

# Open Research Data

---

DANIELA DUCA

26/12/2014



# Summary

---

What is research and open data?

Why research data should be open?

Barriers to opening data

Reports and initiatives

Management of open research data

# Research and open data

---

# Research Data

---

EPSRC:

- *Research data is defined as recorded factual material commonly retained by and accepted in the scientific community as necessary to validate research findings; although the majority of such data is created in digital format, all research data is included irrespective of the format in which it is created.*



EC Consultation (2013):

- *From the perspective of researchers, research data includes **all data from an experiment, study or measurement**, including metadata and details on processing data. For publishers, data linked to publications is part of the publication.*



# Open Data

---

Creative Commons:

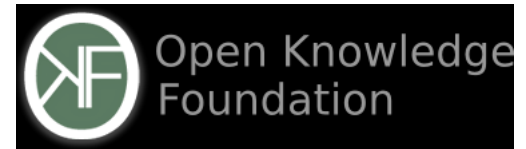
- *we believe scientific data should be freely available to everyone. We call this idea **Open Data**.*

Open Data Institute:

- ***Open data** is data that is made available by organisations, businesses and individuals for anyone to access, use and share.*

Open Knowledge Foundation:

- ***Open knowledge** is what **open data** becomes when it's useful, usable and used.*



# Why should research data be open?

---

# Advantages

---

Accelerates knowledge

Improves transparency

Releases value: your noise can be my signal

Encourage correctness

# Enablers

---

Technology revolution =>

- Capacity to acquire, store and manipulate data
- Capacity to instantaneously distribute information globally, irrespective of location

Funder principles for open research data

Open Data Certificate from ODI

Open Data Handbook from OKF (for more than just research data)



# Principles of Open Research Data

---

US/UK Open Data Summit (Washington DC, 2014)

<http://www.researchinfonet.org/wp-content/uploads/2014/07/Note-of-Washington-meeting.pdf>

- **discoverable** – readily found by online search
- **accessible** – when discovered, can be interrogated
- **intelligible** – can be understood
- **assessable** – for provenance and reliability
- **re-useable** – can be re-used and re-combined with other data as a private preserve

# Principles of Open Research Data

---

Intelligent openness for data from publicly or charitable-funded research

Data citation

Cost should be part of the project

Justifiable limits to openness for data that would be used for commercial properties

Existing processes that prevent sharing should be reformed

# EC's consultation on Open Research Data

---

<http://ec.europa.eu/digital-agenda/en/results-consultation-open-research-data>

Multiple types of research data

Need for Research Data Management Plans

Research data is a public good

# Barriers to opening research data

---

# Barriers

---

Poor infrastructure	Loss
Complexity	Ownership issues
Bureaucracy	Concealment
Destruction	Procrastination

# When to limit openness

---

**Privacy:** biggest issue is with ‘sensitive data’, or what has been called the special category of personal data in EU law

**Commercial use:** intellectual property rights

# Reports and initiatives

---

# Riding the Wave: 2010

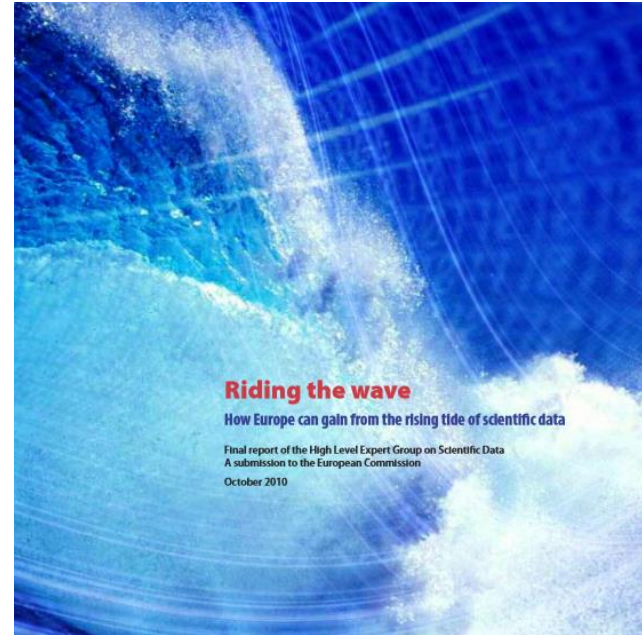
---

How Europe can gain from the rising tide of scientific data

Unlock full value of scientific data

Collaborative data infrastructure

Enormous potential benefits but similar costs





# Surfboard for riding the wave: 2011

---

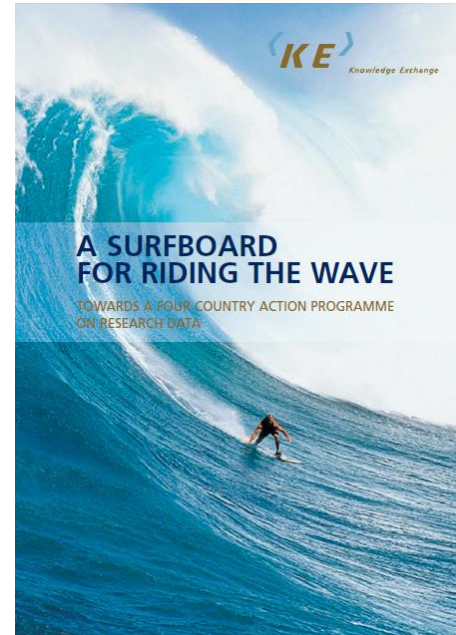
Knowledge Exchange response to EC report

Research data current situation in the 5 countries

Broad outlines for possible actions and a collaborative infrastructure

4 key drivers:

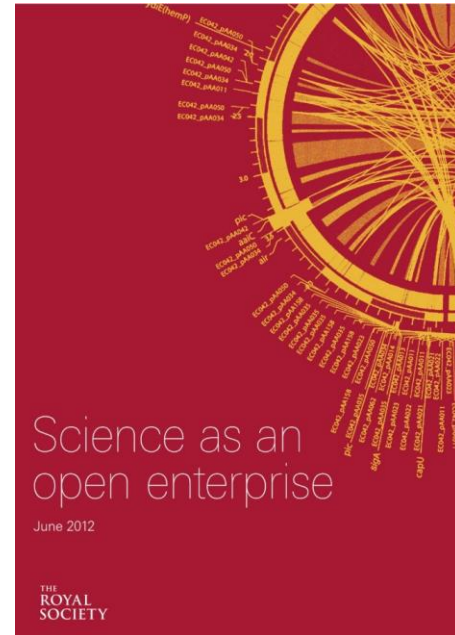
- Incentives
- Training
- Infrastructure
- Funding



# Science as an Open Enterprise: 2012

## The Royal Society UK

- Data is not a private preserve
- Give credit for useful data communication and novel collaborations
- Common standards
- Intelligent openness
- Network of support data scientists
- New software tools for creation and exploitation of data



# Open Data Dialogue: 2012

RCUK, Jisc, RS, Sciencewise-ERC

- Insight on business issues
- Build on prior work
- Engage people meaningfully



# Sowing the seed: 2014

## Knowledge Exchange

- Need for funding
- Promote reuse
- Clear data sharing expectations
- Formally recognise and value data



# Research Funder Policies

---

## RCUK common principles on data policy

- » **Public good:** Publicly funded research data are produced in the public interest should be made openly available with as few restrictions as possible
- » **Planning for preservation:** Institutional and project specific data management policies and plans needed to ensure valued data remains usable
- » **Discovery:** Metadata should be available and discoverable; Published results should indicate how to access supporting data
- » **Confidentiality:** Research organisation policies and practices to ensure legal, ethical and commercial constraints assessed; research process should not be damaged by inappropriate release
- » **First use:** Provision for a period of exclusive use, to enable research teams to publish results
- » **Recognition:** Data users should acknowledge data sources and terms & conditions of access
- » **Public funding:** Use of public funds for RDM infrastructure is appropriate and must be efficient and cost-effective

<http://www.rcuk.ac.uk/research/datapolicy/>

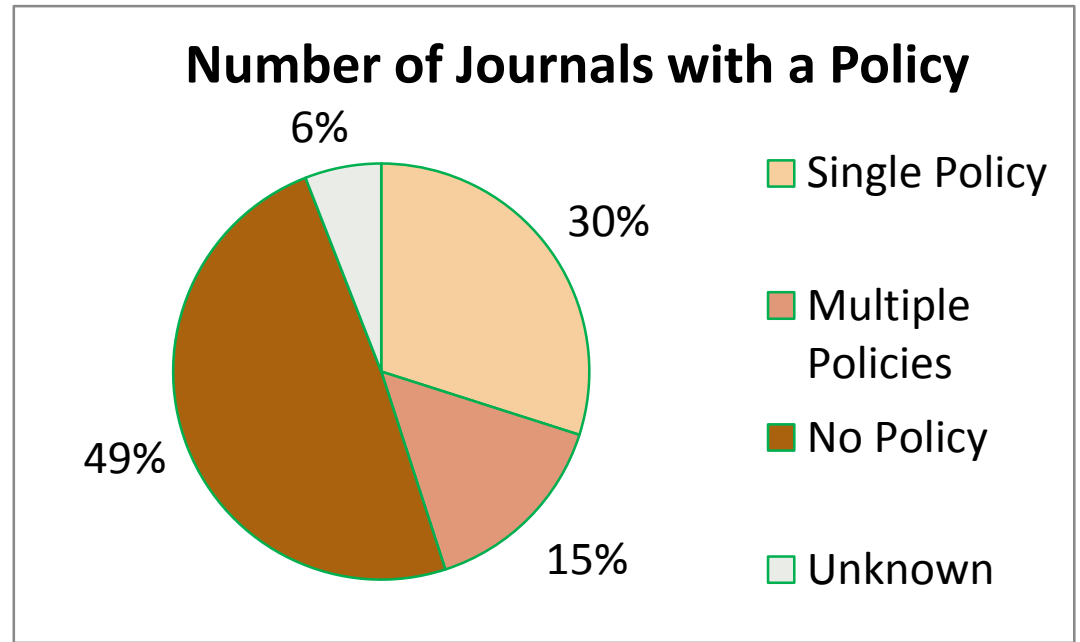
# Journal deposit policies

Jisc journal research data deposit requirements feasibility study:

371 titles, 36 subject areas

Biomedical science leads the way in data sharing practice

A number of journals require research data to be deposited and openly available



# In Horizon 2020...

---

...a limited pilot action on open access to research data will be implemented. Participating projects will be required to develop a Data Management Plan (DMP), in which they will specify what data will be open.

All projects to

- Research and Innovation actions
- Innovation actions

[http://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-data-mgt\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf)

# Managing Open Data

---



# European and global initiatives

---

OpenAIRE: infrastructure project

DANS: data archiving and network services

Knowledge Exchange

DataCite

LIBER – European network of research libraries

# Tools

---

Repository networks and catalogues:

- re3data – global registry of research data repositories
- COAR – confederation of open access repositories

Examples:

- Zenodo
- figshare

re3data.org  
REGISTRY OF RESEARCH DATA REPOSITORIES

COAR  
Confederation of Open Access Repositories

zenodo

Thank you for your  
attention!

---