Revistes i dades en el nou paradigma de la Ciència Oberta

Universitat
Oberta
de Catalunya

@alexandrelopez

II Jornades #BCSalut17 9 juny 2017

KiMO
Research Group

UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

Universidad
Carlos III de Madrid

ingenio
CSIC-UPV
Instituto de gestión de la innovación
y del conocimiento

Universitat
Oberta
de Catalunya

MareData

Red Española sobre Datos de Investigación en Abierto

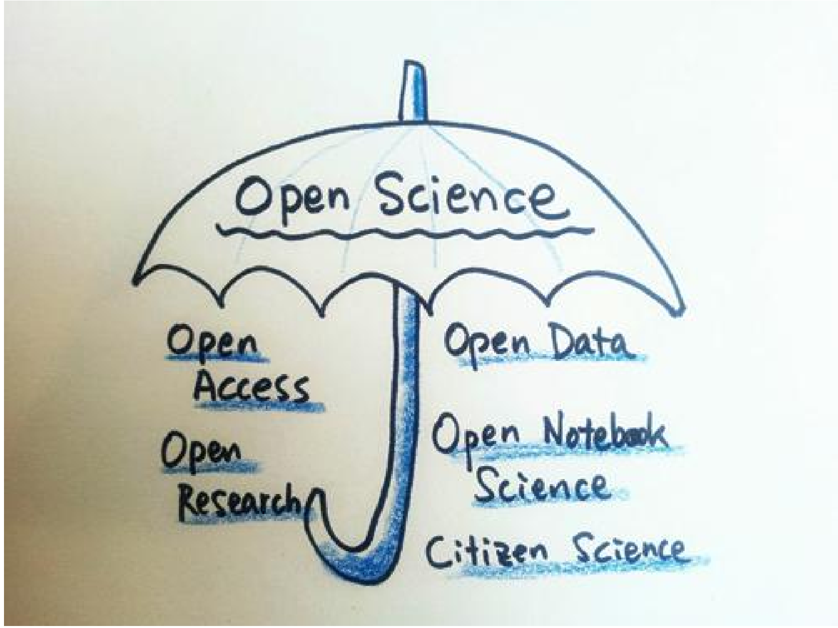
iata
Instituto de Agroquímica
y Tecnología de Alimentos

UNIVERSITAT DE
BARCELONA

Universitat d'Alacant
Universidad de Alicante

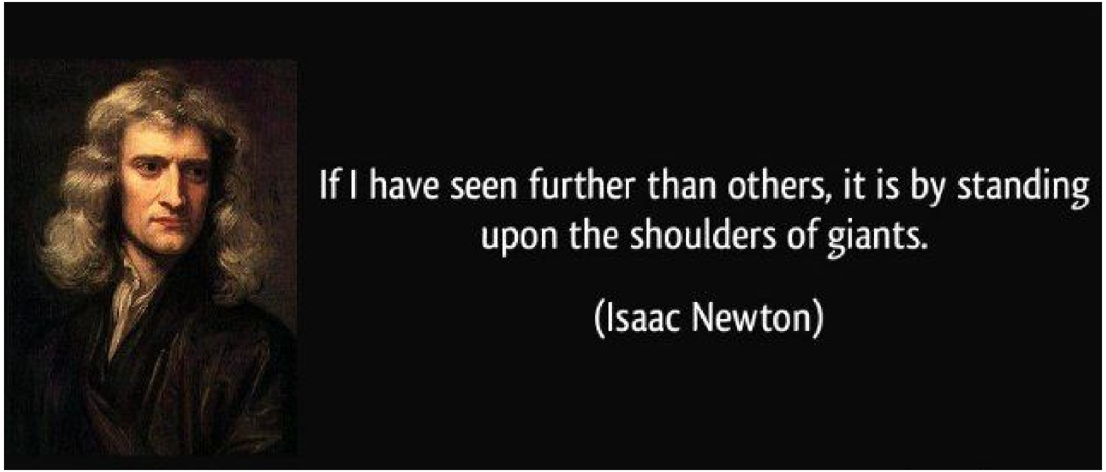
maredata.net **@maredataproject**

On ens trobem els professionals de la informació?



Això de les dades i el Big Data, és per nosaltres?

El per a què, és l'important



If I have seen further than others, it is by standing upon the shoulders of giants.

(Isaac Newton)

Findable Accessible Interoperable Reusable

En Salut, les dades obertes...

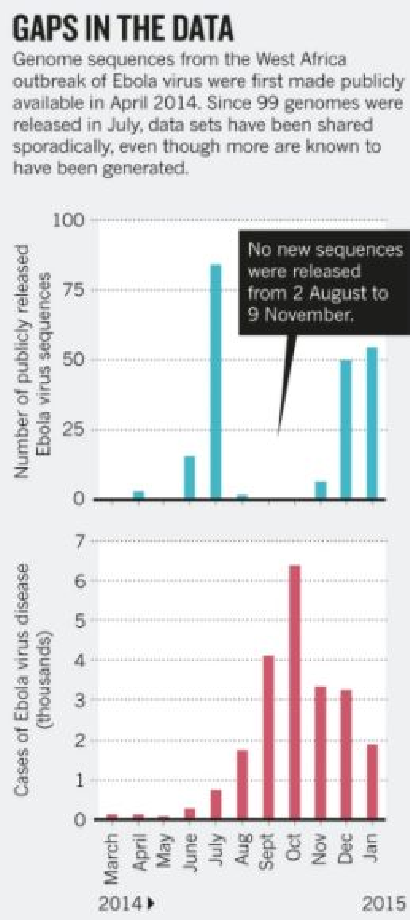
NATURE | COMMENT

Data sharing: Make outbreak research open access

Nathan L. Yozwiak, **Stephen F. Schaffner** & **Pardis C. Sabeti**

25 February 2015

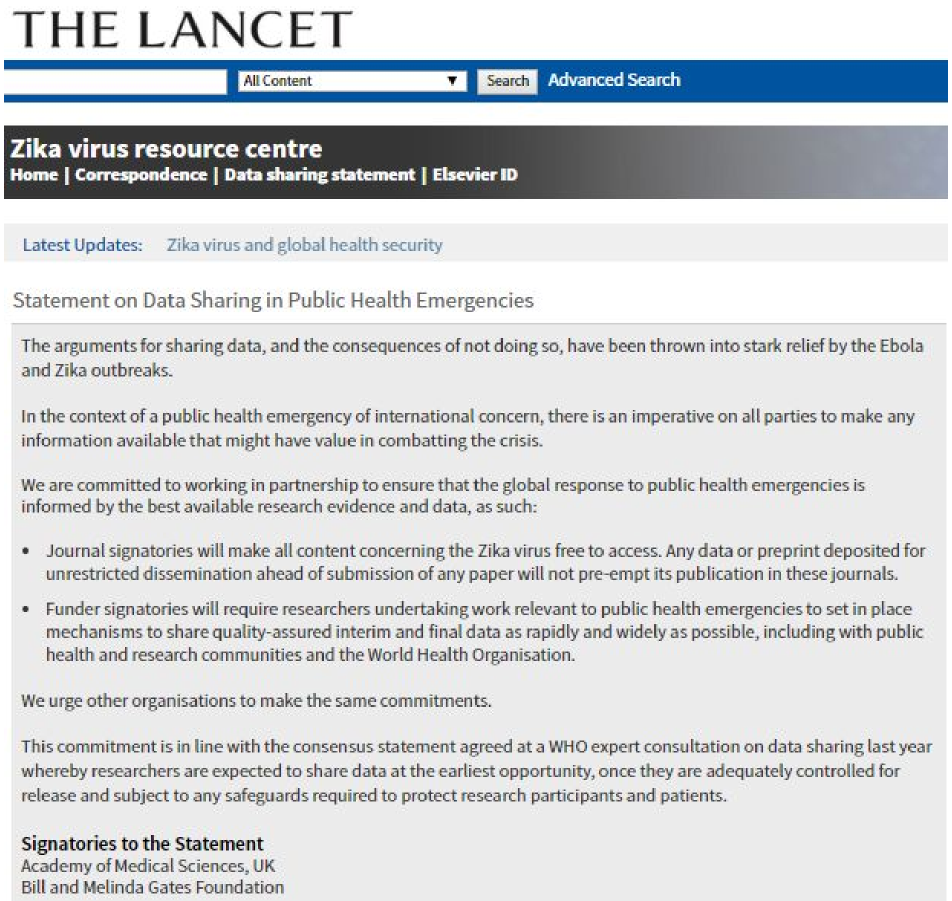
Establish principles for rapid and responsible data sharing in epidemics, urge Nathan L. Yozwiak, Stephen F. Schaffner and Pardis C. Sabeti.



GAPS IN THE DATA

Genome sequences from the West Africa outbreak of Ebola virus were first made publicly available in April 2014. Since 99 genomes were released in July, data sets have been shared sporadically, even though more are known to have been generated.

…marquen la diferència



Zika virus resource centre

Home | Correspondence | Data sharing statement | Elsevier ID

Latest Updates: Zika virus and global health security

Statement on Data Sharing in Public Health Emergencies

The arguments for sharing data, and the consequences of not doing so, have been thrown into stark relief by the Ebola and Zika outbreaks.

In the context of a public health emergency of international concern, there is an imperative on all parties to make any information available that might have value in combatting the crisis.

We are committed to working in partnership to ensure that the global response to public health emergencies is informed by the best available research evidence and data, as such:

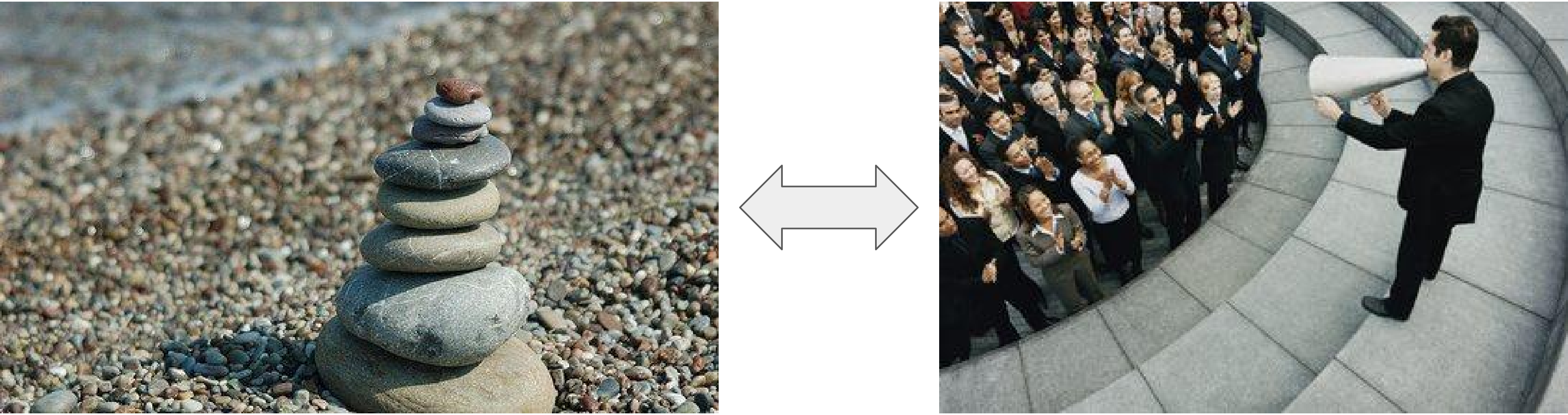
- Journal signatories will make all content concerning the Zika virus free to access. Any data or preprint deposited for unrestricted dissemination ahead of submission of any paper will not pre-empt its publication in these journals.
- Funder signatories will require researchers undertaking work relevant to public health emergencies to set in place mechanisms to share quality-assured interim and final data as rapidly and widely as possible, including with public health and research communities and the World Health Organisation.

We urge other organisations to make the same commitments.

This commitment is in line with the consensus statement agreed at a WHO expert consultation on data sharing last year whereby researchers are expected to share data at the earliest opportunity, once they are adequately controlled for release and subject to any safeguards required to protect research participants and patients.

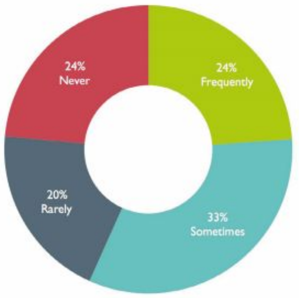
Signatories to the Statement
Academy of Medical Sciences, UK
Bill and Melinda Gates Foundation

Podem comparar amb l’Open Access?



Què diuen els científics?

Approximately **3/4** of respondents have made their research data openly available at some point



Researchers in the social sciences demonstrate the highest level of awareness

Researchers in Asia demonstrate the least familiarity

Researchers admit to gaps in their knowledge of open data and are eager for more information

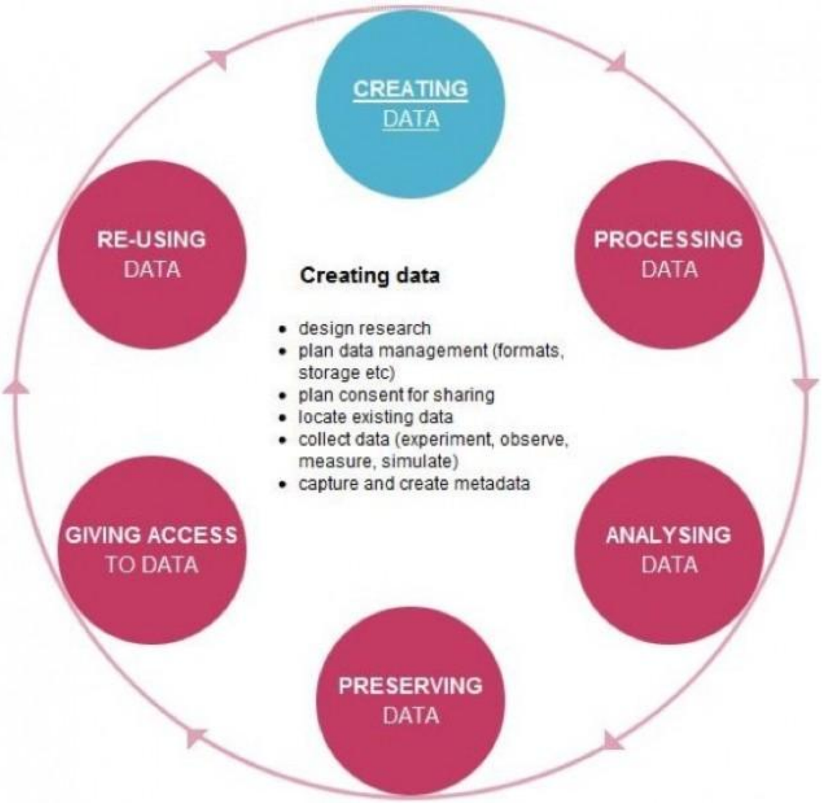
More than **50%** would welcome more guidance on compliance with their funder's policy

Researchers are uncertain who will meet the costs of making data open

The future will be more open

The state of open data: a selection of analyses and articles about open data, curated by Figshare (2016). London: Digital Science. 48 p. Disponible a: https://figshare.com/articles/The_State_of_Open_Data_Report/4036398

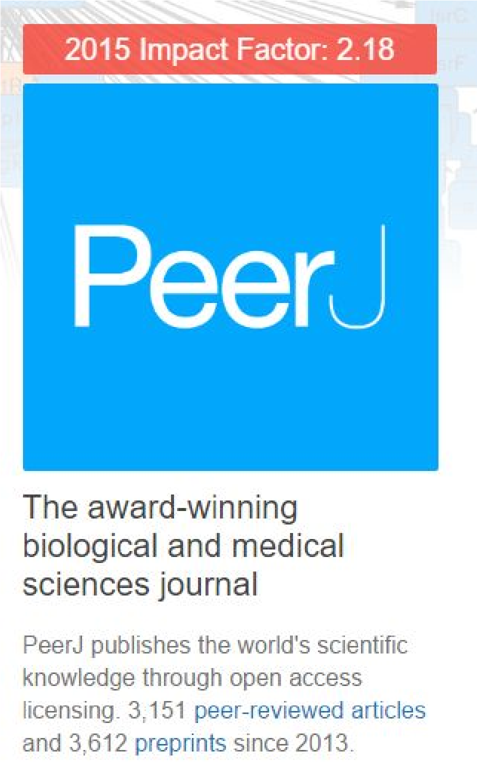
Les dades formen part del cicle de publicació científica



Creating data

- design research
- plan data management (formats, storage etc)
- plan consent for sharing
- locate existing data
- collect data (experiment, observe, measure, simulate)
- capture and create metadata

Noves revistes científiques

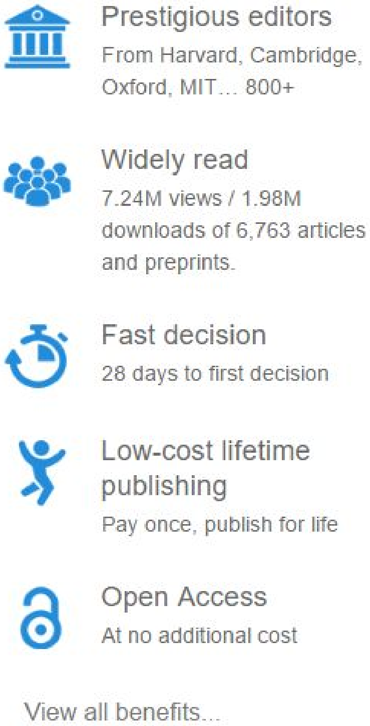


2015 Impact Factor: 2.18

PeerJ

The award-winning biological and medical sciences journal

PeerJ publishes the world's scientific knowledge through open access licensing. 3,151 peer-reviewed articles and 3,612 preprints since 2013.



Prestigious editors
From Harvard, Cambridge, Oxford, MIT… 800+

Widely read
7.24M views / 1.98M downloads of 6,763 articles and preprints.

Fast decision
28 days to first decision

Low-cost lifetime publishing
Pay once, publish for life

Open Access
At no additional cost

View all benefits...

Noves revistes científiques

Diverse bacterial communities exist on canine skin and are impacted by cohabitation and time

Microbiology Veterinary Medicine Dermatology

Sheila Torres[1], Jonathan B. Clayton[2], Jessica L. Danzeisen[2], Tonya Ward[3], Hu Huang[3], Dan Knights[3,4], Timothy J. Johnson[2]

March 9, 2017

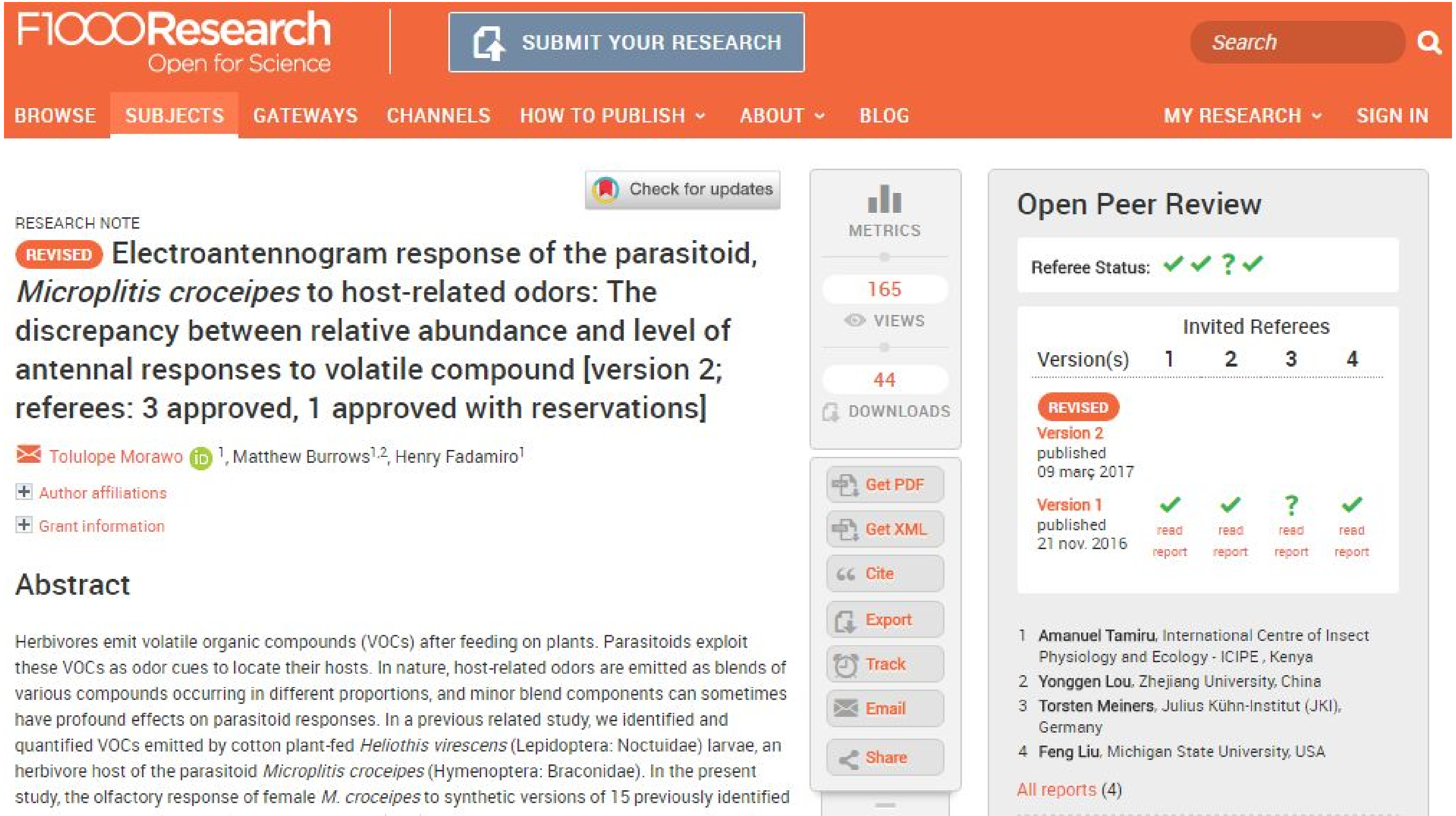
Or Sign up for free and we'll keep you up to date on the latest fee waiver offers and research.

Author and article information

Abstract

It has previously been shown that domestic dogs and their household owners share bacterial populations, and that sharing of bacteria between humans is facilitated through the presence of dogs in the household.

Noves revistes científiques



RESEARCH NOTE

REVISED Electroantennogram response of the parasitoid, *Microplitis croceipes* to host-related odors: The discrepancy between relative abundance and level of antennal responses to volatile compound [version 2; referees: 3 approved, 1 approved with reservations]

Tolulope Morawo [1], Matthew Burrows[1,2], Henry Fadamiro[1]

Author affiliations

Grant information

Abstract

Herbivores emit volatile organic compounds (VOCs) after feeding on plants. Parasitoids exploit these VOCs as odor cues to locate their hosts. In nature, host-related odors are emitted as blends of various compounds occurring in different proportions, and minor blend components can sometimes have profound effects on parasitoid responses. In a previous related study, we identified and quantified VOCs emitted by cotton plant-fed *Heliothis virescens* (Lepidoptera: Noctuidae) larvae, an herbivore host of the parasitoid *Microplitis croceipes* (Hymenoptera: Braconidae). In the present study, the olfactory response of female *M. croceipes* to synthetic versions of 15 previously identified

Open Peer Review

Referee Status: ✓ ✓ ? ✓

Invited Referees

Version(s)	1	2	3	4
REVISED Version 2 published 09 març 2017				
Version 1 published 21 nov. 2016	✓ read report	✓ read report	? read report	✓ read report

1 **Amanuel Tamiru**, International Centre of Insect Physiology and Ecology - ICIPE , Kenya
2 **Yonggen Lou**, Zhejiang University, China
3 **Torsten Meiners**, Julius Kühn-Institut (JKI), Germany
4 **Feng Liu**, Michigan State University, USA

All reports (4)

Noves revistes científiques

Our Publishing Processes

For Articles

Authors are encouraged to publish revised versions of their article. All versions of an article are linked and independently citable. Articles that pass peer review are indexed in external databases such as PubMed, Scopus and Google Scholar.

Estratègies de les revistes respecte els datasets

1. Objecte independent

dins d’un repositori

Pampel, H; Dallmaier-Tiessen, S. (2014) “Open Research Data: From Vision to Practice”. En: Bartling, S.; Friesike, S (eds.). Opening Science (2014). http://www.openingscience.org/get-the-book

Estratègies de les revistes respecte els datasets

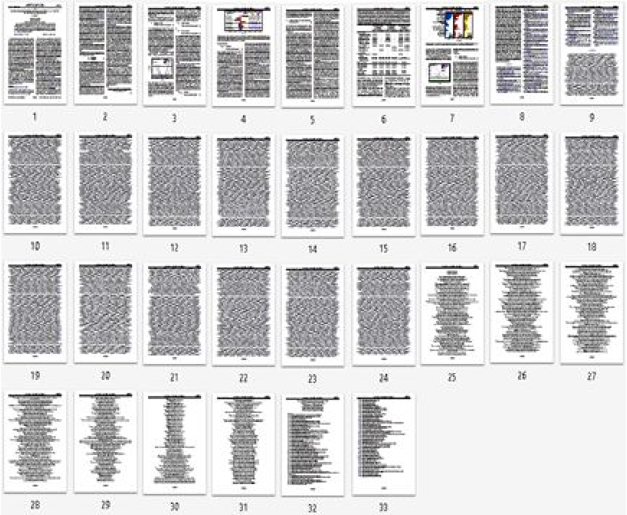
2. *Enriched Publication*

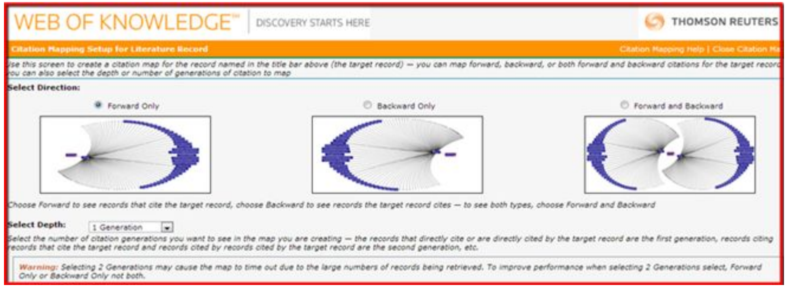
Estratègies de les revistes respecte els datasets

2. *Hi ha altres datasets possibles?*

Ex: 1 article, 5154 autors

Ex: citacions





se this screen to create a citation map for the record named in the title bar above (the target record) — you can map forward, backward, or both forward and backward citations for the target recor
ou can also select the depth or number of generations of citation to map

elect Direction:

● Forward Only ○ Backward Only ○ Forward and Backward

hoose Forward to see records that cite the target record, choose Backward to see records the target record cites — to see both types, choose Forward and Backward

elect Depth: 1 Generation

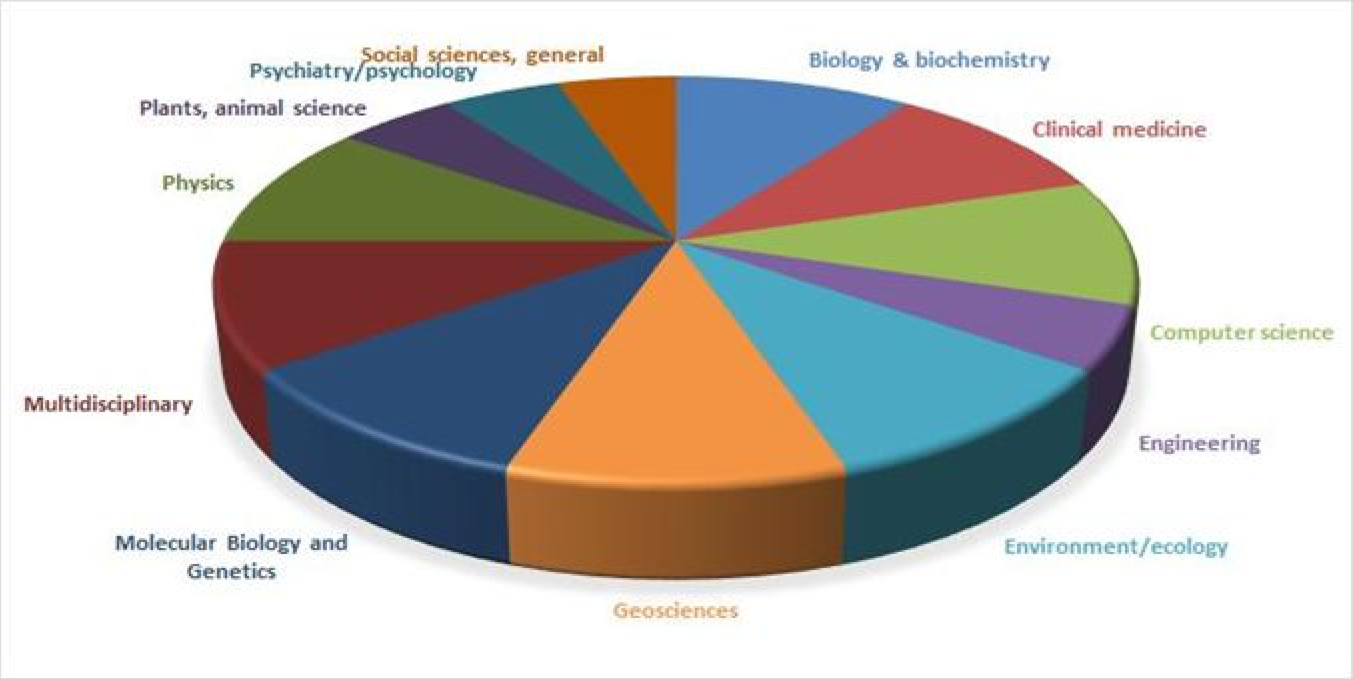
elect the number of citation generations you want to see in the map you are creating — the records that directly cite or are directly cited by the target record are the first generation, records citing ecords that cite the target record and records cited by records cited by the target record are the second generation, etc.

Warning: Selecting 2 Generations may cause the map to time out due to the large numbers of records being retrieved. To improve performance when selecting 2 Generations select, Forward Only or Backward Only not both.

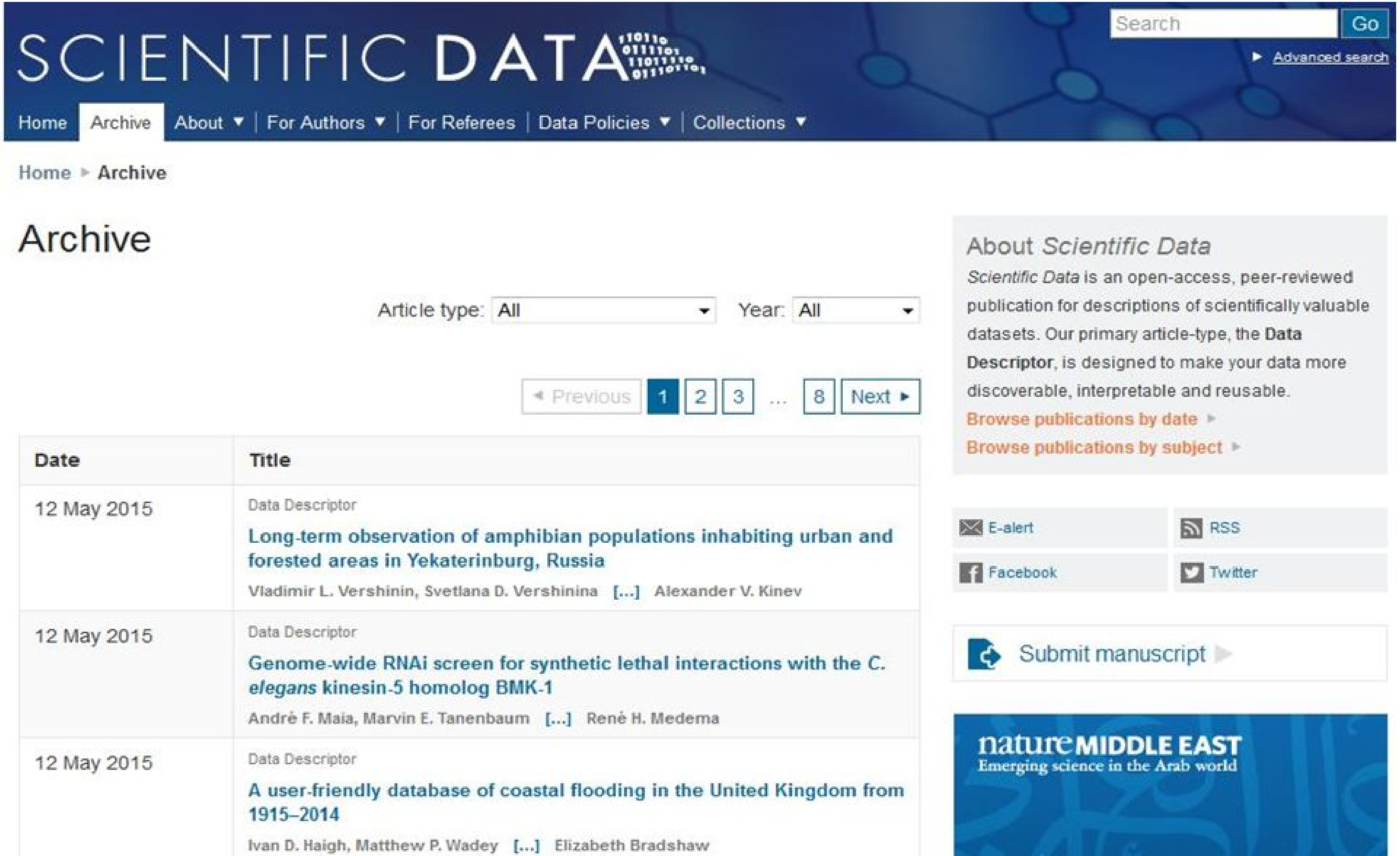
Estratègies de les revistes respecte els datasets

3. Data Paper/Data Journal

Data Journals



Data Papers, Data Journals



Archive

Date	Title
12 May 2015	Data Descriptor Long-term observation of amphibian populations inhabiting urban and forested areas in Yekaterinburg, Russia Vladimir L. Vershinin, Svetlana D. Vershinina [...] Alexander V. Kinev
12 May 2015	Data Descriptor Genome-wide RNAi screen for synthetic lethal interactions with the *C. elegans* kinesin-5 homolog BMK-1 André F. Maia, Marvin E. Tanenbaum [...] René H. Medema
12 May 2015	Data Descriptor A user-friendly database of coastal flooding in the United Kingdom from 1915–2014 Ivan D. Haigh, Matthew P. Wadey [...] Elizabeth Bradshaw

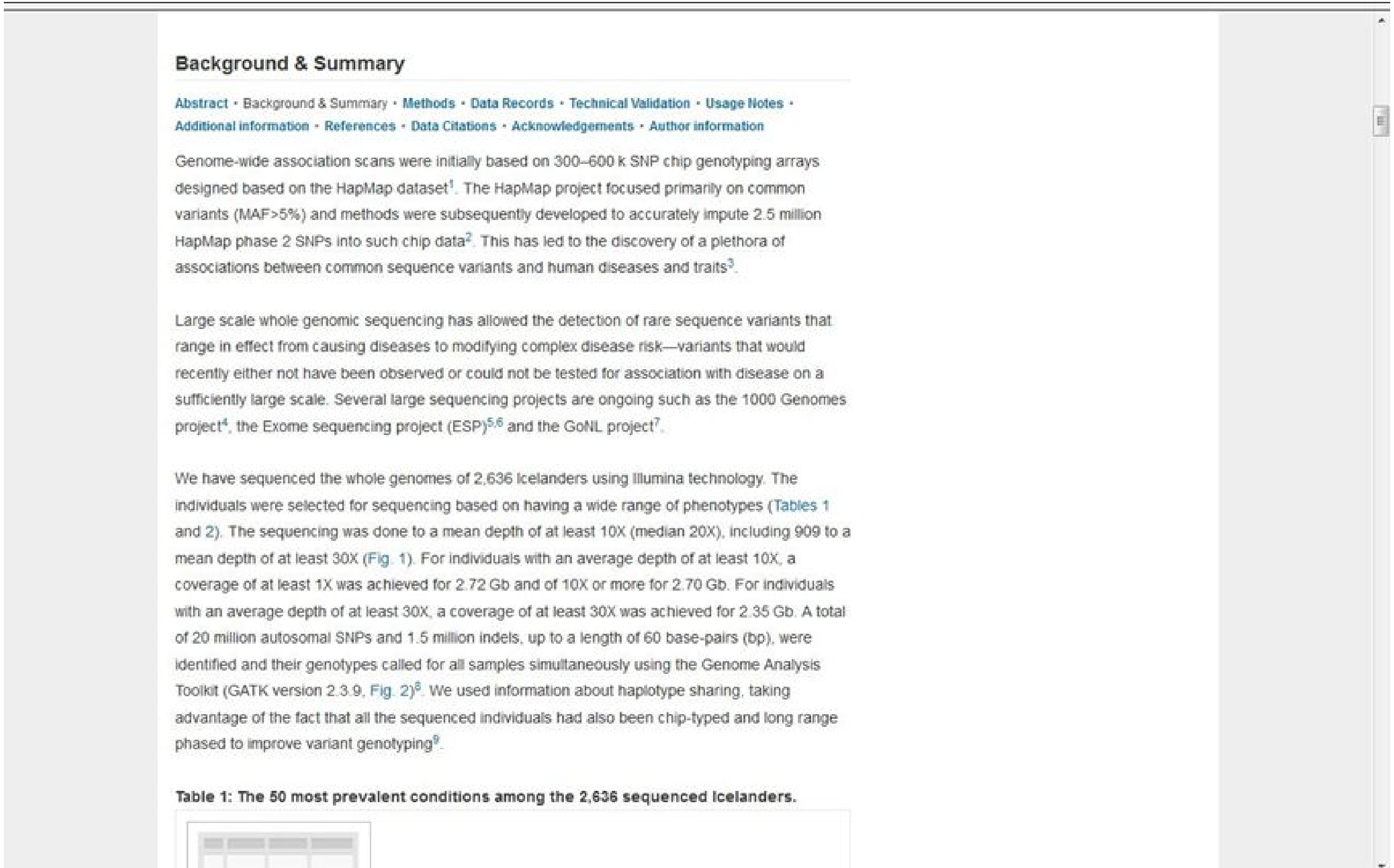
About *Scientific Data*

Scientific Data is an open-access, peer-reviewed publication for descriptions of scientifically valuable datasets. Our primary article-type, the **Data Descriptor**, is designed to make your data more discoverable, interpretable and reusable.

Browse publications by date ▸

Browse publications by subject ▸

Data Papers, Data Journals



Background & Summary

Abstract • Background & Summary • Methods • Data Records • Technical Validation • Usage Notes • Additional information • References • Data Citations • Acknowledgements • Author information

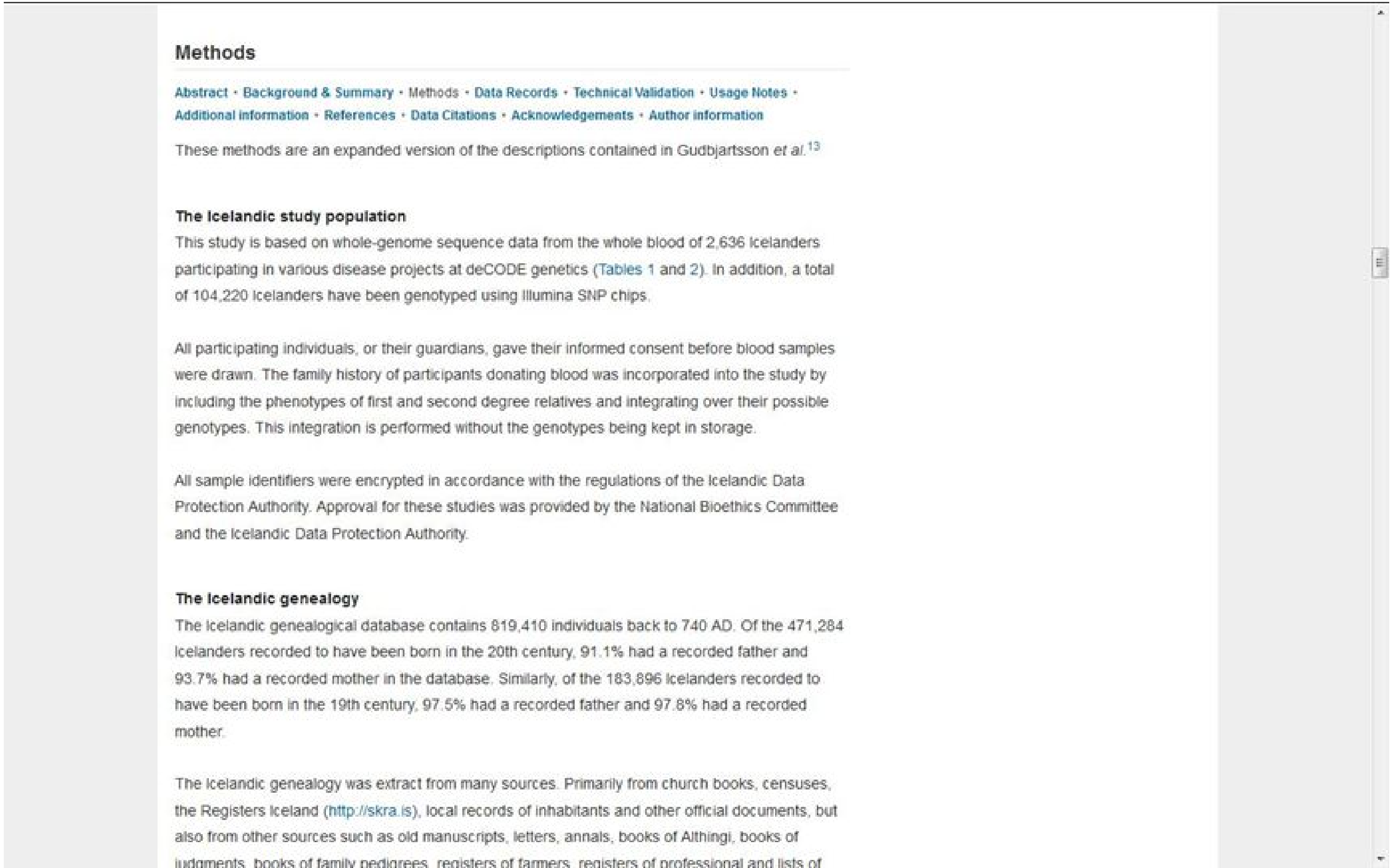
Genome-wide association scans were initially based on 300–600 k SNP chip genotyping arrays designed based on the HapMap dataset[1]. The HapMap project focused primarily on common variants (MAF>5%) and methods were subsequently developed to accurately impute 2.5 million HapMap phase 2 SNPs into such chip data[2]. This has led to the discovery of a plethora of associations between common sequence variants and human diseases and traits[3].

Large scale whole genomic sequencing has allowed the detection of rare sequence variants that range in effect from causing diseases to modifying complex disease risk—variants that would recently either not have been observed or could not be tested for association with disease on a sufficiently large scale. Several large sequencing projects are ongoing such as the 1000 Genomes project[4], the Exome sequencing project (ESP)[5,6] and the GoNL project[7].

We have sequenced the whole genomes of 2,636 Icelanders using Illumina technology. The individuals were selected for sequencing based on having a wide range of phenotypes (Tables 1 and 2). The sequencing was done to a mean depth of at least 10X (median 20X), including 909 to a mean depth of at least 30X (Fig. 1). For individuals with an average depth of at least 10X, a coverage of at least 1X was achieved for 2.72 Gb and of 10X or more for 2.70 Gb. For individuals with an average depth of at least 30X, a coverage of at least 30X was achieved for 2.35 Gb. A total of 20 million autosomal SNPs and 1.5 million indels, up to a length of 60 base-pairs (bp), were identified and their genotypes called for all samples simultaneously using the Genome Analysis Toolkit (GATK version 2.3.9, Fig. 2)[8]. We used information about haplotype sharing, taking advantage of the fact that all the sequenced individuals had also been chip-typed and long range phased to improve variant genotyping[9].

Table 1: The 50 most prevalent conditions among the 2,636 sequenced Icelanders.

Data Papers, Data Journals



Methods

Abstract • Background & Summary • Methods • Data Records • Technical Validation • Usage Notes • Additional information • References • Data Citations • Acknowledgements • Author information

These methods are an expanded version of the descriptions contained in Gudbjartsson *et al.*[13]

The Icelandic study population

This study is based on whole-genome sequence data from the whole blood of 2,636 Icelanders participating in various disease projects at deCODE genetics (Tables 1 and 2). In addition, a total of 104,220 Icelanders have been genotyped using Illumina SNP chips.

All participating individuals, or their guardians, gave their informed consent before blood samples were drawn. The family history of participants donating blood was incorporated into the study by including the phenotypes of first and second degree relatives and integrating over their possible genotypes. This integration is performed without the genotypes being kept in storage.

All sample identifiers were encrypted in accordance with the regulations of the Icelandic Data Protection Authority. Approval for these studies was provided by the National Bioethics Committee and the Icelandic Data Protection Authority.

The Icelandic genealogy

The Icelandic genealogical database contains 819,410 individuals back to 740 AD. Of the 471,284 Icelanders recorded to have been born in the 20th century, 91.1% had a recorded father and 93.7% had a recorded mother in the database. Similarly, of the 183,896 Icelanders recorded to have been born in the 19th century, 97.5% had a recorded father and 97.8% had a recorded mother.

The Icelandic genealogy was extract from many sources. Primarily from church books, censuses, the Registers Iceland (http://skra.is), local records of inhabitants and other official documents, but also from other sources such as old manuscripts, letters, annals, books of Althingi, books of judgments, books of family pedigrees, registers of farmers, registers of professional and lists of

Repositoris

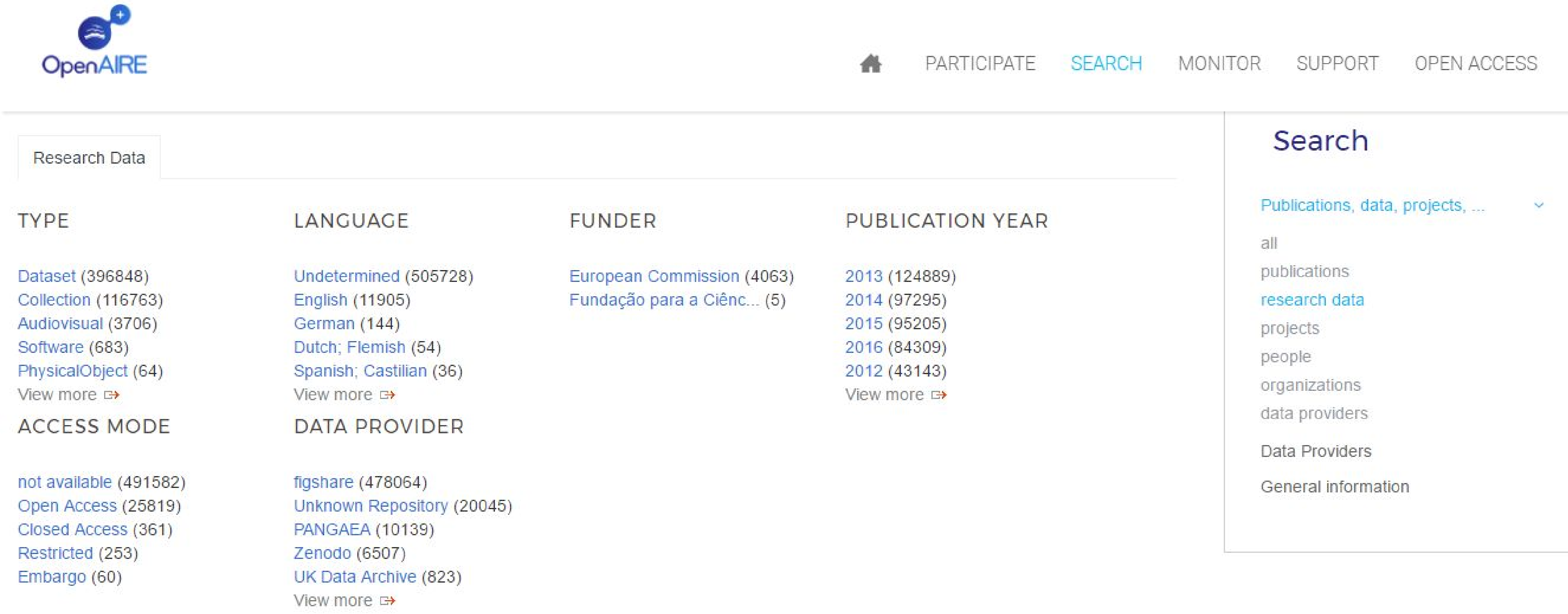
UAB
Universitat Autònoma de Barcelona
Servei de Biblioteques

Repositoris multidisciplinars per a dades en Accés Obert

	zenodo CERN	figshare Digital Science (Macmillan)	DRYAD NCSU	MENDELEY DATA Elsevier	Harvard Dataverse IQSS (Harvard University)
1. Àmbit temàtic i geogràfic	Multidisciplinar Internacional	Multidisciplinar Internacional	Amb èmfasi en ciència i medicina Internacional	Multidisciplinar Internacional	Multidisciplinar Internacional
2. Tipologia de dades (programari, imatges, raw data...)?	Tot tipus de formats i també documents publicats: publicacions, literatura grisa, imatges i audiovisuals, text pla ...	Tot tipus de formats: xifres, conjunts de dades, imatges i audiovisuals, literatura grisa, conjunts d'arxius...	Tot tipus de formats (amb exclusió de software) i amb preferència per formats de codi obert	Tot tipus de formats: text sense format, imatges i audiovisuals, PDF, HTML...	Tot tipus de formats amb funcions específiques per alguns (Stata, SPSS, R, Excel, CSV....)
3. Dades obertes, embargades, restringides o tancades?	Obertes Tancades Restringides Embargades (Policies)	Obertes Tancades (per defecte fins que no es publiquen) Restringides Embargades	Obertes Embargades (per defecte fins el moment de la publicació)	Obertes Tancades (Per defecte fins que no es publiquen) Restringides Embargades	Obertes Restringides La descripció (títol....) d'un conjunt de dades és pública però els arxius poden ser públics o restringits. Sense informació sobre embargaments
4. Mida aproximada dels fitxers a incloure?	Fins a 50 GB per un conjunt de dades (dataset) Per mides superiors es pot contactar amb ells	Fins 5 GB	Fins a 1 GB per fitxer i 20 GB per un conjunt de dades (cost addicional per sobre de 20 GB)	Fins a 10 GB	Fins a 2 GB. Per fitxers més grans contactar amb support@dataverse.org
5. Llicències amb les que es volen difondre les dades?	Permet varietat de llicències encara que recomanen llicències obertes. La descripció de les dades està sota CC0	-CC0 per a dades (*metadata* & *datasets*) -CC-BY per a no dades (audiovisuals, *papers*....) -MIT per a codi i software	CC0 (FAQ)	Permet varietat de llicències : CC-BY 4.0, CC-BY-NC, CC0 1.0, MIT, Apache 2.0, BSD 3 clause, BSD 2 clause, gplv3	Per defecte CC0 però es pot canviar

http://ddd.uab.cat/pub/guibib/150829/repositorisdades_2016m9.pdf

Repositoris



Research Data

TYPE

Dataset (396848)
Collection (116763)
Audiovisual (3706)
Software (683)
PhysicalObject (64)
View more

LANGUAGE

Undetermined (505728)
English (11905)
German (144)
Dutch; Flemish (54)
Spanish; Castilian (36)
View more

FUNDER

European Commission (4063)
Fundação para a Ciênc... (5)

PUBLICATION YEAR

2013 (124889)
2014 (97295)
2015 (95205)
2016 (84309)
2012 (43143)
View more

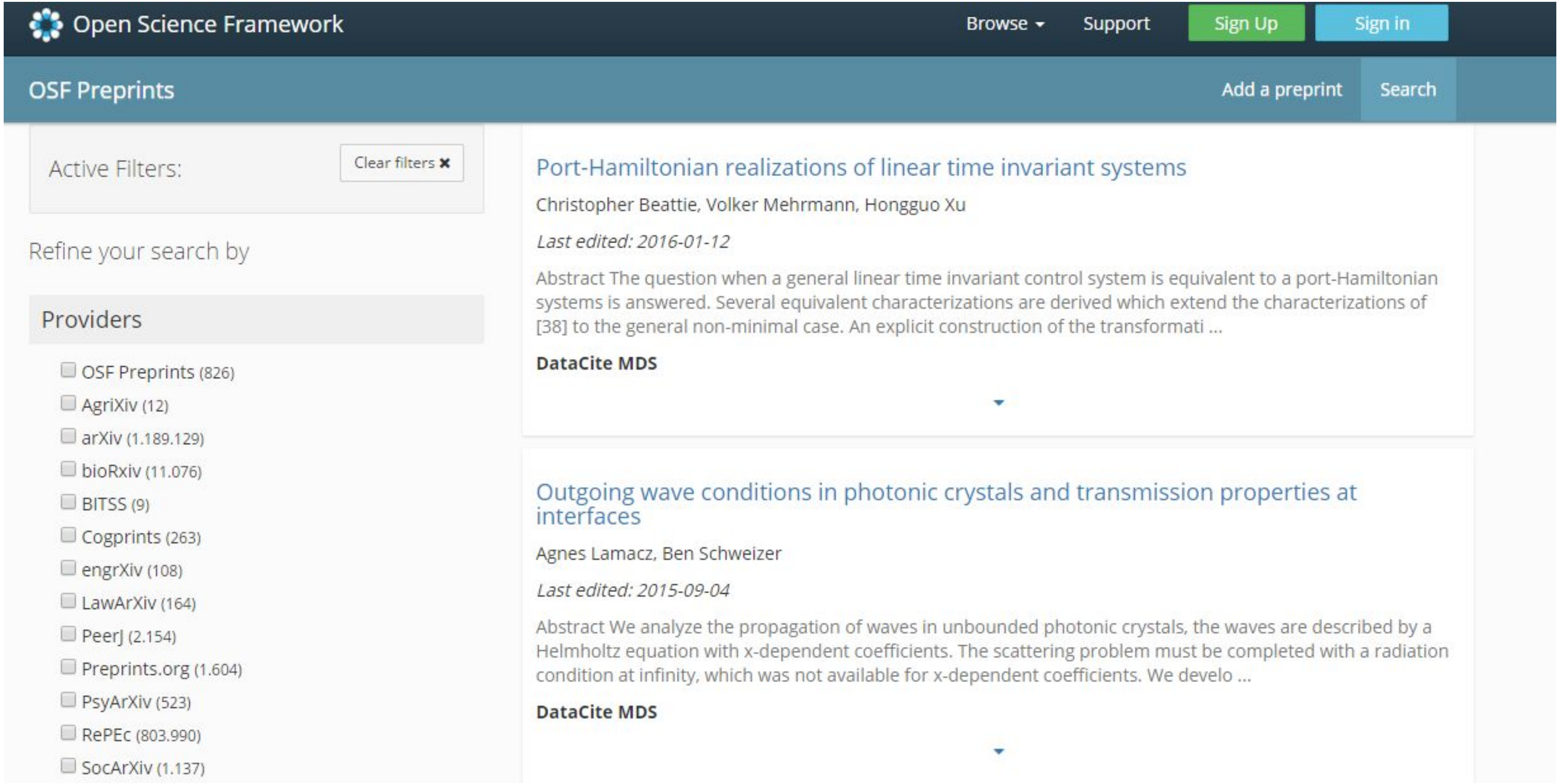
ACCESS MODE

not available (491582)
Open Access (25819)
Closed Access (361)
Restricted (253)
Embargo (60)

DATA PROVIDER

figshare (478064)
Unknown Repository (20045)
PANGAEA (10139)
Zenodo (6507)
UK Data Archive (823)
View more

Repositoris



Active Filters:

Refine your search by

Providers

- ☐ OSF Preprints (826)
- ☐ AgriXiv (12)
- ☐ arXiv (1.189.129)
- ☐ bioRxiv (11.076)
- ☐ BITSS (9)
- ☐ Cogprints (263)
- ☐ engrXiv (108)
- ☐ LawArXiv (164)
- ☐ PeerJ (2.154)
- ☐ Preprints.org (1.604)
- ☐ PsyArXiv (523)
- ☐ RePEc (803.990)
- ☐ SocArXiv (1.137)

Port-Hamiltonian realizations of linear time invariant systems

Christopher Beattie, Volker Mehrmann, Hongguo Xu

Last edited: 2016-01-12

Abstract The question when a general linear time invariant control system is equivalent to a port-Hamiltonian systems is answered. Several equivalent characterizations are derived which extend the characterizations of [38] to the general non-minimal case. An explicit construction of the transformati ...

DataCite MDS

Outgoing wave conditions in photonic crystals and transmission properties at interfaces

Agnes Lamacz, Ben Schweizer

Last edited: 2015-09-04

Abstract We analyze the propagation of waves in unbounded photonic crystals, the waves are described by a Helmholtz equation with x-dependent coefficients. The scattering problem must be completed with a radiation condition at infinity, which was not available for x-dependent coefficients. We develo ...

DataCite MDS

Reflexions pel debat

- Qui cita què?
- Qui avalua què?
- Inflació de la producció científica?
- Paper dels *papers*
- Revisors de datasets?
- De nou, pal o pastanaga?
- Valor afegit dels datasets

Bibliografia

García-García, Alicia; López-Borrull, Alexandre; Peset, Fernanda (2015). “Data journals: eclosión de nuevas revistas especializadas en datos”. El profesional de la información, v. 24, n. 6, pp. 845-854. http://dx.doi.org/10.3145/epi.2015.nov.17

Gómez, Nancy-Diana; Méndez, Eva; Hernández-Pérez, Tony (2016). “Social sciences and humanities research data and metadata: A perspective from thematic data repositories”. El profesional de la información, v. 25, n. 4, pp. 545-555. http://dx.doi.org/10.3145/epi.2016.jul.04

Ollé Castellà, C., López-Borrull, A., and Abadal, E. (2016) "The challenges facing library and information science journals: editors' opinions". Learned Publishing, vol. 29, num. 2, pp. 89-94

Pampel, H; Dallmaier-Tiessen, S. (2014) “Open Research Data: From Vision to Practice”. En: Bartling, S.; Friesike, S (eds.). Opening Science (2014). http://www.openingscience.org/get-the-book

Peset, F.; Aleixandre, R.; Blasco, Y. y Ferrer, A. (2017). “Datos abiertos de investigación. Camino recorrido y cuestiones pendientes”. Anales de Documentación, vol. 20, nº 1. Disponible en: http://dx.doi.org/10.6018/analesdoc.20.1.272101

Serrano-Vicente, R.; Melero, R.; Abadal, E. (2016). "Open Access Awareness and Perceptions in an Institutional Landscape". The Journal of Academic Librarianship vol. 42, p. 595–603

Gràcies!!!

@alexandrelopez