Beyond Open Access, towards Open Research

Pushing publishing standards to make research more visible

Antoine Bocquet
October 24, 2018
# Agenda

<table>
<thead>
<tr>
<th></th>
<th>Open Knowledge</th>
<th>Why do we care about OPEN and why should you?</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Open Research in Action</td>
<td>Future vision</td>
</tr>
<tr>
<td>3</td>
<td>What now?</td>
<td>Challenges and opportunities</td>
</tr>
</tbody>
</table>
Open Knowledge

Why do we care about OPEN and why should you?
Why should you care being OPEN?

*Because it improves the IMPACT of YOUR research*

Independent statistical analysis of articles published in *Nature Communications* found that:

- OA articles are **viewed 3 times more often** than articles that are only available to subscribers.
- OA articles are **cited somewhat more** than subscription articles.

*Research Information Network* on the effect of OA on citations and downloads of *Nature Communications* articles.

http://www.nature.com/press_releases/ncomms-report.html
Gold Open Access article share exceeded 20% in 2017

Share of fully OA APC journals reached 15% - OA in sponsored and hybrid journals add 3% each

Number of Articles in Journal Citation Reports 2017

- Fully OA Journals: Author Pays 15%
- Fully OA Journals: Sponsored 3%
- Subscription & Hybrid Journals 82%

OA articles est. 4% of all articles in subscription and hybrid journals

Total number of articles: 1.5 million

Sources: Journal Citation Reports 2017; English-language primary research journals only
### 2017 Open Access market size and growth

**Approx. 310-320k articles published, and 15% growth**

<table>
<thead>
<tr>
<th></th>
<th># of journals</th>
<th># of articles</th>
<th>Growth 17/16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fully OA Journals: Author Pays</td>
<td>1,900</td>
<td>260k</td>
<td>15%</td>
</tr>
<tr>
<td>Hybrid: Author choice</td>
<td>10,000</td>
<td>35-40k</td>
<td>10%</td>
</tr>
<tr>
<td>Hybrid: Offsetting</td>
<td>10,000</td>
<td>15-20k</td>
<td>20%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>11,900</strong></td>
<td><strong>310-320k</strong></td>
<td><strong>15%</strong></td>
</tr>
</tbody>
</table>

*Fully OA journals / Hybrid journals author choice:* data represent Top 21 publishers only; growth excl. flipped journals and acquisitions
Fully APC OA journals: uptake by country

China is the country with highest number of OA articles and above average OA share
But: Open Access is not a ‘free’ market ...

*China and the USA seek Green – Northern & Western Europe go for Gold*

National funders in Northern Europe encourage **Gold OA**, and many provide dedicated APC funds.

National funders in North America and China require **Green OA**; no dedicated APC funds, although authors may use research funds to pay APCs.

**National research council policies**
- OA mandate + dedicated APC funding
- No mandate, dedicated APC funding
- OA mandate + APC funding via research grants
- Green OA (self-archiving) mandate, no APC funding
- National OA repository scheme

**International**
- OA mandate and APC funding: European Commission, ERC, GRC, WHO
Close to 70% agreed that Open Access is the future of academic and scientific publishing

“Open access is the future of academic and scientific publishing” (n=7790)
- Strongly agree: 36%
- Agree: 31%
- Neutral: 20%
- Disagree: 8%
- Strongly disagree: 5%

“I don’t care whether a journal is open access or not, I will submit to the best journal for my work” (n=8041)
- Strongly agree: 34%
- Agree: 30%
- Neutral: 17%
- Disagree: 13%
- Strongly disagree: 6%

"My primary area of research would benefit from more journals that publish peer-reviewed open access articles" (n=7873)
- Strongly agree: 29%
- Agree: 32%
- Neutral: 24%
- Disagree: 10%
- Strongly disagree: 5%

"All papers should be published via open access" (n=7943)
- Strongly agree: 29%
- Agree: 26%
- Neutral: 23%
- Disagree: 13%
- Strongly disagree: 9%

Weighted 2017 data
Why do we care about being OPEN?
Because it’s part of the solution to a critical problem facing science!

“53 papers were deemed 'landmark' studies. Nevertheless, scientific findings were confirmed in only 6 (11%) cases.”

— *Nature* 483, 531–533 (2012), doi:10.1038/483531a
The reproducibility crisis

We asked researchers about it...

• **52% of the researchers** we surveyed believe there is a “crisis” of reproducibility.

• **70% failed to reproduce** another scientist’s experiment and more than 50% failed to reproduce an experiment of THEIR OWN.

• **Physicists and chemists most confident** in published literature than life scientists.

• **Key culprits** — selective reporting, pressure to publish, low statistical power or poor analysis.

• **1/3 respondents had taken active measures** to improve reproducibility in their labs in the past 5 years; often independent replication, and better attention to documentation.

*Nature* 533, 452–454 (2016); doi:10.1038/533452a
The problem isn’t just fraud

Although the lion share of retractions arise from misconduct, poor reproducibility does not

We are not talking about results that were:

• Falsified — misconduct is a problem but it’s still in the minority.
• Wrong — legitimate observations but subsequent work disproves the hypothesis — that’s how science is meant to work!

We are talking about results that are:

• Poorly described preventing verification — independent experimenter cannot observe the same results under similar conditions.
• Overstated — failure to consider alternative explanations.
• Misrepresented — data claimed to be more robust than they actually are, CHERRY PICKING!
• Sloppy — failure to account for (or even consider) sources of error, poor use of statistics, poor controls.

The problem isn’t restricted to open access journals!

Subscription journals suffer just as much.
An attempt to reproduce key results in eighteen *Nature Genetics* papers that reported microarray-based gene expression data found that:

- Only two could be reproduced fully.
- Six were reproduced partially.
- Ten could not be reproduced at all.

“The main reason for failure to reproduce was data unavailability, and discrepancies were mostly due to incomplete data annotation or specification of data processing and analysis.”


The problem isn’t too much OPEN but too little
How do we fix this?

*Recognize that robust, incremental science is as important as high-impact science*

Ensure that researchers get adequate credit for publishing reliable, reproducible, robust science in journals that aren’t necessarily restricted to publishing high impact science.
Why share data?
Because that’s what science should be about

“Publishing research without data is advertising, not science!”

— Cameron Neylon, Graham Steel, and others.
But what about the results themselves?

The most important currency in science isn’t scientific papers but DATA!
The data descriptor (data paper) article

- Associated Nature Article
- Data at figshare & NCBI GEO
- Integrated figshare data viewer

**Time-resolved gene expression profiling during reprogramming of C/EBPa-pulsed B cells into iPS cells**

Bruno Di Stefano, Samuel Collombet & Thomas Graf

**Abstract**

Methods • Data Records • Technical Validation • Usage Notes • Additional Information • References • Data Citations • Acknowledgements • Author Information

3. Di Stefano, B., Collombet, S., & Graf, T. Figshare
OPEN RESEARCH IN ACTION

Future vision

2.0
Google Dataset Search Beta

Google launches ‘Dataset Search’ to support scientists and journalists (Sep. 2018)

“Dataset Search lets you find datasets wherever they’re hosted, whether it’s a publisher’s site, a digital library, or an author’s personal web page”

Natasha Noy
Research Scientist, Google AI

In today’s world, scientists in many disciplines and a growing number of journalists live and breathe data. There are many thousands of data repositories on the web, providing access to millions of datasets; and local and national governments around the world publish their data as well. To enable easy access to this data, we launched Dataset Search, so that scientists, data journalists, data geeks, or anyone else can find the data required for their work and their stories, or simply to satisfy their intellectual curiosity.


Beyond Open Access, towards Open Research / October 24, 2018
Archaeological fish-bone images
researchdata.ands.org.au
Published Jul 17, 2012

Biological Data for the brilliant pomfret (Eumegistus illustris) off the...
www.seanoe.org
Published 2016

Archaeological fish-bone images
data.wu.ac.at
Updated Oct 10, 2013

Data from: Tropical ancient DNA from bulk archaeological fish bone reveals...
dryad2.lib.ncsu.edu
Published Nov 22, 2016

Data from: Nonlinear fishbone dynamics in spherical tokamaks
www.osti.gov
Published Jan 1, 2017

Archaeological fish-bone images
researchdata.ands.org.au

S scholarly articles cite this dataset (View in Google Scholar)

Dataset published Jul 17, 2012

Dataset provided by
The University of Sydney

Description
The archaeological fish-bone images collection contains images of selected diagnostic anatomical elements from fish taxa commonly found on archaeological sites in the Sydney region, supplemented by taxa of the same family or genus from elsewhere where modern reference skeletons of Sydney taxa were not readily available. Scientific terminology and common names follow the Australian Museum Fish Site. Coding for taxa and anatomical parts were devised by Sarah Colley as part of the Sydney Fish Project archaeological research project. The taxonomy of modern specimens in the Australian National University and University of Sydney reference collections was copied from current collection catalogues. No attempt has been made to verify their accuracy at this stage. Archaeological specimens were matched against examples available in the Australian National University and University of Sydney reference collections, which do not currently contain every possible fish taxon from the Sydney region. Where no exact match could be found the taxon has been classified as 'unknown' at this stage. Some identifications of archaeological specimens are tentative at this stage.

Presentation of the collection is a partnership initiative of the University of Sydney Library and Dr. Sarah Colley, to support archaeological research. Digitisation was funded under the University of Sydney Faculty of Arts Research Support Scheme. Dr. Sarah Colley created and described the content, and additional content was also provided by the Australian National University. Russell Workman photographed the physical specimens. Rowan Brownlee created the site.
Pilot code peer review and publication with ‘Code Ocean’

A mission to make code more readily executable and discoverable

The 6-month author opt-in trial will involve three Nature journals: Nature Methods, Nature Biotechnology, Nature Machine Intelligence
Publishers are working with Katalysis.io on developing a Blockchain for Peer Review project

“Our initiative aims to introduce transparency into peer review by making crucial data about the process available across a shared platform that can comply with the demands of confidentiality and privacy.”

Publishing giants sign up to blockchain pilot

Springer Nature, Taylor & Francis Group and Cambridge University Press have joined a pilot project to test blockchain technologies applications to peer review.

Springer Nature

Beyond Open Access, towards Open Research / October 24, 2018
1. Data extraction from Journal Management System

2. Generation of a unique, anonymous id for the reviewer.

3. Augmentation of ORCID profile with review information.

4. Storage of the anonymised review information on the blockchain.

Publisher

1

Review State

Manuscript

Journal

Reviewer

Reviewer

FADE2341

Reviewer

ORCID

Reviewer

FADE2341

Blockchain

Katalysis

Blockchain for Peer Review
Gold OA in the UK survey: *bringing OA to more disciplines*

Research funding varies significantly by discipline, and this can affect authors’ ability to publish via the gold OA route, particularly if they do not have access to dedicated OA funds for APCs. Interestingly, we are seeing an increasingly high level of gold OA in a number of disciplines that do not traditionally publish large numbers of articles via gold OA.

Source: [https://www.springernature.com/gp/open-research/about/gold-oa-in-the-uk](https://www.springernature.com/gp/open-research/about/gold-oa-in-the-uk)
And finally — *sharing materials* is just as important!

“A condition of publication in a Nature journal is that *authors are required to make unique materials available* to others without undue qualifications.”

http://www.nature.com/authors/policies/availability.html

“After publication, all data and materials necessary to understand, assess, and extend the conclusions of the manuscript must be available to any reader of Science.”

http://www.sciencemag.org/authors/science-editorial-policies?_ga=1.200322094.1663650950.1484202254

“One of the terms and conditions of publishing with Cell Press is that authors be willing to *distribute any materials and protocols used* in the published experiments to qualified researchers for their own use.”

http://www.cell.com/cell/authors
WHAT NOW?

Challenges and opportunities
OA DATA SHARING - EVIDENCE AND CONTEXT
Funder policies and feedback (2016-2017 data)
Growing number of mandates, yet funders share common concerns

Open data mandates
- Forthcoming: 2%
- No mandate: 6%
- Open data encouraged: 10%
- OA mandate for data: 6%
- OA mandate - some data types: 6%
- OA mandate - some projects: 8%
- OA mandate: 60%

Funder concerns
- Inadequate metadata/ lack of metadata expertise
- ID of data to deposit/ preserve
- Lack of incentives
- Disciplinary variation
- Concern over misuse of data by 3rd parties
- Compliance with Federal regulations
- Data protection/ treatment of sensitive data
- Identification of suitable repositories
- Meeting the costs of RDM
- Cultural resistance to sharing

Source: Springer Nature Research, 2016 (n=52) / Springer Nature Research Data Service Funder Survey, April 2017 (n=21)
Beyond Open Access, towards Open Research / October 24, 2018
Higher Education Institutions have similar concerns

- Researchers' need for support to prepare data
- Data protection/treatment of sensitive data
- Lack of incentives (beyond compliance)
- Inadequate metadata/expertise
- Cultural resistance/lack of awareness
- Long-term preservation
- Timing of deposit
- Licensing of data/ownership
- Identifying what to preserve

Springer Nature Research Data Service Higher Education Institute, June 2017 (n=15)
How interested are authors in a service that helps them to deposit their data in a repository?

Authors

<table>
<thead>
<tr>
<th></th>
<th>Positive</th>
<th>Neutral</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of authors</td>
<td>4477</td>
<td>1461</td>
</tr>
</tbody>
</table>

How important is data discoverability to authors?

Authors

<table>
<thead>
<tr>
<th></th>
<th>Important</th>
<th>Unimportant</th>
<th>Neutral</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage</td>
<td>76%</td>
<td>13%</td>
<td>11%</td>
</tr>
</tbody>
</table>

From a Springer Nature researcher survey. Total respondents = 7656
UNDERSTANDING THE CHALLENGES AND OPPORTUNITIES
76% of respondents highly rate the importance of their data being discoverable: most popular ranking was 10/10

Importance of data discoverability

Importance out of 10

- South America: 7.7
- Asia: 7.6
- Europe: 7.3
- North America: 7.2
- Australasia: 6.9

Majority of researchers share their data in some way

63% of respondents stated that they generally submitted data files as supplementary information, deposited the files in a repository, or both
Problems researchers have in sharing data by percentage

From a Springer Nature researcher survey (n=7719). [https://doi.org/10.6084/m9.figshare.5975011](https://doi.org/10.6084/m9.figshare.5975011)
Challenges and opportunities for Japanese researchers

Are researchers creating data management plans, if not, what is preventing them from doing so?

Are researchers sharing data and, if not, what is preventing them from doing so?

- Online survey in English & Japanese that received more than 1,000 responses
- Infographic and data on figshare: DOI: 10.6084/m9.figshare.6328952

Respondents' research subject area

<table>
<thead>
<tr>
<th>Subject Area</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>1393</td>
<td>100.0%</td>
</tr>
<tr>
<td>Physical Sciences</td>
<td>370</td>
<td>27.6%</td>
</tr>
<tr>
<td>Medical Sciences</td>
<td>415</td>
<td>29.8%</td>
</tr>
<tr>
<td>Earth Sciences</td>
<td>74</td>
<td>5.3%</td>
</tr>
<tr>
<td>Biological Sciences</td>
<td>389</td>
<td>27.9%</td>
</tr>
<tr>
<td>Other (inc. Arts &amp; Humanities, social sciences)</td>
<td>145</td>
<td>10.4%</td>
</tr>
</tbody>
</table>

Respondents' career stages (seniority)

<table>
<thead>
<tr>
<th>Career Stage</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>956</td>
<td>100.00%</td>
</tr>
<tr>
<td>First stage researcher (student, PhD, postdoc)</td>
<td>157</td>
<td>16.42%</td>
</tr>
<tr>
<td>Recognised researcher (staff scientist, associate professor/lecturer)</td>
<td>334</td>
<td>34.94%</td>
</tr>
<tr>
<td>Established researcher (senior scientist, PI)</td>
<td>97</td>
<td>10.15%</td>
</tr>
<tr>
<td>Leading researcher (Attending Physician, Lab Director, Professor VP of Research)</td>
<td>295</td>
<td>30.86%</td>
</tr>
<tr>
<td>Other</td>
<td>73</td>
<td>7.64%</td>
</tr>
</tbody>
</table>
Private sharing of data is more common than public sharing of data

<table>
<thead>
<tr>
<th>Country</th>
<th>Public Data Sharing (median)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poland</td>
<td>76%</td>
</tr>
<tr>
<td>Germany</td>
<td>75%</td>
</tr>
<tr>
<td>UK</td>
<td>58%</td>
</tr>
<tr>
<td>USA</td>
<td>55%</td>
</tr>
</tbody>
</table>

The three most common methods of private sharing were:
- Email (65%)
- USB or flash drives (41%)
- File sharing services (39%)

The three most common ways of public sharing were: (n=569)
- Supplementary information to journal articles (51%)
- Lab or personal website (27%)
- Subject specific repository or data archive (25%)

From our global survey levels of public data sharing reported:
- Poland – 76% (highest)
- Germany – 75%
- UK – 58%
- USA – 55%

Beyond Open Access, towards Open Research / October 24, 2018
The current state of awareness of funder requirements

34% of respondents do not know what their main funders’ requirements are in relation to DMPs (n=763)

23% of respondents do not know what their main funders’ requirements are in relation to data sharing (n=723)

Example:
- JST has a data management policy but only 11% of respondents who identified JST as their main research funder correctly identified this. 66% incorrectly identified JST’s requirements and 23% did not know.

<table>
<thead>
<tr>
<th>Funder</th>
<th>None</th>
<th>AMED</th>
<th>JST</th>
<th>JSPS</th>
<th>MHLW</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of respondents</td>
<td>414</td>
<td>68</td>
<td>144</td>
<td>509</td>
<td>20</td>
<td>174</td>
</tr>
</tbody>
</table>
What motivates researchers to share data*?

- **97%** - to accelerate research and its applications¹
- **96%** - increased visibility and discovery of their research data¹,²
- **95%** - increased usability of their research data²
- **>90%** - credit mechanism for deposit of data¹,²
- **88%** - to comply with funder policy¹

---

RESEARCH DATA SUPPORT FOR RESEARCHERS
Springer Nature research data policy initiative (July 2016)

Policy Types

- **Type 1**
  - Data sharing and data citation is encouraged but not required

- **Type 2**
  - Data sharing and evidence of data sharing encouraged

- **Type 3**
  - Data sharing encouraged and statements of data availability required

- **Type 4**
  - Data sharing, evidence of data sharing and peer review of data required

- **Around 1,400** (>50%) Springer Nature journals have adopted a standard research data policy as of March 2018
- Practical and pragmatic
- Preference data archiving in repositories over SI/ESM
- Support community specific policies, mandates and repositories
- Promote use of data availability statements
- Promote data citation
- Offer support from the publisher via our helpdesk

https://www.springernature.com/gp/authors/research-data-policy/journal-policies/15369670

*Standardising and harmonising research data policy in scholarly publishing*

Iain Hrynaszkiewicz, Aliaksandr Birukou, Mathias Astell, Sowmya Swaminathan, Amye Kenall, Varsha Khodiyar, bioRxiv 122929; doi: https://doi.org/10.1101/122929

Beyond Open Access, towards Open Research / October 24, 2018
## Research Data Helpdesk

<table>
<thead>
<tr>
<th>Queries are answered within two business days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Run by members of the Springer Nature Data Publishing team</td>
</tr>
<tr>
<td>Expertise in data curation and management, archiving and digital preservation, copyright and licensing, Open Access publishing</td>
</tr>
<tr>
<td>Always encourage best practices, e.g. the use of community repositories for specific data types</td>
</tr>
</tbody>
</table>

**Email:** researchdata@springernature.com
Introducing Springer Nature Research Data Support

To help Springer Nature authors and journals follow good practice in sharing and archiving of research data, we’re offering optional data deposition and curation services.

Researchers submit their data files securely  

The Research Data team curates the data and metadata  

The data are published and linked to the author’s paper

More information is available on our website here:  
http://go.nature.com/ResearchDataServices
Our research data editors:

- Enhance metadata of dataset(s) to improve findability and encourage reuse
- Check for presence of sensitive information and human identifiers
- Apply DOIs to provide unique persistent links to dataset(s) and enable citation of them
- Link data to their associated article(s) and coordinate publication with the article
- Store data in the Springer Nature portal in the figshare repository
TYPES OF RESEARCH DATA WE CAN HELP TO ORGANISE AND SHARE

Research data are any files that have been generated as part of your research that are not your research manuscript.

- Archived data (zip, rar, iso)
- Audio (mp3, wav, aif)
- Spreadsheet (csv, xls, tsv)
- Documents (doc, pdf, odt)
- Text (txt, rtf, bib)
- 3D graphic (obj, stl, ply)
- Notebook (ipynb)
- Molecular (cif, pdb, xyz)
- Image (jpg, png, svg)
- Presentation (ppt, pptx, pptm)
- Visualisations (gexf, gephi)
- Geographical & map (keyhole, GIS, gif)
- Code (python, r, java)
- Video (mp4, mov, avi)
Springer Nature Research Data Support

Benefits to researchers
- Enhance peer-reviewed publications with consistent links to research data
- Aligns with the FAIR (Findable, Accessible, Interoperable and Reusable) data principles
- Improve discoverability of research data and associated publications
- Provide secure private storage until researchers are ready to publish

Benefits to librarians and other research support staff
- Complement institutional research data management solutions
- Save researchers and support staff time
- Apply consistent editorial standards to data archiving and publishing
- Covers all scholarly research disciplines and digital data (including code, videos, images, text, raw and processed data)

Benefits to institutions and funders
- Enable compliance with funding agency policy, institutional and journal policies
- Enables easier compliance monitoring and reporting
- Provides reliable, scale-able service for all disciplines
- Can complement or augmenting existing institutional data support teams
Research Data Support

After curation, data is not just more DISCOVERABLE, it’s more USEFUL!

Data Support Services help researchers make their open data more valuable, with:

- Links to associated, peer-reviewed publications,
- Consistent titles and author names,
- Clear citation information,
- Files preview-able in browser,
- Metadata for each file in the archive,
- Contextual information,
- Clear license/terms of use,
- Dataset description/abstract,
- Rich usage statistics.
Example output of Research Data Support

**Paper** published in *Nature* (https://doi.org/10.1038/nature23654)

**Dataset** published in the Springer Nature figshare repository (https://doi.org/10.6084/m9.figshare.c.3814360)

---

**Data availability statement included with the paper**

---

**Data availability**

The CT data that support the findings of this study, as well as 3D surface files of described material, are available in figshare with the identifier https://doi.org/10.6084/m9.figshare.c.3814360. Other data files are included in the Supplementary Information.

---

**References**

Researcher feedback on Research Data Support

I had a really great experience using the Springer Nature Research Data Support. The process was straightforward, and the team were very helpful at guiding me through the process and dealing with my enquiries. The resultant data package is very easy to access and navigate. I opted to use the service because I wanted to make sure that the data accompanying my paper are as accessible as possible, and this presents an ideal way of facilitating access.

Dr Samantha Giles, Research Fellow, University of Oxford
(https://doi.org/10.1038/nature23654)

Research Data Support provided an uncomplicated yet highly efficient way to share and archive my data. The whole process went smoothly and the team was always available and helpful. I really enjoyed the minimal effort and time I spent on the data submission, which resulted in high quality outputs.

Dr Hasina Josué Rakotoianiaina, University of Göttingen
(https://doi.org/10.1186/s12898-017-0140-1)
In July 2018, the Wellcome Trust partnered with Springer Nature on a pilot to make our Research Data Support service available to all Wellcome-funded researchers at no cost to the researcher.

- Removes the barrier of lack of funding for research data sharing
- Available to anyone with a Wellcome grant or affiliated to a Wellcome centre
- Researchers remain in control of what they publish and choose licence for their data
- Available to all research disciplines and file formats
- Makes the service freely accessible to researchers at many UK institutions
- Available for research data supporting any publication in any journal or publisher
- Provides deposition and curation for all datasets without a community repository

CONCLUSIONS
Open research is more productive research!

Researchers who make their data open, publish more papers

Data archiving can double the publication output of studies

A study of 7,000 NSF and NIH research projects in social sciences found that:

- Those with archived data resulted in 10 (median) publications
- Those without archived data resulted in 5 publications

Principal investigators who archived their data were more likely to publish more articles per project, and to see others build on their work

Research articles with open data are cited up to 50% more

Analysis shows that articles with data available are cited 9-50% more, depending on the field

5. Sears et al (2011) https://figshare.com/articles/Data_Sharing_Effect_on_Article_Citation_Rate_in_Paleoceanography/1222998/1

Beyond Open Access, towards Open Research / October 24, 2018
Open research benefits the economy

CASE STUDY: The Human Genome Project

US government invested a total of $14.5 billion in the Human Genome Project and other project related to sequencing the human genome. The data obtained during the project — that is, the entire human genome — is open to anyone who wants to use it.

A report by the Battelle Memorial Institute estimates that the project generated more than $1 trillion dollars in economic revenue to the US economy in the decade since its completion.

This represents a 178-to-1 return on investment — not to mention the social and scientific benefit of the project. It is unlikely that this return would have been anywhere near as large had the genome not been made open.
Open research has **more impact**!
Papers published with open data are cited more often

Gene expression microarray papers that link to open data receive **9% to 30% more citations** than those that don’t.

Astronomy papers that link to open data receive **20% more citations** than those that don’t.

Astrophysics papers that link to open data receive **28% to 50% more citations** than those that don’t.

Paleoceanography papers that link to open data receive **35% more citations** than those that don’t.
— Sears et al (2011) https://figshare.com/articles/Data_Sharing_Effect_on_Article_Citation_Rate_in_Paleoceanography/1222998/1

Open has more impact than closed
CONCLUSION

Open Science is growing globally in diversity, and awareness is peaking, with the continuous support from academic institutions, funders and publishers. Startups and publishers are among the front runners of this new developing ecosystem.
Beyond Open Access, towards Open Research / October 24, 2018

The story behind the image

Marie Curie (1867–1934)
In a scientific world still dominated by men, Marie Curie shone not only as an extraordinary pioneer in the field of radioactivity, but also as a trailblazing female scientist. A French-Polish chemist and physicist, Curie discovered two new elements, polonium and radium, and revolutionised our understanding of radioactivity, the process by which unstable atoms decay by emitting energy in the form of radiation. The first person of either gender to win or share two Nobel Prizes, Curie is one of the most renowned scientists of a generation, whose influences can be seen throughout many areas of modern science, from particle physics to medicine.

Thank you

Antoine Bocquet
Vice President Sales Japan, Southeast Asia and Oceania
Managing Director, Nature Japan K.K / Springer Japan K.K

Springer Nature
Shiroyama Trust Tower 5F, 4-3-1 Toranomon, Minato-Ku, Tokyo 105-6005, Japan
T    +81 (0)3 4533 8116 | M +81 (0)90 6797 0089
antoine.bocquet@springernature.com | www.springernature.com

For further reading:
https://go.nature.com/OpenResearch
http://go.nature.com/ResearchDataServices

Slide acknowledgements:
Iain Hrynaszkiewicz, Edmund Gerstner, Mithu Lucraft, Natalia Timiraos, Katie Allin, Ludivine Allagnat

For further reading:
https://go.nature.com/OpenResearch
http://go.nature.com/ResearchDataServices

Slide acknowledgements:
Iain Hrynaszkiewicz, Edmund Gerstner, Mithu Lucraft, Natalia Timiraos, Katie Allin, Ludivine Allagnat