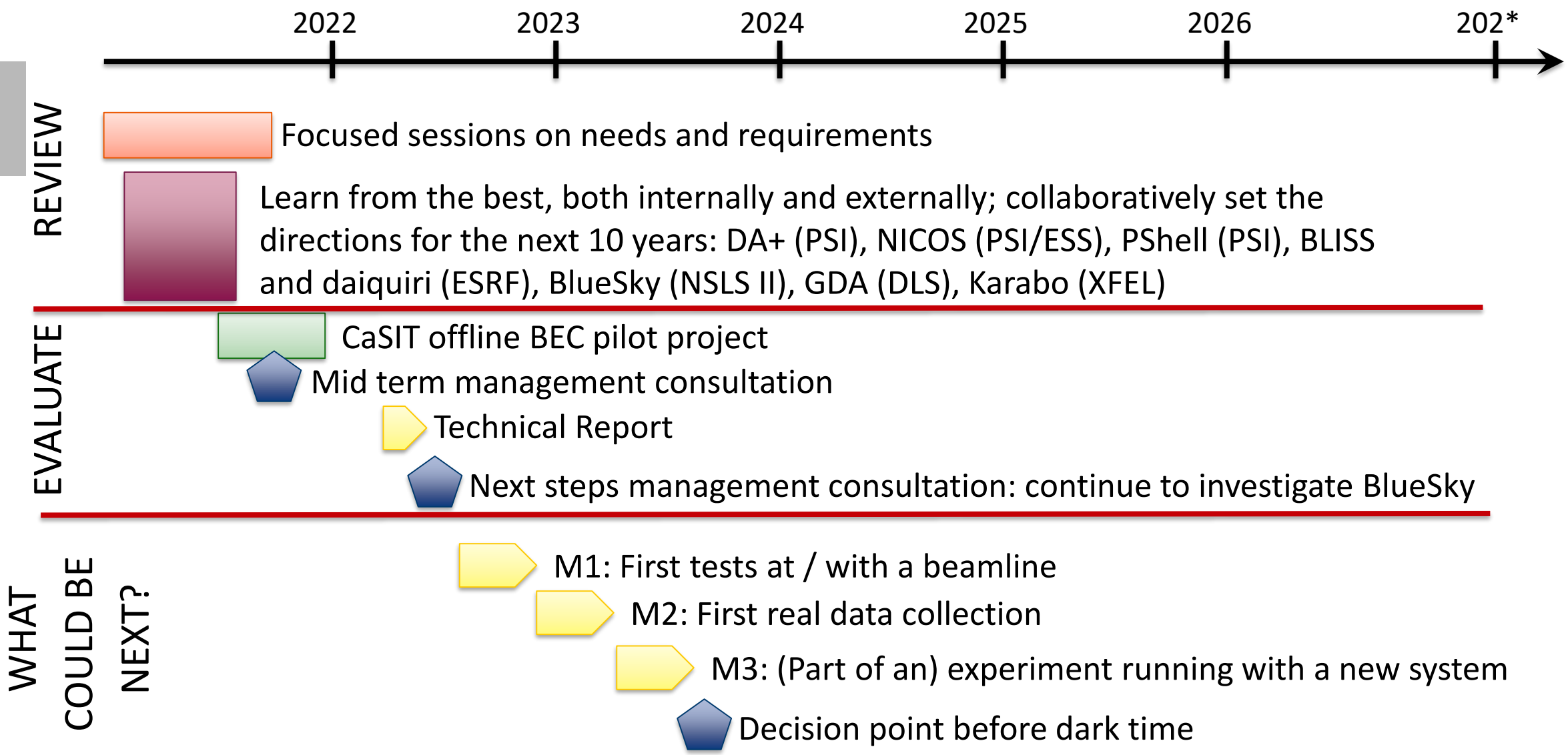
Beamline Experiment Control (BEC)

BEC Scope

- " … the layer between EPICS and the beamline operator…."



Beamline Experiment Control (BEC)



Bluesky is a Toolbox, not an Application

Current work at PSI is to establish the tools to use, feasibility of migrating, developing and supporting a new service layer with little change in resources

- **Assignment of responsibilities between three stakeholders**
  - **Controls groups**
  - **Science IT**
  - **Beamline staff**
- **Managing expectations, capacity, enthusiasm and/or anxiety.**

Beamline Experiment Control

- Eventual Decision
  - Responsibility: PSD Lab Heads with SLS 2.0 beamline responsibilities, AEK Department Head, AWI Department Head, SLS 2.0 Project Head
- Decision Criteria
  - Suitability (based on requirements, goals and risks outlined in CDR)
  - Ease of migration
    - how easily the beamlines will be able to migrate their existing functionality
  - Level/model of operational support
    - the level of investment, training and new resources needed for operational support and how that can be shared between different skillsets
  - Broad community and good support to reduce risks associated with single point of failure and improve shared development costs and sustainability
- Essential for e.g. transferable experience on smart data acquisition, remote access/control, data standards, data reduction, processing and analysis, open data…..

PAUL SCHERRER INSTITUT

PSI

Data and Data Treatment

Data Reduction, Reconstruction and Analysis

From the requirement gathering

- Optimize existing **data processing software and hardware** to prepare for increased data volume and speed, and faster experiments, for both online processing (for live visualization and feedback to a running experiment) and offline processing (toward publishable data).
- Dedicate **personnel resources and data processing services** to provide a more complete package of larger user facilities, data centre, computing infrastructure and data processing services to improve the competitiveness of the SLS 2.0 and ensure its long-term success.
- **Link experiment measurements and findings to advanced computational methods** such as simulation and modelling, data mining, machine learning and artificial intelligence.

Data Reduction and Compression.

Example activities

- MX: Task specific hardware
- Imaging: lossy compression
- HDF5 or native compressions
- Local beamline initiatives joint with detector group
- Region of interests
- Memory based camera hardware compression

- LEAPS-INNOV
- REDML: collaboration between PSI, the SDSC and the CSCS to reduce high volume experimental data using Machine Learning.

- Current/perceived drawbacks,
  - Data loss
  - Data Feature loss/degradation in scientific data
  - Computationally expensive
  - Still challenging to put into data streams/operations
  - No centrally coordinated activity
  - Data usability by data owners

PAUL SCHERRER INSTITUT

Data Reduction, Reconstruction and Analysis

DA+ data acquisition and analysis software at the Swiss Light Source macromolecular crystallography beamlines

Justyna Aleksandra Wojdyla, Jakub W. Kaminski, Ezequiel P
Simon Ebner, Xiaoqiang Wang, Jose Gabadinho and Meitian

Swiss Light Source, Paul Scherrer Institute, 5232 Villigen, Switzerland.
*Correspondence e-mail: meitian.wang@psi.ch

Data ac crystallo synchro acquisit beamlin Python The maj processing distribu explorati beamlin and dev

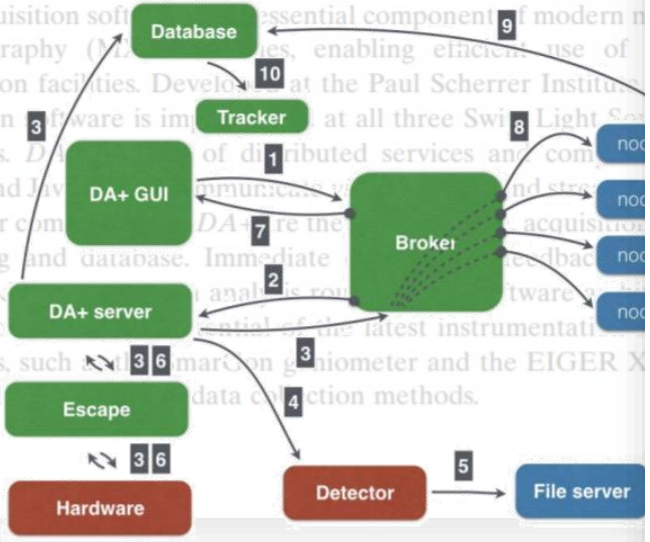


Figure 1
Schematic representation of the software infrastructure at the SLS MX be components are shown in green boxes, hardware components in red boxes, a computing nodes in blue boxes. Lines indicate interactions between different numbers show the order of workflow (a detailed description is given in in §3.1) message broker is a major communication hub used by *DA+* daq software compon experiment parameters in the *DA+* GUI, while *DA+ server* carries out da communicates with detector and hardware *via* basic state machine *escape*. *Adp* message from the broker, start data processing and send results to the *mxdb* database. Results of *adp* are displayed in the web-based *adp-tracker*.

PAUL SCHERRER INSTITUT

Research at PSI | Research Divisions and Labs | Facilities and Instruments

cSAXS beamline software packages

Base package

Basic functionalities for file reading and radial integration and plot provides a lot of functions used in the other packages.

– cSAXS base package

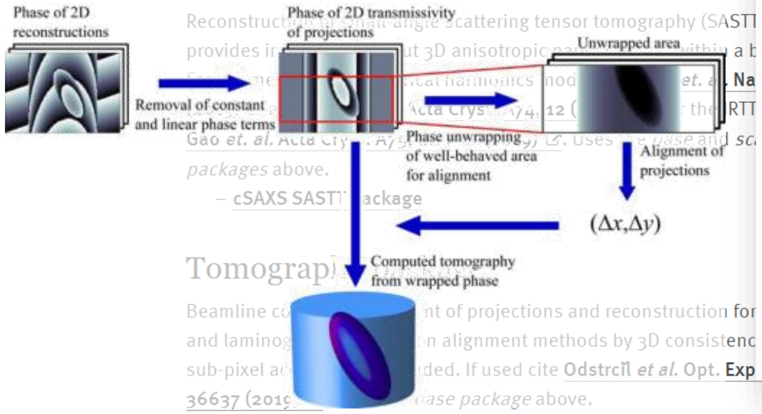
Scanning SAXS package

Analysis and plotting of scanning SAXS, main orientation of scatter orientation. Please cite Bunk *et. al.* New J. Phys. 11, 123016 (2009) *base package* above.

– cSAXS scanning SAXS package

SASTT package

Reconstruction ... scattering tensor tomography (SASTT provides ... 3D anisotropic ... *packages* above.



– cSAXS SASTT package

Tomography

Beamline c... nt of projections and reconstruction for and laminog... n alignment methods by 3D consistenc sub-pixel ac... ded. If used cite Odstrcil *et al.* Opt. Exp 36637 (2019 ... *ase package* above.

– cSAXS tomography package

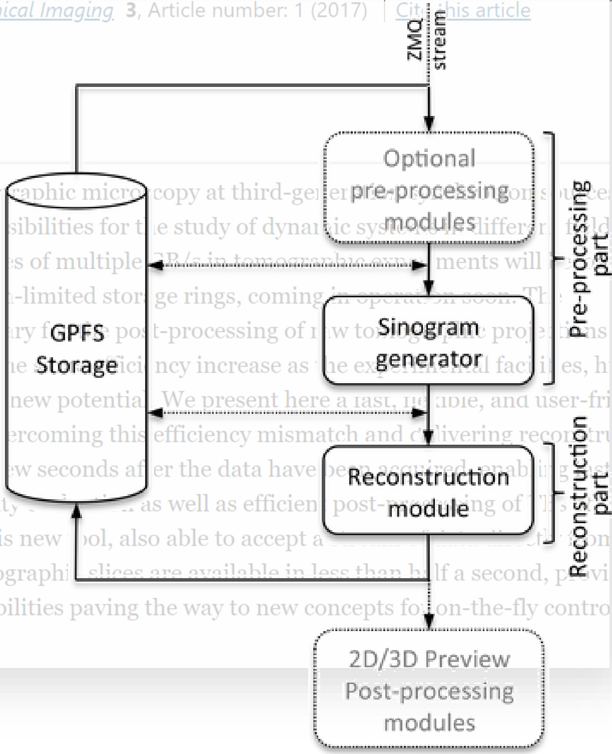
Towards on-the-fly data post-processing for real-time tomographic imaging at TOMCAT

Federica Marone, Alain Studer, Heiner Billich, Leonardo Sala & Marco Stampanoni

Advanced Structural and Chemical Imaging 3, Article number: 1 (2017) | Cite this article

Abstract

Sub-second full-field tomographic microscopy at third-generation ... reality, opening up new possibilities for the study of dynamic systems ... Sustained elevated data rates of multiple GB/s in tomographic ... more common at diffraction-limited storage rings, coming ... computational tools necessary for post-processing of raw ... generally not experienced ... efficiency increase as ... optimal exploitation of this new potential. We present here a fast, flexible, and user-fri... post-processing pipeline overcoming this efficiency mismatch and delivering ... tomographic datasets just a few seconds after the data have been acquired ... parameter and image quality ... as well as efficient post-processing of ... tomographic data. With this new tool, also able to accept a ... detector, few selected tomographic slices are available in less than half a second, provi... advanced previewing capabilities paving the way to new concepts for on-the-fly contro... dynamic experiments.



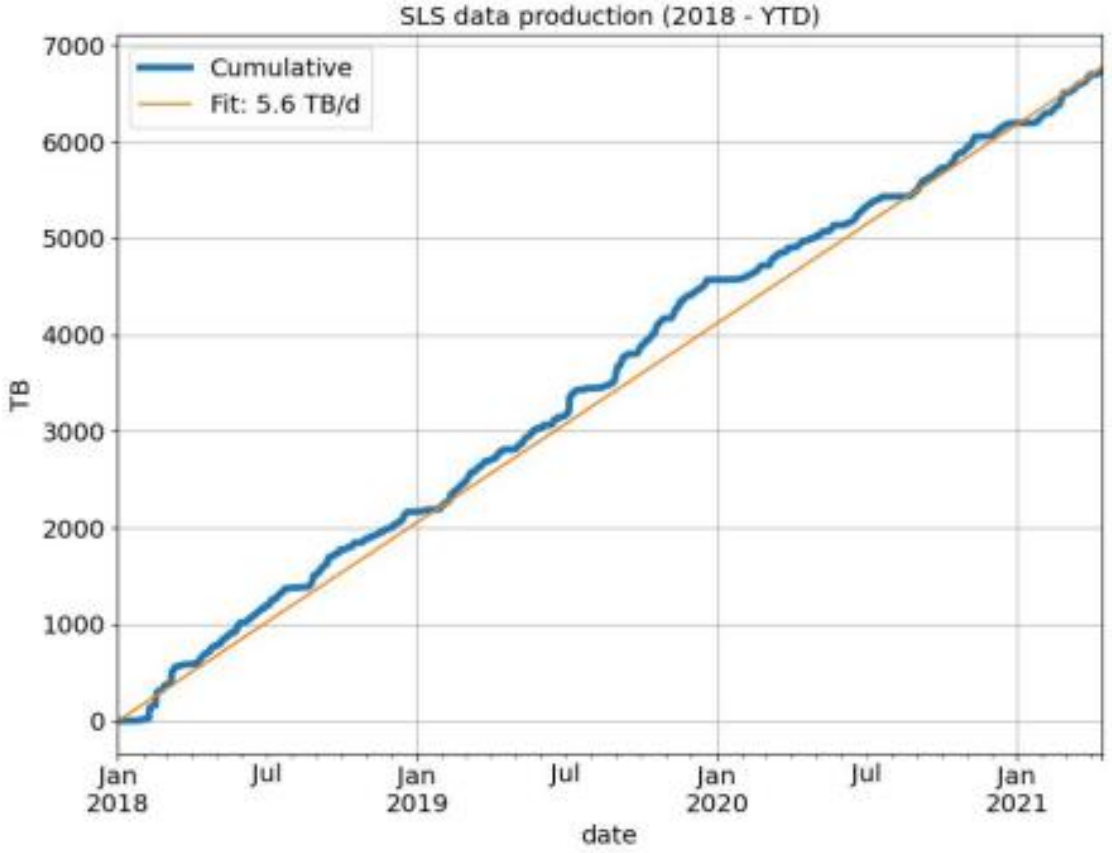
PAUL SCHERRER INSTITUT

PSI

Data Storage and Computing

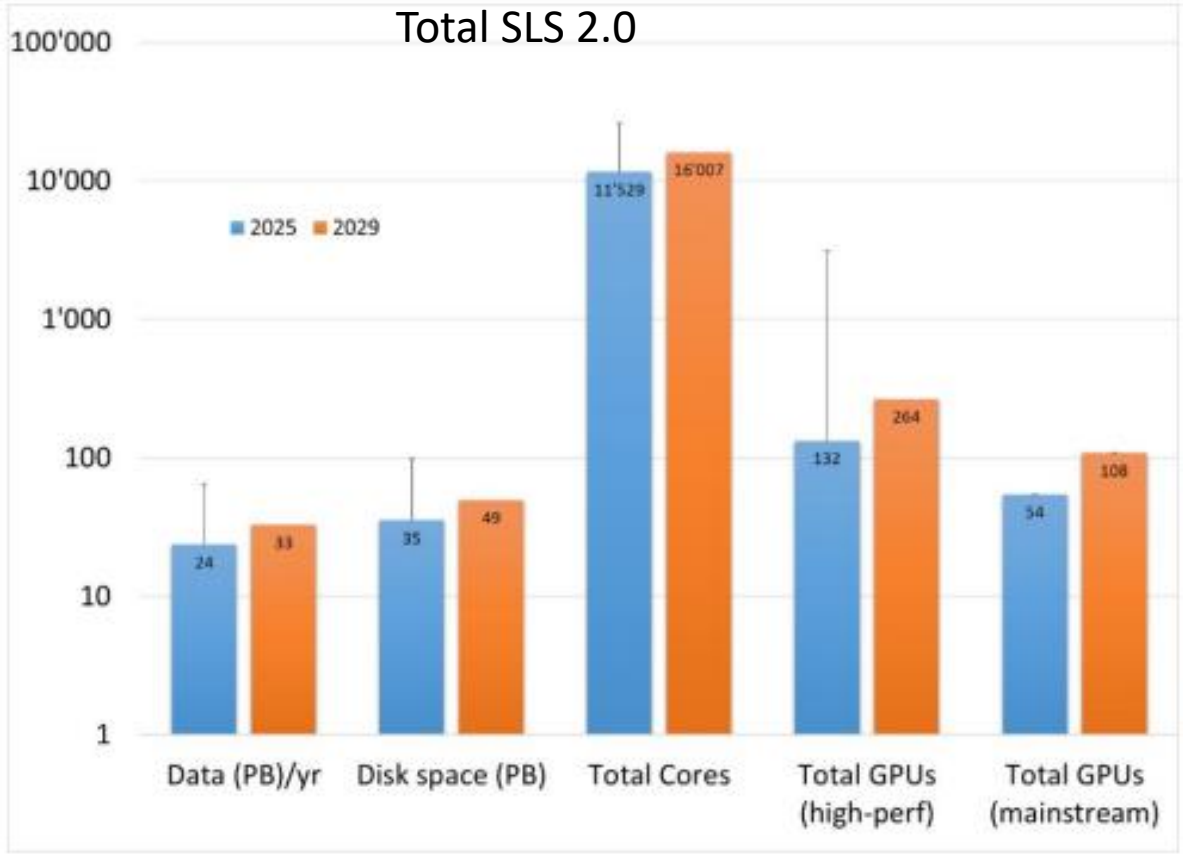
Data Storage and Computing (summer 2021)

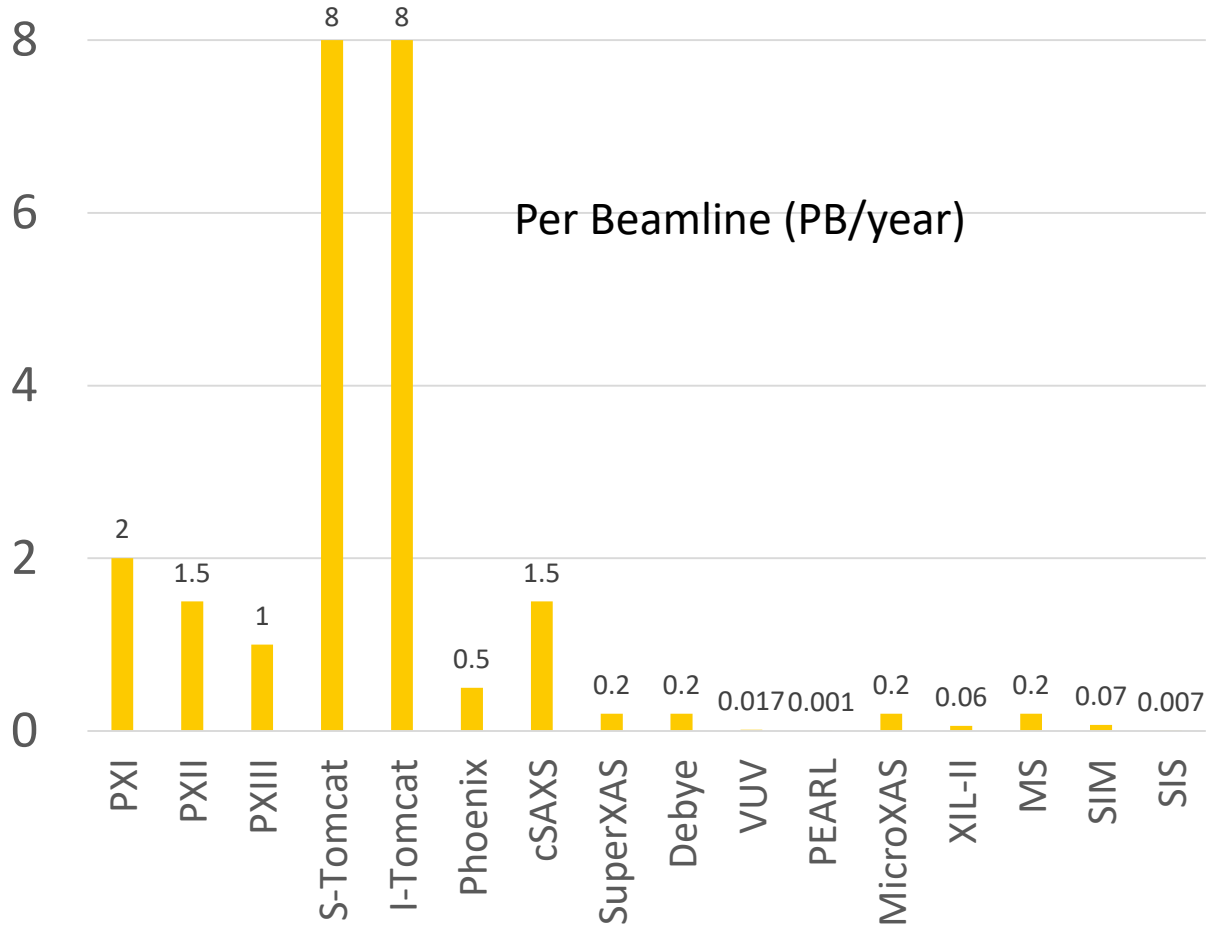
- Historically split into:
  - online - dedicated compute, storage
  - offline / data analysis - DaaS,
    - shared between SLS and SwissFEL



	SLS	DaaS (SLS + SwissFEL
Storage: PB	3.2	11
Storage: GB/s	25	90
Compute: cores	2360	2000
Compute: RAM (TB)	2	18
Compute: GPUs	4	12

Forecast for 2025

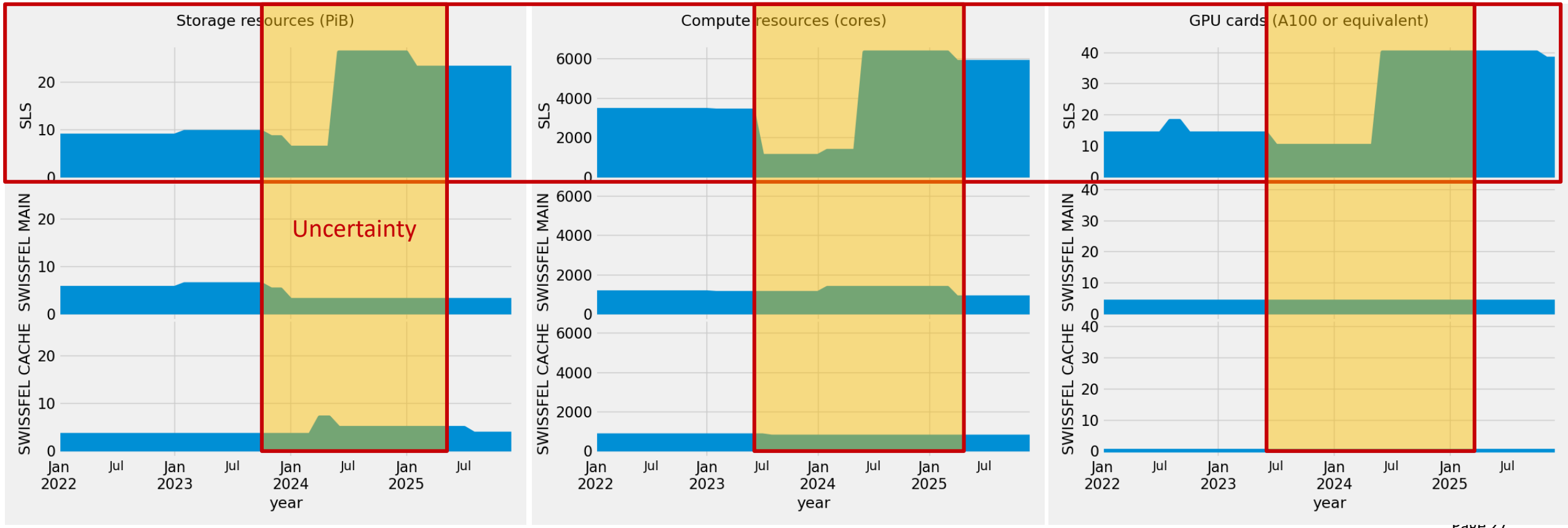




Storage and compute updates

IT resources planning with SLS2 project- and PSD baseline-budget (*build to budget – first draft*)

- SLS resources reduced during dark period, but analysis and tests still possible
- Detailed numbers will depend on requirements, technology (e.g. A100 vs A10 cards)
- Shared resources with SwissFEL and need to move server rooms



Long term offline data processing services.

2022 2023 2024

CSCS Collaboration

CSCS ALPS Project

ALPS mi

CSCS
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

ETH zürich

ABOUT SERVICES COMPUTERS USER LAB SCIENCE PUBLICATIONS EVENTS

USER PORTAL

Science | Computer Science &...

CSCS, HEWLETT PACKARD ENTERPRISE AND NVIDIA ANNOUNCE WORLD'S MOST POWERFUL AI-CAPABLE SUPERCOMPUTER

"Alps" system to advance research across climate, physics, life sciences with 7x more powerful AI capabilities than current world-leading system for AI on MLPerf.

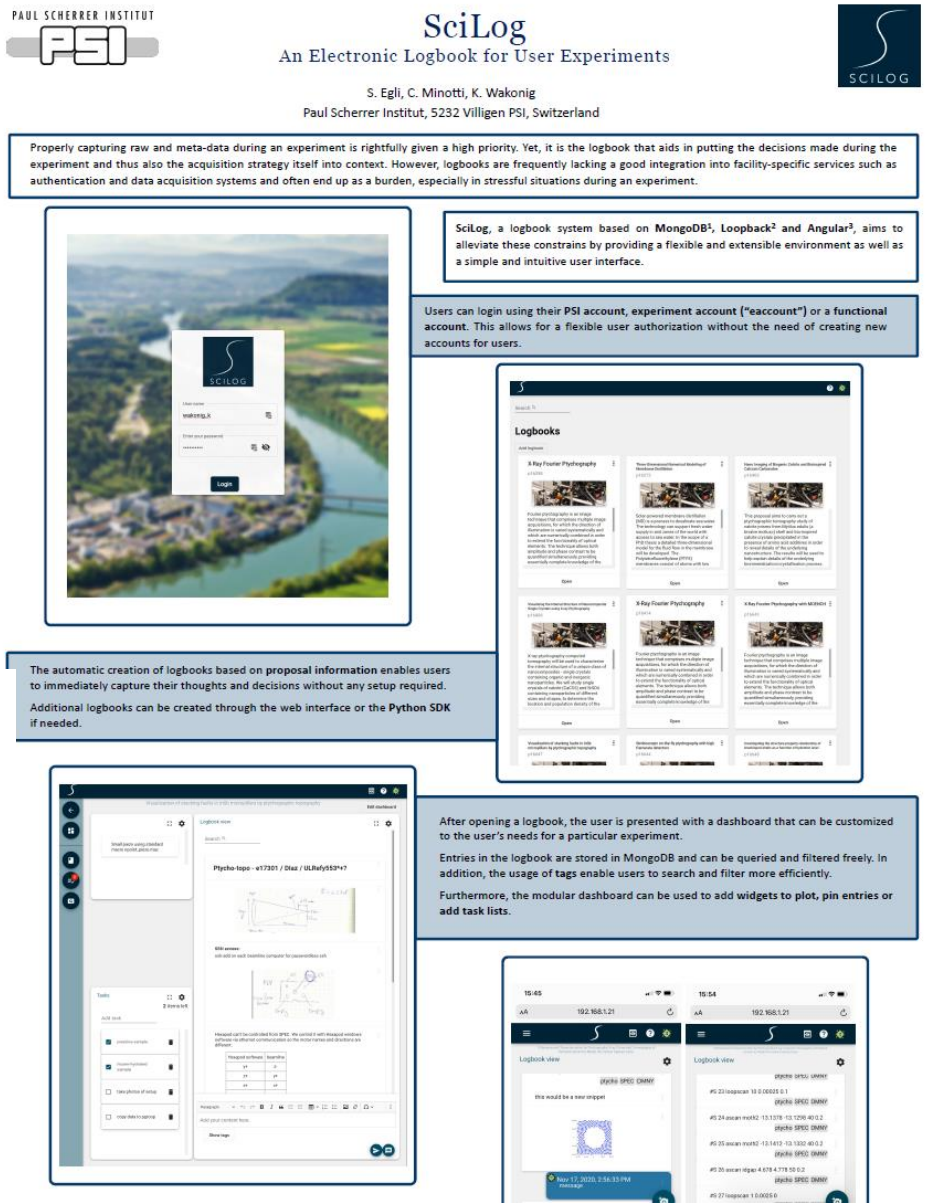
PAUL SCHERRER INSTITUT

PSI

Experiment Information Management and Data Curation

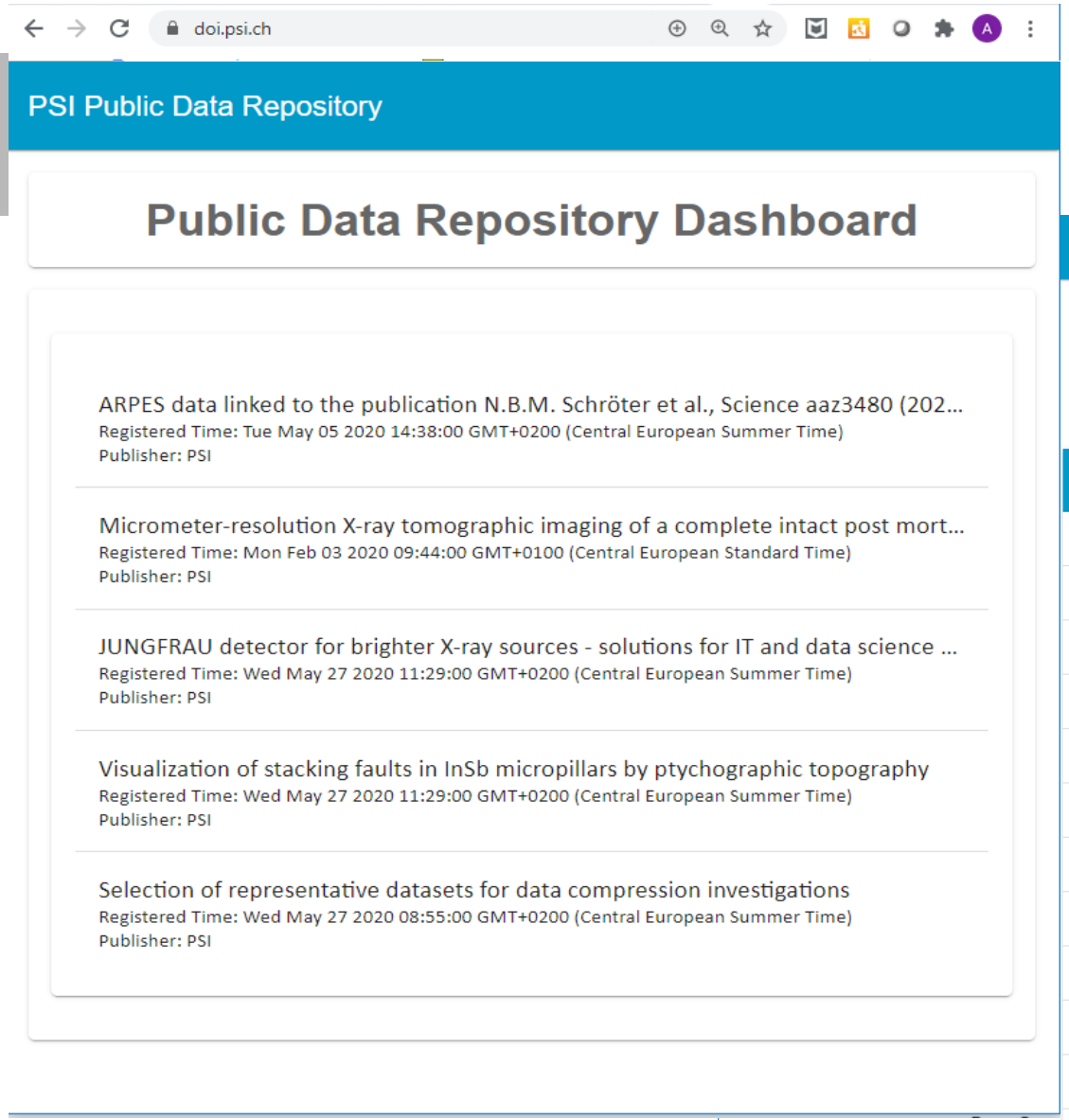
Experiment Information Management

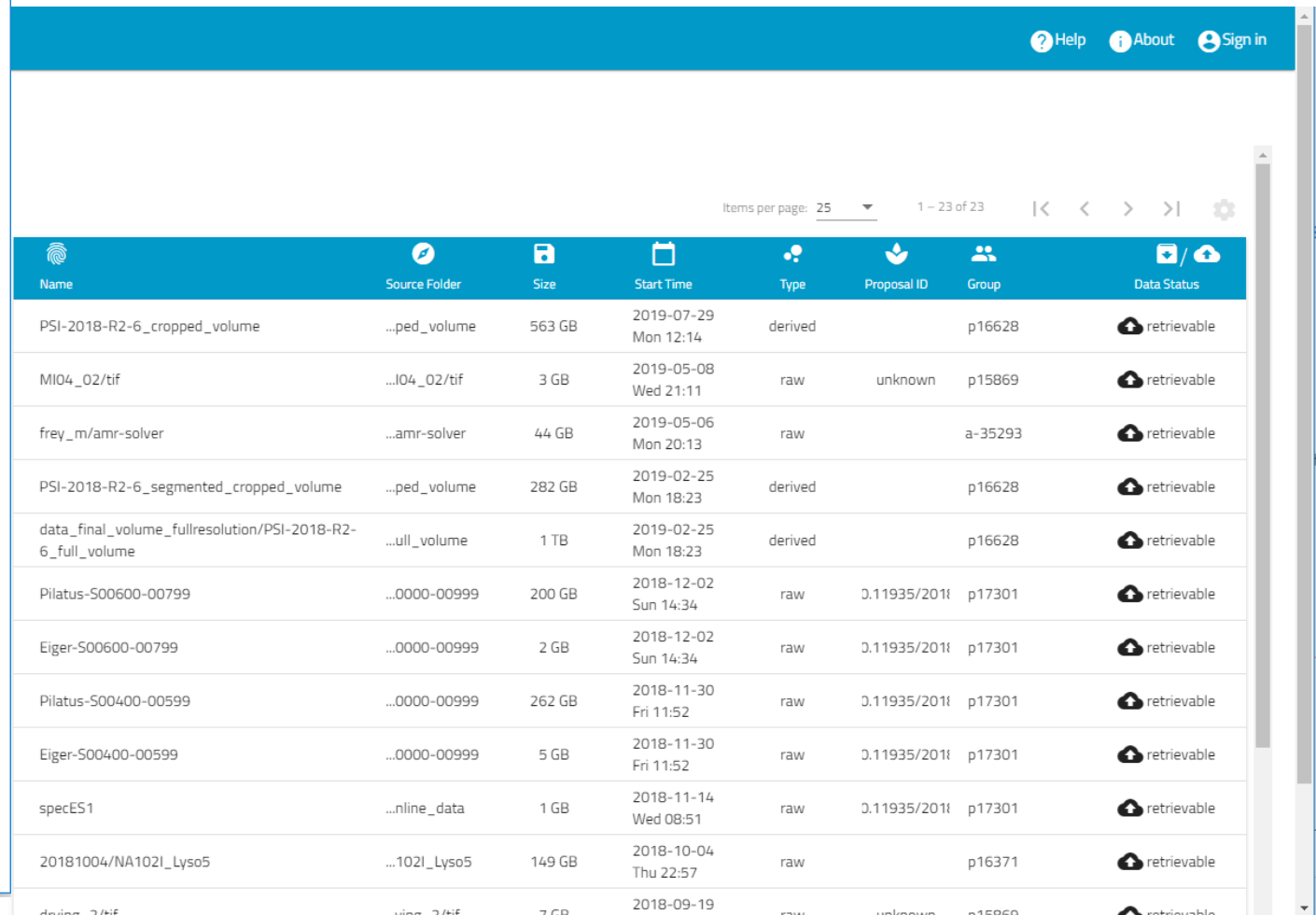
- SciLog – An Electronic Logbook for User Experiments
- https://github.com/paulscherrerinstitute/scilog
- Initiatives (funded) within the wider ETH domain to explore synergies and platform sharing in preparation.



PAUL SCHERRER INSTITUT

Data Curation: see Carlo Minotti presentation





PAUL SCHERRER INSTITUT

PSI

Thanks to

- CaSIT and SLS 2.0 project team
- PSD/SLS beamline staff
- AEK, AIT, AWI…
- National and international collaborators/partners