



Open Data: Delivering the benefits

Ben Johnson


Research policy adviser

RDMF workshop

27 April 2016

Shape of this presentation

- The changing role of outputs in research
- Benefits of open data
- Support for open data and evidence of progress
- How we are getting there



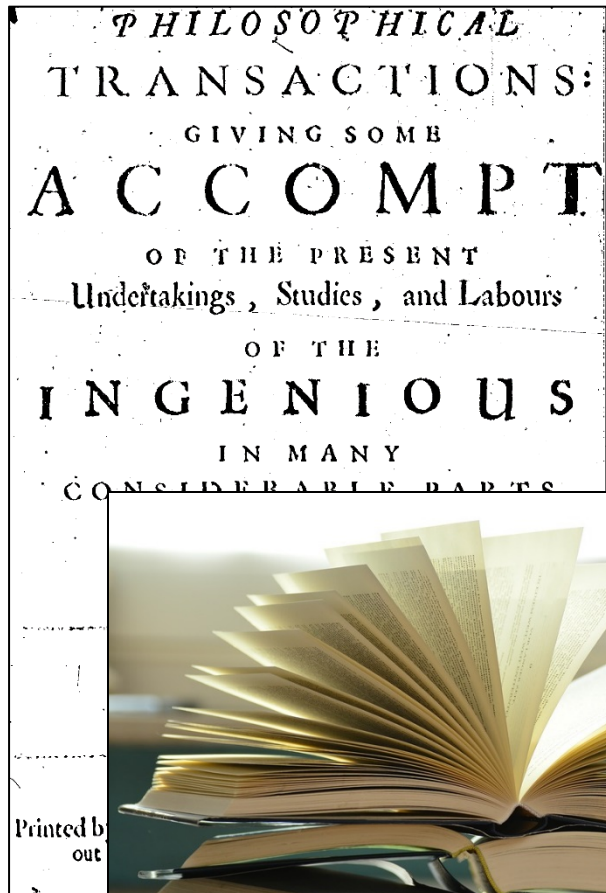
**<open
data>**

The role of outputs in research



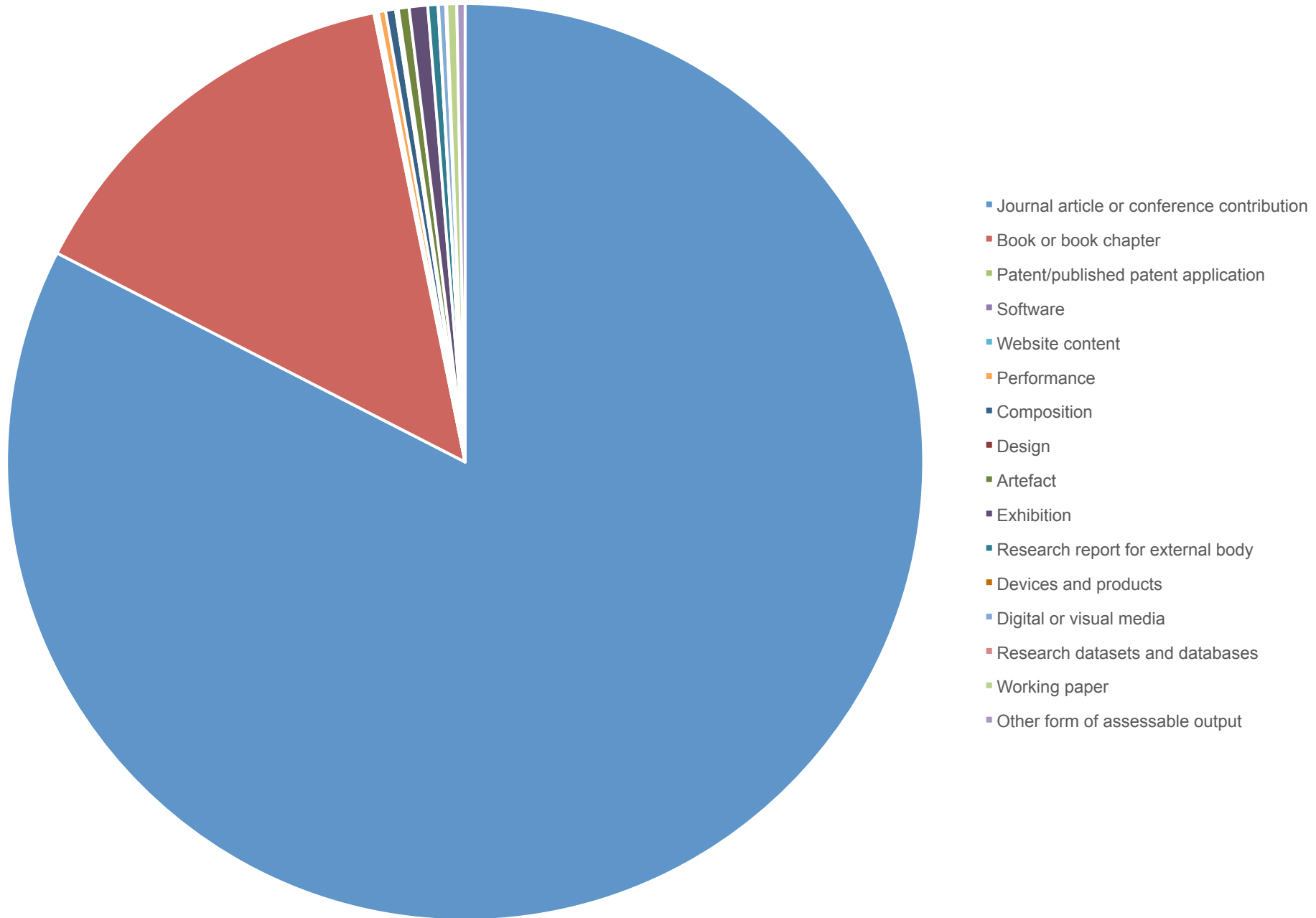
- Communicate results
- Increase standing

The classic way – books and articles



- Results
- Arguments
- Discussion

Output types submitted to REF2014



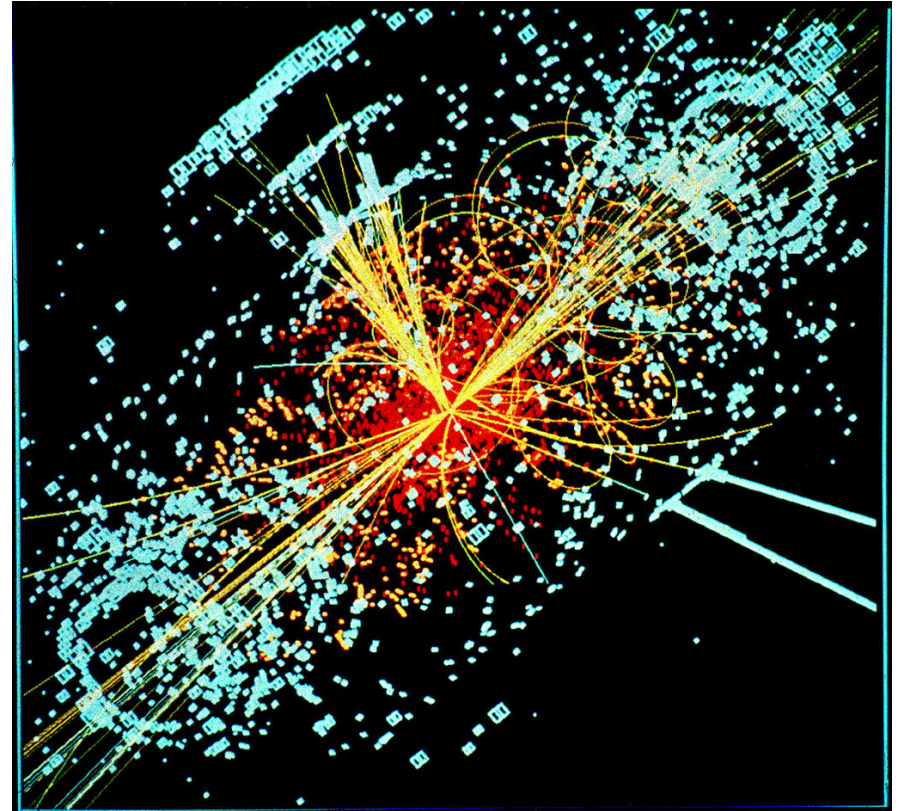
The classic way – books and articles



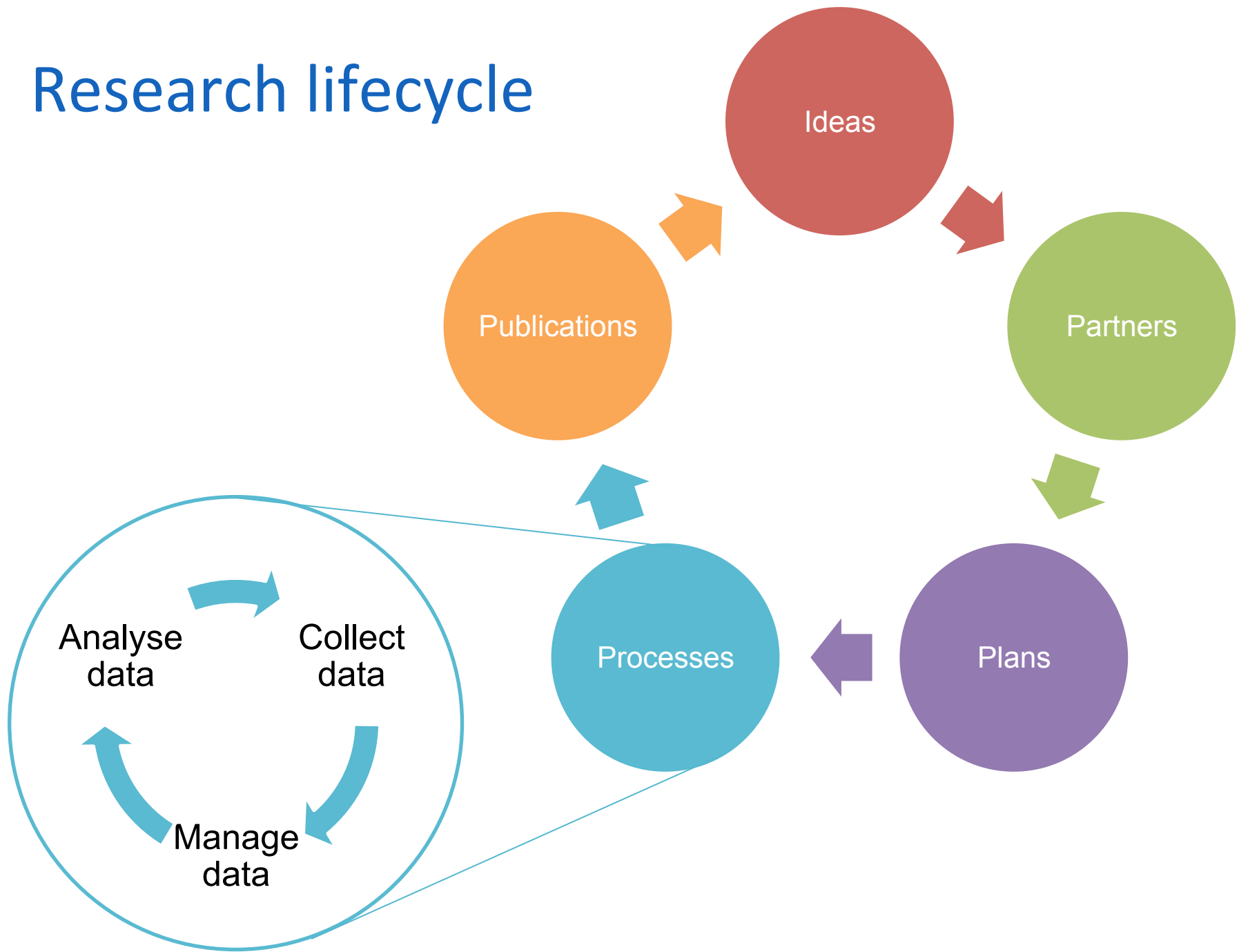
- Selectively descriptive
(I only publish what matters)
- Carefully contextualised
(I only publish what fits)
- A form of advertising
(I only publish what makes me look good)

“I”? Or “we”?

- Research is getting ever faster and bigger
- “Team science”
- Researchers are part of a bigger global ‘club’
- We need better ways to communicate at all parts of the lifecycle



Research lifecycle



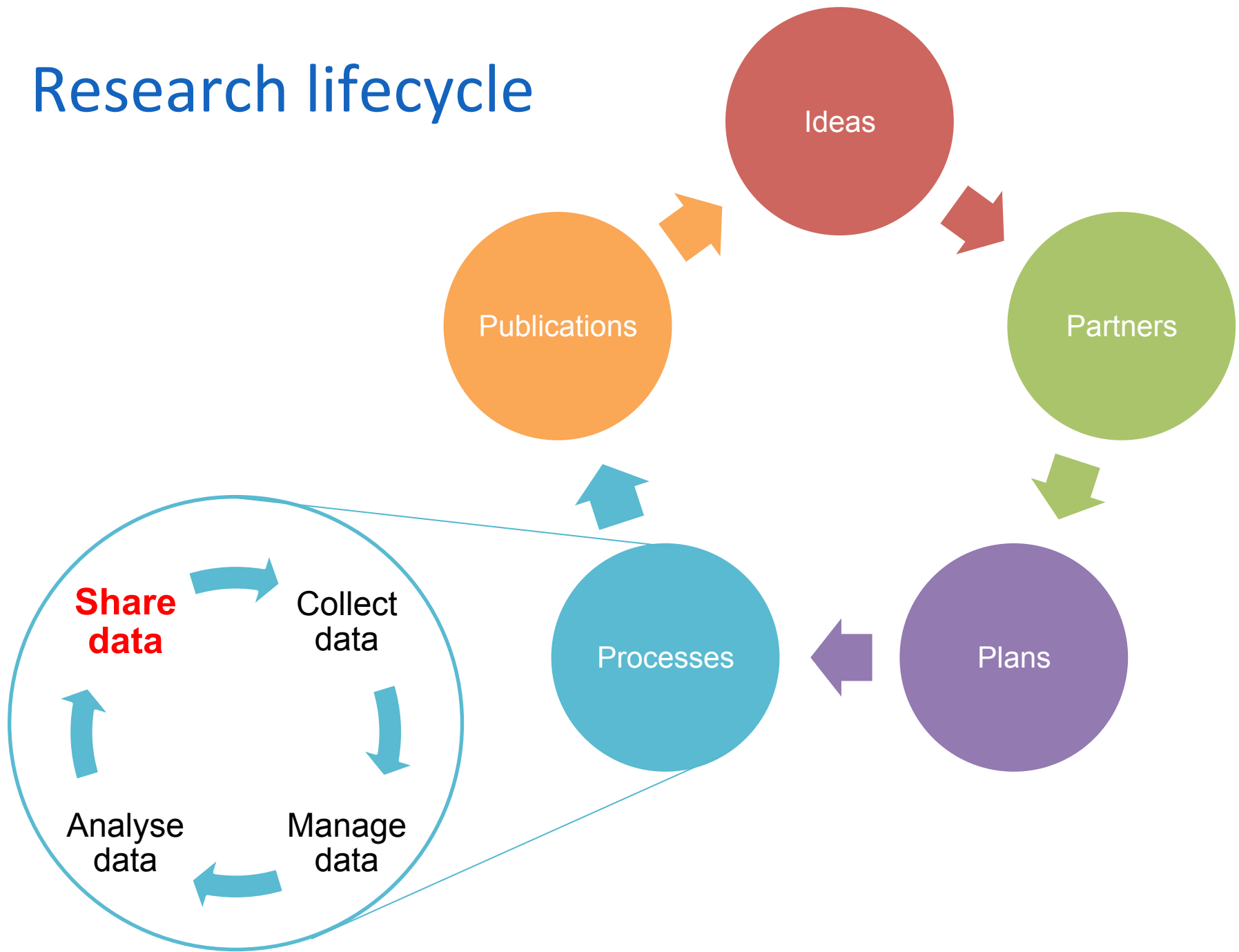
Where ideas, partners and publications are online and global...

Shouldn't data be an online, globally useful good?

For researchers? For business? For everyone?



Research lifecycle



Benefits of open research data

For researchers

- Ensuring that data are properly managed in the first place
- Avoiding duplication (e.g. two teams unknowingly collecting same data)
- Enabling collaboration and new uses for existing data
- Increasing reproducibility
- Making data ‘prestigious’

80% of datasets are lost within two decades [1]

Willingness to share research data is related to the strength of the evidence and the quality of reporting of statistical results [2]

Benefits of open research data

For the public

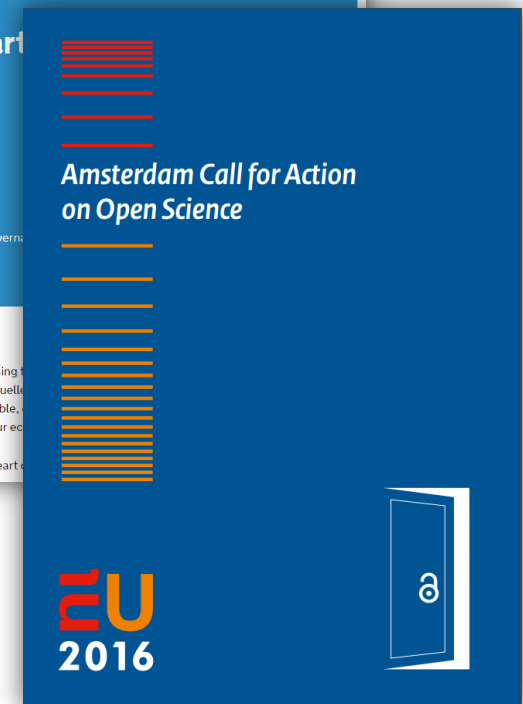
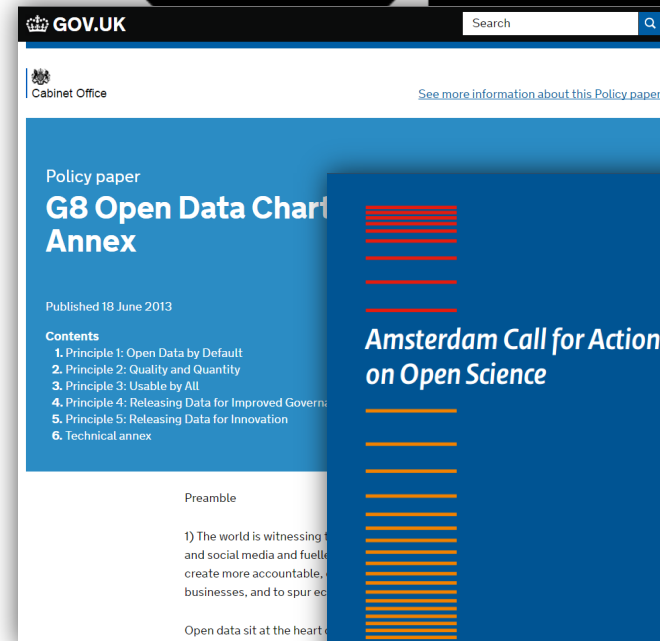
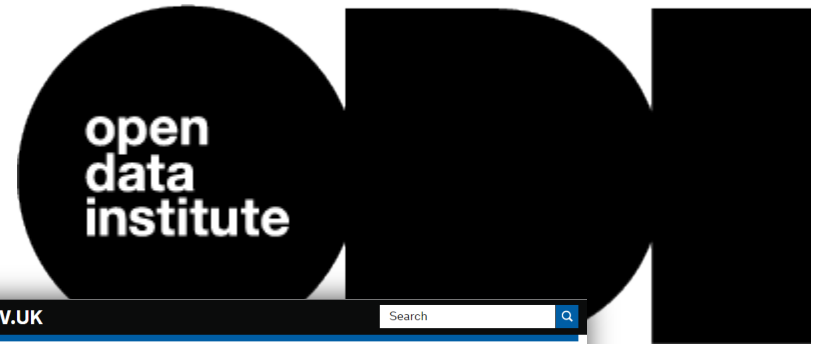
Open data will result in £10 return for every £1 invested [3]

Patterns of database citation in patents indicate long-term industry value of biological data resources [4]

- Innovation through openness
- Impact on business
- Impact on our economy and society
- Public trust in research
- Public inquiry / 'citizen science'

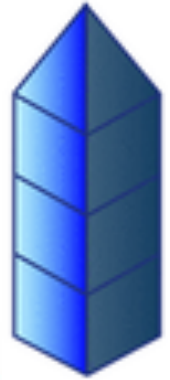
Examples of support and progress

- Strong advocacy for open data
- G8 commitment to open data (2013)
- Dutch presidency of the EU – focus on open science



Examples of initiatives

- Human Genome Project
- www.crystallography.net
- Blue Obelisk
- Environmental science
- <http://opendata.cern.ch/>
- ESRC business and local government data sets



ESRC Business and Local Government
Data Research Centre

Challenges along the way

- Concerns over “data parasites”
- Concerns over patient data (*care.data*)
- Other barriers (e.g. cost)



Work to be done

Infrastructure – we need:

- Tools for managing, curating and sharing data
- Standards
- Metrics for tracking scholarly/societal impact of data



Work to be done

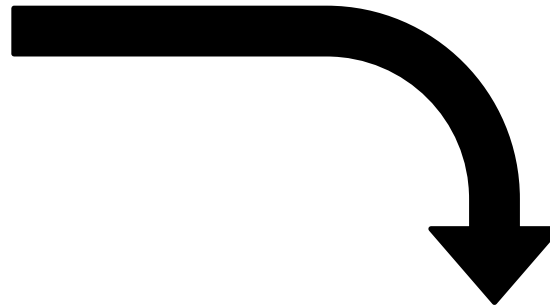
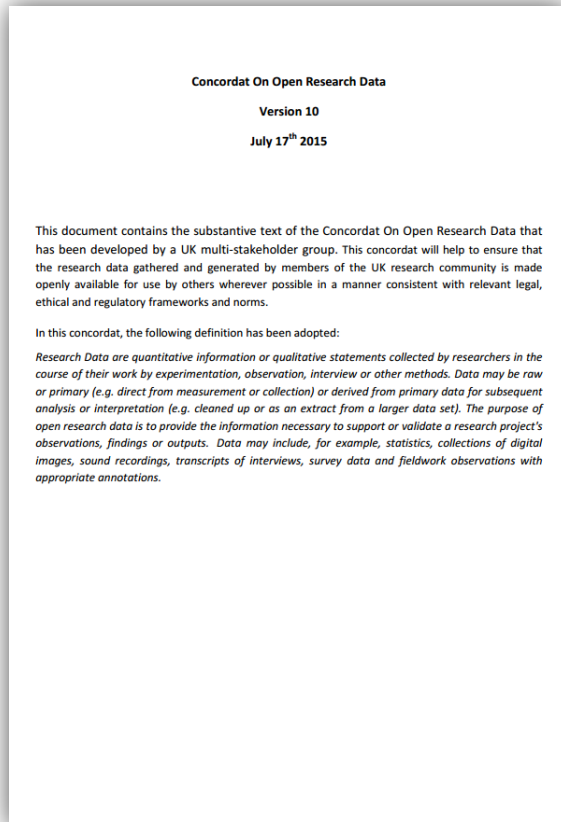
Policies and processes

- **Funders**
 - RCUK
 - REF?
 - Wellcome Trust
- **Publishers**
 - Policies (e.g. PLOS)
 - Journals (e.g. Scientific Data)



Collective efforts

- UK Open Research Data Forum
 - **“Concordat to support open research data”**
<http://www.rcuk.ac.uk/documents/documents/concordatopenresearchdata-pdf/>



- **Institutional practice**
- **Disciplinary arrangements**
- **Skills and training**

Come in WE'RE

OPEN



Thank you for listening

b.johnson@hefce.ac.uk

References

- [1] Vines, T.H. et al. (2014). The Availability of Research Data Declines Rapidly with Article Age. *Current Biology* 24(1). <http://dx.doi.org/10.1016/j.cub.2013.11.014>
- [2] Wicherts, J.M. et al. (2011). Willingness to Share Research Data Is Related to the Strength of the Evidence and the Quality of Reporting of Statistical Results. *Plos ONE*. <http://dx.doi.org/10.1371/journal.pone.0026828>
- [3] Tennison, J. (2015). The Economic Impact of Open Data: What Do We Already Know? *Huffington post blog*. Blog post. Accessed 26 April 2016. http://www.huffingtonpost.co.uk/jeni-tennison/economic-impact-of-open-data_b_8434234.html
- [4] Bousfield D, McEntyre J, Velankar S et al. Patterns of database citation in articles and patents indicate long-term scientific and industry value of biological data resources [version 1; referees: 3 approved]. *F1000Research* 2016, 5(ELIXIR):160 doi: 10.12688/f1000research.7911.1 <http://f1000research.com/articles/5-160/>

Copyright and credits

© LHC photo: Lucas Taylor / CERN (<http://cdsweb.cern.ch/record/628469>) [CC BY-SA 3.0 (<http://creativecommons.org/licenses/by-sa/3.0>)], via Wikimedia Commons

© Pipeline photo: By USAF photo (US Air Force Public Affairs [1]) [Public domain], via Wikimedia Commons

© Rusty tools photo: By Biser Todorov (Own work) [GFDL (<http://www.gnu.org/copyleft/fdl.html>) or CC BY 3.0 (<http://creativecommons.org/licenses/by/3.0>)], via Wikimedia Commons

Open sign: Aged Come In We're Open by Czarina Alegre
<https://www.flickr.com/photos/czarinaalegre/9400527612> [CC BY 2.0]

The copyright in this presentation is held either by the Higher Education Funding Council for England (HEFCE) or by the originating authors.

Due to use of ShareAlike content in this presentation, the slides are offered under a CC BY-SA 3.0 licence, enabling free distribution and adaption (as long as the authors are credited and the resultant work is shared under a similar licence).

Please contact customerservices@hefce.ac.uk for further information and re-use requests.