Argonne NATIONAL LABORATORY

… for a brighter future

U.S. Department of Energy

THE UNIVERSITY OF CHICAGO

Office of Science U.S. DEPARTMENT OF ENERGY

A U.S. Department of Energy laboratory managed by The University of Chicago

The Scientific Software Initiative at the APS

Kenneth Evans, Jr.

Presented at the EPICS Collaboration Meeting

October 13 - 14, 2007

ICALEPCS Meeting, Knoxville, TN

Current X-Ray Scientific Software Situation

- Best described as “Uncoordinated”
- Wide variety of languages
  - FORTRAN, C, C++, Perl, Tcl/Tk, Python, Java, …
- Visualization relies on (different) commercial products
  - IDL, IGOR, Matlab, …
- Each beamline tends to do its own thing
- Modeling and Analysis is not well integrated with Data Acquisition
- Lack of real-time data reduction
- Little high-performance computing
- Little remote access
- No common data format

- A Scientific Software Section was formed to help remedy this situation

Scientific Software Section

- Specific goals:
  - Combine existing analysis and visualization codes with beamline data acquisition software and transform these codes into easy-to-use software
  - Provide a scientific workbench program that is easy to use and learn and from which users can access all the software that is necessary to manage the entire scientific work flow
  - Create new analysis and visualization applications that can be used on all beamlines and that are easily integrated into the standard workbench
  - Develop a software framework, perhaps more than one, that provides tested and debugged scientific routines, such as fitting and visualization, which can be used by developers to create applications
  - Provide high-performance computing
  - Provide documentation, distribution, maintenance, and support

Why Java ?

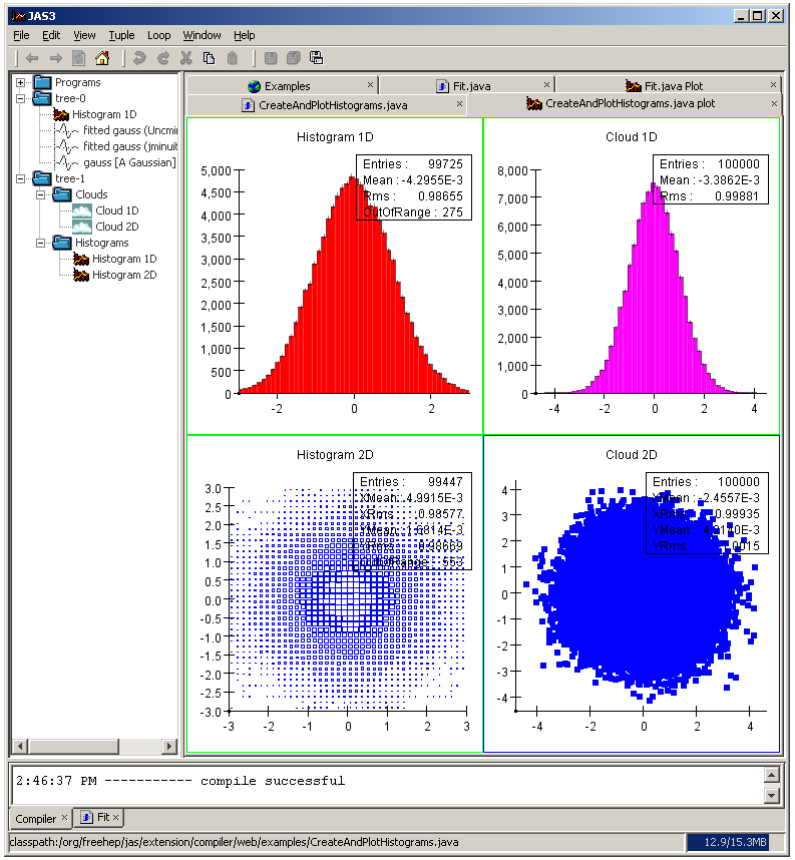
- Java has become a major and important language
- The reason is that most commercial development uses J2EE
  - It is running the financial world
  - There is money to be made improving Java and its tools
- Applications have performance approaching applications written in C
- There is already extensive scientific development in Java
- Advantages
  - Huge API
  - Write once, run anywhere
  - Easy to code (compared to C or C++, anyway)
  - Good performance
  - Excellent development tools

Java Development Tools

- Spell checks as you go
  - Eliminates most typos and all syntax errors
- Provides content assist
  - It tells you what to type next and provides documentation, etc.
- Compiles as you write
  - No “write – compile – load – run – figure out what happened” cycle
  - Cycle is now “write – run”
- Massive refactoring
  - E.g. Change a variable name in all your files in all your projects
- Wizards and Tools to help at every stage
  - E.g. Generate getters and setters for all your properties
- The above are just a small sample
  - Some of these are available for other languages
  - But generally not at the level they are for Java

Java Analysis Studio (JAS3)

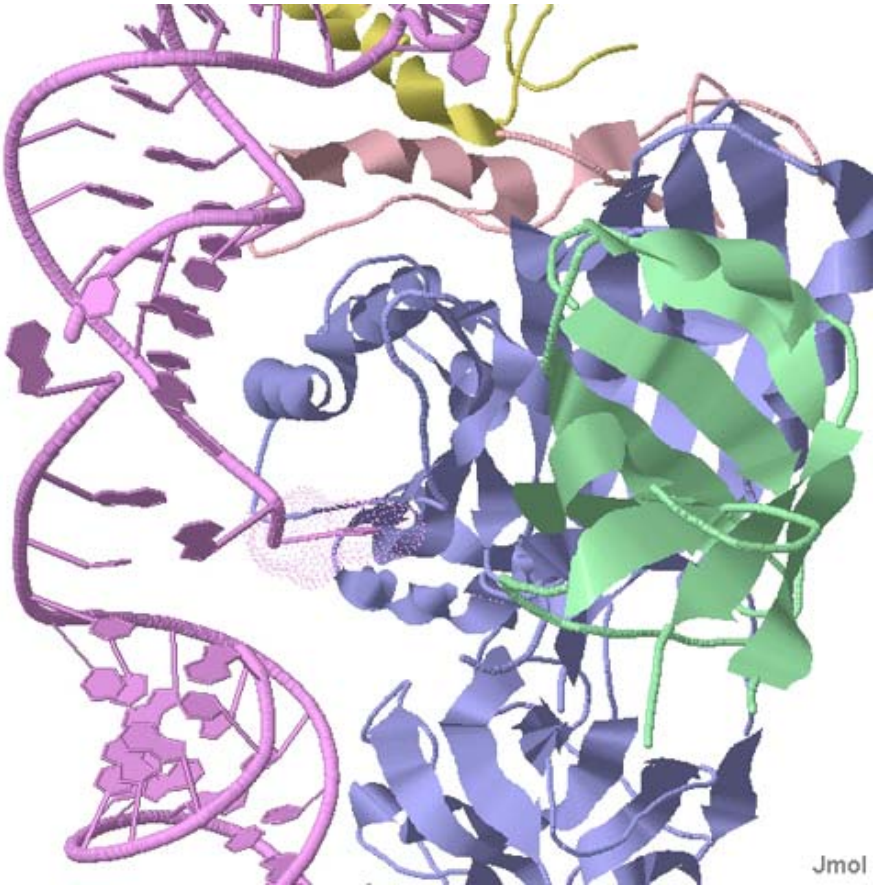
- Developed by and for the High-Energy physics community
- Plotting of 1d, 2d, 3d Histograms, XY plots, Scatter plots, *etc*.
- Open source
- Attractive plotting
- Fitting, other mathematical analysis
  - Primarily from CERN
- Highly modular structure
  - Uses plug-ins



JMol – Molecular Viewer

- Commonly used as an applet that can be integrated into web pages to display molecules in a variety of ways
- Also has a standalone application and a development tool kit that can be integrated into other Java applications
- Interactive, 3D
- Free, Open Source

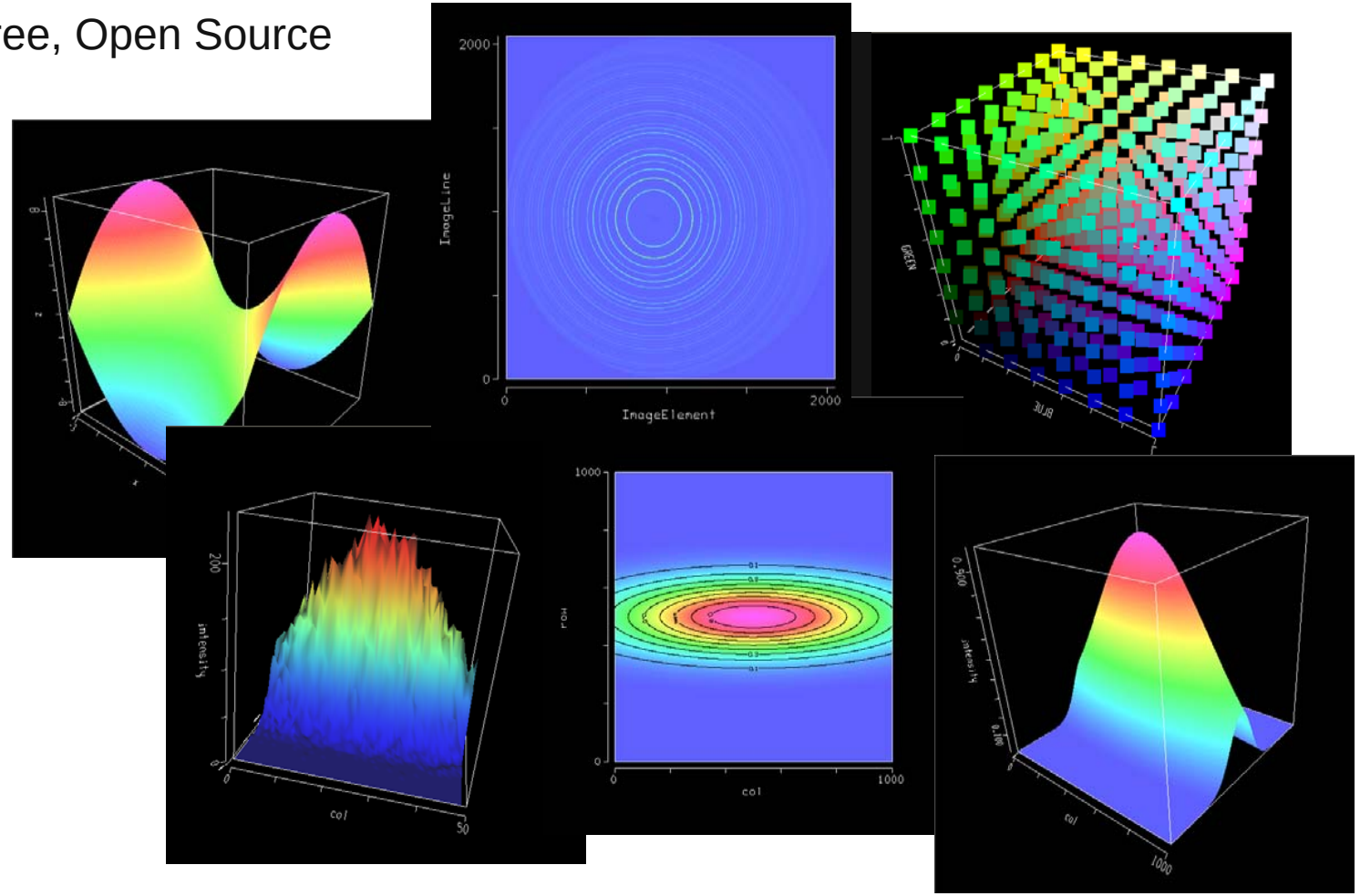
- One of several Java Molecular Graphics packages



Crystal structure of an H/ACA box RNP from Pyrococcus furiosus (PDB CODE: 2HVY)

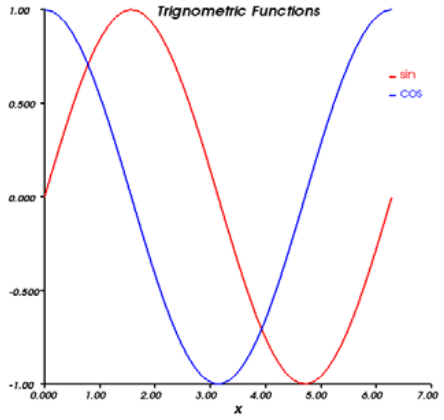
VisAD

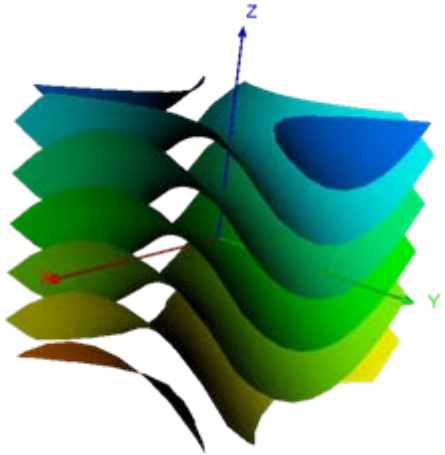
- Space Sciences and Engineering Center (SSEC), and others
- Extensive 2D and 3D visualization package
- Free, Open Source

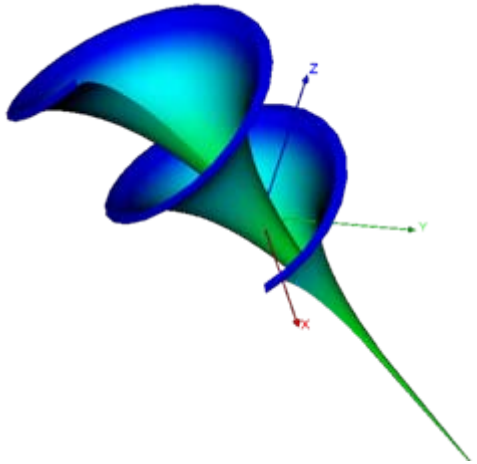


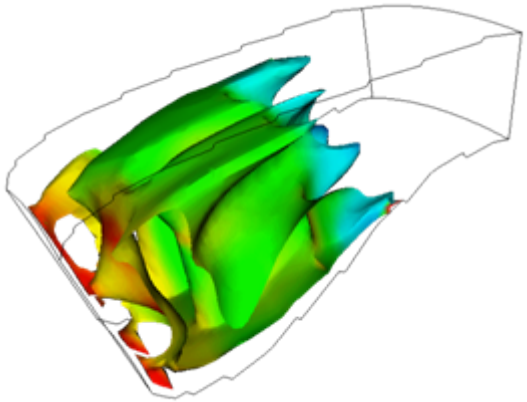
VTK

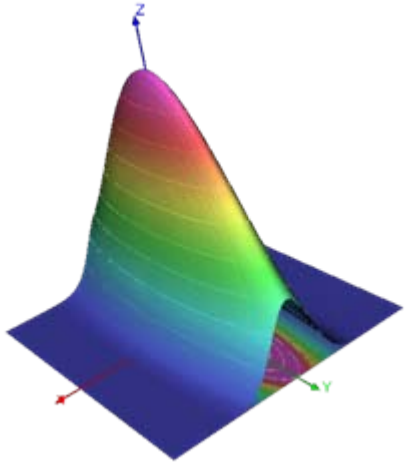
- Software system for 3D computer graphics, image processing, and visualization
- Used by thousands of researchers and developers around the world
- Written in C++
- Has Java wrappers
  - Also, Tcl/Tk, Python
- Free, Open Source





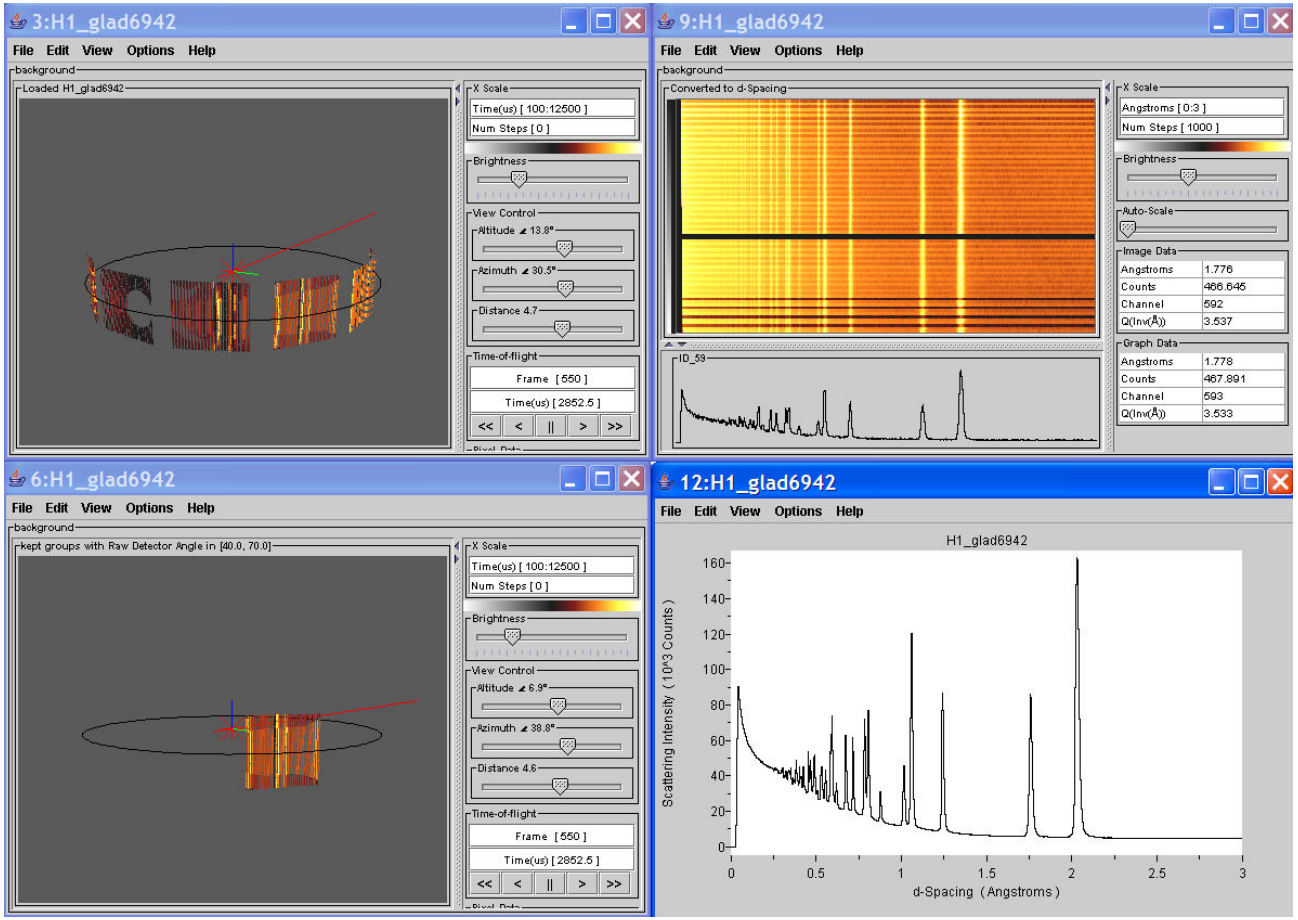






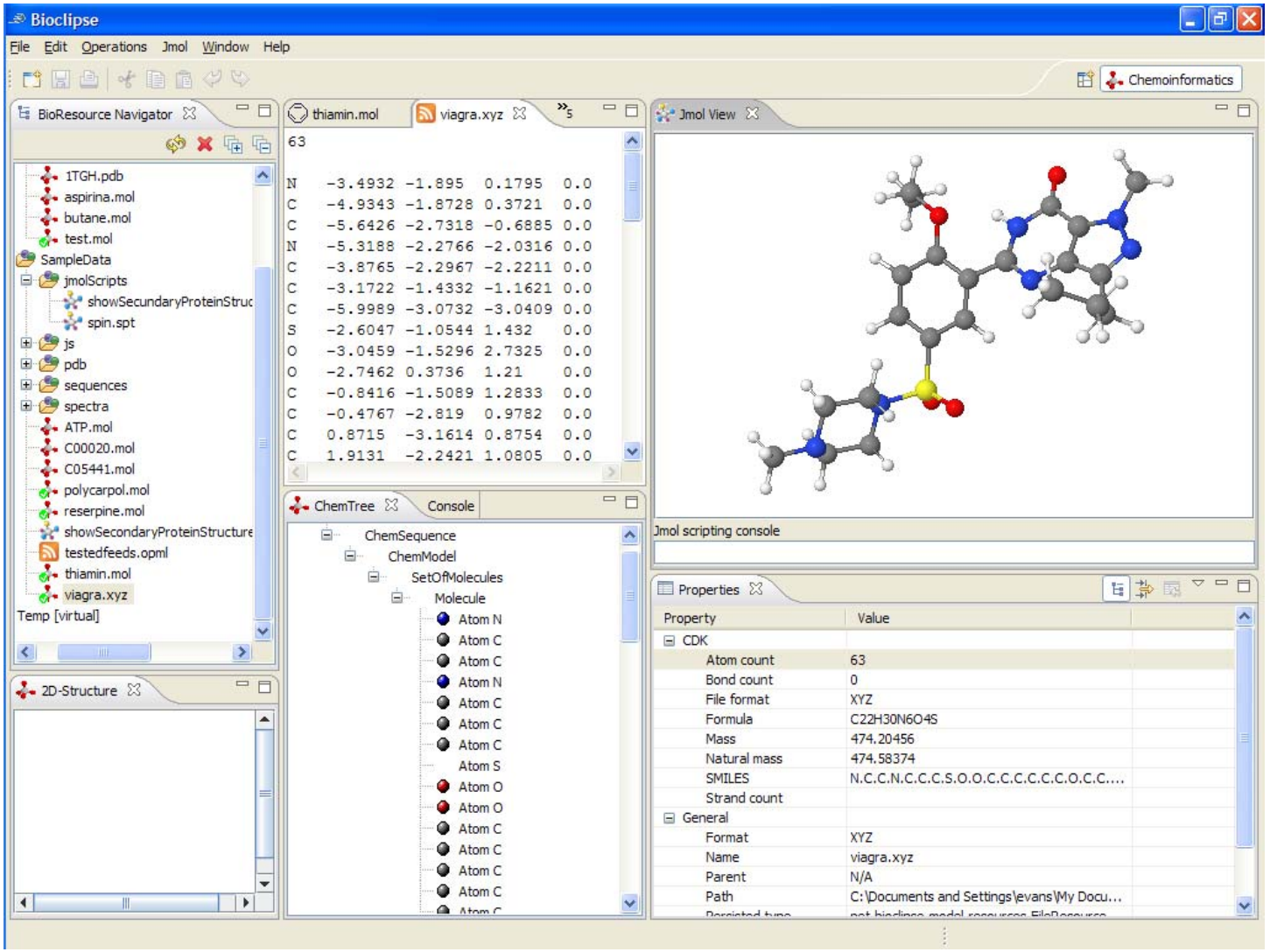
ISAW

- The primary tool for analyzing neutron scattering data at the IPNS
- Has an extensive and sophisticated interface

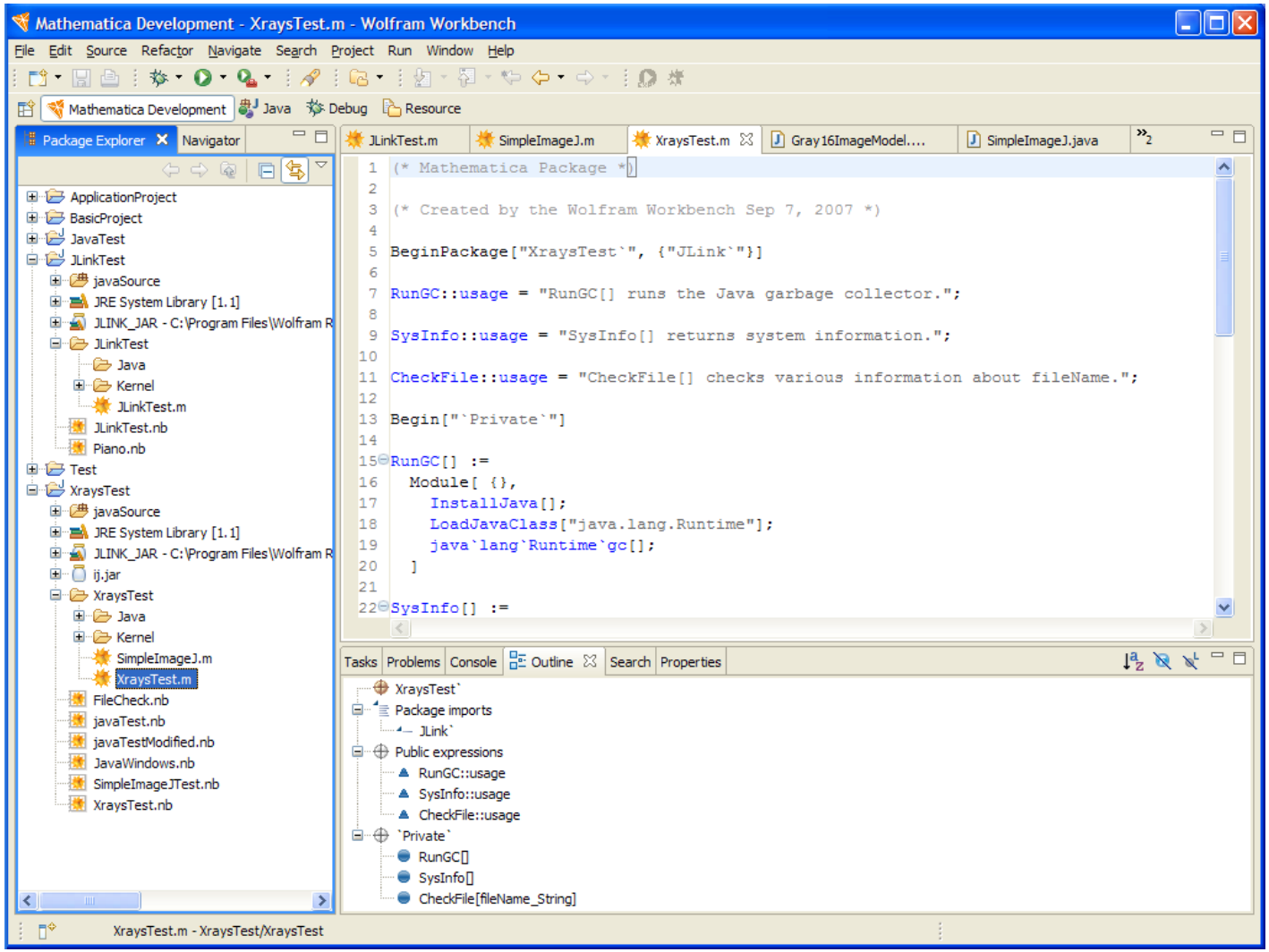


From: John Hammonds, IPNS

Bioclipse

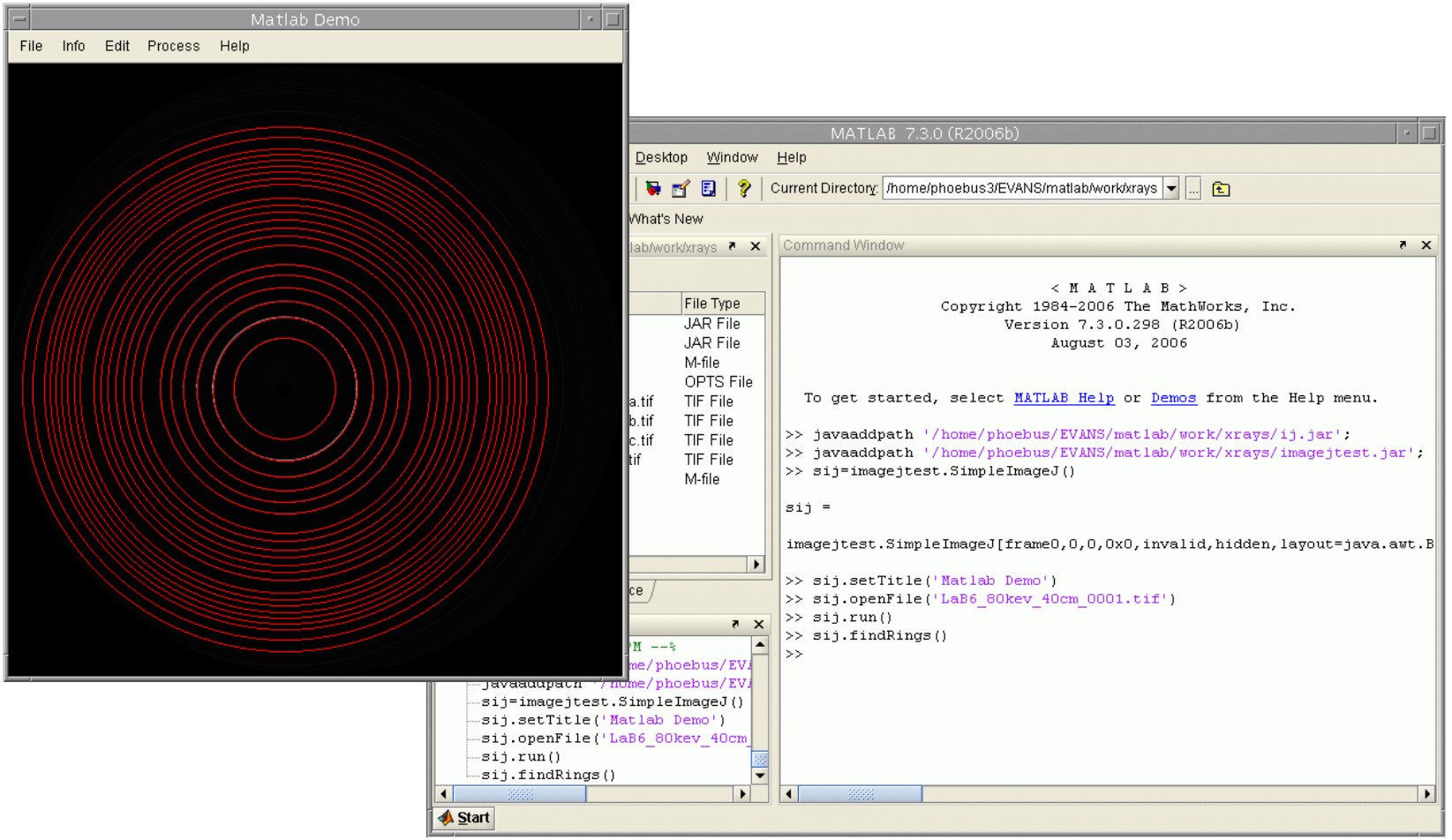

Bioclipse
File Edit Operations Jmol Window Help
Chemoinformatics
BioResource Navigator
1TGH.pdb
aspirina.mol
butane.mol
test.mol
SampleData
jmolScripts
showSecundaryProteinStruc
spin.spt
js
pdb
sequences
spectra
ATP.mol
C00020.mol
C05441.mol
polycarpol.mol
reserpine.mol
showSecondaryProteinStructure
testedfeeds.opml
thiamin.mol
viagra.xyz
Temp [virtual]
thiamin.mol
viagra.xyz
63
N -3.4932 -1.895 0.1795 0.0
C -4.9343 -1.8728 0.3721 0.0
C -5.6426 -2.7318 -0.6885 0.0
N -5.3188 -2.2766 -2.0316 0.0
C -3.8765 -2.2967 -2.2211 0.0
C -3.1722 -1.4332 -1.1621 0.0
C -5.9989 -3.0732 -3.0409 0.0
S -2.6047 -1.0544 1.432 0.0
O -3.0459 -1.5296 2.7325 0.0
O -2.7462 0.3736 1.21 0.0
C -0.8416 -1.5089 1.2833 0.0
C -0.4767 -2.819 0.9782 0.0
C 0.8715 -3.1614 0.8754 0.0
C 1.9131 -2.2421 1.0805 0.0
Jmol View
Jmol scripting console
ChemTree
Console
ChemSequence
ChemModel
SetOfMolecules
Molecule
Atom N
Atom C
Atom C
Atom N
Atom C
Atom C
Atom C
Atom S
Atom O
Atom O
Atom C
Atom C
Atom C
Atom C
2D-Structure
Properties
Property
Value
CDK
Atom count 63
Bond count 0
File format XYZ
Formula C22H30N6O4S
Mass 474.20456
Natural mass 474.58374
SMILES N.C.C.N.C.C.C.S.O.O.C.C.C.C.C.C.O.C.C....
Strand count
General
Format XYZ
Name viagra.xyz
Parent N/A
Path C:\Documents and Settings\evans\My Docu...

Wolfram Workbench (Mathematica)



Java in Matlab (or IDL or Mathematica or…)

- Matlab has extensive support for Java
  - A scientific software framework can also be used in these apps



Why Eclipse?

- Powerful, flexible, extensible
- Open-source
- Has a huge community with many projects
- Java development environment leads to high productivity
- Deployment via plug-ins appears to solve many problems

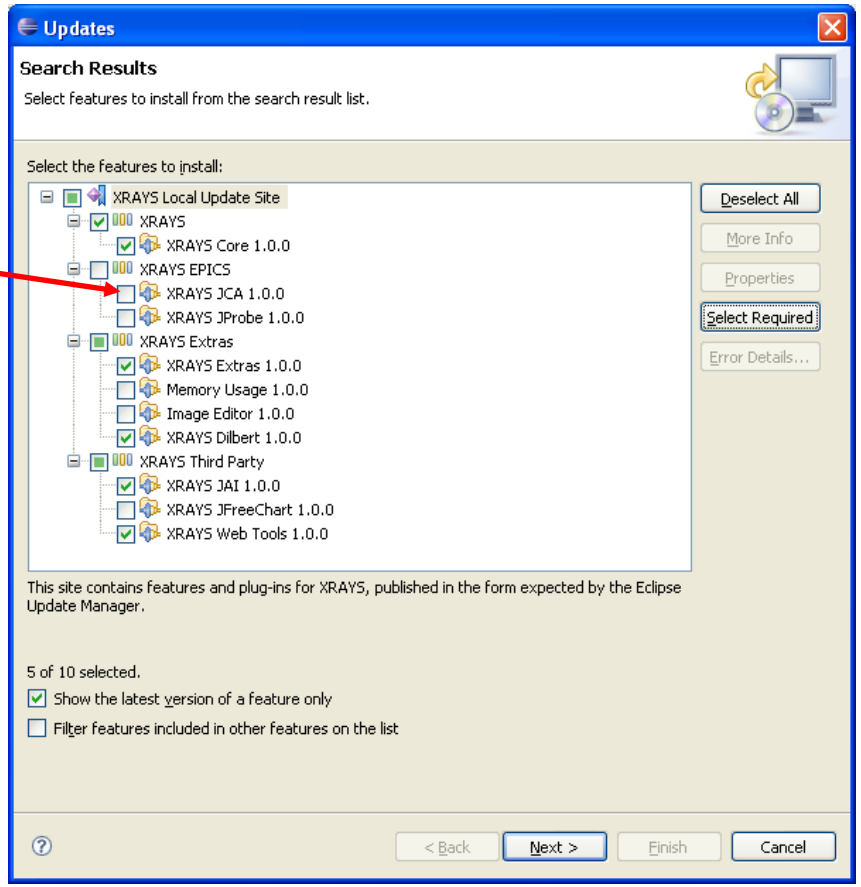
- We intend to use Eclipse, not as an IDE, but as a workbench (IE)
  - Something users will use
  - A way to organize their work

- Downsides
  - Most x-ray beamline staff and users are not using Eclipse now
  - 95% will be unhappy [with anything we do]

Deployment is a Major Reason for Using Eclipse

- Both Java and Eclipse are multi-platform
- Updates are easily made through the Eclipse update mechanism
- You can wrap 3rd party applications in your own plug-ins
  - For example:

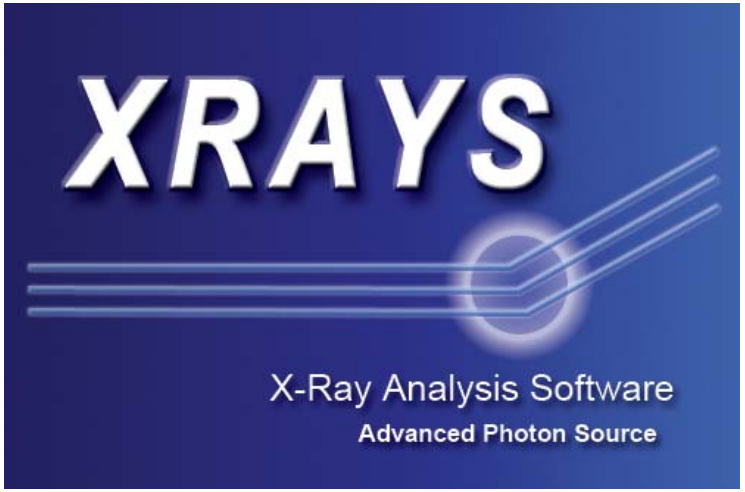
    The Feature “XRAYS JCA” contains org.scatteringsw.xrays.jca which wraps JCA
  - Including DLLs and Shared Objects
- Guarantees they are versions that work with your applications on all platforms you support
- Makes it easy for the user to install and update both your stuff and the 3rd party stuff



Eclipse for Users, not Developers

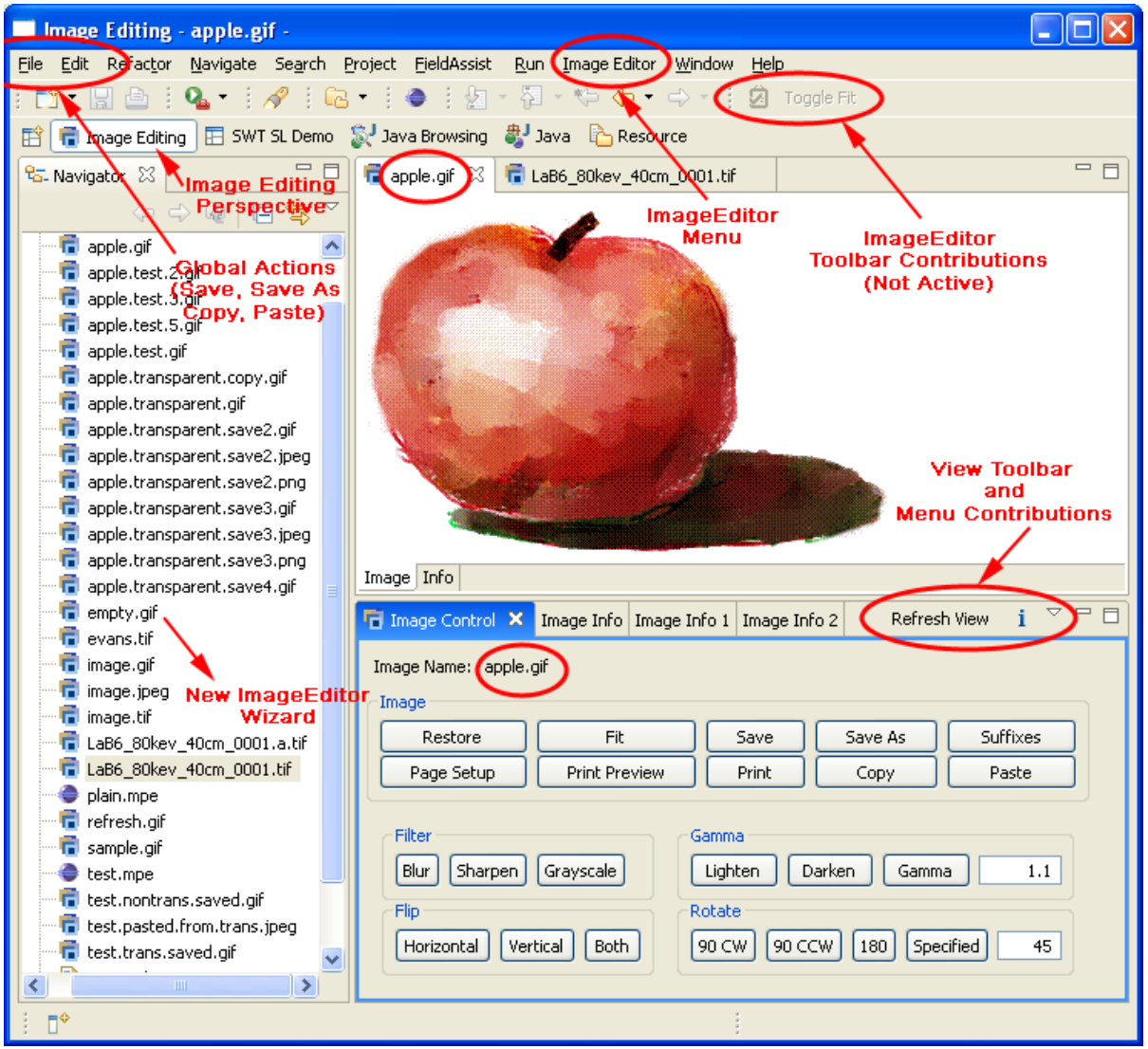
- We intend to use Eclipse as a workbench
- Something a user can come in and be up and running with in a short time
  - Probably with community help
- Each user can use and customize it in his or her own way
- They will probably use it for more than one thing
  - That is why the layout by Perspective is important
  - You just switch perspectives to change tasks
- I think this paradigm is better than using RCP applications (such as CSS)
  - Especially for our less-controlled environment
  - You provide the plug-ins
  - The user manages his Workbench as he or she pleases

XRAYS



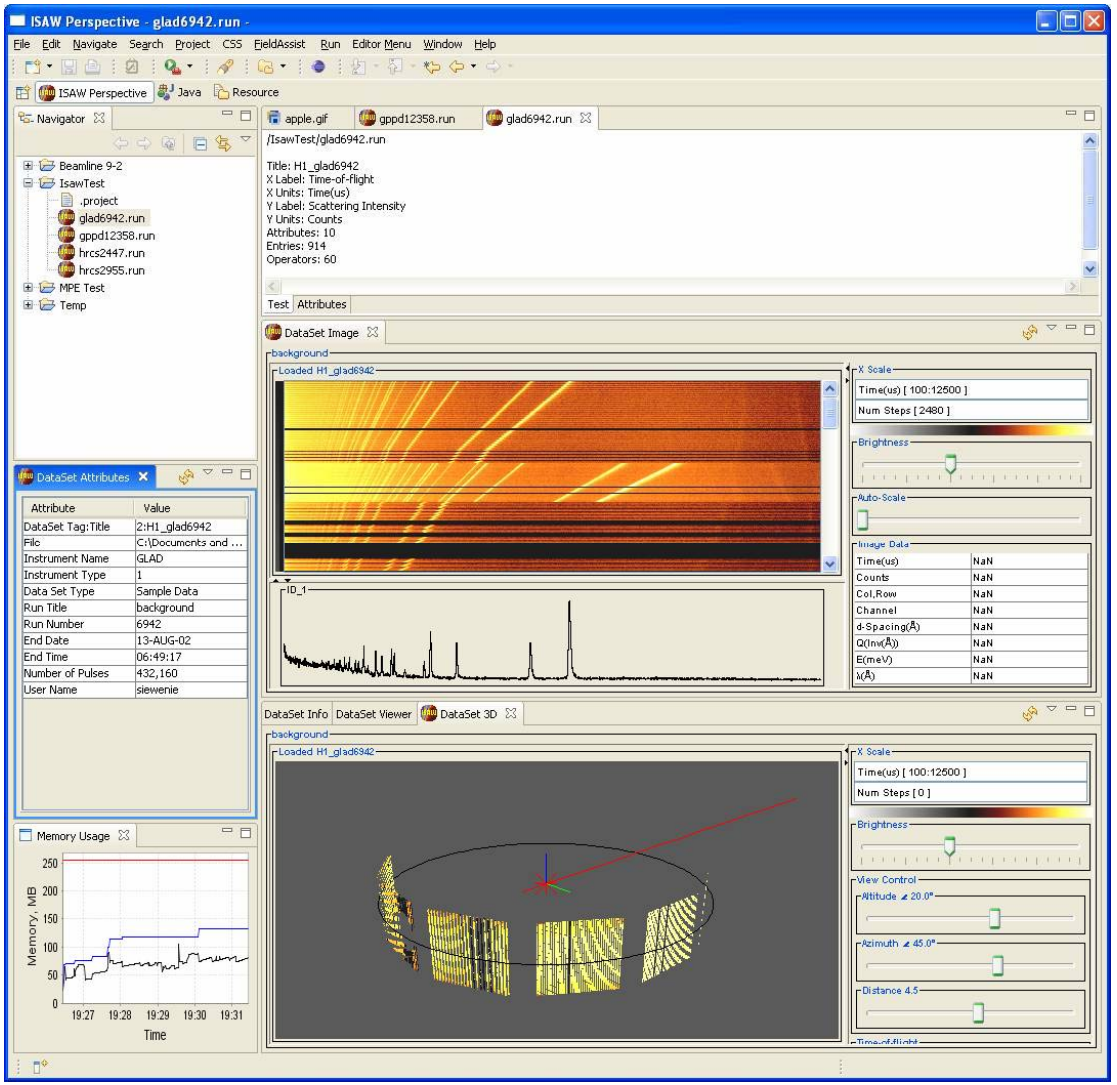
- Stands for X-Ray Analysis Software
  - (or X-Ray Software)
- It is expected to grow into a large suite of analysis and visualization applications
- These will include:
  - Scientific workbench program
  - New analysis and visualization applications
  - Updating and coordination of existing analysis and visualization applications
  - A framework of software routines that developers can use to write applications
- It currently consists mostly of exploration and prototype applications
  - This is the groundwork for what we really want to do
  - Currently more than 3000 Java source files in 120 projects
  - 38 Java projects intended for distribution (org.scatteringsw.xrays.xxx)
  - 10 ready-to-deploy features (collections of plug-ins) in 4 categories

Image Editor as an Experiment Prototype

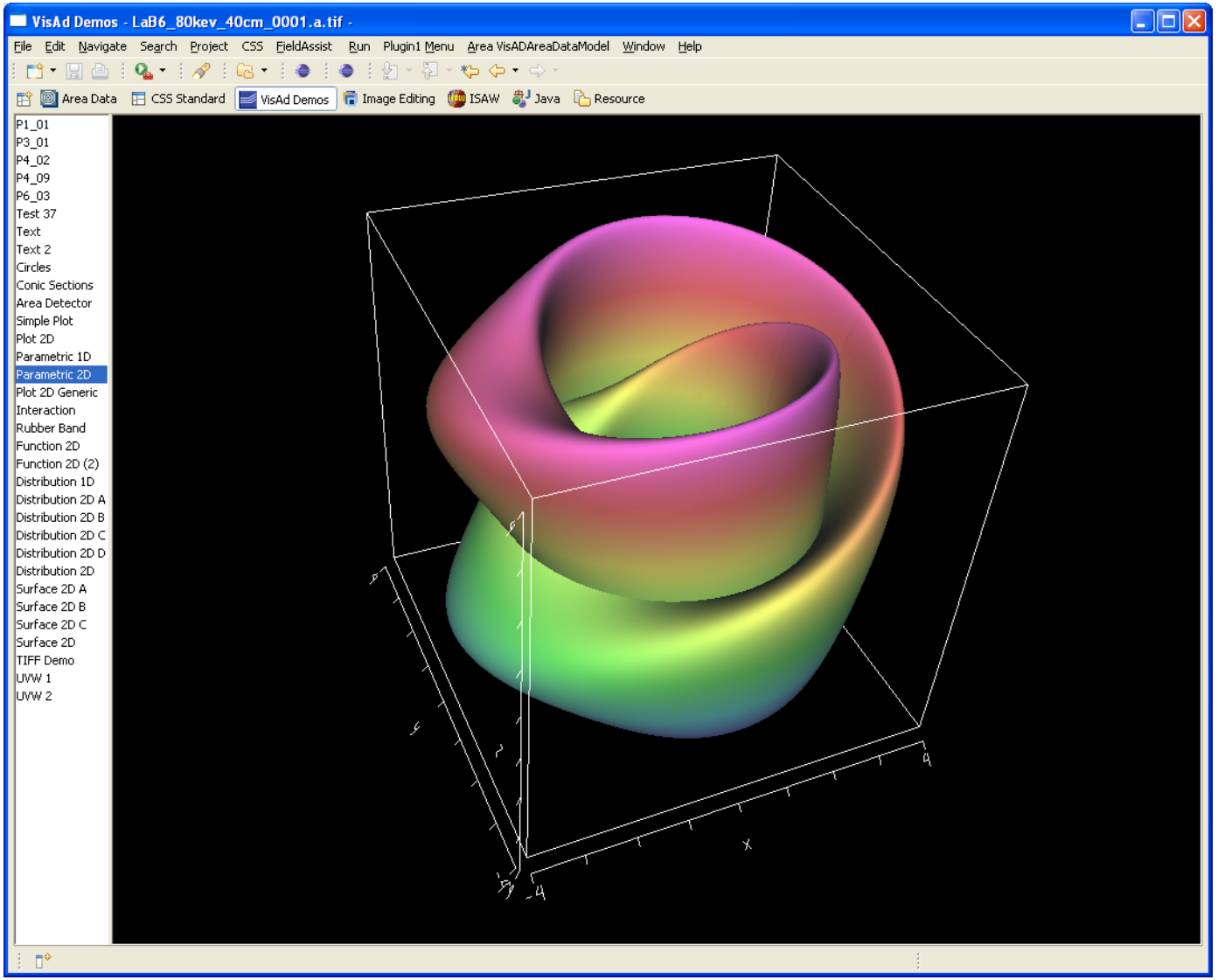


Prototype Implementation of ISAW

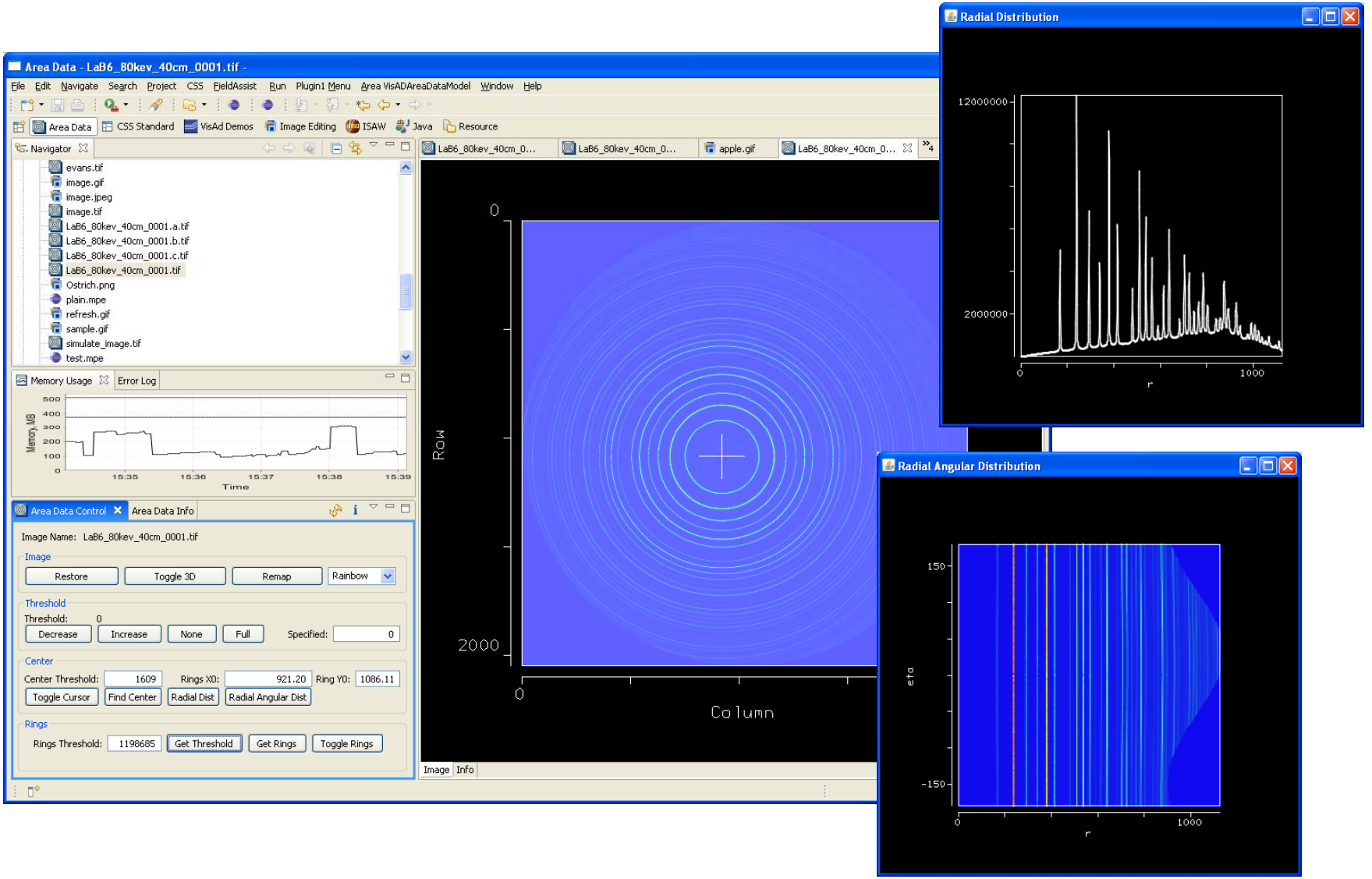
- Includes:
  - A Perspective
  - An Editor for ISAW DataSets
    - *.run, .isd*
  - Some Views
- All work together
  - Views change when the edited file changes



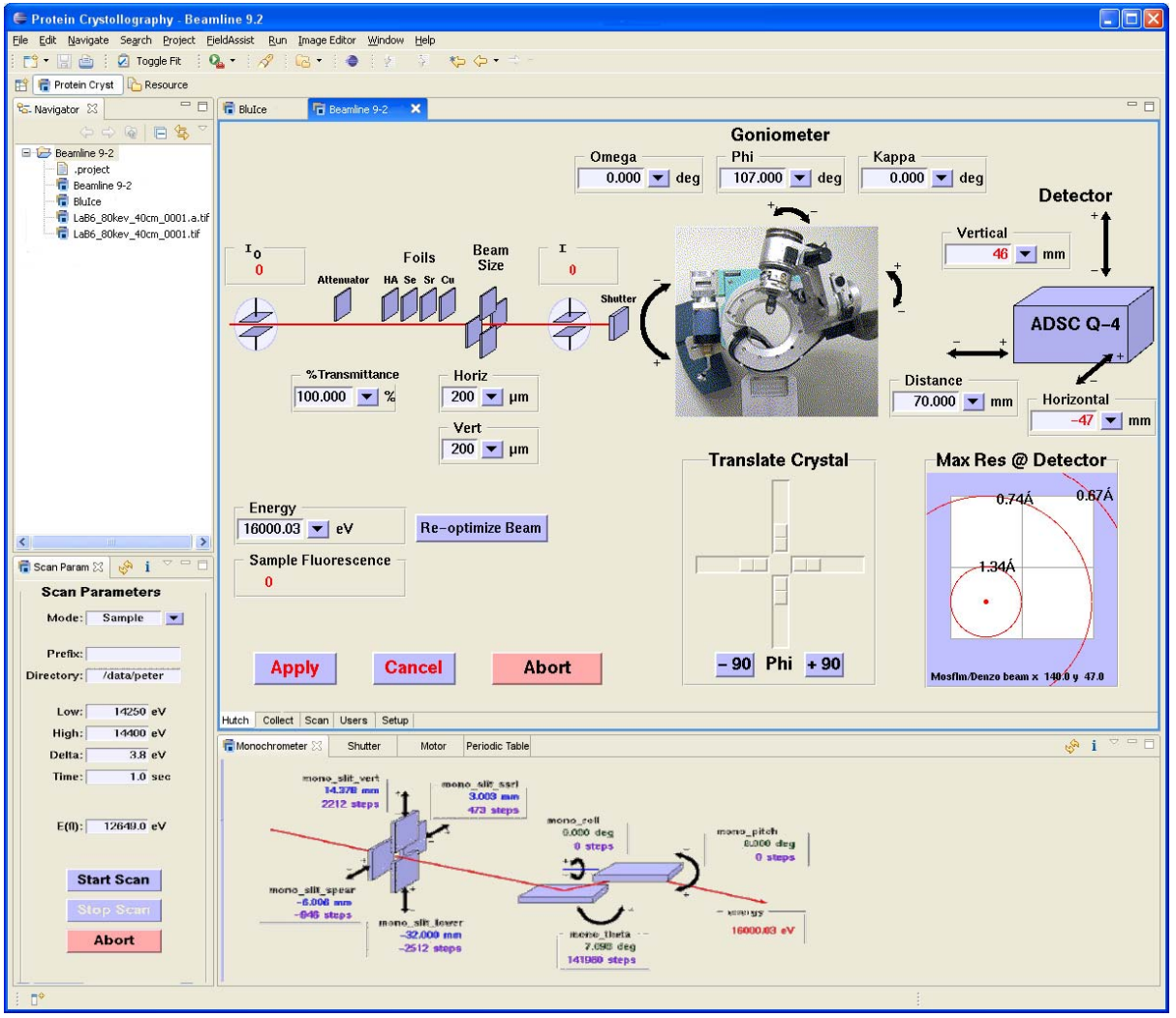
A Perspective Can be a Single Application



Area Data Editor



X-Ray Experiment

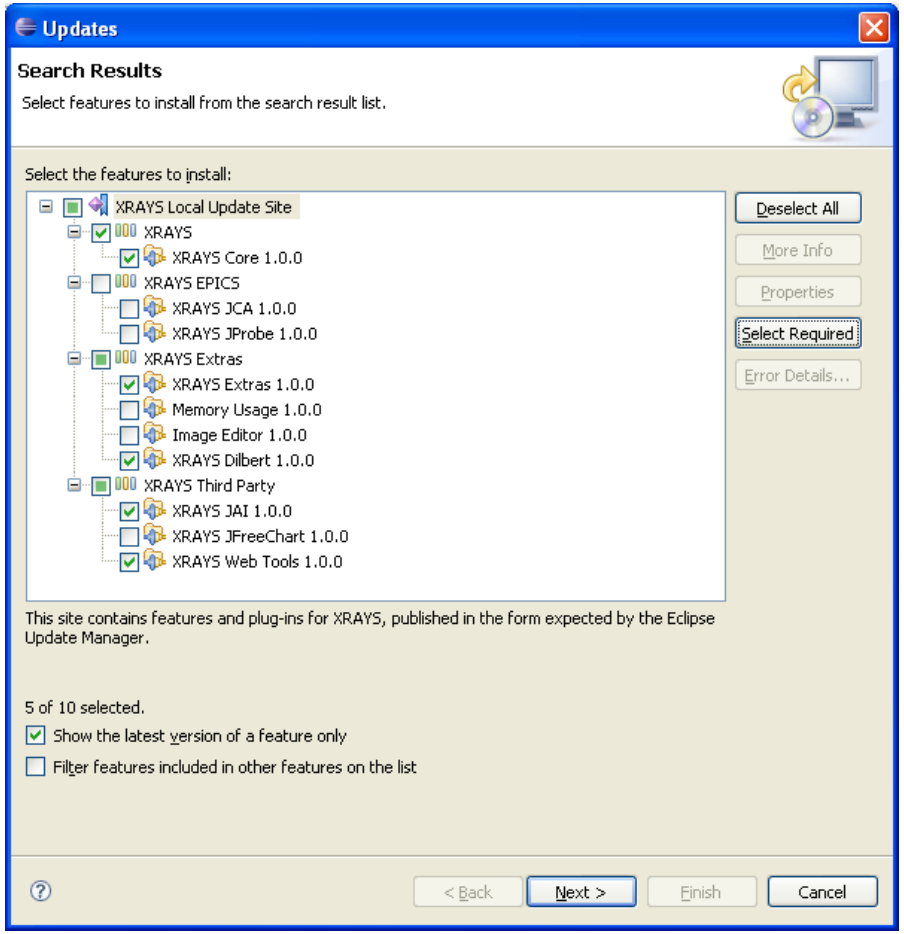


Images from: BLU-ICE and the Distributed Control System, NOBUGS III, January 2000

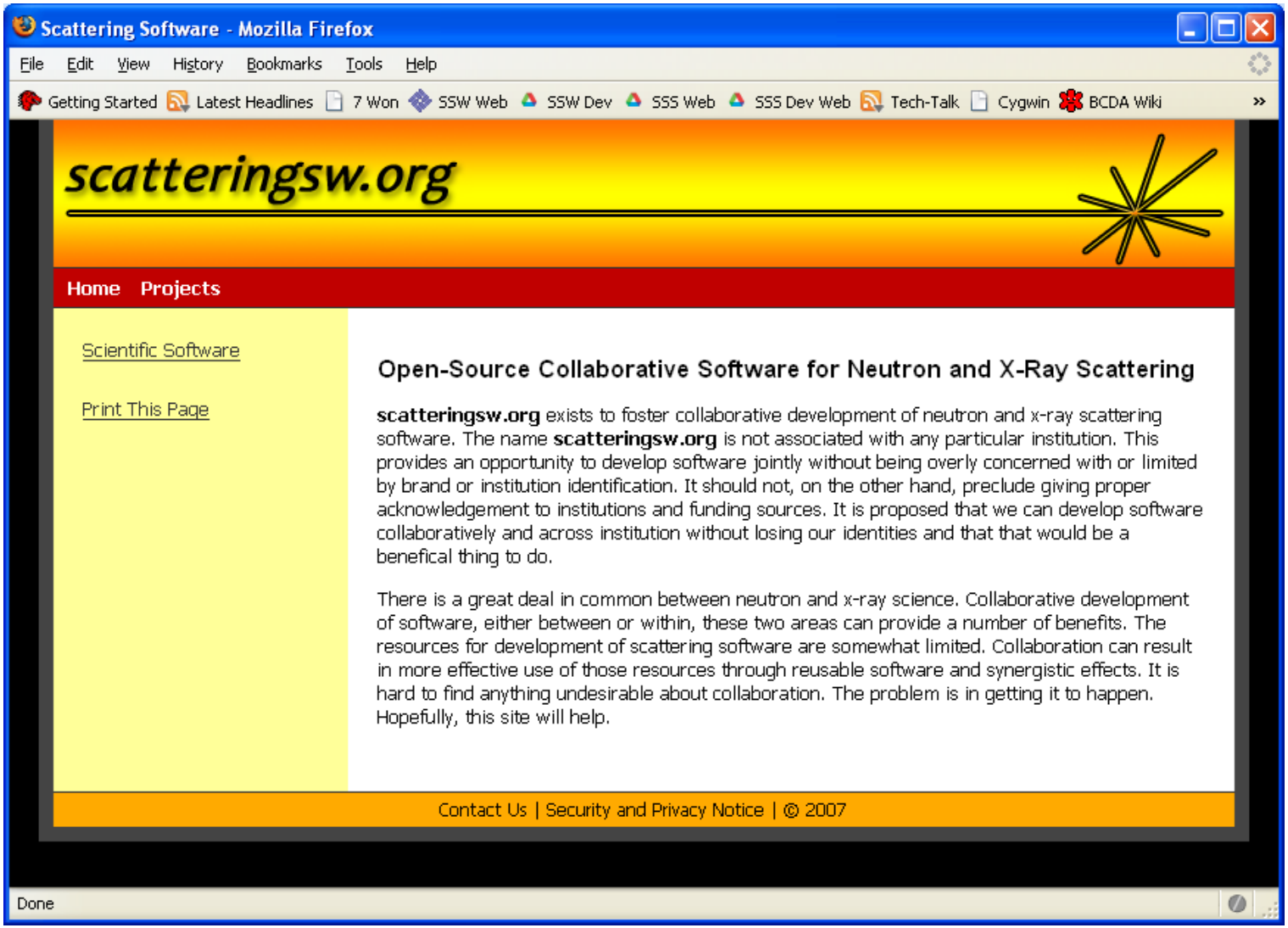
Deployable Plug-ins

- These Eclipse projects are available through the Eclipse update mechanism at http://www.scatteringsw.org/XRAYS/updates

- 10 features in 4 categories consisting of ~25 plug-ins
- Work in Progress
  - Not thoroughly tested
  - Not fully documented
- Not really scientific software
  - But potentially useful



scatteringsw.org



Acknowledgements

- Pete Jemian – Group Leader of the APS BCDA group
  - Is invaluable in helping provide the direction for this effort.

Thank You

This has been an

APS Scientific Software Section Presentation

Thank You

This has been an

APS Scientific Software Section Presentation