



FOSTER

Overview and definitions of Open Science, Open Access and Open Data

Iryna Kuchma, EIFL Open Access Programme Manager, Twitter: @irynakuchma

Dr. Nancy Pontika, Open Access Aggregation Officer, CORE, Twitter: @nancypontika



What is Open Science?

Open science is the movement to make scientific research, data and dissemination accessible to all levels of an inquiring society, amateur or professional

[https://en.wikipedia.org/wiki/Open_science]

Scope:

- **Transparency** in experimental methodology, observation, and collection of data
- **Public** availability and reusability of **scientific data**
- **Public** accessibility and transparency of **scientific communication**
- Using web-based tools to facilitate scientific **collaboration**

[The OpenScience Project, What exactly is open science <http://www.openscience.org/blog/?p=269>]





Open Access, Open Data, Open Science

EIFL Train-the-trainer programme

Bianca Kramer & Jeroen Bosman, Utrecht University Library



(except logos)



@MsPhelps
@jeroenbosman

Opening up the research workflow

Assessment

- Comment / peer review
- Determine impact of research output
- Determine impact of researchers

Outreach

- Archive/share posters
- Archive/share presentations
- Tell about research outside academia
- Researcher profiles/networks

Publication

- Archive / share publications; data & code
- Publish in OA journal

Preparation

- Define & crowdsource research priorities
- Organize project, team, collaborations
- Get funding / contract

Discovery

- Search literature / data / code / ...
- Get access; Get alerts / recommendations
- Read / view
- Annotate

Analysis

- Collect, mine, extract data / experiment
- Share protocols / notebooks / workflows
- Analyze

Writing

- Write / code
- Visualize
- Cite
- Translate

Open Science is ...

Open to participation

Open to (re)use

Open to the world

Open Science practices

involve public / patients
in drafting
research proposals

openly share
project proposals

share hypothesis before
starting research
(if possible/relevant)

having open discovery
of open access
materials

extensively search for
existing data before
generating your own

use easily attainable
software to allow
anyone to reproduce
your results

Open Science practices

sharing protocols
openly, online

store data in the most
open format possible

cite OA versions of
literature & provide
data and code citations

acknowledge
contributor roles
in a publication

translate research objects
in world languages

publish preprints,
encourage feedback /
open peer review

Open Science practices

publish pre-publication
history (version + reviews)

making conflicts of
interest transparent

using academic social
networks to find and
communicate with
other researchers

refuse to be part of
all male or all white
panels

having all types of
review openly
available

use metrics of
commercial /social
applications to
assess research

Reproducibility & integrity

“Mostly due to current methods capture and data malpractice, approximately 50% of all research data and experiments is considered **not reproducible**, and the vast majority (likely over 80%) of data never makes it to a trusted and sustainable repository.”

Reproducibility and replication

Unreliable research

Trouble at the lab

Scientists like to think of science as self-correcting. To an alarming degree, it is not

Oct 19th 2013 | From the print edition

Like 17k Tweet 1,723



- [Economist article](#):
Low reproducibility
- [Psychology replication study](#):
only 36/97 results
reproducible
- [Cancer biology project](#)

Menu <

- Project Wiki Pages
 - Home
 - Advisory Board
 - Core Project Team
 - Frequently Asked Questions (F...
 - Funding and Supporting Organ...
 - Press
 - Studies
- Component Wiki Pages

View Wiki Version: (Current) Tim Errington: 2017-07-13 14:42:36+00:00 UTC

The Reproducibility Project: Cancer Biology is a collaboration between [Science Exchange](#) and the [Center for Open Science](#), and is independently replicating a subset of experimental results from a number of high-profile papers in the field of cancer biology published between 2010-2012 using the Science Exchange network of expert scientific labs.

Replication Study Results

All published articles related to the project can be found on the [Reproducibility Project: Cancer Biology collection](#) at *eLife*. Each replication is being organized on the [Open Science Framework](#) (OSF). The OSF is a free service and is where all the experimental protocols, materials, data, analysis, and results will be made openly available to the public. Furthermore, quality assurance will be maintained with the [Registered Reports](#) format, in which peer review of proposed experimental designs and protocols will be conducted prior to data collection, in conjunction with *eLife* with the eventual results published in a Replication Study.

At this time, seven replication studies have been published:

Replication Study: Coadministration of a Tumor-Penetrating Peptide Enhances the Efficacy of Cancer Drugs

- View the [Replication Study results](#) published in *eLife*.
- Access the [OSF Project page](#) to view all data, methods, and materials pertaining to this Replication Study.
- View the [Registered Report](#), which contains detailed, peer-reviewed, protocols for this Replication Study.

Replication Study: BET Bromodomain Inhibition as a Therapeutic Strategy to Target c-Myc

- View the [Replication Study results](#) published in *eLife*.
- Access the [OSF Project page](#) which contains all data, methods, and materials pertaining to the replication of this Replication

Science-Based Medicine

Exploring issues and controversies in the relationship between science and medicine

[About SBM](#) [Reference](#) [Links](#) [Recent Comments](#)

Basic Science Cancer

How reproducible is basic lab research in cancer biology?

Last week, a review of the reproducibility of several highly cited cancer biology papers was published. The results were mixed and demonstrate how difficult reproducing published results can be at times—and how scientists need to do better.

David Gorski on January 23, 2017

When deciding what to write about this week, I had thought about expounding on, for instance, my concerns regarding vaccine policy given the new administration, but I think I've done enough of that for the moment at my not-so-super-secret other blog. Besides, there will be plenty of time and many opportunities to return to my concerns in that area

Support science-based medicine

[Donate](#)

Join [The Society for SBM](#)

Buy an e-book: [SBM: The Science of Medicine](#)

Reproducibility

Method Reproducibility

the provision of enough detail about study procedures and data so the same procedures could, in theory or in actuality, be exactly repeated.

Result Reproducibility (aka replicability)

obtaining the same results from the conduct of an independent study whose procedures are as closely matched to the original experiment as possible

What does research reproducibility mean? Steven N. Goodman, Daniele Fanelli, John P.A. Ioannidis *Science Translational Medicine* 8 (341), 341ps12.

[doi: 10.1126/scitranslmed.aaf5027]

<http://stm.sciencemag.org/content/scitransmed/8/341/341ps12.full.pdf>

Reproducibility in the research workflow

assessment

- check (statistical) methods /reporting
- welcome replication studies

publication

- use executable/forkable publications
- use IDs for preregistrations, data, methods, materials, contributors

preparation

- pre-register (can be embargoed)

experimenting / analysis

- share protocols, scripts
- use materials ids (RRIDs)
- use open hardware
- document steps, file management
- share data



Help support open science today.

Donate Now



Preregistration makes your science better.



If you have a project that is entering the planning or data collection phase, we'd like you to try out a preregistration. Through our **\$1 Million Preregistration Challenge**, we're giving away \$1,000 to 1,000 researchers who preregister their projects before they publish them. It's straightforward to complete and will really enhance your research output.

Get Started Now

Pre-registering, e.g. at OSF or AsPredicted



The screenshot shows the OSF Registries website. At the top, the logo consists of a blue flower-like icon followed by the text "OSF REGISTRIES" in blue, with "The open registries network" in white below it. A search bar with the placeholder text "Search registrations..." is present, with a count of "147,382 searchable registrations as of April 20, 2017" below it. A link "See an example" is located below the search bar. The main content area is titled "Browse Recent Registrations" with a "See more" link. Three registration entries are listed, each with a title and author names.

OSF REGISTRIES
The open registries network

Search registrations...

147,382 searchable registrations as of April 20, 2017

[See an example](#)

Browse Recent Registrations [See more](#)


Local conditions explain variation in plant phenology within species
Margaret Kosmala

The Role of Framing Effects, the Dark Triad and Empathy in Predicting Behavior in a One-shot Prisoner's Dilemma
Paul Michael Deutchman, Jess Sullivan

Promoting School Belongingness and Academic Performance: A Multisite Effectiveness Trial of a Scalable Student Mindset Intervention
Geoffrey Borman, Jon Baron

Make it easy to verify your hypothesis and analysis plans. Prevent p-hacking

Pre-registering, e.g. at OSF or AsPredicted

 Open Science Framework Browse ▾ Support 🔍 [Sign Up](#) [Sign In](#)

Promoting School Belongingness and A... [Files](#) [Wiki](#) [Analytics](#) [Forks](#)

Study Information

Title

Provide the working title of your study. It is helpful if this is the same title that you submit for publication of your final manuscript, but it is not a requirement.

Promoting School Belongingness and Academic Performance: A Multisite Effectiveness Trial of a Scalable Student Mindset Intervention

Authors

The author who submits the preregistration is the recipient of the award money and must also be an author of the published manuscript. Additional authors may be added or removed at any time.

Geoffrey Borman, Jon Baron

Research Questions

Please list each research question included in this study.

Though undergoing any transition from the familiar to the unknown may cause the experience of belonging uncertainty, a psychological state whereby people perceive the surrounding environment as potentially threatening, the middle-school transition, more so than other school displacements, is fraught with risks for students. In the context of such changes, a variety of indicators of academic performance, including grade point average, tend to decline during middle school for all students. To mitigate belonging uncertainty and improve students' academic outcomes, we will administer a social-belonging intervention consisting of two 15-minute in-class reading and writing exercises that ask beginning middle-school students to consider and respond to a specific school experience in which they might feel belonging uncertainty. The control exercise includes the same amount of reading and writing but asks students to write about neutral middle-school experiences that are not related to belonging uncertainty, including dealing with a loud lunchroom and learning about politics. After participation, treatment students may benefit socially and psychologically and may realize improved academic outcomes.

Hypotheses

For each of the research questions listed in the previous section, provide one or multiple specific and testable hypotheses. Please state if the hypotheses are directional or non-directional. If directional,

Study Information

- Title
- Authors
- Research Questions
- Hypotheses

Sampling Plan

- Existing Data
- Explanation
- Data collection procedures
- Sample size
- Sample size rationale
- Stopping rule

Variables

- Manipulated
- Measured
- Indices

Design Plan

- Study type
- Blinding
- Study design

Variables

Manipulated

Measured

Indices

Design Plan

Study type

Blinding

Study design

Randomization

Analysis Plan

Statistical models

Transformations

Follow-up analyses

Inference criteria

Data exclusion

Missing data

Exploratory analysis

Pre-registering, e.g. at OSF or AsPredicted



[About COS](#) ▾ [Our Products](#) ▾ [Our Services](#) ▾ [Our Communities](#) ▾ [Blog](#) [Contact Us](#) [Donate](#) [Q](#)



Registered Reports: Peer review before results are known to align scientific values and practices.

Aspredicted.org



Create a new AsPredicted pre-registration

CREATE

See your existing AsPredicteds (e.g. approve, make public)

Your email address (used in AsPredicted)

SEE OWN



Sharing methods and materials, e.g. at Protocols.io or RRID



Explore protocols.io

Discover free, up-to-date research protocols and useful content in your field of interest



Version, modify, and discuss existing protocols

You can "clone" protocols in order to be able to modify existing protocols from other scientists. You can also ask questions and comment on step-level or on the entire protocols.

MyExperiment – research workflow

The screenshot shows the MyExperiment website home page. At the top is a navigation bar with the MyExperiment logo, a 'Home' button, and links for 'Users', 'Groups', 'Workflows', 'Files', and 'Packs'. There is also a search bar with a 'Go' button and a 'Log in' link. Below the navigation bar is a 'Home' breadcrumb. The main content area is divided into three columns. The left column has 'My News' with a 'Log in to view your News' link, and 'Activity' with a list of workflows by Barbara Zdrzil. The middle column has 'Announcements' with one announcement about SSL. The right column has 'Content Stats' with statistics for members, groups, workflows, files, and packs.

my experiment Home Users Groups Workflows Files Packs Search Go Log in

Home

My News

[Log in to view your News](#)

Activity

- Workflows for scaffold trend analyses by **Barbara Zdrzil** (about one day ago)
- KNIME workflows from Zdrzil et al, MedChemComm, 2016: "From linked open data to molecular interaction: studying selectivity trends for ligands of the human serotonin and dopamine transporter" by **Barbara Zdrzil** (about one day ago)
- Scaffolds_trends_workflow_1 by **Barbara Zdrzil** (about one day ago)
- Scaffolds trends workflow 2 by **Barbara Zdrzil** (about one day ago)

Announcements

SSL on myExperiment
5 months ago by Finn Bacall

Content Stats

- 10532 members**
- 394 groups**
- 3860 workflows**
- 1225 files**
- 472 packs**



sharing notebooks e.g. at ONSNetwork or OSF

Computing – Oly BGI GBS Reproducibility; fail?

OK, so things have improved since [the last attempt at getting this BGI script to run](#) and demultiplex the raw data.

I played around with the index.lst file format (based on the error I received last time, it seemed like a good possibility that the file formatting was incorrect) and actually got the script to run to completion! Granted, it took over 16hrs (!!), but it completed!

See the Jupyter notebook link below.

Results:

Well, although the script finished and kicked out all the demultiplexed FASTQ files, the contents of the FASTQ files don't match (the read counts differ between these results and the BGI files) the original set of demultiplexed files. I'm not entirely sure if this is to be expected or not, since the script allows for a single nucleotide mismatch when demultiplexing. Is it possible that the mismatch could be interpreted slightly differently each time this is run? I'm not certain.

Get feedback from peers, help form your thoughts, feel less alone while doing the analyses. Spot mistakes early on.

Open Notebook Science Network

Open Notebook Science Network

HOME

NETWORK ACTIVITY

WHAT IS OPEN NOTEBOOK SCIENCE?

#SCIFUND UNIVERSITY

ABOUT ONS NETWORK

Welcome! to a network of open science notebooks. Questions? tweet us at [@ONScience](https://twitter.com/ONScience).



Source: <http://onsnetwork.org/>

Sharing data, e.g. at Dryad, Figshare or Zenodo

A composite image showing two website screenshots. The top screenshot is the Dryad website, featuring a green tree logo, navigation links like "About", "For researchers", and "Submit data now" button. The bottom screenshot is the Zenodo website, featuring a blue header with the "zenodo" logo, a search bar, and "Upload" and "Communities" links. A central white box with a dark background contains the text "store, share, discover research" and "get more citations for all of the outputs of your academic research over 5000 citations of figshare content to date".

DRYAD About ▾ For researchers ▾ For organizations ▾ Contact us Log in Sign up

DataDryad.org is a curated general-purpose repository that makes the data underlying scientific publications discoverable, freely reusable, and citable. Dryad has integrated data submission for a growing list of journals; submit your data and get the credit you deserve. Welcome!

Submit data now

[How and why?](#)

Search for data

store, share, discover **research**

get more citations for all of the outputs of your academic research over 5000 citations of figshare content to date

zenodo Search [input] [Q] Upload Communities

Sharing research data, e.g. at Dryad, Figshare or Zenodo

F
indable



A
ccessible



I
nteroperable



R
eusable



Sharing code e.g. at GitHub with GNU OR MIT license

harthur-org / brain.js

Watch 63 Star 787 Fork 83

Code Issues 19 Pull requests 0 Projects 0 Wiki Insights

Neural networks in JavaScript <http://brainjs.com>

neural-network brain recurrent-neural-networks easy-to-use api web nodejs browser convolutional-neural-networks node stream

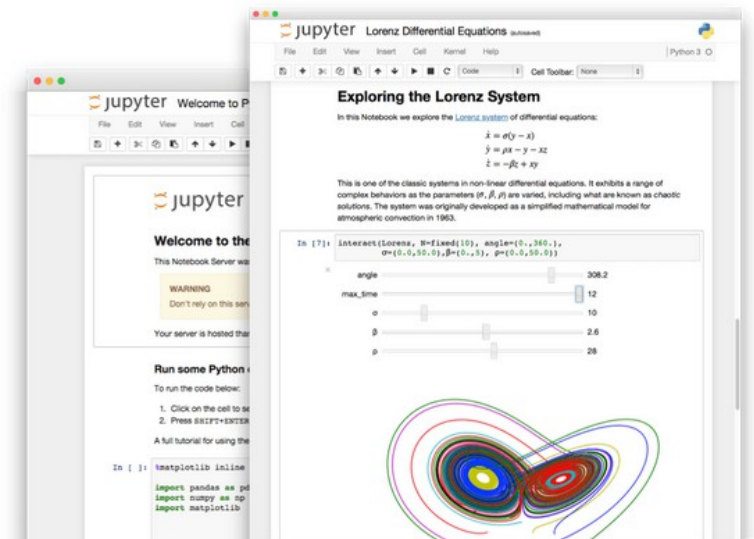
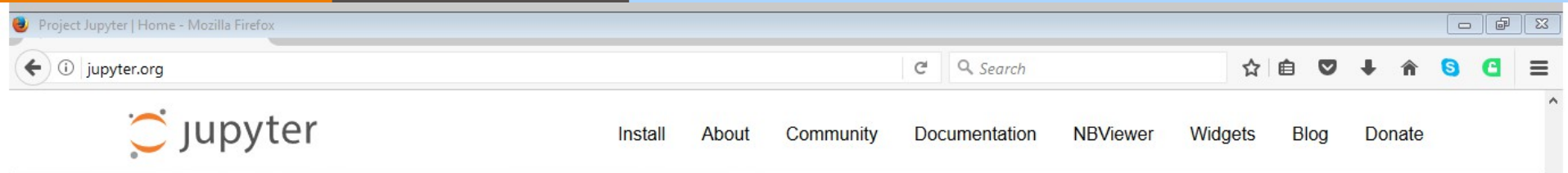
330 commits 13 branches 0 releases 22 contributors MIT

Branch: master New pull request Create new file Upload files Find file Clone or download

robortleeplummerjr committed on GitHub Merge pull request #70 from abhisheksoni27/speed/unique Latest commit b3bbbbc 2 days ago

dist	rename "Vocab" to "DataFormatter"	3 months ago
examples/cli	Fixes #40 and preps for release of 1.0.0	5 months ago
src	Speed 🚀 using Set when removing duplicate elements from array	5 days ago
test	array check and tests	5 days ago

Get people
to check,
contribute
to and use
and build
on your
code



The Jupyter Notebook

The Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and explanatory text. Uses include: data cleaning and transformation, numerical simulation, statistical modeling, machine learning and much more.



Language of choice



Share notebooks



Interactive widgets



Big data integration

Supporting the spread of open research practices

RECENT POSTS

PRO Initiative media

RECENT COMMENTS

Richard Morey on Making your materials public

Richard Morey on Making your materials public

Matt Superdock on Making your materials public

Tomasz Witkowski on Signatories

Graeme Armstrong on Join the Initiative

ARCHIVES

November 2015

CATEGORIES

Uncategorized

META

Log in

Entries [RSS](#)

Comments [RSS](#)

WordPress.org



We believe that openness and transparency are core values of science. For a long time, technological obstacles existed preventing transparency from being the norm. With the advent of the internet, however, these obstacles have largely disappeared. The promise of open research can finally be realized, but this will require a cultural change in science. The power to create that change lies in the peer-review process.

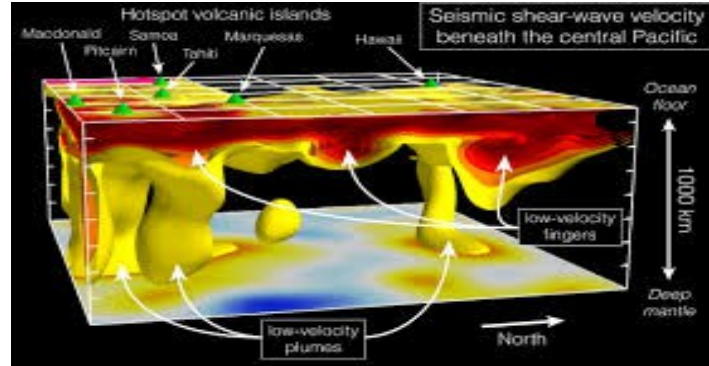
We suggest that beginning January 1, 2017, **reviewers make open practices a pre-condition for more comprehensive review**. This is already in reviewers' power; to drive the change, all that is needed is for reviewers to collectively agree that the time for change has come.



Read the paper about the PRO initiative

- [Read the Initiative](#)
- [Read our guidelines for reviewers](#)
- [PRO-related media](#)
- [See the signatories](#)

Variety of disciplines



Disciplinary variety and Open Science

	ARTS & HUMANITIES	SOCIAL SCIENCE	LIFE/HEALTH	PHYSICAL SCIENCES
Research types	often exploratory research	often confirmatory research	often confirmatory research	often confirmatory research?
Data	often textual data	also qualitative data, sometimes sensitive data	sensitive patient data / big datasets	big datasets
Publ. Types	books, chapters, articles	mostly articles and chapters	mostly articles, (syst.) reviews	preprints, conf papers, articles
Collaboration	typically 1	typically 1-4	typically 3-10	typically 3-many
Languages	native language & some English	English, some native languages	English	English
Funding	small scale funding	small & medium scale funding	large scale funding	large scale funding
Review	double blind	double + single blind	single blind	single blind

Research characteristics and Open Science options/issues

Characteristic	Open Science options/issues
Research types	Preregistration different for exploratory research
Data	Costs of archiving large datasets / conderations of anonimity/sensitiveness / Patentable code/outcomes
Publ. Types	Limited OA Book options / Book publishers small and many / Preprints
Collaboration	Reaching agreements with co-authors
Languages	Not all languages always accepted / Non-native English researcherds at disadvantage
Funding	Large projects have funding but may 'dictate' way of archiving/publishing/communicating
Review	Closed and blind variants of peer review are deeply rooted

Advantages of Open Science for innovation and economic growth

early feedback



adds to **quality** of outcomes

results become available earlier



so problems can be solved **earlier**

having data & code freely available for (re)use



saves time & input costs in research projects of other

no barriers to (re)use data/code, no patents



more people and (small) companies can use research

having more research outcomes available



enhances intellectual creativity by confrontation with alternatives

having more research outcomes and project information available







makes larger projects possible through **collaboration**

Connecting outcomes to social goals

e.g. UN sustainable development goals in RIO journal






Tracking Invasive Alien Species (TriAS): Building a data-driven framework to inform policy

 Sonia Vanderhoeven,  Tim Adriaens, Peter Desmet, Diederik Strubbe, Thierry Bacheljau, Yvan Barbier,  Dimitri Brosens, Julien Cigar, Maxime Coupremagne, Rozemien De Troch, Hilde Eggermont, André Heughebaert, Kris Hostens, Pierre Huybrechts, Anne-Laure Jacquemart, Luc Lens, Arnaud Monty, Jean-Yves Paquet, Céline Prévot, Tim Robertson, Piet Termonia,  Ruben Van De Kerchove, Gert Van Hoey, Bert Van Schaeybroeck, Diemer Vercayie, Thomas Verleye, Sarah Welby, Quentin Groom

Grant Proposal doi: 10.3897/rio.3.e13414

 02-05-2017  Unique: 360 | Total: 602  Reprint: € 9,60



by Sustainable Development Goals (SDG) ▲

- No poverty
- Zero hunger
- Good health & well-being
- Quality education
- Gender equality
- Clean water & sanitation
- Affordable & clean energy
- Decent work & economic growth
- Industry, innovation & infrastructure
- Reduced inequalities




4th European Biodiversity Observation Network (EU BON) Stakeholder Roundtable: Pathways to sustainability for EU BONs network of collaborators and technical infrastructure

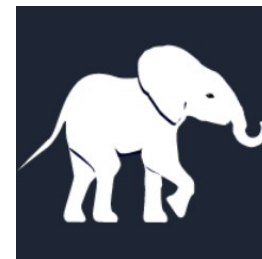
Florian Wetzel, Katherine Despot Belmonte, Heather Bingham, Evelyn Underwood, Anke Hoffmann, Christoph Häuser, Piotr Mikolajczyk, Katrin Vohland

Workshop Report doi: 10.3897/rio.3.e11875

 24-01-2017  Unique: 279 | Total: 530  Reprint: € 4,80

[See collection](#)



Open Science and innovation, contribution to social and economic goals



Institut et hôpital neurologiques de Montréal
Montreal Neurological Institute and Hospital

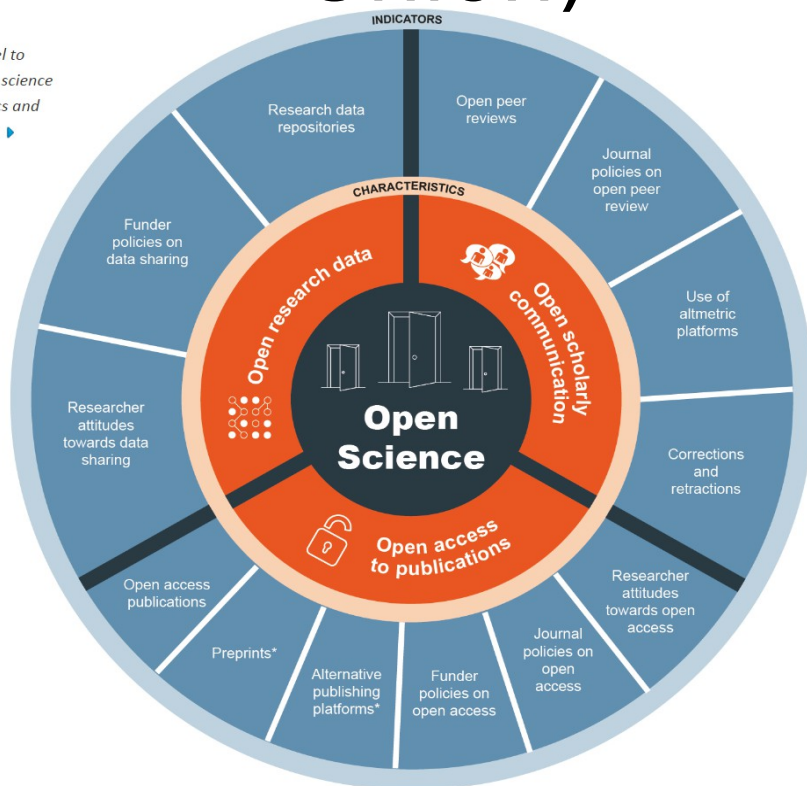
At The Neuro, all findings will be patent-free and freely accessible to other scientists worldwide – making it the first academic institute in the world to fully embrace open science. The Neuro can afford this experiment thanks to a \$20-million (Canadian) **donation** from the family of Larry Tanenbaum, the philanthropist and chairman of Maple Leaf Sports and Entertainment Ltd. As a savvy businessman, he is convinced that openness will accelerate research and discovery. “What we are celebrating here today is the transformation of research, the removal of barriers, the breaking of silos and, most of all, the courage of researchers to put patients and progress ahead of all other considerations,” Mr. Tanenbaum said at Friday’s announcement.

From:
[The Globe and Mail, 20161220](#)

Denmark

Open Science monitor (European Union)

Use the wheel to explore open science characteristics and indicators. ▶▶



<http://ec.europa.eu/research/openscience/index.cfm?pg=home§ion=monitor>

Saving wasted time

Open Access helps to reduce time spent finding/accessing material: “If around 60 minutes were characteristic for researchers (the average time spent trying to access the last research article they had difficulty accessing), then in the current environment the time spent dealing with research article access difficulties might be costing around DKK 540 million (EUR 72 million) per year among specialist researchers in Denmark alone.”

Access to research and technical information in Denmark, Houghton, Swan & Brown (2011)

<http://eprints.ecs.soton.ac.uk/22603>

Open Science contributes to Economic Growth

19% of the processes developed would have been delayed or abandoned without access to research

a 2.2 years delay would cost around EUR 5 million per firm in lost sales



Source: Houghton, J., Swan, A. & Brown, S. Access to research and technical information in Denmark. (2011) <http://eprints.soton.ac.uk/272603>

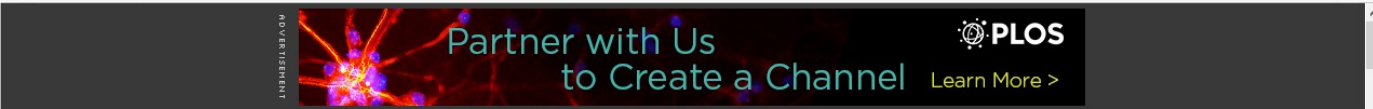
Acceleration of the research process

“As more papers are deposited and more scientists use the repository, the time between an article being deposited and being cited has been shrinking dramatically, year upon year. This is important for research uptake and progress, because it means that in this area of research, where articles are made available at – or frequently before – publication, the research cycle is accelerating.”

Open Access: Why should we have it? Alma Swan www.keyperspectives.co.uk

Benefits of Open Access

Involve citizens and society: Making research openly available is potentially beneficial not only for the individual citizen but also for NGOs and other non-for profit organisations, which often cannot afford subscriptions to a large number of academic journals but for whom academic research is nevertheless very important, e.g. in the field of climate change. A recent study has shown that health NGO staff utilize more research in the course of their work as a result of increasing Open Access to research.



OPEN ACCESS | PEER-REVIEWED

RESEARCH ARTICLE

In an Age of Open Access to Research Policies: Physician and Public Health NGO Staff Research Use and Policy Awareness

Laura L. Moorhead, Cheryl Holzmeyer, Lauren A. Maggio, Ryan M. Steinberg, John Willinsky

Published: July 22, 2015 • <https://doi.org/10.1371/journal.pone.0129708>

20 Save	2 Citation
3,124 View	23 Share

Article	Authors	Metrics	Comments	Related Content
---------	---------	---------	----------	-----------------

- Abstract
- Introduction
- Methods
- Results
- Discussion
- Future Research
- Limitations
- Conclusions
- Supporting Information
- Acknowledgments
- Author Contributions
- References

Abstract

Introduction

Through funding agency and publisher policies, an increasing proportion of the health sciences literature is being made open access. Such an increase in access raises questions about the awareness and potential utilization of this literature by those working in health fields.

Methods

A sample of physicians (N=336) and public health non-governmental organization (NGO) staff (N=92) were provided with relatively complete access to the research literature indexed in PubMed, as well as access to the point-of-care service UpToDate, for up to one year, with their usage monitored through the tracking of web-log data. The physicians also participated in a one-month trial of relatively complete or limited access.

Download PDF | Print | Share

Check for updates

Included in the Following Collection

Open Access

ADVERTISEMENT | PLOS ONE TENTH ANNIVERSARY | In the Media



Open Science can Multiply Serendipity in research ...

Find a Job | M&S Wine | Feedback | Like (334) | Follow @MailOnline | Tuesday, Sep 03 2013 3PM 18°C @ 6PM 16°C @ 5-Day Forecast

MailOnline

Home News U.S. | Sport | TV&Showbiz | Femail | Health | Science | Money | RightMinds | Coffee Break

News Home Arts | Headlines | Pictures | Most read | News Board



AWAITING INSTRUCTIONS

Visit **endabuse.org** NOW.



Ads by Google

Cancer Research Reagents
www.sinobiological.com
1000 Cancer Biomarkers
Antibodies, ELISA Kits, cDNA, Affordat Reagents
Sko på Zalando Zalando.dk

'This is a Cinderella moment for a science nerd like me' The high school student who devised a cure for cancer (that we could be using in as little as 15 years)

By HANNAH ROBERTS
UPDATED: 11:07 GMT, 15 January 2012

53 View comments

Most teenage girls spend their free time gossiping with their friends and figuring out how to outwit their parents. But Angela Zhang spends her time somewhat more productively—the 17-year-old has found a possible cure for cancer.

The extraordinary high school senior from Cupertino has now been rewarded with a scholarship for \$100,000.



Prodigy: 17 year old Angela Zhang has found a possible cure for cancer
At first glance the first generation Chinese schoolgirl, who is learning to drive, seems in many ways an average Californian teenager, CBS

BBC NEWS MAGAZINE

Home UK Africa Asia Europe Latin America Mid-East US & Canada Business Health Sci/Environment

Magazine In Pictures Also in the News Editors' Blog Have Your Say World News TV World Service Radio

LIVING ONLINE

US teen invents advanced cancer Google



20 August 2012 Last updated at 23:34 GMT
Fifteen-year-old high school student Jack Andracka likes to kayak and watch Glee.

And when time permits, he also likes to do advanced research in one of the most respected cancer laboratories in the world.

g+ 117 | Recommend 698 | Share 26 | Tweet 84

FRANCE - Article published the Saturday 05 January 2013 - Latest update : Sunday 06 January 2013

French teenager's research published in Nature

By RFI

A 15-year old school boy from Strasbourg has had his research on astronomy published in the latest issue of the prestigious scientific journal, Nature.



Neil Ibatra, 15-year-old French high school student and the son of an astrophysicist at the Strasbourg Observatory
Reuters/Jean-Marc Loos



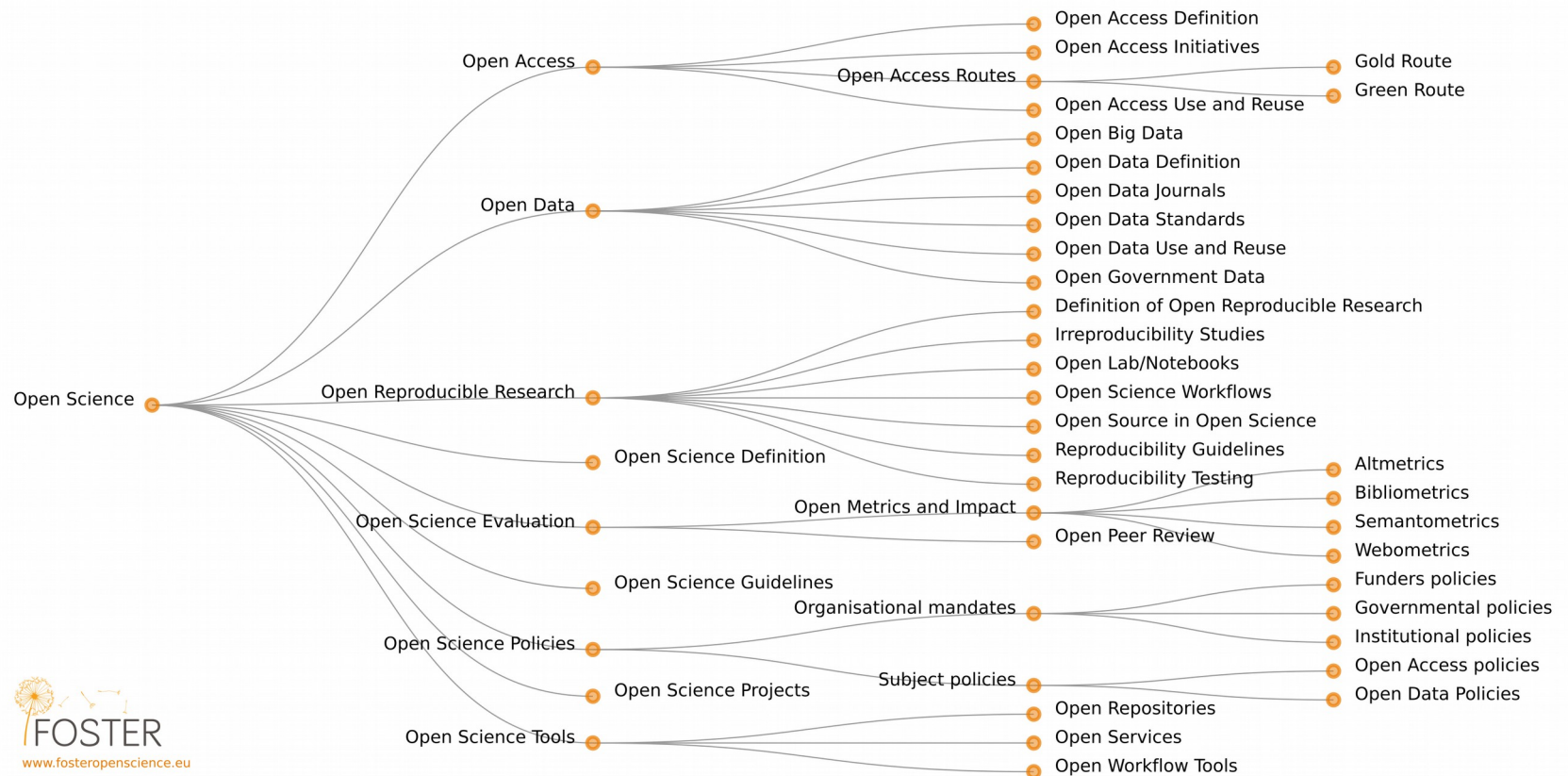


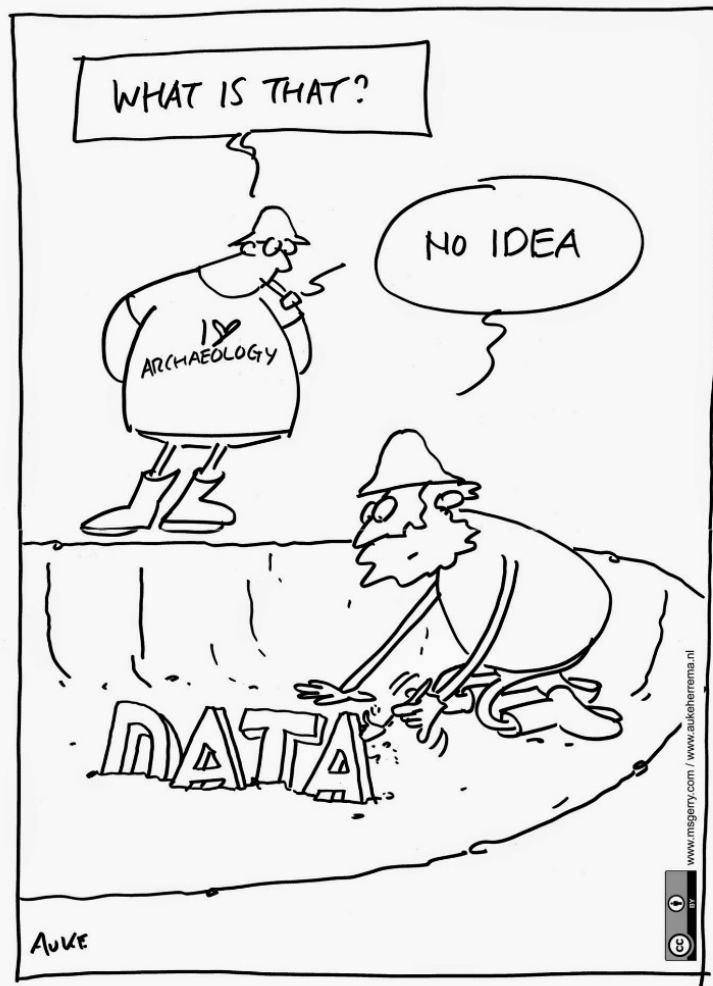
“If we wait 5 years for (Arctic) data to be released, the Arctic is going to be a very different place”

Parsons, Arctic Research Scientist

Source: <http://www.nature.com/nature/journal/v461/n7261/index.html>

Open Science taxonomy





DATA FOR FUTURE GENERATIONS

Open data

“Open data and content can be freely used, modified and shared by anyone for any purpose”

<http://opendefinition.org>

Tim Berners-Lee’s proposal for five star open data - <http://5stardata.info>

- make your stuff available on the Web (whatever format) under an open licence
- make it available as structured data (e.g. Excel instead of a scan of a table)
- use non-proprietary formats (e.g. CSV instead of Excel)
- use URIs to denote things, so that people can point at your stuff
- link your data to other data to provide context

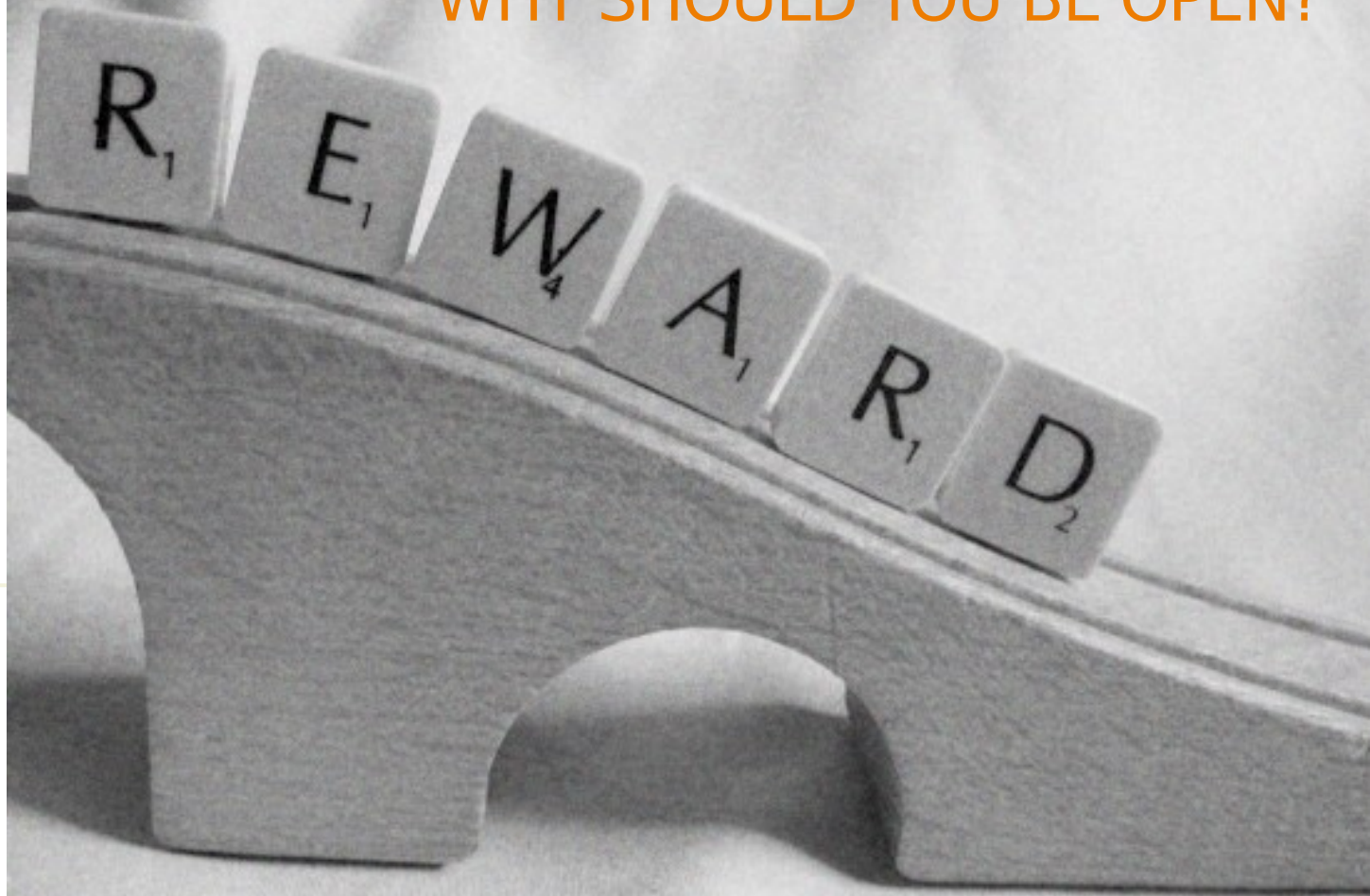
How to make data open?



<https://okfn.org>

1. Choose your dataset(s)
 - What can you open? You may need to revisit this step if you encounter problems later.
2. Apply an open license
 - Determine what IP exists. Apply a suitable licence e.g. CC-BY
3. Make the data available
 - Provide the data in a suitable format. Use repositories.
4. Make it discoverable
 - Post on the web, register in catalogues...

WHY SHOULD YOU BE OPEN?





PUBLICATIONS AND DATA

It's part of good research practice

"It was *never* acceptable to publish papers without making data available."

- Ewan Birney

#OpenData
#OpenScience



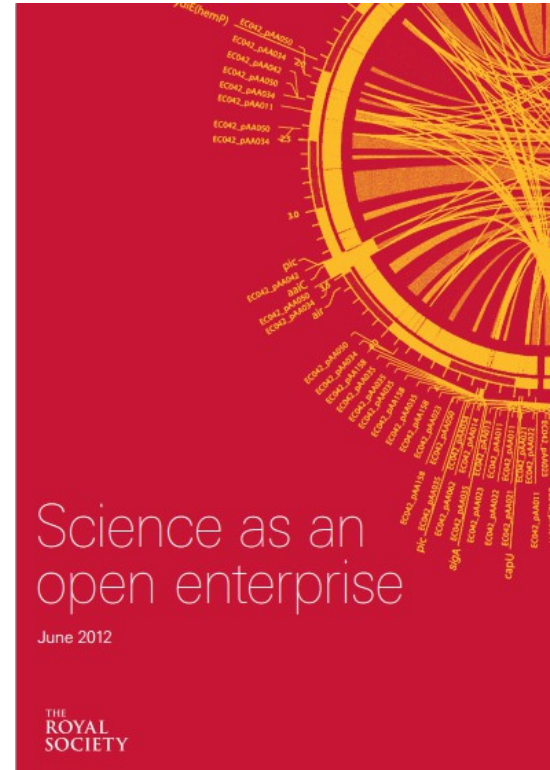
Original image via doi:10.1038/461145a. "Research cannot flourish if data are not preserved and made accessible. Data management should be woven into every course in science." - *Nature* 461, 145

Science as an open enterprise

“Much of the remarkable growth of scientific understanding in recent centuries is due to open practices; open communication and deliberation sit at the heart of scientific practice.”

Royal Society report calls for ‘intelligent openness’ whereby data are accessible, intelligible,

assessable and usable.



<https://royalsocietypublishing.org/doi/10.1098/rsos.120101>

Cut down on academic fraud

The screenshot shows a news article on the Nature website. The article title is "Report finds massive fraud at Dutch universities". The sub-headline is "Investigation claims dozens of social-psychology papers contain faked data." The author is Even Callaway. The article text discusses an investigation into Dutch psychologist Diederik Stapel, who is shown in a photograph. The text states that colleagues called his work good to be true, but it was a preliminary investigative report released on 31 October. The report details years of data manipulation and fabrication by Stapel. A quote from Pim Levelt, chair of the committee that investigated Stapel, is included. The article also mentions that Stapel's studies on social behaviour, such as power and stereotyping, have garnered wide press coverage. A recent Science paper is cited, which identified Stapel as fraudulent. The article concludes by mentioning that Stapel reported that untidy environments encouraged discrimination.

nature International weekly journal of science

[nature news home](#) [news archive](#) [specials](#) [opinion](#) [features](#) [news blog](#) [nature journal](#) [Login](#)

[comments on this story](#)

Published online 1 November 2011 | *Nature* **479**, 15 (2011) | doi:10.1038/479015a
Updated online: 1 November 2011
Updated online: 8 December 2011

News

Report finds massive fraud at Dutch universities

Investigation claims dozens of social-psychology papers contain faked data.

Even Callaway

When colleagues called the work of Dutch psychologist Diederik Stapel too good to be true, they meant it as a compliment. But a preliminary investigative report (go.nature.com/tamp5c) released on 31 October gives literal meaning to the phrase, detailing years of data manipulation and blatant fabrication by the prominent Tilburg University researcher.



Dutch psychologist Diederik Stapel.
Persbureau van Eindhoven

"We have some 30 papers in peer-reviewed journals where we are actually sure that they are fake, and there are more to come," says Pim Levelt, chair of the committee that investigated Stapel's work at the university.

Stapel's eye-catching studies on aspects of social behaviour such as power and stereotyping garnered wide press coverage. For example, in a recent *Science* paper (which the investigation has not identified as fraudulent), Stapel reported that untidy environments encouraged discrimination ([Science 332, 251-253; 2011](#)).

Related stories

- [Seven days: 9-15 September 2011](#)
14 September 2011
- [Chaos promotes stereotyping](#)
07 April 2011

Naturejobs

Tenure-Track Faculty Positions (Assistant / Associate / Full Professor) Yale University, Department of Genetics
Yale University School of Medicine

Assistant Professor
Harvard Medical School

- [More science jobs](#)
- [Post a job for free](#)

Resources

- [PDF Format](#)
- [Send to a Friend](#)
- [Reprints & Permissions](#)
- [RSS Feeds](#)

external links

- [Tilburg University](#)
- [Interim investigation report](#)

Stories by subject

- [Brain and behaviour](#)
- [Lab life](#)

Stories by keywords

- [Diederik Stapel](#)
- [Tilburg University](#)
- [Academic fraud](#)
- [Retractions](#)
- [Social psychology](#)

This article elsewhere

- [Blogs linking to this article](#)

[Add to Digg](#)
[Add to Facebook](#)
[Add to Newsvine](#)
[Add to Del.icio.us](#)
[Add to Twitter](#)

Validation of results

“It was a mistake in a spreadsheet that could have been easily overlooked: a few rows left out of an equation to average the values in a column.

The spreadsheet was used to draw the conclusion of an influential 2010 economics paper: that public debt of more than 90% of GDP slows down growth. This conclusion was later cited by the International Monetary Fund and the UK Treasury to justify programmes of austerity that

The error that could subvert George Osborne's austerity programme

The theories on which the chancellor based his cuts policies have been shown to be based on an embarrassing mistake

Charles Arthur and Phillip Inman
The Guardian, Thursday 18 April 2013 21.10 BST



George Osborne says that Ken Rogoff, the man whose economic error has been uncovered, has strongly influenced his thinking. Photograph: Stefan Wermuth/PA

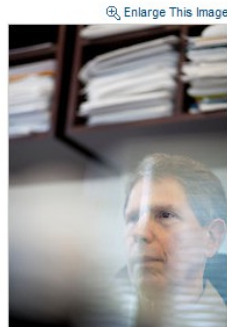
www.guardian.co.uk/politics/2013/apr/18/uncovered-error-george-osborne
terity

More scientific breakthroughs

Sharing of Data Leads to Progress on Alzheimer's

By GINA KOLATA
Published: August 12, 2010

In 2003, a group of scientists and executives from the [National Institutes of Health](#), the [Food and Drug Administration](#), the drug and medical-imaging industries, universities and nonprofit groups joined in a project that experts say had no precedent: a collaborative effort to find the biological markers that show the progression of [Alzheimer's disease](#) in the human brain.



[Enlarge This Image](#)

Now, the effort is bearing fruit with a wealth of recent scientific papers on the early diagnosis of Alzheimer's using methods like PET scans and tests of spinal fluid. More than 100 studies are under way to test drugs that might slow or stop the disease.

And the collaboration is already serving as a model for similar efforts against [Parkinson's disease](#). A \$40 million project to look for biomarkers for Parkinson's, sponsored by the [Michael J. Fox Foundation](#), plans to enroll 600 study subjects in the United States and Europe.

"It was unbelievable. Its not science the way most of us have practiced in our careers. But we all realised that we would never get biomarkers unless all of us parked our egos and intellectual property noses outside the door and agreed that all of our data would be public immediately."

www.nytimes.com/2010/08/13/health/research/13alzheimer.html?pagewanted=all&_r=0 University of Pennsylvania

A citation advantage

A study that analysed the citation counts of 10,555 papers on gene expression studies that created microarray data, showed:

“studies that made data available in a public repository received 9% more citations than similar studies for which the data was not made available”



Data reuse and the open data citation advantage,
Piwowar, H. & Vision, T. <https://peerj.com/articles/175>

Increased use and economic benefit

The case of NASA Landsat satellite imagery of the Earth's surface:

Up to 2008

- Sold through the US Geological Survey for US\$600 per scene
- Sales of 19,000 scenes per year



Since 2009

- Freely available over the internet
- Google Earth now uses the images
- Transmission of 2,100,000 scenes per year.
- Estimated to have created value for the environmental management industry of \$935 million, with direct benefit of more than \$100 million per year to the US economy
- Has stimulated the development of applications from a large number of companies worldwide

<http://earthobservatory.nasa.gov/IOTD/view.php?id=83394&src=v>

BE PART OF THE NEW ERA OF OPEN SCIENCE



reach more
people,
have greater
impact



avoid
duplication
of efforts



preserve data
for future
researchers



simplify final
Horizon 2020
reporting
thanks to an
up-to-date DMP

BE PART OF THE NEW ERA OF OPEN SCIENCE

here's one example of the gains
arising from open research data

Bioinformatics Institute

€1.3 billion per year

Benefits identified by the European
Bioinformatics Institute to users and
their funders just by making scientific
information freely available to the
global life science community...



equivalent to **more
than 20 times**
the direct operational
cost of the Institute

Source: Charles Beagrie Ltd. for EMBL-EBI



With thanks to Bianca Kramer & Jeroen Bosman, Utrecht University Library

**for re-using their slides presented at Open Access, Open Data
Open Science EIFL Train-the-trainer programme**



(except logos)



@MsPhelps
@jeroenbosman

With thanks to

Marjan Grootveld:

marjan.grootveld@dans.knaw.nl

Sarah Jones:

sarah.jones@glasgow.ac.uk

Acknowledgements:

Thanks to DANS and DCC for reuse of

slide www.eudat.eu

www.openaire.eu

Thank you! Questions?

iryna.kuchma@eifl.net

[@irynakuchma](#)

<https://www.fosteropenscience.eu>

[@fosterscience](#)

[facebook.com/fosteropenscience](https://www.facebook.com/fosteropenscience)