

# **Collaboration Experiences in the DNA and EDNA projects**

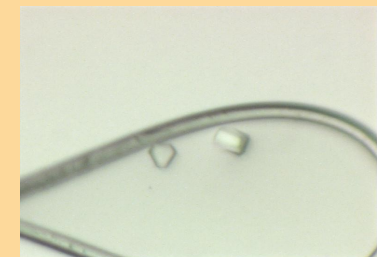
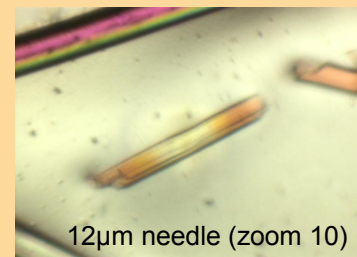
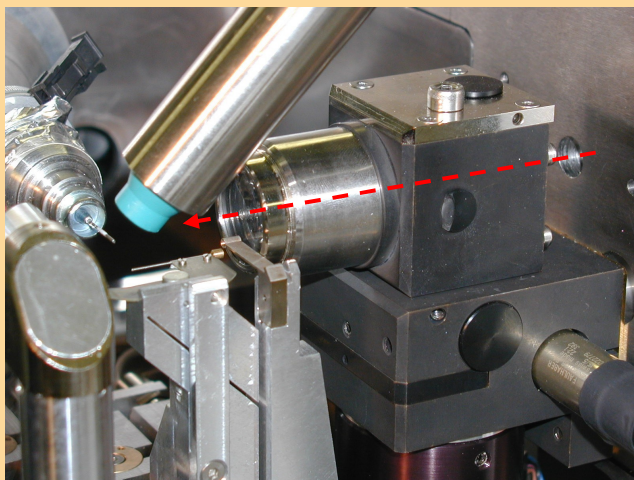
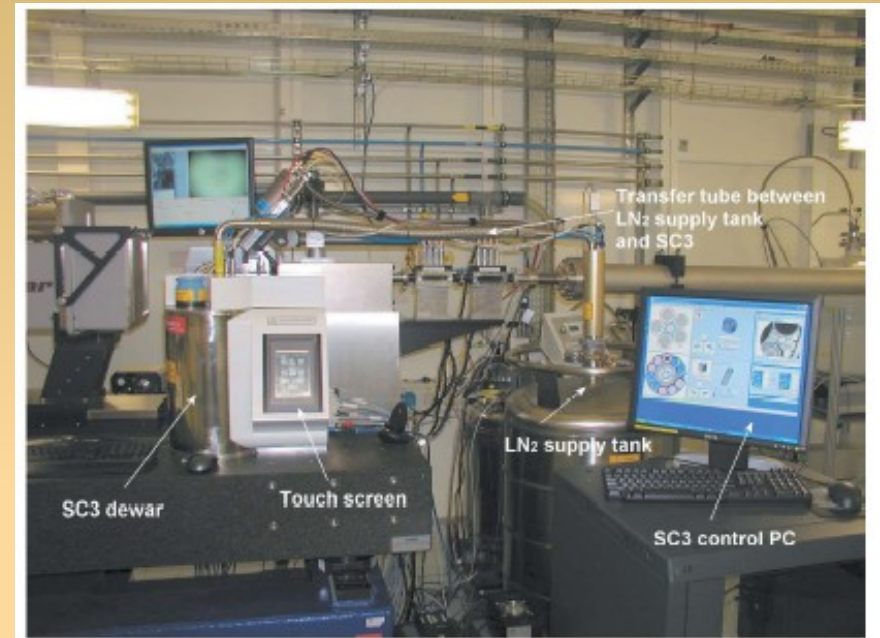
Olof Svensson  
EDNA Project Manager  
ESRF

# Introduction

- When starting a new collaboration :
  - Constraints
  - Slooow advances...
- Therefore, if you want to collaborate :
  - Be patient (within reasonable limits)
  - Be understanding – try to see the problem with different eyes
  - Make the software modular
  - Implement tests
- Well-managed collaboration leads to better software

# Automation of MX crystallography

- Motivation behind the MX collaboration :
  - Same experiments (different hardware)
  - Same data processing techniques
  - System too complex for a single facility to develop

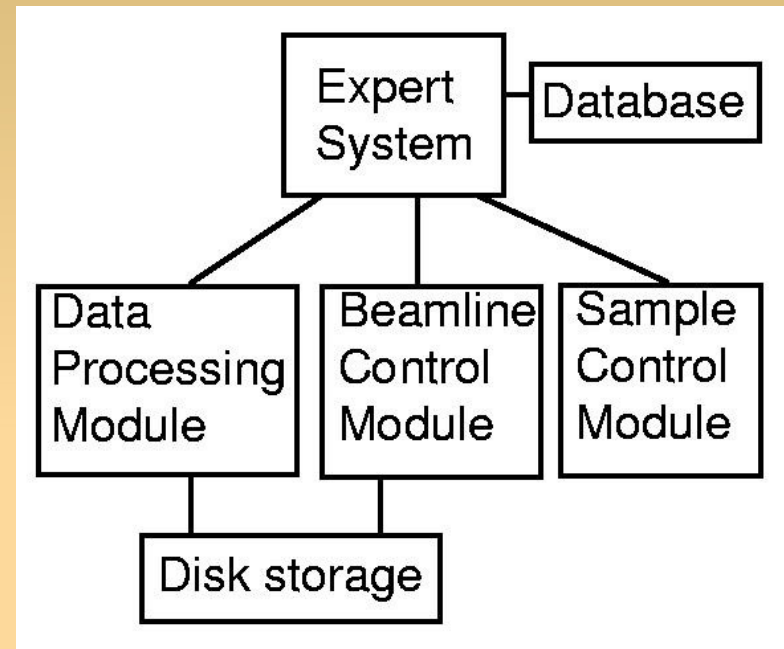


10 µm sample

12µm needle (zoom 10)

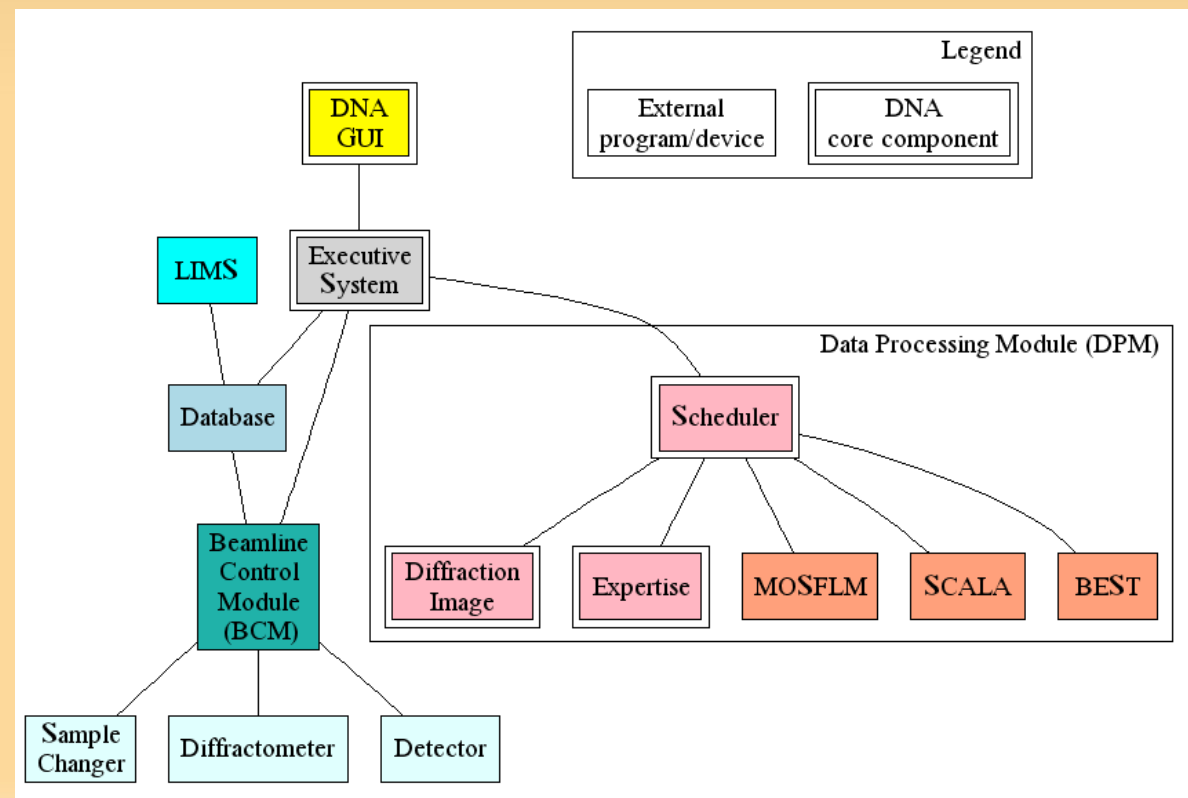
# Creation of the DNA collaboration

- Kick-off meeting in 2001
- Initial collaborators :
  - ESRF
  - Daresbury SRS
  - MRC LMB Cambridge
- Initially no external funding
- Meaning of "DNA" :
  - 1<sup>st</sup> version: "**DNA** is **Not** **Auto**struct" (like "GNU is Not Unix")
  - 2<sup>nd</sup> version: "automated**D** collection**N** of dat**A**"



# Evolution of the DNA collaboration

- Main development period 2001 – 2005
- More collaborators and more developers entered the project, mainly thanks to external fundings : BioXHIT and e-HTPX.
- Since 2005 an integral part of the ESRF "data collection pipeline"
- Installed and used at :
  - ESRF
  - Diamond
  - NSLS (Brookhaven)



# DNA collaborators in 2007

## DNA Collaborators

| Home institute                                | Name and link to email                     | DNA related work funded by         | DNA tasks  |
|---|--|------------------------------------|--|
| <a href="#">Brookhaven Nat'l Lab.</a>         | <a href="#">Alex Soares</a>                | <a href="#">BNL</a>                |  |
|   | <a href="#">Bob Sweet</a>                  |                                    | DNA co-ordinator at BNL  |
|   | <a href="#">John Skinner</a>               |                                    |  |
| <a href="#">Diamond Light Source</a>          | <a href="#">Alun Ashton</a>                | <a href="#">DLS</a>                | Acting DNA project co-ordinator  |
|   | <a href="#">Colin Nave</a>                 |                                    |  |
|   | <a href="#">Elizabeth Duke</a>             |                                    | DNA co-ordinator at DLS  |
|   | <a href="#">Karl Levik</a>                 |                                    | DNA developer  |
|   | <a href="#">Katherine McAuley</a>          |                                    | Testing and implementation of DNA at Diamond Light Source              |
| <a href="#">EMBL Grenoble</a>                 | <a href="#">Raimond Ravelli</a>            | <a href="#">EMBL Grenoble</a>      | DNA co-ordinator at the EMBL Grenoble                                  |
|   | <a href="#">Sandor Brockhauser</a>         | <a href="#">BioXHIT</a>            | Working on introducing kappa geometry strategy into DNA                |
| <a href="#">EMBL Hamburg</a>                  | <a href="#">Alexander Popov</a>            | <a href="#">BioXHIT</a>            | DNA co-ordinator at EMBL Hamburg                                       |
|   | <a href="#">Gleb Bourenkov</a>             | <a href="#">DESY</a>               |  |
|   | <a href="#">Venkataraman Parthasarathy</a> | <a href="#">SPINE</a>              | Integration of the BEST strategy software                              |
|   | <a href="#">Sean McSweeney</a>             | <a href="#">ESRF</a>               | DNA co-ordinator at the ESRF   |
|   | <a href="#">Darren Spruce</a>              |                                    | Responsible for the ESRF BCM (ProDC) and the DNA - LIMS connection     |
| <a href="#">ESRF</a>                          | <a href="#">Olof Svensson</a>              |                                    | Responsible for the DNA Executive System and DNA 2.0 Project Manager   |
|   | <a href="#">Marie Françoise Incardona</a>  | <a href="#">BioXHIT</a>            | Working on DNA 2.0   |
|   | <a href="#">Romeu Pieritz</a>              |                                    | Responsible for developing the ranking module and working on DNA 2.0   |
| <a href="#">Global Phasing</a>                | <a href="#">Gérard Bricogne</a>            | <a href="#">Global Phasing</a>     | DNA co-ordinator at Global Phasing                                     |
|   | <a href="#">Peter Keller</a>               | <a href="#">BioXHIT</a>            | Working on DNA 2.0   |
| <a href="#">MRC LMB Cambridge</a>             | <a href="#">Andrew Leslie</a>              | <a href="#">MRC</a>                | DNA co-ordinator at Cambridge  |
|   | <a href="#">Harry Powell</a>               | <a href="#">CCP4</a>               | Responsible for the DNA DPM based on MOSFLM                            |
| <a href="#">MRC France - BM14 at the ESRF</a> | <a href="#">Ludovic L auner</a>            | <a href="#">e-htpx</a>             | Working together with Darren on the DNA - beamline database connection |
| <a href="#">STFC Daresbury</a>                | <a href="#">Graeme Winter</a>              | <a href="#">e-htpx</a>             | Responsible for the DNA Scheduler                                      |
| <a href="#">SLS - PSI</a>                     | <a href="#">Takashi Tomizaki</a>           | <a href="#">SLS - PSI</a>          | DNA co-ordinator at SLS  |
| <a href="#">Synchrotron Soleil</a>            | <a href="#">Andrew Thompson</a>            | <a href="#">Synchrotron Soleil</a> | DNA co-ordinator at Soleil   |
|   | <a href="#">Lucile Roussier</a>            |                                    | Working on DNA - database connection for Soleil                        |
|   | <a href="#">Eric Girard</a>                |                                    | Working on offline tests of DNA  |
|   | <a href="#">Pierre Legrand</a>             |                                    | Working on integrating XDS as a DPM in the DNA system                  |

# Problems faced in 2005

- The new developers who entered the project found it to be hard if not impossible to contribute to the existing code
- The knowledge of the DNA system was (and still is...) shared by only two developers
- Also for the original developers it became more and more costly to implement new features because of the software structure, i.e. it is very hard to change the DNA workflow
- It became more and more difficult to agree in which direction to go with the project
- Therefore a proposal was made by five DNA developers in 2005 for re-develop a new system from zero



# Start of DNA 2.0 → EDNA

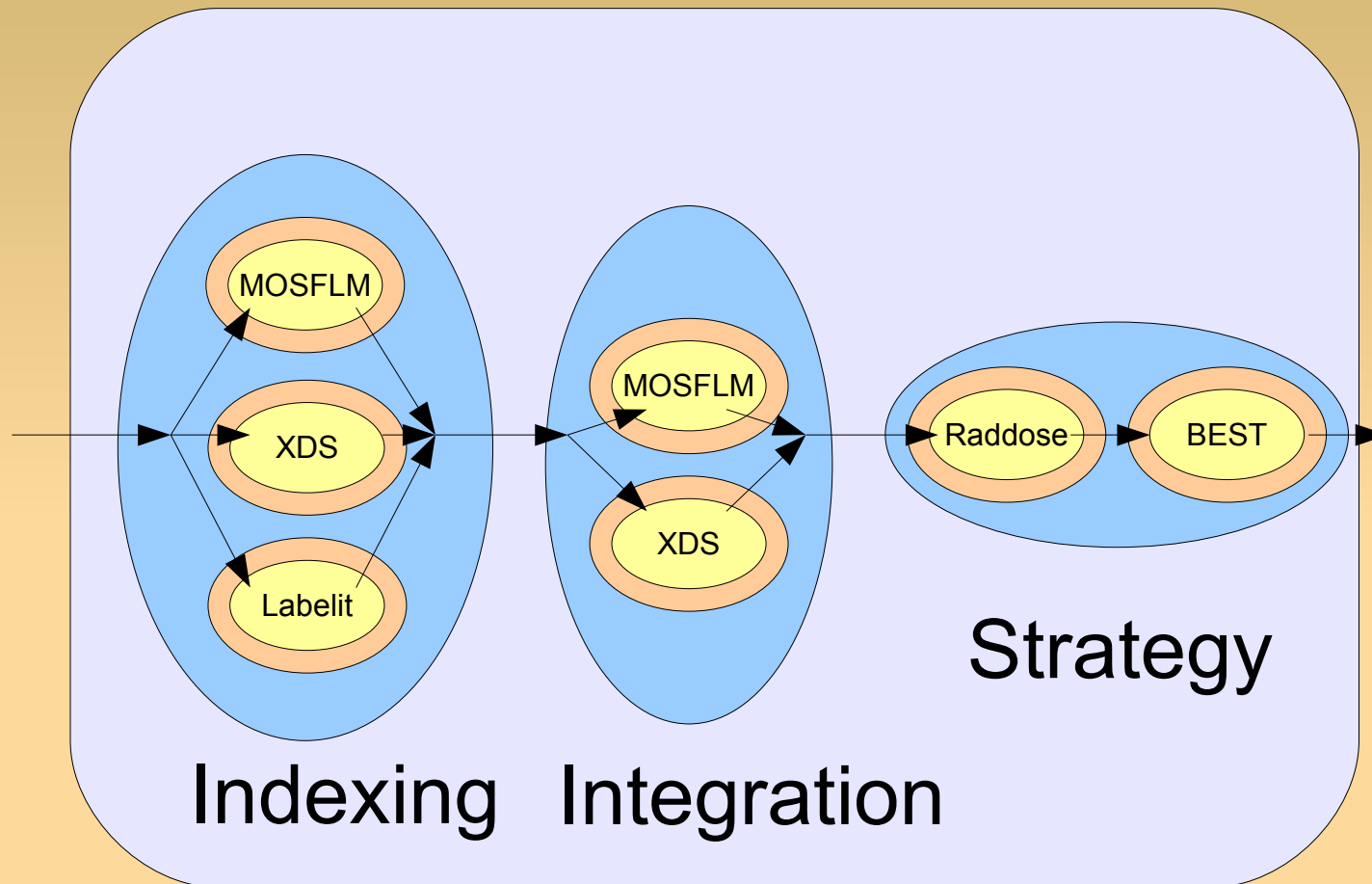
- The proposition for a new project was accepted by the DNA executive committee in Autumn 2005. The initial name was DNA 2.0.
- In 2006 BioXHIT funding for a DNA 2.0 project manager was obtained.
- The post was announced in Autumn 2006 and was assigned to me in January 2007. The money set aside for the project manager was used for hiring Marie-Françoise as a DNA 2.0 Technical Manager.
- The DNA 2.0 project started officially at a DNA meeting held at Diamond in February 2006.
- The name EDNA was adopted in the Project Agreement meeting at the ESRF in October 2007.



# EDNA – Current status

- Release of the prototype to EDNA members in July 2008
  - DNA characterisation (indexing, integration and strategy)
  - No limitation on the number of reference images (in DNA = 2)
  - Strategy calculation taking into radiation damage (RADDPOSE)
- Currently being tested at the ESRF, Diamond, NSLS, Soleil and EMBL Hamburg
- Second EDNA developers workshop October 22<sup>nd</sup> -24<sup>th</sup> 2008:
  - 14 participants from 10 institutes (ESRF, Diamond, EMBL Grenoble, EMBL Hamburg, Soleil, SLS, BESSY, NSLS, MAX-LAB, Global Phasing)
  - Workshop topic : "How to develop an EDNA plugin"

# EDNA Characterisation Module



Generic Data Model

Specific Data Model

Translation layer

# EDNA Project Management

- From DNA :
  - EDNA executive committee
  - Project coordinator : Alun Ashton
- Improved in comparison with DNA :
  - Project manager
  - Clearly defined project goals
  - Project agreement (in preparation) with licence (GPL 3)
  - Modular software structure :
    - EDNA Framework kernel
    - "Kernel" plugins
    - Other plugins

# EDNA Development Best Practices

- EDNA testing framework
  - Kernel unit tests
  - Plugin unit tests
  - Plugin execution tests
- Code review
  - Every code change except trivial bug fixes reviewed
- Coding conventions
  - Assure that code written by different developers can be read by other developers
- Wiki documentation
  - Currently being developed
  - Should evolve with the code

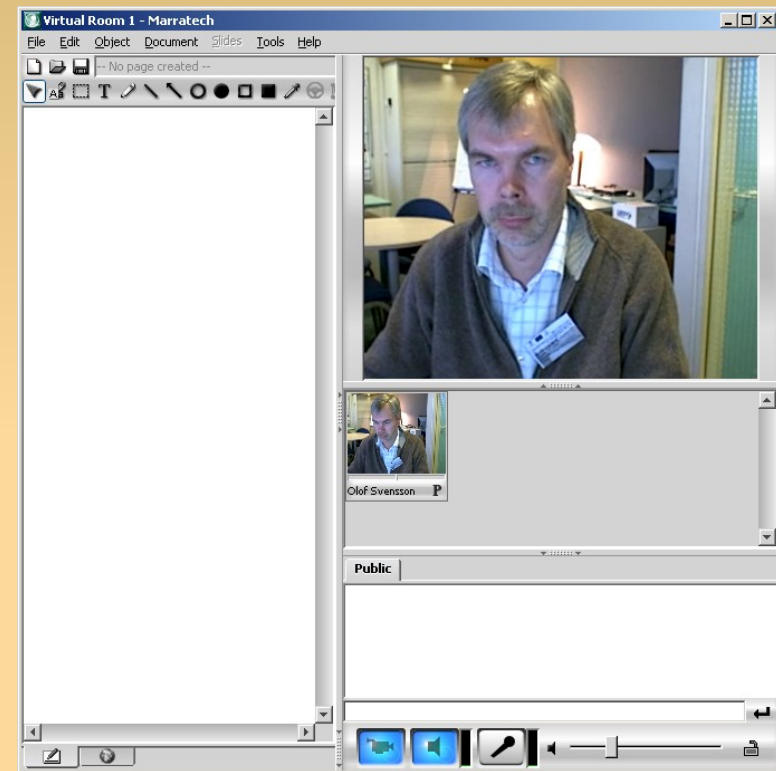
# Example of EDNA Plugin Execution Tests

```
Total Cumulated Tests: [ 16 ]  
Total Cumulated Test SUCCESS: [ 16 ]  
Total Cumulated Test FAILLURE: [ 0 ]  
Total Cumulated Test Methods: [ 0 ]
```

```
[SUCCESS] [ 1 ] [ EDTestCasePluginExecuteMOSFLMv01Indexing.execute ] [15.8376750946]  
[SUCCESS] [ 2 ] [ EDTestCasePluginExecuteMOSFLMv01IndexingWithForcedSymmetry.execute ] [11.7330498695]  
[SUCCESS] [ 3 ] [ EDTestCasePluginExecuteMOSFLMv01Integration.execute ] [5.33224701881]  
[SUCCESS] [ 4 ] [ EDTestCasePluginExecuteMOSFLMv01GeneratePrediction.execute ] [5.28757691383]  
[SUCCESS] [ 5 ] [ EDTestCasePluginExecuteRaddosev01.execute ] [4.47688508034]  
[SUCCESS] [ 6 ] [ EDTestCasePluginExecuteBestv01.execute ] [8.4916009903]  
[SUCCESS] [ 7 ] [ EDTestCasePluginExecuteStrategyv01.execute ] [6.90980195999]  
[SUCCESS] [ 8 ] [ EDTestCasePluginExecuteStrategyv01WithoutChemComposition.execute ] [6.02249383926]  
[SUCCESS] [ 9 ] [ EDTestCasePluginExecuteControlIndexingv01.execute ] [15.7771389484]  
[SUCCESS] [ 10 ] [ EDTestCasePluginExecuteControlIndexingv01withForcedSymmetry.execute ] [17.3605411053]  
[SUCCESS] [ 11 ] [ EDTestCasePluginExecuteControlIntegrationv01.execute ] [7.62139487267]  
[SUCCESS] [ 12 ] [ EDTestCasePluginExecuteControlCharacterisationv01.execute ] [26.129956007]  
[SUCCESS] [ 13 ] [ EDTestCasePluginExecuteControlCharacterisationv01withForcedSpaceGroup.execute ] [26.9026]  
[SUCCESS] [ 14 ] [ EDTestCasePluginExecuteControlGeneratePredictionv01.execute ] [7.64759492874]  
[SUCCESS] [ 15 ] [ EDTestCasePluginExecuteReadImageHeaderv01.execute ] [3.60145902634]  
[SUCCESS] [ 16 ] [ EDTestCasePluginExecuteControlCharacterisationv01With2Sweep.execute ] [45.3589470387]
```

# EDNA Collaboration Tools

- Video conferences :  
Marratech
- Mailing lists
- Project dedicated server :
  - Joomla
  - Wiki
  - Discussion forum
- Development tools :
  - Eclipse IDE
  - Bugzilla
  - Subversion



# Conclusion

- The DNA collaboration was successful in proving the usefulness of an automated data collection system, regularly used at the ESRF, Diamond and the NSLS
- The lessons learned from the DNA collaboration were taken into account when setting up the EDNA collaboration
- Thanks to the EDNA management structure and the modular software structure (plugins) the collaboration is now about to take off
- Comparison with EVUO collaboration :
  - Automation → less pressure to provide fast a working solution
  - Developers sometimes not software engineers
  - Sometimes head of developers not part of collaboration



# Acknowledgements – EDNA members

Alexander Popov<sup>(d)</sup>

Alun Ashton<sup>(e)</sup>

Andrew Leslie<sup>(h)</sup>

Andrew McCarthy<sup>(b)</sup>

Andrew Thompson<sup>(k)</sup>

Clemens Schulze<sup>(j)</sup>

Clemens Vornrhein<sup>(f)</sup>

Darren Spruce<sup>(d)</sup>

Elsbeth Gordon<sup>(d)</sup>

Ezequiel Panepucci<sup>(j)</sup>

G rard Bricogne<sup>(f)</sup>

**Gerrit Langer**<sup>(b)</sup>

**Gleb Bourenkov**<sup>(b)</sup>

Gordon Leonard<sup>(d)</sup>

Harry Powell<sup>(h)</sup>

Johan Turkenburg<sup>(m)</sup>

**Johan Unge**<sup>(g)</sup>

**John Skinner**<sup>(i)</sup>

**Karl Levik**<sup>(e)</sup>

Katherine McAuley<sup>(e)</sup>

Lucile Roussier<sup>(k)</sup>

**M.-F. Incardona**<sup>(d)</sup>

**Mark Basham**<sup>(e)</sup>

Meitian Wang<sup>(j)</sup>

**Michael Hellmig**<sup>(a)</sup>

Olga Roudenko<sup>(k)</sup>

**Olof Svensson**<sup>(d)</sup>

**Peter Keller**<sup>(f)</sup>

Peter Turner<sup>(l)</sup>

**Pierre Legrand**<sup>(k)</sup>

Robert Sweet<sup>(i)</sup>

**Romeu Pieritz**<sup>(d)</sup>

**Sandor Brockhauser**<sup>(b)</sup>

Sean McSweeney<sup>(d)</sup>

**Takashi Tomizaki**<sup>(j)</sup>

Thomas Schneider<sup>(b)</sup>

Uwe Mueller<sup>(a)</sup>

(a) BESSY, Berlin, Germany

(b) EMBL, Grenoble, France

(c) EMBL, Hamburg, Germany

(d) ESRF, Grenoble, France

(e) Diamond Light Source, UK

(f) Global Phasing, Cambridge, UK

(g) MAX LAB, Lund, Sweden

(h) MRC LMB, Cambridge, UK

(i) NSLS, Brookhaven, U.S.

(j) SLS, Villigen, Switzerland

(k) Synchrotron Soleil, France

(l) University of Sydney, Australia

(m) University of York, UK

**Names in bold:** Participants in the next developers' workshop

Underlined names : Members of the EDNA executive committee