

# DOIs at STFC

Sri Nagella and Michael Wilson  
STFC eScience

Tom Griffin  
STFC ISIS Facility  
8<sup>th</sup> Feb 2010



Science & Technology Facilities Council

ISIS

# What and Why

- Digital Object Identifiers
- “10.1000/182”
- Persistent identifier
- Widely used for journal articles
  - Accepted standard
- Citation, tracking, access, visibility



# Issuing DOIs

- STFC are using DataCite/British Library
- First 5000 for free
- Then £250 / thousand
- At investigation level = 800/year
- Option to issue at any ICAT level
- Certain required public metadata
- Landing page

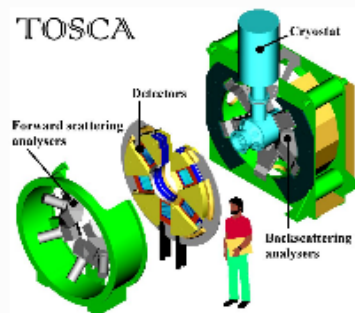


# Public metadata

- DOI = 10.5286/ISIS.V.1010256
- URL of landing page
- Creator = Investigators/PI
- Title = Title
- Publisher = STFC ISIS Facility
- Date = 3<sup>rd</sup> December 2010
- Data Format = ISIS Neutron RAW/ NXS
- Data Citation =
  - Griffin, T. et al; (2010): RB12345, STFC ISIS Facility,  
doi:10.1107/S0021889810034357



TOSCA



Data collected on the  
Tosca instrument  
at the ISIS facility

ISIS Data Data Set

**Title:** RB12345.

**Investigation title:** Hitting a material with neutrons to determine its structure.

**Creator:** *Griffin, T.*

**Creator:** *Smith, T.*

**DOI:** 10.1107/S0021889810034357

**Date of Experiment:** January 2004 BCE

**Publisher:** STFC ISIS Facility

**Data format:** RAW

Select the data format above to find out more about it.

### Data Citation

The recommended format for citing this dataset in a research publication is as:  
[author], [date], [title], [publisher], [doi]

For Example:

Griffin, T. et al; (2010): RB12345, STFC ISIS Facility, doi:10.1107/S0021889810034357

### Abstract

abstract text for data sets which are older than the embargo period. In the beginning God created the heaven and the earth. And the earth was without form, and void; and darkness was upon the face of the deep. And the Spirit of God moved upon the face of the waters. And God said, Let there be light: and there was light. And God saw the light, that it was good: and God divided the light from the darkness. And God called the light Day, and the darkness he called Night. And the evening and the morning were the first day. And God said, Let there be a firmament in the midst of the waters, and let it divide the waters from the waters. And God made the firmament, and divided the waters which were under the firmament from the waters which were above the firmament: and it was so. And God called the firmament Heaven. And the evening and the morning were the second day. And God said, Let the waters under the heaven be gathered together unto one place, and let the dry land appear: and it was so. And God called the dry land Earth; and the gathering together of the waters called he Seas: and God saw that it was good. And God said, Let the earth bring forth grass, the herb yielding seed, and the fruit tree yielding fruit after his kind, whose seed is in itself, upon the earth: and it was so. And the earth brought forth grass, and herb yielding seed after his kind, and the tree yielding fruit, whose seed was in itself, after his kind: and God saw that it was good. And the evening and the morning were the third day. And God said, Let there be lights in the firmament of the heaven to divide the day from the night; and let them be for signs, and for seasons, and for days, and years: And let them be for lights in the firmament of the heaven to give light upon the earth: and it was so. And God made two great lights; the greater light to rule the day, and the lesser light to rule the night: he made the stars also. And God set them in the firmament of the heaven to give light upon the earth, And to rule over the day and over the night, and to divide the light from the darkness: and God saw that it was good.



download  
the dataset

# Formats

- DataCite assigns a prefix to STFC
  - 10.5286
- We can create 5000 DOIs under this
- 10.5286/<Facility>.<E/D/F>.<Id>
- E-Investigation, D- Dataset, F-Datafile
  
- 10.5286/ISIS.V.1010256



# How it happens

- 3 days before an experiment
- Metadata ingested into ISIS ICAT
- Calls STFC webservice with metadata
- Maintains a small database of DOIs we have issued
- (Create also overwrites!)
- [dx.doi.org](http://dx.doi.org) dereferences to TopCAT page



# MySQL ICAT

Richard Sinclair  
STFC eScience

Tom Griffin  
STFC ISIS Facility  
8<sup>th</sup> Feb 2010



Science & Technology Facilities Council

ISIS



# Basic Tables

- MySQL Conversion Toolkit.
- connects to the Oracle database and builds the same table structures in MySQL.
- tables, indexes, foreign key constraints



# Moving Parts

- ***Triggers***

- **AIR\_DATASET\_TRG** - Updates ICAT\_AUTHORISATION table when a new a dataset is inserted
- **AIR\_INVESTIGATION\_TRG** - Updates ICAT\_AUTHORISATION table when a new a investigation is inserted
- **INVESTIGATION\_DATE\_TRG** - Sets days\_until\_public\_release when a new investigation is inserted
- 

- ***Procedures***

- **SET\_ICAT\_AUTH\_DATA\_ENTITIES** - grants permissions to view investigations and datasets
- **SET\_INVESTIGATION\_DATES** - Updates investigation times
- **WRITE\_LOG\_INTERNAL** - writes error messages to LOG\_TABLE
- **WRITE\_LOG** - Calls WRITE\_LOG\_INTERNAL with specific flag for non-exception set
- **WRITE\_EXCEPTION** - Calls WRITE\_LOG\_INTERNAL with specific flag for exception set



# PL/SQL -> SQL:2003

- Looping and cursors different in PL/SQL and SQL:2003 – most moving parts rewritten
- *“For example one of the procedures grabs a subset of data into a cursor using the SELECT FOR UPDATE clause and then does an update on the row in the cursor using WHERE CURRENT OF. In MySQL cursors are read only so no update can occur, this had to be worked around by ditching the SELECT FOR UPDATE and updating the row based on the id currently held in the cursor.”*



# Sequences

- ICAT uses Oracle Sequences for PKs
- No Sequences in MySQL
- Simplest and most portable solution is sequence tables
- Similar performance due to JPA allocation/caching



# Where are we?

- Completed *very* basic testing
  - Looks OK
- Need to run full test suite



# Why?

- Solves problem of Oracle licence for virtual machine distribution
- Very limited work
- SQL2000 so should be portable



# Want more?

- Ask me or (better) Richard Sinclair

[richard.sinclair@stfc.ac.uk](mailto:richard.sinclair@stfc.ac.uk)



Science & Technology Facilities Council

ISIS