Towards Lensfield: data management, processing and semantic publication for vernacular e-science

Nick Day, Jim Downing, Lezan Hawizy, Nico Adams and Peter Murray-Rust
Unilever Centre for Molecular Science Informatics, University of Cambridge

This presentation: CC-By-SA Jim Downing
Linked Data

CC-By-SA-NC jmelchio

CC Images from Flickr
Selling Linked Data

Make it transparent
Make it easy

CC-By mrslogic
Selling Linked Data

Citations
Selling Linked Data

- Visualizations
- Data management
- Automation
Demo

http://code.google.com/p/lensfield/
Lensfield Principles

• Make it easier to do the right thing
• Vernacular
• KISS and Embrace constraints
Constraints

• Work on the desktop without infrastructure installation

• Processing tasks could be anything and aren’t predictable
Re-use

Lensfield

Jumbo Converters

Mercurial
Jumbo-Converters

- Library of chemistry file format converters, semantifiers and enhancers
- Part of the CML Java libraries
- http://sourceforge.net/projects/cml/
Version Control

- Mercurial
  - Excellent support for experimentation
  - Backup to remote machine
  - P2P sharing

- Track script changes with data
- Automatically ignore deterministic intermediates
Build metaphor

- Describing state transitions rather than process better for provenance tracking
- Alternative to graphical programming languages / workflow packages
- Hard problems are re-use and comprehension
• Strong on concurrency
• Functional
• Software Transactional Memory
• Lisp
• Snapshots, pause and resume, continuations
Future Development

• Templated Parameter Sweeps & sensitivity analysis
• Design of Experiments
• Multicore performance testing
• Grid processing
Users

- CLARION project
  - Embargo management and publication of Electronic Lab Notebook data.
- OREChem
  - Distributed chemistry eScience using Linked Data.
- Computational Chemical engineering
Users

You?

... to use Lensfield!

http://code.google.com/p/lensfield/

CC-By-NC ilonameagher
Thanks

Colleagues
Nick Day
John Aspden
Lezan Hawizy
Peter Murray-Rust

Collaboration and Inspiration
Nico Adams (Dept of Genetics, Cambridge)
Jerry Winter (Unilever)
Noel Ruddock (Unilever)
Markus Kraft, Weerapong Phadungsukanan (Chemical Engineering, Cambridge)

Funds
JISC
Unilever