Making research data the new water

Dr Marta Teperek
Office of Scholarly Communication, University of Cambridge
@martateperek
mt446@cam.ac.uk

22 November
CBU Open Science Workshop 2016
Analogy: data is the new oil

Why Data Is The New Oil

by Jonathan Vanian


DATA IS THE NEW OIL OF THE DIGITAL ECONOMY


http://www.leighpartnership.com/tlpweb/?p=122
But is it really the new oil?

Oil: finite quantity

And there are some negative associations:

• Oil is about the money
• Oil is about exploitation and not about social justice
So perhaps data should be the new water?

- Water is essential
- Water is transparent
- Water can be re-used

*Image from Pixabay, CC0 licence*
Assumptions behind data being like water

Data producer

Shares data FAIRly:
- Findable
- Accessible
- Interoperable
- Reusable

Data re-user

New hypotheses and new research

Data validation

Funding for new research

Data citation

Recognition and reward

Image from Pixabay, CC0 licence
Assumptions behind data being like water

Shares data FAIRly:
• Findable
• Accessible
• Interoperable
• Re-usable

Data validation
New hypotheses and new research

Data producer

Funding for new research

Data re-user

Data citation
Recognition and reward

Image from Pixabay, CC0 licence
How to make data recyclable?

The problem:
Everyone is saying this for a looong time now

Data producers and data re-users need to feel rewarded for doing the right thing

Image from Pixabay, CC0 licence
Time to stop believing in fairy tales?

But: it’s not all bad
let’s see what’s already good
and what’s missing
So where are we now and what can be done?

Activities and roles for:

- Funders
- Publishers
- National/international organisations
- Institutions
- Individual researchers

Image from Pixabay, CC0 licence
Funders – NWO call

NWO makes 3 million available for Replication Studies pilot

Funders – Horizon 2020

- 2015: 65.4% of 431 signed projects opted to share data
- From 2017: the current Open Research Data Pilot expands to cover all areas of Horizon 2020, with the same rules

“Publicly funded research data are a public good (...), which should be made openly available with as few restrictions as possible...”

www.rcuk.ac.uk/research/datapolicy

In April 2016 we started exploring whether open research should be one of Wellcome’s strategic priorities.

https://wellcome.ac.uk/what-we-do/our-work/open-research
Perhaps researchers who share could be better incentivised?

- All previous work made available?
- Commitment to openness assessed at the application stage?
Data Availability

The following policy applies to all PLOS journals, unless otherwise noted.

PLOS journals require authors to make all data underlying the findings described in their manuscript fully available without restriction, with rare exception.

http://journals.plos.org/plosone/s/data-availability

OF SCHEMES AND MEMES BLOG

Promoting research data sharing at Springer Nature

05 Jul 2016 | 14:25 GMT | Posted by Lillienne Zen | Category: Data policy, Featured, Science communication and outreach

This blog was written by Iain Hrynaszkiewicz, Head of Data Publishing at Springer Nature, in support of the newly launched, company-wide Springer Nature Research Data Policies.

http://blogs.nature.com/ofsciencesandmemes/2016/07/05/promoting-research-data-sharing-at-springer-nature
Publishers – data citation requirements

ENDORSE THE DATA CITATION PRINCIPLES

Open Access Project Board of the University of Cambridge

https://www.force11.org/datacitation/endorsements
Publishers – ensuring good practice

- Data citation?
- In house data experts?
National and international organisations

Integrated, collaborative, cost-effective approaches

European Open Science Cloud

11 October 2016 – first report from the High Level Expert Group

http://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud
National and international organisations

Integrated, collaborative, cost-effective approaches

Research data shared service
Advancing research data management through collaboration

Cambridge is part of this project

https://www.jisc.ac.uk/rd/projects/research-data-shared-service
Could Open Research benefit Cambridge University researchers?

This blog is part of the recent series about Open Research and reports on a discussion with Cambridge researchers held on 8 June 2016 in the Department of Engineering. Extended notes from the meeting and slides are available at the Cambridge University Research Repository. This report is written by Lauren Cadwallader, Joanna Jasiewicz and Marta Teperek (listed alphabetically by surname).

At the Office of Scholarly Communication we have been thinking for a while about Open Research ideas and about moving beyond mere compliance with funders’ policies on Open Access and research data sharing. We thought that the time has come to ask our researchers what they thought about opening up the research process and sharing more: not only publications and research data, but also protocols, methods, source code, theses and all the other elements of research. Would they consider this beneficial?

Working together with researchers – democratic approach to problem-solving

https://unlockingresearch.blog.lib.cam.ac.uk/?p=817
Institutions – removing the barriers

Image from Pixabay, CC0 licence
Data sharing in Cambridge is easy

Apollo - University of Cambridge Repository

This repository holds the research output of members of the University of Cambridge. It is delivered and managed by the University Library’s Office of Scholarly Communication team.

School of Arts and Humanities
School of Clinical Medicine
School of Technology

School of the Biological Sciences
School of the Humanities and Social Sciences
School of the Physical Sciences

Colleges
Other Communities

www.repository.cam.ac.uk
Each dataset gets its own record

Research data supporting “Hierarchical Self-Assembly of Cellulose Nanocrystals in a Confined Geometry”

Citation

Description
Supporting data for the article titled “Hierarchical Self-Assembly of Cellulose Nanocrystals in a Confined Geometry”. The article was accepted for publication in 2016 in the journal “ACS Nano”. Electronic supporting information is available from the publisher (ACS). The data is provided with a description in pdf and a structured set of seven folders compressed in zip, each correlating to a specific data type presented in the published article. The spreadsheet used to plot Figure S12 in the ES1 is also included. Folder 1: Phase diagram - containing png of the pictures, as well as a txt, xlsx and pdf of the measurements (Figure S12). Folder 2: Polariod optical microscopy (POM) - containing png pictures with scale bars. Folder 3: Scanning electron microscopy (SEM) - original images in tiff (Figure S11). Folder 4: Cholesteric pitch determination - data in xlsx and pdf of the Figures 3, S9, S10, S16. Folder 5: Atomic force microscopy (AFM) - svgs of the AFM and png of the length distribution analysis (Figure S1). Folder 6: Droplet size determination - data in xlsx and pdf of the droplet size distribution (Figure S17) and evaporation rate (Figure S18). Folder 7: Simulation of cholesteric droplets - png of the simulation with computing parameters in tiff (Figure 2), data in xlsx and pdf of optical indices from Bruggeman method (Figure S19), matlab folder with codes in m format.

File formats and software
Microsoft Office 2007 or later, Acrobat reader, Matlab

Subjects
liquid crystals, microfluidics, colloidal self-assembly, cellulose nanocrystals, hierarchical architecture

Relationships
Publication Reference:
http://dx.doi.org/10.1021/acsnano.6b03355

Sponsorship
BBSRC [BB/K014617/1], Isaac Newton Trust Cambridge [76933], European Research Council [ERC-2014-STG H2020 639088]

Identifiers
This record's DOI:
http://dx.doi.org/10.17863/CAM.1233

http://dx.doi.org/10.17863/CAM.1233
Should commitment to sharing be a criterion for academic promotion and employment?

Changing hiring practices towards research transparency: The first open science statement in a professorship advertisement

January 6, 2016 Filed in: Open Science | Science

Engaging in open science practices increases knowledge as a common good, and ensures the reproducibility, verifiability and credibility of research. But some have the fear that on an individual strategic level (in particular from an early career perspective) engaging in research transparency could reduce a researcher's chance to get a tenured

http://www.nicebread.de/open-science-hiring-practices/
Ideas and results need to be shared to move human knowledge forward.

Publications without supporting data (and code!) are just claims.

We live in a digital world, so research data (and code!) can be shared.
Researchers – lead by example

Manage your data well
Share your research

Image from Pixabay, CC0 licence
Researchers – lead by example

#OpenInAction: I pledge to publish only in openaccess journals, + share my preprints, code, data, etc.
emckiernan.wordpress.com/pledge/ #OAWeek

I pledge to:

- edit and review only for open access journals
- publish only in open access journals
- openly share my working manuscripts
- openly share my code, when possible
- openly share my data, when possible
- openly share my notebooks, when possible
- ask my professional societies to support open research
- speak out in support of open research

https://emckiernan.wordpress.com/pledge/
New Year’s Resolutions: My Open Pledge

JANUARY 5, 2016 / HEIDI LAINE

Ever since I heard Erin McKiernan telling about her Open Pledge at the 2015 OpenCon, I have been thinking about making my own. It seems only appropriate to publish it at the beginning of a new year. In addition to McKiernan I have been inspired for example by the Open Science Peer Review Oath and the responsible conduct of research guidelines I’m studying.

My research is open by default.*

I will not be open at the cost of my research subjects privacy; when in doubt I will refrain from publishing.

https://thehonestbrokerblog.org/2016/01/05/new-years-resolutions-my-open-pledge/
Researchers – help others

Would you like to become a Data Champion?

We are looking for people who would like to volunteer to become Data Champions within their departments, institutes and colleges to provide local expertise about Research Data Management (RDM) issues.

Who can become an Data Champion?

Anyone interested in research data management and sharing who would like to become the local expert on RDM. We welcome researchers, students, librarians, IT managers, data

http://www.data.cam.ac.uk/datanews/call-data-champions
Cultural change is achievable

Image from Pixabay, CC0 licence
Cultural change is achievable

Wellcome Trust funded researchers share because it is a good thing

FIGURE 14. REASONS FOR MAKING RESEARCH DATA AVAILABLE

- My funder requires me to share my data (N=273)
- Journal expects data underpinning findings to be... (N=277)
- My research community expects data sharing (N=274)
- It is good research practice to share research data (N=277)
- It enables collaboration and contribution by other... (N=277)
- It has public health benefits, e.g. disease... (N=269)
- Ability to respond rapidly to public health... (N=269)
- Ethical obligation towards research participants to... (N=277)
- Contributes to academic credentials (N=273)
- Improved visibility for my research (N=273)
- I can get credit and more citations by sharing data (N=267)

https://figshare.com/articles/Survey_of_Wellcome_researchers_and_their_attitudes_to_open_research/4055448/1
Everyone can contribute to sharing

Shares data FAIRly:
- Findable
- Accessible
- Interoperable
- Re-usable

Data producer

Funding for new research

Data re-user

Data validation

New hypotheses and new research

Recognition and reward

Data citation

Image from Pixabay, CC0 licence
Cambridge & Wellcome Trust are calling for Open Research experiment participants

Start today

Open Research Pilot Project - call for participants - applications close on Sunday 11 December

Are you in favour of more transparency in research? Are you concerned about research reproducibility? Would you like to get better recognition and credit for all outputs of your research process? Would you like to open up your research and make it more available to others?

If you responded 'yes' to any of these questions, we would like to invite you to participate in the Open Research Pilot Project, organised jointly by the Open Research team at the Wellcome Trust and the Office of Scholarly Communication at the University of Cambridge.

What is the Open Research Pilot Project?

The Open Research Pilot Project is a two year long project which will look at gaining an

http://www.data.cam.ac.uk/datanews/call-participants-open-research-pilot-project
Thank you

Questions/keep in touch

@martateperek

mt446@cam.ac.uk