



Using the ICAT API to ingest business and experiment metadata

Tom Griffin, STFC ISIS Facility
NOBUGS 2012 ICAT Workshop

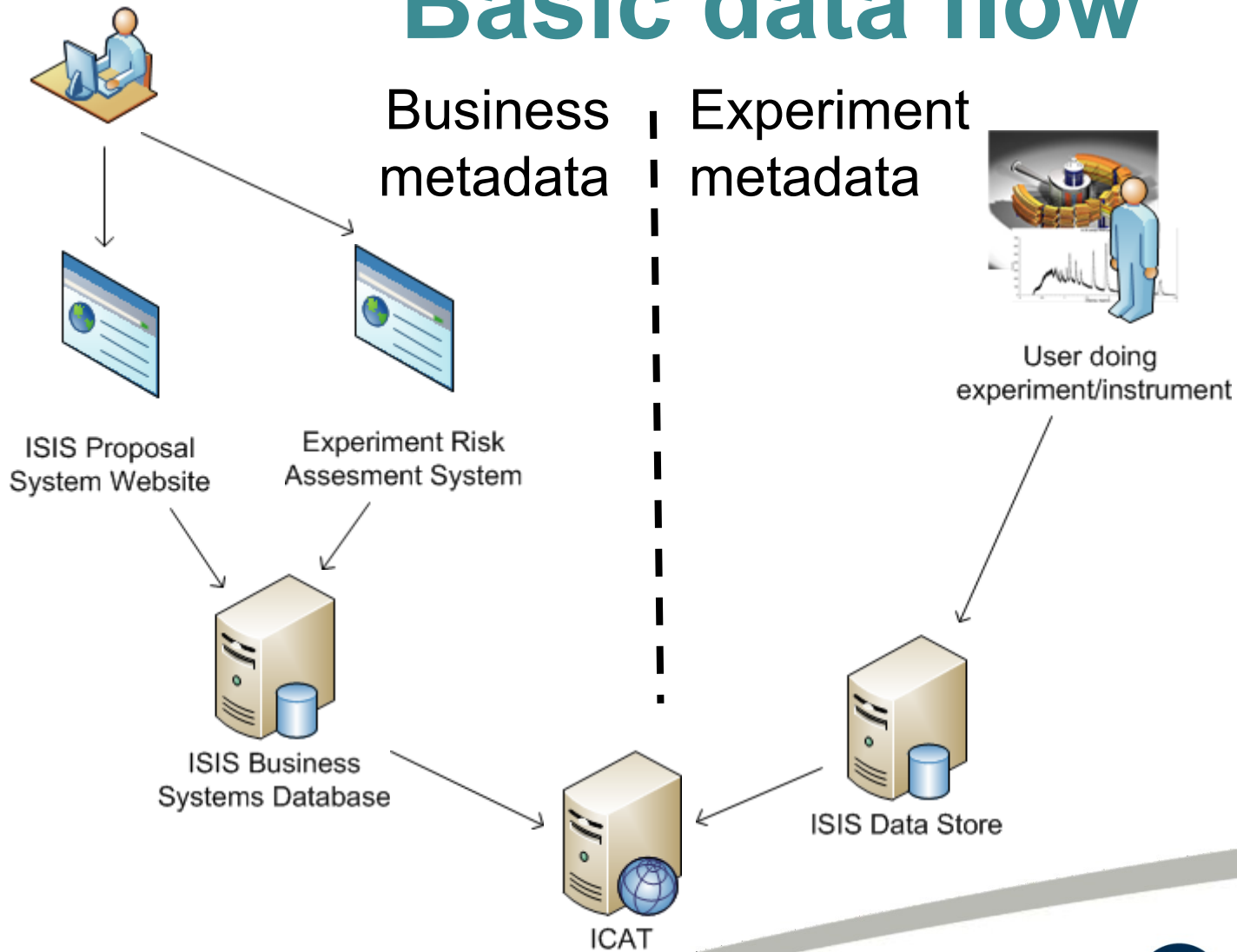
tom.griffin@stfc.ac.uk

Overview

- Assumes the user has permissions (see later)
- Basic process
- 'Business' metadata
- 'Experiment' metadata
- Tools for ingesting Nexus and XML
- 'Post experiment' file
- Exceptions



Basic data flow



'Business' metadata - Sources

- Proposal system
- User database
- Risk assessment / COSHH / Sample tracking
- Visits system

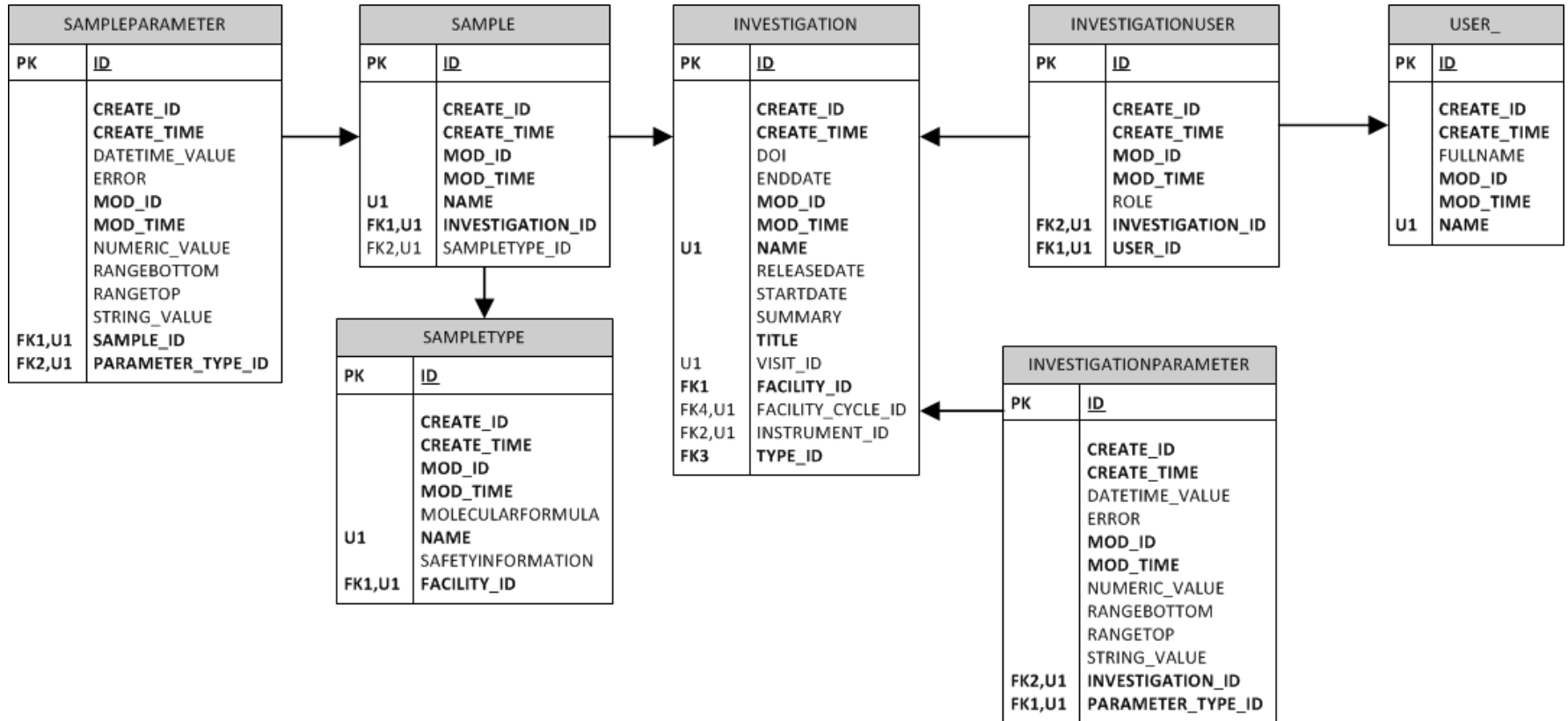


'Business' metadata - Creation

- Investigation
- InvestigationParameters
- (Users)
- Sample
- SampleParameters



'Business' metadata - Creation



'Business' metadata - Creation

- Code will be mainly facility specific
- Probably a common general flow
- Will assume users are in the User table already



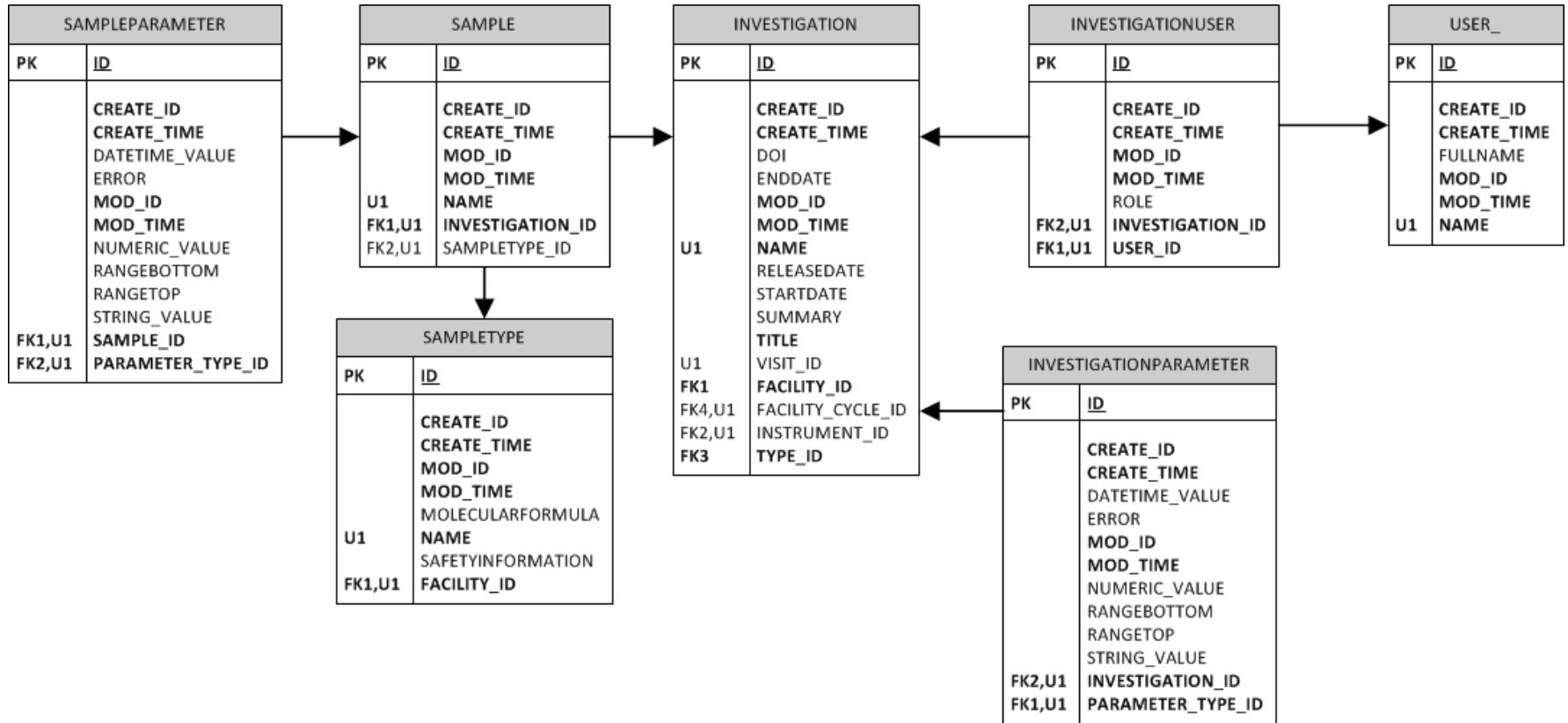
'Business' metadata - Creation

```
Investigation invOne = new Investigation();
inv.instrument = (Instrument)port.search(sessionId, "Instrument[name='" +
proposal.instrument + "']").get(0);
inv.title = proposal.title
inv.facility = (Facility)port.search(sessionId, "Facility[name='ISIS']").get(0);
inv.type = (InvestigationType)port.search(sessionId,
"InvestigationType[name='experiment']").get(0);
inv.name = investigationTitle
inv.releaseDate = Now + inv.facility.daysUntilRelease;
inv.id = port.create(sessionId, invOne);
```

```
for (User proposalUser : experimenters)
{
    InvestigationUser iu= new InvestigationUser();
    iu.investigation = inv;
    iu.user = (User)port.search(sessionId, "User[name='" + proposalUser .name +
        "']").get(0);
    iu.role = proposalUser.role ; //e.g. Principal Investigator
    iu.id = port.create(sessionId, iu);
}
```



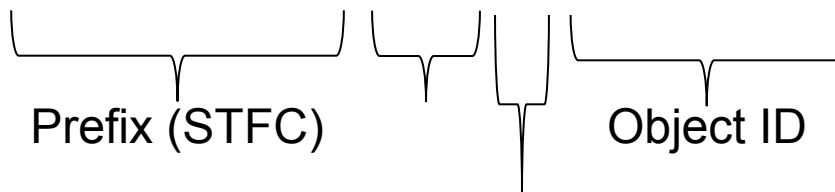
'Business' metadata - Creation



'Business' metadata - DOIs

- See talk by Brian Matthews/Micheal Wilson
- Issued by DataCite/British Library
- Previously issued on demand
- Now issued at start of experiment
- Defined by DOI issuer – us

doi:10.5286/ISIS.E.24088959



E = Experiment

D = Dataset

F = DataFile



Experiment metadata - Sources

- Datafiles
 - Nexus
 - Binary



Experiment metadata - Extraction

- ISIS – ‘WriteRaw’
 - Custom binary reader – outputs ICAT friendly XML
- ISIS – ‘NXIngest’
 - Flexible Nexus reader – outputs ICAT friendly XML
- SNS – ‘ingestNexus’
 - Simple Python script – direct input to ICAT

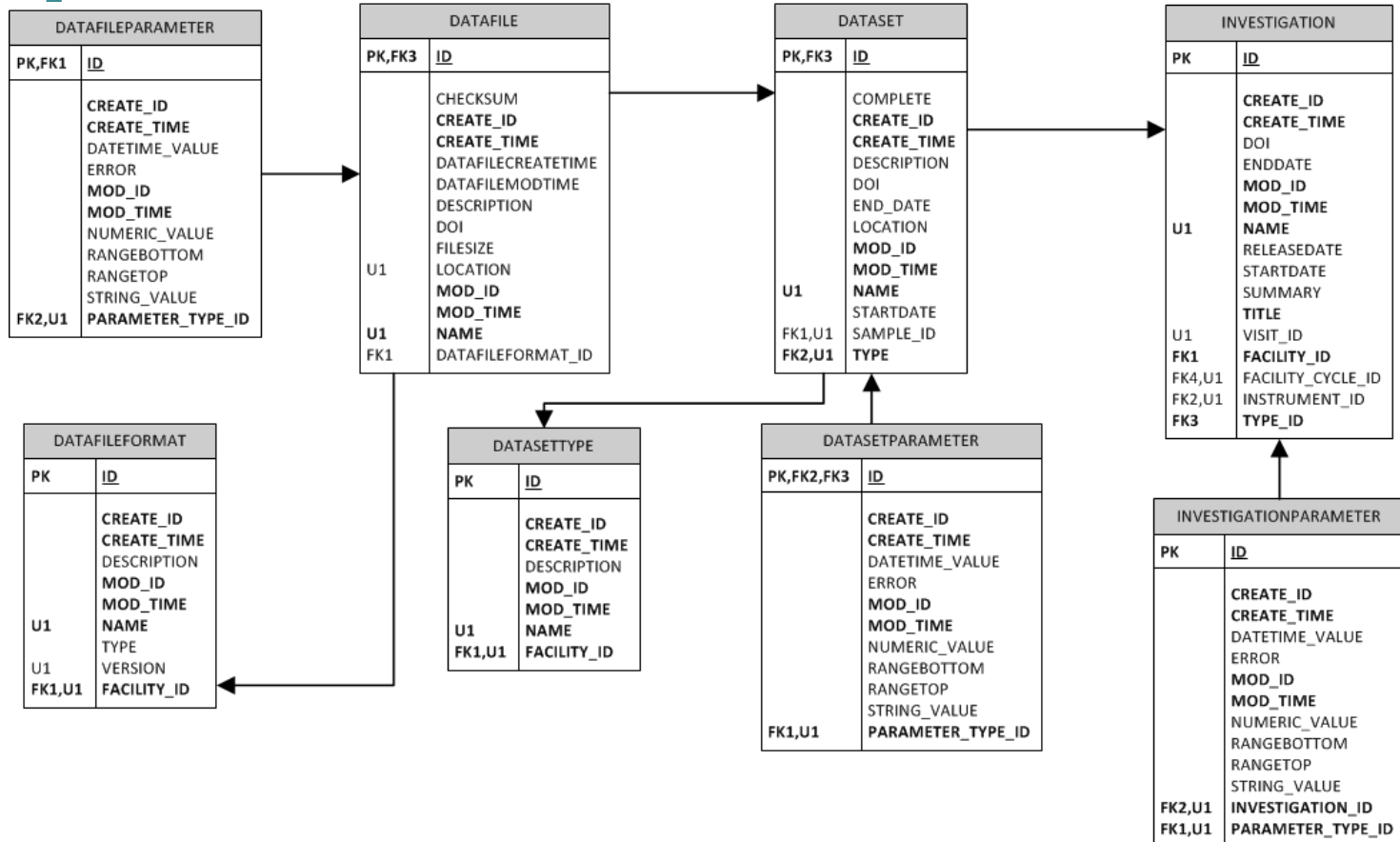


Experiment metadata - Creation

- Investigation update
- Dataset
- Datafile
- Dataset and Datafile Parameters
- Updates to Sample



Experiment metadata - Creation



Python - ingestNexus

Go to code....

<http://code.google.com/p/icatproject/source/browse/#svn%2Fcontrib%2Fscripts>



XMLIngest – a good idea?

- ICAT 3.3 exposed an XML schema
- ingestMetadata(string sessionId, string xml)
- Removed in ICAT 4.0




```
<?xml version="1.0" encoding="utf-8"?>
<icat version="1.0 RC6" xsi:noNamespaceSchemaLocation=" ../icatXSD/icatXSD.xsd" xmlns:xsi="http://www.w3.org/2001/XMLSchema-
<study>
  <investigation>
    <inv_number>1010613</inv_number>
    <visit_id>1</visit_id>
    <instrument>MAPS</instrument>
    <title>Some title would go here</title>
    <investigator>
      <user_id>damian</user_id>
      <affiliation_id>9932</affiliation_id>
      <role>experimenter</role>
      <privilege>READER</privilege>
    </investigator>
    <investigator>
      <user_id>doug</user_id>
      <affiliation_id>9933</affiliation_id>
      <role>principal_experimenter</role>
      <privilege>ADMIN</privilege>
    </investigator>
    <inv_abstract>abstract would go here and would be very interesting to read</inv_abstract>
    <inv_type>Experiment</inv_type>
    <keyword>
      <name>abstract</name>
    </keyword>
    <keyword>
      <name>proposal</name>
    </keyword>
    <keyword>
      <name>hydrogen</name>
    </keyword>
    <sample>
      <name>Carbon</name>
      <chemical_formula>C</chemical_formula>
      <safety_information>Safe</safety_information>
      <parameter>
        <name>Sample Type</name>
        <string_value>Solid</string_value>
        <units>N/A</units>
      </parameter>
      <parameter>
        <name>Weight</name>
        <numeric_value>50</numeric_value>
        <units>g</units>
      </parameter>
      <parameter>
        <name>Sample Environment Requested</name>
        <string_value>CCR</string_value>
        <units>N/A</units>
      </parameter>
    </sample>
    <sample>
      <name>Hydrogen</name>
      <chemical_formula>H</chemical_formula>
```

```
<?xml version="1.0" encoding="utf-8"?>
<icat version="1.0 RC6" xsi:noNamespaceSchemaLocation="../icatXSD/icatXSD.xsd" xmlns:xsi="http://www.w3.org/2001/XMLSchema-
<study>
  <investigation>
    <inv_number>1010613</inv_number>
    <visit_id>1</visit_id>
    <instrument>MAPS</instrument>
    <title>new proposal</title>
    <inv_type>EXPERIMENT</inv_type>
    <bcat_inv_str>MAPS team -</bcat_inv_str>
    <keyword>
      <name>Davide</name>
    </keyword>
    <keyword>
      <name>-</name>
    </keyword>
    <keyword>
      <name>Quiet</name>
    </keyword>
    <keyword>
      <name>count-PH</name>
    </keyword>
    <keyword>
      <name>Mode</name>
    </keyword>
    <dataset>
      <name>Default</name>
      <dataset_type>EXPERIMENT_RAW</dataset_type>
      <dataset_status>COMPLETE</dataset_status>
      <description>WIS17323 MAPS team Davide - Quiet count-PH
08-JUL-2010 14:05:29 0.00</description>
      <sample>
        <name>Davide - Quiet count-PH Mode</name>
        <safety_information>No information
available</safety_information>
      </sample>
      <datafile>
        <name>MAPS00017323.raw</name>
        <location>
file:///S:/Cycles$/cycle_10_2/NDXMAPS//MAPS00017323.raw</location>
        <description>Davide - Quiet count-PH Mode</description>
        <datafile_format>isis neutron raw</datafile_format>
        <datafile_format_version>2</datafile_format_version>
        <datafile_create_time>
2010-07-08T14:05:29</datafile_create_time>
        <datafile_modify_time>
2010-08-11T11:32:00</datafile_modify_time>
        <file_size>451466240</file_size>
        <parameter>
          <name>run_number</name>
          <numeric_value>17323</numeric_value>
          <units>N/A</units>
        </parameter>
        <parameter>
```

NXIngest

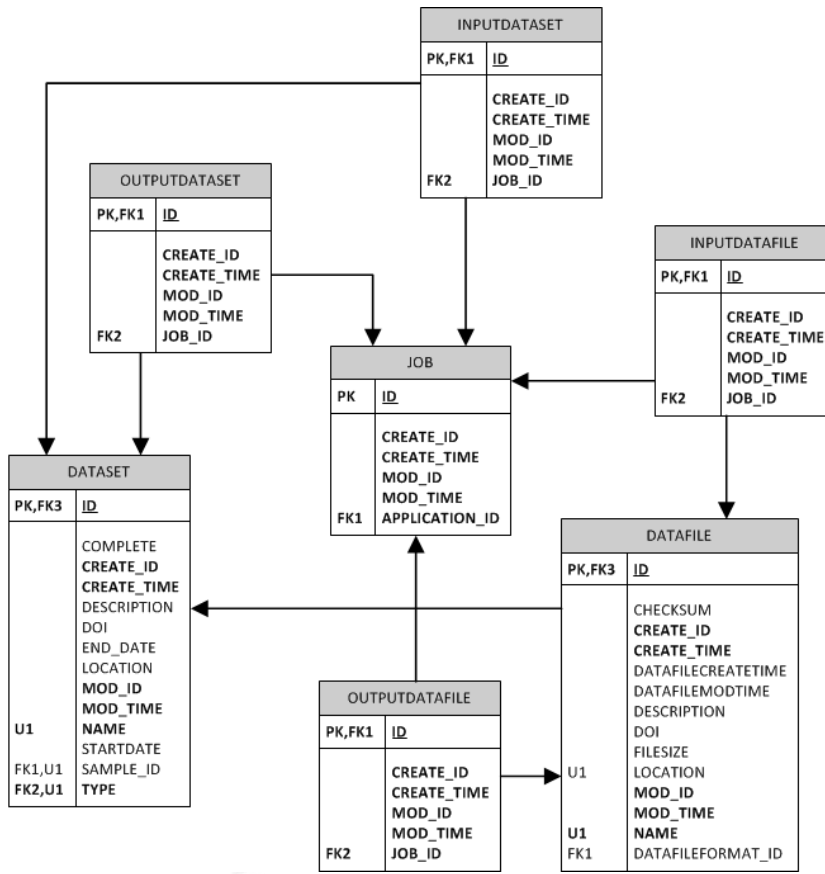
- Part of the nexus package
- Highly customizable by a mapping file
- Different mapping file per nexus 'style'

- Go to XML files.....



Post-Experiment

- Reduced, Derived, Analysed Data
- All can be created in the same way as raw data/metadata



Exceptions

- Single exception: `IcatException_Exception`
- A wrapper around the real exception which in turn includes an enumerated code to identify the kind of exception and the usual message.

Enum Constant Summary

Enum Constants

Enum Constant and Description

<code>BAD_PARAMETER</code>
<code>INSUFFICIENT_PRIVILEGES</code>
<code>INTERNAL</code>
<code>NO_SUCH_OBJECT_FOUND</code>
<code>OBJECT_ALREADY_EXISTS</code>
<code>SESSION</code>
<code>VALIDATION</code>

Field Summary

Fields

Modifier and Type	Field and Description
<code>protected String</code>	<code>message</code>
<code>protected int</code>	<code>offset</code>
<code>protected IcatExceptionType</code>	<code>type</code>



Exceptions

- **BAD_PARAMETER**
 - generally indicates a problem with the arguments made to a call.
- **INTERNAL**
 - may be caused by network problems, database problems, glassfish problems or bugs in ICAT.
- **INSUFFICIENT_PRIVILEGES**
 - indicates that the authorization rules have not matched your request.
- **NO_SUCH_OBJECT_FOUND**
 - thrown when something is not found.
- **OBJECT_ALREADY_EXISTS**
 - thrown when type to create something but there is already one with the same values of the constraint fields.



Exceptions

- **SESSION**
 - is used when the sessionId you have passed into a call is not valid or if you are unable to authenticate.
- **VALIDATION**
 - marks an exception which was thrown instead of placing the database in an invalid state.



Exceptions

- Operations which work on a list of objects, such as `createMany`, may fail because of failure to process one of the objects.
- In this case the state of the database will be rolled back
- *offset* is set to the index of the entry in the list causing the error



Exceptions

```
String sessionId;  
try {  
    sessionId = icat.login("db", credentials);  
}  
catch (IcatException_Exception e) {  
    IcatException ue = e.getFaultInfo();  
    System.out.println("IcatException " + ue.getType() +  
        " " + ue.getMessage() + (ue.getOffset() >= 0 ? "  
    at offset " + ue.getOffset() : ""));  
}
```



Summary

- Explained my ingest process (typical?)
- Creating Investigations, Parameters, Samples
- Creating Datasets, Datafiles, Parameters
- Existing tools available to help
- Consider XML ingest
- Post experiment data
- Exceptions





Questions?

www.icatproject.org

code.google.com/p/icatproject

<http://groups.google.com/group/icat-developers/>

<http://groups.google.com/group/icatgroup/>

icat-developers@googlegroups.com

icatgroup@googlegroups.com