

# Synchrotron SOLEIL

## The Data Access Challenge



*Alain BUTEAU (on behalf of S. Poirier and G. Viguier)*

## Data Access Challenge

Motivations and new ideas to manage the  
problem

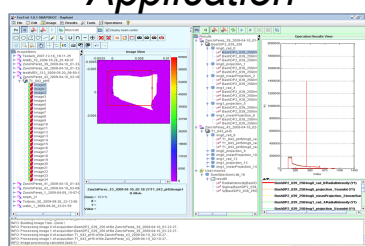
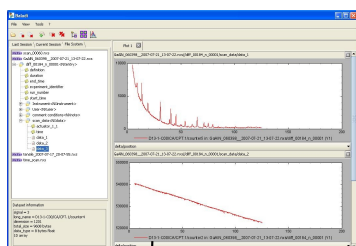
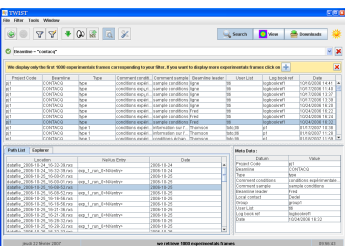
# NeXus Files choice : Is it Enough ?



File retrieval

File browsing

foxtrot Data Analysis Application

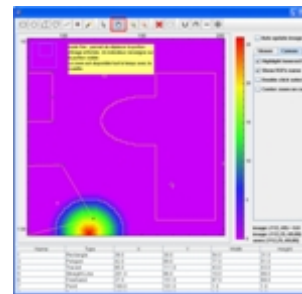
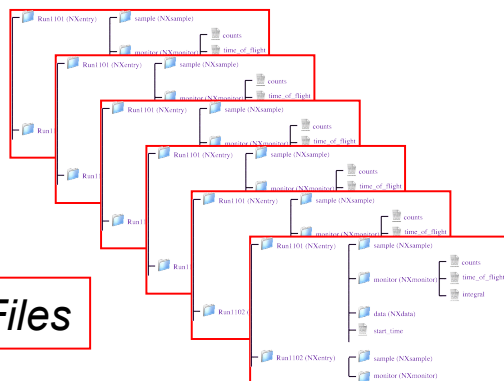


•Thanks to a unique API and a « SOLEIL standardized internal data organization », we have :

- ✓ developed common software solutions for all beamlines
- ✓ Decoupled the development of Acquisition software from Data Analysis software

NeXus Application Interface

SOLEIL NeXus Files



The COMETE library of data visualization components  
<http://comete.sourceforge.net>

# NeXus Files choice : is it enough ?

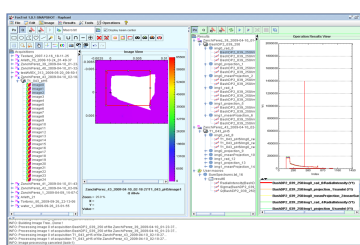
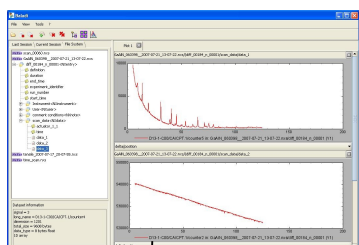
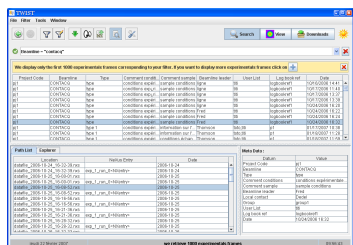
File retrieval

File browsing

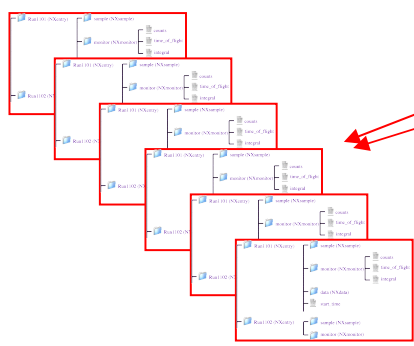
foxtrot Data Analysis Application

Data Analysis Application B

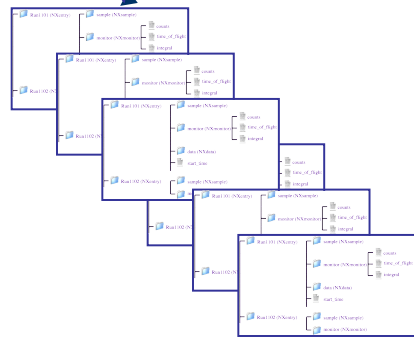
Data Analysis Application C



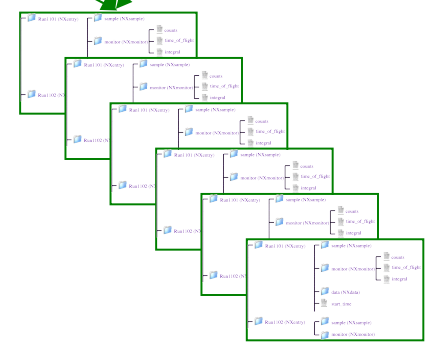
NeXus Application Interface



SOLEIL NeXus Files



ESRF Files



DESY Files

❖ The real needs for scientists are to :

- ✓ Find solutions to data format issues from the **data analysis point of view** ?
- ✓ Put in **common** different **algorithms** for analyzing data ?
- ✓ Find the most suitable ways to **exchange data** ?

## ❖ The foreseen solutions :

- ✓ Choose Nexus/HDF5 data format

**Why not ? But it's not enough**

- ✓ Define a standard internal data file structure for experimental data storage

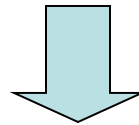
- ✓ **It's a complex process, involving :**

- ✓ **Many institutes**

- ✓ **many software developers**

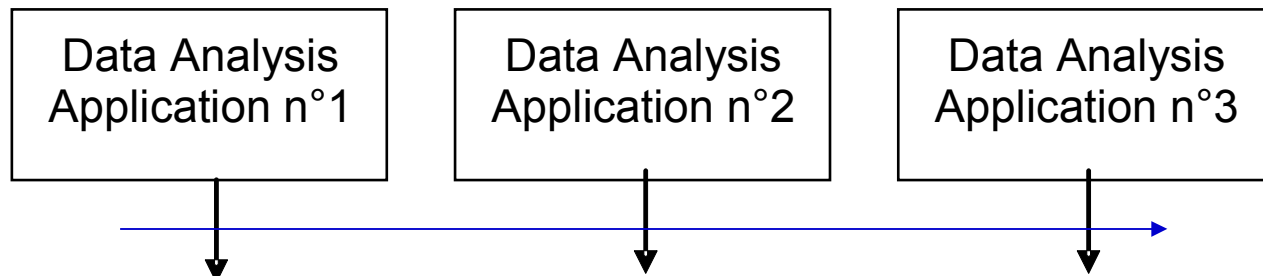
- ✓ **many existing data format and files**

- ❖ Proposal from the MAHID Group :
  - ✓ Access data thanks to **attributes/tags**
  - ✓ Define a standard way to access synchrotron experimental data

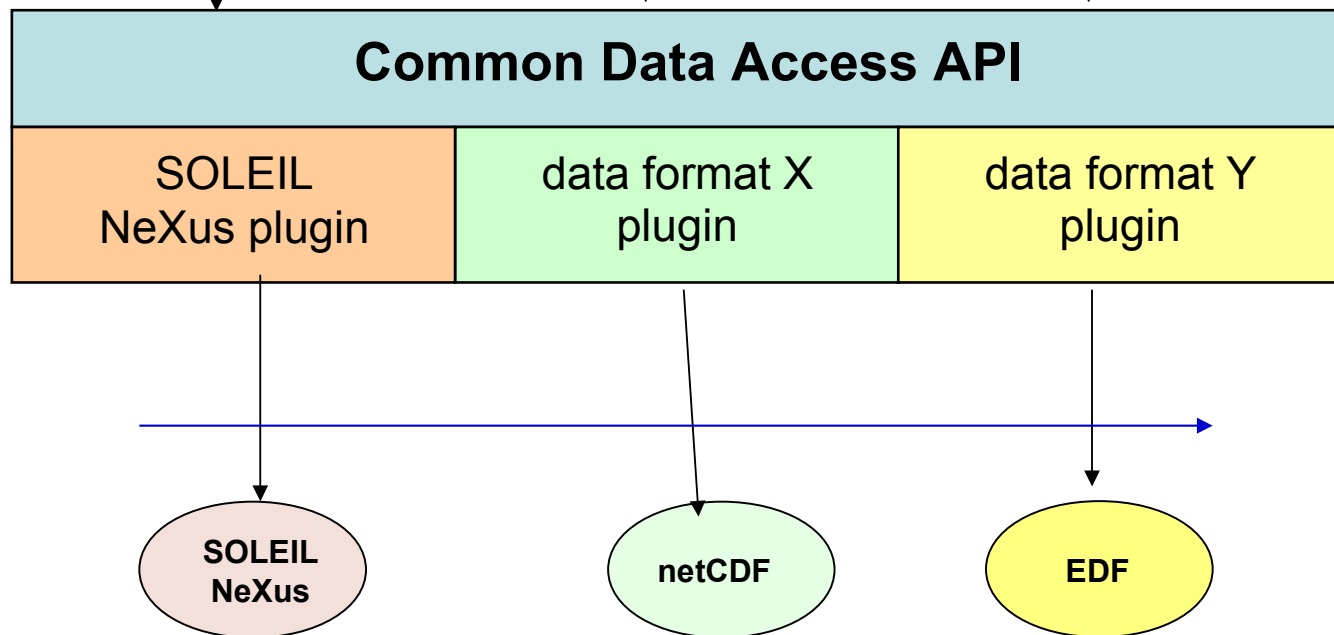


- ❖ We only agree on an “**Abstract Software Interface**” defining a Data Access Model
- ❖ We let **Institutes implement this Software Interface** to deal with their own data files

**Adapt each application to the Common DataAccess API**



**Implement CommonDataModel plugin for each data format**



# What is the role of a CommonDataModel plugin ?

- ✓ Provide Data thanks to keywords
- ✓ Implement the standard definition of these keywords

Data Analysis Application

Common Data Access API

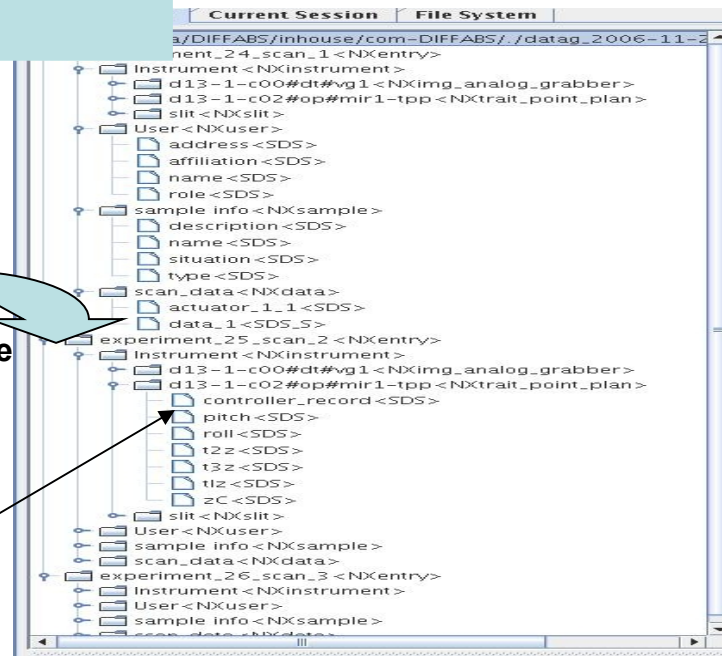
Let scientists agree on keyword definitions

getDataItem(« pitch »)

SOLEIL NeXus plugin

Nexus File

Look in dictionary



Pitch ->experiment/instrument/d13/pitch  
Roll ->experiment/instrument/d13/Roll

\*  
\*  
\*  
\*

```

Double getDataItem(« pitch »)
{
    pitch = dictionary.get(« pitch »)
    return pitch *100;
}
    
```

- It's of course a very long process
  - But it has already been done by the imgCIF community (at least) for crystallography
  - Let's start from this point rather than redoing the job in the NeXus community

```

imgCIF - Recherche Google  DRAFT CIF (imgCIF DEFINITION)  of_img_cif.v1.1.3
http://www.bernstein-plus-sorns.com/software/CBF/doc/cif_img_1.1.3.html
Mantis  Controle  Tableau de bord [Indu...  QRAO - DC  Sonatype Nexus Rev...  0013298: Migration g...  Annuaire
# UC# CBF# CBF# #
# # CBF #
# # imgCIF #
# # #
# # # Extensions Dictionary #
#####
# #
# # Image CIF Dictionary (imgCIF)
# # and Crystallographic Binary File Dictionary (CBF)
# # Extending the Macromolecular CIF Dictionary (mmCIF)
# #
# # Version 1.1.3
# # of 2001-04-19
# #
# # Adapted from the
# # imgCIF Workshop, BNL Oct 1997
# # and
# # Crystallographic Binary File Format Draft Proposal
# # by Andy Hammersley
# #
#####
# #

```

```

imgCIF - Recherche Google  DRAFT CIF (imgCIF DEFINITION)  of_img_cif.v1.1.3
http://www.bernstein-plus-sorns.com/software/CBF/doc/cif_img_1.1.3.html
Mantis  Controle  Tableau de bord [Indu...  QRAO - DC  Sonatype Nexus Rev...  0013298: Migration g...  Annuaire
category:
sample
__item.name      '_diffn_radiation.diffn_id'
__item.mandatory_code  yes
save_
save__diffn_radiation.div_x_source
__item.description.description
Beam crossfire in degrees parallel to the laboratory X axis
(see AXIS category).

This is a characteristic of the xray beam as it illuminates
the sample (or specimen) after all monochromatation and
collimation.

This is the esd of the directions of photons in the X-Z plane
around the mean source beam direction.

Note that some synchrotrons specify this value in milliradians,
in which case a conversion would be needed. To go from a
value in milliradians to a value in degrees, multiply by 0.180
and divide by Pi.

__item.name      '_diffn_radiation.div_x_source'
__item.category_id  diffn_radiation

```

- ❖ It is a light process
  - ❖ *Developing a plugin costs a few weeks of work*
  - ❖ *Adapting an application costs a few weeks of work*
- ❖ It allows to deal with existing files
- ❖ It is an open process
  - ❖ *Newcomers have only to implement the standardised interface*



Well , well

These are beautiful ideas but where is the  
source code ?

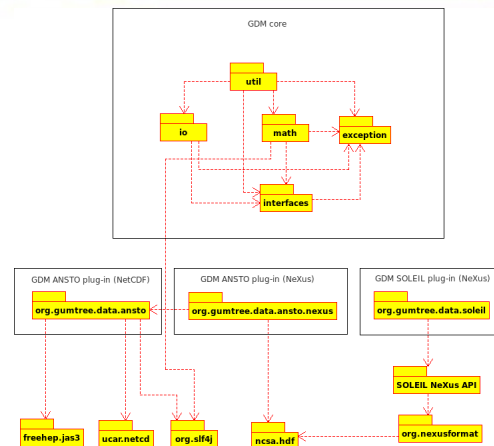
- Data file access abstraction in collaboration with ANSTO (Common Data Model)
  - Project initiated by ANSTO  
(<http://gumtree.codehaus.org/GumTree+Data+Model/>)
  - The CDM API allows to create generic data analysis tools
    - without need to know about experimental data files format(s)
  - A set of data format plugins have to be developed by each facility so a generic application become able to read and process data files from these facilities
    - Current plugins: NetCDF (ANSTO), NeXus (Soleil, V1)
  - CDM API is currently available in Java



# SOLEIL/ANSTO Collaboration goals and milestones

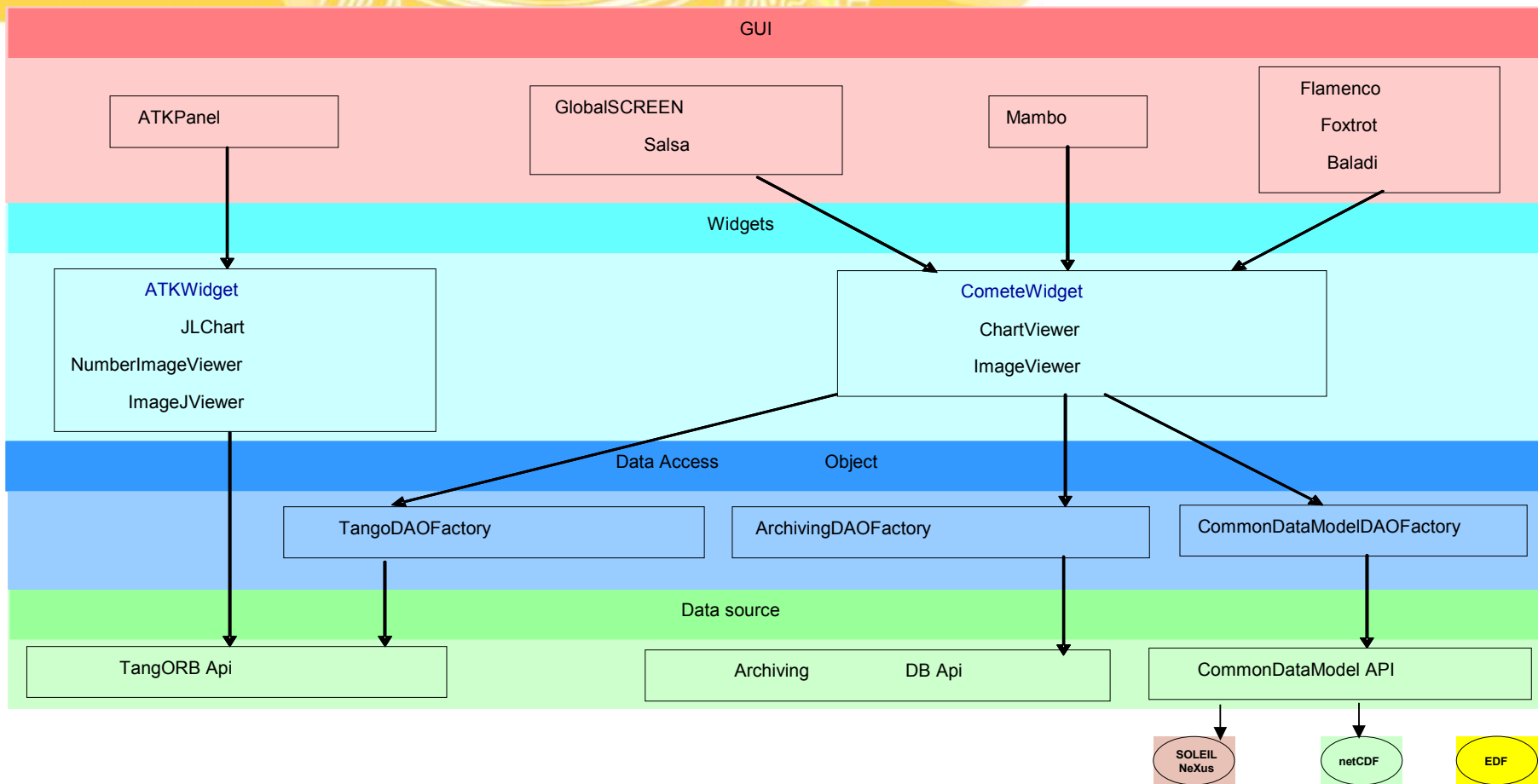
First share the Common Data Model API and then

- ✓ the graphical components (COMETE and GumTree UI) made on top of the “CommonDataDataModel API”
- ✓ the graphical applications made on top of these graphical components
- ✓ the data analysis algorithms used in the GumTree and COMETE frameworks



<i>Who</i>	<i>What</i>	<i>When</i>
SOLEIL	Sends ANSTO the feedback on the current GumTree java API	Done
SOLEIL-ANSTO	Discussion on the evolution of the current java API and agreement on the interface architecture to have an official V1.0 of the “Common DataModel”	Done
SOLEIL	C++ implementation of the API	In Progress by SOLEIL
SOLEIL	Java implementation of EDF file plugin for foxtrot application	To be done (could ESRF help?)
SOLEIL/ANSTO	Propose the CDM to the HDF group	Summer 2010

# Just a reminder of the overall GUI application architecture



# Conclusion

## We need you



- ❖ Good wills are welcome to :
  - ❖ At the CDM level :
    - ❖ Help doing the C++ port
    - ❖ Write CDM plugins for their own institutes format
  - ❖ At COMETE level :
    - ❖ Writing new GUI components
    - ❖ Writing (or enhancing) applications of common interest (NeXusExplorer, foxtrot, ..)