

21st Century Research: How e-Research Will Reconfigure Access to Information and Expertise

Bill Dutton and Ralph Schroeder
E-Horizons
James Martin 21st Century School

Outline

- E-Horizons
- Focus on e-Research
- A Transformative Perspective:
Reconfiguring Access
- Case Studies
- Implications for the 21st Century School
and the University of Oxford

E-Horizons

an institute of the 21st Century School

- Aim: Stimulate and inform debate about
the future of communication and
information technologies (ICTs) and
their implications for society
- Initial Focus: e-Science (e-Research) as
a window on future developments
- Strategy: Cross-cutting divisions

E-Horizons

Directors

- Professor Bill Dutton, OII
- Professor Paul Jeffreys, OUCS

Executive Director

- Professor Anne Trefethen, OeRC, Executive Director

James Martin Fellows:

- Dr Ralph Schroeder, OII
- Dr Marina Jirotko, OeRC and ComLab
- Dr Matthijs den Besten, OeRC

Terms, Definitions...

e-Science

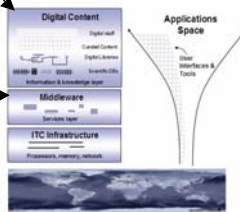
"Large-scale science carried out through distributed collaborations – often leveraging access to large-scale data and computing"
(John Taylor/ National Grid Service, <http://www.nsl.gov.au/e-science> and Taylor, J.M. and e-Science <http://www.e-science.org/>)

e-Infrastructure

"the integrated ICT-based infrastructure" with its key components being networking infrastructure (connectivity, CPU and storage), middleware and organisation (enabling deep integration of individual components across the network, and working collaboratively), and various types of resources (research outputs)
The e-IRG roadmap on Infrastructure

Cyberinfrastructure

"At the heart of the Cyberinfrastructure vision is the development of a cultural community that supports peer-to-peer collaboration and new modes of education based upon broad and open access to leadership computing, data and information resources, online instruments and observatories, and visualization and collaboration services"
NIST's report: Cyberinfrastructure vision for the 21st Century Discovery

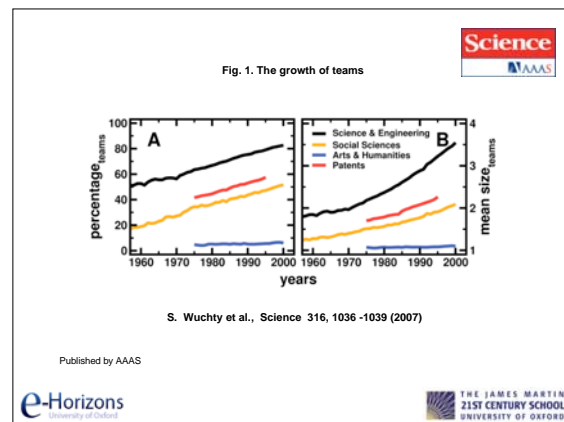


E-Research?

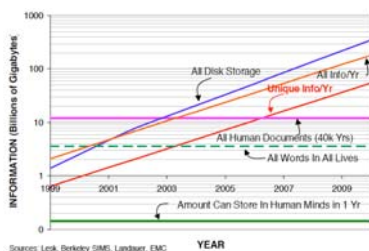
- E-Science, Cyber-Infrastructure
- E-Research: advanced ICT-enabled
research (e.g., sensor networks,
networking tools 2.0, visualization, Web
metrics, modeling and simulation, ...)

'World Wide Science: The Promises, Threats and Realities of e-Research'

- Edited book drawing on e-Horizons and related projects across Oxford
- Covers:
 - Enabling Technologies
 - Applications: Sciences and Humanities
 - Social Shaping of e-Research
 - Institutional and Societal Implications



The Data Deluge



Perspectives on e-Research

- Technological Determinism

Internet and Society

Figure 1-1: A Deterministic Model



Perspectives on e-Research

- Technological Determinism?
- Research Method?
- Infrastructure?
- Reconfiguring Access

Transformative Impact of the Internet

Figure 1-3: Reconfiguring Access

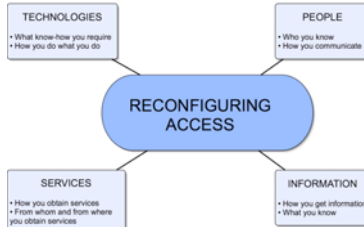
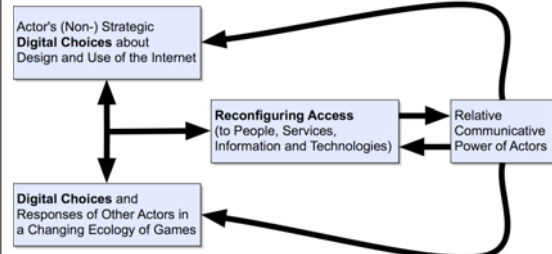


Figure 1-2: Strategies Reconfiguring Access

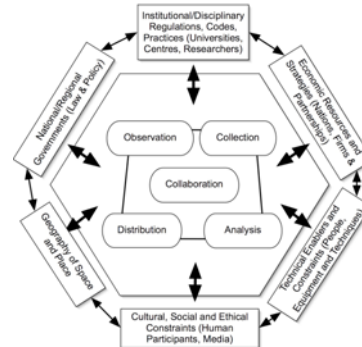


ICTs Reshaping Research

- Collaboration
- Observation
- Collection
- Analysis
- Distribution

'Will e-Research enable scientists to reconfigure the processes and outcomes of discovery – reshaping what researchers can observe, collect, analyze, and distribute? Will researchers collaborate in more powerful ways with teams that cross broader disciplinary, institutional and national boundaries?'

Figure 1-4. Social Shaping of e-Research



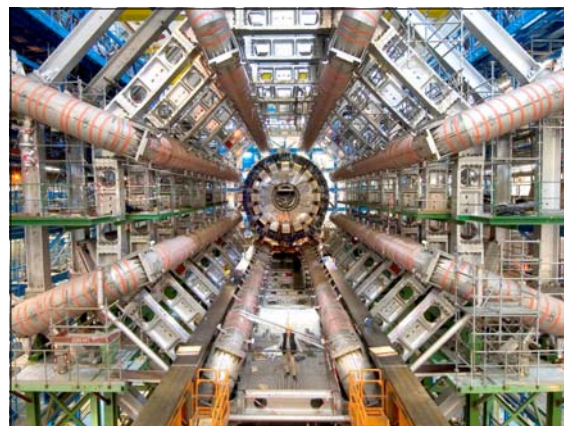
So what?

- Quality of Research
- Nature of Research: Artisan or Knowledge Worker; Embedded or Mediated Observer
- Privacy and Confidentiality
- Ownership, IPR, and Openness
- Distribution of Expertise: Greater Diversity or a Winner-Takes-All?

Cases Illustrating e-Research

Particle Physics to the Humanities

Particle Physics and EGEE: The world's largest e-Science collaboration



Particle Physics and EGEE



- LHC computing grid highly distributed and multi-tiered
- Petabytes of data, 100,000s CPUs
- Memoranda of understanding about the uses of computing resources

Particle Physics and EGEE

- The largest e-Science collaboration worldwide, organizationally and technically
- The Large Hadron Collider, the most powerful particle accelerator
- Searching for Higgs Boson – “1 person in 1000 worlds, or a needle in 20 million haystacks”
- Enabling Grids for E-Science (EGEE): a European Grid moves beyond Europe and beyond physics
- Does the model of physics transfer to other forms of research collaboration?
- Reshapes the nature of collaboration

e-Research in Sweden – New ways of sharing data in the social and health sciences

e-Research in Sweden

- Sweden has a major e-Research initiative
- 'Universal' personal identification
- Uniquely powerful datasets (e.g. twin registry)
- UK (ID cards, NHS) and US parallels?
- Significance: If Swedes can't do it, no one can?
- Future possibilities: public health via mobile phones?

Preventing Flu via Mobile



e-Research in Sweden

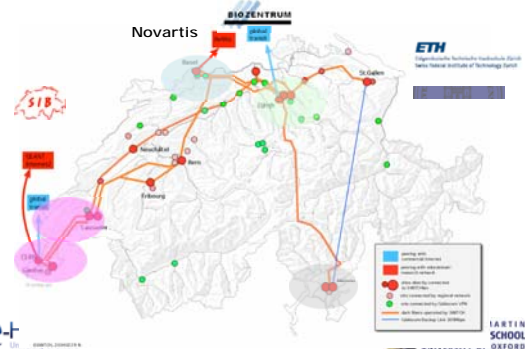
- Use of population data in a 'transparent' society with high trust between people, authorities and researchers...
- ...but, implementation of secure distributed access and 'incidents' creating public concerns
- Reshapes how data are collected

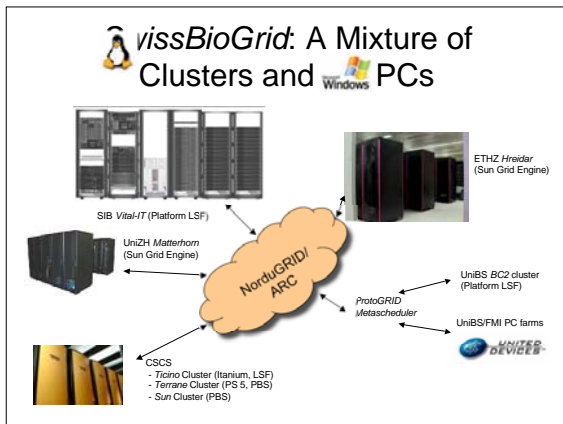
SwissBioGrid - Shared computing power for biomedicine

SwissBioGrid

- Aims: high throughput analysis of proteomics data, virtual screening of possible drugs for dengue fever
- Collaborators: Swiss Institute of Bioinformatics, Novartis, Swiss National Supercomputing Centre
- Using the spare capacity of Linux clusters and PCs

Swiss BioGrid





SwissBioGrid

- Working across the academic – commercial divide
- Demonstrates that PC clusters can usefully be deployed in biomedicine...
- ...but a challenge to embed shared computing resources without a larger national Grid
- Reshapes how data is analysed

e-Horizons University of Oxford

THE JAMES MARTIN 21ST CENTURY SCHOOL UNIVERSITY OF OXFORD

The Virtual Observatory for the Study of Online Networks (VOSON): Tracking Web Presence over Time

e-Horizons University of Oxford

THE JAMES MARTIN 21ST CENTURY SCHOOL UNIVERSITY OF OXFORD

VOSON

- Rob Ackland, Australian National University and Visiting James Martin Research Fellow at the Oxford Internet Institute
- Hyperlink analysis and other tools (web page text analysis)
 - Conducting empirical social science research into online networks
 - Combines web mining, data visualisation, and social network

e-Horizons University of Oxford

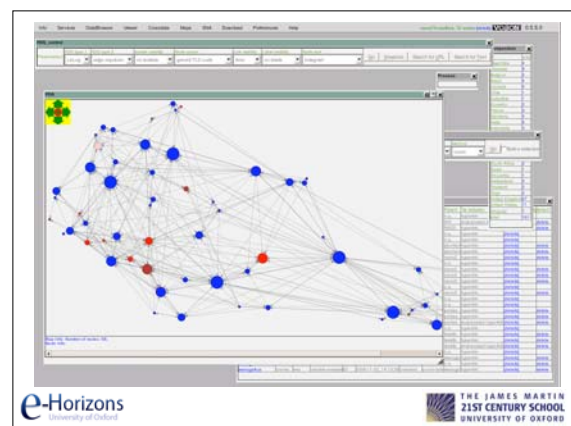
THE JAMES MARTIN 21ST CENTURY SCHOOL UNIVERSITY OF OXFORD

VOSON

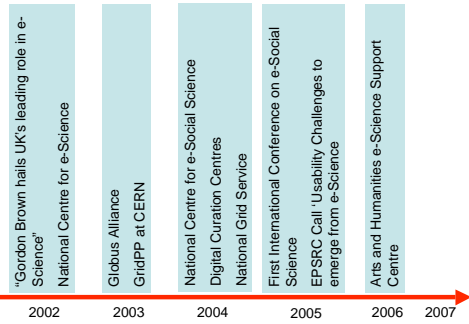
- A study of changes in the UK e-Science web presence
- Use of National e-Science Centre and National Centre for e-Social Science to identify projects
 - Sample consists of 51 projects (42/7) to constitute seed sites
- Use of Internet Archive's (IA) Wayback Machine
 - Historical hyperlink patterns going back to 2002
 - Based on hyperlinks & text from archived top-level pages (and all internal pages linked to directly from the top-level pages)
- A time series of UK web networks from 2002-2007

e-Horizons University of Oxford

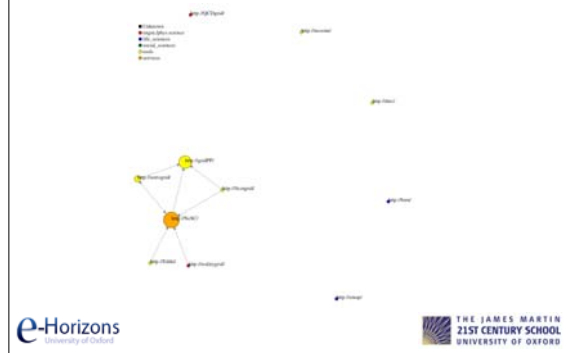
THE JAMES MARTIN 21ST CENTURY SCHOOL UNIVERSITY OF OXFORD



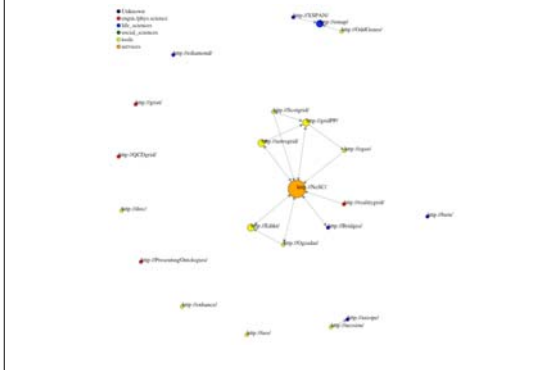
UK e-Research Timeline



UK e-Research Web Network 2002



UK e-Research Web Network 2003



UK e-Research Web Network 2004



UK e-Research Web Network 2005



UK e-Research Web Network 2006



UK e-Research Web Network 2007



UK e-Research Network Patterns

	Seeds - average number of hyperlinks to (from) other sites					
	2002	2003	2004	2005	2006	2007
Disciplinary grouping						
Engin. / phys. sciences	20.0 (0.0)	14.3 (0.0)	18.8 (0.0)	19.3 (0.3)	24.0 (1.0)	31.0 (1.3)
Life sciences	45.5 (0.0)	17.8 (0.7)	18.6 (1.0)	15.1 (1.2)	18.3 (0.9)	24.4 (1.0)
Social sciences	n.a.	n.a.	n.a.	0.0 (0.0)	7.3 (0.1)	17.4 (1.0)
Tool development	23.3 (0.7)	24.2 (0.9)	25.2 (1.1)	30.2 (1.5)	31.9 (1.3)	35.1 (1.6)
Services/infrastructure	25.0 (4.0)	40.0 (8.0)	32.5 (6.0)	48.0 (7.5)	63.0 (10.0)	62.0 (10.0)
All	26.9 (0.7)	21.4 (1.0)	23.3 (1.2)	25.0 (1.6)	25.2 (1.4)	31.0 (1.9)
Network measures						
Network inclusiveness	0.55	0.64	0.73	0.81	0.82	0.97
Network density	0.073	0.048	0.043	0.044	0.038	0.059

VOSON analysis of UK e-Science

- The steady growth of a network
- E-Science and e-Social Science are still distinctive networks
- Interconnectedness as a function of organizational factors, rather than coordination of network around an e-infrastructure
- Reshapes what we can observe, online presence on a global scale

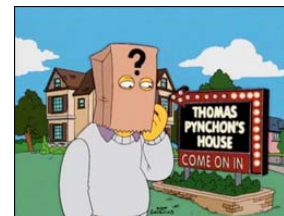
A Collaborative Wiki for Literary Annotation: The Pynchon Wiki

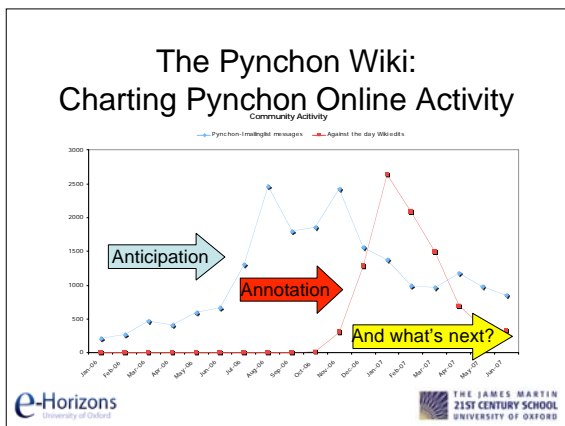
The Pynchon Wiki

- A Wiki for annotating a contemporary American novel
- A 1085 page novel is annotated between November 2006 and early 2007
- The equivalent single author annotation in book form takes longer than a decade
- A flexible, highly motivating, distributed collaborative effort – a model for other forms of online collaboration?

The Pynchon Wiki

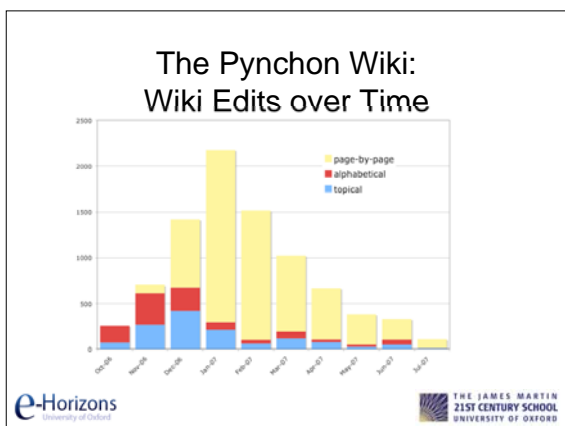
- A notoriously reclusive novelist;
- Author of
 - Gravity's Rainbow, annotated in book form
 - Against the Day, annotated in Wiki form
 - Arcana integral to story-lines





The Pynchon Wiki: Comparison of Annotation efforts by Weisenburger and on the Wiki

Size (no. of words)	Entries (topical+alphabetical + page-by-page)	Contributors	Annotation
162000	904	1(22)	Weisenburger's Gravity's Rainbow Book Form Annotation
455057	120+1358+4067	235	Against the Day Wiki



- ### The Pynchon Wiki
- A race to finish the 'detective work'
 - Encouraging amateur contribution and learning from other contributors
 - A model for self-organized collaboration?
 - 'Finalization' of reference work or endless discussion?
 - Reshaping how scholarly resources are distributed, and how we collaborate

Opportunities for:

- Projects Across the James Martin 21st Century School
- The University of Oxford and 21st Century Research
- e-Research as a strategic objective for the School and the wider University?
- The value of cross-cutting university divisions – building a unique capacity

21st Century Research:

How e-Research Will Reconfigure Access to
Information and Expertise

Bill Dutton and Ralph Schroeder
E-Horizons

James Martin 21st Century School