Assonant

WE3BCO06 - A beamline-agnostic event processing engine for data collection and standardization

Paulo Baraldi Mausbach, Eduardo X. Miqueles, and Allan Pinto

ICALEPCS CAPE TOWN SOUTH AFRICA 2023

LNLS Laboratório Nacional de Luz Síncrotron

CNPEM Centro Nacional de Pesquisa em Energia e Materiais

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO

GOVERNO FEDERAL BRASIL UNIÃO E RECONSTRUÇÃO

About me

- Paulo Baraldi Mausbach – 27 years
- (2015 - 2019): Degree in Computer Engineering by Pontifical Catholic University of Campinas (PUC-Campinas)
- (2020 – in progress): Pursuing a master degree in Computer Science at Institute of Computing (IC) from University of Campinas (Unicamp)
- Working with Data Science and Machine Learning since 2020
- Intern at Sirius in 2019 -> Sirius Control group and Scientific Computing group (GCC)
- Back to Sirius in July 2022 at GCC and currently part of the recently created (~8 months old) Data Science and Management group (GCD)

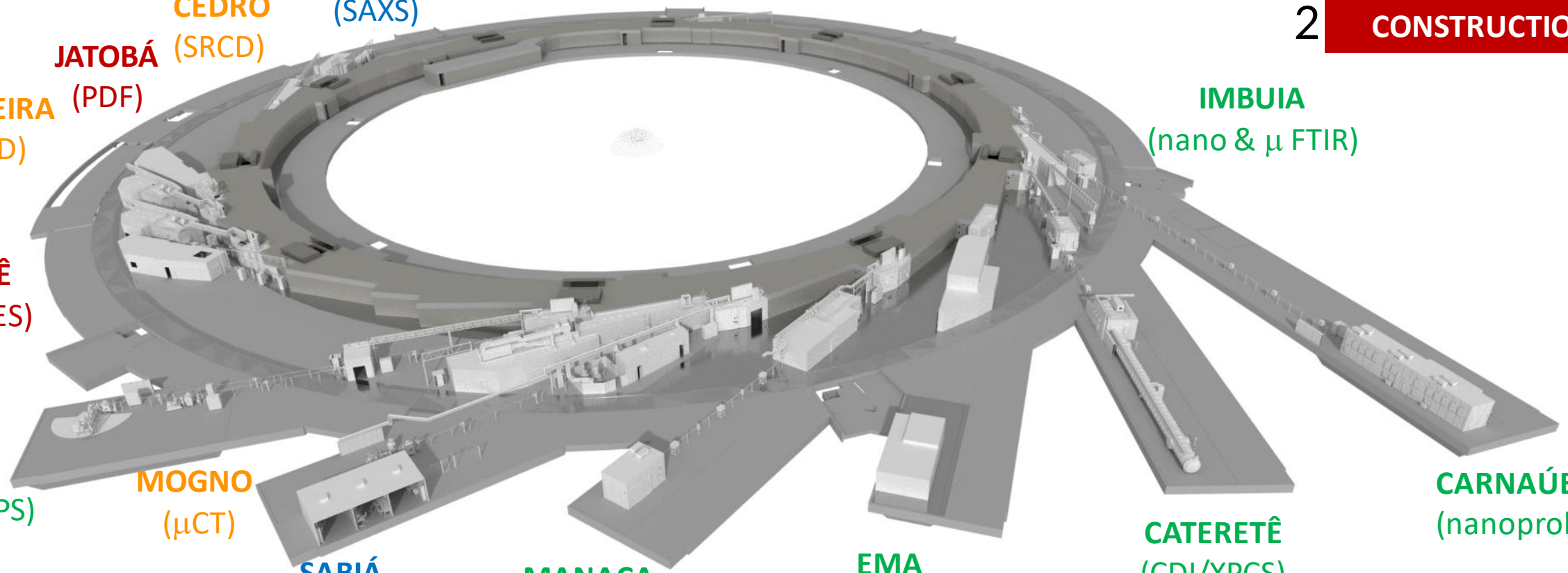
Schedule

- About Sirius
- Current workflow and its problems
- Assonant
- Experiments and results
- Future work

About Sirius – Stage 1

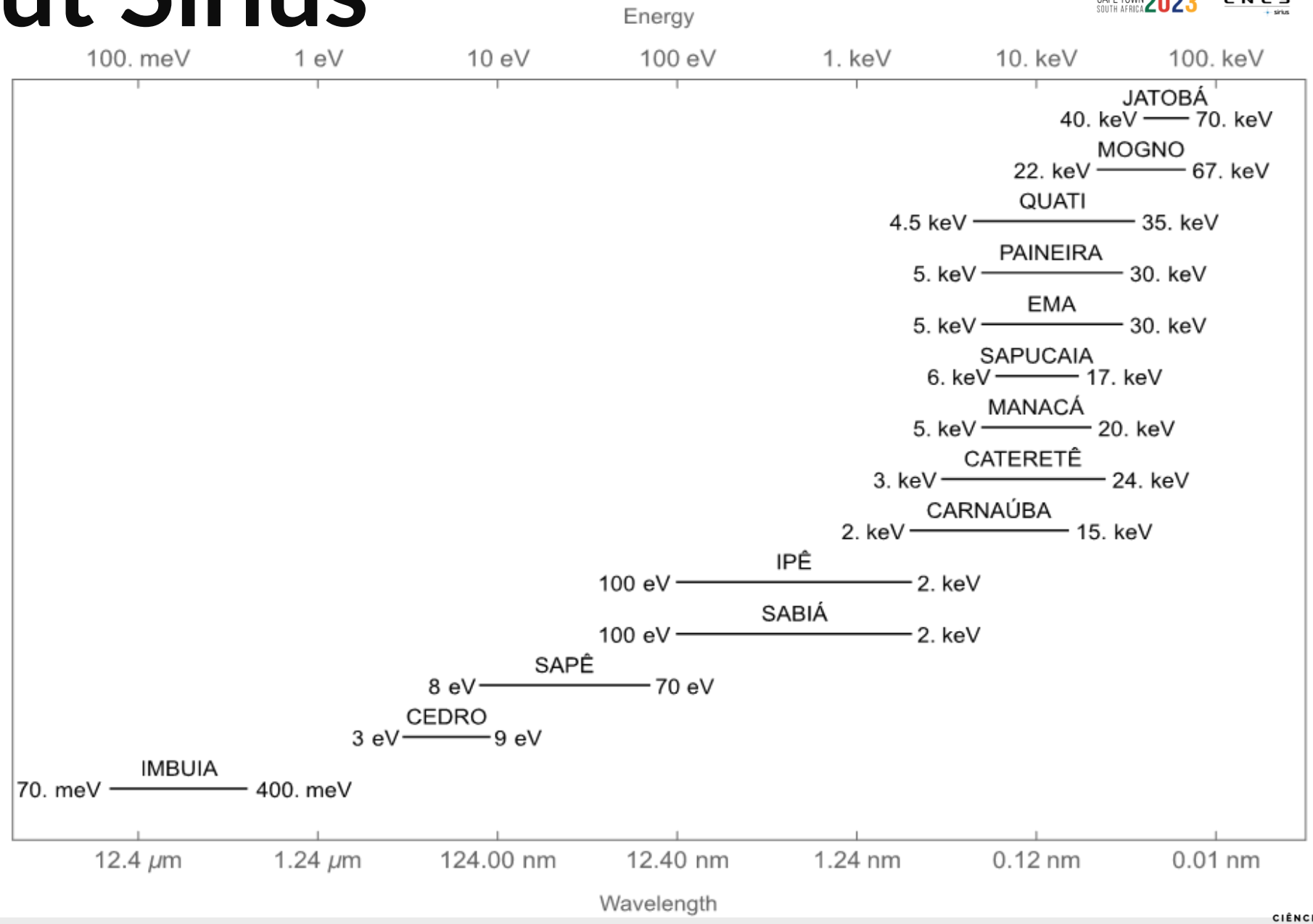
- 4th-generation Brazilian synchrotron light source

6	OPERATION
3	COMMISSIONING
3	INSTALLATION
2	CONSTRUCTION



Courtesy from: Harry Westfahl Junior

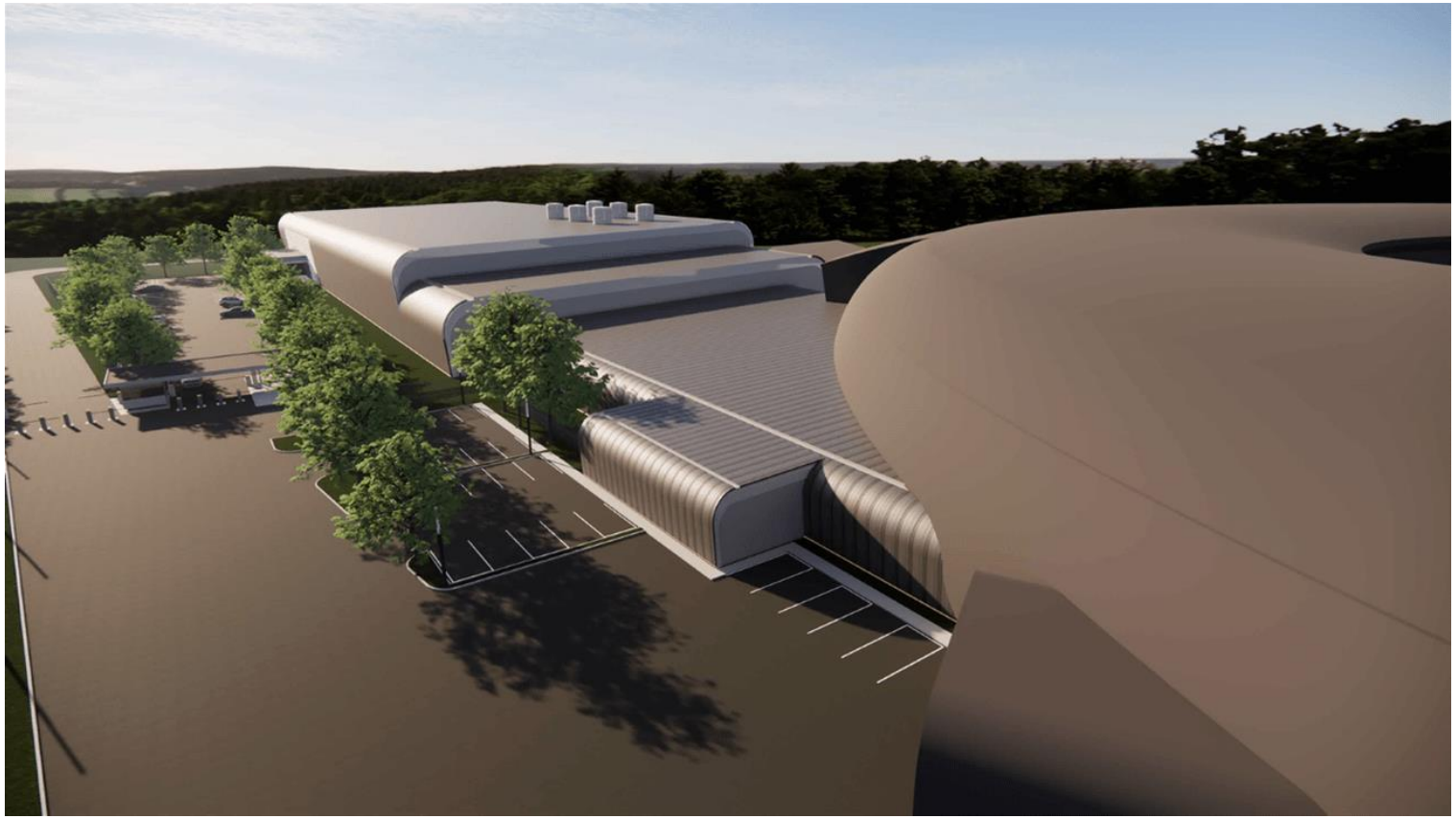
About Sirius



Courtesy from: Harry Westfahl Junior

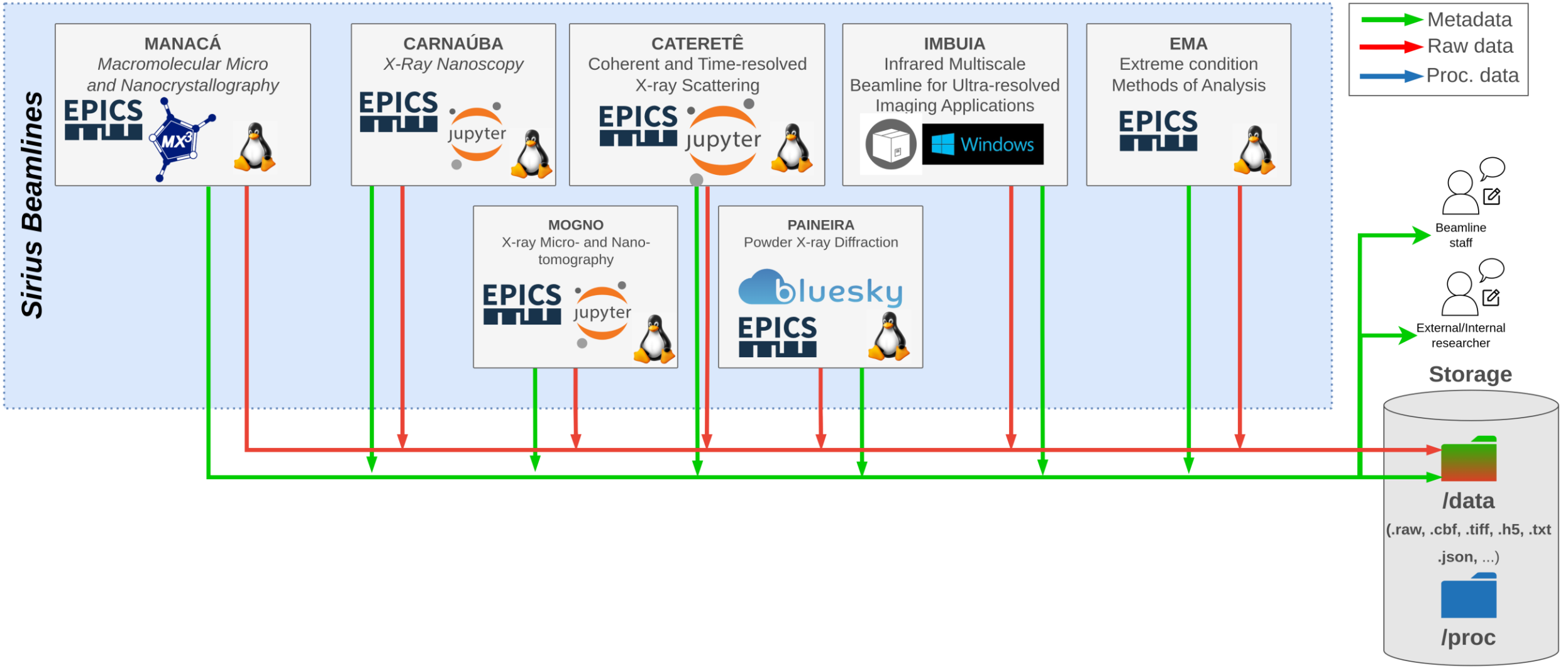
About Sirius – Stage 2 + Orion coming

- Construction of +10 beamlines and
- Orion, **a First of its kind**, Maximum Biosafety Lab (BSL-3 & BSL-4) with 3 beamlines integrated on it

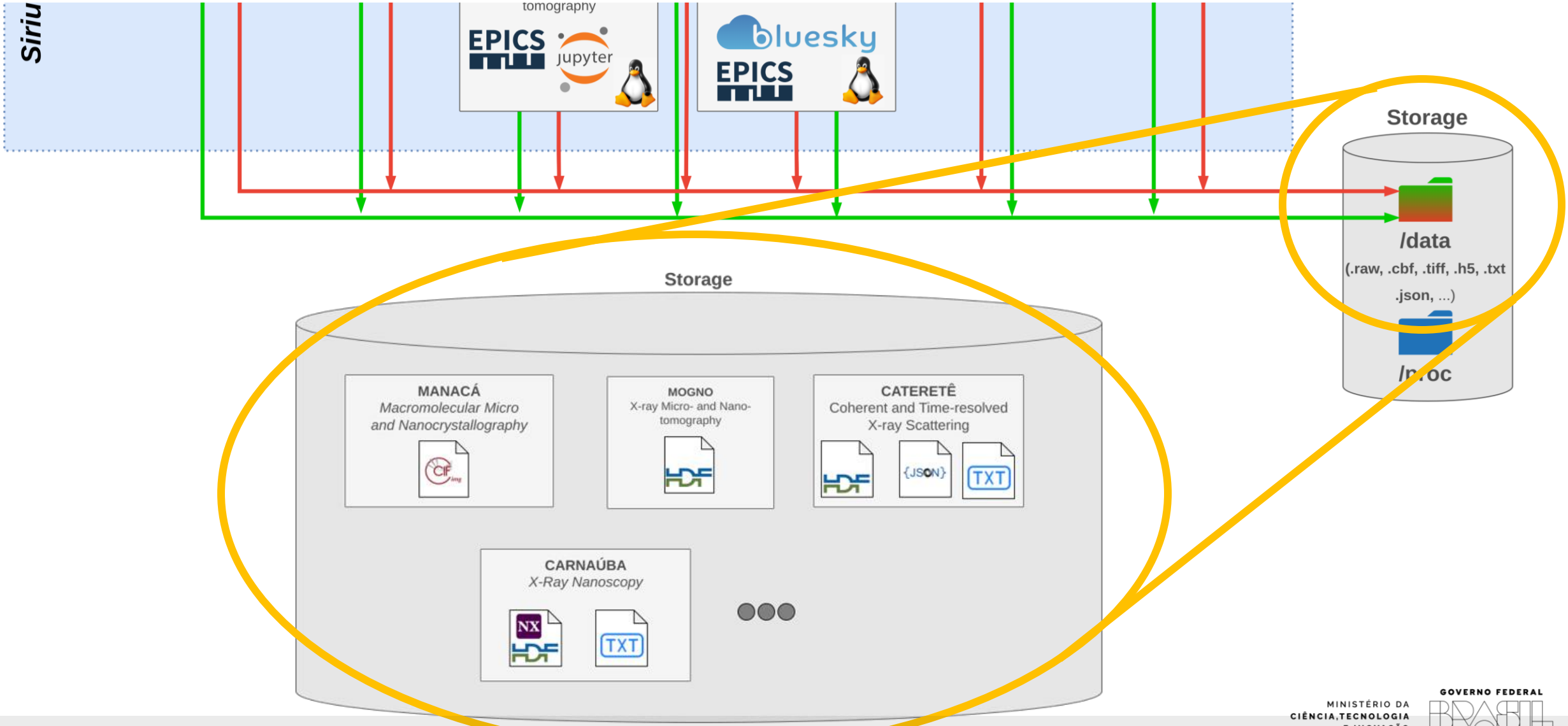


https://cnpem.br/category/clipping-cnpem/projeto-orion-na-midia/
https://earthsky.org/astronomy-essentials/star-sirius-torchbearer-of-the-new-year/

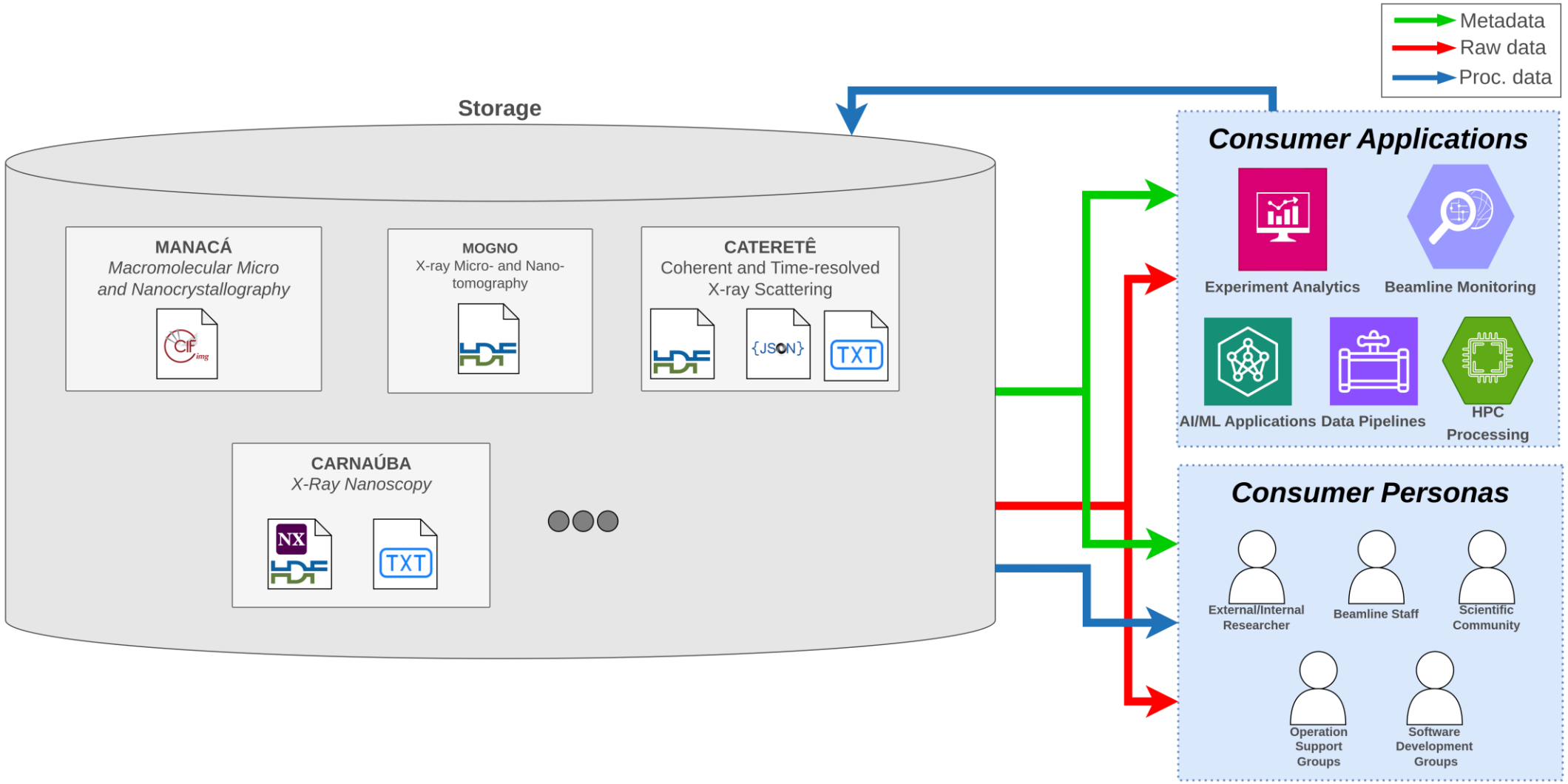
Current dataflow and its problems



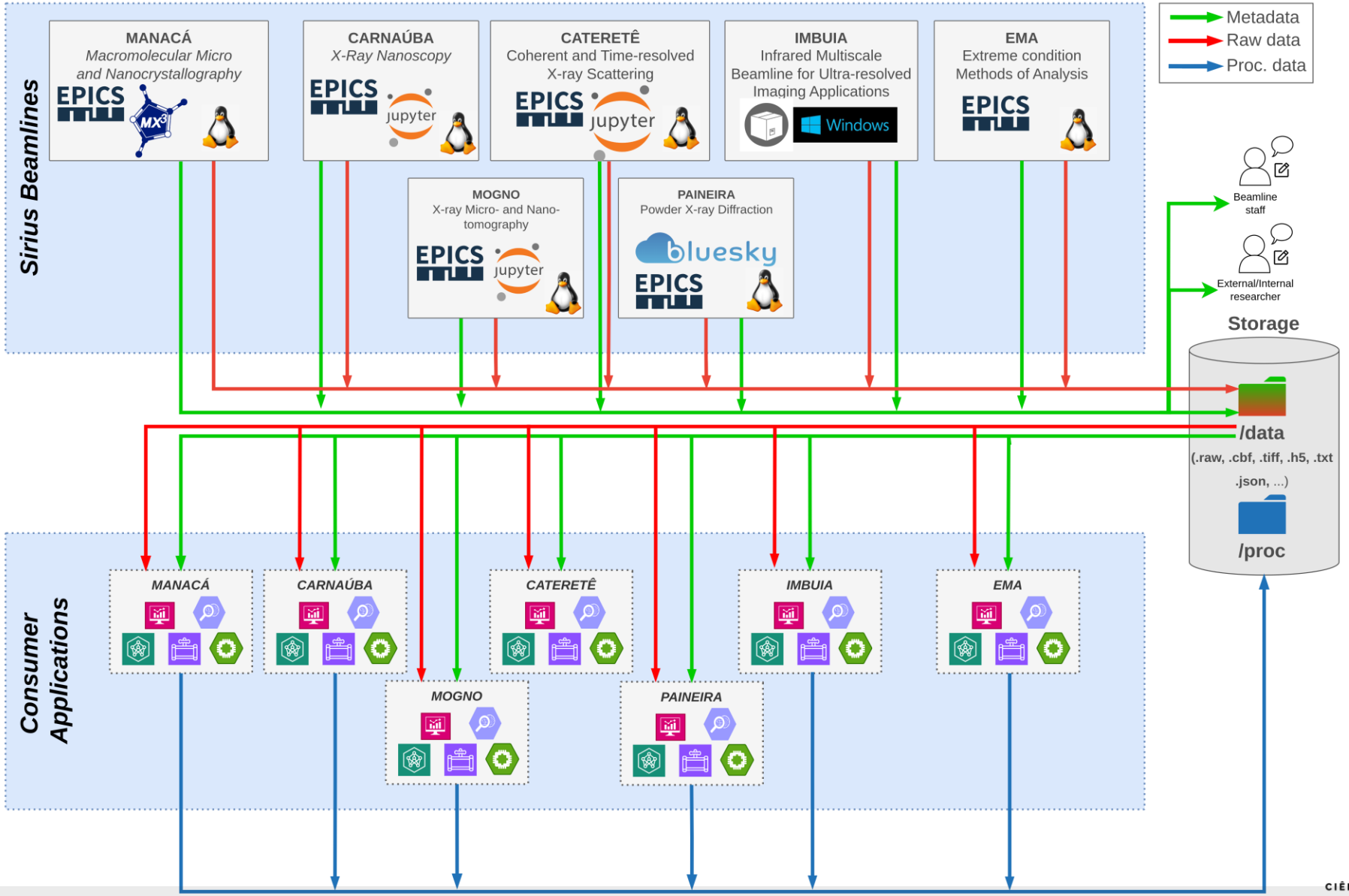
Current dataflow and its problems



Current dataflow and its problems



Current dataflow and its problems

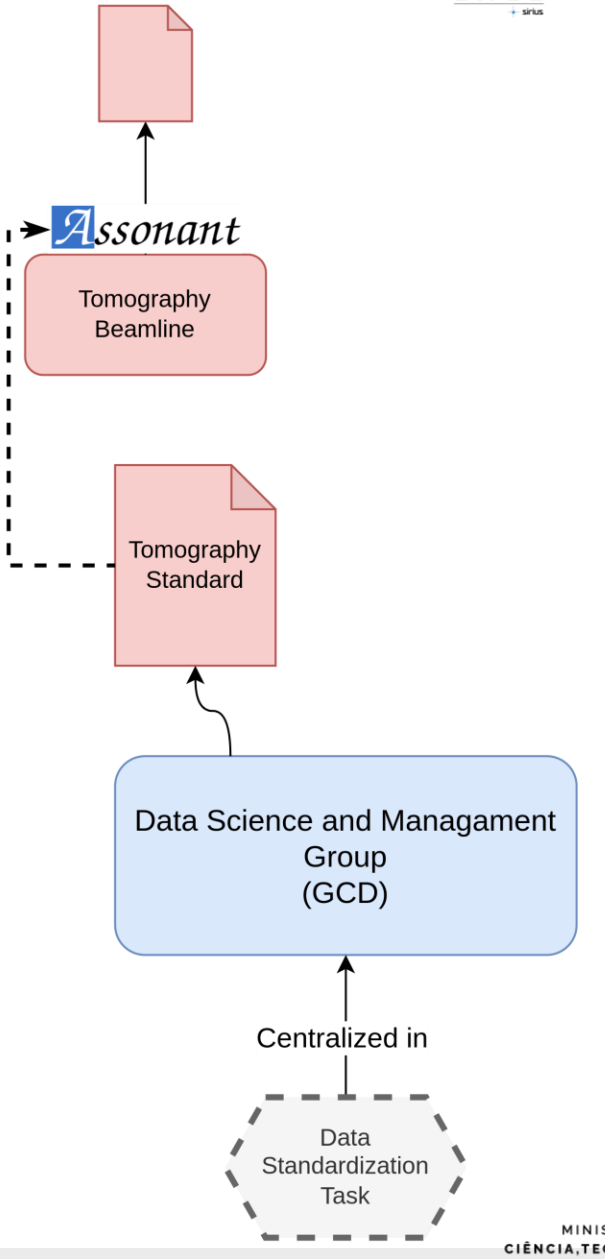



Assonant

Assonant – Design premises

1. Data Standardization

- Centralize it as a Data Science and Management task
- Abstract it from beamlines
- Technique-centric



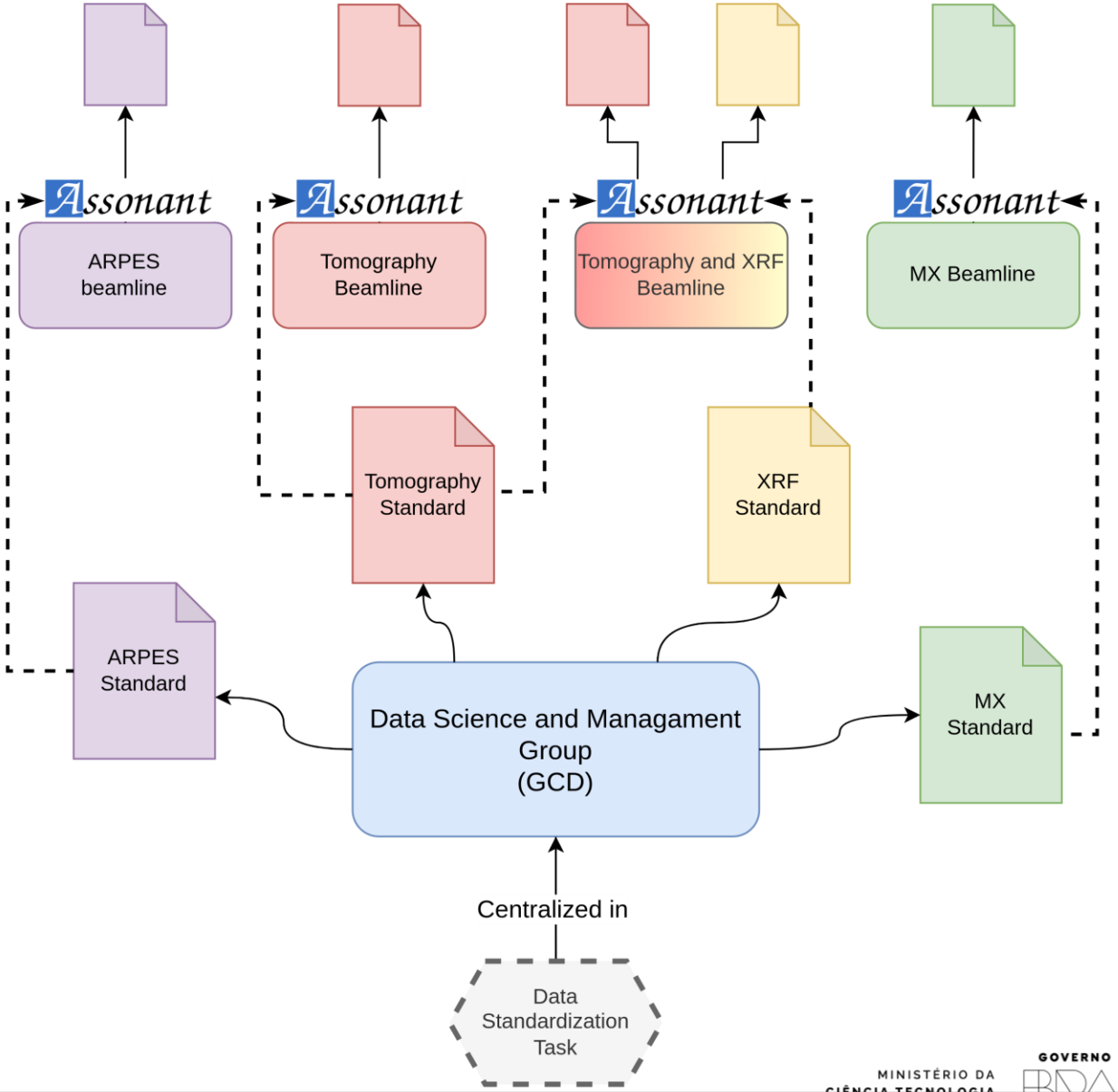
Assonant – Design premises

1. Data Standardization

- Centralize it as a Data Science and Management task
- Abstract it from beamlines
- Technique-centric

2. Beamline Agnostic

- Usable at any beamline
- Low invasive:
  - Impact the less possible the control logic from beamline



Assonant – Design premises

1. Data Standardization

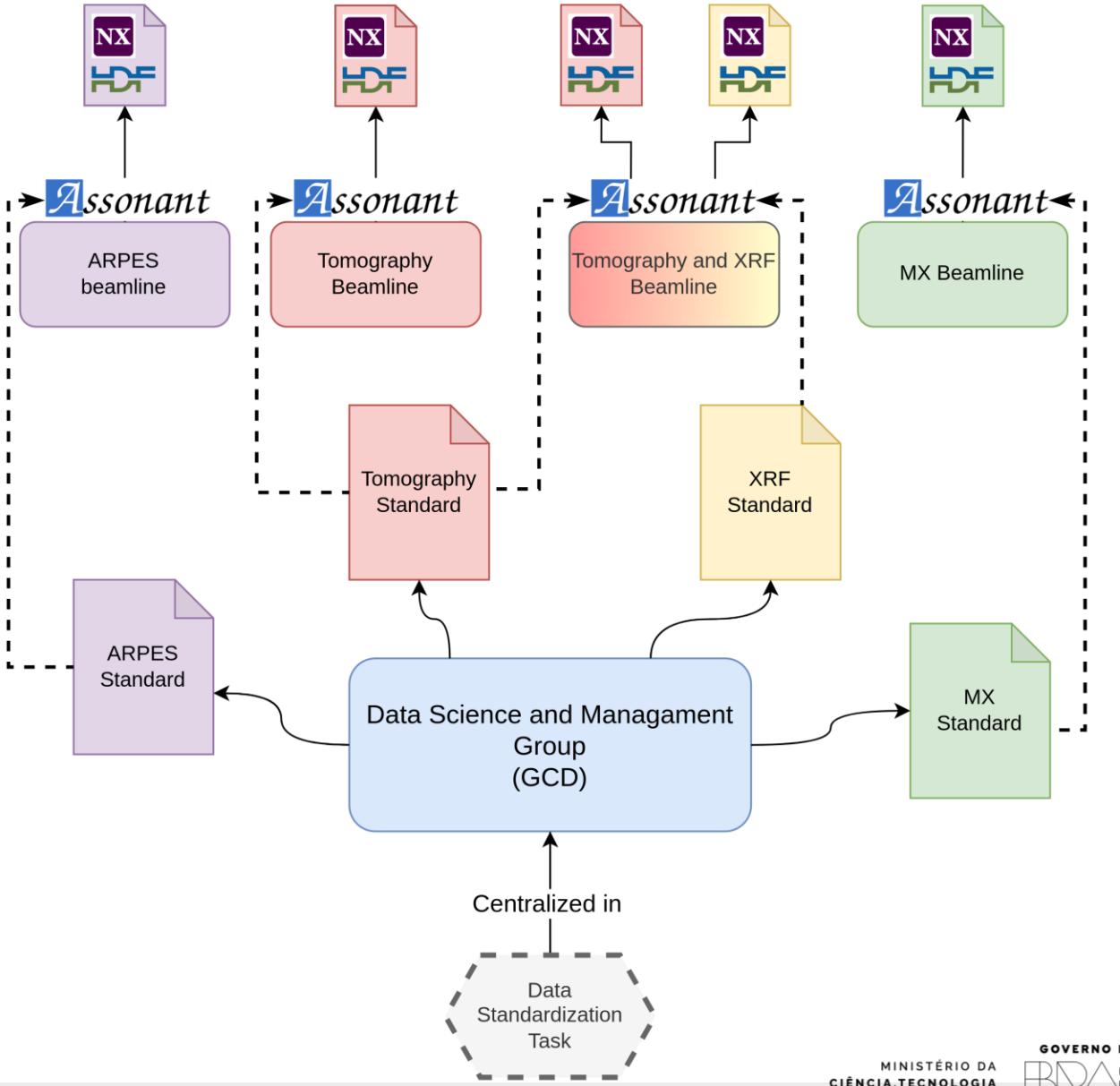
- Centralize it as a Data Science and Management task
- Abstract it from beamlines
- Technique-centric

2. Beamline Agnostic

- Usable at any beamline
- Low invasive:
  - Impact the less possible the control logic from beamline

3. Data Format

- Handle the heterogeneous nature of synchrotron data



Assonant – Design premises

1. Data Standardization

- Centralize it as a Data Science and Management task
- Abstract it from beamlines
- Technique-centric

2. Beamline Agnostic

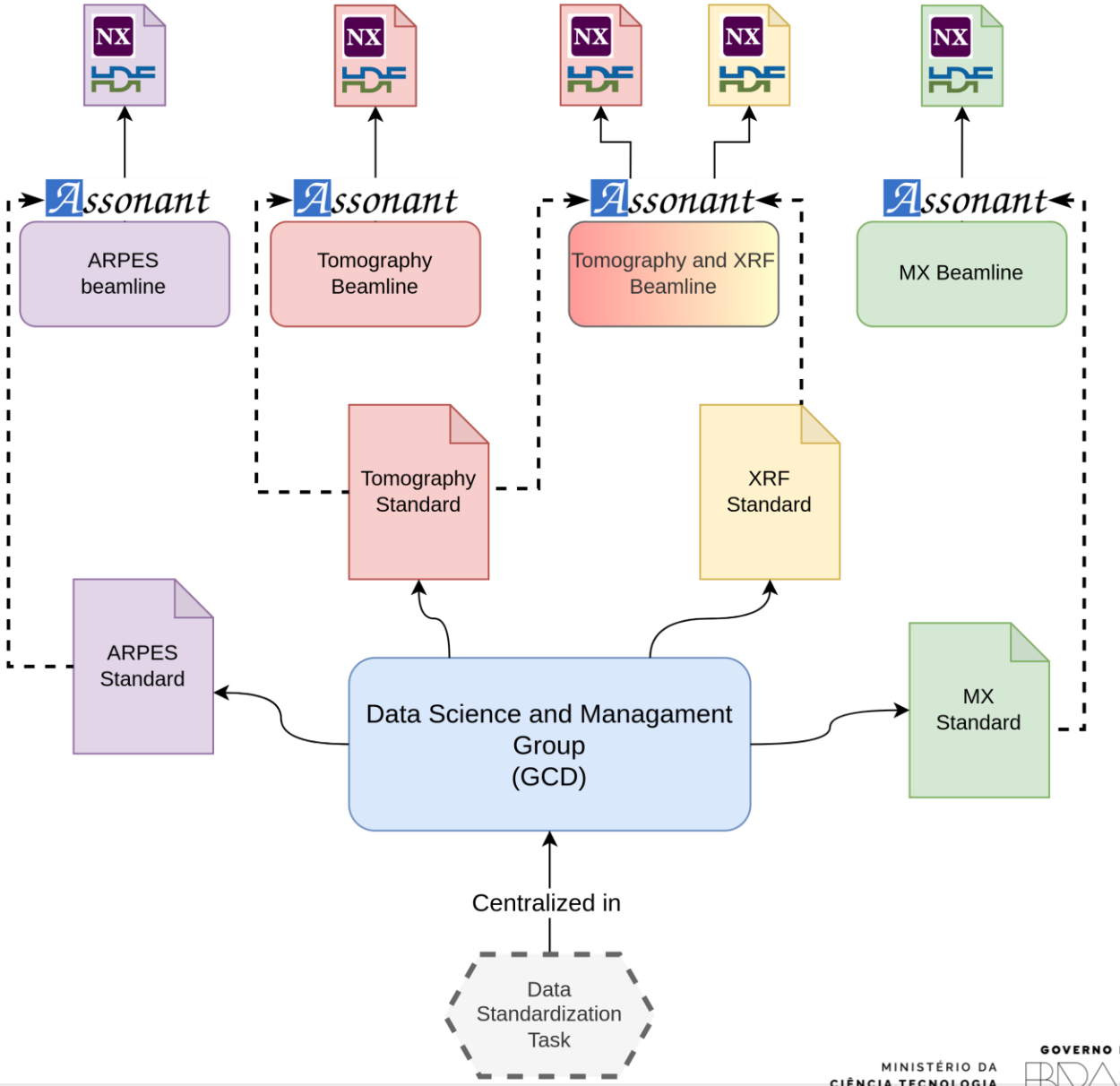
- Usable at any beamline
- Low invasive:
  - Impact the less possible the control logic from beamline

3. Data Format

- Handle the heterogeneous nature of synchrotron data

4. Development

- Extensible modules developed in Python



Assonant – Modules

1. Data Classes

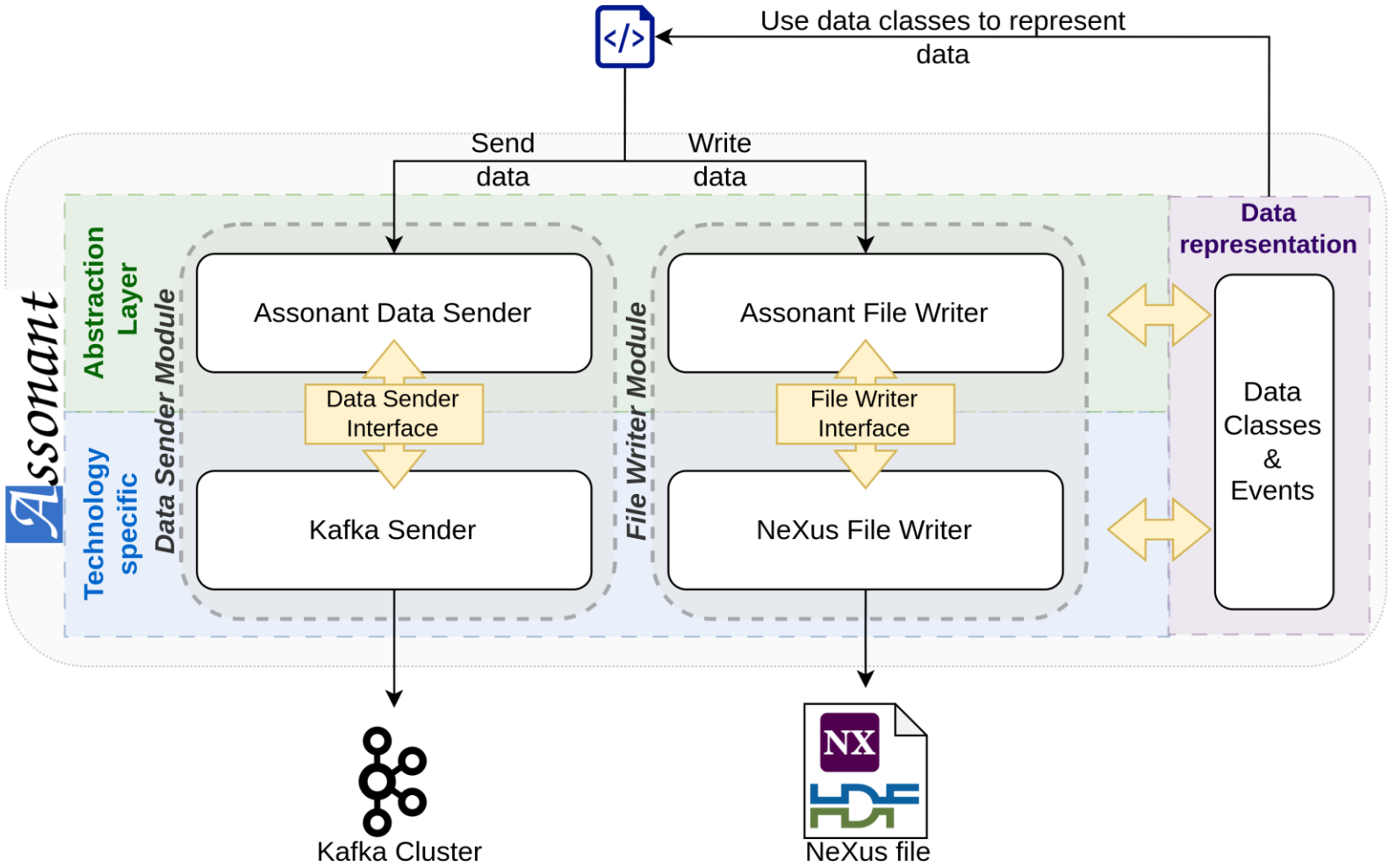
- Allow representing data as the defined data model
- Standardize how data is exchanged between modules

2. File Writer

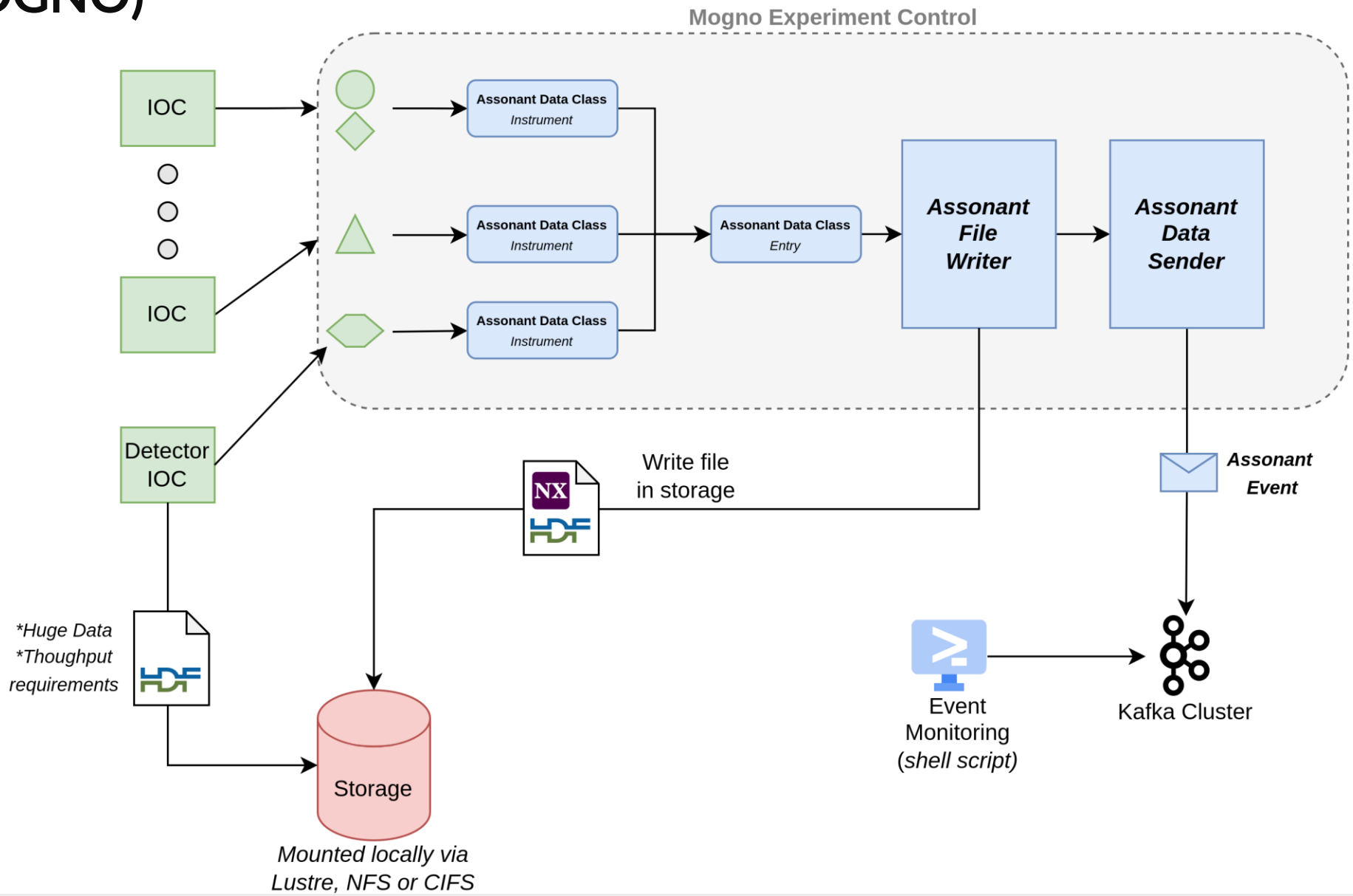
- Write data/metadata contained in Data Classes into a file

3. Data Sender

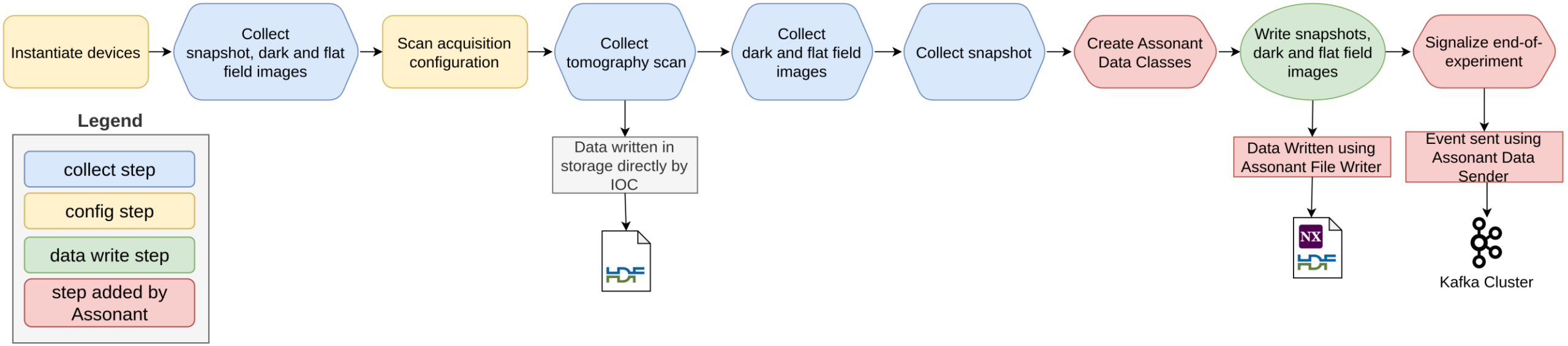
- Send data/metadata contained in Data Classes through a communication interface
  - Event Queue, REST API, ...



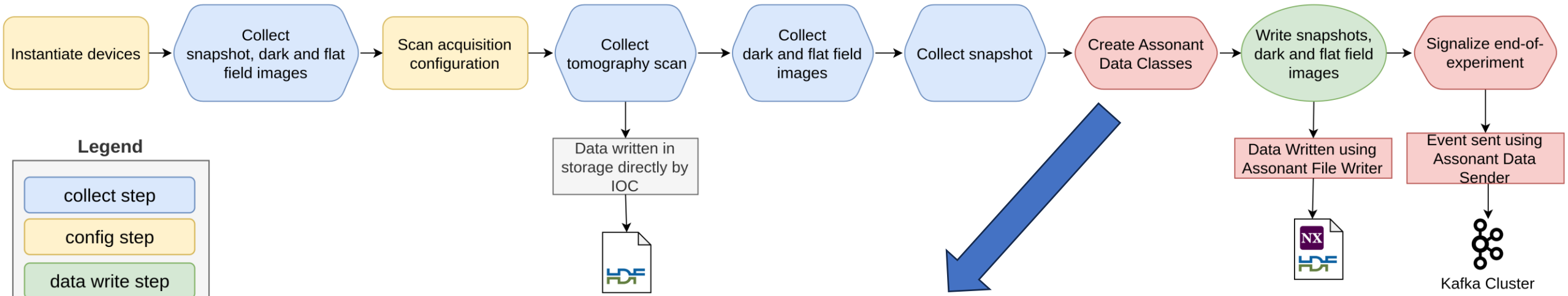
Experiment – Tomography Beamline (MOGNO)



Experiment – Tomography Beamline (MOGNO)



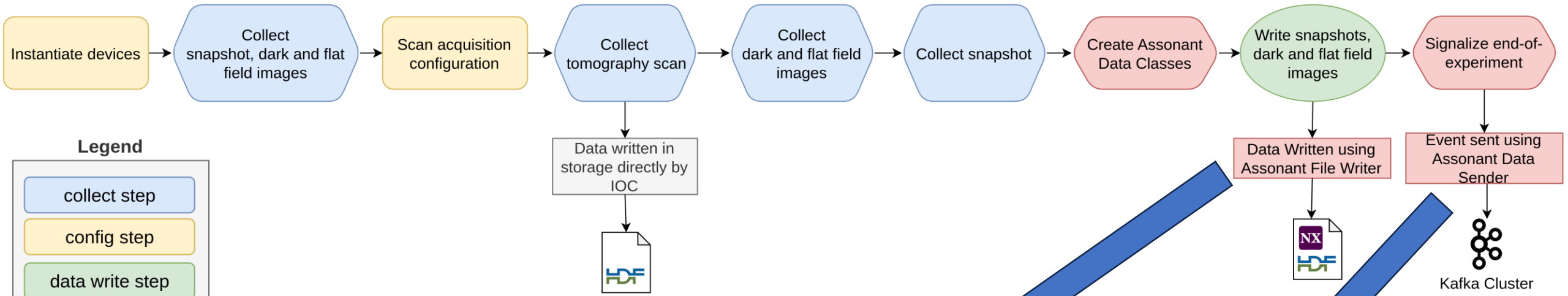
Experiment – Tomography Beamline (MOGNO)




```
sample_image = np.zeros((256,256))

mono_timeseries = Monochromator(name="mono_timeseries",
                        wavelength=TimeSeries(
                            value=DataField(
                                value=[sample_image,sample_image,sample_image,sample_image],
                                unit="db"
                            ),
                            timestamps=DataField(value=[1,2,3,4])
                        ),
                        energy=TimeSeries(
                            value=DataField(
                                value=[1,2,3,4],
                                unit="keV",
                            ),
                            timestamps=DataField(
                                value=[10,20,30,40],
                                extra_metadata={
                                    "attr1": 10,
                                    "attr2": "atributo 2",
                                    "attr3": 0.765,
                                    "attr4": [1,2,3,4,5]
                                }
                            ),
```

Experiment – Tomography Beamline (MOGNO)




```
file_writer = AssonantFileWriter('nexus')
file_writer.write_data(filepath, filename, entry)
```

```
producer = AssonantDataSender(comm_method="kafka", configs=configs)
producer.send_data(event)
```

Experiment – Tomography Beamline (MOGNO)

Before Assonant

Name	Description	Type	Shape
assonant_mgn_000.hdf5			
after			
before			
scan			
beamline			
detector			
data	3D data	uint16	100 × 2048 ×
metadata			
ad_default			
motor			
Rotional_stage_gather	1D data	float64	100
snapshot			
after			
beamline-state			
beam-optics			
position			
detector			
device			
host	"10.39.5…	string	scalar
precision	0	float32	scalar
pvname	"MGN:C…	string	scalar
timestamp	1.69211…	float64	scalar
units	0	float32	scalar
value	0	int64	scalar
fov-x			
fov-y			
microscope-objective-lens			
pixel-size-x			
pixel-size-y			
kb-measured-focus-z			
nano-station			
sample-change			
motion			
sensor			

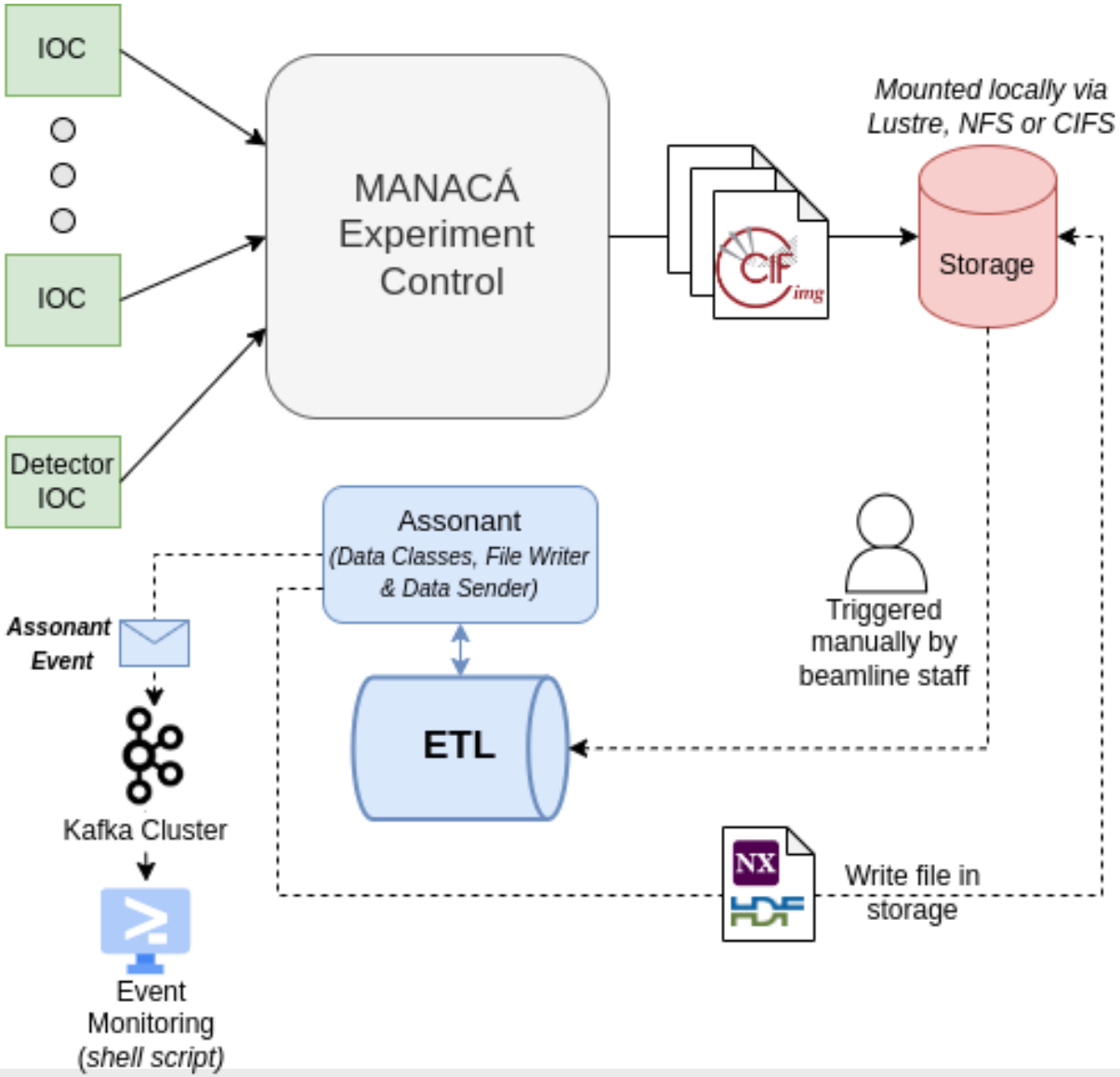
mogno_tomography.h5

After Assonant

Name	Description	Type	Shape
assonant_mgn_000_assonant.nxs			
during_exposure		NXentry	
instrument		NXinstrument	
detector		NXdetector	
data	"/ibira/l..	string	scalar
transformations		NXtransformations	
long-stroke-x	2.638	float64	scalar
long-stroke-y	-0.660002	float64	scalar
long-stroke-z	2469.99	float64	scalar
sample	"my_sa…	NXsample	
post_exposure		NXentry	
instrument		NXinstrument	
dvf1		NXdetector	
extra_info		NXcollection	
breaker	0	int64	scalar
transformations		NXtransformations	
slot-position	-25.8	float64	scalar
dvf2		NXdetector	
foe		NXslit	
transformations		NXtransformations	
foe_slit_1_ux_actuator	1.32086	float64	scalar
foe_slit_1_uy_actuator	1.01772	float64	scalar
foe_slit_2_ux_actuator	0.37378	float64	scalar
foe_slit_2_uy_actuator	0.12758	float64	scalar
pink		NXslit	
transformations		NXtransformations	
pbs1_bottom_actuator	0.57995	float64	scalar
pbs1_left_actuator	0.17	float64	scalar
pbs1_right_actuator	0.113	float64	scalar
pbs1_top_actuator	0.1297	float64	scalar
pre_exposure		NXentry	
instrument		NXinstrument	

mogno_tomography.nxs

Experiment - MX beamline (MANACÁ)



Experiment - MX beamline (MANACÁ)

After Assonant

Name	Description	Type	Shape	Link
manaca_nexus.nxs				
during_exposure		NXentry		
instrument		NXinstrument		
monochromator		NXmonochromator		
wavelength	0.97721	float64	scalar	
PILATUS 2M		NXdetector		
data	1D data	string	3600	
extra_info		NXcollection		
sample	"2972-1…	NXsample		
name	"2972-1…	string	scalar	
transformations		NXtransformations		
Rx	1D data	float64	3600	

manaca_nexus.nxs

Experiment – Final files comparison

Name	Description	Type	Shape	Link
manaca_nexus.nxs				
during_exposure		NXentry		
instrument		NXinstrument		
monochromator		NXmonochromator		
wavelength	Ⓥ 0.97721	float64	scalar	
PILATUS 2M		NXdetector		
data	Ⓥ 1D data	string	3600	
extra_info		NXcollection		
sample	Ⓝ "2972-1…	NXsample		
name	Ⓥ "2972-1…	string	scalar	
transformations		NXtransformations		
Rx	Ⓥ 1D data	float64	3600	

manaca_nexus.nxs

Name	Description	Type	Shape
assonant_mgn_000_assonant.nxs			
during_exposure		NXentry	
instrument		NXinstrument	
detector		NXdetector	
data	Ⓥ "/ibira/l…	string	scalar
transformations		NXtransformations	
long-stroke-x	Ⓥ 2.638	float64	scalar
long-stroke-y	Ⓥ -0.660002	float64	scalar
long-stroke-z	Ⓥ 2469.99	float64	scalar
sample	Ⓝ "my_sa…	NXsample	
post_exposure		NXentry	
instrument		NXinstrument	
dvf1		NXdetector	
extra_info		NXcollection	
breaker	Ⓥ 0	int64	scalar
transformations		NXtransformations	
slot-position	Ⓥ -25.8	float64	scalar
dvf2		NXdetector	
foe		NXslit	
transformations		NXtransformations	
foe_slit_1_ux_actuator	Ⓥ 1.32086	float64	scalar
foe_slit_1_uy_actuator	Ⓥ 1.01772	float64	scalar
foe_slit_2_ux_actuator	Ⓥ 0.37378	float64	scalar
foe_slit_2_uy_actuator	Ⓥ 0.12758	float64	scalar
pink		NXslit	
transformations		NXtransformations	
pbs1_bottom_actuator	Ⓥ 0.57995	float64	scalar
pbs1_left_actuator	Ⓥ 0.17	float64	scalar
pbs1_right_actuator	Ⓥ 0.113	float64	scalar
pbs1_top_actuator	Ⓥ 0.1297	float64	scalar
pre_exposure		NXentry	
instrument		NXinstrument	

mogno_tomography.nxs

Future Work

1. Enhacements

- NeXus compliance
- Application definition

2. Expansion

- Cover more:
  - Devices
  - Beamlines
  - Experimental techniques (SAXS, XRF, ARPES, ...)

3. Data Services

- Data Catalog
- Data Enrichment or reduction
- Custom Data views
- Post-processing automatization
- And more...

4. Collaboration

- I am open to talk about possible collaborations, similar projects, previous experiences, ...

Thanks for your attention!

Paulo Baraldi Mausbach
paulo.mausbach@lnls.br

Check our paper: WE3BCO06

cnpem.br

ICALEPCS CAPE TOWN SOUTH AFRICA 2023

LNLS + sirius

CNPEM Centro Nacional de Pesquisa em Energia e Materiais

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO

GOVERNO FEDERAL BRASIL UNIÃO E RECONSTRUÇÃO