Enhanced and flexible software tools for X-ray CT at the SYRMEP beamline of Elettra

Francesco Brun

Elettra - Sincrotrone Trieste S.C.p.A, Basovizza, Trieste, Italy

University of Trieste, Trieste, Italy

Elettra Sincrotrone Trieste

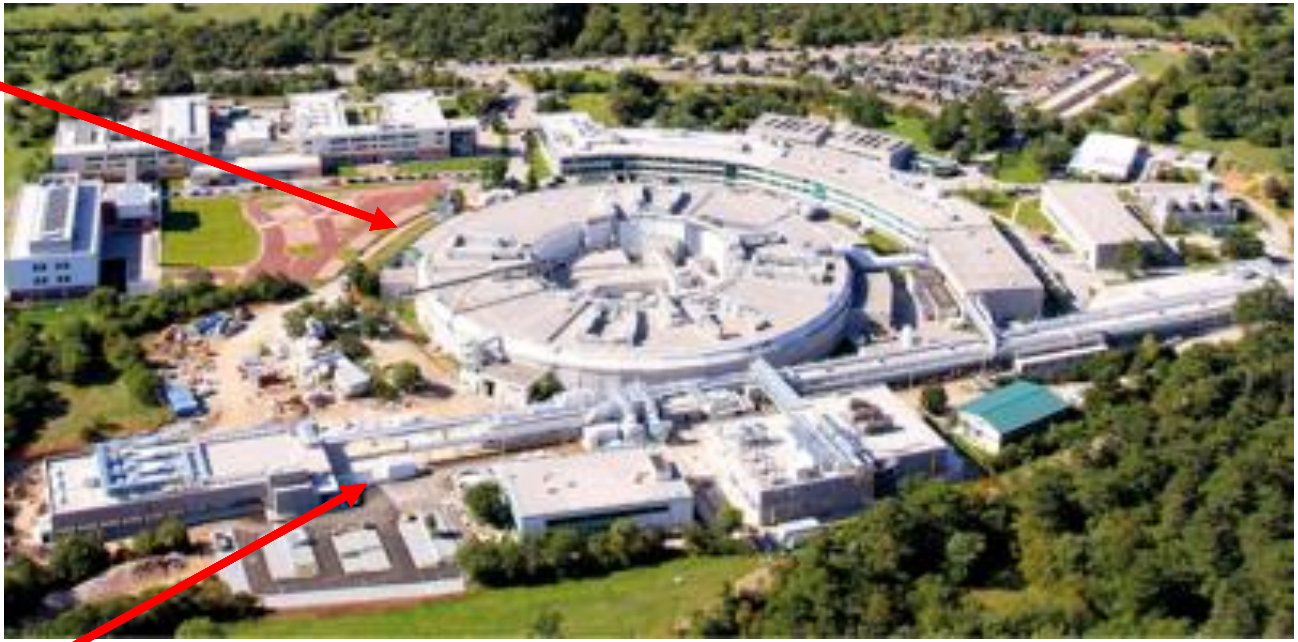
Szeged – January, 27th 2016

Elettra – Sincrotrone Trieste S.C.p.A.

Elettra is a non-profit company established in 1986

Two laboratories: **Elettra** and **Fermi**

3rd generation 2.4 GeV synchrotron radiation laboratory (**Elettra**)



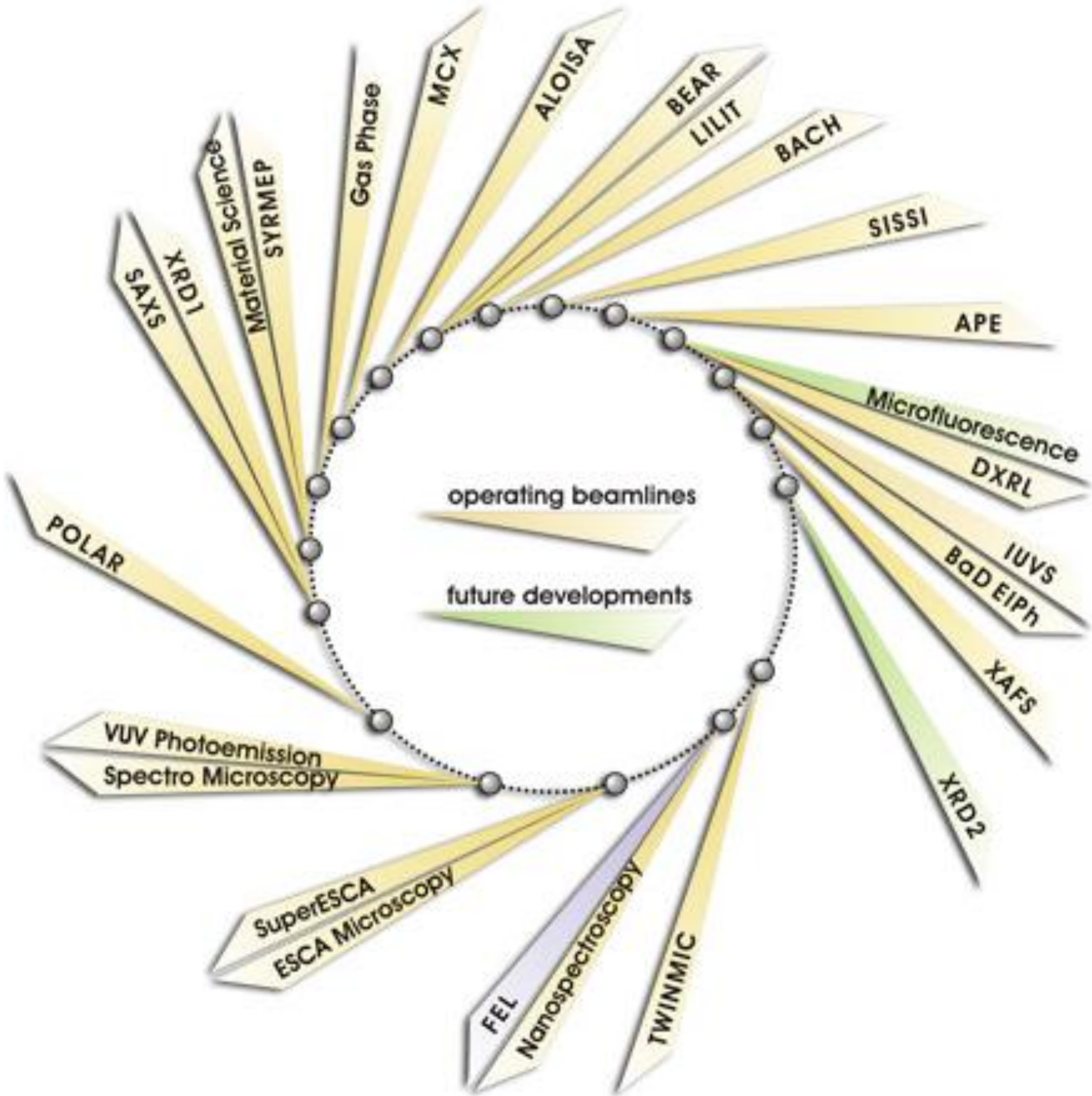
4th generation light source free-electron laser (FEL): **Fermi**

The Elettra laboratory

Several beamlines with different techniques

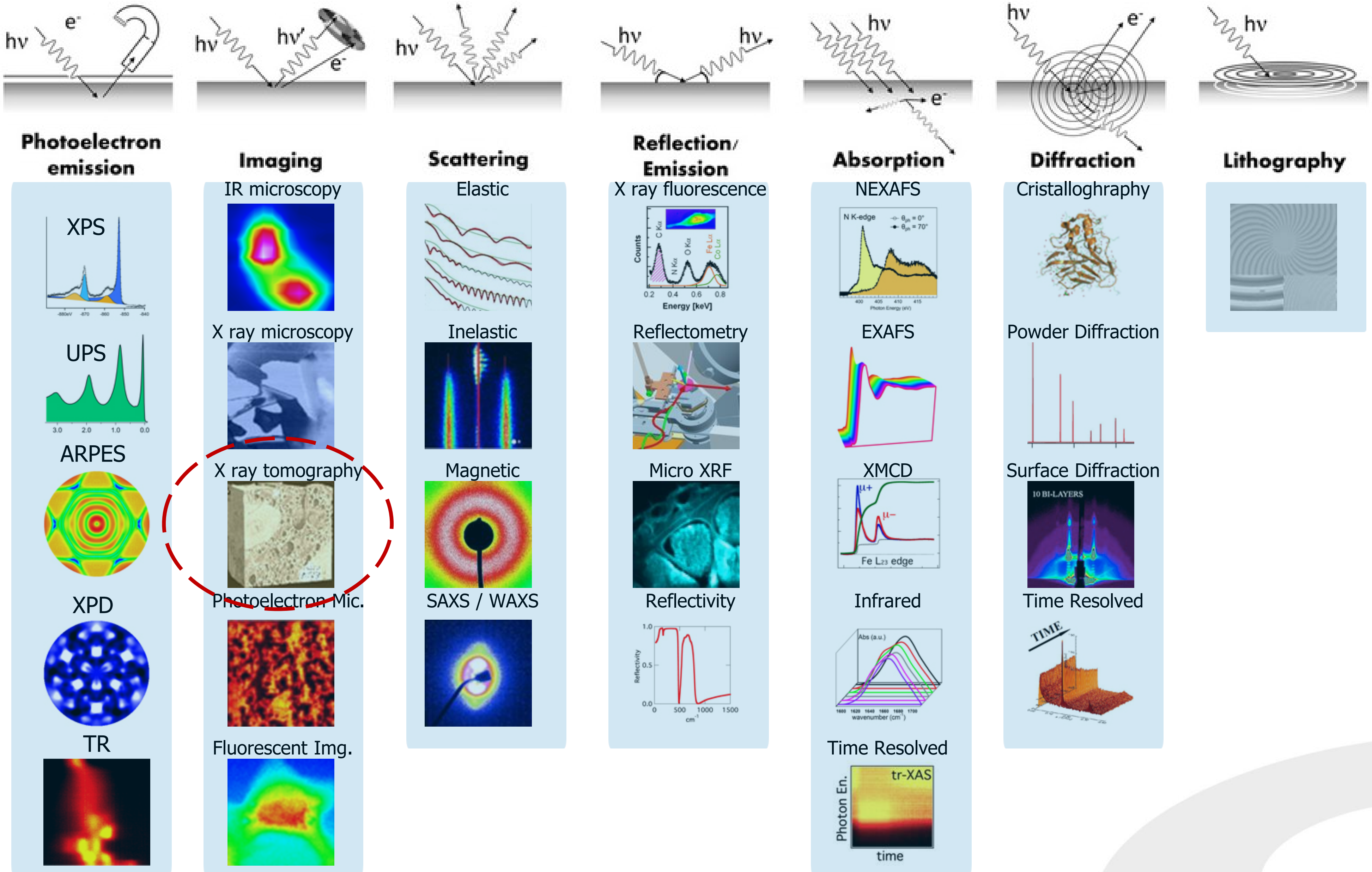
- diffraction
- scattering
- photoemission
- …

Several support laboratories



Elettra
Sincrotrone
Trieste

The techniques



SYRMEP

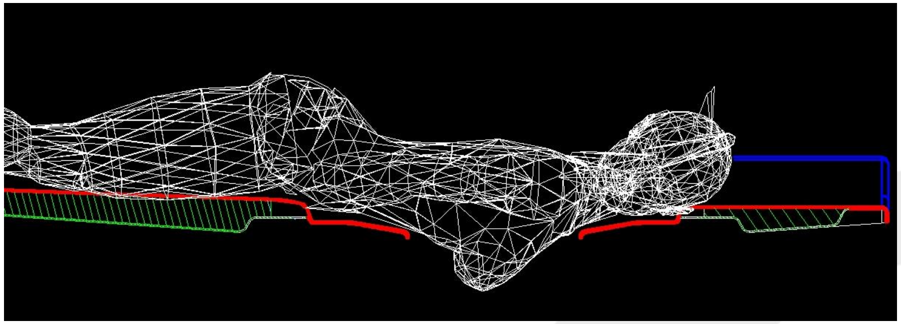
SYnchrotron **R**adiation for **ME**dical **P**hysics (SYRMEP) beamline[1]

Designed originally for a specific project

The aim is to show the possibility to have tomographic images (3D volume) with the **same radiation dose** of clinical mammography (2D projection)

Formal authorization to scan live patients

Other activities are performed



[1] http://www.elettra.eu/elettra-beamlines/syrmep.html

SYRMEP and its users

In addition to *in house* research, SYRMEP is opened to users

A significant portion of the proposals is in the **biomedical field**

Users want to exploit **phase contrast µ-CT**

↓

Most of SYRMEP users are e.g. biologists, medical doctors, neuroscientists, ...

In general, “bio-scientists” don’t have strong computer skills

Radiation dose is a concern in most of the experiments

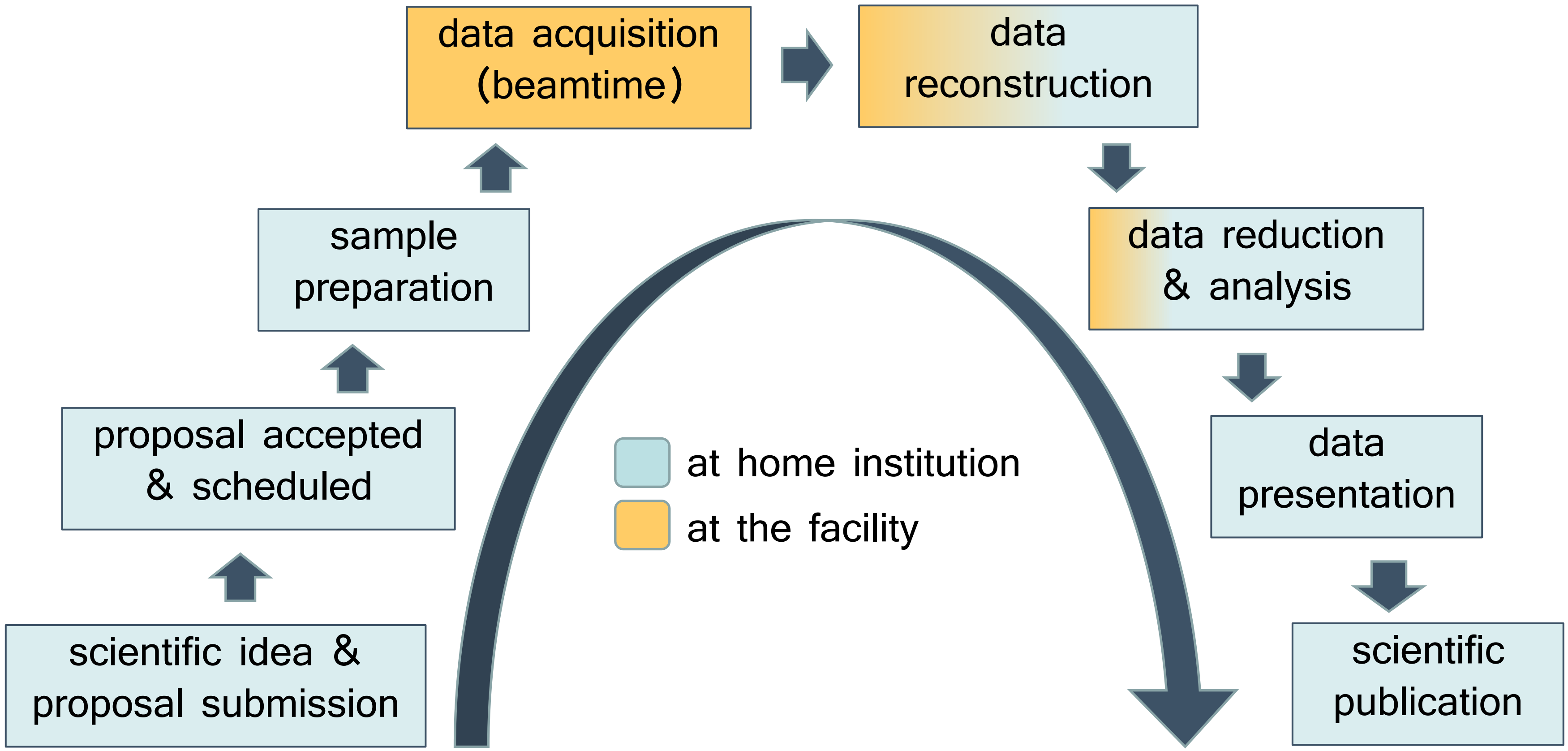
↓

“Not that big” data producer (e.g. no ultra-fast 4D experiments)…

… but we want to be ready for the future

The whole workflow at a CT beamline

The beamline staff supports (collaborative) users in each of the following steps:



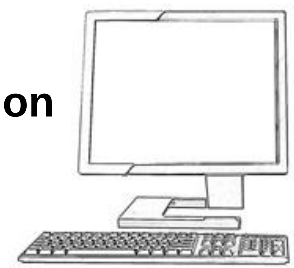
Usually each step is considered as an independent “black box”

An integrated approach (from the idea to the publications) is desirable

Elettra
Sincrotrone
Trieste

The whole workflow at a CT beamline

image reconstruction



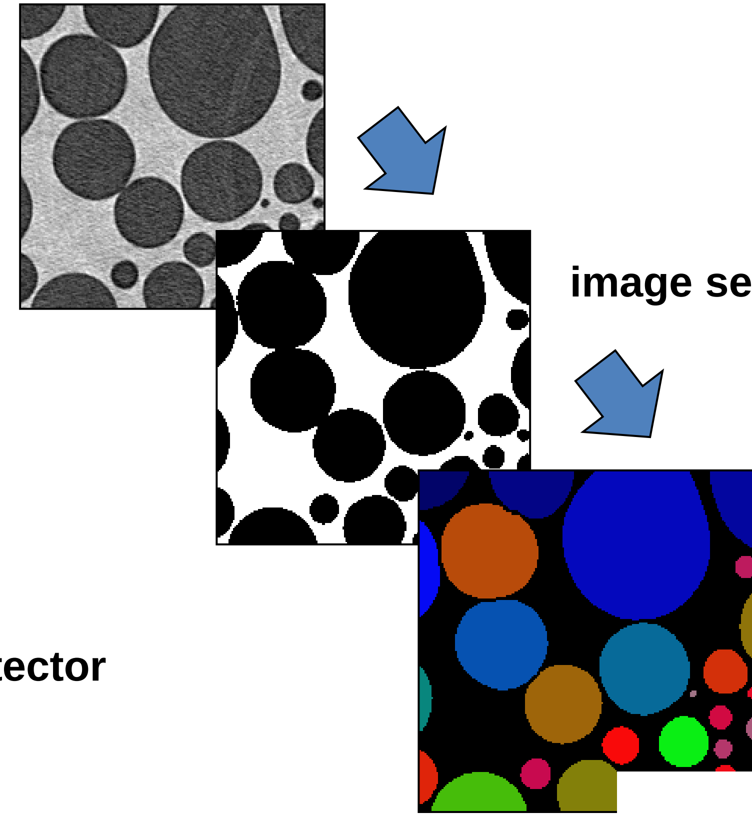
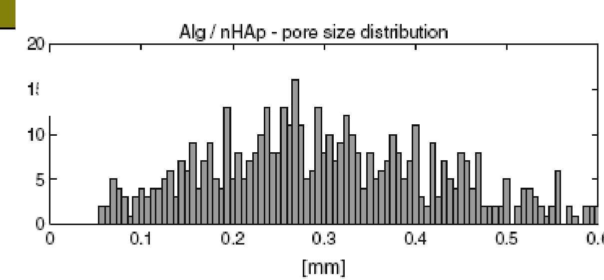


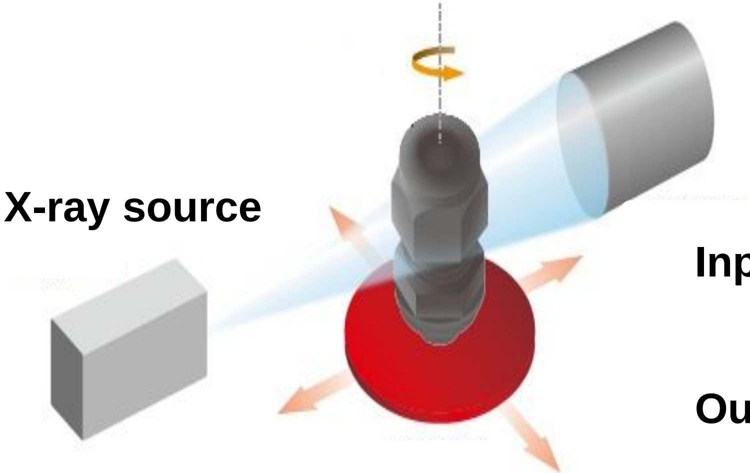
image segmentation

pores quantification



detector

X-ray source



Input:

X-ray photons

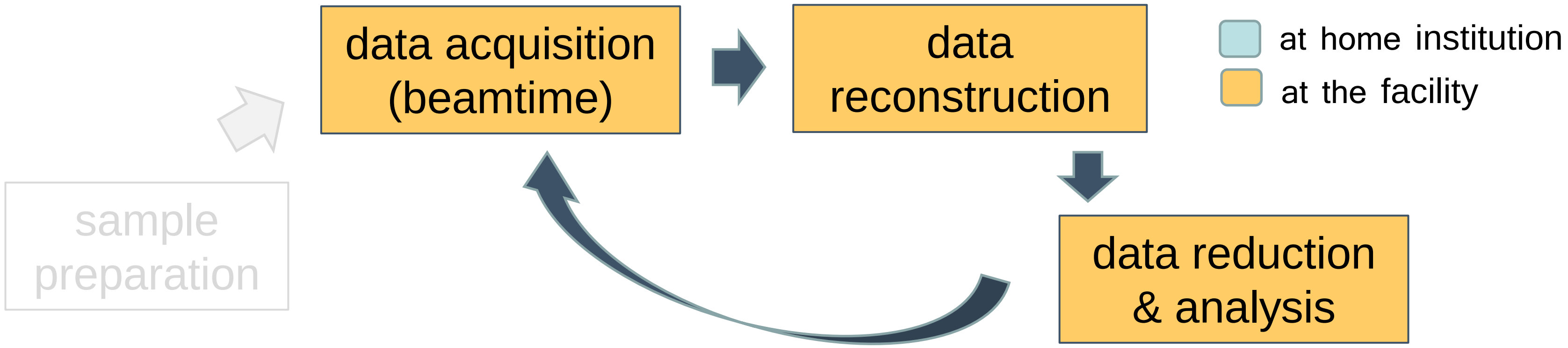
Output:

quantitative characterization of one or more physical properties of the imaged sample

The ideal scenario

Ideally it would be nice to have tools to optimize all the steps at the facility

Reconstruction and a preliminary data analysis give feedback to the acquisition



This integrated approach could lead to the more effective results

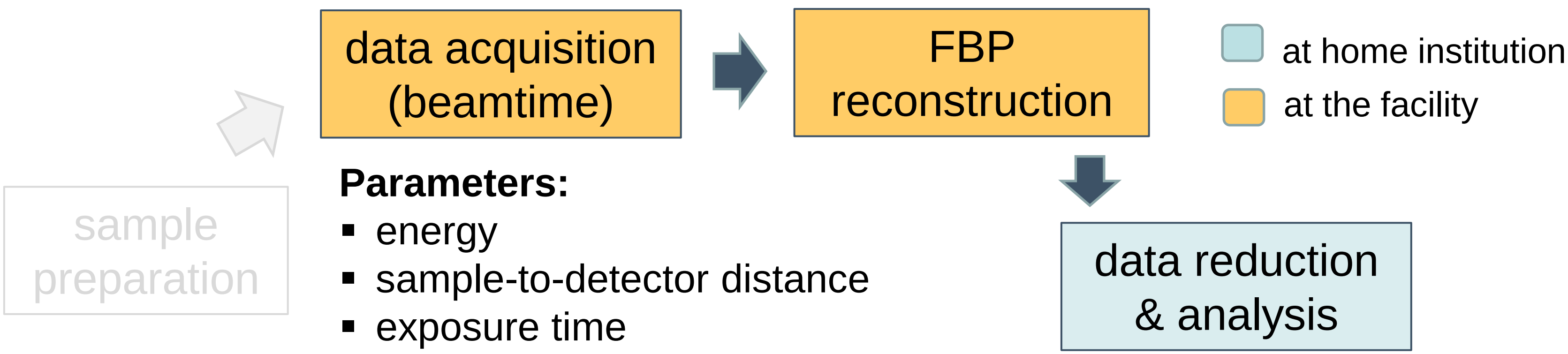
Investments in “computing” (computer vision, image processing, …) are required

Synchrotrons have to favor investments in “physics” and instrumentation

The common scenario

A “standard” fast reconstruction (FBP or GridRec) is available at the facility

A “standard” (i.e. equally spaced angles) acquisition is usually performed



Parameters:

- energy
- sample-to-detector distance
- exposure time
- …

Users prefer to focus onto a fine tuning of the acquisition parameters

Fine tuning of other (time consuming) reconstruction techniques is not performed

Sometimes users go back home with **sub-optimal reconstructed data**

In most of the experiments **users do not have an analysis protocol** yet

Elettra Sincrotrone Trieste

An example of big data application

A 4D *in situ* experiment performed at the TOMCAT beamline of SLS

~10 TB of reconstructed data on external hard drives

Two weeks after the experiment we received an email from the local contact

Data were removed from the SLS data storage

No backup of raw data ➡ no chances to apply refined reconstruction in the future

nature COMMUNICATIONS

ARTICLE

Received 23 Mar 2012 | Accepted 13 Sep 2012 | Published 16 Oct 2012 DOI: 10.1038/ncomms2134

A four-dimensional X-ray tomographic microscopy study of bubble growth in basaltic foam

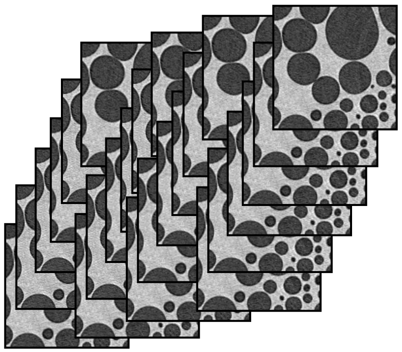
Don R. Baker[1], Francesco Brun[2,3], Cedrick O'Shaughnessy[1], Lucia Mancini[3], Julie L. Fife[4,5] & Mark Rivers[6]

McGill UNIVERSITY

PAUL SCHERRER INSTITUT PSI

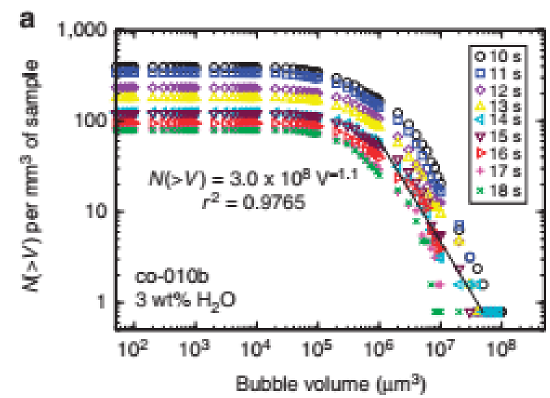
Elettra Sincrotrone Trieste

An example of data reduction



~10 TB of images

data reduction & analysis



~340 KB MS Excel file

There was the need to develop a 4D data reduction protocol

There was the need to implement the 4D data reduction protocol

There was the need to execute the 4D data reduction protocol

Elettra Sincrotrone Trieste

Quantitative results **one year after** the beamtime

The challenges

Open issues in practical SR μ-CT:

- What can be/has to be done at/by the facility (during the beamtime)?
- What can be/has to be done at/by the home institution?

- To which extent is the facility responsible to the **storage** of users' data?

  (The data policy requires to keep users' raw data for years…)

  (… what about the reconstructed and processed data)?
- What about the **computational power** and **software** for data reduction?

These issues have to be faced in the big data (TBs/day) scenario

The answers are basically application-dependent

Elettra Sincrotrone Trieste

Big *meaningful* data

Not all the scans at the beamline lead to meaningful data

To optimize a beamtime more feasibility studies are required

Pre-characterization of the samples is often useful

A custom easy-to-access desktop cone beam micro-focus CT device has been designed and built at Elettra (TOMOLAB)

https://www.elettra.trieste.it/it/lightsources/labs-and-services/tomolab/tomolab.html

Data-gravity

The general trend to face the “big data” issue is:

“The data cannot be moved, let’s move the software”

With this approach the data producer facility has to offer software for remote:

- data reconstruction
- data visualization
- data reduction and analysis

An interesting application of this approach is the Australian MASSIVE project

MASSIVE | Multi-modal Australian ScienceS Imaging and Visualisation Environment

Australian Synchrotron CSIRO MONASH University NCI Victoria VPAC

Elettra Sincrotrone Trieste

The *Pore3D* project

Pore3D is a Software-as-a-Service (SaaS) consisting in several state-of-the-art functions and procedures for performing reconstruction, filtering, segmentation and quantitative analysis of 3D µ-CT

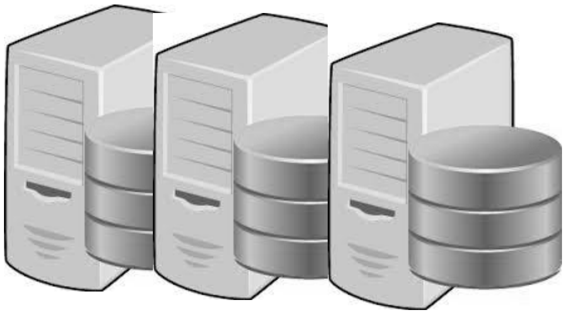
http://www.elettra.eu/pore3d/index.html

Connection to a remote Linux desktop via *NoMachine*

IDL session can then be executed

IDL has an additional package called *Pore3D*

More computational power than common hardware



IDL pore3d

remote reconstruction and analysis of porous media

Multi-facility (multi-modal) experiments

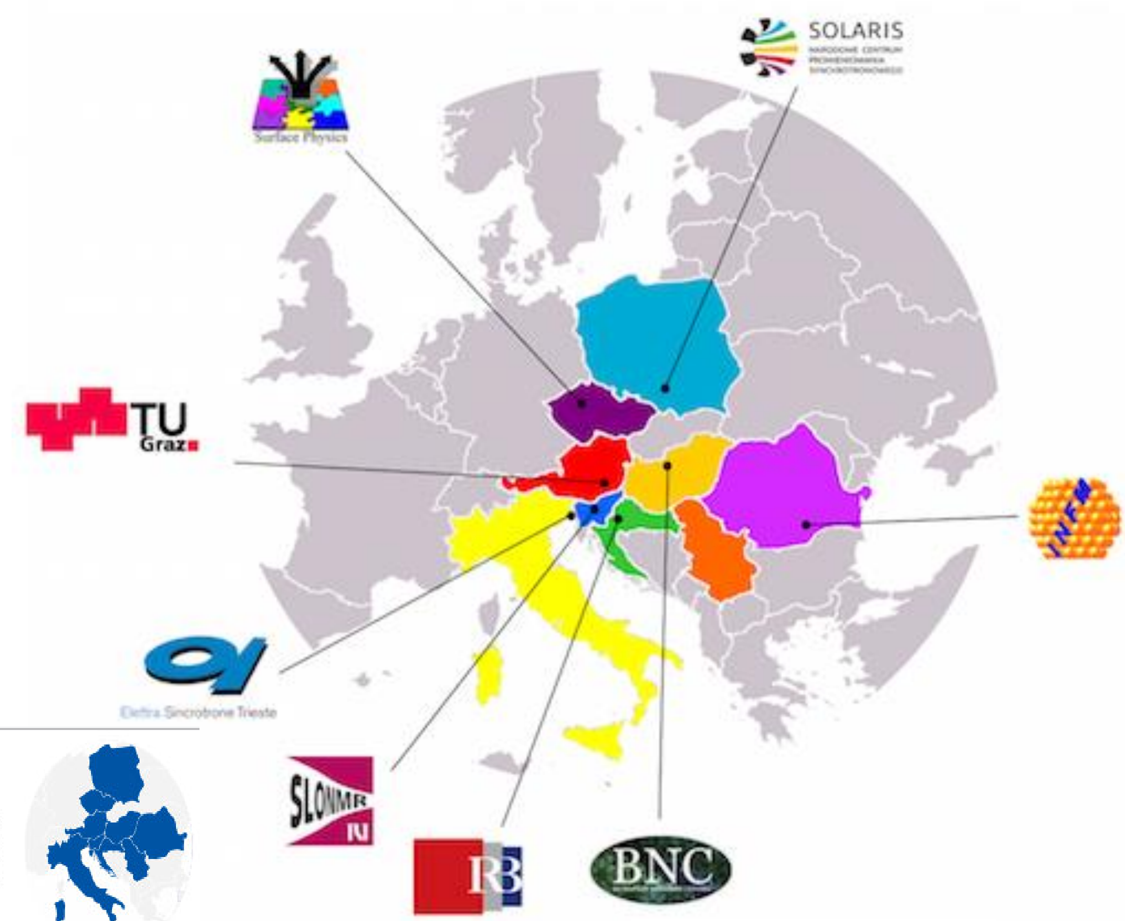
Another trend in science deals with **multi-modal experiments**

Different big data producers might be involved in a single project

Elettra is partner of Ceric-Eric: a single proposal to access multiple facilities

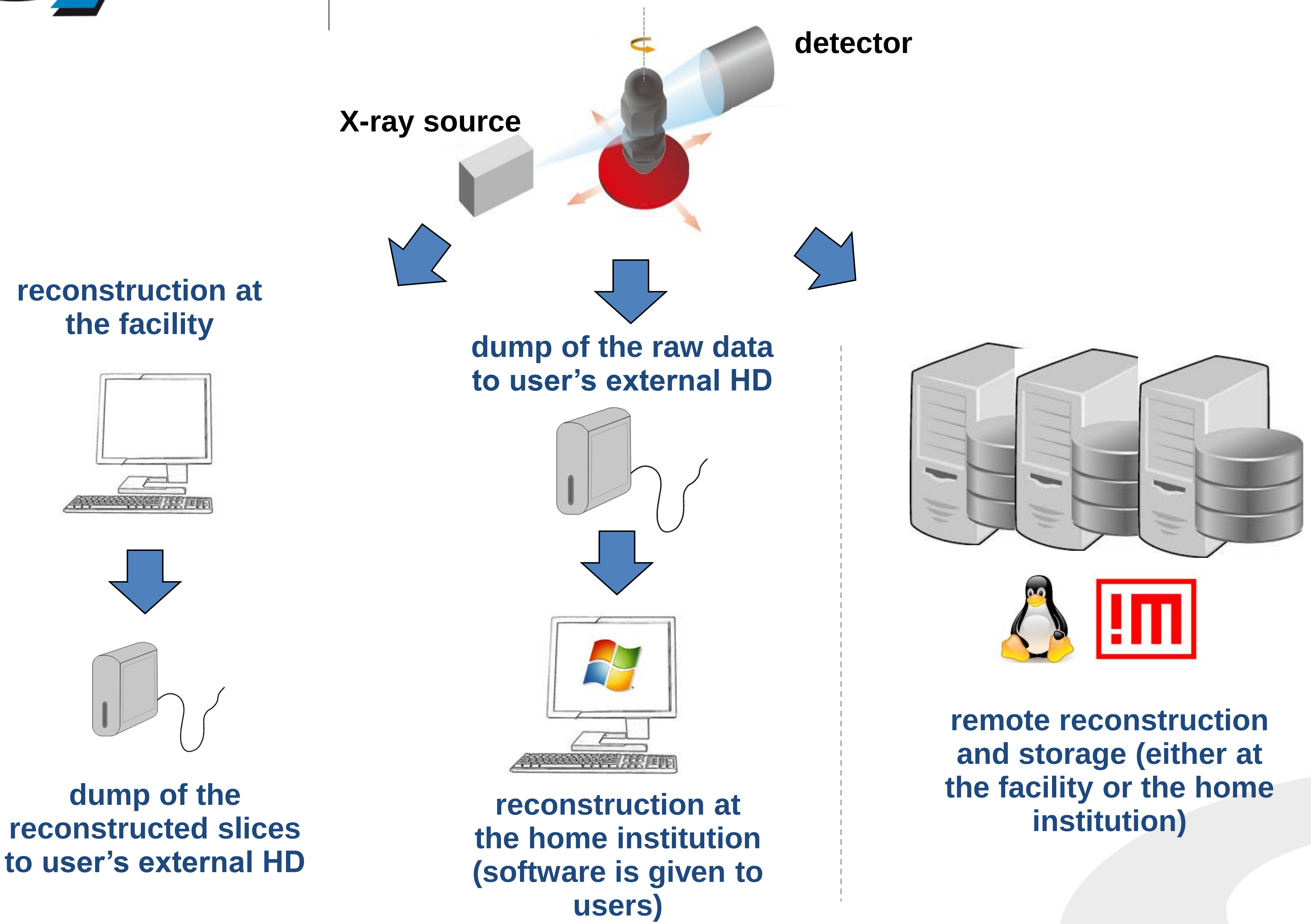
A user can e.g. do µ-CT in Trieste and µ-MRI in Ljubljana

Then co-registration/correlation of the data might be required

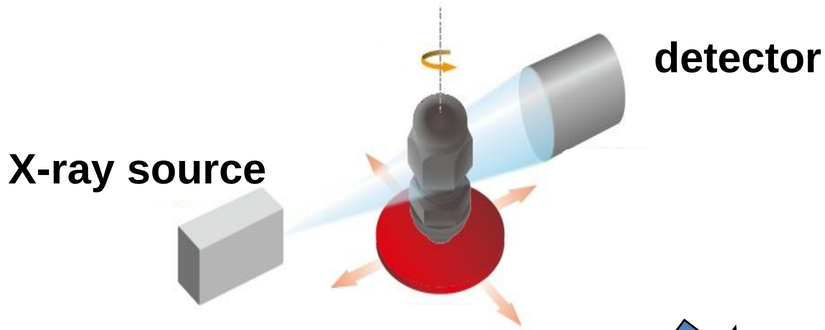


Central European
Research Infrastructure
Consortium

Flexibility at SYRMEP



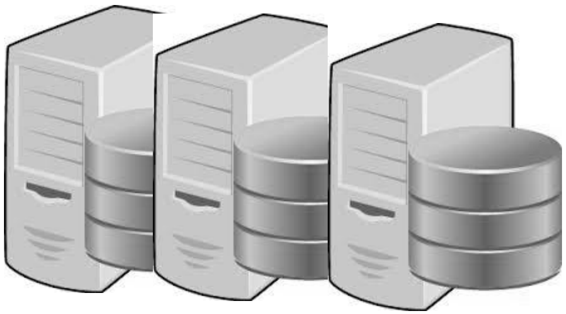
Remote computation: pros and cons



Pros and cons:

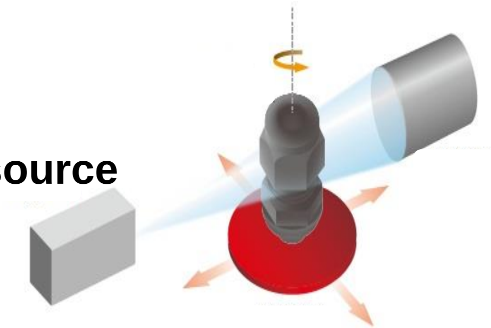
- Faster and “big data-ready” workflow
- Less expensive for users
- More expensive for the facility (remote storage and computational infrastructure is required)

- Users cannot exploit their software licenses, e.g. VG Studio MAX, Avizo, Matlab, ... (unless data download to their home institution is performed)

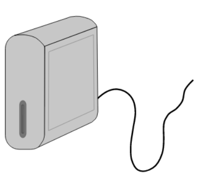


remote reconstruction and storage (either at the facility or the home institution)

Local computation: pros and cons

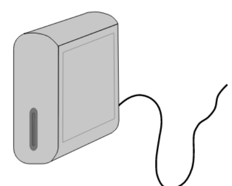


reconstruction at the facility



dump of the reconstructed slices to user's external HD

dump of the raw data to user's external HD



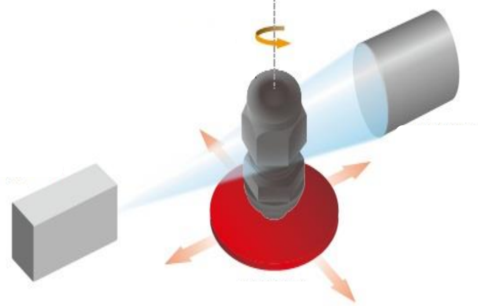
reconstruction at the home institution (software is given to users)

Pros and cons:

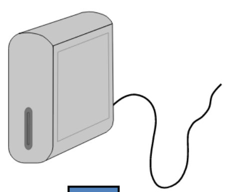
- Simpler for the facility (simple hardware and storage is required)
- Easier when users want to exploit later their software licenses (e.g. VG Studio MAX, Avizo, Matlab, ...)
- No network data transfer

- Slower reconstruction workflow

Local reconstruction



dump of the raw data to user's external HD



reconstruction at the home institution (software is given to users)

At SYRMEP we care about post-beamtime support

Almost all users are collaborative users

SSDs and GPUs are now affordable for users

Raw data is smaller than reconstructed data

Reconstruction of relatively big data in a reasonable time with common hardware is feasible

We give users for free a user-friendly software for the reconstruction of phase contrast CT data with common hardware

Users can re-reconstruct past acquired data (with refined tools not available at the time of acquisitions)

The *SYRMEP Tomo Project* software

The goal:

- reconstruct SR phase contrast μ-CT data with **common hardware**
- **user friendly** reconstruction tool
- it includes the most common image processing pre- and post- reconstruction
- a (cheap) PC with a SSD and a GPU reconstructs in a reasonable time
- developed also for **teaching/educational** purposes

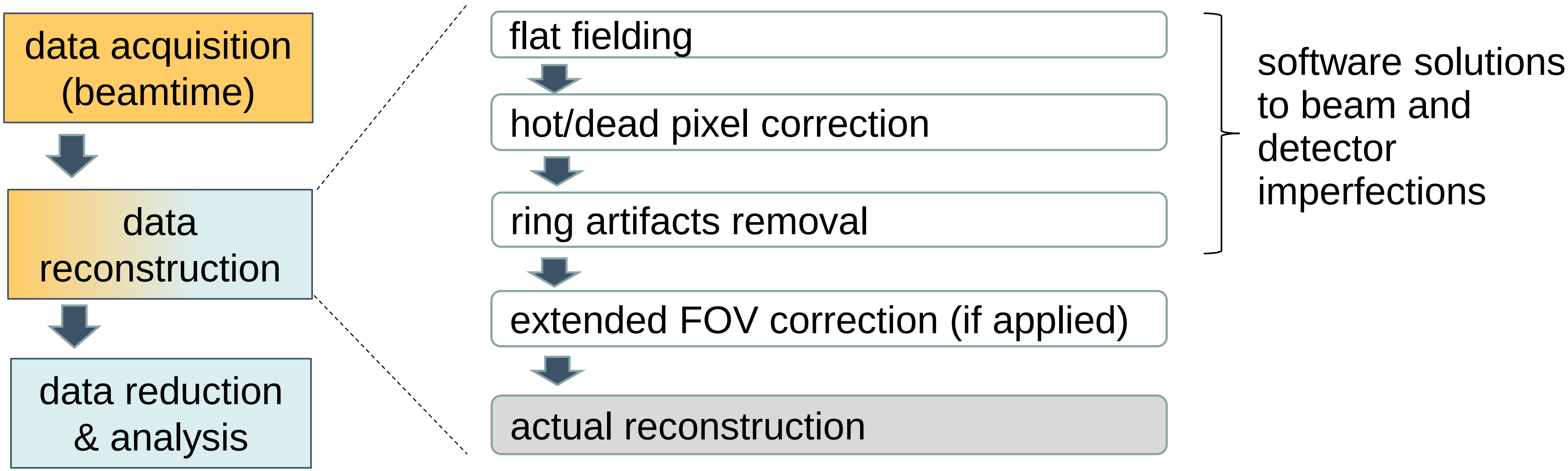
History:

- originally developed in IDL
- recently re-designed to exploit the *ASTRA Toolbox* and *TomoPy*
- GUI for Windows wrapping Python scripts

Available for free to SYRMEP users (available for free to everyone very soon)

The pre-reconstruction pipeline

An essential **reconstruction** pipeline includes



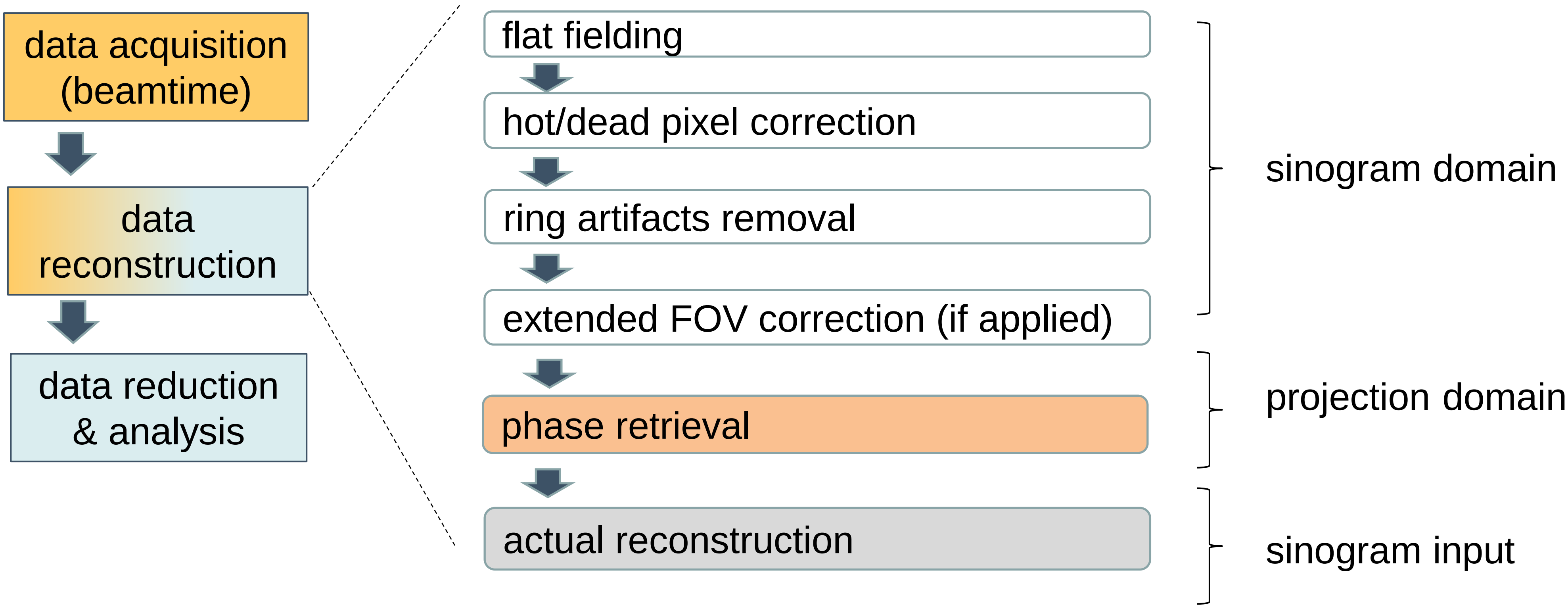
All these steps have parameters

Reconstruction is not a “black box” (fine tuning of the parameters is desirable)

Pre-reconstruction and phase retrieval

Elettra Sincrotrone Trieste

With phase retrieval an essential **reconstruction** pipeline is:



Dynamic flat fielding [1] has been recently proposed which operates with projections

[1] V. Van Nieuwenhove, J. De Beenhouwer, F. De Carlo, L. Mancini, F. Marone, and J. Sijbers, *Dynamic intensity normalization using eigen flat fields in X-ray imaging,* Optics Express, vol. 23, issue 21, pp. 27975-27989, 2015.

File format issue

The past: thousands of **2D projection files**

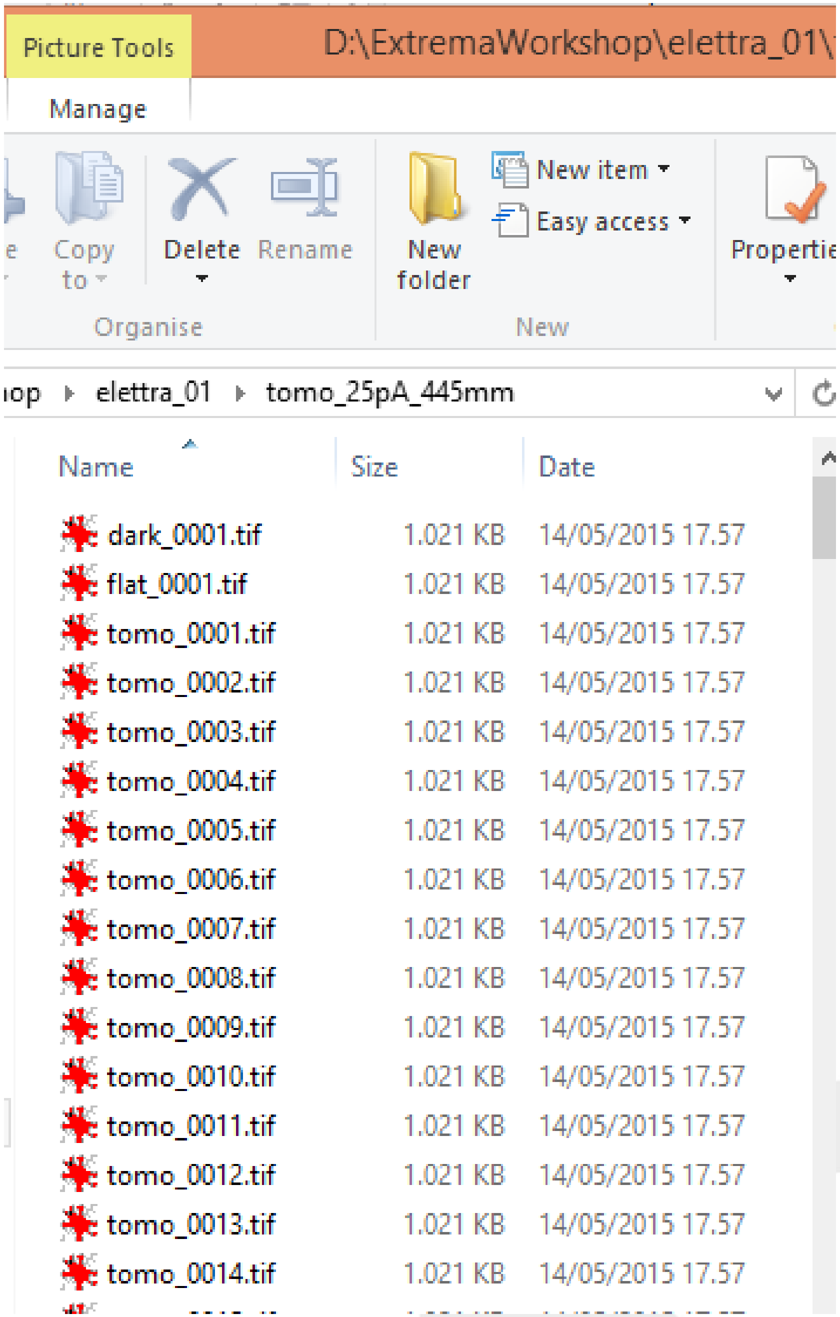
- at e.g. Elettra and SLS: TIFF files
- at e.g. ESRF: custom EDF files

Metadata in a separate text file

(a few metadata into the EDF)

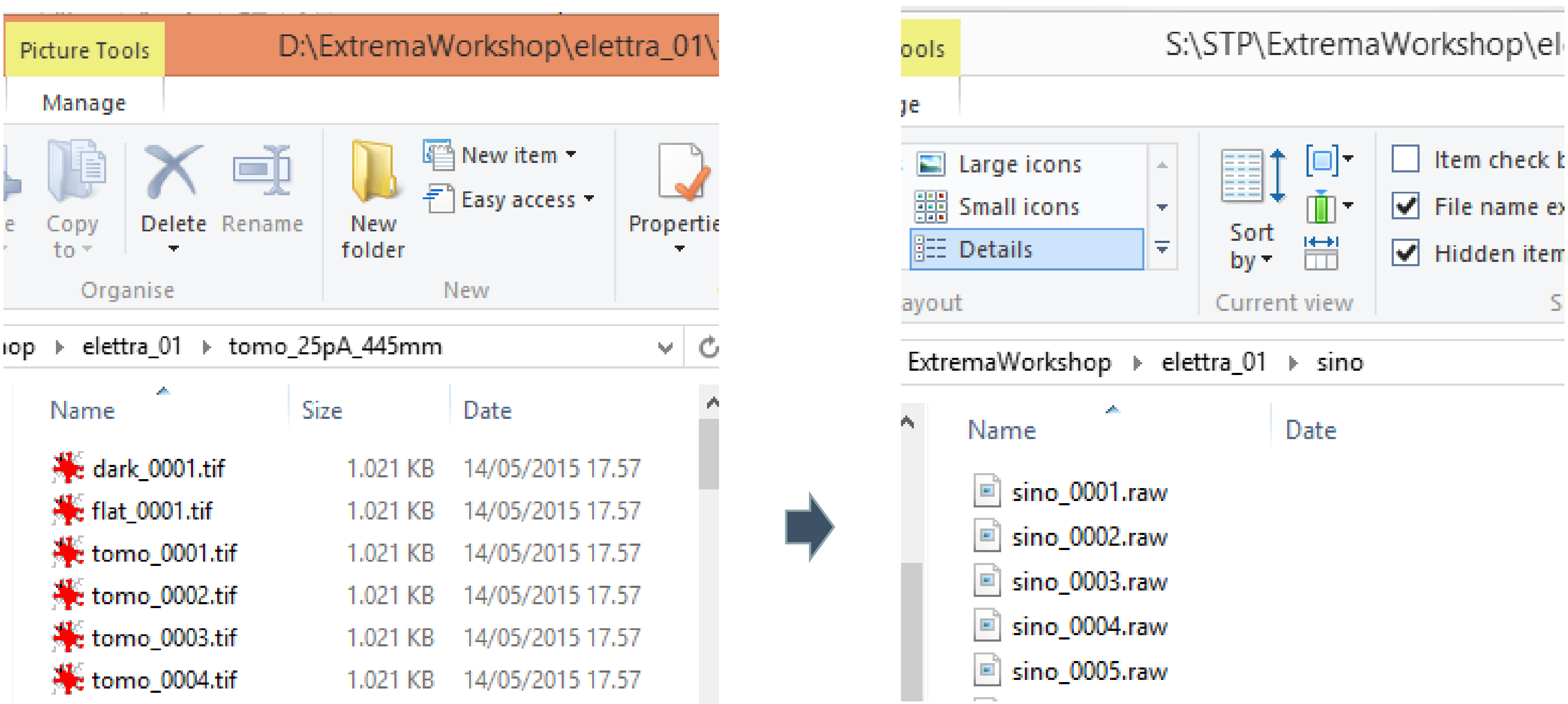
The motivations:

- **32 bit file systems** (file size < 4GB)
- Most detector control software uses TIFF



The Elettra scenario in the past

Sinogram files were created in parallel during the acquisition



Sinograms are the actual input of the reconstruction process

The HDF5 format

The scientific community is proposing HDF5 as a “standard” format

The idea: a single “volume” file with all the acquired projections and metadata

Advantages:

- data and **metadata** in the same file
- easy **switch between sinogram- and projection-** perspective
- no redundant data (projection files **and** sinogram files)
- better lossless **compression**

Disadvantages:

- **parallel writes** are still a concern
- not yet widely supported by commercial (image viewer) software

The *DataExchange* initiative

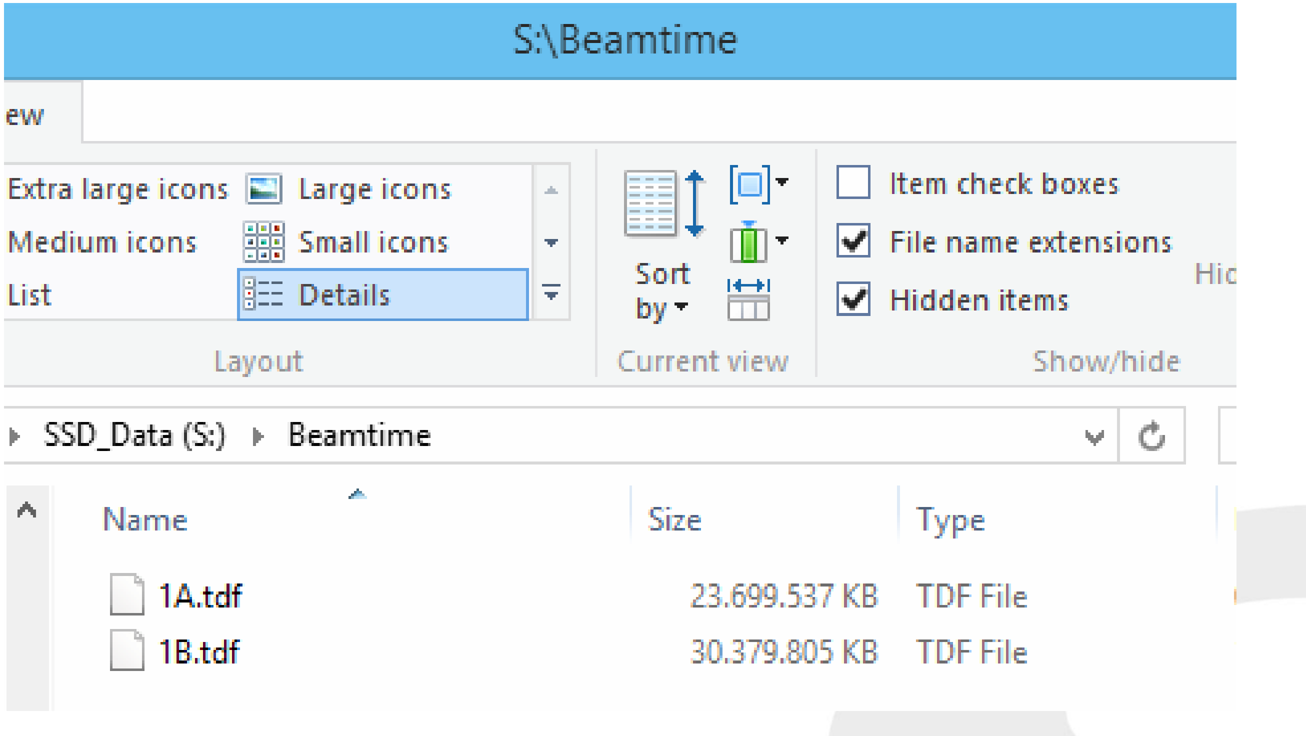
The APS promoted the *DataExchange* initiative

Roughly, a way to standardize the “internal” fields of a HDF5 file

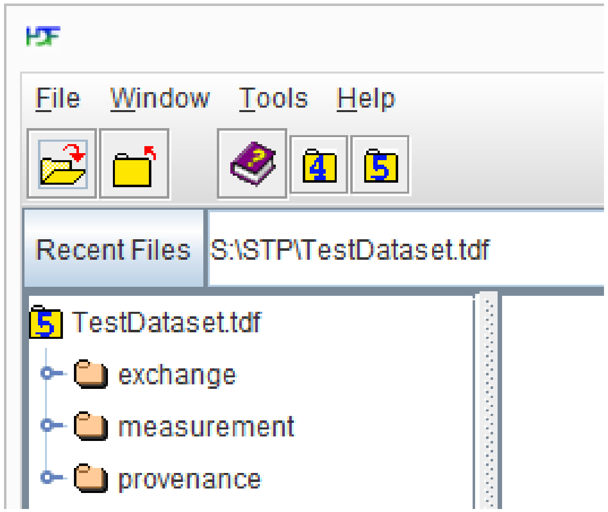
The initiative is now followed by **APS, SLS** and **Elettra**

At Elettra we use .tdf instead of .h5 (or .hdf5) for the file extension

TDF stands for *Tomographic Data Format*

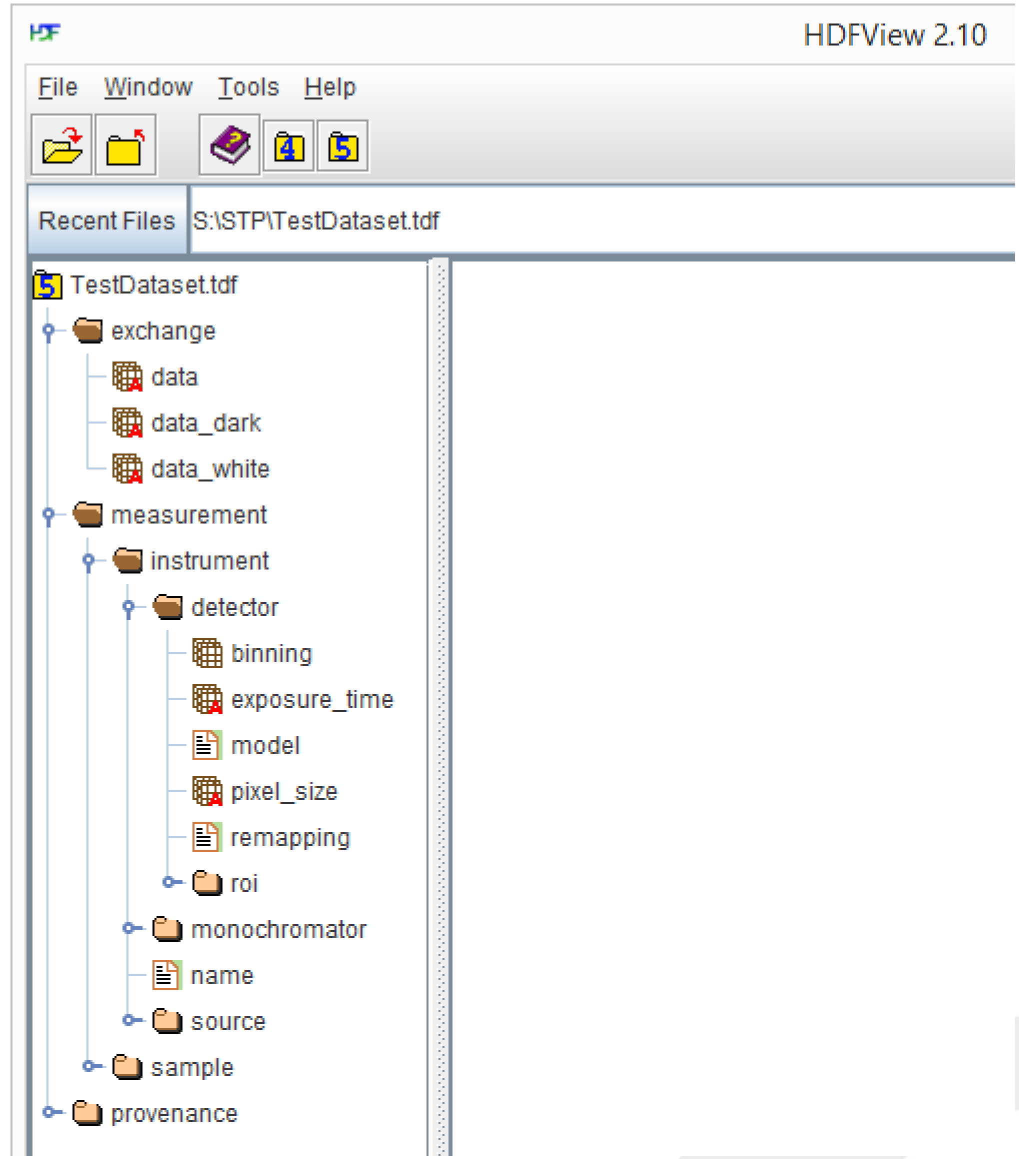


HDF5 for tomography



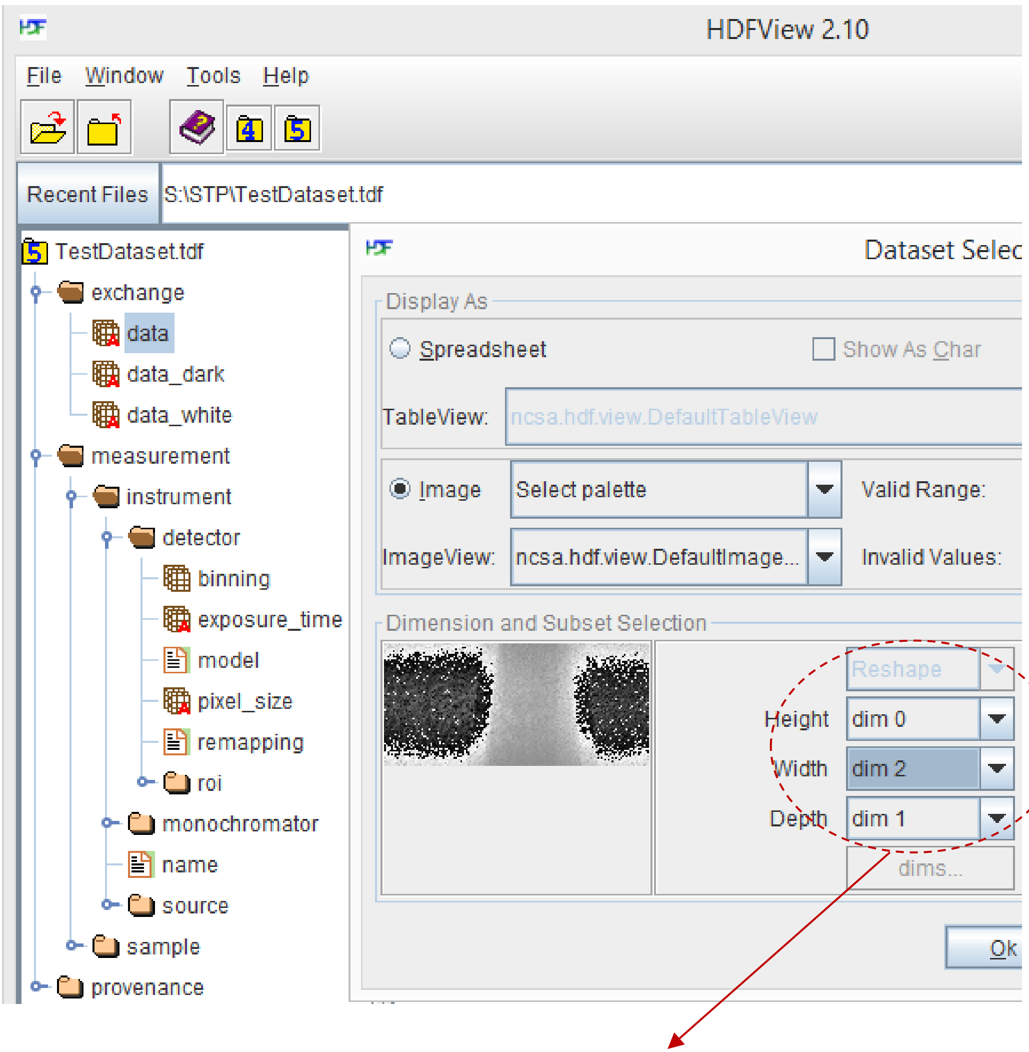
Acquisition metadata saved into the HDF5 file

Pre-processing as well as reconstruction parameters can be also saved within the HDF5

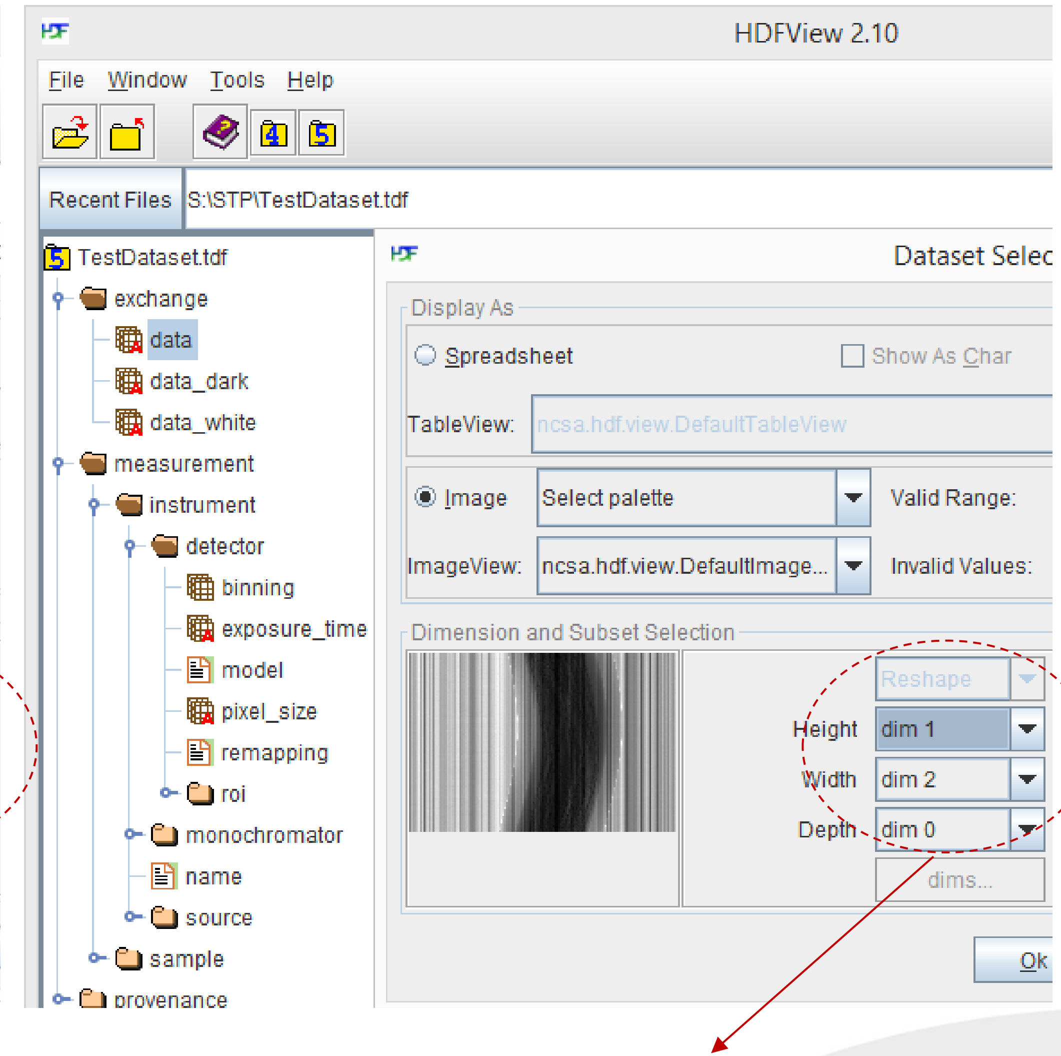


Elettra Sincrotrone Trieste

Sinogram- and projection- perspective



Projection (non-contiguous)

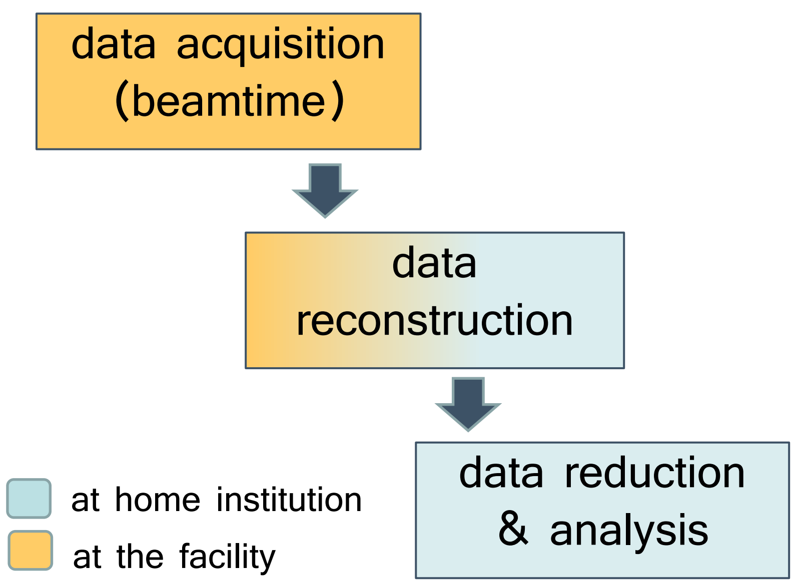


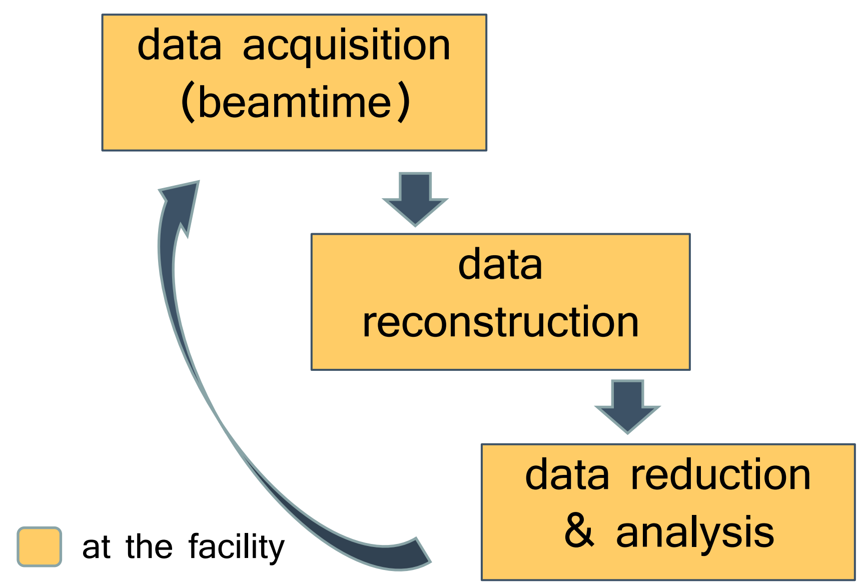
Sinogram (contiguous with faster I/O)

Conclusion

A challenge for SR CT is to have **more feedback as possible** at the facility

To tackle the big data issue, data reduction has to be done as soon as possible





The development of an effective integrated approach is **application dependent**

For instance, the direct reconstruction of segmented data is interesting

An interesting application

Mammo-CT with extremely low dose (a few and noisy projections)

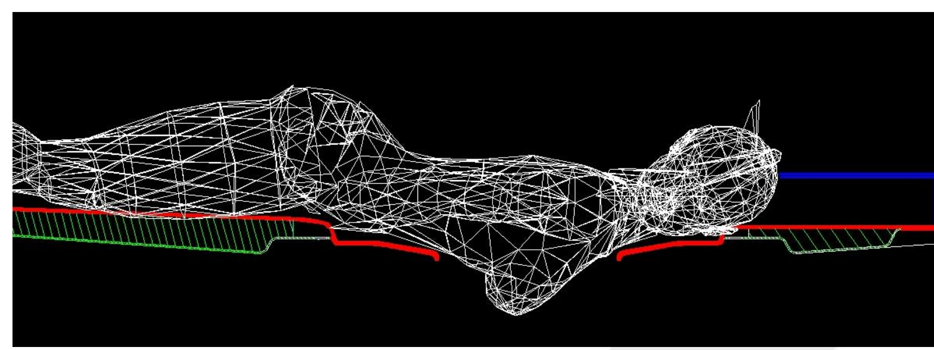
The "standard" approach, i.e.

- equally spaced angles
- Paganin's phase retrieval
- FBP or GridRec

leads to pore results

A contribution from the community of mathematicians is more than welcome

(Not a pure big data application)



Elettra Sincrotrone Trieste

Acknowledgements

Many thanks to the SYRMEP family:

- Giuliana Tromba
- Lucia Mancini
- Franco Zanini
- Diego Dreossi
- Nicola Sodini
- Francesco Brun

Thanks also to the Scientific Computing:

- Georgos Kourousias
- Fulvio Billè
- Roberto Pugliese

www.elettra.eu