

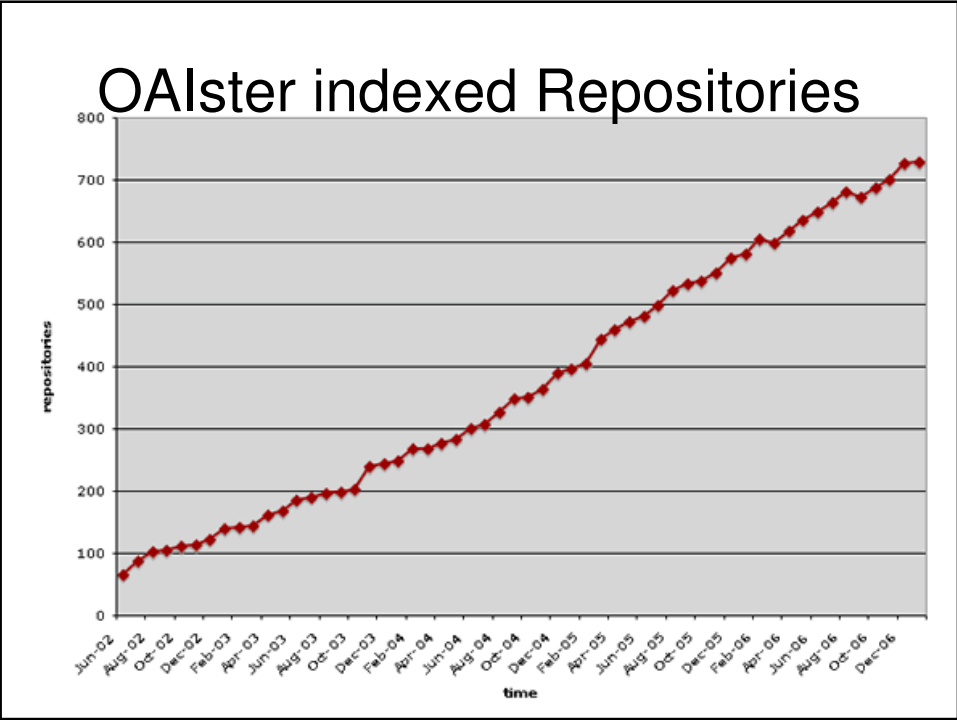
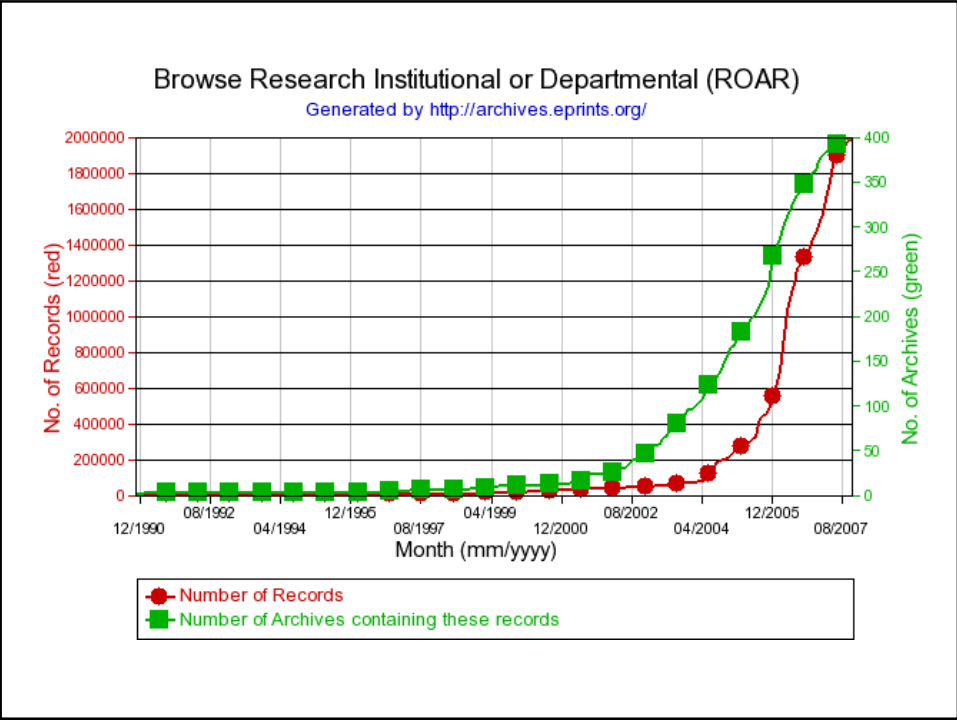
# Repositories, Use and Impact

Tim Brody

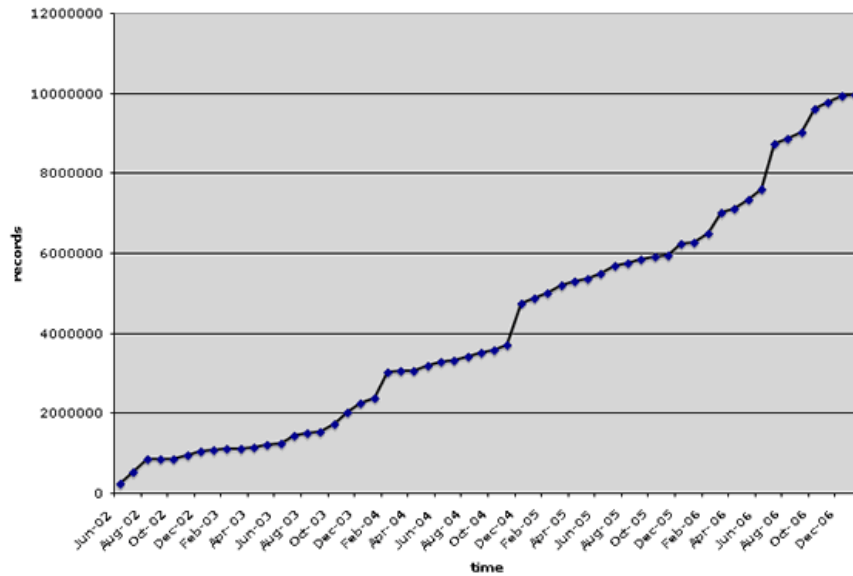
## Content

- Repositories, a linear story
- Introducing IRStats:
  - picking apart repository usage



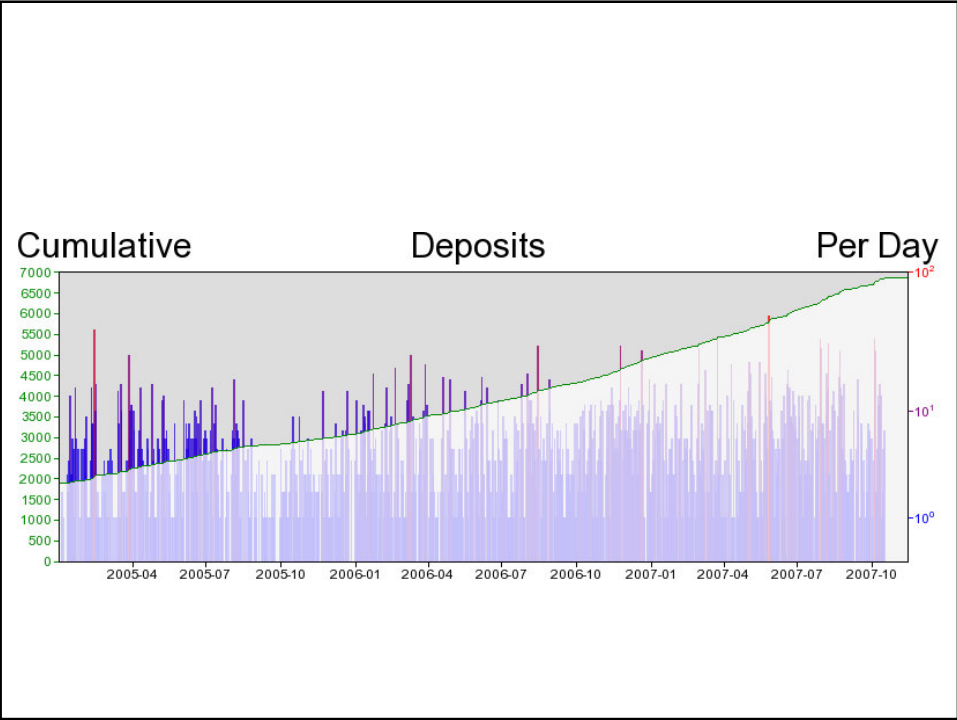
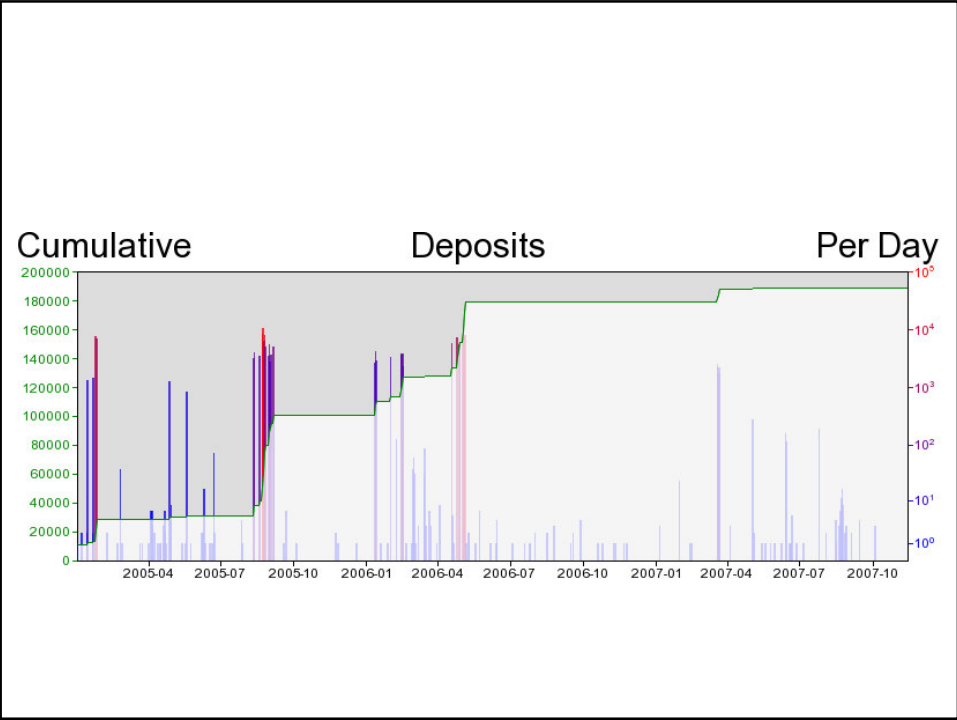


## OAlster indexed records



So far, so linear ...

- What does the **pattern** of records tell us?
- 10,000 records on one day
- 100 records over 100 days

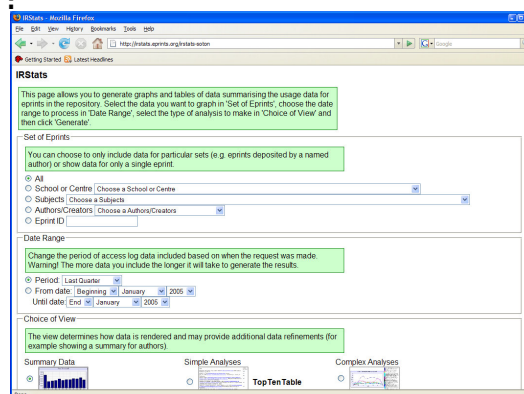


## Size Isn't Everything

- Some of the 'largest' repositories are built from existing databases/aren't primary research
- Successful embedding of repository = higher activity
- To see more: <http://roar.eprints.org/>

## Introducing IRStats

- Repository usage analyser
- i.e. How many fulltexts downloaded? By-author? By-group?



The screenshot shows the IRStats web interface. At the top, there is a navigation bar with 'Getting started' and 'Latest Headlines'. Below this is a main heading 'IRStats' and a brief description: 'This page allows you to generate graphs and tables of data summarising the usage data for eprints in the repository. Select the data you want to graph in 'Set of Eprints', choose the date range to process in 'Date Range', select the type of analysis to make in 'Choice of View' and then click 'Generate'.

The interface is divided into several sections:

- Set of Eprints:** A section where users can choose to include data for particular sets (e.g. eprints deposited by a named author) or show data for only a single eprint. Options include:  All,  School or Centre (with a dropdown menu),  Subjects (with a dropdown menu),  Authors/Creators (with a dropdown menu), and  Eprint ID (with a text input field).
- Date Range:** A section where users can change the period of access log data included based on when the request was made. A warning states: 'Warning! The more data you include the longer it will take to generate the results.' Options include:  Period (with a dropdown menu),  From date (with dropdowns for 'Beginning', 'January', and '2005'), and  Until date (with dropdowns for 'End', 'January', and '2005').
- Choice of View:** A section where users can determine how data is rendered and may provide additional data refinements (for example showing a summary for authors). It includes a 'Summary Data' section with a bar chart icon, and two radio buttons for 'Simple Analyses' and 'Complex Analyses'. Under 'Simple Analyses', there is a 'Top Ten Table' option.

## Using IRStats #1

- Download IRStats
- Install dependencies:
  - Perl Logfile::EPrints, Date::Calc, Geo::IP
  - awstats (robots.pm)
  - Optional: ChartDirector, Maxmind Org Database
- Install IRStats
  - ./configure, make install

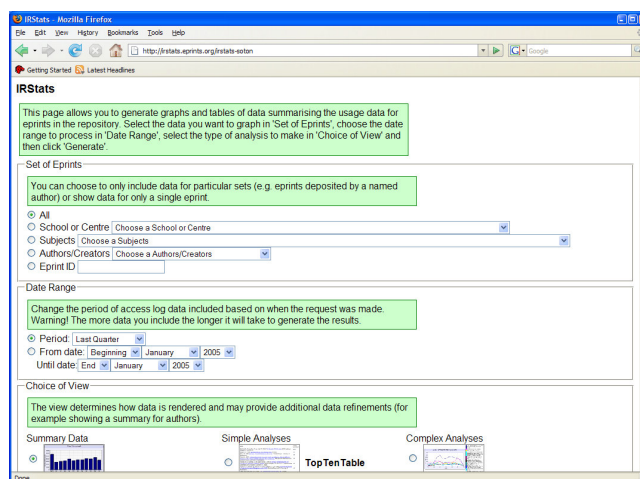
## Using IRStats #2

- Configure settings (irstats.conf)
  - Repository name/type
  - Database access
  - Sets
- Apache
  - Add irstats.cgi
  - Add static path (CSS+graphs)

## Using IRStats #3

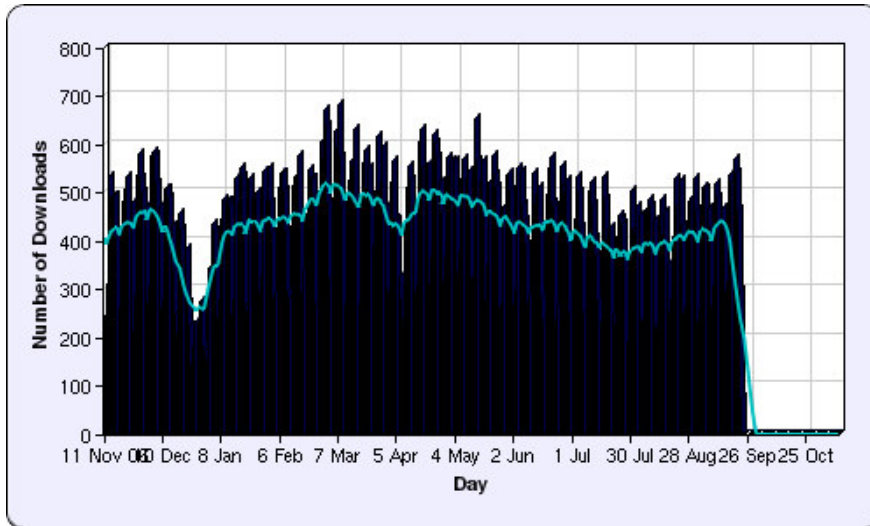
- Run `update_table`
  - EPrints access/accesslog → `irstats_true`
  - DSpace access.log → `irstats_true`
- Run `export_metadata_*`
  - EPrints supported, DSpace to do
- Run `import_metadata`
- <http://yourrepository/cgi/irstats.cgi>

## Repository Admin Tool

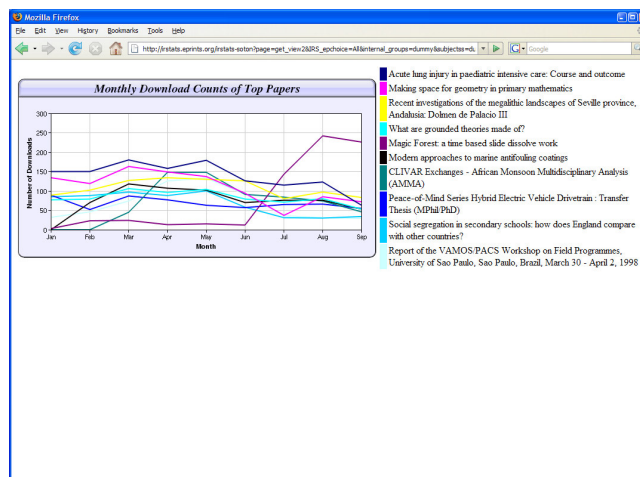




# Daily Fulltext Downloads



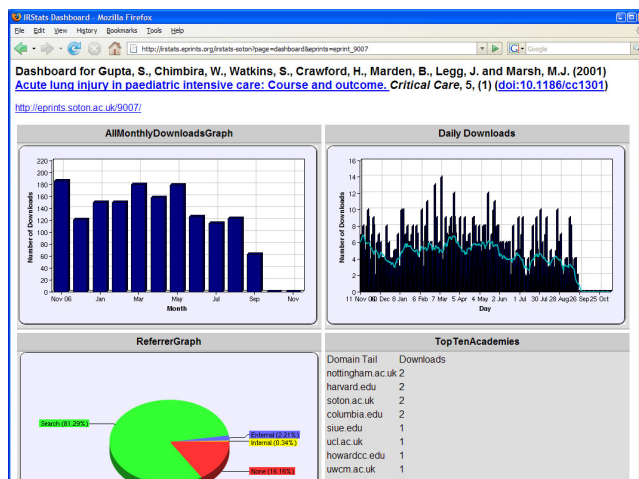
# Top-Ten Eprints Trend



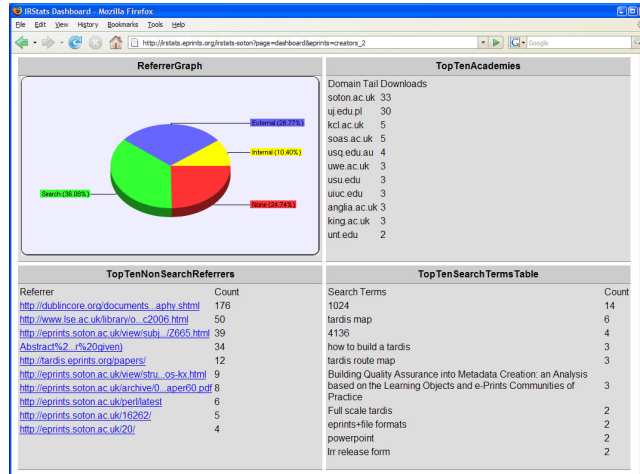
# Top-Ten Eprints

Eprint	Fulltext Downloads
Gupta, S., Chimbara, W., Watkins, S., Crawford, H., Marden, B., Legg, J. and Marsh, M.J. (2001) <a href="#">Acute lung injury in paediatric intensive care: Course and outcome</a> . <i>Critical Care</i> , 5, (1) ( <a href="#">doi:10.1186/cc1301</a> )	1244
Jones, Keith and Mooney, Claire (2003) <a href="#">Making space for geometry in primary mathematics</a> . In: Thompson, I. (ed.) <i>Enhancing primary mathematics teaching</i> . Maidenhead, UK, Open University Press, 3-15.	990
García Sainza, Leonardo and Woodley, David (2006) <a href="#">Recent investigations of the metallic landscapes of Sicily province</a> . <i>Andalusia Dolores de Palacios III</i> . In: Jonsamne, R., Laporte, I. and Scare, C. (eds.) <i>Origine et Développement du Mégalithisme de L'ouest de L'Europe. Actes du Colloque International, 26-30 Octobre 2002, Bougon, France</i> . Niort, France, Conseil Général des Deux-Sèvres, 473-484.	968
Warburton, W.I. (2005) <a href="#">What are grounded theories made of?</a> . In: <i>2005 University of Southampton LASS Faculty Post-graduate Research Conference, Southampton, UK, 6-7 Jun 2005</i> . Southampton, UK, Faculty of Law, Arts and Social Sciences (LASS), 10pp, 1-10.	745
Carnie, Andrew <a href="#">Magic Forest: a time based slide dissolve work</a> .	701
Chambers, Lily D., Stokes, Keith R., Walsh, Frank C. and Wood, Robert J.K. (2006) <a href="#">Modern approaches to marine antifouling coatings</a> . <i>Surface and Coatings Technology</i> , 201, (6), 3642-3652 ( <a href="#">doi:10.1016/j.surfcoat.2006.08.129</a> )	672
Cattle, H. (ed.) (2007) <a href="#">CLIVAR Exchanges - African Monsoon Multidisciplinary Analysis (AMMA)</a> . Southampton, UK, International CLIVAR Project Office, 28pp. (CLIVAR Exchanges, No. 41 (Vol. 12(2)))	636
Doerfl, Dennis (2003) <a href="#">Peace-of-Mind Series Hybrid Electric Vehicle Drivetrain - Transfer Thesis (MPhil/PhD)</a> . Southampton, UK, University of Southampton, School of Engineering Sciences, 118pp.	611
Jenkins, Stephen P., McKeown, John and Schaeff, Sylke V. (2006) <a href="#">Social segregation in secondary schools: how does England compare with other countries?</a> (A06/01)	611
(1998) <a href="#">Report of the FAMOS/PACS Workshop on Field Programmes, University of Sao Paulo, Sao Paulo, Brazil, March 30 - April 2, 1998</a> . Southampton, UK, International CLIVAR Project Office (ICPO Publication Series, 20)	599

# Eprint 'Dashboard'

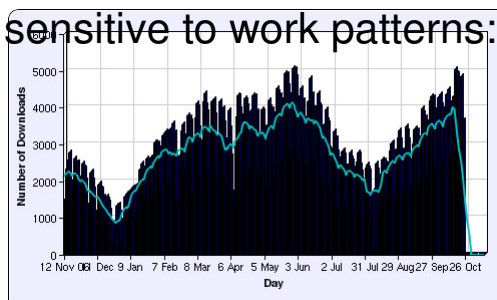


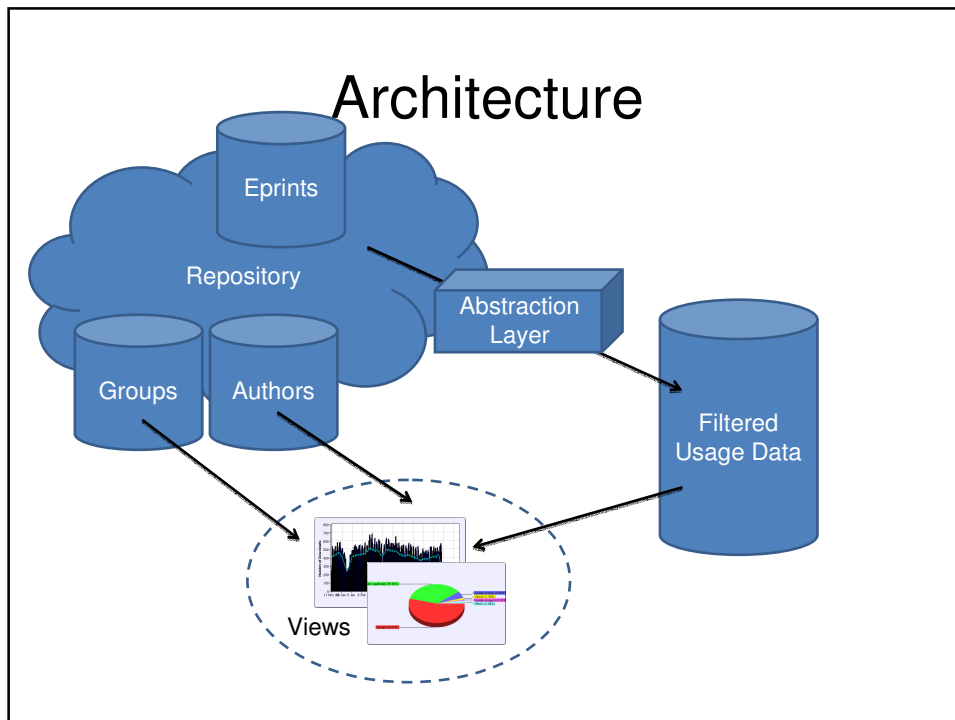




## IRStats (Stats)

- Looked at four repositories:
  - CADAIR (Aberystwyth), Minho, Southampton, UCL (London)
- Average downloads per day 1/3 total fulltexts
- Highly sensitive to work patterns:





## IRStats

- <http://trac.eprints.org/projects/irstats/>

Sorry, not much on citations but

...

- Usage statistics are fun too
- Thanks for listening!

## SWORD Profile: Machine deposit

- Atom Publishing Protocol: Simple Publishing Protocol
- Create a deposit bundle:
  - Metadata in XML
  - Document objects
- HTTP POST the deposit bundle to APP interface
- Object appears in 'inbox' – approve/publish

## SWORD Profile

- Draft implementations for EPrints, DSpace, Fedora ...
- 'Why do I have to submit my work to multiple locations'
- SWORD/APP mechanism
- Metadata compatibility ... ?