



## Open access – open for whom, access to what?

Sally Wyatt (sally.wyatt@ehumanities.knaw.nl)  
ASIST European Workshop  
Turku, 5-6 June 2013



## *Technologies might have been otherwise (Bijker & Law, 1992)*

- Alternatives to communicating data in 1980s
- Today's internet – based on TCP/IP (transmission control protocol/internet protocol) – allowing file transfer, email, web, other applications – all developed within academic contexts
- Alternative histories – so maybe alternative futures

***Do artefacts have politics?*** (Winner, 1980)

***Do politics have artefacts?*** (Joerges, 1992)

- Why is bandwidth important for understanding distribution of internet?
- Does it matter where servers are located?
- How do search engines work?
- How are categories constructed? What are the design principles underlying databases?



Design of internet, of digital tools, of databases  
all have political implications

***New technologies & old social forms***  
(Raymond Williams, 1990)

- Scholarly publishing
- Academic reward systems
- Scholarly divisions of labour

## **Users matter!**

Merete Lie & Knut Sørensen (eds) (1996) *Making Technology Our Own? Domesticating Technology into Everyday Life*, Scandinavian University Press.

Nelly Oudshoorn & Trevor Pinch (eds) (2003) *How Users Matter. The Co-construction of Users and Technology*, MIT Press.

## **Non-users also matter**

- Resisters – never used, don't want to
- Rejecters – stopped using voluntarily
- Excluded – never used but would like to
- Expelled – stopped using involuntarily

S Wyatt, G Thomas, T Terranova (2002) 'They came, they surfed, they went back to the beach. Conceptualising use and non-use of the Internet' in S Woolgar (ed) *Virtual Society?* Oxford: OUP

## Extending the principle of symmetry

<b>Bloor on science</b>	<b>Pinch &amp; Bijker on technology</b>	<b>Callon on socio-technology</b>	<b>Wyatt on method in STS</b>
Impartial to statement being true or false	Impartial to machine being success or failure	Impartial to actor being human or non-human	Impartial to actor being identified by actor or analyst
Symmetrical with respect to explaining truth & falsity	Symmetrical with respect to explaining success & failure	Symmetrical with respect to explaining the social world & the technical world	Symmetrical with respect to using concepts from analysts & actors
'Nature' is result & not cause of a statement becoming a true fact	'Working' is result & not cause of machine becoming successful artefact	Distinction between 'social' & 'technical' is result & not cause of stabilisation	'Success' is result & not cause of machine becoming working artefact

first three columns adapted from Bijker (1995: 275)

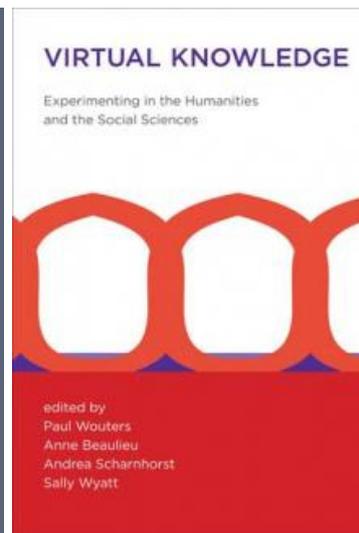
What's in a term?

## Open access to research data

*virtual*  
*cyber-*  
*data-driven*  
*e (electronic)*  
*e (enhanced)*  
*e (executable)*  
*i (interactive)*  
*computer (mediated)*  
*online*  
*distance*  
*tele-*  
*computational*  
*p (personalised)*  
*digital*

science  
 research  
 knowledge  
 scholarship  
 social sciences  
 humanities  
 infrastructure  
 methods  
 tools  
 models  
 objects  
 publications  
 hermeneutics

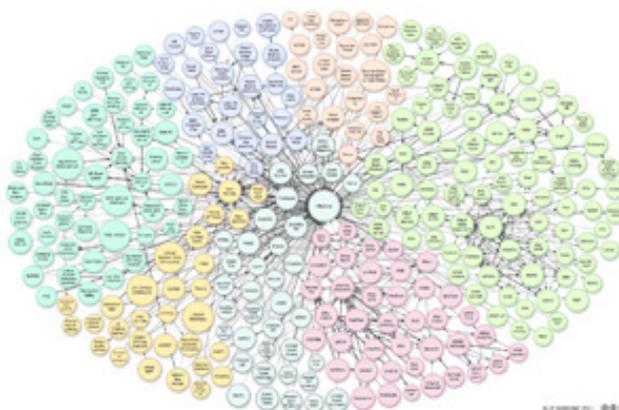
- Always inscribed in & by instruments (e.g. telescopes, microscopes, calculators, computers)
- Deeply social – in contexts of discovery & certainly in contexts of justification & use (e.g. labs, universities, publication practices)
- Mutual influence between systems/ infrastructures of knowledge production & practices of knowledge production



With acknowledgement of influence of  
 Crombie (1994 – *Styles of Scientific  
 Thinking*), Hacking (various), Kwa (2011  
 – *Styles of Knowing*), Radder (1997)

## OPEN

- Access  
Gold vs. Green
- Content
- Source
- Hardware
- Data



Linking Open Data cloud diagram by R Cyganiak & A Jentzsch.  
<http://lod-cloud.net>

## Open access to research data

- Free internet access to and use of publicly-funded (scientific publications and) data (EC, 2012:13)
- Including original scientific research results, raw data & metadata, sources materials, digital representations of pictorial & graphical materials, scholarly multimedia material (Berlin Declaration)
- Numerical/quantitative, descriptive/ qualitative or visual, raw or analyzed, experimental or observational. Examples are digitized primary research data, photographs & images, films, etc. (EC 2012: 45)

## Berlin Declaration, 2003

- Open access is 'a comprehensive source of human knowledge and cultural heritage that has been approved by the scientific community.'
- Internet most important tool for making 'original scientific research results, raw data and metadata, source materials, digital representations of pictorial and graphical materials and scholarly multimedia material.'

## Open data must be:

- Accessible
- Usable
- Assessable
- Able to be evaluated

Royal Society, 2012: 12

## Clearly a good thing

For researchers & for science:

- Data re-use - avoid costly duplication & facilitate complex interdisciplinary enquiry
- Validation of results – quality control
- Validation of results – reduce fraud, integrity
- For teaching purposes
- Increase impact of research (if properly cited by others)

For policy:

- Inform decision making in health, environment, etc
- Inform science policy decision making - altmetrics

## And it just gets better and better

For industry:

- In development of new products & services

For civil society:

- To become informed about important developments
- To participate in scientific debates
- To contribute to scientific knowledge production (citizen science, crowdsourcing, wisdom of crowds, etc)

## So what's the problem?

- 'Accessible', 'usable', 'assessable' & 'able to be evaluated' not so easy in practice:
  - Many stakeholders sometimes with conflicting interests
  - (lack of) availability of infrastructure
  - (international) legal complexities
  - 'Curation' – whose responsibility?
  - Implications for academic careers
  - 'Data gap' (between context of discovery, context of justification, context of application)

## So what is research data? And what is not?

### From RIN (Research Information Network)

- Observational – real-time, irreplaceable
- Experimental – often reproducible, expensive
- Simulation – model & metadata perhaps more important
- Derived/compiled – text & data mining
- Reference datasets

### From OECD

- Factual records – numbers, texts, images & sounds – but NOT lab notebooks, preliminary analyses, drafts, personal communications, physical objects

And what about descriptive, administrative & structural METADATA

## Who pays? Reader or author?

Costs of publishing in the *Journal of Neuroscience*

- \$100 submission fee
- \$850 publication fee
- \$1000 for each colour figure
- \$2500 open access fee (optional)

To say institutions rather than individuals usually pay (e.g. Peter Suber, 2012) misses the point

## Barriers to use of digital resources, even if open, free...

- Lack of awareness of tools, and of the potential of standard software
- Lack of standardisation of databases & archives
- Inadequate annotation tools
- Difficult and unstable access to remote resources
- Lack of institutional training and support
- Irregular use – repeated learning curves

source: M Bulger et al 2011, *Reinventing research? Information practices in the humanities*.  
London: RIN [www.rin.ac.uk](http://www.rin.ac.uk)

## More potential barriers

- Ethics of data sharing
- Skills needed by researchers (training)
- Recognising production & curation of data in evaluations
- Involving new stakeholders – publishers, repositories, research libraries, etc.

## Conclusions & future research

- Defining key terms not so easy
- Practice even harder
- Distributional implications – between individuals, institutions, disciplines, countries
- Lessons from philosophy of science and from STS – data do not travel easily; rational reasons for non-use; look at successful & less successful instances of making data open; pay attention to technical infrastructure

## More information:

- RECODE project: <http://recodeproject.eu>
- eHumanities group: <http://ehumanities.nl>
  
- EC (2012) *Commission Recommendation on access to & preservation of scientific information.*
- Budapest Open Access Initiative (2002)
- Royal Society (2012) *Science as an open enterprise.*
- OECD (2007) *Principles & Guidelines for Access to Research Data from Public Funding*
- Leonelli (2010) Commodification of knowledge exchange In Radder (ed) *Commodification of Academic Research*