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.

Why share your data? Reuse.

NanoCommons
Nano-Knowledge Community

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 OAIster, 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

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

Data reuse and the open data citation advantage

Heather A. Piwowar[1,2] and Todd J. Vision[1,2,3]

[1] National Evolutionary Synthesis Center, Durham, NC, USA
[2] Department of Biology, Duke University, Durham, NC, USA
[3] Department of Biology, University of North Carolina - Chapel Hill, Chapel Hill, NC, USA

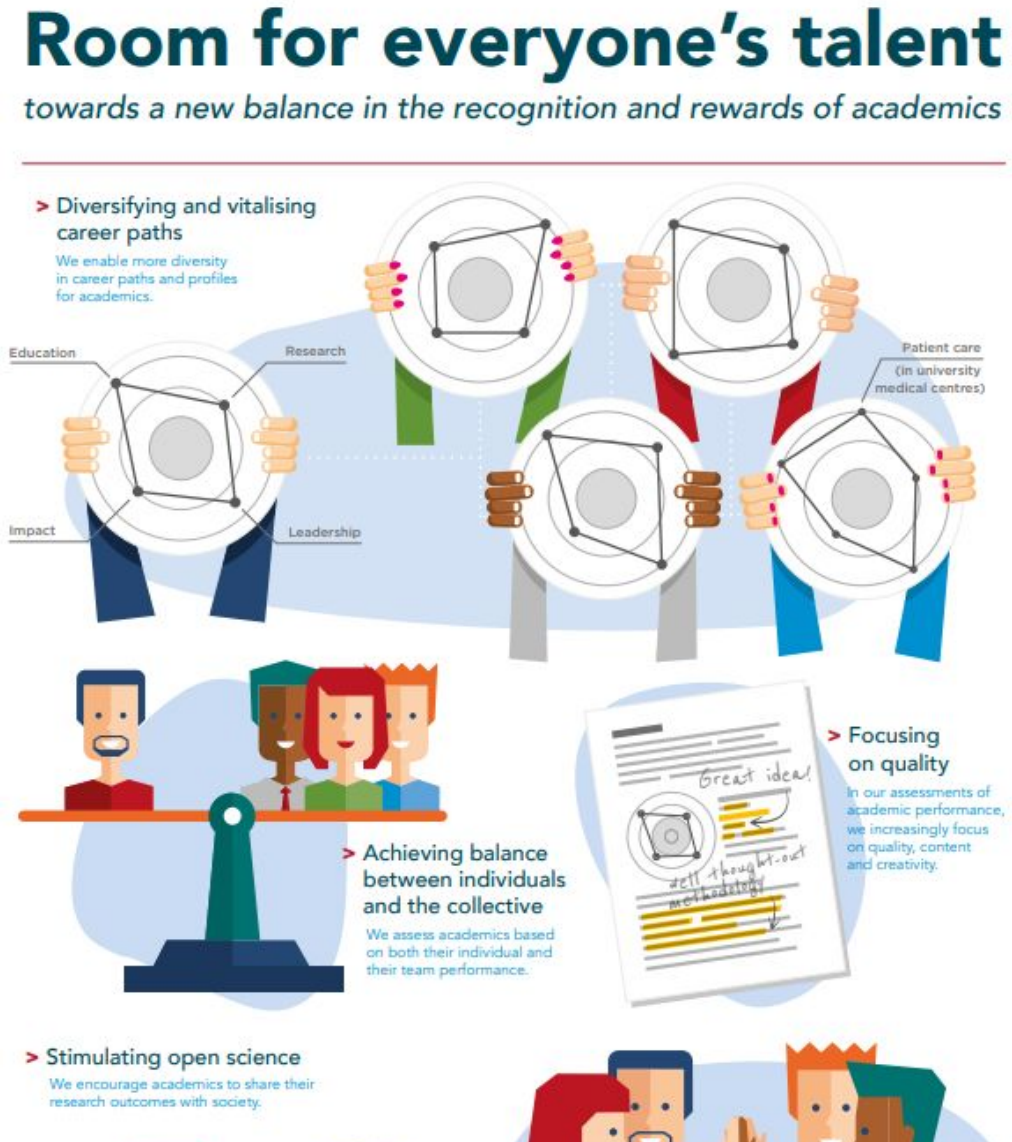
ABSTRACT

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

Data as primary research output

NanoCommons
Nano-Knowledge Community



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

NanoCommons
Nano-Knowledge Community

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

Article Views	Altmetric	Citations
935	4	3

LEARN ABOUT THESE METRICS

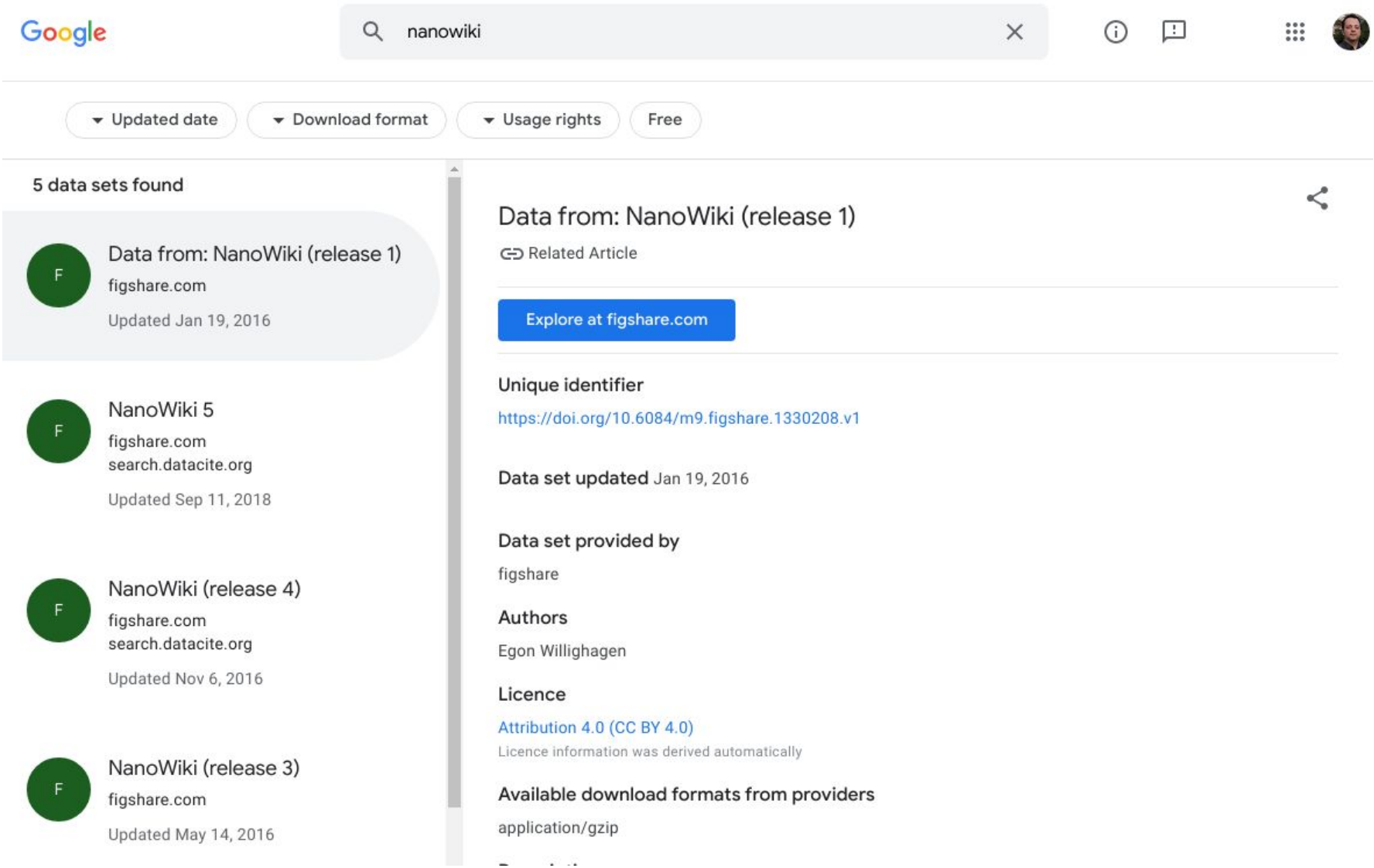
PDF (3 MB) | Loading Institutional Login Options... | More Access Options | SI Supporting Info

SUBJECTS: Coating materials, Assays, Toxicity Nanoparticles

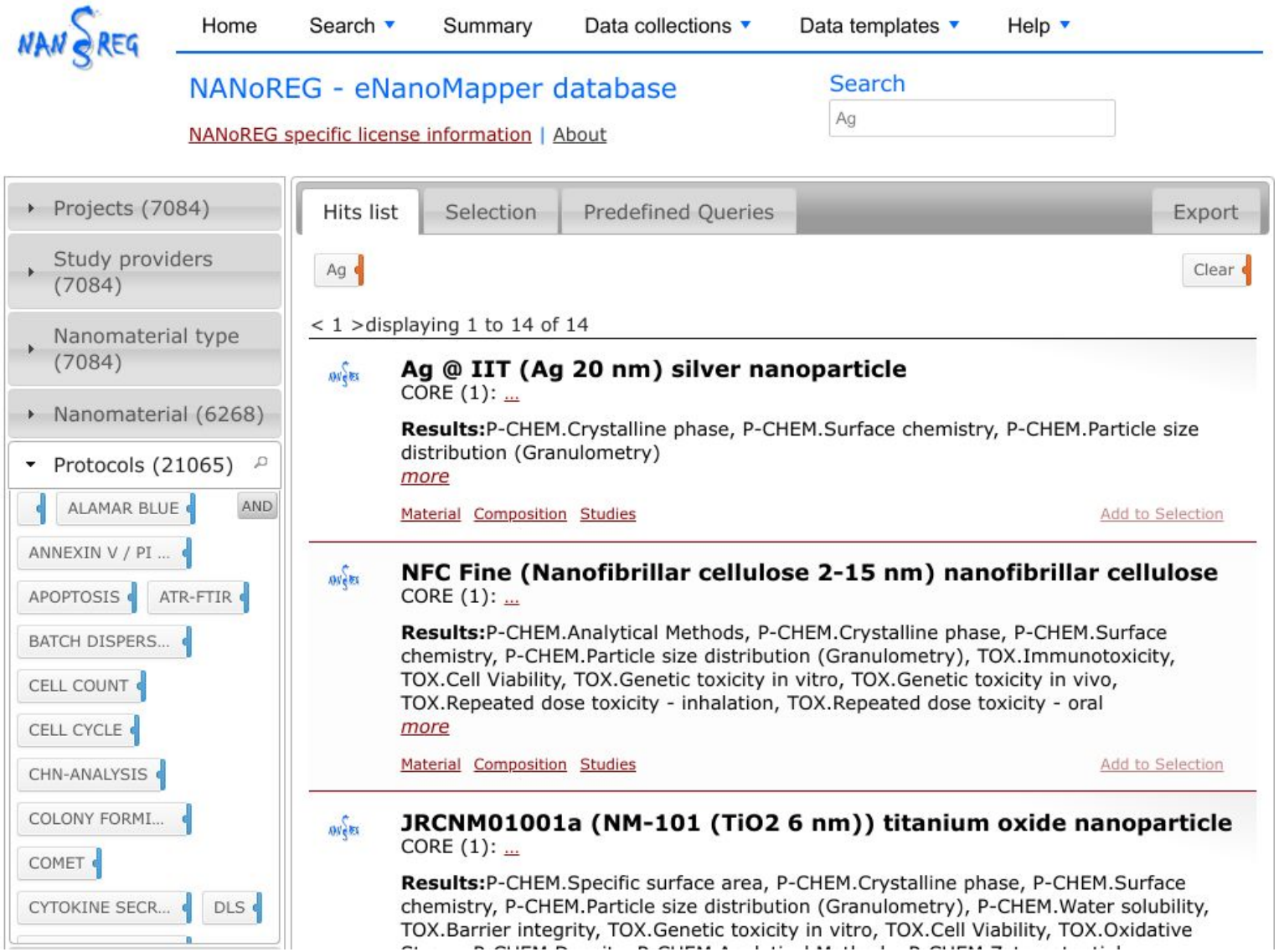
Abstract

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

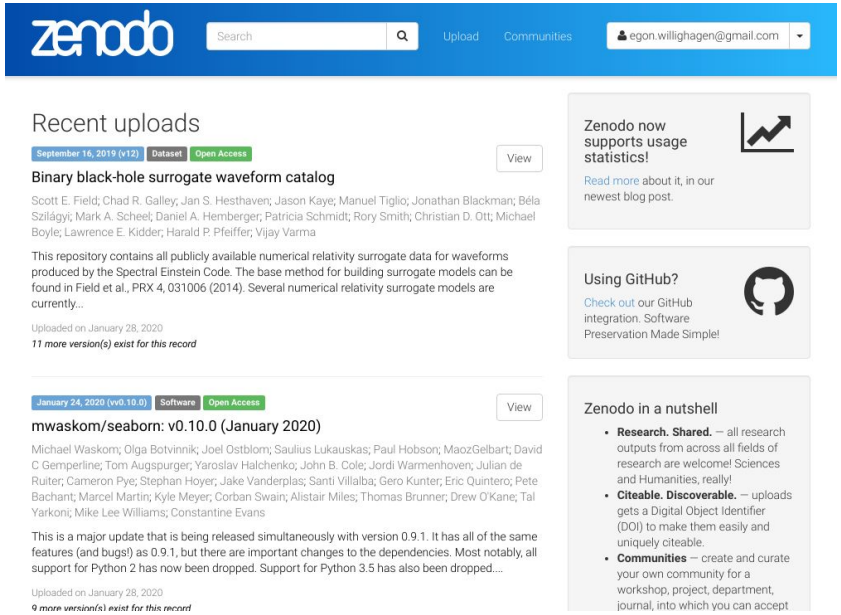


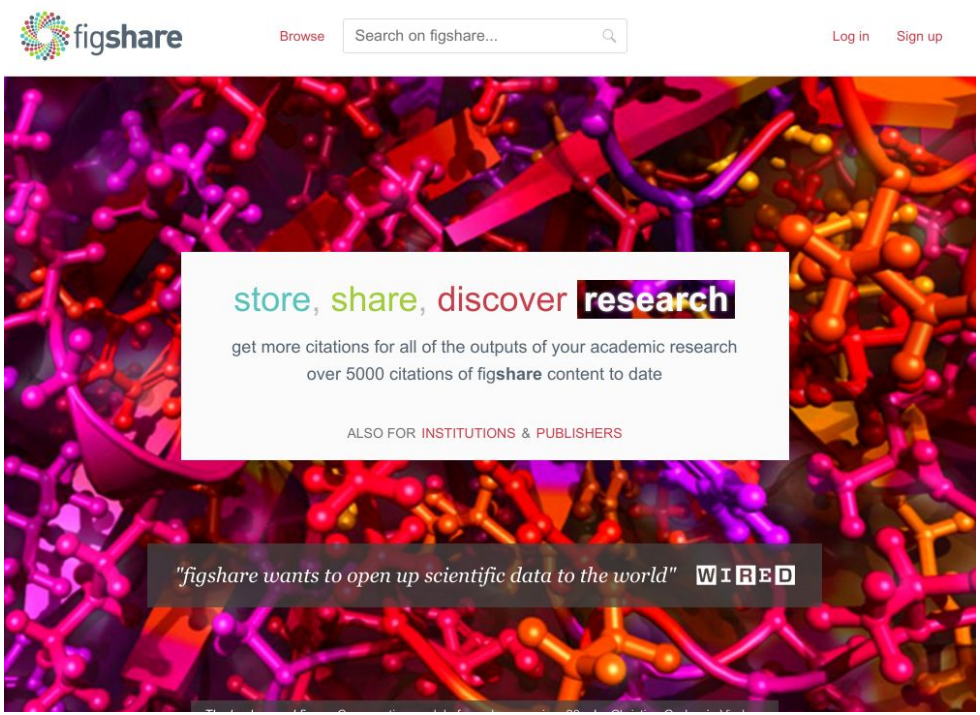
Searching on Chemistry



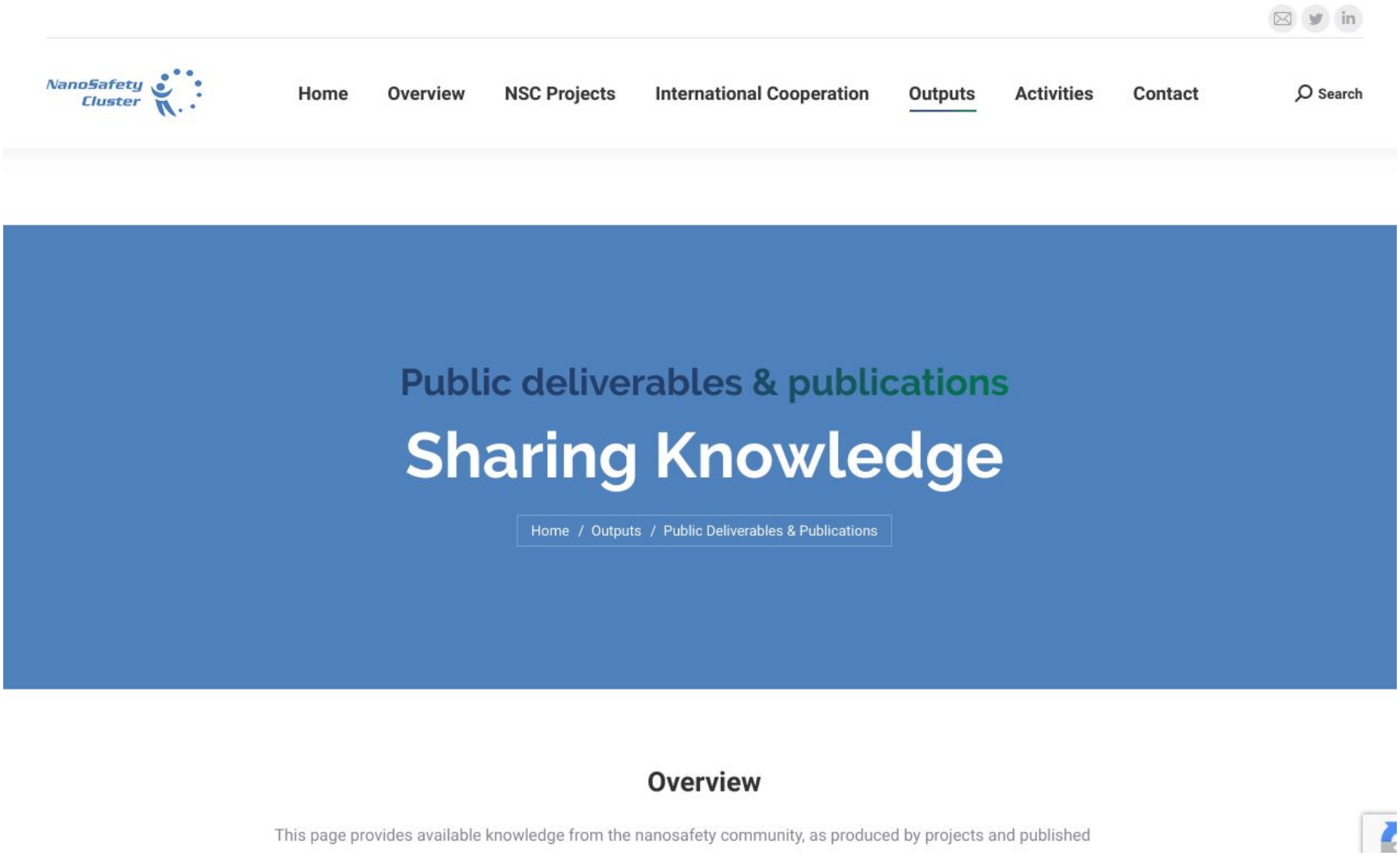
Easy way: Figshare, Zenodo

NanoCommons
Nano-Knowledge Community





Advertising your data (#nanocommons)



Wikidata / Scholia

NanoCommons
Nano-Knowledge Community

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

topic

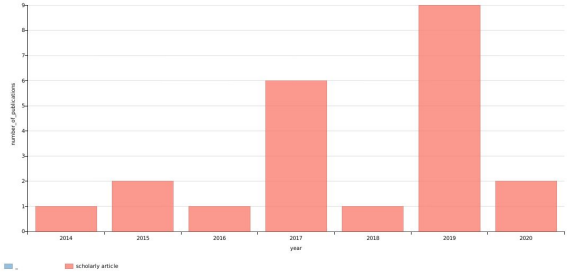
JRC representative nanomaterial (Q47461491)

Recently published works on the topic RSS

Show 10 entries Search:

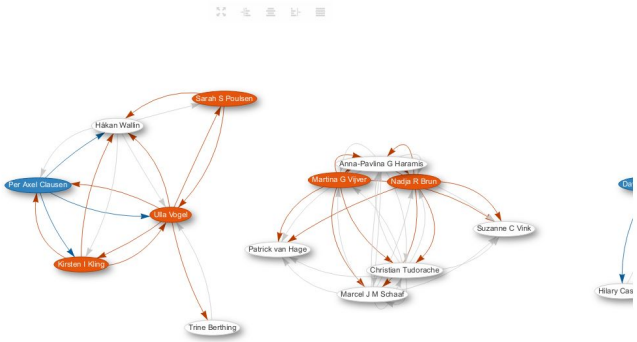
Date	Work	Topics
2020-02-26	Agglomeration of titanium dioxide nanoparticles increases toxicological responses in vitro and in vivo	toxicology // JRCNM10202a // JRCNM10200a
2020-02-08	Assessment of strategies for the formation of stable suspensions of titanium dioxide nanoparticles in aqueous media suitable for the analysis of biological fluids	suspension // JRCNM10200a
2019-11-21	Applicability and Limitations in the Characterization of Poly-Dispersed Engineered Nanomaterials in Cell Media by Dynamic Light Scattering (DLS)	JRCNM01001a // JRCNM02102a // JRCNM02000a
2019-11-19	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	hepatotoxicity // JRCNM01005a // JRCNM01101a
2019-10-18	Polystyrene nanoplastics disrupt glucose metabolism and cortisol levels with a possible link to behavioural changes in larval zebrafish	Danio rerio // cortisol // glucose metabolic process // JRCNM01005a

Publications per year



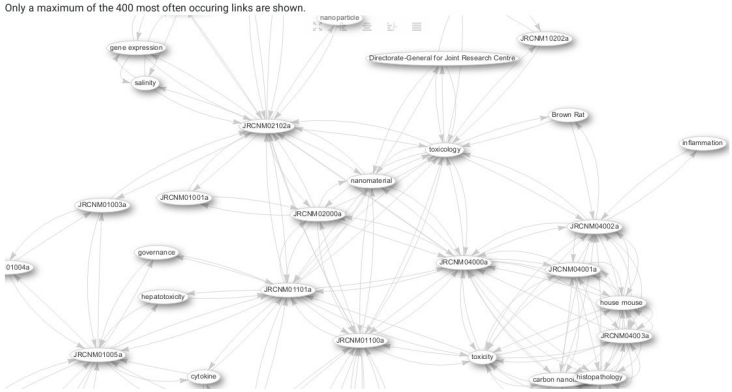
Co-author graph

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



Co-occurring topics graph

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



DataCite

NanoCommons
Nano-Knowledge Community

DOI Citation Formatter

Paste your DOI:

10.6084/m9.figshare.7075214

For example 10.1145/2783446.2783605

Select Formatting Style:

apa

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

Select Language and Country:

en-US

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

DataCite Crossref mEDRA 中文DOI Japan Link Center KiSTi www.kisti.re.kr

BioSchemas 4 custom databases

NanoCommons
Nano-Knowledge Community

Examples

Property	Expected Type	Description	CD	Controlled Vocabulary	Example
Marginality: Minimum.					
description	Text	**Schema:** A description of the item. **Bioschemas:** A short summary describing a dataset.	ONE		
identifier	PropertyValue Text URL	**Schema:** 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 background notes for more details.	MANY		
keywords	Text	**Schema:** Keywords or tags used to describe this content. Multiple entries in a keywords list are typically delimited by commas. **Bioschemas:** These keywords provide a summary of the dataset.	MANY		
name	Text	**Schema:** The name of the item.	ONE		

elixir

Bioschemas

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