The power and the promise of Infrastructure for Open Research

Mark Hahnel, CEO

@figshare
Step 1. Getting the files on the web
NESS, REPRODUCIBILITY AND ECOLOGICAL CONSISTENCY

DEMARICATION OF OPERATIONAL TAXONOMIC UNITS

Z. Schmidt, J. Matias Rodrigues, C. von Mering

Department of Molecular Life Sciences, University of Zürich, Switzerland

J robustness, reproducibility & ecological consistency

Zherma Palca, Maria Johnson, Wesley Davidson, Zoe (2015): WT Tm 1 moves away from the nascent T0 after APC stimulation. figshare.

https://dx.doi.org/10.5072/FK2.figshare.2001585

Retrieved 15:24, Aug 14, 2015 (GMT)
co-occurrence networks

A notebook creates co-occurrence networks and exports them to .gexf-files, either at the start of a new book or chapter, or as a network for all selected books or chapters.

ser variables

which problem does co-occurrence networks for a given passage solve? Is the last chapter/verse useful when selecting a passage?

what range do the co-occurrence networks show?

Explores the context of Biblical narrative

Scanned with a Nextengine Desktop 3D Scanner and Scan Studio Pro (NextEngine) on high resolution settings. Model composed of 72,969 vertices and 145,758 faces. Saved as an .stl file in MeshLab (v1.3.2).

Mapping the Structure of the Archaeological Web

Version 2
29.04.2014, 17:03 by Shawn Graham

A fileset to accompany an article in a special issue of Internet Archaeology. In this article, I map the structure of the web to understand the context of archaeological blogging.
Why Publishers?

• Path of least resistance
• Author Services

Why figshare?

• Compliance
• Technology
• Pace of change
• APIs
you don’t have an API, you may as well not exist

cs.figshare.com

n clone whole public site
All metadata (JSON-LD, OAI-PMH, RDF)
Can pull all files

metrics available on everything
Time-stamped
Geotagged

Notable Users
- PubMed
- Clarivate Data Citation Index
- Research Data Australia
- SHARE
FAIR
A Comparative Review of Various Data Repositories

https://docs.google.com/spreadsheets/d/1KptHzDHIdB3s1v5m1mMwphecwXhOVVdkRYdjEWW1dc/edit#gid=355072175
The State of Open Data
State of Open Data 2017

Public Funder
Jean-Claude Burgelman, European Commission

Private Funder
Robert Kiley & David Carr, Wellcome Trust

Repository Space
Mark Hahnel, Figshare

University
Dale Peters, University of Cape Town

Publisher
Grace Baynes, Springer Nature

2359 responses
Researchers awareness of Open Data

<table>
<thead>
<tr>
<th>Region</th>
<th>2016</th>
<th>2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asia &amp; Middle East</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australasia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Europe</td>
<td></td>
<td></td>
</tr>
<tr>
<td>North &amp; Central America</td>
<td></td>
<td></td>
</tr>
<tr>
<td>South America</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Percentage values for each region in 2016 and 2017.
What motivates researchers to share the data?

- 24% Increased impact and visibility of my research
- 20% Public benefit
- 13% Other
- 11% Transparency and re-use
- 9% Getting proper credit for sharing data
- 8% Funder requirement
- 8% Journal/publisher requirement
- 7% Trust the person requesting data
Who owns the data?

- My institution
- Myself
- Colleagues
- Funder
- Publisher
- Other
- Unsure

- After publication
- Before publication
Other Takeaways

tatus
79% of researchers have made data openly available
Majority of researchers support a national open data mandate

nectives
Researchers are most motivated to share data by increasing their profile and doing things for the public benefit
transparency and re-use over publisher and funder mandates
77% of researchers value a data citation as much or more than an article citation (validating last years findings
92% of researchers motivated to share their data if it leads to citation

ucation
Majority of researchers think they don’t have a publisher, funder or institutional mandate to share data
Majority of researchers think they own their data
Majority of researchers don’t understand creative commons licenses
Only approx 25% of researchers look to their library for advice on data sharing
Majority of researchers still using physical media to store and archive their data
83% of reported data loss happened as a result of using physical media
out of 3 ‘Aint Bad.
\textbf{Will Open Research Make it Through Those Generating Research?}
tep 2. Education as priority?
tep 2. Curation?

**Mechanical Chameleon through Dynamic Real-Time Plasmonic Tuning**

15.02.2015, 13:40 by Hua Wang, Xu Fan, Xu Wang

The development of camouflage methods, often through a general
FAIR

FAIR
| Mission (EC) | 1,413 | 2.89 | 1.01 | 73.6 | 5.5 |
| Search Council (ERC) | 1,024 | 3.50 | 1.17 | 82.9 | 10.0 |
| Environment Research Council (NERC) | 173 | 3.39 | 1.16 | 93.6 | 14.5 |
| None, Industry and Competitiveness (MI) | 155 | 2.78 | 0.98 | 78.7 | 8.0 |
| Research Foundation (DFG) | 136 | 3.00 | 1.09 | 72.1 | 7.0 |
| Science Foundation (SNF) | 128 | 4.03 | 1.15 | 88.3 | 10.5 |
| Physical Sciences Research Council (EPSRC) | 121 | 2.78 | 0.83 | 68.1 | 3.0 |
| Life (WT) | 120 | 4.70 | 2.48 | 91.7 | 10.0 |
| ANR | 111 | 3.54 | 1.21 | 77.5 | 6.0 |
| and Biological Sciences Research Council | 104 | 5.01 | 1.60 | 88.5 | 13.0 |
| EC | 469 | 381 | 87 | 56 | 0 | 62 | 0 | 0 | 39 |
| ERC | 339 | 193 | 56 | 42 | 0 | 29 | 0 | 0 |
| NERC | 214 | 81 | 25 | 25 | 0 | 10 | 0 | 0 |
| MINECO | 63 | 95 | 19 | 10 | 0 | 15 | 0 | 0 |
| DFG | 40 | 83 | 0 | 0 | 12 | 0 | 0 | 9 |
| SNF | 63 | 50 | 0 | 0 | 16 | 0 | 0 | 7 |
| EPSC | 75 | 31 | 0 | 0 | 7 | 8 | 0 | 10 |
| WT | 52 | 82 | 0 | 0 | 11 | 6 | 7 | 15 |
| Plant Biology | 66 | 9 | 0 | 0 | 3 | 0 | 0 | 3 |
| Microbiology | 46 | 27 | 10 | 9 | 5 | 5 | 5 | 0 |
| 104 Statistics | 26 | 37 | 7 | 5 | 5 | 5 | 0 | 4 |
| 1051 Ecological Applications | 32 | 26 | 0 | 6 | 3 | 4 | 5 | 2 |
| 1108 Medical Microbiology | 32 | 15 | 10 | 5 | 0 | 0 | 0 | 0 |
| 0299 Other Physical Sciences | 25 | 25 | 0 | 5 | 5 | 5 | 7 | 0 |
| 1117 Public Health and Community Medicine | 31 | 10 | 0 | 0 | 0 | 0 | 0 | 9 |
| 0306 Physical Chemistry | 23 | 26 | 0 | 0 | 0 | 0 | 15 | 0 |
| 0803 Computer Software and Systems | 32 | 6 | 0 | 0 | 3 | 0 | 7 | 0 |
| 0705 Forestry Sciences | 25 | 16 | 10 | 4 | 0 | 0 | 0 | 0 |
| 0403 Geology | 22 | 18 | 0 | 5 | 0 | 0 | 4 | 0 |
e expanding universe
New Universe
National Non-Traditional Research Output Portal

Region or subject specific 'Data' Portals

Portals for each grant code

normalisation and mapping to Funded Grant Code and Institution using GRIB, UberResearch and FundRef

Funded Researcher Uploads
Academic Publisher Datasets
from e.g. SpringerNature, ACS-Valley, etc.

Migration or ingest of data from other Funded Repositories, or Institutional Data Repositories

Migration or ingest of data from popular repositories e.g. Pashare, Dryad, Zenodo
Mark Hahnel
Function: CEO
email: mark@figshare.com
@markhahnel
@figshare
Some of team figshare’s beliefs

Academic research outputs should be as open as possible, as closed as necessary

Academic research outputs should never be behind a paywall

Academic research outputs should be human and machine readable/query-able

Academic infrastructure should be interchangeable

Academic researchers should never have to put the same information into multiple systems at the same institution

Identifiers for everything

The impact of research is independent of where it is published and what type of output it is