



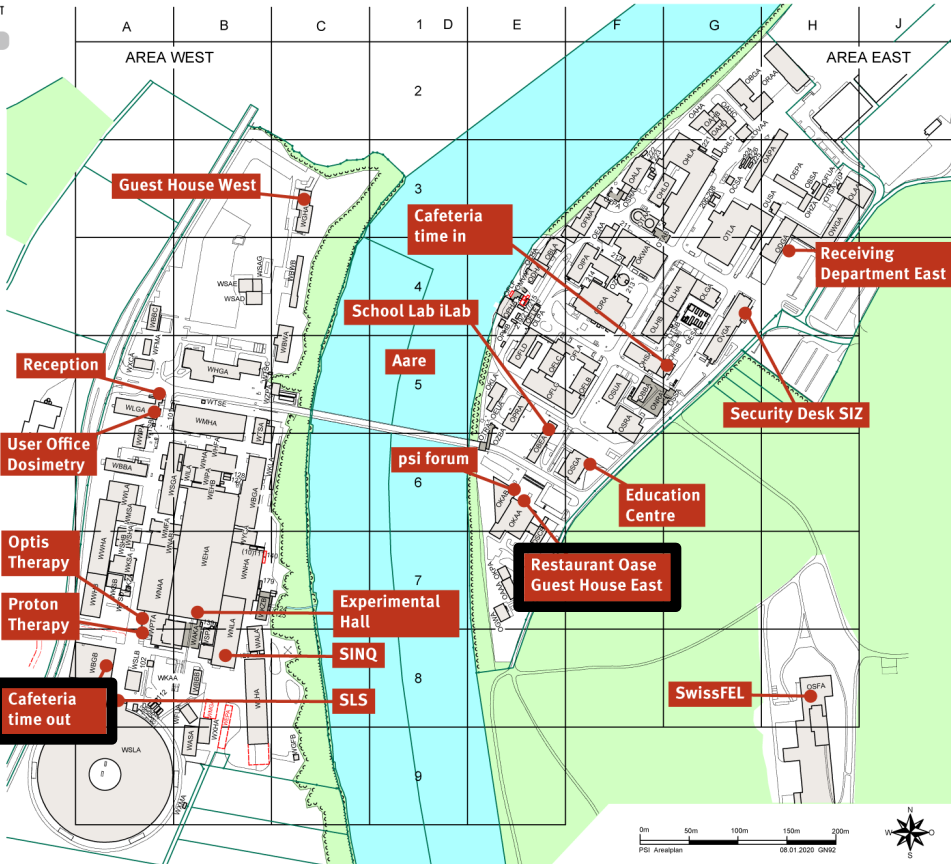
2024 EM Standards Workshop

22–23 February 2024

Paul Scherrer Institute, Villigen, Switzerland

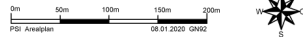
Legende PSI-West

- B8 WALA
- B8 WASA
- B6 WEGA
- A8 WRGB
- A6 WBBA
- B8 WBBB
- C5 W8HA
- B4 WBWB
- A6 WWSA
- A4 WBBC
- B7 WEHA
- B8 WFLA
- A5 WFMA
- B6 WEHB
- C3 WCHA
- C9 WGF8
- B5 WIGA
- A7 WWSB
- B6 WPA
- B6 WHFA
- B6 WHA
- A7 WWSB
- B6 WPA
- C6 WKLA
- A7 WWSA
- B8 WKA
- B8 WKA
- A7 WKZA
- B7 WKZB
- A5 WLGA
- B8 WLHA
- A7 WFLA
- A9 WXMA
- B5 W8HA
- B7 W8HA
- B7 W8LA
- C7 WYCA
- A7 WKA
- A8 WSLA
- A8 WSLB
- B4 WSGE
- B4 WAG
- B4 WSAD
- A6 WSGA
- B8 WSP
- A7 WSP
- A7 WSA
- B5 WTS
- A5 WTS
- B5 WTS
- A7 WTS
- B5 WTS
- A7 WTA
- A7 WNA
- A5 WCA
- A6 WMA
- A7 WWA
- A7 WWB
- A6 WWA
- C5 WZPA



Legende PSI-Ost

- F3 OALA
- G2 OAHA
- G2 OAHG
- G2 OAHD
- H3 OAPA
- G2 OBKA
- H3 OBBA
- E6 OZBA
- H2 ORKA
- F5 OFLA
- F5 OFLB
- E5 OFLG
- F5 OHSA
- F5 OSUA
- F5 OSBA
- F5 OHSB
- G2 OAHB
- OZBA
- G3 OCCA
- G3 OCCS
- E4 OCHA
- F3 OTAB
- H3 ODGA
- F4 ORA
- E5 OSLA
- H3 OPEA
- G4 OESA
- E4 OSLA
- E5 OFLG
- H3 OFJA
- F3 OFMA
- F5 OFLA
- F5 ONRA
- H3 OHZA
- G3 OHLA
- F3 OHLD
- G3 OHLE
- E5 OFLD
- E7 OKPA
- E7 OGWA
- E7 OMAA
- F4 OKWA
- E5 OKLA
- E4 OKHA
- E5 OFLF
- G4 OLGA
- G4 OLHA
- F4 OLHB
- E4 OFLE
- E4 OLPA
- H4 OLAA
- E4 OFLE
- E4 OLVA
- F3 OGAA
- F3 OGKA
- E5 OPRA
- E6 OKAB
- E4 OPHA
- F4 OIPA
- G5 OSSA
- F3 OSBA
- E4 OPBH
- F3 OSBA
- F6 OSGA
- E7 OSGB
- H3 OSFA
- F3 OTAA
- G3 OTLA
- E5 OTBA
- H3 OTSA
- G4 OTSB
- H3 OUSA
- G2 OVA
- G4 OVGA
- H3 OVKA
- F4 OZFA
- F3 OZPA



12:00	Lunch (East Campus)	OASE Restaurant
13:30	Walk to west campus, get badges, coffee	WBGB/019
Session 1: Motivation		
14:00	Welcome	Spencer Bliven
14:15	Swiss OpenEM Project	Yves Tittes
14:30	INSTRUCT	Jose Maria Carazo
14:45	Summary of goals and progress	Jose Maria Carazo
Session 2: Existing data models		
15:00	Repositories (EMDB)	Kyle Morris
15:20	Facilities (ISPyB)	Alex de Maria, Jirka Novacek
15:40	Workflows (Scipion)	Carlos Oscar Sanchez Sorzano
16:00	Coffee Break	
Session 3:		
16:30	Planning session <ul style="list-style-type: none"> • Define scope • Approve name • Set workshop expectations 	Unified discussion
17:00- 18:00	Split into working groups	Working groups
18:30	Dinner	

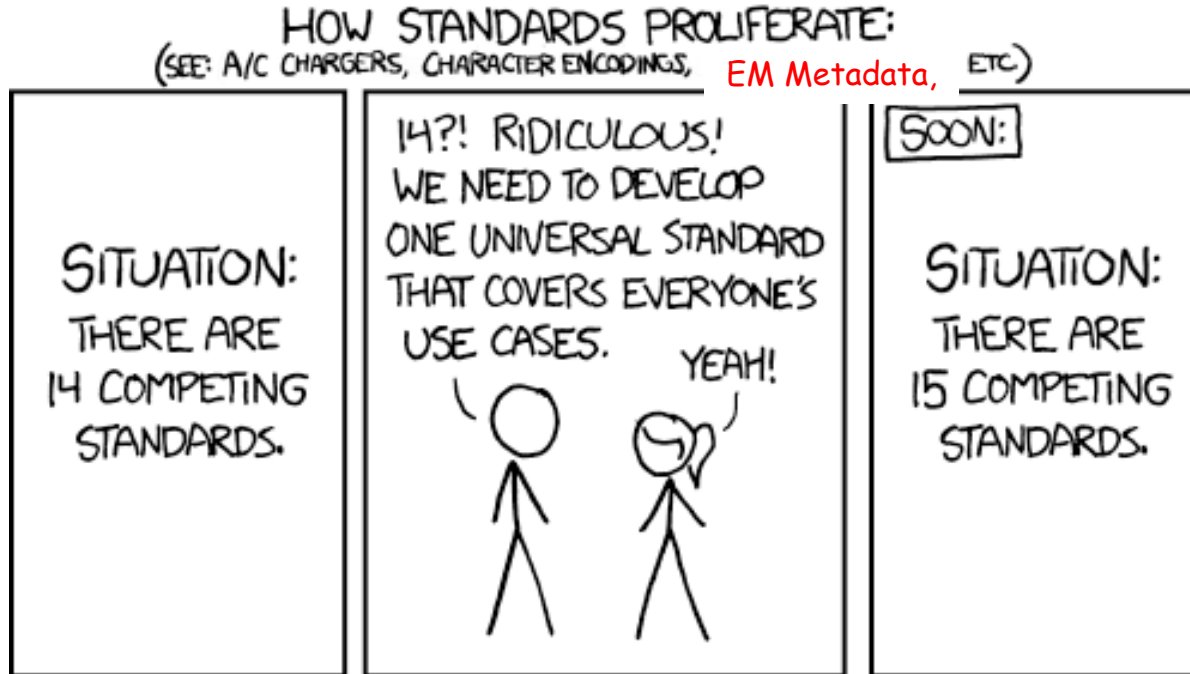
Day 2

Day 2		
8:30	Summary of working group	WBGB/019
9:15	Resume working groups	Working groups
10:45	Coffee	
11:00	Summary of working groups	Unified discussion
11:30	Conclusion	Unified discussion
12:00-13:30	Lunch & Coffee	OASE Cafeteria
14:00-15:30	SLS Tour (Optional)	

- Event page: <https://indico.psi.ch/e/em-standards-2024>
 - Speakers, please upload slides here
- Collaborative Document: <https://bit.ly/emstandards2024>
 - Everyone please add notes!
- Mailing list: data_management_swissplus_effort@cnb.csic.es
 - Managed by Oscar Saiz Gutierrez oscar.saiz@cnb.csic.es
- Contacts:
 - Spencer Bliven spencer.bliven@psi.ch +41 56 310 36 03
 - Viviana Sabatini viviana.sabatini@psi.ch +41 56 310 25 92
 - Jose Maria Carazo carazo@cnb.csic.es



Words of warning



Standardization levels



Physical file format and encoding
(JSON vs XML, TIFF vs MRC, LZW vs gzip,
JSON-LD vs RDF)



Data model (schema)



Conventions (ontologies)

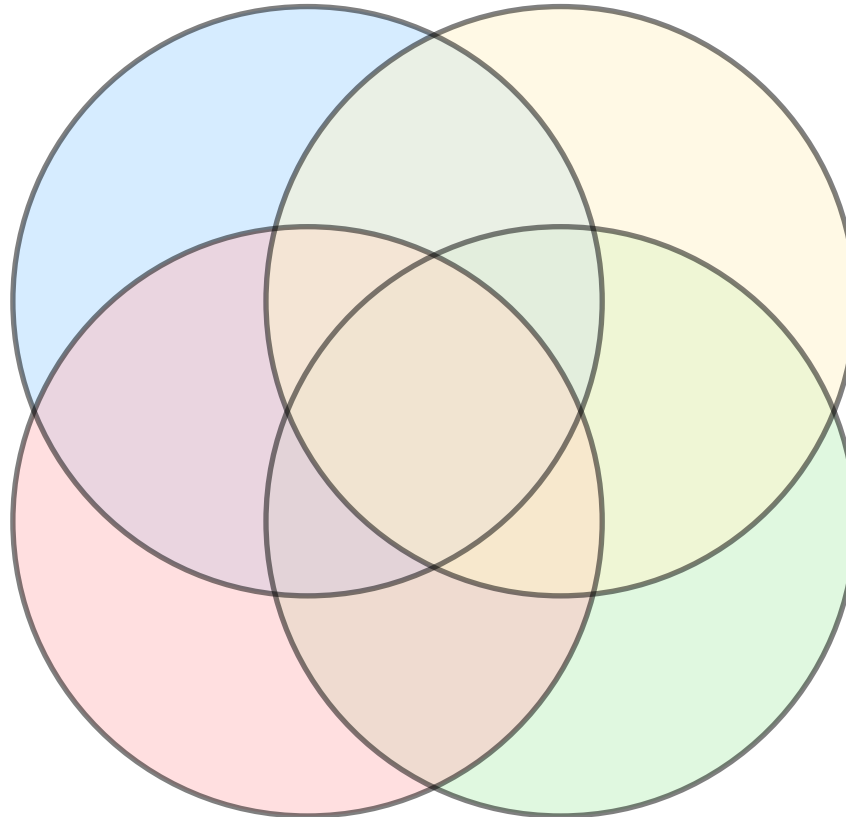
Diverse Stakeholders

Facilities

INSTRUCT
ESRF
Diamond
Swiss Facilities
...

Processing

Scipion
Relion
CryoSparc
EMinsight
Pipelinr
Aida
...



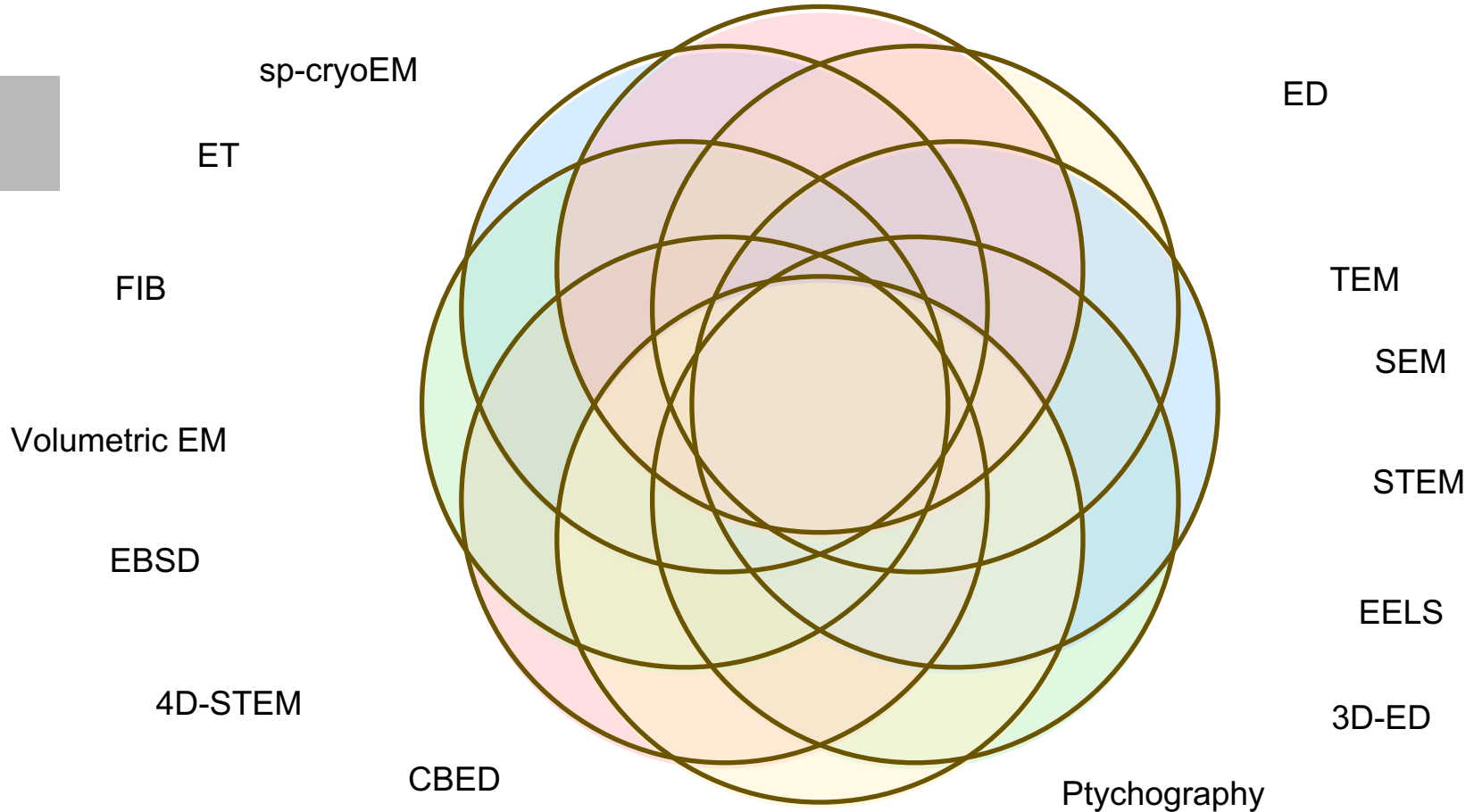
Repositories

EMDB
EMPIAR
PDB
Materials Cloud
SciCat

Data Consumers

Search engines
APIs
Journals

Diverse Experimental Techniques



Standards for diverse communities: NeXus



- Common data exchange format for neutron, X-ray, and muon experiments.
- Hierarchical Data Format (HDF5)
 - Combines data and metadata
- Base definitions provide the common data model for all NeXus files
- Domain-specific *application definitions* specify compulsory fields to enable specific use cases
 - 34 accepted
 - NXarchive (used by ICAT), NXmx, NXtomo
 - More than 100 community-contributed definitions
 - NXem, NXem_ebsd

Consider an application definition as a *contract* between a data provider (such as the beam line control system) and a data consumer (such as a data analysis program for a scientific technique) that describes the information is certain to be available in a data file. – [nexus manual](#)

Challenges

- Don't want to reinvent the wheel, but also need to support more use cases
- Diverse community with heterogeneous needs
- Diverse experimental techniques
- Unified data management workflows

Session 1: Motivation		
14:00	Welcome	Spencer Bliven
14:15	Swiss OpenEM Project	Yves Tittes
14:30	INSTRUCT	Jose Maria Carazo
14:45	Summary of goals and progress	Jose Maria Carazo

- Event page: <https://indico.psi.ch/e/em-standards-2024> →
- Collaborative Document: <https://bit.ly/emstandards2024> →
- Mailing list: data_management_swissplus_effort@cnb.csic.es
 - Managed by Oscar Saiz Gutierrez <oscar.saiz@cnb.csic.es>

