



**NanoCommons**

Nano-Knowledge Community

***FAIR starts with findable: data sets and nanomaterials***

***Egon Willighagen, orcid:0000-0001-7542-0286, @egonwillighagen  
NanoCommons Workshop, 2020-04-28, #nanocommons***

***This project has received funding from the European Union Horizon 2020  
Programma (H2020) under grant agreement no. 731032.***

***CC-BY 4.0 International***

# Why share your data? Reuse.



Journal of the American Society for Information Science and Technology

Research Article

## The citation advantage of open-access articles

Michael Norris  Charles Oppenheim  Fytton Rowland 

First published: 09 July 2008 | <https://doi.org/10.1002/asi.20898> | Citations: 88

Maastricht University [find full text](#)

SECTIONS

 PDF  TOOLS  SHARE

### Abstract

Four subjects—ecology, applied mathematics, sociology, and economics—were selected to assess whether there is a citation advantage between journal articles that have an open-access (OA) version on the Internet compared to those articles that are exclusively toll access (TA). Citations were counted using the Web of Science, and the OA status of articles was determined by searching OALster, OpenDOAR, Google, and Google Scholar. Of a sample of 4,633 articles examined, 2,280 (49%) were OA and had a mean citation count of 9.04 whereas the mean for TA articles was 5.76. There appears to be a clear citation advantage for those articles that are OA as opposed to those that are TA. This advantage, however, varies between disciplines, with sociology having the highest citation advantage, but the lowest number of OA articles, from the sample taken, and

PeerJ

## Data reuse and the open data citation advantage

Heather A. Piwowar<sup>1,2</sup> and Todd J. Vision<sup>1,2,3</sup>

<sup>1</sup> National Evolutionary Synthesis Center, Durham, NC, USA

<sup>2</sup> Department of Biology, Duke University, Durham, NC, USA

<sup>3</sup> Department of Biology, University of North Carolina - Chapel Hill, Chapel Hill, NC, USA

### ABSTRACT

**Background.** Attribution to the original contributor upon reuse of published data is important both as a reward for data creators and to document the provenance of research findings. Previous studies have found that papers with publicly available datasets receive a higher number of citations than similar studies without available data. However, few previous analyses have had the statistical power to control for the many variables known to predict citation rate, which has led to uncertain estimates of the “citation benefit”. Furthermore, little is known about patterns in data reuse over time and across datasets.

**Method and Results.** Here, we look at citation rates while controlling for many known citation predictors and investigate the variability of data reuse. In a multivariate regression on 10,555 studies that created gene expression microarray data, we found that studies that made data available in a public repository received 9% (95% confidence interval: 5% to 13%) more citations than similar studies for which the data was not made available. Date of publication, journal impact factor, open access status, number of authors, first and last author publication history, corresponding author country, institution citation history, and study topic were included as covariates. The citation benefit varied with date of dataset deposition: a citation benefit was most clear for papers published in 2004 and 2005, at about 30%. Authors published most papers using their own datasets within two years of their first publication on the dataset, whereas data reuse papers published by third-party investigators continued

Submitted 4 April 2013  
Accepted 13 September 2013  
Published 1 October 2013



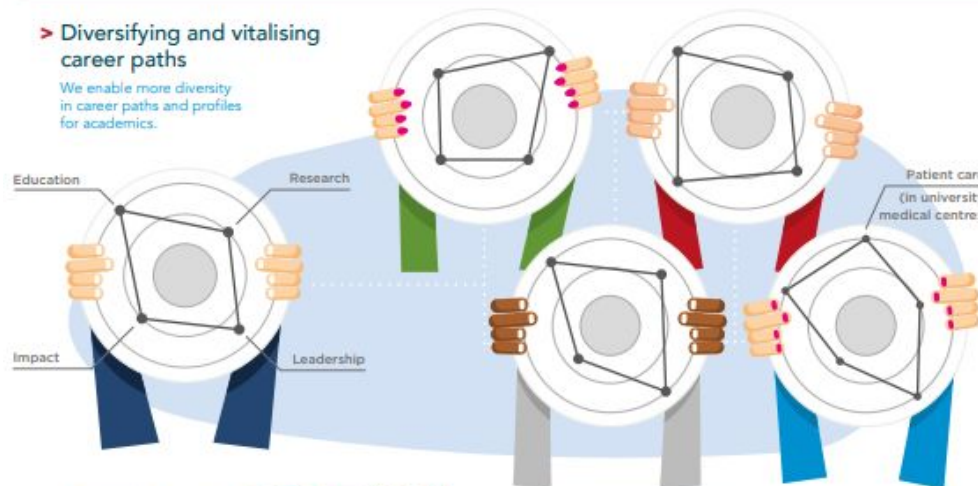
# Data as primary research output

## Room for everyone's talent

*towards a new balance in the recognition and rewards of academics*

### > Diversifying and vitalising career paths

We enable more diversity in career paths and profiles for academics.



### > Achieving balance between individuals and the collective

We assess academics based on both their individual and their team performance.



### > Focusing on quality

In our assessments of academic performance, we increasingly focus on quality, content and creativity.

### > Stimulating open science

We encourage academics to share their research outcomes with society.



# The FAIR data principles

---



F1: identifiers, F2: rich metadata, F3: registered or indexed, F4: specify identifiers

A1: standard protocols, A2: metadata persistent

I1: common language, I2: FAIR vocabularies, I3: references other FAIR

R1: clear license, provenance, community standards

SCIENTIFIC DATA

[Comment](#) | [OPEN](#) | [Published: 15 March 2016](#)

The FAIR Guiding Principles for scientific data management and stewardship

[Mark D. Wilkinson](#), [Michel Dumontier](#) [...] [Barend Mons](#) ✉

*Scientific Data* **3**, Article number: 160018 (2016) | [Download Citation](#) ↓

# The F is for Findable

---

- Traditional: central databases
  - Chemical Abstracts
  - Ensembl/UniProt, NCBI
  - PubChem
  - ...
- Moving to a decentralized world
  - Search.data.enanomapper.net
  - Google Dataset Search
  - DataCite
  - ...
- Archives versus databases
- Data versus dataset
- Data versus information

*Image: CC-BY 3.0, WikiCommons, File:Duke Humfrey's Library Interior 6, Bodleian Library, Oxford, UK - Diliff.jpg*



# Supplementary Information

## Meta-Analysis of Nanoparticle Cytotoxicity via Data-Mining the Literature

Hagar I. Labouta\*, Nasimeh Asgarian, Kristina Rinker and David T. Cramb\*

✓ **Cite this:** *ACS Nano* 2019, 13, 2, 1583-1594

Publication Date: January 28, 2019 ✓

<https://doi.org/10.1021/acsnano.8b07562>

Copyright © 2019 American Chemical Society

[RIGHTS & PERMISSIONS](#)

Article Views

935

Altmetric

4

Citations

3

[LEARN ABOUT THESE METRICS](#)

 PDF (3 MB)



Loading Institutional Login Options...

**More Access Options**



Supporting Info

**SUBJECTS:** [Coating materials](#), [Assays](#), [Toxicity Nanoparticles](#)

### Abstract

Developing predictive modeling frameworks of potential cytotoxicity of engineered nanoparticles is critical for environmental and health risk analysis. The complexity and the heterogeneity of available data on potential risks of nanoparticles, in addition to interdependency of relevant influential attributes, makes it challenging to develop a generalization of nanoparticle toxicity behavior. Lack of

# Google Dataset Search



nanowiki



Updated date

Download format

Usage rights

Free

5 data sets found

**Data from: NanoWiki (release 1)**  
figshare.com  
Updated Jan 19, 2016

**NanoWiki 5**  
figshare.com  
search.datacite.org  
Updated Sep 11, 2018

**NanoWiki (release 4)**  
figshare.com  
search.datacite.org  
Updated Nov 6, 2016

**NanoWiki (release 3)**  
figshare.com  
Updated May 14, 2016

## Data from: NanoWiki (release 1)

Related Article

[Explore at figshare.com](#)

### Unique identifier

<https://doi.org/10.6084/m9.figshare.1330208.v1>

**Data set updated** Jan 19, 2016

### Data set provided by

figshare

### Authors

Egon Willighagen

### Licence

[Attribution 4.0 \(CC BY 4.0\)](#)

Licence information was derived automatically

### Available download formats from providers

application/gzip

# Searching on Chemistry



Home Search Summary Data collections Data templates Help

NANOREG - eNanoMapper database

[NANOREG specific license information](#) | [About](#)

Search


Ag


- Projects (7084)
- Study providers (7084)
- Nanomaterial type (7084)
- Nanomaterial (6268)
- Protocols (21065)
- ALAMAR BLUE AND
- ANNEXIN V / PI ...
- APOPTOSIS ATR-FTIR
- BATCH DISPERS...
- CELL COUNT
- CELL CYCLE
- CHN-ANALYSIS
- COLONY FORMI...
- COMET
- CYTOKINE SECR... DLS


Hits list Selection Predefined Queries Export

Ag Clear

< 1 > displaying 1 to 14 of 14

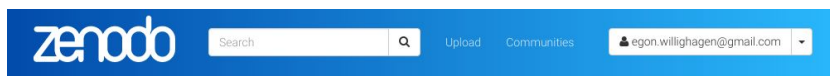
 **Ag @ IIT (Ag 20 nm) silver nanoparticle**  
CORE (1): ...  
**Results:**P-CHEM.Crystalline phase, P-CHEM.Surface chemistry, P-CHEM.Particle size distribution (Granulometry)  
[more](#)  
[Material](#) [Composition](#) [Studies](#) [Add to Selection](#)

 **NFC Fine (Nanofibrillar cellulose 2-15 nm) nanofibrillar cellulose**  
CORE (1): ...  
**Results:**P-CHEM.Analytical Methods, P-CHEM.Crystalline phase, P-CHEM.Surface chemistry, P-CHEM.Particle size distribution (Granulometry), TOX.Immunotoxicity, TOX.Cell Viability, TOX.Genetic toxicity in vitro, TOX.Genetic toxicity in vivo, TOX.Repeated dose toxicity - inhalation, TOX.Repeated dose toxicity - oral  
[more](#)  
[Material](#) [Composition](#) [Studies](#) [Add to Selection](#)

 **JRCNM01001a (NM-101 (TiO2 6 nm)) titanium oxide nanoparticle**  
CORE (1): ...  
**Results:**P-CHEM.Specific surface area, P-CHEM.Crystalline phase, P-CHEM.Surface chemistry, P-CHEM.Particle size distribution (Granulometry), P-CHEM.Water solubility, TOX.Barrier integrity, TOX.Genetic toxicity in vitro, TOX.Cell Viability, TOX.Oxidative



# Easy way: Figshare, Zenodo



zenodo Search Upload Communities [egon.willighagen@gmail.com](#)

## Recent uploads

September 16, 2019 (v12) Dataset Open Access

View

### Binary black-hole surrogate waveform catalog

Scott E. Field; Chad R. Galley; Jan S. Hesthaven; Jason Kaye; Manuel Tiglio; Jonathan Blackman; Béla Szilágyi; Mark A. Scheel; Daniel A. Hemberger; Patricia Schmidt; Rory Smith; Christian D. Ott; Michael Boyle; Lawrence E. Kidder; Harald P. Pfeiffer; Vijay Varma

This repository contains all publicly available numerical relativity surrogate data for waveforms produced by the Spectral Einstein Code. The base method for building surrogate models can be found in Field et al., PRX 4, 031006 (2014). Several numerical relativity surrogate models are currently...

Uploaded on January 28, 2020

11 more version(s) exist for this record

January 24, 2020 (v0.10.0) Software Open Access

View

### mwaskom/seaborn: v0.10.0 (January 2020)

Michael Waskom; Olga Botvinnik; Joel Ostblom; Saulius Lukauskas; Paul Hobson; Maoz Gelbart; David C Gempferline; Tom Augspurger; Yaroslav Halchenko; John B. Cole; Jordi Warmenhoven; Julian de Rutter; Cameron Pye; Stephan Hoyer; Jake Vanderplas; Santi Vialla; Gero Kunter; Eric Quintero; Pete Bachant; Marcel Martin; Kyle Meyer; Corban Swain; Alistair Miles; Thomas Brunner; Drew O'Kane; Tal Yarkoni; Mike Lee Williams; Constantine Evans

This is a major update that is being released simultaneously with version 0.9.1. It has all of the same features (and bugs!) as 0.9.1, but there are important changes to the dependencies. Most notably, all support for Python 2 has now been dropped. Support for Python 3.5 has also been dropped...

Uploaded on January 28, 2020

9 more version(s) exist for this record

Zenodo now supports usage statistics!



Read more about it, in our newest blog post.

Using GitHub?



Check out our GitHub integration. Software Preservation Made Simple!

Zenodo in a nutshell

- **Research. Shared.** – all research outputs from across all fields of research are welcome! Sciences and Humanities, really!
- **Citeable. Discoverable.** – uploads gets a Digital Object Identifier (DOI) to make them easily and uniquely citeable.
- **Communities** – create and curate your own community for a workshop, project, department, journal, into which you can accept



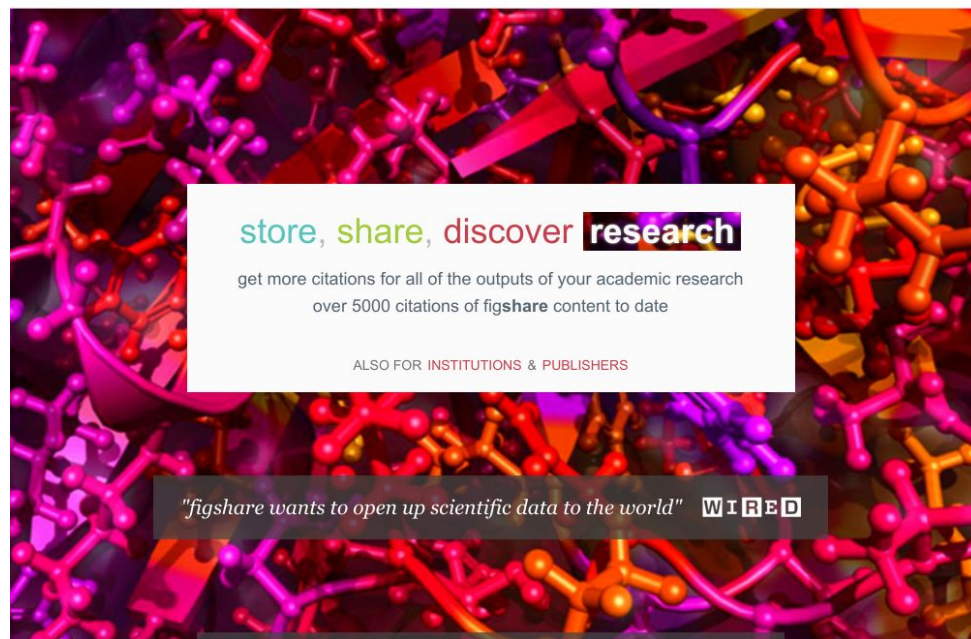
Browse

Search on figshare...



Log in

Sign up



store, share, discover **research**

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

ALSO FOR INSTITUTIONS & PUBLISHERS

"figshare wants to open up scientific data to the world" **WIRED**

# Advertising your data (#nanocommons)



[Home](#)

[Overview](#)

[NSC Projects](#)

[International Cooperation](#)

[Outputs](#)

[Activities](#)

[Contact](#)

[Search](#)

Public deliverables & publications

## Sharing Knowledge

[Home](#) / [Outputs](#) / [Public Deliverables & Publications](#)

### Overview

This page provides available knowledge from the nanosafety community, as produced by projects and published



# Wikidata / Scholia

**SCHOLIA** Author Work Organization Location Event Project Award Topic Tools Help Search...

topic

## JRC representative nanomaterial (Q47461491)

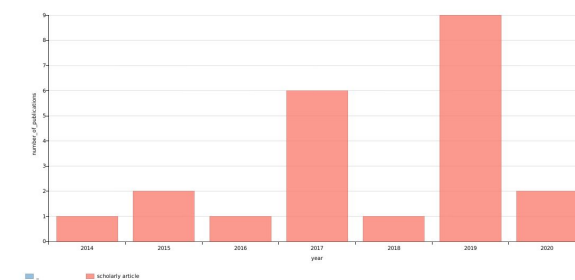
### Recently published works on the topic

Show 10 entries

Search:

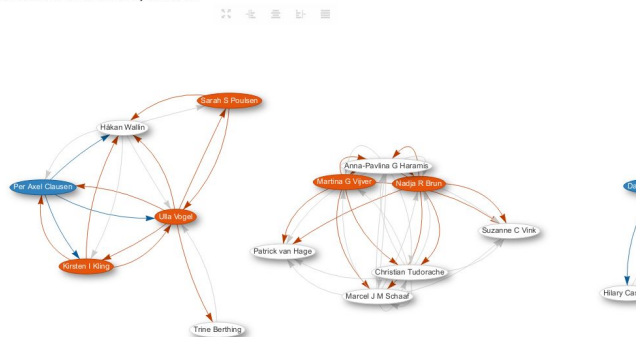
Date	Work	Topics
2020-02-26	<a href="#">Agglomeration of titanium dioxide nanoparticles increases toxicological responses in vitro and in vivo</a>	toxicology // JRCNM10202a // JRCNM10200a
2020-02-08	<a href="#">Assessment of strategies for the formation of stable suspensions of titanium dioxide nanoparticles in aqueous media suitable for the analysis of biological fluids</a>	suspension // JRCNM10200a
2019-11-21	<a href="#">Applicability and Limitations in the Characterization of Poly-Dispersed Engineered Nanomaterials in Cell Media by Dynamic Light Scattering (DLS)</a>	JRCNM01001a // JRCNM02102a // JRCNM02000a
2019-11-19	<a href="#">Assessment of nanomaterial-induced hepatotoxicity using a 3D human primary multi-cellular microtissue exposed repeatedly over 21 days - the suitability of the in vitro system as an in vivo surrogate</a>	hepatotoxicity // JRCNM01005a // JRCNM01101a
2019-10-18	<a href="#">Polystyrene nanoplastics disrupt glucose metabolism and cortisol levels with a possible link to behavioural changes in larval zebrafish</a>	Danio rerio // cortisol // glucose metabolic process // IRCNM01005a

Publications per year



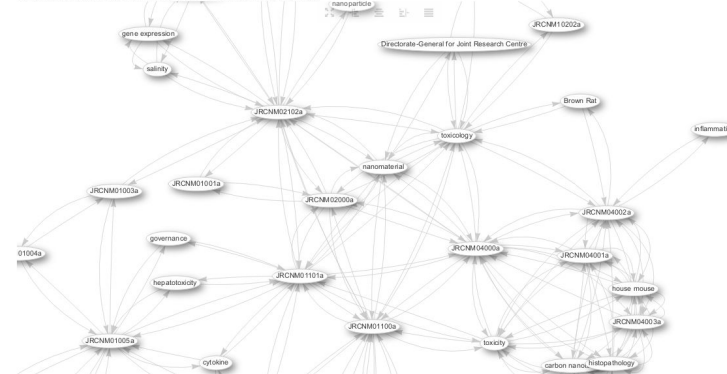
### Co-author graph

The 25 most prolific authors and some of their key co-authors.



### Co-occurring topics graph

Only a maximum of the 400 most often occurring links are shown.



## DOI Citation Formatter

Paste your DOI:

For example 10.1145/2783446.2783605

Select Formatting Style:

Begin typing (e.g. Chicago or IEEE.) or use the drop down menu.

Select Language and Country:

Begin typing (e.g. en-GB for English, Great Britain) or use the drop down menu.

Format

Willighagen, E. (2018). NanoWiki 5 [Data set]. <https://doi.org/10.6084/M9.FIGSHARE.7075214>

Copy to clipboard





Do you want to integrate this service? Check the [Documentation](#)

## DOI Registration Agencies



# BioSchemas 4 custom databases

## Examples

Property	Expected Type	Description	CD	Controlled Vocabulary	Example
<b>Marginality: Minimum.</b>					
<u><a href="#">description</a></u>	<u><a href="#">Text</a></u>	<p><b>Schema:</b> A description of the item.</p> <p><b>Bioschemas:</b> A short summary describing a dataset.</p>	ONE		
<u><a href="#">identifier</a></u>	<u><a href="#">PropertyValue</a></u> <u><a href="#">Text</a></u> <u><a href="#">URL</a></u>	<p><b>Schema:</b> The identifier property represents any kind of identifier for any kind of Thing, such as ISBNs, GTIN codes, UUIDs etc. Schema.org provides dedicated properties for representing many of these, either as textual strings or as URL (URI) links. See <u><a href="#">background notes</a></u> for more details.</p>	MANY		
<u><a href="#">keywords</a></u>	<u><a href="#">Text</a></u>	<p><b>Schema:</b> Keywords or tags used to describe this content. Multiple entries in a keywords list are typically delimited by commas.</p> <p><b>Bioschemas:</b> These keywords provide a summary of the dataset.</p>	MANY		
<u><a href="#">name</a></u>	<u><a href="#">Text</a></u>	<p><b>Schema:</b> The name of the item.</p>	ONE		

# Conclusion

---

- Data is an essential research output
- Sharing your data increases the impact of your work
- Making data available is easy
  - If using Figshare or Zenodo (or similar)
- Metadata can be indexed
- Data can be indexed