

ICAT Status

Alistair Mills

Alistair.Mills@stfc.ac.uk

Project Manager
Scientific Computing Department



What is ICAT?

- A metadata catalogue to manage data
 - Database (JPA+ support e.g. Oracle, MySql);
 - Web Service API (developed in Java, running in Glassfish, usable in various languages).
- To allow high-quality
 - Registration of data as it is collected in experiments;
 - Discovery of data based on what it was collected for;
 - Access to the data, according to some policy;
 - Association of data with other resources.
- Moving towards the ICAT Toolkit
 - Pluggable component architecture;
 - Supporting different tasks.



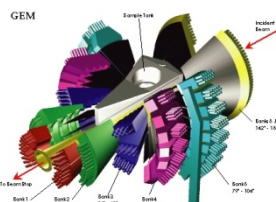
ICAT Vision

What? What? What? What? **What is ICAT?**

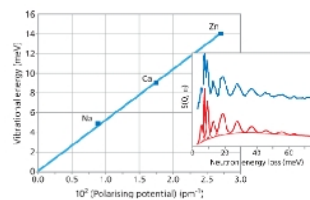
ICAT is a database (with a well defined API) that provides a uniform interface to experimental data and a mechanism to link all aspects of research from proposal through to publication.

- Access data anywhere via the web
- Annotate your data
- Search for data in a meaningful way e.g. taxonomy, Sample, temperature, pressure etc
- Share data with colleagues
- Access data via your own programs (C++, Fortran, Java etc.) via the ICAT API
- Identify potential collaborations
- Utilise integrated e-Science High-Performance Computing and Visualisation resources
- Link to data from your publications
- Etc.

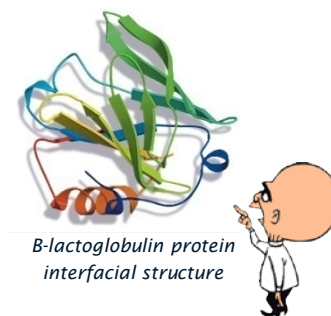
Example ISIS Proposal



GEM - High intensity, high resolution neutron diffractometer



H2-(zeolite) vibrational frequencies vs polarising potential of cations



B-lactoglobulin protein interfacial structure



Proposals

Once awarded beamtime at ISIS, an entry will be created in ICAT that describes your proposed experiment.

Experiment

Data collected from your experiment will be indexed by ICAT (with additional experimental conditions) and made available to your experimental team

Analysed Data

You will have the capability to upload any desired analysed data and associate it with your experiments.

Publication

Using ICAT you will also be able to associate publications to your experiment and even reference data from your publications.



The ICAT family of parts

The core, supported, documented parts:

- Icat.ear – the core of all deployments;
- Authn_db.ear – the database authenticator, since icat 4.2;
- Authn_ldap.ear – the ldap authenticator; since icat 4.2;
- Icat-client.jar – the client application interface;
- Topcat.war – a web based ICAT browser;
- A set of scripts to assist in simple deployments.

IDS – the integrated data server – expected in May 2013

A set of user contributed parts such as:

- Icat-api-examples.jar – an easy to use set of examples;
- icat-setup.jar – an easy to use set of rules for ICAT;
- Restful ICAT;
- PyICAT – a webservice client;
- Client demo – an example of a client application;
- Icat xml client.



Active collaboration

Telco every other Thursday at 15:00 UT/BST:

- ISIS/STFC/RAL;
- DLS/RAL;
- SCD/STFC/RAL;
- SNS/ORNL;
- ELETTRA.

Part-time members:

- ESRF;
- ILL;
- DESY;
- ALBA;
- PSI;
- Soleil;
- Jcns;

The meetings are documented and on the web site:

<http://code.google.com/p/icatproject/wiki/MeetingActionsRecord>



Release summary

Release (date)	Comments	Status
3.3 (2007)	Big API - many variants	End of life in 2013
4.0 (1/2012)	Small API – technology preview	Should no longer be used
4.1 (6/2012)	Production use for new users	Should no longer be used
4.2 (8/2012)	Pluggable authentication	In production use
4.3 (4/2013)	Under discussion	April 2013



Recent success stories

Organisation	Comments	Date
Pandata	Adoption of ICAT as preferred catalogue	July '12
ICAT	Availability of ICAT 4.2	August '12
CRISP	Adoption of ICAT as preferred catalogue	November '12
CLF/STFC	Deployment of ICAT 4.2 in production	December '12
Pandata	Successful service verification [0], [1]	December '12
ISIS/DLS	Deployment of ICAT 4.2 in production	February '13
Pandata	Integration of Virtual Lab Data	April '13
ICAT	Release of ICAT 4.3	April '13



Frequent requests for ICAT



- Make it easier to install;
- Provide set-up tools to define common usage rules;
- Provide ingestion tools for common data formats;
- More authentication methods;
- Improve the documentation;
- Provide a larger collection of supported parts;
- Provide better integration with the User Office;
- Provide a lightweight implementation on a cloud service;
- Simplify import, export, synchronisation.



Candidates for ICAT 4.3



- Remove the ICATCompat interface;
- Improve error messages for unusual errors;
- Add a entity for provenance called “run”;
- Improve notifications for usage reporting;
- Allow a rule to specify that an authz plugin apply;
- Introduce a namespace for types instead of facility;
- Experimental features.



Down the line



- More standard:
 - Vocabulary for searching/reporting;
Keywords, facilities, instruments, techniques;
Parameters and parameter sets;
Samples;
Dataset types;
Investigation types;
Datafile formats (mime-types);
Groups and rules.
- Integration with other software:
 - Mantid, Dawn;
Workflow Engine;
Umbrella;
SLS;
PanSoft software catalogue.
- Programming language bindings:
 - Java, Python, Bash
- Extend the component architecture:
 - Light-weight laboratory use;
Simple web interface;
ICE;
- Xml Ingest.
- CRISP.
- Pandata
- More Service Verifications.
- Integration with Tardis.
- More deployments.
- Example data: script to set-up with virtual data, Virtual machine.
- Linked data.
- Notification and reporting examples.
- Handle more than one facility in one ICAT.
- Common data server interface (IDS2).





Feedback

- What do you need?
- What would you like to see?
- What can you contribute?

<http://www.icatproject.org/>

<http://code.google.com/p/icatproject/>

