

ICAT : capturing and cataloguing facility metadata



Freddie Akeroyd

Tom Griffin

NOBUGS

October 2010



Science & Technology Facilities Council

ISIS

Why-CAT?
ICAT



- Increasing numbers of files
- Provenance (Creation, Ownership, History)
- Data Access requirements (Sharing and Restriction)
- Relating to Proposals and Publications
- Governments want return on investment

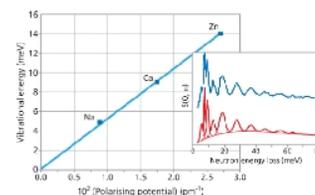
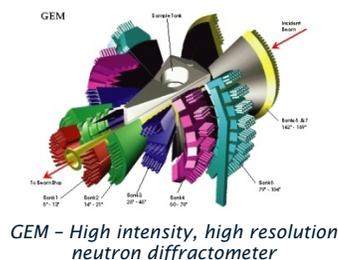
What is ICAT?

ICAT

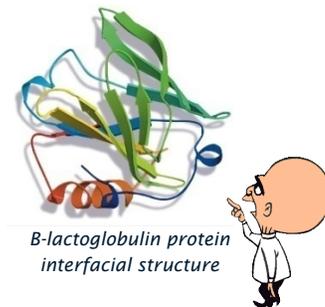
What? What? What? What? What is ICAT?

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.

Example ISIS Proposal



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



ICAT

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

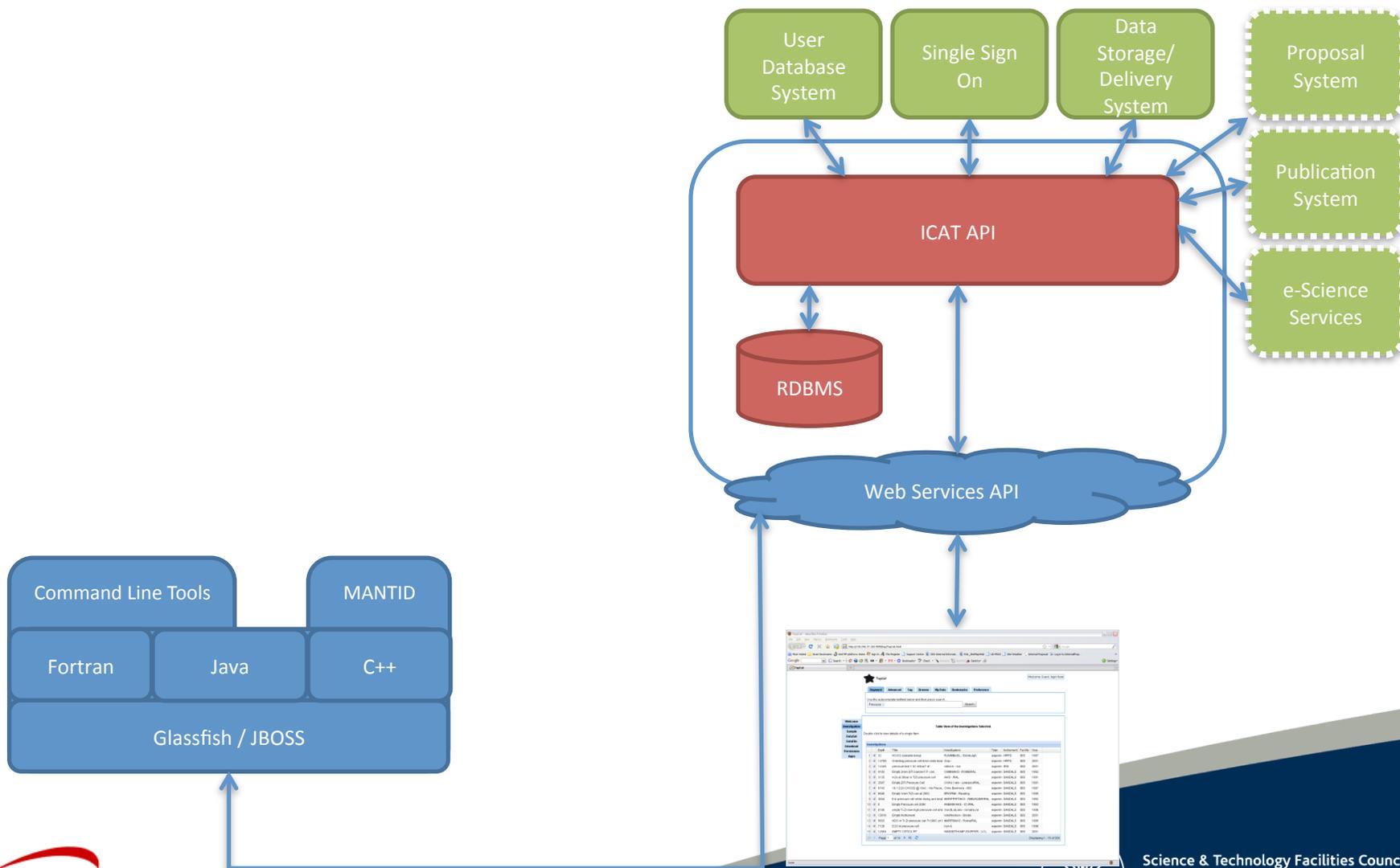
Analysed Data

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

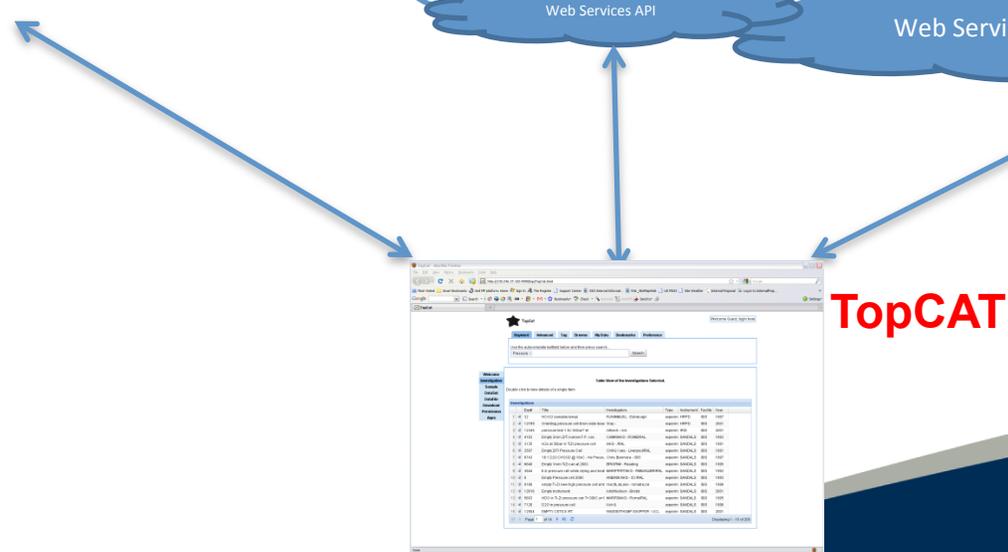
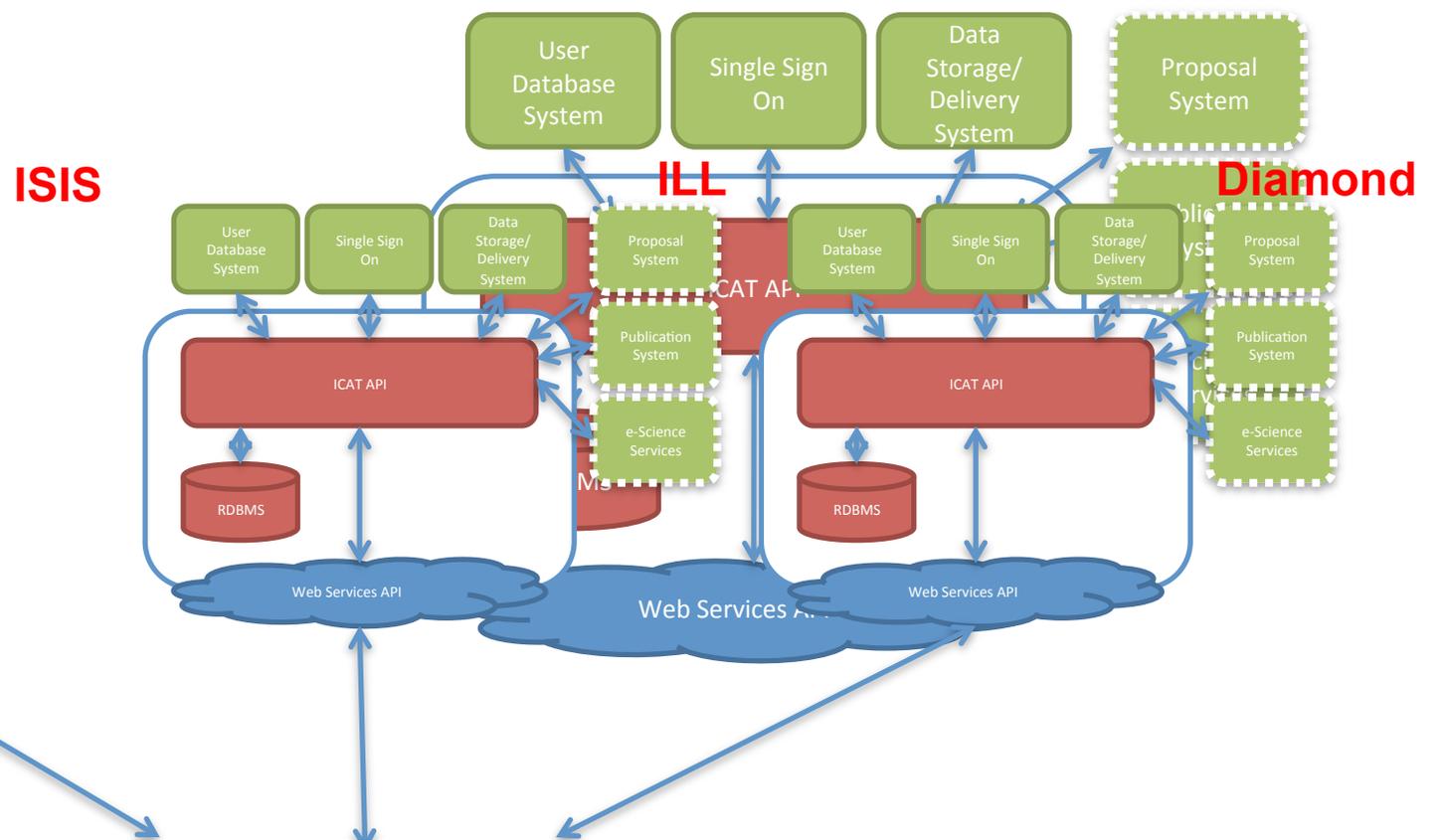
Publication

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

Overview ICAT



Federation ICAT



TopCAT – Browsing all Data ICAT

- ▶ GH342 Sr single Se RT(Id:14559)
- ▶ GH-299 SMOCS_2K(Id:14559)
- ▶ hjo27b3 Nd2 298 K(Id:14559)
- ▶ hjo2603(Id:14559)
- ▶ sb004a(Id:14559)
- ▶ HC004a 298K(Id:14559)
- ▶ Formate containing apatite D19 RT(Id:15261)
- ▶ empty cryostat OC50-02-MAXI(Id:15261)
- ▶ ZnV2O4 (sm47) cooling(Id:14588)
- ▶ Mg2NF room temperature scan(Id:15096)
- ▶ Ba2NCI Room Temperature scan(Id:15096)
- ▶ Ba2NF Room Temperature scan(Id:15096)
 - ▶ POL36473.RAW
 - ▶ POL36473.LOG
 - ▶ POL36483.RAW
 - ▶ POL36474.RAW

Datafile: POL36473.RAW

Export

Name	Units	Value
seter	N/A	
seter1	N/A	
seter2	N/A	
seter3	N/A	
settes	N/A	
short_title	N/A	Ba2NF Room Temperature
skthly	LOGOHMS	
sttc1	N/A	
sttc2	KELVIN	
socs	N/A	
srot	KELVIN	
ssamp	KELVIN	
start_date	yyyy-MM-dd HH:mm:ss	2004-11- 5 13:37:17
state	N/A	
svti	KELVIN	
temp	KELVIN	
temp1	K	

- ▶ Refined coarse PSD UMB room temperature(Id:15445)
- ▶ Coarse densified Mix (start material) room temperature(Id:15445)

TopCAT – Searching – Keyword completion

ICAT

The screenshot displays the TopCAT search interface. At the top, there is a 'Keywords' search bar containing the text 'jaw' and a 'Search' button. Below this, a dropdown menu is open, listing various search suggestions starting with 'jaw', such as 'jaw0', 'jaw1', 'jaw19', 'jaw2', 'jaw29.5', 'jaw3', 'jaw4', 'jaw5', 'jawas', 'jawas=45x45mm', 'jawd', 'Jaworska', 'JAWS', 'jaws', 'jAWS', 'Jaws', 'jaws(both)', 'Jaws(CH=44)', and 'Jaws(F=30x30)'. The 'Search All Data' radio button is selected. The 'Advanced Search' panel is visible, featuring a 'Case Sensitive' checkbox and an 'End Date' field. The 'Facilities Search' panel at the bottom shows a table with columns for 'Facility Name', 'Investigation Number', 'Title', 'Start Date', and 'End Date'.

Facility Name	Investigation Number	Title	Start Date	End Date
---------------	----------------------	-------	------------	----------

TopCAT – Downloading files from search results

ICAT

Investigation

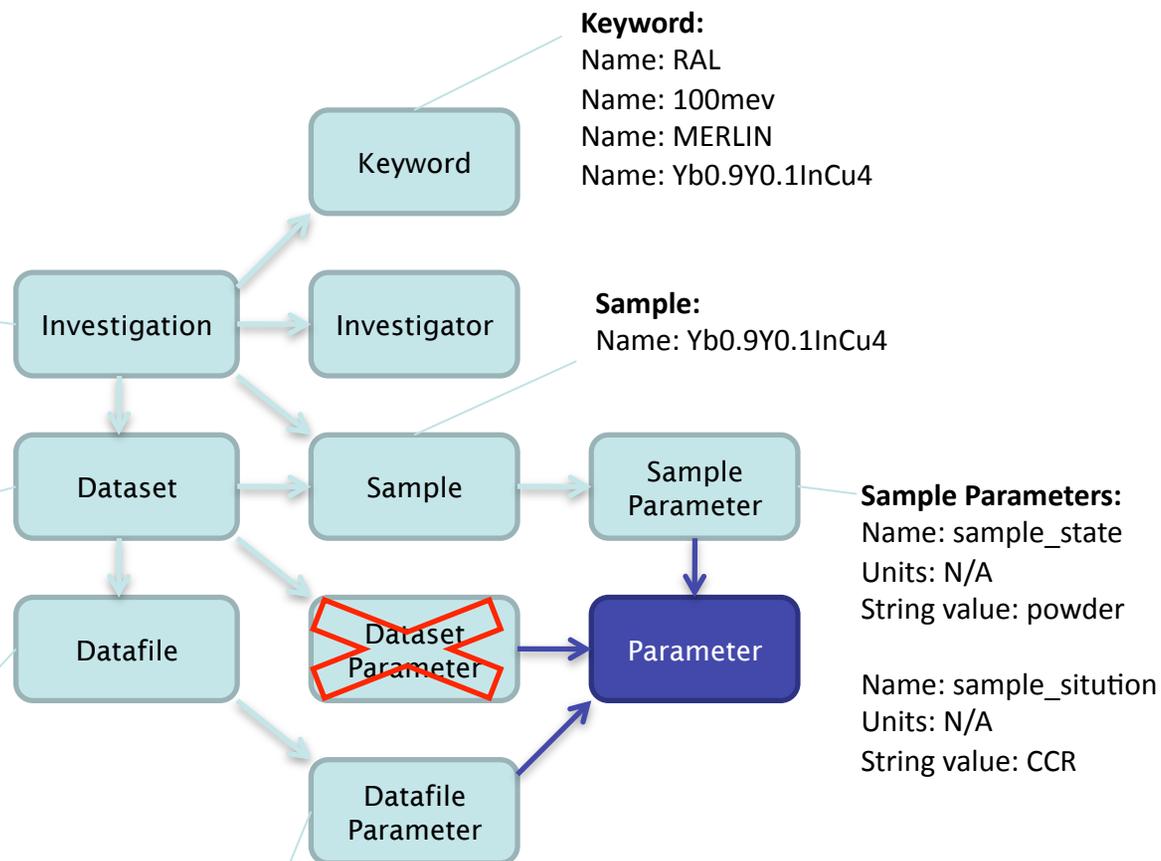
Facility: ISIS
Instrument: MERLIN
Title: Yb0.9Y0.1InCu4 8s 300k in CCR 45x45mm
inv_type: experiment
Bcat_inv_string: Mark Dr A - UniversitDr A,,

Dataset

Name: Default,
Type: experiment_raw
Dataset_Status: complete
Description: MER03766 Mark Dr A
Yb0.9Y0.1InCu4 100mev 8s 300k

Datafile

Name: MER03790.raw
Desc: Yb0.9Y0.1InCu4 15meV 4S 40K 3Kbar
CuBe cell 10x22mm
Format: isis neutron raw



Keyword:
Name: RAL
Name: 100mev
Name: MERLIN
Name: Yb0.9Y0.1InCu4

Sample:
Name: Yb0.9Y0.1InCu4

Sample Parameters:
Name: sample_state
Units: N/A
String value: powder

Name: sample_situation
Units: N/A
String value: CCR

Datafile parameter:
Name: total_proton_charge
Units: uAmpHours
Value: 0.233844

ICAT API

ICAT

- Service Oriented Architecture
 - Services exposed as Web Services
 - User required to authenticate in order to obtain Session Token
 - Token is used in all subsequent API calls to for authorisation
- The API is modular in order to fit the needs of the facilities
 - Plugin own user database
 - Plugin data delivery system
- Characteristics
 - Platform independent [*Java*]
 - Application Server independent [*EJB3*]
 - Database Independent (Almost!) [*JPL*]
 - Language independent [*Web Services*]
- Internals
 - Core functionality implemented as POJOs using JPA
 - For deployment EJB3 Session Beans bind the core API, user db and data delivery aspects together
 - Services are unit tested using JUNIT
 - Services are logged at every interaction point using LOG4J



- All raw data in ICAT and available within 30 seconds
- Links to:
 - Proposal system database
 - User Office Systems (SharePoint)
 - Safety systems (MS SQL Server database)
- Available from within Mantid

- 1 Contact Details
- 2 Samples & Handling
- 3 Sample Preparation
- 4 Beamline Experiment & Sample Environment**
- 5 After the Experiment



Beamline Experiment and Sample Environment

This section aims to assess the safety of your beamline neutron or muon experiment. It also deals with the sample environment and the conditions you will be putting your samples under.

ISIS equipment

 Ticking these boxes does not book this equipment

What ISIS equipment will you use on the beamline?

- | | | |
|--|--|---|
| <input type="checkbox"/> Helium Cryostat | <input type="checkbox"/> CCR | <input type="checkbox"/> T < 1K cryostat |
| <input type="checkbox"/> T < 0.3K cryostat | <input type="checkbox"/> Cryofurnace | <input type="checkbox"/> Gas Handling |
| <input type="checkbox"/> Furnace | <input type="checkbox"/> High Pressure | <input type="checkbox"/> 1.5 Tesla Magnet |
| <input type="checkbox"/> 7.5 Tesla Magnet | <input type="checkbox"/> 10 Tesla Magnet | <input type="checkbox"/> Water Bath |
| <input type="checkbox"/> Sample Changer | <input type="checkbox"/> RF / High Voltage | |

Details of any other ISIS equipment

Your own equipment

Will you be bringing any of your own equipment for this beamline experiment?

Sample environment conditions

Across all of your samples, what ranges of conditions will you use during your experiment?

Temperature range
from to

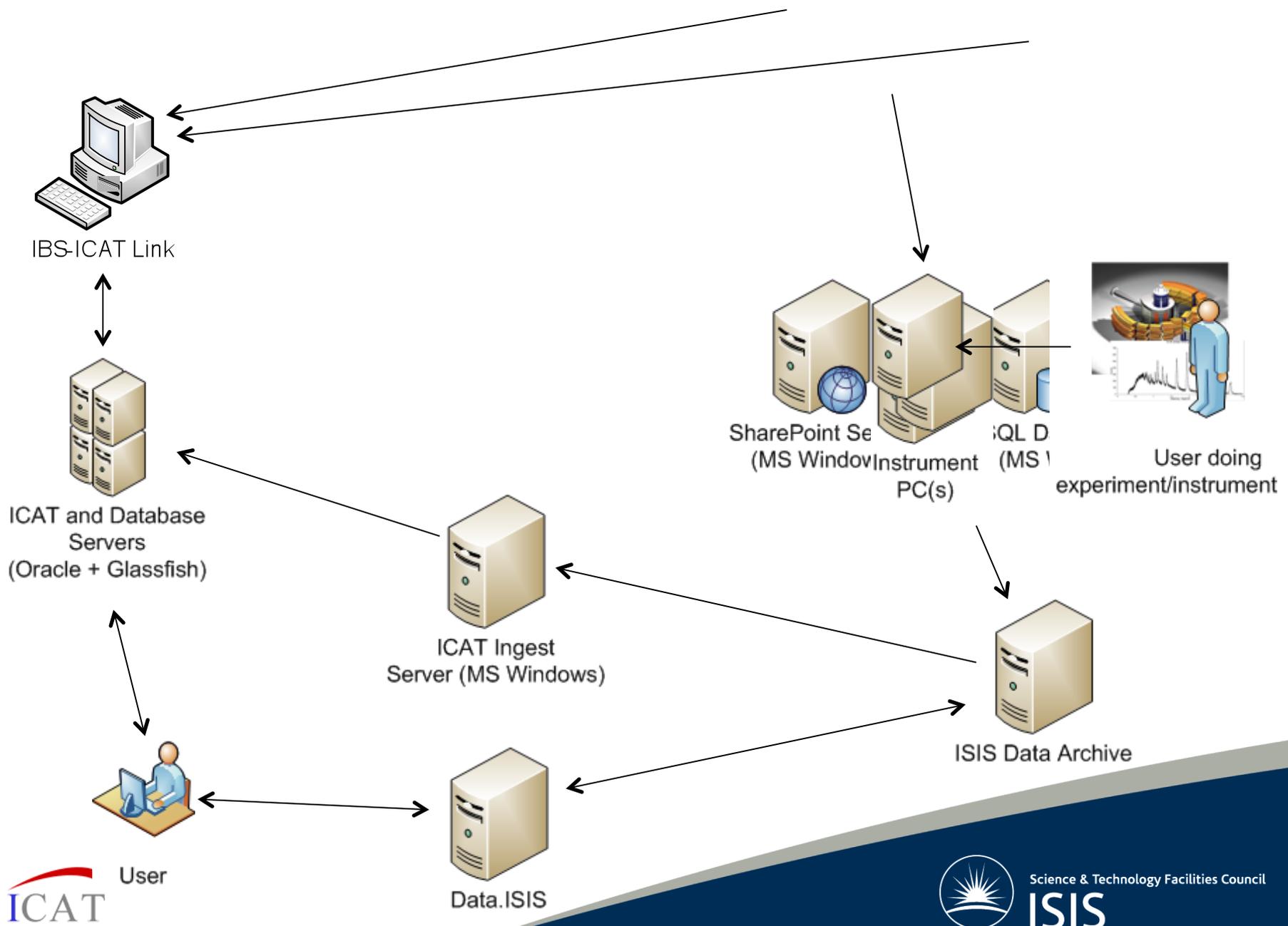
Pressure range
from to

Magnetic field strength
from to

Reactions & changes in state

Do you expect any reactions to occur during your experiment?

- | | |
|--------------------------------------|--|
| <input type="checkbox"/> Exothermic | <input type="checkbox"/> Evolution of gas |
| <input type="checkbox"/> Endothermic | <input type="checkbox"/> Absorption of gas |



- Users don't reliably enter experiment numbers
- User Office systems don't always match what goes on – can't automate
- Parameter names

(temp, temp1, temp2, temp_1, temp_...,
temperat, Temperature) – 89 in total

- Can't rely on users to enter information correctly
- Do as much as you can automatically
- Be pragmatic but robust in solutions and linking

- Installed, populated, interface ready
- Building continuous ingest tools
- Same problems with parameters
- Same problems with users

- Installed, populated, continuous ingestion
- Reliability / data rate issues on ingest

- ISIS, STFC eScience, ILL, (Diamond)
- Open source
- <http://code.google.com/p/icatproject/>
- Regular VCs
- Watching Mantid for project management ideas
- Continuous integration beneficial as the project grows

- ISIS: March 2010 – DOIs for data, full release, default data access method
- ILL – Release. Parameter searching. TopCAT as interface
- CLF – Central Laser Facility

- Continue iterative interface development
- Rules based permissions
- Rewrite ingest systems to handle partial failure
- Cross ICAT linking
- Better usage statistics

Acknowledgements

ICAT

- **ISIS**
 - Tom Griffin, (Damian Flannery), Robert McGreevy, (Kenneth Shankland), Stuart Ansell
 - Freddie Akeroyd, Chris Moreton-Smith, Matt Clarke, Kevin Knowles
- **e-Science**
 - Sri Nagella, Brian Matthews, (Glen Drinkwater, Shoaib Sufi, Kerstin Kleese Van Dam, Laurent Lerusse, Rik Tyer, Phil Couch)
 - Gordon Brown, Kier Hawker, Carmine Coiffe
 - Roger Downing
- **ILL**
 - Holger Gebhard, Najor Cruz-Cruz,
 - Jean-Francois Perrin

Questions
ICAT



www.icatproject.org

tom.griffin@stfc.ac.uk



Science & Technology Facilities Council

ISIS