

The JISC logo is positioned in the top left corner of the slide. It consists of the letters 'JISC' in a bold, orange, sans-serif font. The background of the slide features a close-up of a piece of weathered, light-colored wood with a dark, irregular hole cut into it. The hole reveals a blue, textured surface underneath. The wood grain is clearly visible, and the lighting creates soft shadows and highlights on the wood's surface. The top of the slide has a blue gradient bar.

JISC

inspiring innovation

Open Scholarship: The Web as the Platform for Scientific Communication

Dr Neil Jacobs

Programme Director, JISC

- The **scholarly record** is the set of information that describes the inputs and outputs of academic research and scholarship.
 - It underpins research, scholarship and innovation
- Traditionally the scholarly record has focused on research **works** (books, papers and, increasingly, data) and their **authors**.
 - Newer forms of **work**, such as software, simulations, interactive and dynamic web environments, blogs and tweets...
 - Other forms of **contribution**, including from data managers, but also by facilities and instruments used (and their calibrations), funding sources, host and associated organisations...
- Also attention / use data – this is a dynamic graph...

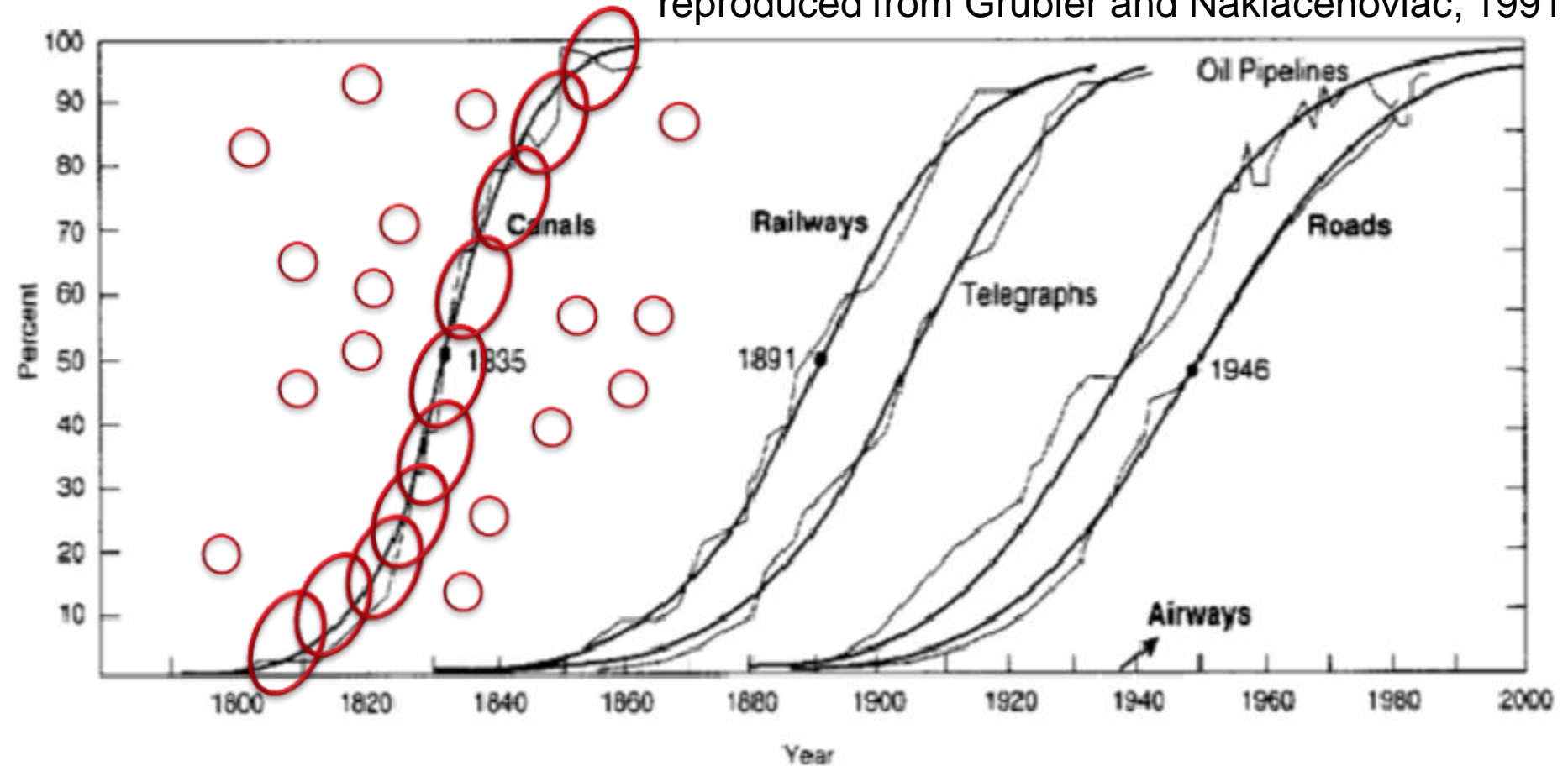
- The **scholarly record** is the set of information that describes the inputs and outputs of academic research and scholarship.
 - It underpins research, scholarship and innovation
- Traditionally the scholarly record has focused on research **works** (books, papers and, increasingly, data) and their **authors**.
 - Newer forms of **work**, such as software, simulations, interactive and dynamic web environments, blogs and tweets...
 - Other forms of **contribution**, including from data managers, but also by facilities and instruments used (and their calibrations), funding sources, host and associated organisations...
- Also attention / use data – this is a dynamic graph...

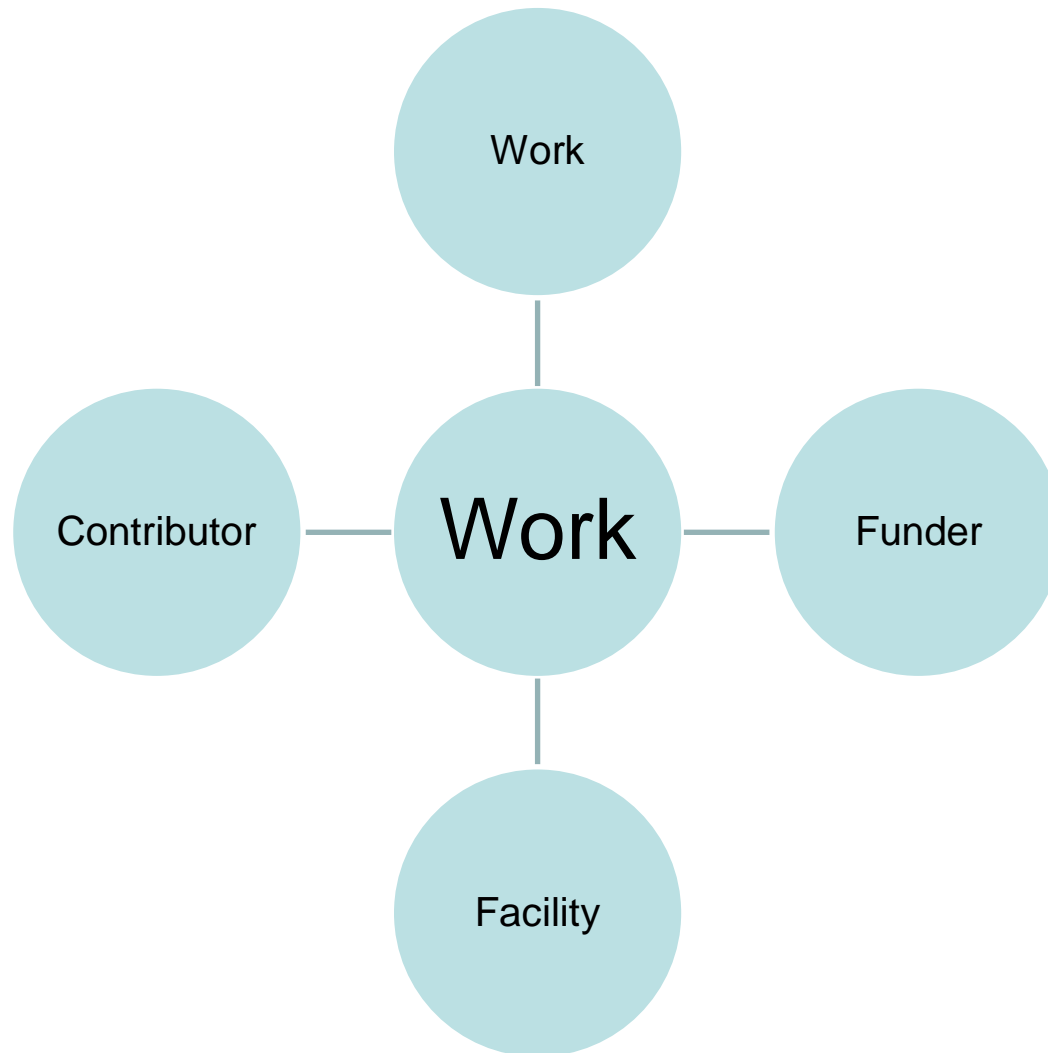
1. The scholarly record is data

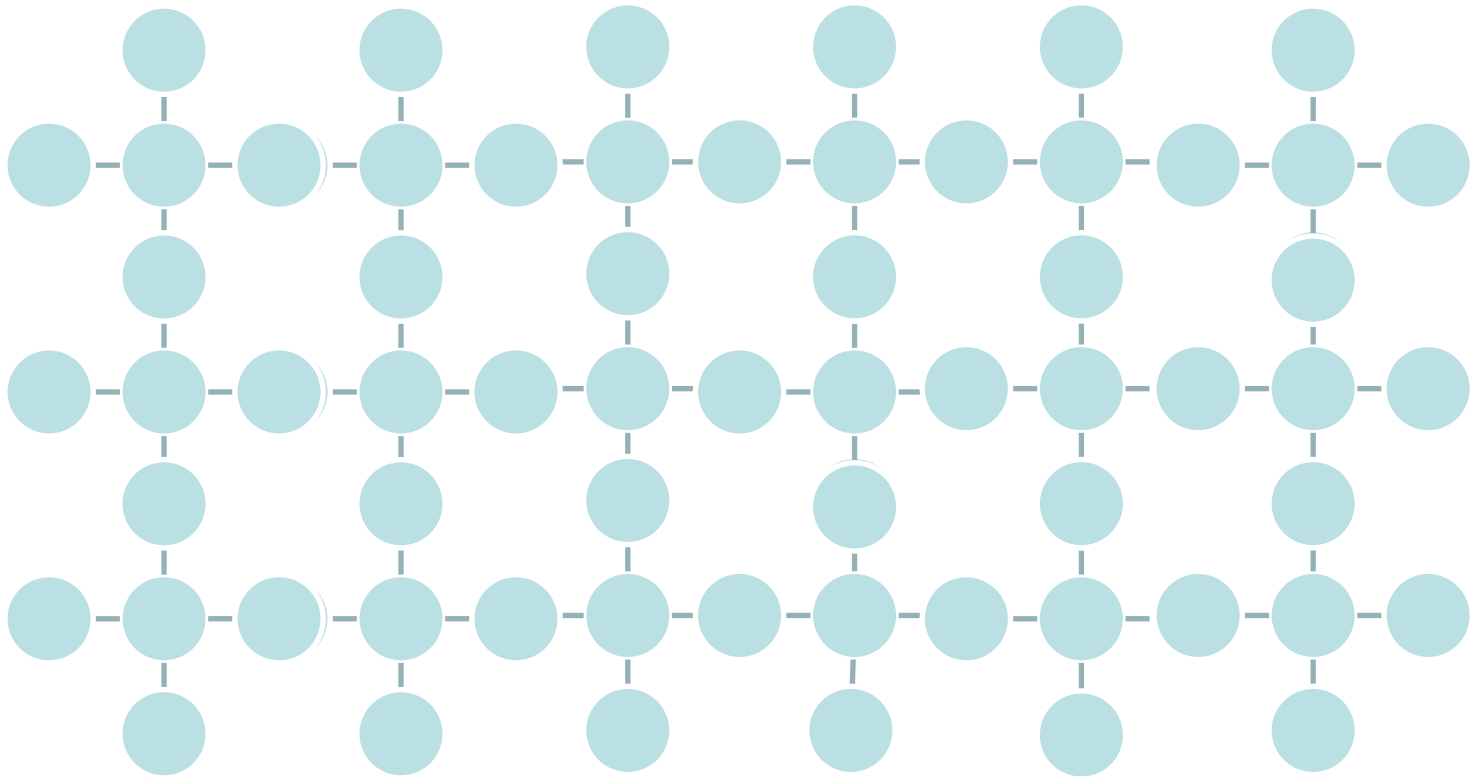
2. For science, data is infrastructure

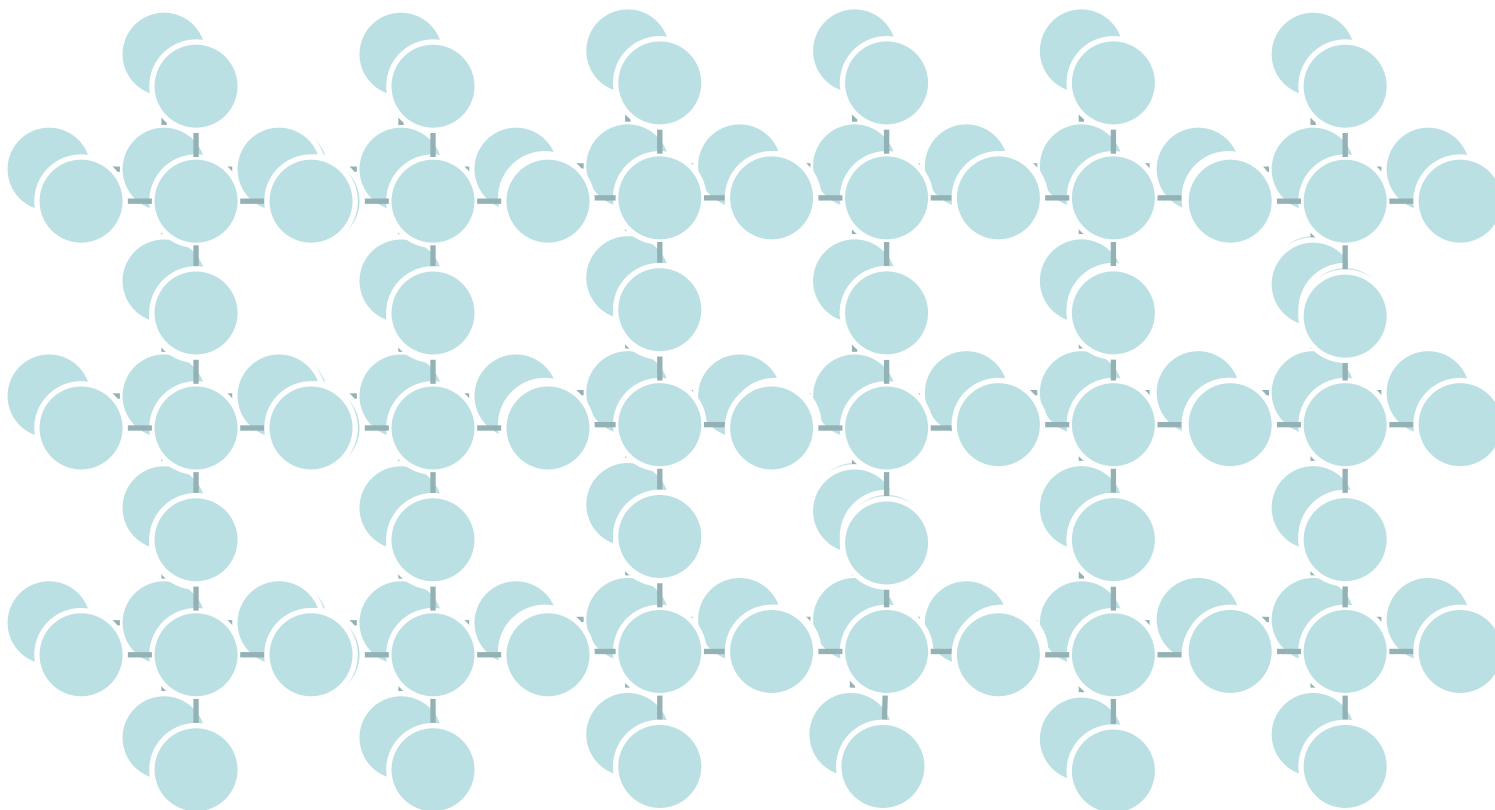
**Infrastructure is evolved over decades not years...
...and it is not “built” as a linear process...**

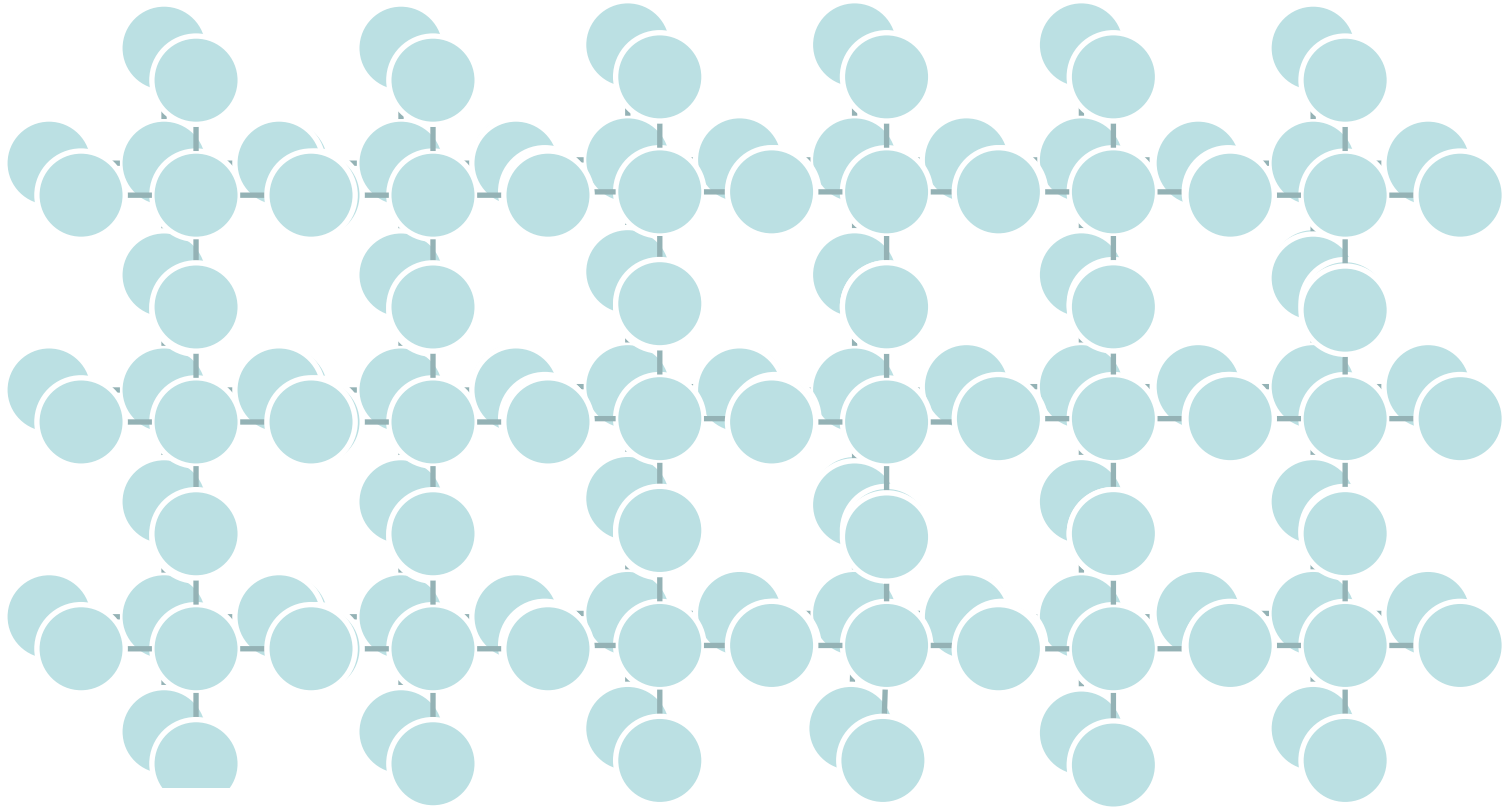
reproduced from Grübler and Nakiâcenoviâc, 1991











- The aim is to have a **scholarly record** that is more
 - **Complete**
 - An adequate basis for research, operational and statistical purposes
 - **Authoritative**
 - Data has provenance, claims are authored, identity is trusted, bad science is excluded
 - **Available**
 - ...to those who need it, when they need it, with the right permissions and cost/benefits
 - **Sustainable**
 - For components and for the system as a whole, includes adequate business models, planned resilience, balance between innovation and stability, cost-effectiveness at all levels, etc

- The aim is to have a **scholarly record** that is more
 - **Complete**
 - An adequate basis for research, operational and statistical purposes
 - **Authoritative**
 - Data has provenance, claims are authored, integrity is trusted, bad science is excluded
 - **Available**
 - ...to those who need it, when they need it, with the right permissions and cost/benefits
 - **Sustainable**
 - For components and for the system as a whole, includes adequate business models, planned resilience, balance between innovation and stability, cost-effectiveness at all levels, etc

Given these desirable attributes...
 how much of the scholarly record
 should be surfaced on the open web?

1. Institutional benchmarking

- Closed: financial information, citation data, some publication data
- Open: some publication data, researchers' names, some research assessment outcomes, some grant information

1. Institutional benchmarking

- Closed: financial information, citation data, some publication data
- Open: some publication data, researchers' names, some research assessment outcomes, some grant information

Some information is closed because of certain business models, but is this sustainable in the wider sense, and does it compromise other desirable attributes of the scholarly record?

1. Institutional benchmarking

- Closed: financial information, citation data, some publication data
- Open: some publication data, researchers' names, some research assessment outcomes, some grant information

Some information is closed because of certain business models, but is this sustainable in the wider sense, and does it compromise other desirable attributes of the scholarly record?

2. Peer review (PLoS-One model)

- Closed: assessment of scientific integrity
- Open: assessment of importance and relevance

1. Institutional benchmarking

- Closed: financial information, citation data, some publication data
- Open: some publication data, researchers' names, some research assessment outcomes, some grant information

Some information is closed because of certain business models, but is this sustainable in the wider sense, and does it compromise other desirable attributes of the scholarly record?

2. Peer review (PLoS-One model)

- Closed: assessment of scientific integrity
- Open: assessment of importance and relevance

Information closed to ensure the scholarly record is authoritative, but open to ensure it is complete and available.

1. Institutional benchmarking

- Closed: financial information, citation data, some publication data
- Open: some publication data, researchers' names, some research assessment outcomes, some grant information

Some information is closed because of certain business models, but is this sustainable in the wider sense, and does it compromise other desirable attributes of the scholarly record?

2. Peer review (PLoS-One model)

- Closed: assessment of scientific integrity
- Open: assessment of importance and relevance

Information closed to ensure the scholarly record is authoritative, but open to ensure it is complete and available.

**I will argue for more "open".
But these kinds of questions will often
arise as the infrastructure grows**

1. Enables reuse

- easier discovery and access, clearer and more permissive rights position, easier curation (so greater longevity and so reuse over time), ability to recombine material (new types of reuse)

2. Supports innovation and agility

- promotes change and enables organisations to respond well to change, innovation in technologies, business models, etc

3. Increases cost-effectiveness

- because it enables collective / shared approaches, reduces duplication of effort, removes friction from transactions, saves time, supports collaboration

4. Improves quality

- by ensuring visibility of material, uses peoples' concern with their reputation to see better quality work shared, supports review of that work (and the reviews themselves are subject to the same incentives, so leading to a “virtuous circle”)

5. Enables better risk management

- including easier compliance with legal/regulatory requirements

Putting the scholarly record on the web:

- A – Open Access
- B – Open Bibliography
- C – Open Citation
- D – Open Data

Implies shifting boundaries, responsibilities, rights, etc across the scholarly record and associated value chains.

Dr Leslie Carr

Senior Lecturer in
Intelligence, Agents,
Multimedia
University of
Southampton

Open Access Vision:
More entwined
international scholarly
teams working together.



- Reuse
 - Innovation
 - Cost-effective
 - Quality
 - Risk
- New kinds of reuse
 - eg text-mining:
Neurocommons,
UKPubMed Central...

- Reuse
 - Innovation
 - Cost-effective
 - Quality
 - Risk
- New business models
 - APCs, PLoS-One, Bloomsbury Academic, submission charging, Faculty of 1000...

- Reuse
 - Innovation
 - Cost-effective
 - Quality
 - Risk
- Need to:
 - Manage a transition
 - Explore how OA benefits the wider economy and society

- Reuse
- Innovation
- Cost-effective
- Quality
- Risk

[Dr Peter Murray Rust](#)

Department of Chemistry
University of Cambridge

**Vision for Open
Bibliography:** A
comprehensive map of the
scholarly world.



- Reuse
- Innovation
- Cost-effective
- Quality
- Risk



- “Discovery” programme, collecting evidence...
 - OpenBib, 2m records from the BL and Cambridge open as linked data
 - Building aggregations
 - Piloting services...

Dr David Shotton

University Reader in Image
Bioinformatics,
University of Oxford

Vision for Open Citation:
Quality assurance and
awareness of key ideas.



■ Reuse

■ Innovation

■ Cost-effective

■ Quality

■ Risk



Open Citation: why and how

■ Citation as-is:

- DOARC, Oldenburg
- Repository citation sharing, Southampton

■ Semantic:

- OpenCit, Oxford

■ Data:

- Datacite, etc

■ Web:

- Webtracks, UK
- CAPret, MIT

- Reuse
- Innovation
- Cost-effective
- Quality
- Risk



Open Citation: why and how

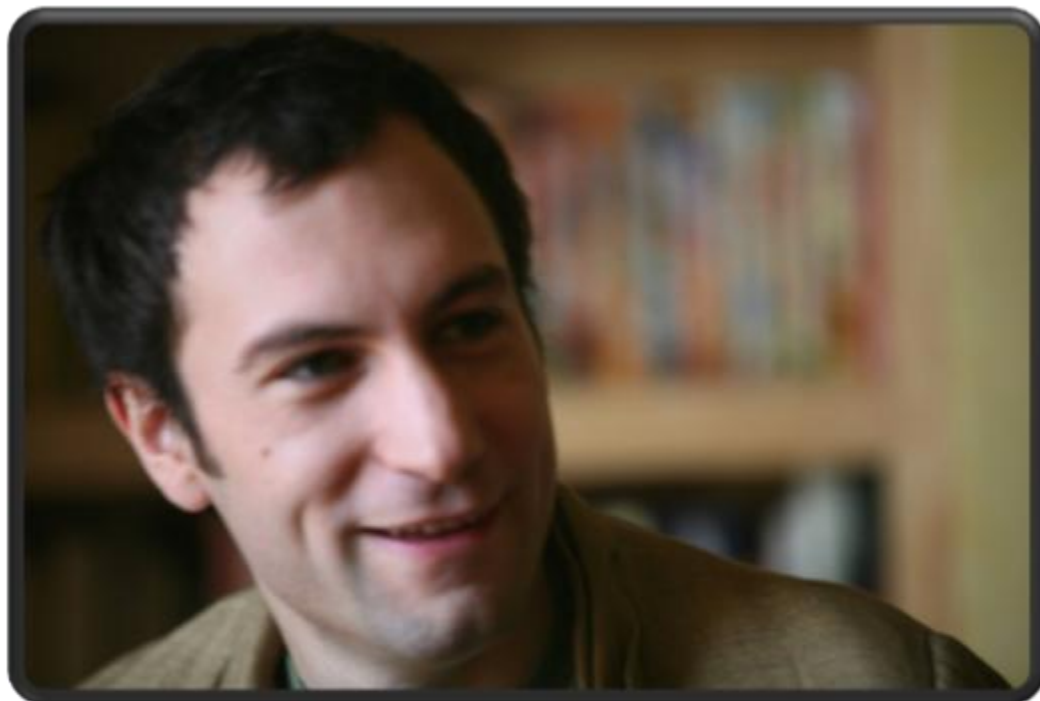
- Citation as-is:
 - DOARC, Oldenburg
 - Repository citation sharing, Southampton
- Semantic:
 - OpenCit, Oxford
- Data:
 - Datacite, etc
- Web:
 - Webtracks, UK
 - CAPret, MIT

Dr Rufus Pollock

Fellowship for the
Shuttleworth Foundation,
Open Knowledge Foundation

Vision for Open Datasets:

Reduce tedium to allow for
more time spent on analysis
and hypothesis.



- Reuse

“Overall, usage of data centres is high, with most centres supporting thousands of researchers and millions of downloads each year.”

- Innovation

“mixed evidence about the importance of data centres in stimulating new research questions”

- Cost-effective

- Quality

“Data centres make research quicker, easier and cheaper, and ensure that work is not repeated unnecessarily.”

- Risk

“Research quality was another important benefit, although not rated quite as highly as efficiency.”

<http://www.rin.ac.uk/data-centres>

- Reuse

“Overall, usage of data centres is high, with most centres supporting thousands of researchers and millions of downloads each year.”

- Innovation

“mixed evidence about the importance of data centres in stimulating new research questions”

- Cost-effective

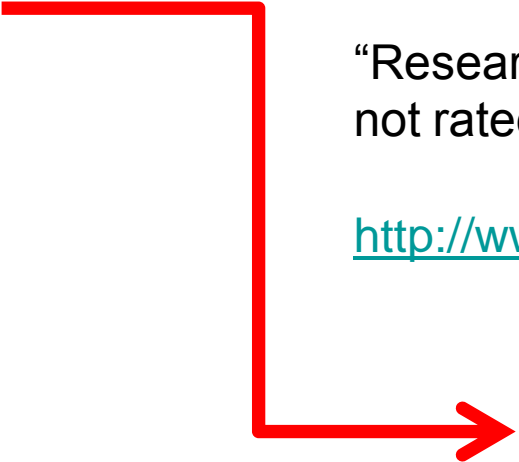
- Quality

“Data centres make research quicker, easier and cheaper, and ensure that work is not repeated unnecessarily.”

- Risk

“Research quality was another important benefit, although not rated quite as highly as efficiency.”

<http://www.rin.ac.uk/data-centres>



“climate scientists should take steps to make available all the data that support their work (including raw data) and full methodological workings (including the computer codes)”.

JISC work:

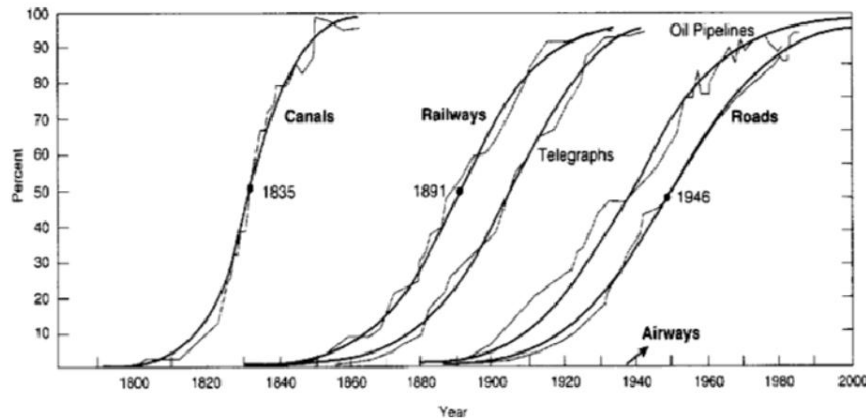
- Data infrastructure for universities (technical and organisational)
- National data infrastructure, data centres, etc
- Shared services (Data Management Planning tool, registry, perhaps “RoMEO for data”?)...
- Data citation projects, data publication projects (Dryad-UK, Datacite..)

Future

- Sim4RDM – Sharing lessons across Europe

Evolution, not revolution

No-one now knows what the map will look like



We only have pointers:

- Toward a scholarly record (= data) that is more complete, authoritative, available and sustainable
- Benefits in reuse, innovation, cost-effectiveness, quality, risk

Thank you.

n.jacobs@jisc.ac.uk