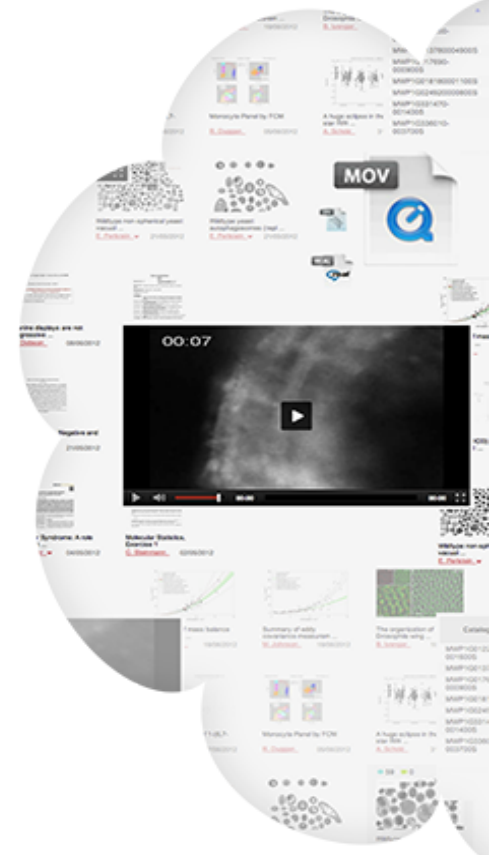
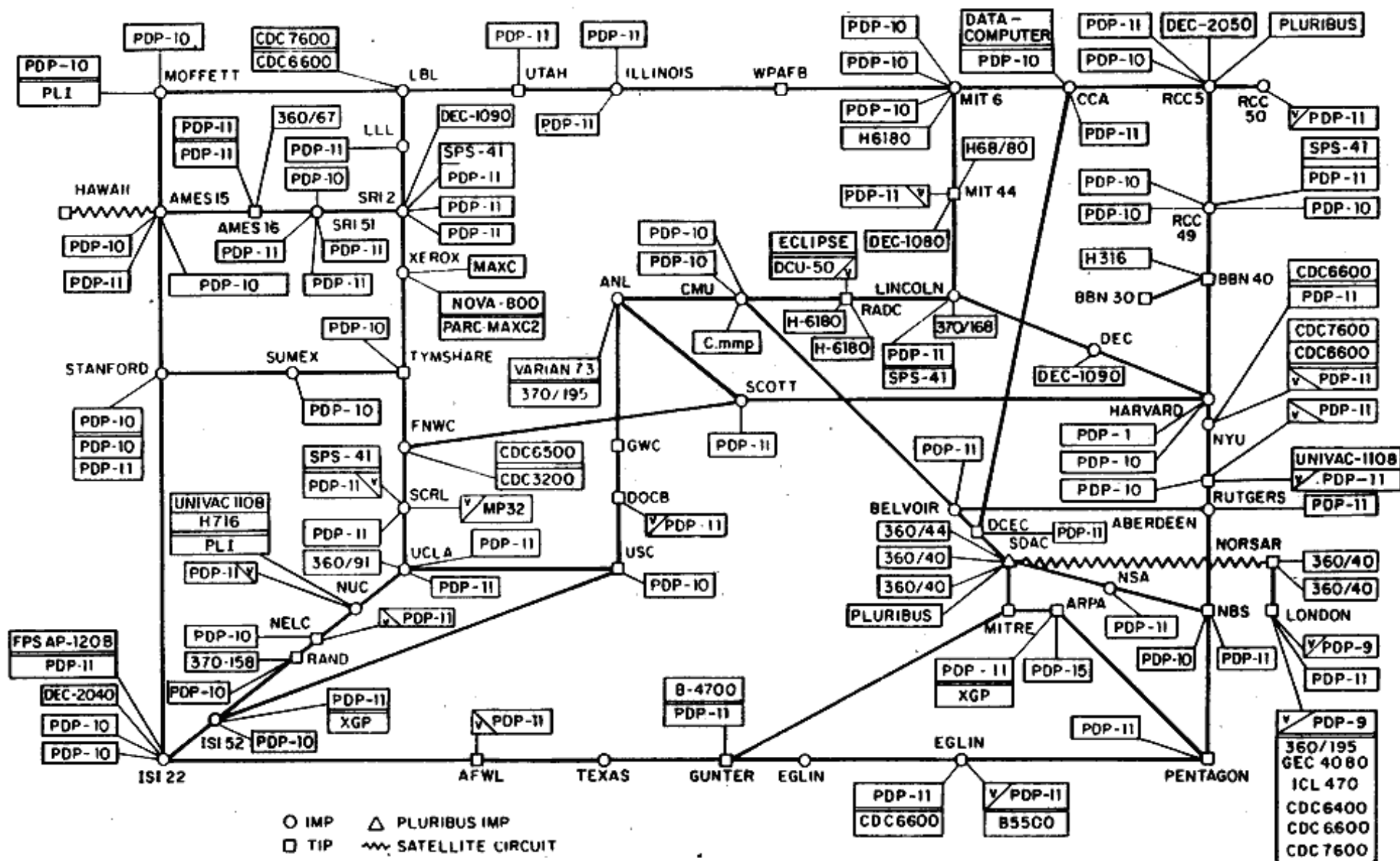


# Turning Funder Data Mandates Into Opportunities

@MarkHahnel  
@figshare



# ARPANET LOGICAL MAP, MARCH 1977



(PLEASE NOTE THAT WHILE THIS MAP SHOWS THE HOST POPULATION OF THE NETWORK ACCORDING TO THE BEST INFORMATION OBTAINABLE, NO CLAIM CAN BE MADE FOR ITS ACCURACY)

NAMES SHOWN ARE IMP NAMES, NOT (NECESSARILY) HOST NAMES



**Institutions**  
**Generate the world's knowledge**

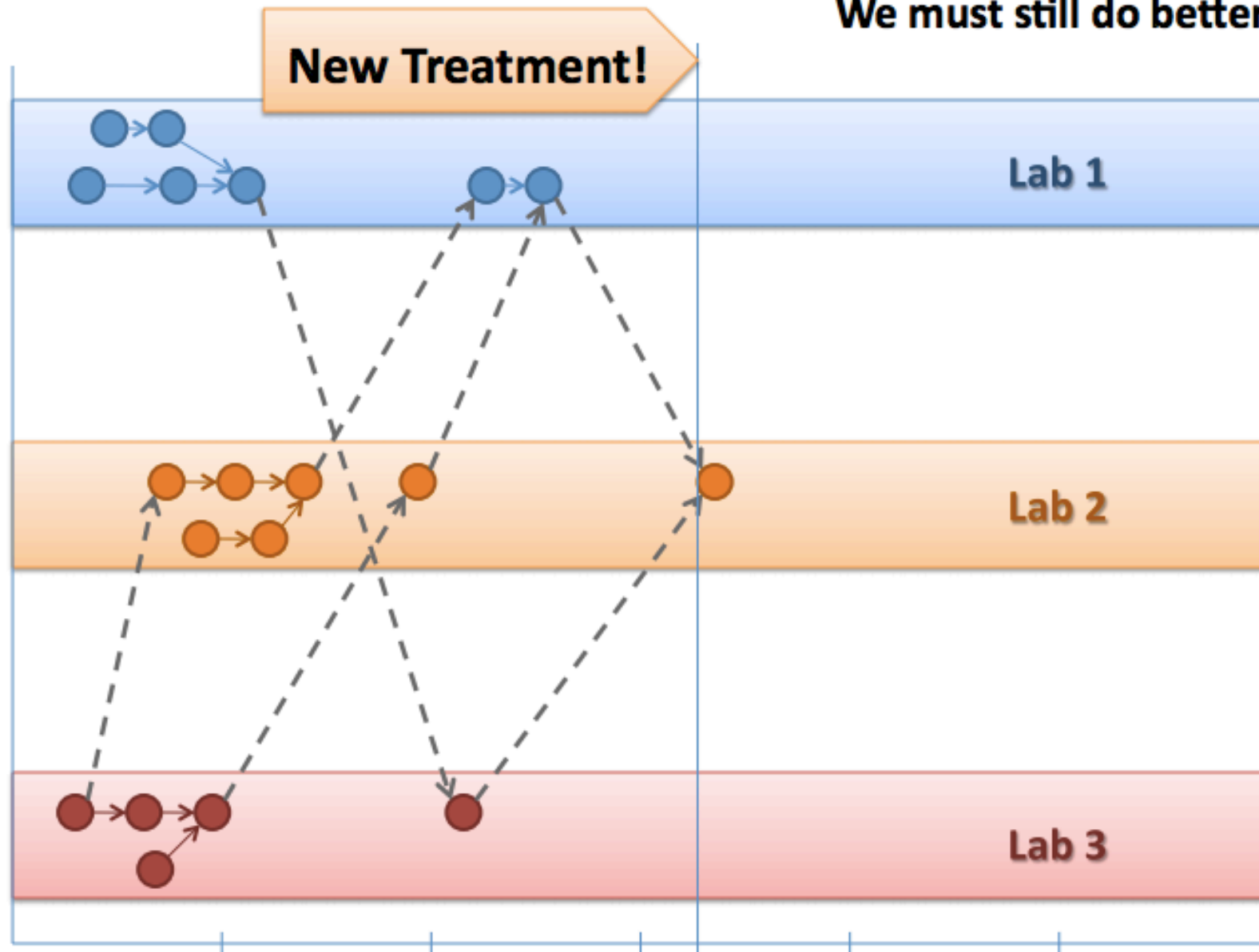
# Who are the stakeholders and what is their motivation?

- Government
- Funders
- Societies
- Publishers
- Institutions
- Academics

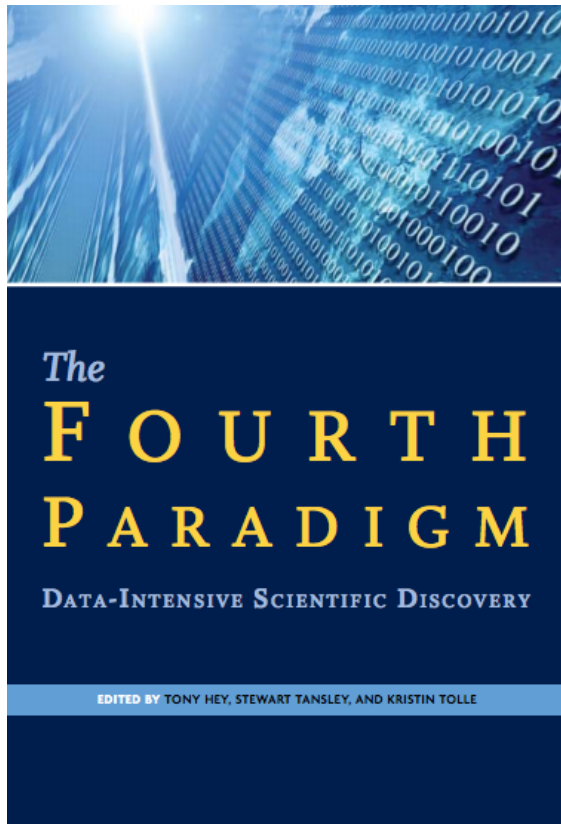




**We must still do better**

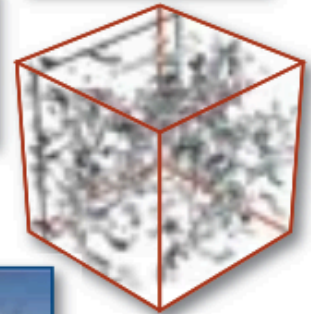


# Science Paradigms



- Thousand years ago:  
science was **empirical**  
*describing natural phenomena*
- Last few hundred years:  
**theoretical** branch  
*using models, generalizations*
- Last few decades:  
a **computational** branch  
*simulating complex phenomena*
- Today: **data exploration** (eScience)  
*unify theory, experiment, and simulation*
  - Data captured by instruments or generated by simulator
  - Processed by software
  - Information/knowledge stored in computer
  - Scientist analyzes database/files using data management and statistics

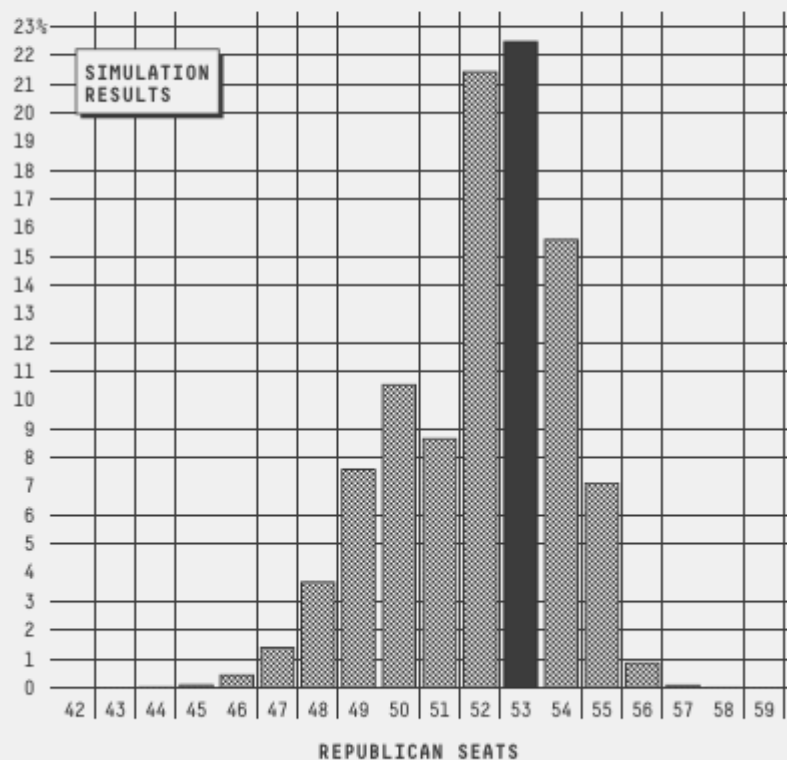
$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{4\pi G\rho}{3} - K\frac{c^2}{a^2}$$



[OVERVIEW](#)
[ELECTIONS](#)
[POLL](#)
[MATH](#)

## The Model's Output

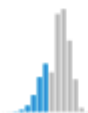
DATE OF RUN: 11/04/14  
 POLLS ANALYZED: 1670  
 DAYS UNTIL ELECTION: 0



## What It Means



Republicans have a **76.2%** chance of winning a majority.



Democrats have a **23.8%** chance of keeping the majority.



There is a **22.5%** chance Republicans will control **53** seats and Democrats will control **47** seats.

## How It's Changed



Article outline ☐ Show full outline

## Abstract

1. Introduction
  2. The ATLAS detector
  3. Signal and background simulation sa...
  4.  $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$  channel
  5.  $H \rightarrow \gamma\gamma$  channel
  6.  $H \rightarrow WW^{(*)} \rightarrow e\nu\mu\nu$  channel
  7. Statistical procedure
  8. Correlated systematic uncertainties
  9. Results
  10. Conclusion
- Acknowledgements  
Open access  
References

## Figures and tables


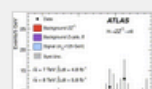
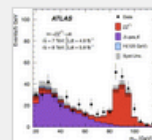
 Table 1

 Table 2



ScienceDirect is currently experiencing a technical issue within our User Registration service whereby access to this piece of content is temporarily denied. We are working to restore this service as soon as possible and apologize for this inconvenience!



## Physics Letters B

Volume 716, Issue 1, 17 September 2012, Pages 1–29



## Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC ☆

This paper is dedicated to the memory of our ATLAS colleagues who did not live to see the full impact and significance of their contributions to the experiment.

### ATLAS Collaboration\*

G. Aad<sup>48</sup>, T. Abajyan<sup>21</sup>, B. Abbott<sup>111</sup>, J. Abdallah<sup>12</sup>, S. Abdel Khalek<sup>115</sup>, A.A. Abdelalim<sup>49</sup>, O. Abdinov<sup>11</sup>, R. Aben<sup>105</sup>, B. Abi<sup>112</sup>, M. Abolins<sup>88</sup>, O.S. AbouZeid<sup>158</sup>, H. Abramowicz<sup>153</sup>, H. Abreu<sup>136</sup>, B.S. Acharya<sup>164a, 164b</sup>, L. Adamczyk<sup>38</sup>, D.L. Adams<sup>25</sup>, T.N. Addy<sup>56</sup>, J. Adelman<sup>176</sup>, S. Adomeit<sup>98</sup>, P. Adragna<sup>75</sup>, T. Adye<sup>129</sup>, S. Aefsky<sup>23</sup>, J.A. Aguilar-Saavedra<sup>124b, a</sup>, M. Agustoni<sup>17</sup>, M. Aharrouche<sup>81</sup>, S.P. Ahlen<sup>22</sup>, F. Ahles<sup>48</sup>, A. Ahmad<sup>148</sup>, M. Ahsan<sup>41</sup>, G. Aielli<sup>133a, 133b</sup>, T. Akdogan<sup>19a</sup>,

[+ Show more](#)

doi:10.1016/j.physletb.2012.08.020

[Get rights and content](#)

[Open Access](#)

### ▼ Recommended articles

[Observation of a new boson at a mass of 125 GeV wi...](#)

2012, Physics Letters B [more](#)

[Combined search for the Standard Model Higgs boso...](#)

2012, Physics Letters B [more](#)

[Broken symmetries, massless particles and gauge fie...](#)

1964, Physics Letters [more](#)

[View more articles »](#)

### ► Citing articles (1888)

### ► Related book content

Table 1.

Event generators used to model the signal and background processes. "PYTHIA" indicates that PYTHIA6 and PYTHIA8 are used for simulations of  $\sqrt{s} = 7$  TeV and  $\sqrt{s} = 8$  TeV data, respectively.

Process	Generator
ggF, VBF	POWHEG [57] and [58] + PYTHIA
$WH, ZH, t\bar{t}H$	PYTHIA
$W + \text{jets}, Z/\gamma^* + \text{jets}$	ALPGEN [59] + HERWIG
$t\bar{t}, tW, tb$	MC@NLO [60] + HERWIG
$tqb$	AcerMC [61] + PYTHIA
$q\bar{q} \rightarrow WW$	MC@NLO + HERWIG
$gg \rightarrow WW$	gg2WW [62] + HERWIG
$q\bar{q} \rightarrow ZZ$	POWHEG [63] + PYTHIA
$gg \rightarrow ZZ$	gg2ZZ [64] + HERWIG
$WZ$	MadGraph + PYTHIA, HERWIG
$W\gamma + \text{jets}$	ALPGEN + HERWIG
$W\gamma^*$ [65]	MadGraph + PYTHIA
$q\bar{q}/gg \rightarrow \gamma\gamma$	SHERPA

Table options ▲

View in workspace

Download as CSV

The following parton distribution function (PDF) sets are used: CT10, gg2WW and gg2ZZ samples; CTEQ6L1 [79] for the PYTHIA8, ALPGEN and SHERPA samples; and MRSTMCa1 [80] for the PYTHIA6 samples.

Acceptances and efficiencies are obtained mostly from full simulations of the ATLAS detector [81] using

[next table »](#)

Table 1.


Event generators used to model the signal and background processes. "PYTHIA" indicates that PYTHIA6 and PYTHIA8 are used for simulations of  $\sqrt{s} = 7$  TeV and  $\sqrt{s} = 8$  TeV data, respectively.

Process	Generator
ggF, VBF	POWHEG [57] and [58] + PYTHIA
$WH, ZH, t\bar{t}H$	PYTHIA
$W + \text{jets}, Z/\gamma^* + \text{jets}$	ALPGEN [59] + HERWIG
$t\bar{t}, tW, tb$	MC@NLO [60] + HERWIG
$tqb$	AcerMC [61] + PYTHIA
$q\bar{q} \rightarrow WW$	MC@NLO + HERWIG
$gg \rightarrow WW$	gg2WW [62] + HERWIG
$q\bar{q} \rightarrow ZZ$	POWHEG [63] + PYTHIA
$gg \rightarrow ZZ$	gg2ZZ [64] + HERWIG
$WZ$	MadGraph + PYTHIA, HERWIG
$W\gamma + \text{jets}$	ALPGEN + HERWIG
$W\gamma^*$ [65]	MadGraph + PYTHIA
$q\bar{q}/gg \rightarrow \gamma\gamma$	SHERPA

[View in article](#)



SEARCH

 Home



BUSINESS MODELS, COMMERCE, ECONOMICS, INNOVATION, OPEN ACCESS, TOOLS, WORLD OF TOMORROW

## Ask The Chefs: What Do You Think Will Have The Biggest Impact On Scholarly Publishing In 2015?

“According to the Chefs, we’re looking at a year of mergers and acquisitions, the continuing growth of open access both in number of opportunities and in scale, **the publication of data and objects (like multimedia, application code, etc.)**, and more start-ups.”



# **FUNDER COMPLIANCE**



**“But taxpayers who are paying for that research will want to see something back. Directly – through open access to results and data. And indirectly – through making science work better for all of us.**




**That’s why we will require open access to all publications stemming from EU-funded research. That’s why we will progressively open access to the research data, too. And why we’re asking national funding bodies to do the same.”**




































































































**Neelie Kroes.**

**Vice President for the European Commission**





 Full Coverage
  Partial Coverage
  No Coverage

	Policy Coverage		Policy Stipulations					Support Provided			
Research Funders	Published outputs	Data	Time limits	Data plan	Access/sharing	Long-term curation	Monitoring	Guidance	Repository	Data centre	Costs
AHRC											
BBSRC											
CRUK											
EPSRC											
ESRC											
MRC											
NERC											
STFC											
Wellcome Trust											



“The Obama Administration is committed to the proposition that citizens deserve easy access to the results of scientific research their tax dollars have paid for. That’s why, in a policy memorandum released today, OSTP Director John Holdren has directed Federal agencies with more than \$100M in R&D expenditures to develop plans to make the published results of federally funded research freely available to the public within one year of publication and **requiring researchers to better account for and manage the digital data resulting from federally funded scientific research.**”

February 22nd 2013

“NIH intends to make public access to digital scientific data the standard for all NIH-funded research, as well as exploring ways to advance data as a legitimate form of scholarship through data citation and other means.”

#### The AHRQ Public Access Plan

“AHRQ will require all primary data collected by grant, contract, or intramural research to be submitted to a data repository, unless confidentiality restrictions prevent the data from being made public.”

#### The FDA Public Access Plan

“The FDA has plans for data storage, archiving, public access, and long-term preservation, including a description of the way in which shared digital data will be discoverable, retrievable, and analyzable”

#### The ASPR Public Access Plan

“All ASPR-funded researchers will be required to make the data underlying the conclusions of peer-reviewed scientific research publications freely available in public repositories at the time of initial publication in machine readable formats.”

#### The CDC Public Access Plan

“CDC has shared and continues to share data in accordance with its policy will encourage the use of established public repositories and community based standards.”

Q

NIH

x

Projects	Funders	Researchers	Organizations	Places	Categories
----------	---------	-------------	---------------	--------	------------

Results | Visualizations

10,825

Projects

$\Sigma$

\$ 22.2 B

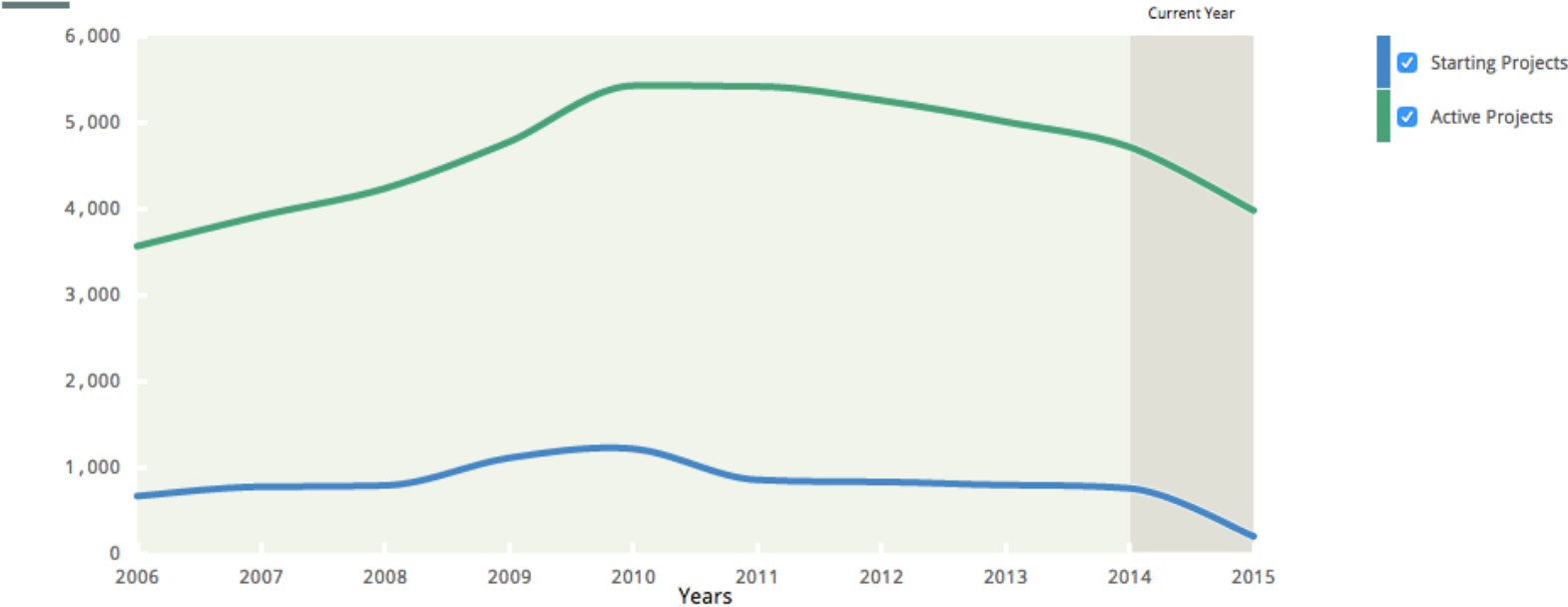
Aggregated funding amount

$\bar{x}$

\$ 2.0 M

Average funding amount




Active and starting projects ▾
















































































































































“NIH intends to make **public access to digital scientific data the standard for all NIH-funded research**, as well as exploring ways to advance data as a legitimate form of scholarship through data citation and other means.”

# Overview of OSTP Responses

 Full Coverage
  Partial Coverage
  No Coverage

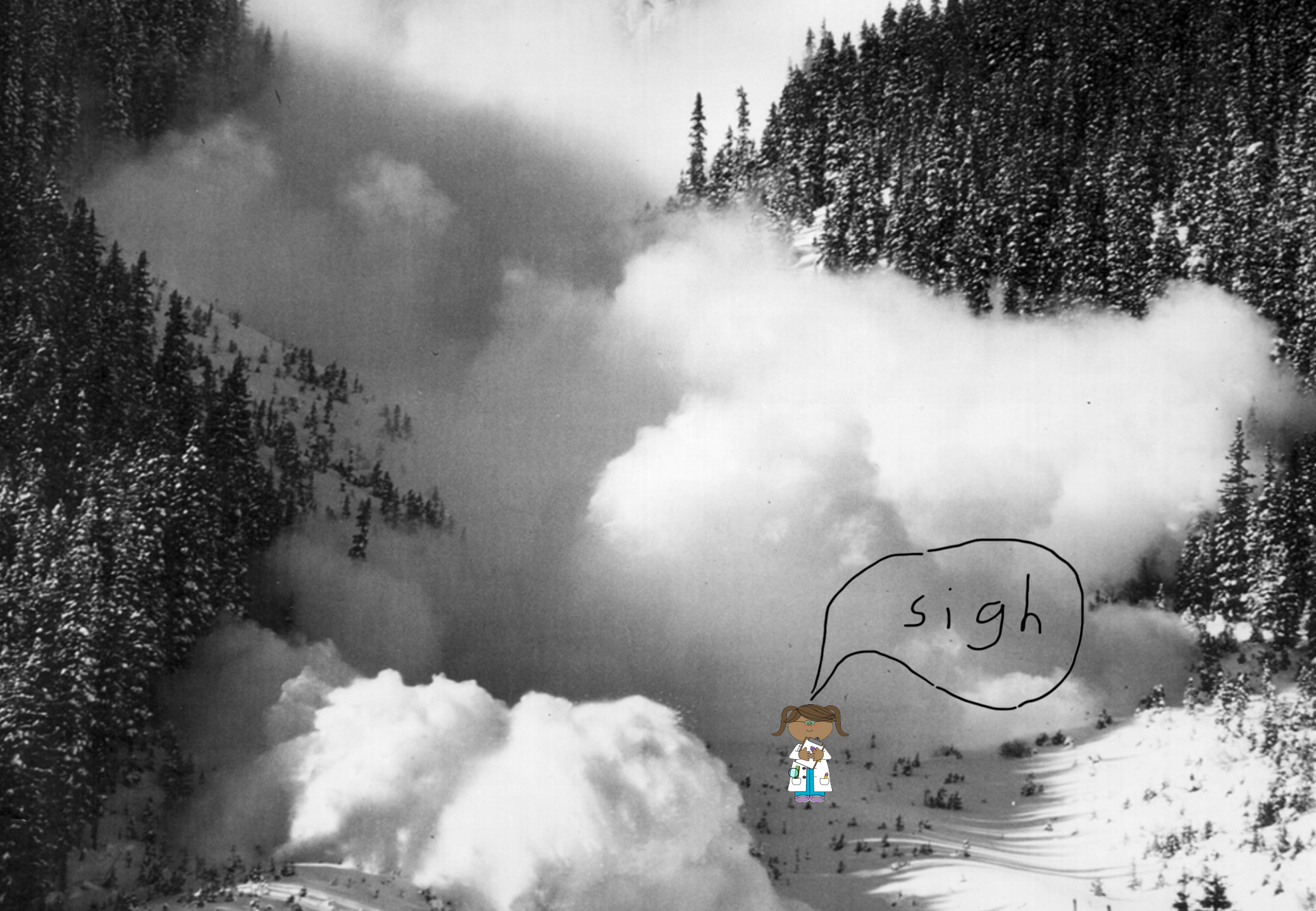
	Policy Coverage		Policy Stipulations					Support Provided			
Funder	Published Outputs	Data	Time Limits	DMP	Access / Sharing	Long-term curation	Monitoring	Guidance	Article Repository	Data Repository	Costs
AHRQ											
ASPR											
CDC											
DOD											
DOE											
DOT											
FDA											
NASA											
NIH											
NIST											
NOAA											
NSF											
USDA											
USAID											
USGS											
VA											

Last updated: May 7, 2015



## The Open Academic Tidal Wave

1. **Recommended** open access to **scholarly papers** of publicly funded research
2. **Recommended** open access to all **digital outputs** of publicly funded research
3. **Mandated** open access to **scholarly papers** of publicly funded research
4. **Mandated** open access to all **digital outputs** of publicly funded research
5. **Enforced**, mandated open access to **scholarly papers** of publicly funded research
6. **Enforced**, mandated open access to all **digital outputs** of publicly funded research



sigh







# Dreadnoughtus schrani 3D PDF images - Lacovara et al., 2014, A Gigantic, Exceptionally Complete Titanosaurian Sauropod Dinosaur from Southern Patagonia, Argentina, Scientific Reports,

[News](#)[Blogs](#)[Twitter](#)[Facebook](#)[Google+](#)[Video](#)[Demographics](#)[Help](#)

So far Altmetric has seen 5 stories from 4 outlets.

## Score in context

Is one of the highest ever scores in this journal (ranked #5 of 4,522)

[show more...](#)

## Mentioned by

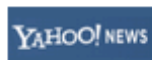
- 4 news outlets
- 2 blogs
- 154 tweeters
- 2 Facebook users
- 2 Google+ users
- 1 video uploader

## Readers on

- 0 Mendeley
- 0 CiteULike

## Track this article

- Get email updates when this article is shared

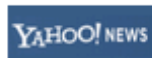


### Gigantic Dreadnoughtus Dinosaur Weighed Whopping 65 Tons, Feared Nothing

Yahoo! News

A gargantuan, long-necked dinosaur as big as a two-story house and weighing as much as 12 elephants once stalked a flower-dotted ..

2014-09-04T14:59:34+01:00



### Dreadnoughtus Dinosaur Weighed Whopping 65 Tons, Feared Nothing

Yahoo! News

A gargantuan, long-necked dinosaur as big as a two-story house and weighing as much as 12 elephants once stalked a flower-dotted ..

2014-09-04T14:59:34+01:00



### Dino-hunters discover Dreadnoughtus - maybe the biggest creature to have walked

The Register

Finders make 3D scans of bones available for download Paleontologists have discovered the remains of a new type of dinosaur, dub ..

2014-09-04T18:22:59+01:00

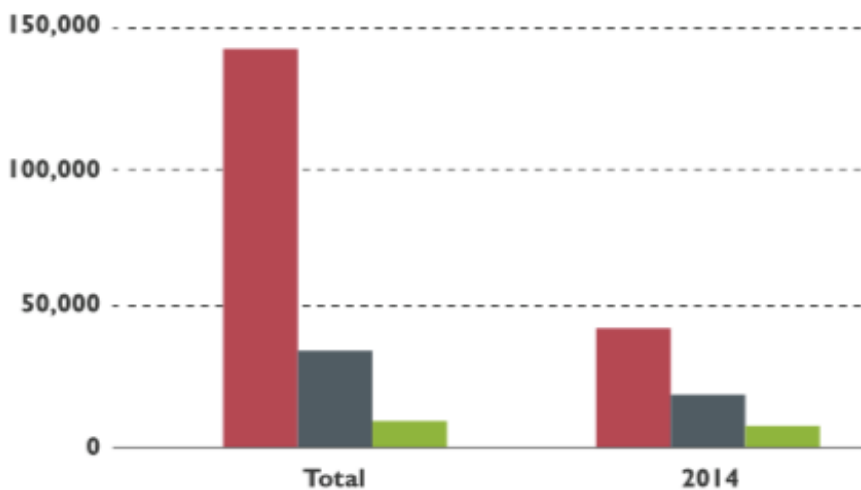


### Giant dinosaur discovered! Dreadnoughtus weighed as much as 7 T-Rexes

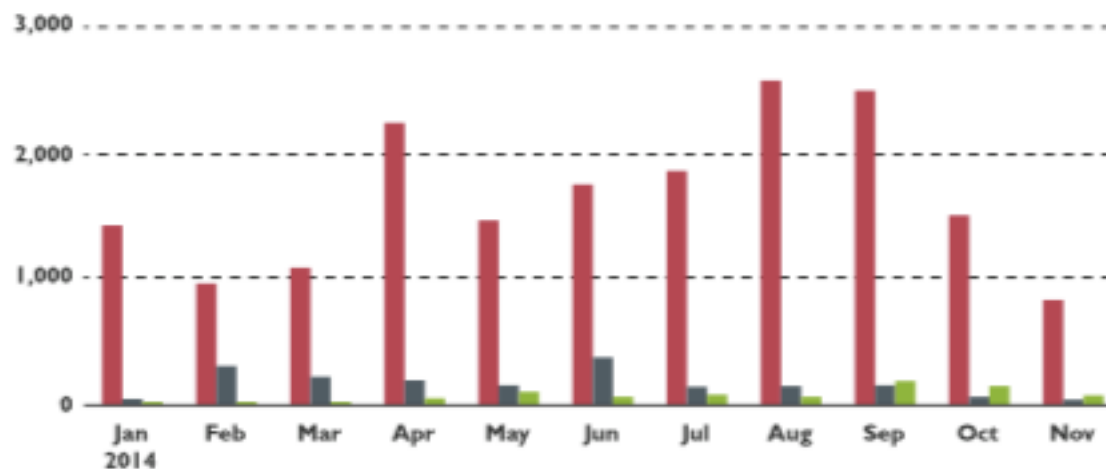
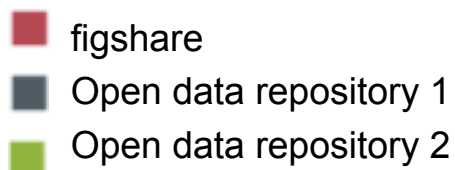
Los Angeles Times

Scientists have just announced the discovery of Dreadnoughtus schrani, a hulking dinosaur that stood two stories tall at its sho ..

# figshare content on DataCite



Number of DataCite DOI registrations



figshare-hosted file views

Source is DataCite (<http://stats.datacite.org>), covers all DOI registrations until December 18th, 2014. We calculated the median value for DOIs registered by the Top 20 research data centers registered in DataCite

Median views of the top ten data files from each platform according to DataCite DOI viewing statistics





figshare for institutions

## Get more citations & impact for your work

Melbourne University has invested in this infrastructure to help your research data go further.



Google

10.6084/m9.figshare.104629

Scholar

11 results (0.05 sec)

Articles

Case law

My library

Any time

Since 2015

Since 2014

Since 2011

Custom range...

Sort by relevance

Sort by date

☒ include patents

☒ include citations

**Altmetrics: Value all research products**

[H Piwowar](#) - *Nature*, 2013 - [nature.com](#)

... Priem, J., Costello, K. & Dzuba, T. Figshare

**104629** (2012). Fausto, S. et al. PLoS ONE 7,

Cited by 112 Related articles All 10 versions

**[PDF] When will librarians start research**

[M Readership](#) - [researchgate.net](#)

... Poster at Metrics 2011: Symposium on Inform

LA, USA, October 12. doi: **10.6084/m9.figshare**

Eysenbach G. Can Tweets Predict Citations? I

Cite Save More

**The role of Twitter in the life cycle of**

[ES Darling](#), [D Shiffman](#), [IM Côté](#), [JA Drew](#) - ar

Page 1. 1 The role of twitter in the life cycle of

David Shiffman2, Isabelle M. Côté1, Joshua A

Group, Department of Biological Sciences, Sin

Cited by 38 Related articles All 14 versions

**From Attention to Citation, What an**



## Geoscience Data Journal

© 2013 John Wiley & Sons Ltd.

Each article is made available under the term:



Edited By: Dr Rob Allan  
Online ISSN: 2049-6060  
Associated Title(s): Atmospheric Climatology, Meteorological Society,



# Article Processing Charges

Below are the Article Processing Charges (APCs) for *F1000Research*.\*

ARTICLE TYPE	CHARGE
Research Articles, Method Articles, Study Protocols, Reviews, Systematic Reviews	US \$1000 <sup>†</sup>
Software Tool Articles, Research Notes, Antibody Validation Articles, Observation Articles, Opinion Articles, Clinical Practice Articles	US \$500 <sup>†</sup>
Data Notes, Case Reports, Correspondence	US \$250 <sup>#</sup>

\* UK customers will also be subject to pay VAT on these fees.

<sup>†</sup> These fees include up to 1 GB of data.

<sup>#</sup> The fee includes up to 10 MB of data; for up to 1 GB of data, we will need to charge an additional US \$100 to cover the storage costs.





	A	B	C
Breast Miller (GSE3494)		$P=3.588 \times 10^{-74}$	
Gene Symbol		FC (vs G1 vs vs G2)	Probe Set
MMP1		2.4602	204475_at
TOP2A		1.9777	201291_s
S100P		1.932	204361_at
RRM2		1.8386	201890_at
CXCL11		1.8003	211122_s
CDK1		1.7674	203213_at
CDC20		1.7632	202870_s

## 4 Key Modules

1

### **Research Data Management**

Private, controlled storage and collaborative spaces for every academic at the institution.

2

### **Public Digital Research Repository**

A customizable public portal with all digital files made public at an institutional, departmental and group level.

3

### **Administrative Workflow Portal**

A portal where administrators can manage curation of files to be made public, storage space allocation and user rights.

4

### **Reporting Dashboard**

Impact and Usage Reporting.



# Admin **all groups** (chose a group)

**GROUPS** | [USERS](#)

GROUPS' STRUCTURE ☐ collapse all

USERS

PROJECTS

STORAGE

ACTIONS

LILLIPUT UNIVERSITY

36 | 7

0

1.48 TB

8.70 TB



☐ Aeronautical, Automotive...

8 | 2

0

240.00 GB

3

Configure

Create subgroup

Aeronautical and Autom...

2

0

0 B

8

Chemical Engineering

2

0

0 B

80.00 GB



Materials

2

0

0 B

80.00 GB



Civil and Building Engine...

2

0

0 B

80.00 GB



☐ Humanities

6 | 2

0

160.00 GB

240.00 GB



Economics

2

0

0 B

80.00 GB



English and Drama

2

0

0 B

80.00 GB







## The FAIR data Guiding Principles

*DRAFT for final editing by the wider community September 1, 2014*

For all parties involved in Data Stewardship, the facets of FAIRness, described below, provide incremental guidance regarding how they can benefit from moving toward the ultimate objective of having all concepts referred-to in Data Objects (Meta data or Data Elements themselves) unambiguously resolvable for machines, and thus also for humans.

By adopting all FAIR facets, Data Objects become fully:

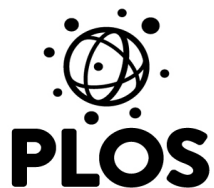
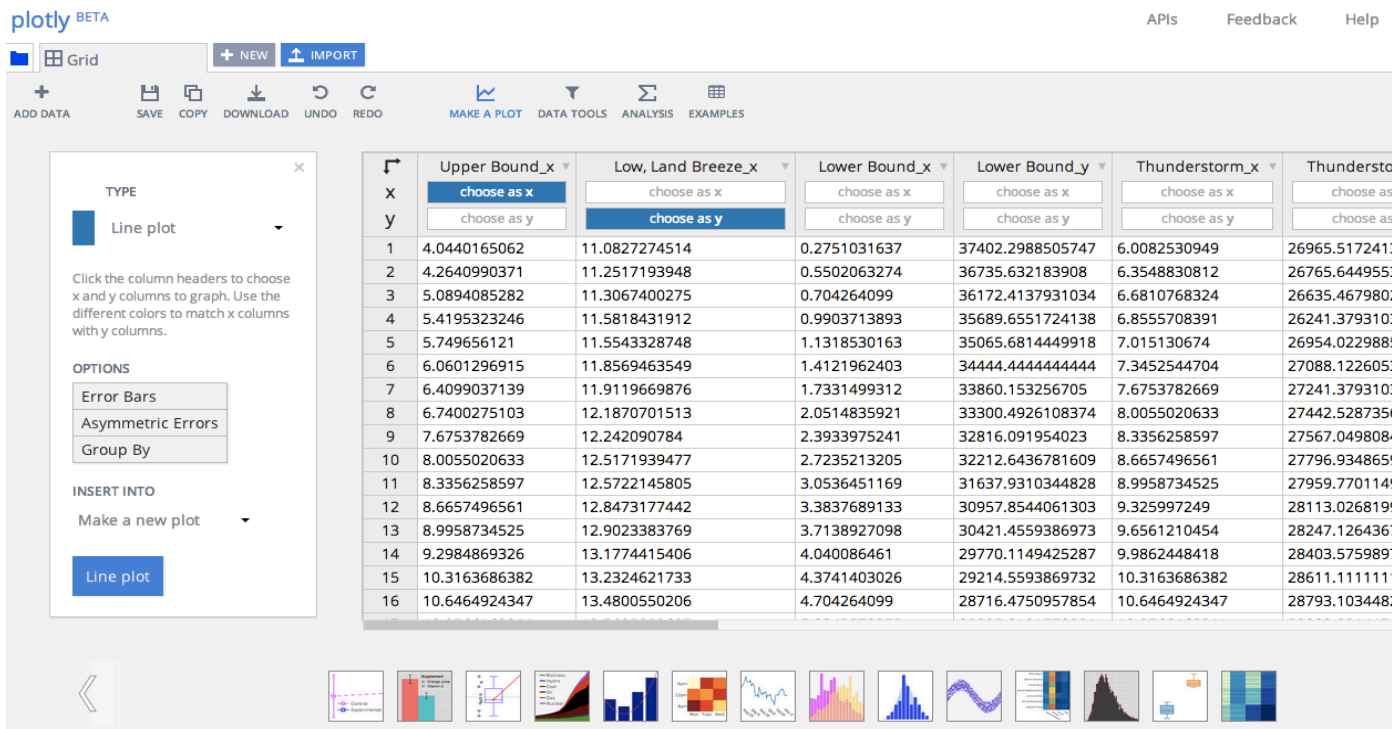
### Findable, Accessible, Interoperable, and Reusable

#### Definitions

- A **Concept** is any defined 'unit of thought' to which we refer in our digital formats [1]
- A **Data Object** is defined for the purpose of the principles below as: An **Identifiable Data Item with Data elements + Metadata + an Identifier** [2]
- When we use the term **(Meta) data** here, we intend to indicate that the principle is true for **Metadata** as well as for the actual, collected **Data Elements** in the Data Object, but that the principle in question can be independently implemented for each of them [3].

#### FAIR Guiding Principles

1. To be **Findable** any Data Object should be uniquely and persistently identifiable [4]
  - 1.1. The same Data Object should be re-findable at any point in time, thus Data Objects should be **persistent**, with emphasis on their metadata, [see [JCDP 4](#) and [6](#)]
  - 1.2. A Data Object should minimally contain basic machine readable metadata that allows it to be distinguished from other Data Objects [see [JCDP 5](#)]
  - 1.3. Identifiers for any concept used in Data Objects should therefore be **Unique** and **Persistent** [5 and [JCDP 4](#) and [6](#)].
2. Data is **Accessible** in that it can be always obtained by machines and humans
  - 2.1 Upon appropriate authorization [6]
  - 2.2 Through a well-defined protocol [7 and [JCDP 5](#)]
  - 2.3 Thus, machines and humans alike will be able to judge the actual accessibility of each Data Object.
3. Data Objects can be **Interoperable** only if:
  - 3.1. (Meta) data is machine-readable [8]
  - 3.2. (Meta) data formats utilize shared vocabularies and/or ontologies [9]
  - 3.3 (Meta) data within the Data Object should thus be both syntactically parseable and semantically machine-accessible [10]
4. For Data Objects to be **Re-usable** additional criteria are:
  - 4.1 Data Objects should be compliant with principles 1-3
  - 4.2 (Meta) data should be sufficiently well-described and rich that it can be automatically (or with minimal human effort) linked or integrated, like-with-like, with other data sources [11 and [JCDP 7](#) and [8](#)]
  - 4.3 Published Data Objects should refer to their sources with rich enough metadata and provenance to enable proper citation (ref to [JCDP 1-3](#)).



# Institutional **API**

The figshare API allows you to push data to figshare, or pull data out

This allows you to build applications on top of academic research.





My data

search figshare (titles, tags, authors, etc.)



Browse

Upload



Dan Valen



figshare makes it easy to sync with your ORCID profile



your content that is publically available on figshare and can be synced with ORCID is listed below

<< SYNC ALL >>

Send all >>

<< Get all

or select individually below...

ORCID

CONNECT

you must connect and authorize with ORCID before syncing

### Biography

I currently handle all product-related everything for figshare for institutions in North America. Frequent flier and big fan of mountains, paperback books, cheeseburgers, and analog watches.

### My data (public only)

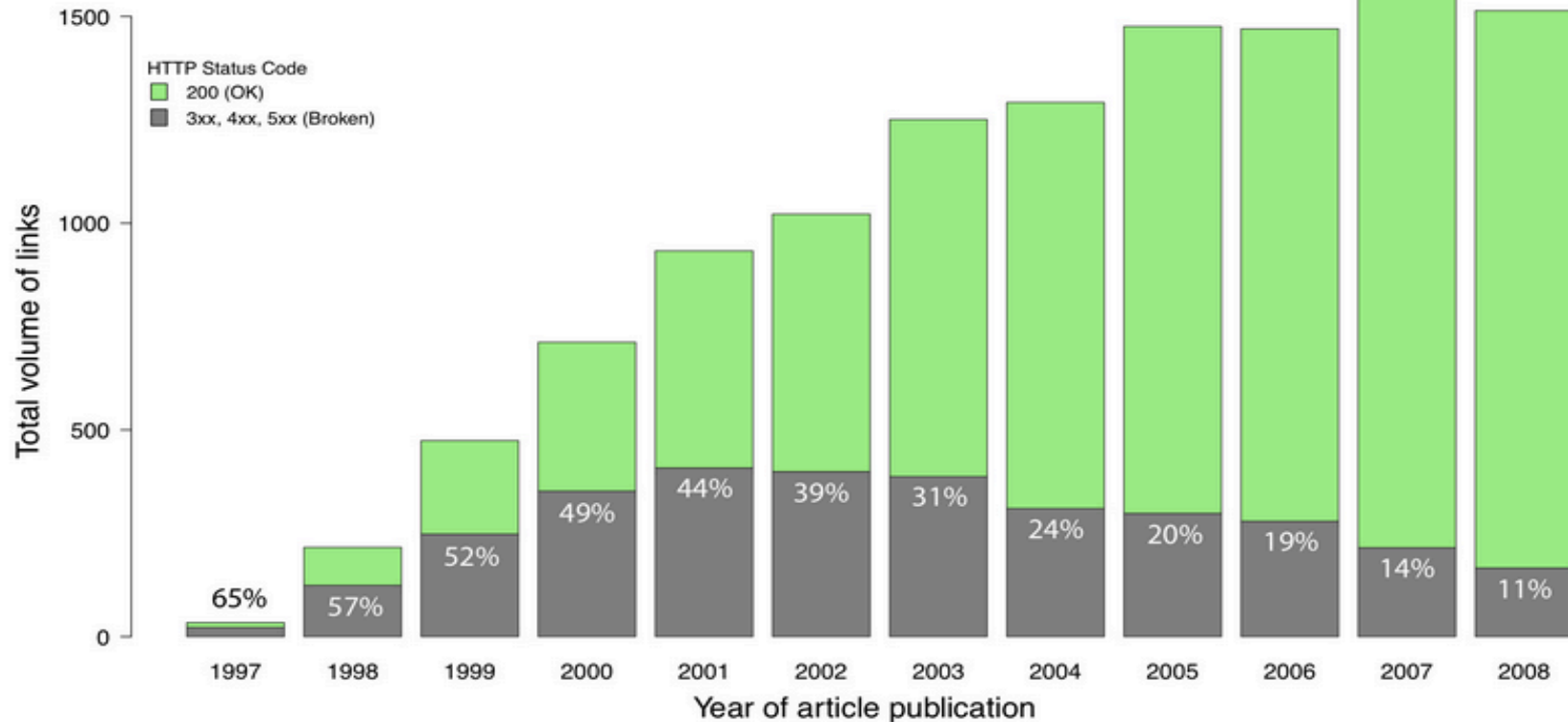
Overview of OSTP Responses  
DOI: 10.6084/M9.FIGSHARE.1367165

You have to **connect** to your ORCID profile first.

Your content will be displayed here afterwards



# Persistent identifiers are essential



**Fig. 2. Volume of potential data links in astronomy publications.** Total volume of external links in all articles published between 1997 and 2008 in the four main astronomy journals, color coded by HTTP status code. Green bars represent accessible links (200), grey bars represent broken links.



# Persistent identifiers are essential

Domain	links (broken)	.html	.txt	.dat	.gz	.tar	.fits	tilde
cxc.harvard.edu	802 (110)	336 (70)	0	0	4 (2)	5 (4)	1	0
heasarc.gsfc.nasa.gov	640 (33)	423 (27)	1	0	0	0	0	0
www.stsci.edu	498 (61)	205 (29)	3	0	0	0	0	15 (10)
asc.harvard.edu	471 (152)	212 (99)	0	0	0	0	0	1 (1)
ssc.spitzer.caltech.edu	427 (194)	125 (76)	3 (3)	0	0	0	0	0
cfa-www.harvard.edu	352 (68)	277 (52)	1	0	0	0	0	54 (17)
archive.stsci.edu	308 (58)	57 (9)	2	1 (0)	0	0	0	0
www.ipac.caltech.edu	285 (14)	209 (12)	0	0	0	0	0	0
www.atnf.csiro.au	211 (21)	12 (6)	0	0	0	0	0	7 (5)
space.mit.edu	193 (10)	58 (5)	1	0	0	0	0	2 (1)
www.astro.psu.edu	186 (4)	103 (1)	1	10	1	1	0	2
www.eso.org	186 (58)	54 (22)	1 (1)	0	0	0	0	4 (1)
irsa.ipac.caltech.edu	163 (5)	38	0	0	1	0	0	0
www.sdss.org	156 (2)	106 (1)	0	0	0	0	0	0
hea-www.harvard.edu	125 (37)	42 (17)	1	0	0	1	0	26 (16)
physics.nist.gov	125 (3)	63 (2)	0	0	0	0	0	0
www.noao.edu	120 (3)	50 (2)	0	0	0	0	0	0
xmm.vilspa.esa.es	118 (35)	23 (19)	0	0	8 (1)	0	0	1 (1)
www.astro.princeton.edu	115 (31)	43 (14)	0	0	0	0	0	53 (12)
ad.usno.navy.mil	110 (27)	98 (22)	3 (3)	0	0	0	0	1 (1)

# APIs are essential. Open Access?



**OPEN**  **ACCESS**

# TAKING CONTROL

ctrl



# Thanks for your time.

@markhahnel

@figshare

figshare.com

api.figshare.com

institutions.figshare.com

mark@figshare.com

