

# OPEN ACCESS TO SCIENTIFIC DATA: WHY RESEARCHERS ARE RELUCTANT TO SHARE

Dr Dale Peters  
Academic Computing  
ICT

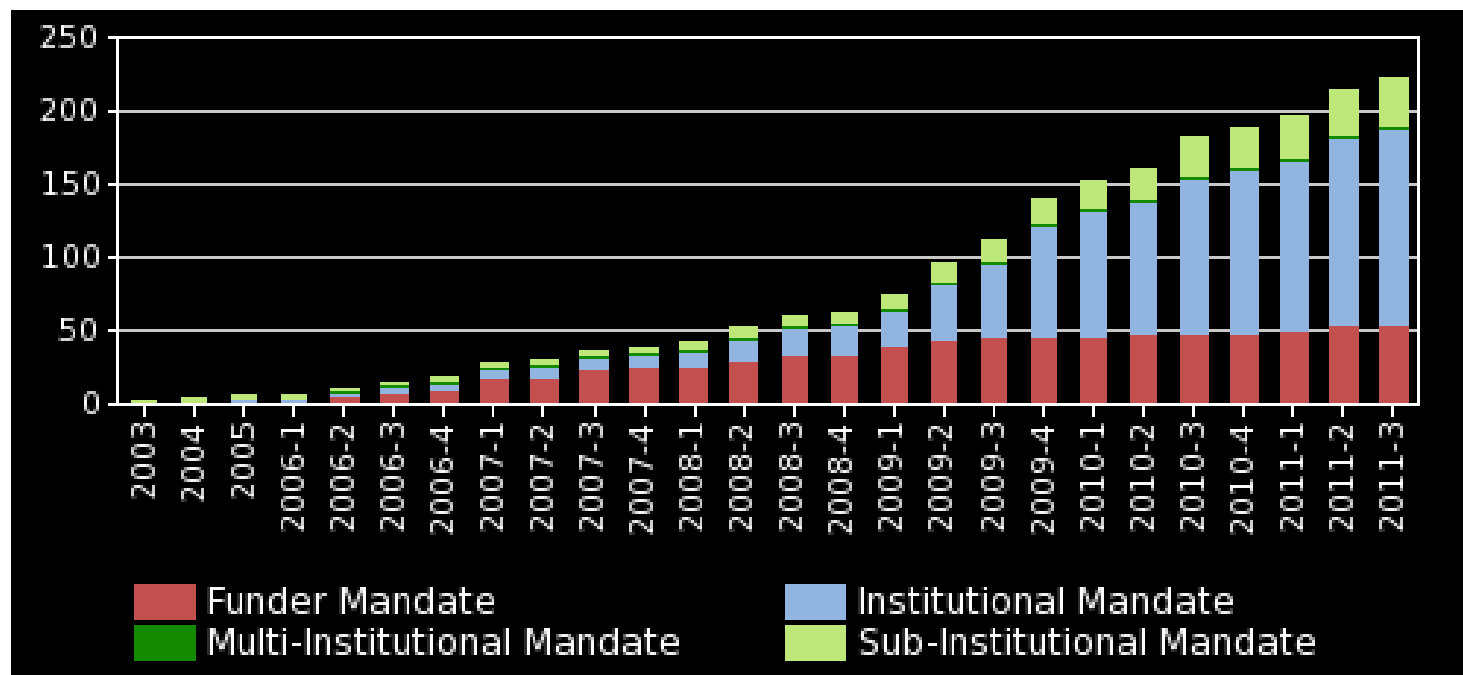


# Outline

- Increasing acceptance of open access policies as a way to enhance global research
- Concerns of academics that encumber the full potential of OA
- Priority action areas for future development of global research data infrastructures

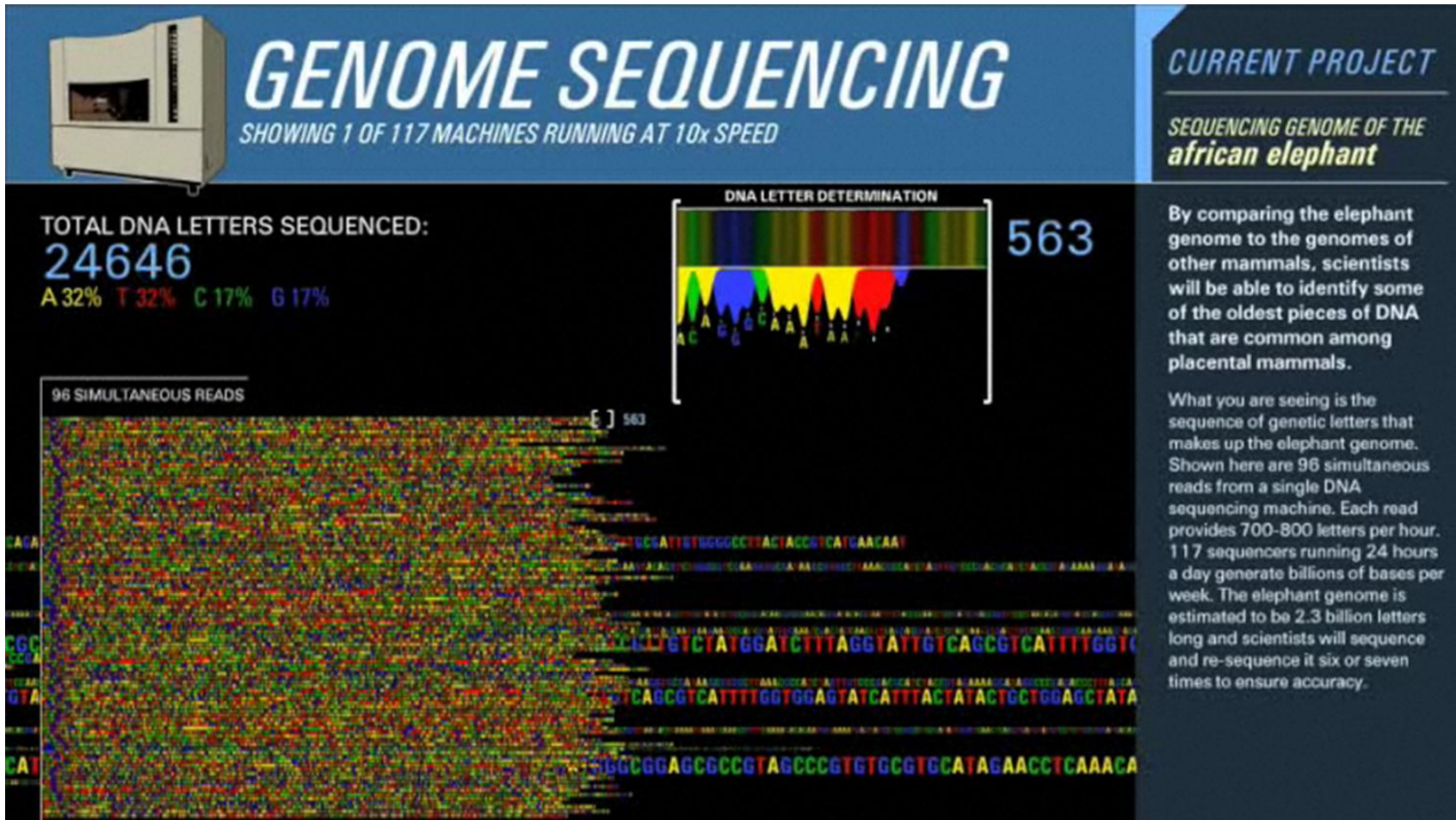


# OA mandates



<http://roarmap.eprints.org/>

# The genome generation

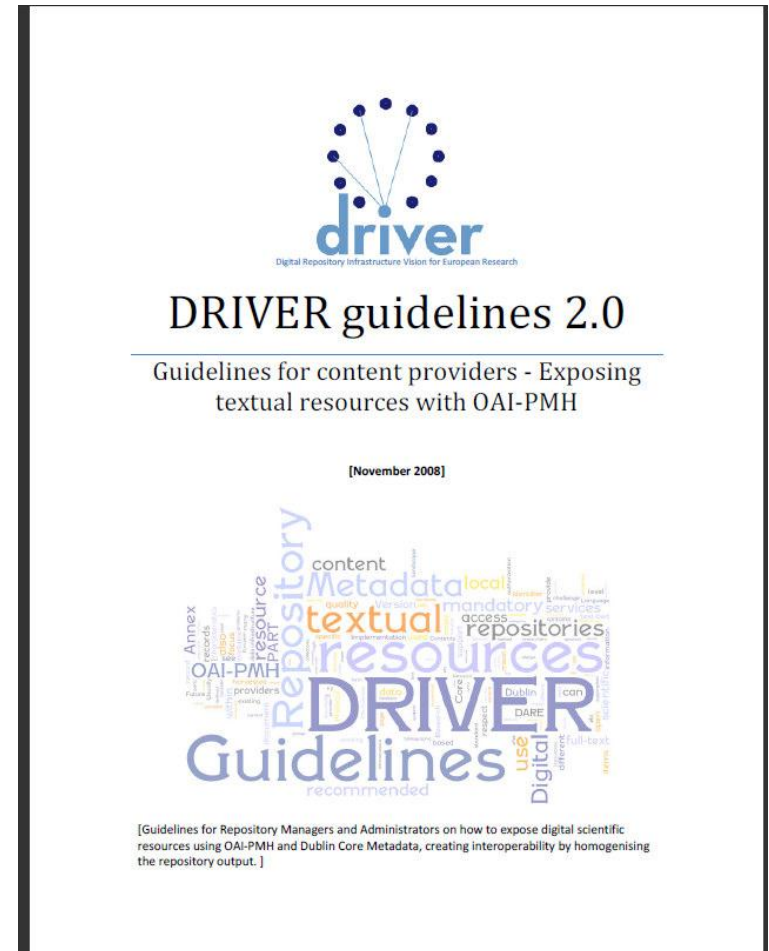


# Concerns of scholars

- Self- archiving an administrative burden
- Few rewards
- Separate from IRMA
- No services on top of repositories
- Usage metrics as quality research indicators
- No legal requirement – yet
- Responsible data management of primary and secondary research data

# Repository Services

- OAI-PMH / OAI-ORE
- Reliable Metadata OAI\_DC
- Best Practices for OAI\_DC
- Compound Object Wrapping
- Vocabularies and Semantics
- Quality Labels
- Long Term Preservation
- Persistent Identifiers
- Usage Statistics Exchange
- Intellectual Property Rights (IPR)



# Usage metrics

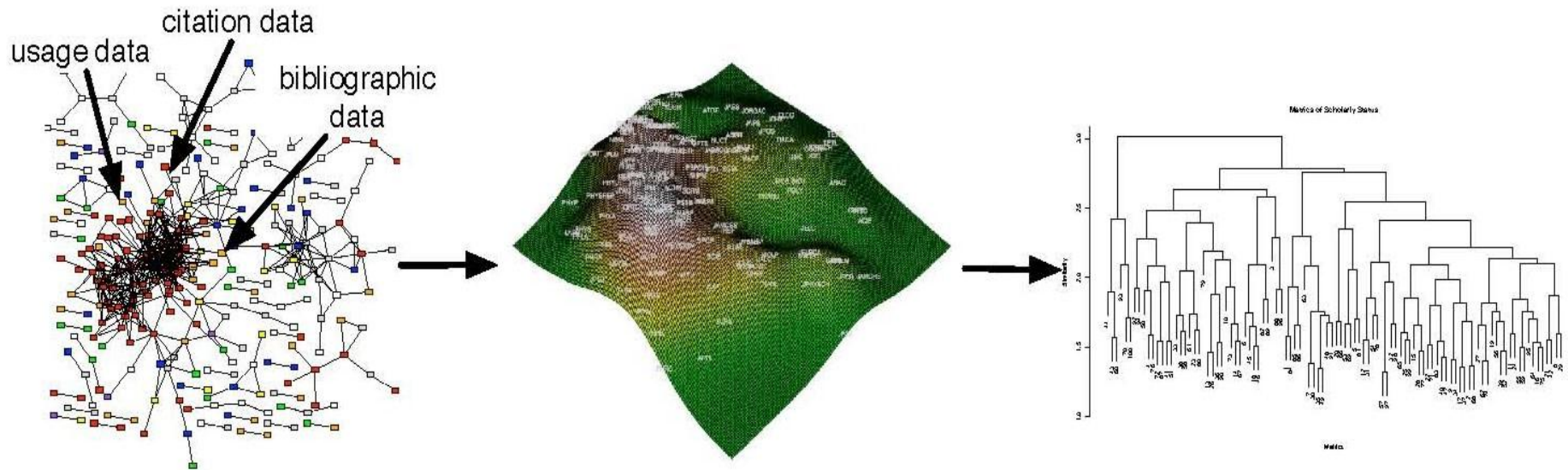
- Repository usage
  - Content ( extent, author, subject, publication type)
  - Hits vs. downloads
  - COUNTER/ PIRUS / SUSHI
- Research impact
  - Journal Impact Factor
  - Citation analysis
- Research assessment
  - Enhance research performance?

citebase Search





# Metrics of scholarly status



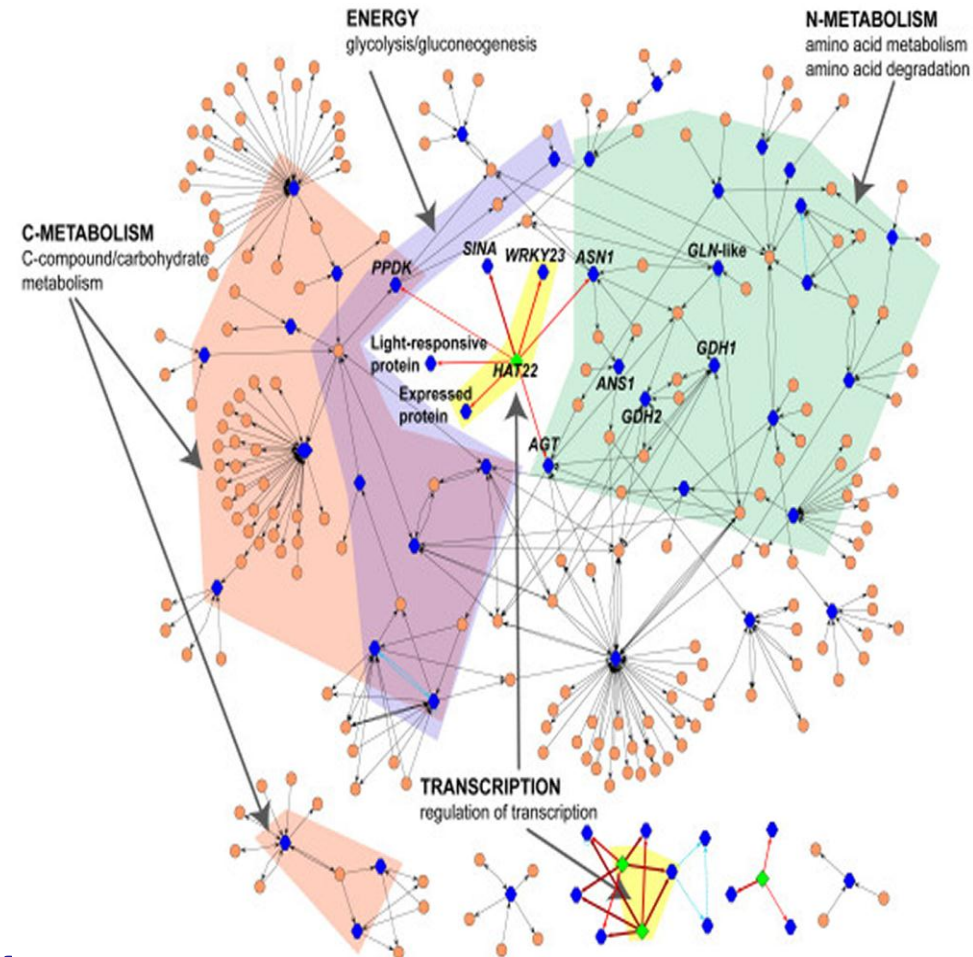
Bollen, J., RODRIGUEZ, M., & vd SOMPEL, H. (2007) MESUR: usage-based metrics of scholarly impact.

# Complexity : what are we citing?

- Journal
- Article
- Workflow
- Data
- Visualisation
- Model
- Annotation
- Concept



## Attribution granularity



# Legal requirement

LEAFLET



EUROPEAN  
COMMISSION

## Open Access Pilot in FP7

*"If I have seen further it is by standing  
on the shoulders of giants."* Isaac Newton, 1676



- Support the diffusion and adoption of the European Commission Open Access mandates
- Focus on publications in seven pilot disciplines
  - including datasets
- Orphan works repository supported by CERN
- Global dissemination



# National Institutes of Health Public Access

*The Public Access Policy ensures that the public has access to the published results of NIH funded research to help advance science and improve human health.*

## Home

- 1. Determine Applicability
- 2. Address Copyright
- 3. Submit paper to PMC
- 4. Include PMCID in Citations
- Identify Submission Method
- Policy Details
- Training/Communications
- Glossary/FAQs
- For Sponsored Programs
- For NIH Employees
  - For NIH Employee Authors
  - For NIH Extramural Staff (NIH Access Only)
- Contact Us

## Overview

The [NIH Public Access Policy](#) ensures that the public has access to the published results of NIH funded research. It requires scientists to submit final peer-reviewed journal manuscripts that arise from NIH funds to the digital archive [PubMed Central](#) upon acceptance for publication. To help advance science and improve human health, the Policy requires that these papers are accessible to the public on PubMed Central no later than 12 months after publication.

## How to Comply

All of your papers that fall under the NIH Public Access Policy, whether in press or in print, must include [evidence of compliance](#) in all of your NIH applications and reports.

1. [Determine Applicability](#)  
Does the NIH Public Access Policy apply to your paper?
2. [Address Copyright](#)  
Ensure your publishing agreement allows the paper to be posted to PubMed Central in accordance with the NIH Public Access Policy.
3. [Submit paper to PMC](#)  
Submit papers to PubMed Central (PMC) and approve public release. Enter your journal name into the box on the right side of the screen to determine how you paper will be posted to PMC.
4. [Include PMCID in Citations](#)  
Include the PMCID at the end of the full citation in your application or report.



# Data management

- Policies and practice
- Creating & organising data
- Storage and access
- Back-up
- Preservation
- Sharing and re-use

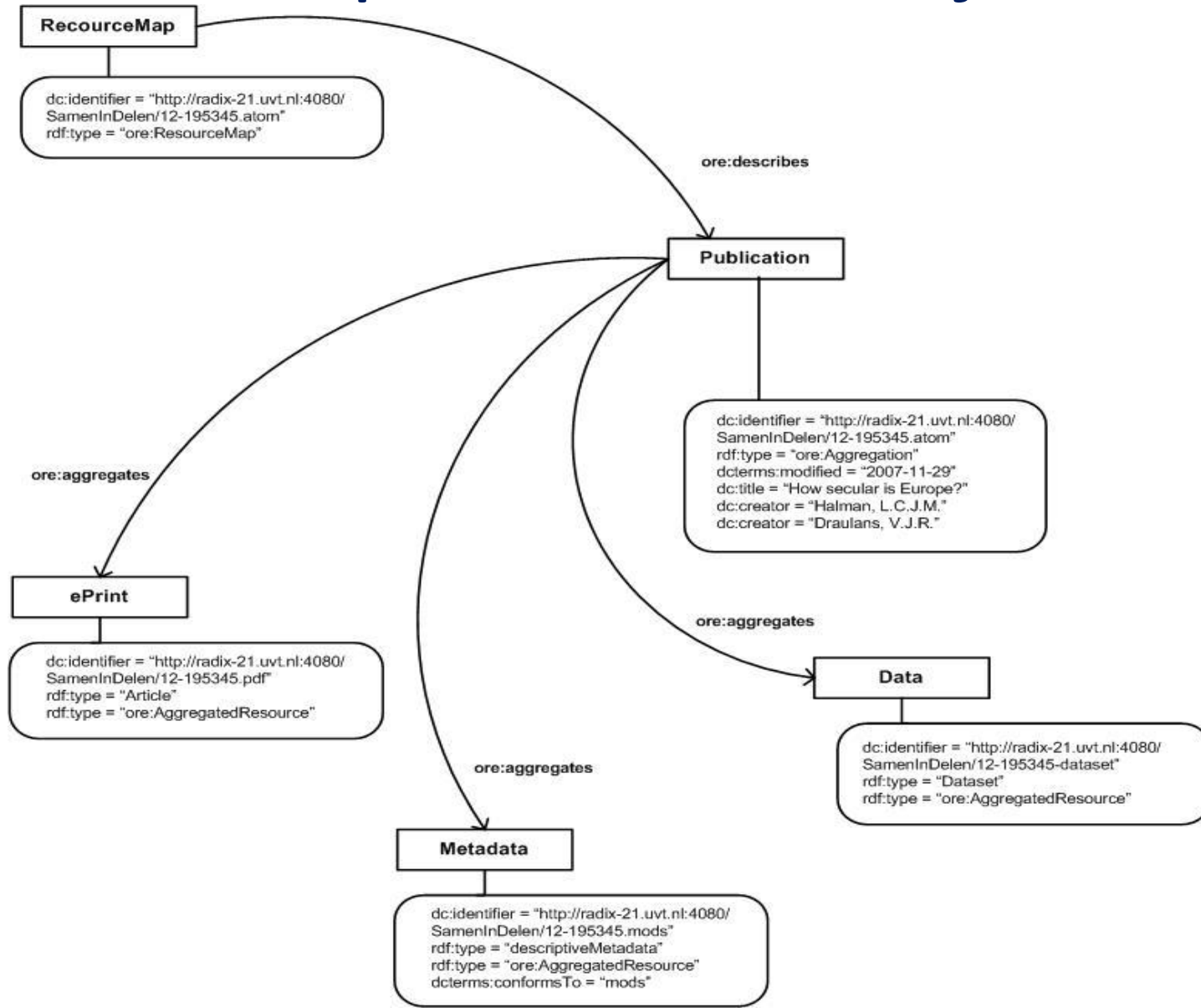


# Enhanced Publications

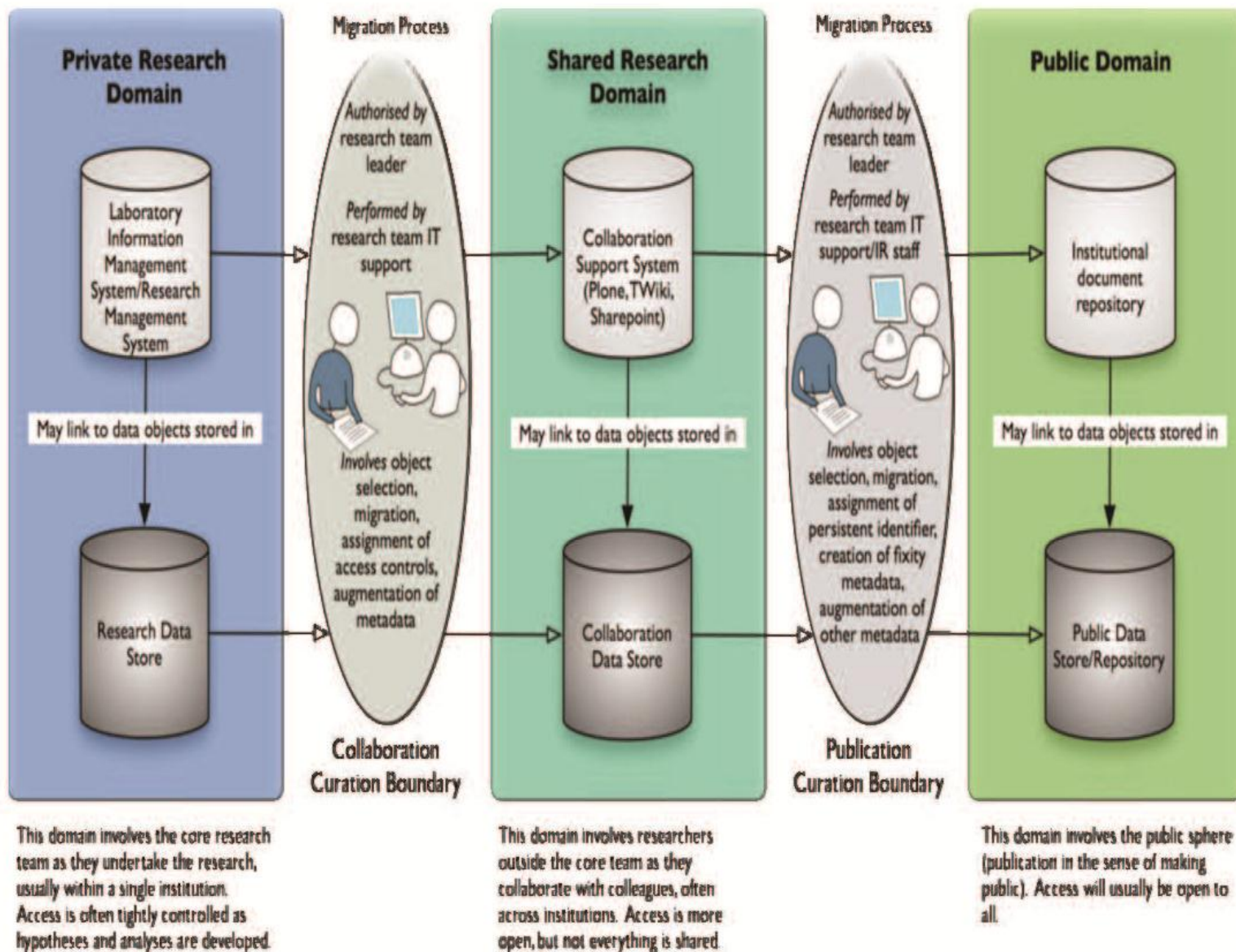
An enhanced publication is a compound data object which combines an “ePrint article ” with:

- one or more metadata records
- one or more data resources
- or any combination of these

# Compound Data Object







*“While many researchers are positive about sharing data in principle, they are almost universally reluctant in practice. ....*

*using these data to publish results before anyone else is the primary way of gaining prestige in nearly all disciplines.”*

*“Data sharing was more readily discussed by early career researchers.”*





Misha Angrist



George Church



Esther Dyson



Jay Flatley



Henry Louis Gates



Rosalynn Gill



Seong-Jin Kim



Greg Lucier



James Lupski



Stephen Quake



Dan Stoicescu



James Watson

They  
have  
shared  
their  
data....



23andMe genetics just got personal.



West Family



PERSONALGENOMES.ORG

Personal  
Genome  
Project



Navigenics®

deCODEme



## Desmond Tutu's genome sequenced as part of genetic diversity study

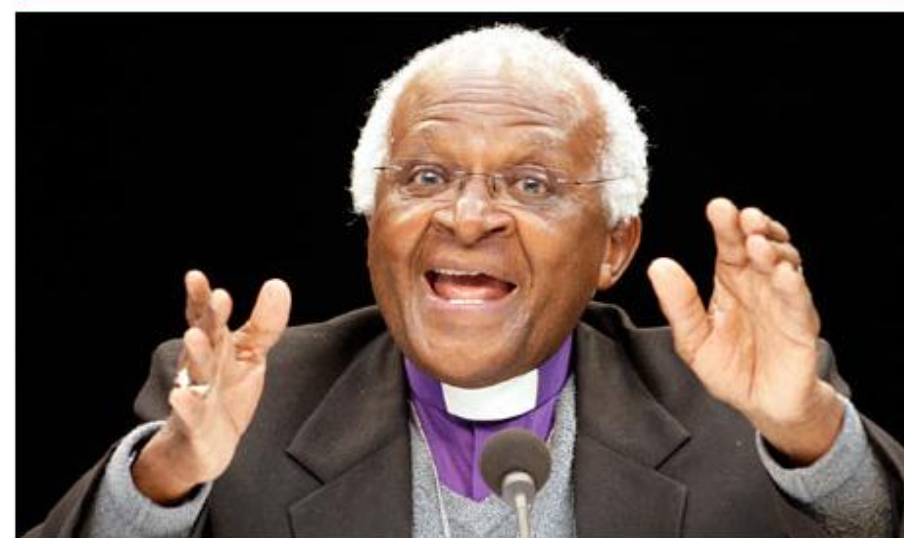
Archbishop Desmond Tutu has had his genome sequenced in research to reveal the true breadth of human genetic diversity

Ian Sample, science correspondent

guardian.co.uk, Wednesday 17 February 2010 18.02 GMT

Article history

guardian.co.uk



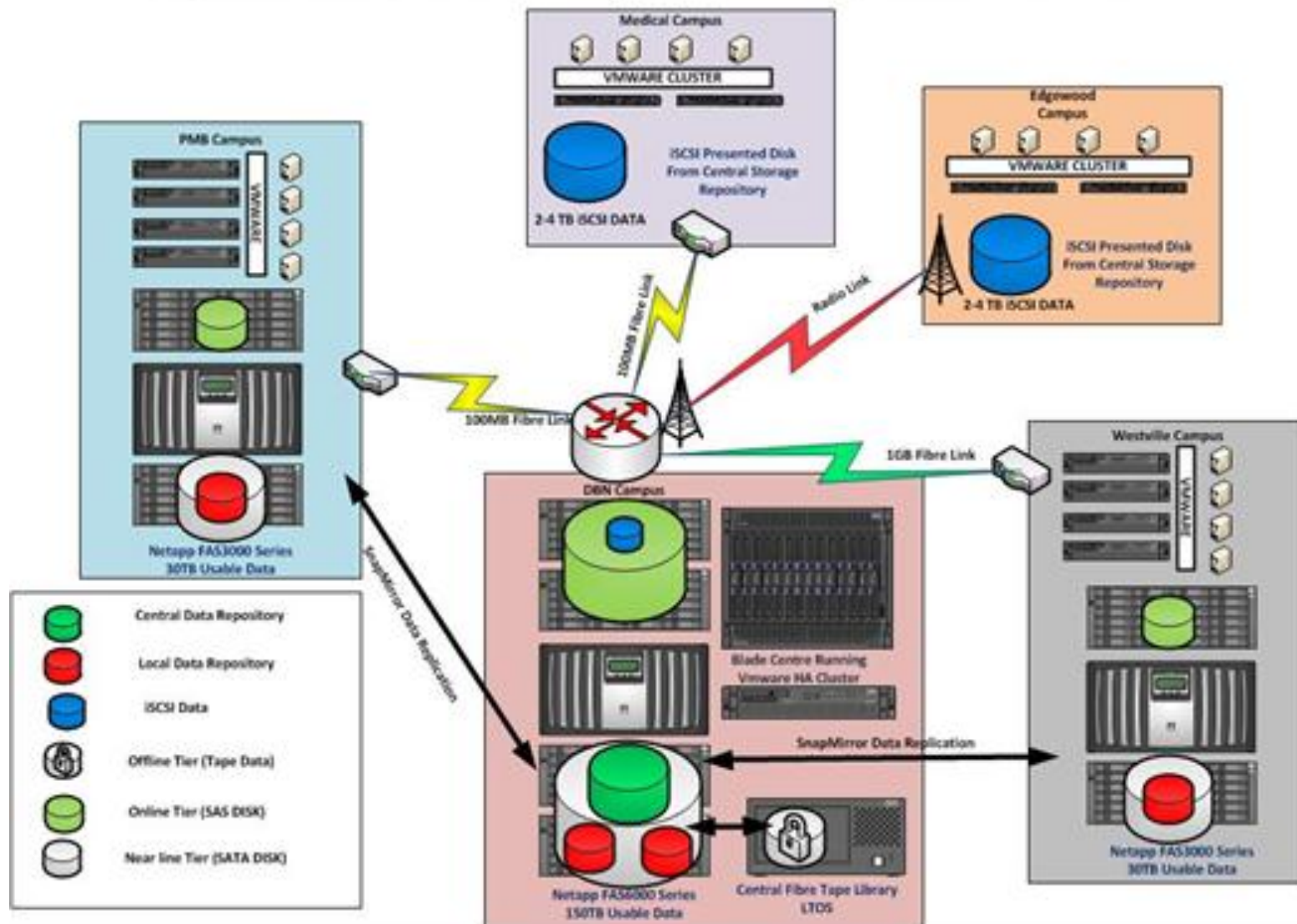
*The majority of people felt that some form of policy or guidance was needed....*

# e-Infrastructure policy @ UKZN

- Research networks
- e-Infrastructure governance
- Authentication, authorisation, access
- Energy management
- Exa-scale computing
- e-Infrastructure services
- Data management



## High Level Over View of Centrally Managed Storage Utilization for UKZN

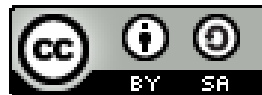


# Priority action areas

- Open Access mandate
- Establish an e-Infrastructure governance model
- Enable technical solutions for open data platforms
- Recognise and promote data ownership, data management and attribution granularity
- Develop research skills in data analysis, data access and data sharing

# References

- Bollen, J., Rodriguez, M., & vd Sompel, H. (2007) MESUR: usage-based metrics of scholarly impact.
- LYON, L. (2010) Mind the gap: reflections on data policies and practice. JISC-CNI Conference, Edinburgh, July 2010
- ROARMAP (Registry of Open Access Repository Material Archiving Policies)  
<http://www.eprints.org/openaccess/policysignup/>
- CHAN, L (et al) .(2009) Open access: a giant leap towards bridging health inequities. Bulletin of the World Health Organization (BLT). <http://www.who.int/bulletin/volumes/87/8/09-064659/en/index.html>
- Registry of Open Access Repositories (ROAR).  
[http://roar.eprints.org/cgi/roar\\_search/advanced?location\\_country=&software=&type=&order=-recordcount%2F-date](http://roar.eprints.org/cgi/roar_search/advanced?location_country=&software=&type=&order=-recordcount%2F-date)
- Manola, N., Manghi, P., Horstmann, W., and Peters, D. An Infrastructure for Managing EC Funded Research Output - The OpenAIRE Project. The Grey Journal, 6 (1), 2010.  
[http://www.greynet.org/images/Contents\\_TGJ.V6.N1.pdf](http://www.greynet.org/images/Contents_TGJ.V6.N1.pdf)
- Open Access Pilot in FP7. [http://www.google.co.za/search?sourceid=navclient&ie=UTF-8&rlz=1T4ADFA\\_enZA373ZA374&q=Open+Access+Pilot+in+FP7](http://www.google.co.za/search?sourceid=navclient&ie=UTF-8&rlz=1T4ADFA_enZA373ZA374&q=Open+Access+Pilot+in+FP7)
- Vernooy-Gerritsen, M. (ed).(2009) Enhanced Publication : linking publications and research data in digital repositories. Amsterdam University Press. <http://dare.uva.nl/aup/en/record/316849>





This page is available in the following languages:

Afrikaans বাংলা Català Dansk Deutsch English English (CA) English (GB) English (Hong Kong) English (Singapore) English (US) Esperanto Castellano Castellano (AR) Español (CL) Castellano (CO) Español (Ecuador) Español (Guatemala) Castellano (MX) Castellano (PE) Euskara Suomi français français (CA) Galego עברית hrvatski Magyar Italiano བོད་སྐད་ Macedonian Melayu Nederlands Norsk Sesotho sa Leboa polski Português română slovenski jezik српски srpski (latinica) Sotho svenska isiZulu



## Attribution-Noncommercial 2.5 South Africa

### You are free:



to Share — to copy, distribute and transmit the work



to Remix — to adapt the work

### Under the following conditions:



**Attribution** — You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).



**Noncommercial** — You may not use this work for commercial purposes.