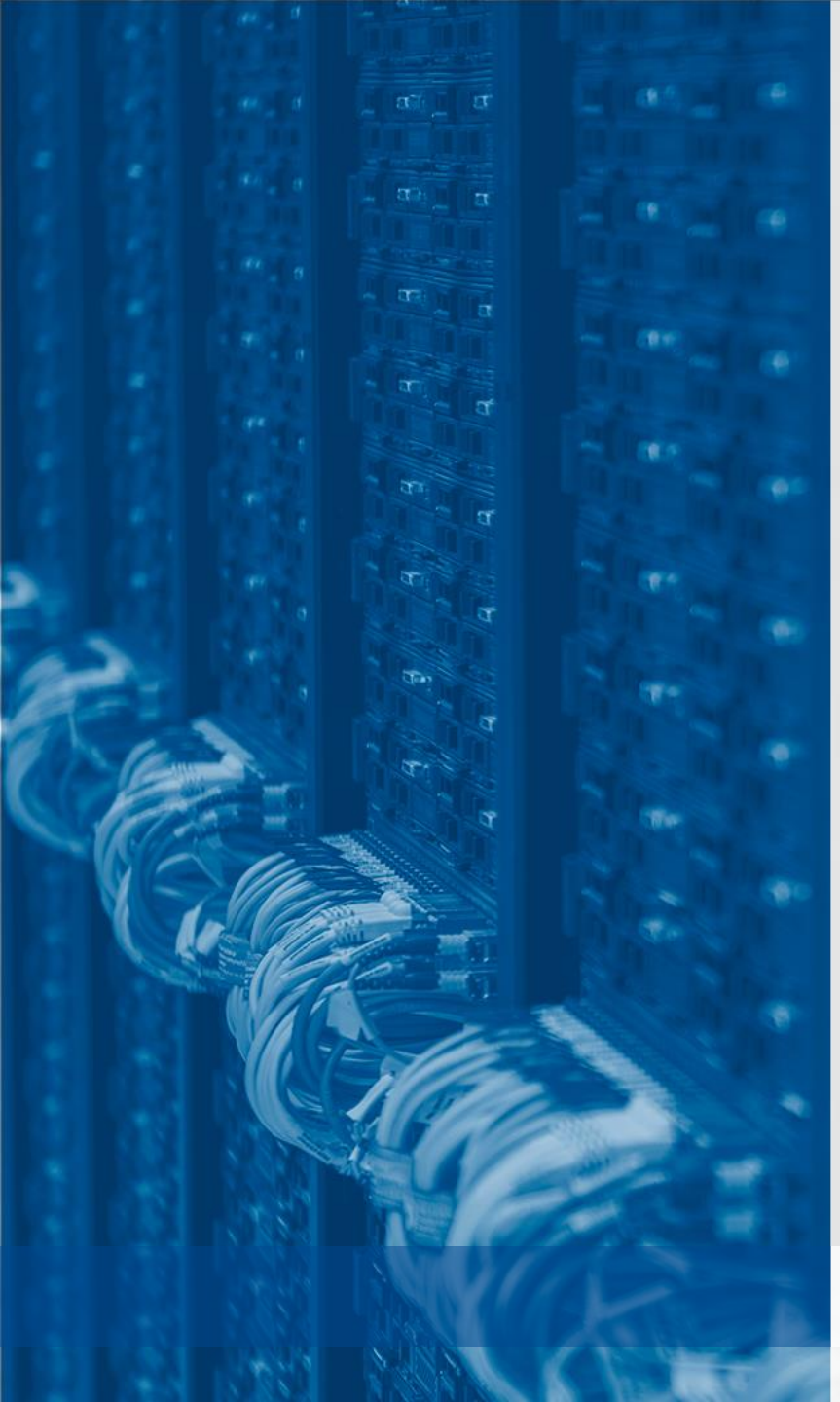


Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Data Sharing across Scientific Domains and Sectors

Mercè Crosas, Ph.D.
Head of Computational Social Sciences Program, bsc.es
President of CODATA, codata.org

Jornada de Ciència Oberta a la UIB, 26 novembre, 2023

´The Care and Feeding of Data´ across scientific domains, sectors, continents

Physics
Astrophysics

Theoretical Modeling, Computing

Engineering

Research Software

Data Analysis, Computing

Software Development

Information Management Systems

Biotechnologies

Quantitative Social Sciences

Data Sharing

Data Repositories

Open-Source

Open Science

Data Management

FAIR Data principles

Data Science

Data Commons

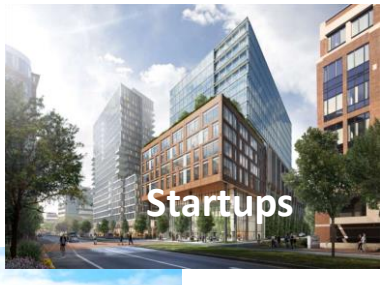
Open Government
Open Data

Computational Social Sciences

Data Science, AI

Data Policies

COMMITTEE ON DATA
CODATA
INTERNATIONAL SCIENCE COUNCIL



Barcelona Supercomputing Center
Centro Nacional de Supercomputac

The Dataverse® Project

BSC Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Science, Data, Computation, Open

A Life of Data Sharing Experiences:

1. **Foundations, Principles, Implementation**
2. **Open Data Across Sectors and Domains**
3. **Science with Computation and AI, and Openness in mind**

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Data Sharing Experiences:

1. **Foundations, Principles, Implementation**
2. Open Data Across Sectors and Domains
3. Science with Computation and AI, and Openness in mind

The Care and Feeding of Scientific data

PLoS Comput Biol. 2014 Apr; 10(4): e1003542. PMCID: PMC3998871
Published online 2014 Apr 24. doi: 10.1371/journal.pcbi.1003542 PMID: 24763340

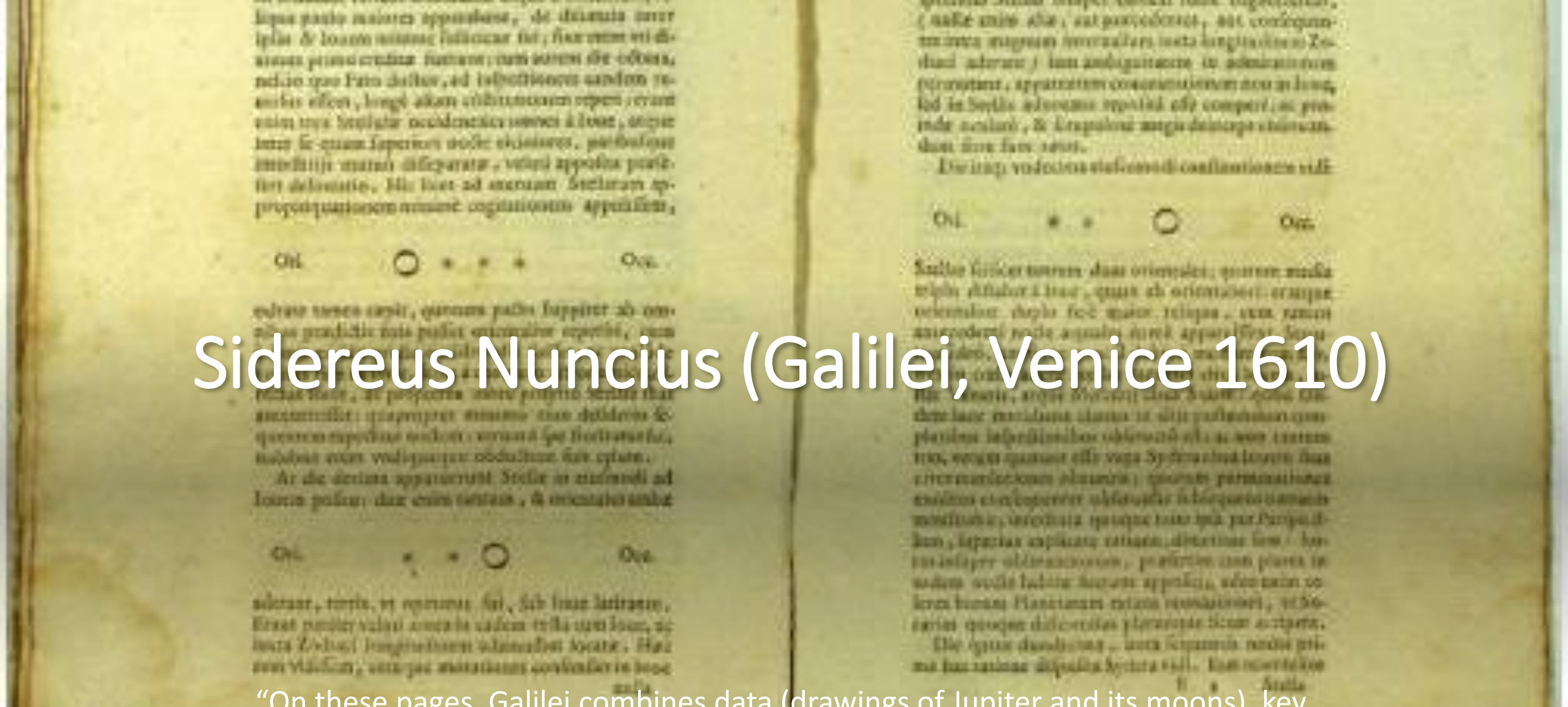
Ten Simple Rules for the Care and Feeding of Scientific Data

Alyssa Goodman, [1] Alberto Pepe, [1,*] Alexander W. Blocker, [1] Christine L. Borgman, [2] Kyle Cranmer, [3] Merce Crosas, [1] Rosanne Di Stefano, [1] Yolanda Gil, [4] Paul Groth, [5] Margaret Hedstrom, [6] David W. Hogg, [3] Vinay Kashyap, [1] Ashish Mahabal, [7] Aneta Siemiginowska, [1] and Aleksandra Slavkovic [8]

Philip E. Bourne, Editor

- **Rule 1**. Love Your Data, and Help Others Love It, Too
- **Rule 2.** Share Your Data Online, with a Permanent Identifier
- **Rule 3.** Conduct Science with a Particular Level of Reuse in Mind
- **Rule 4.** Publish Workflow as Context
- **Rule 5.** Link Your Data to Your Publications as Often as Possible
- **Rule 6.** Publish Your Code (Even the Small Bits)
- **Rule 7.** State How You Want to Get Credit
- **Rule 8.** Foster and Use Data Repositories
- **Rule 9.** Reward Colleagues Who Share Their Data Properly

Sidereus Nuncius (Galilei, Venice 1610)

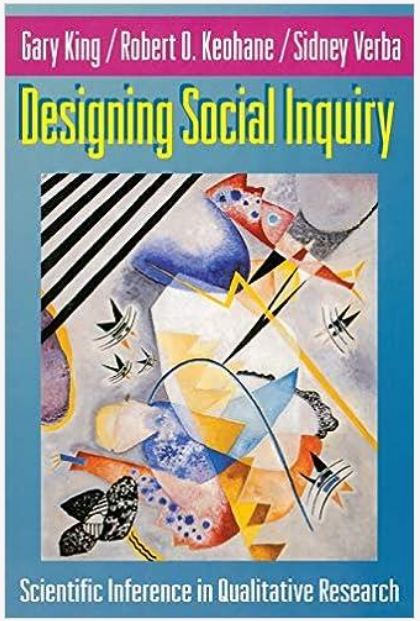


"On these pages, Galilei combines data (drawings of Jupiter and its moons), key metadata (timing of each observation, weather, and telescope properties), and text (descriptions of methods, analysis, and conclusions)."

Goodman A, Pepe A, Blocker AW, Borgman CL, Cranmer K, Crosas M, et al. (2014) Ten Simple Rules for the Care and Feeding of Scientific Data. PLoS Comput Biol 10(4): e1003542. https://doi.org/10.1371/journal.pcbi.1003542

Why data sharing?

- Scientific publications serve as the process to find and remove error from research
- But the materials and methods sections of scientific papers cannot support all the information (data and code) to verify science as it is done today
- **Data and code sharing are part of being scientific**



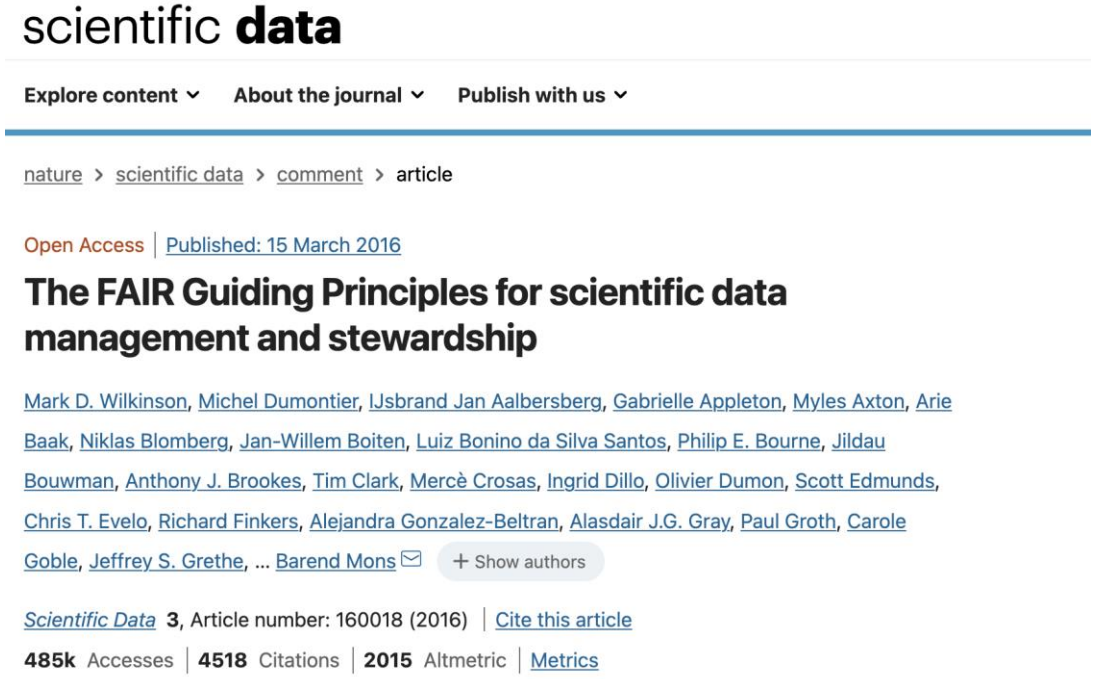
Scientific inference in Qualitative and Quantitative Research

Science?

- The goal is inference
- **The procedures are public**
- The conclusions are uncertain
- `The unity of all science consists alone in its method, not in its material´ (Karl Pearson, 1892)

"Science at its best is a *social enterprise*"

FAIR: Principles and Implementation



scientific **data**

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature › scientific data › comment › article

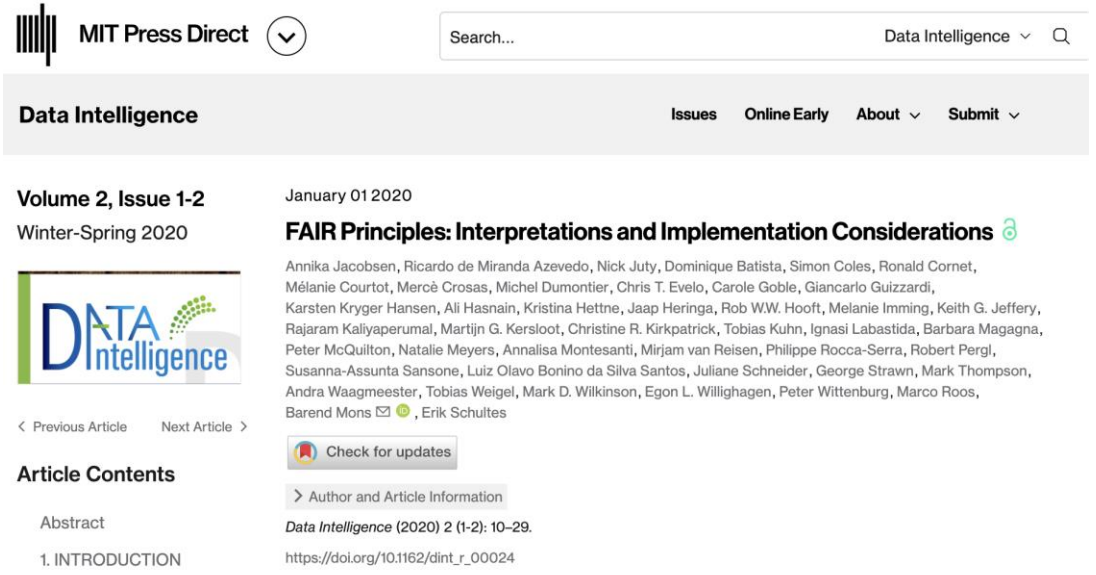
Open Access | Published: 15 March 2016

The FAIR Guiding Principles for scientific data management and stewardship

Mark D. Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E. Bourne, Jildau Bouwman, Anthony J. Brookes, Tim Clark, Mercè Crosas, Ingrid Dillo, Olivier Dumon, Scott Edmunds, Chris T. Evelo, Richard Finkers, Alejandra Gonzalez-Beltran, Alasdair J.G. Gray, Paul Groth, Carole Goble, Jeffrey S. Grethe, ... Barend Mons + Show authors

Scientific Data **3**, Article number: 160018 (2016) | Cite this article

485k Accesses | **4518** Citations | **2015** Altmetric | Metrics



MIT Press Direct ⌄ Search... Data Intelligence ⌄

Data Intelligence Issues Online Early About ⌄ Submit ⌄

Volume 2, Issue 1-2
Winter-Spring 2020

January 01 2020

FAIR Principles: Interpretations and Implementation Considerations

Annika Jacobsen, Ricardo de Miranda Azevedo, Nick Juty, Dominique Batista, Simon Coles, Ronald Cornet, Mélanie Courtot, Mercè Crosas, Michel Dumontier, Chris T. Evelo, Carole Goble, Giancarlo Guizzardi, Karsten Kryger Hansen, Ali Hasnain, Kristina Hettne, Jaap Heringa, Rob W.W. Hooft, Melanie Imming, Keith G. Jeffery, Rajaram Kaliyaperumal, Martijn G. Kersloot, Christine R. Kirkpatrick, Tobias Kuhn, Ignasi Labastida, Barbara Magagna, Peter McQuilton, Natalie Meyers, Annalisa Montesanti, Mirjam van Reisen, Philippe Rocca-Serra, Robert Pergl, Susanna-Assunta Sansone, Luiz Olavo Bonino da Silva Santos, Juliane Schneider, George Strawn, Mark Thompson, Andra Waagmeester, Tobias Weigel, Mark D. Wilkinson, Egon L. Willighagen, Peter Wittenburg, Marco Roos, Barend Mons, Erik Schultes

‹ Previous Article Next Article ›

Check for updates

Author and Article Information

Data Intelligence (2020) 2 (1-2): 10–29.

https://doi.org/10.1162/dint_r_00024

Article Contents

Abstract

1. INTRODUCTION

- 15 principles
- It is not a standard
- The principles apply to:
  - **Metadata:** the descriptor
  - **Data:** the digital object being described (e.g., tabular, domain or instrument-specific format, images, text, code)

FAIR Implementation with the Dataverse open-source platform for data repositories

- Dataverse is repository platform to share research data
- Supports the implementation of FAIR Data:
  - Global Unique Identifiers: Digital Object Identifiers (DOIs)
  - Data licenses, Data User Agreements:
    - Open Data (CC0 or CC-By) when possible
  - Semantic Artifacts:
    - Standard metadata schemas: schema.org, DC
    - Ontologies, taxonomies, vocabularies
    - Data Documentation Initiative (DDI)

The Dataverse® Project

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

HARVARD UNIVERSITY

IQSS
The Institute for Quantitative Social Science

Opening Data while Preserving Privacy with the OpenDP toolkit

OpenDP is an open-source project that provides Differential Privacy (DP) tools:

- A differentially private algorithm:
  - introduces a **minimum amount of noise** to released statistics
  - to **mathematically guarantee the privacy of any individual** in a dataset
- Aims to [Dwork, McSherry, Nissim, Smith, '06]:
  - enable statistical analysis of datasets - **utility**
  - while protecting individual level data – **privacy**
- In the last years, DP has moved from theory to practice

An Introduction to the Joint Principles for Data Citation

by Micah Altman, Christine Borgman, Mercè Crosas and Maryann Martone

NOTE: This article summarizes and extends a longer report published as [1]. Contributors are listed in alphabetical order. We describe contributions to the paper using a standard taxonomy described in [2]. Micah Altman and Mercè Crosas were the lead authors, taking equal responsibility for revisions and authoring the first draft of the manuscript from which this is derived. All authors contributed to the conception of the Force 11 principles discussed, to the methodology, to the project administration and to the writing through critical review and commentary.

Data citation is rapidly emerging as a key practice supporting data access, sharing and reuse, as well as sound and reproducible scholarship. Consensus data citation principles, articulated through the *Joint Declaration of Data Citation Principles* [3], represent an advance in the state of the practice and a new consensus on citation.

Lowering the barrier to research data discovery and use, coupled with an increased ability to link data with publications, could enable new forms of scholarly publishing, promote interdisciplinary research, strengthen the linkage between policy and science and lower the costs of replicating and extending previous research. For this reason, the submission requirements for *Science* – one of the most cited, read and respected journals in the sciences – stipulate that "all data necessary to understand, assess and extend the conclusions of the manuscript must be available to any reader of *Science*" and that "*citations to unpublished data* [emphasis added] and personal communications cannot be used to support claims in a published paper" [4]. Too often, however, this proscription and others like it have been honored only in the breach. Few research articles provide access to the data on which they are based, nor specific citations to data on which the findings rely, nor protocols, algorithms, code or other technology necessary to reproduce, reuse or extend results.

The practice of bibliographic citation to supporting materials was formalized in scholarly publishing more than a century ago. In this tradition, a "bibliographic citation" refers to a formal, structured reference to another scholarly work. In most fields, citations are made in the body of the work. Full references typically appear at the end of the main text, providing more detailed bibliographic information for each work referenced. Following the establishment of the first scientific digital data archives in the late 1960s, bibliographic standards for data were developed and refined over the next decades but never widely used in practice.

The theory and practice of data citation have advanced considerably over the last five years, and these parallel efforts led to concern for a unified approach.

Declaration of Data Citation Principles

On this page:

Data Citation Principles

Cite as: *Data Citation Synthesis Group: Joint Declaration of Data Citation Principles. Martone M. (ed.) San Diego CA: FORCE11; 2014*
https://doi.org/10.25490/a97f-egyk

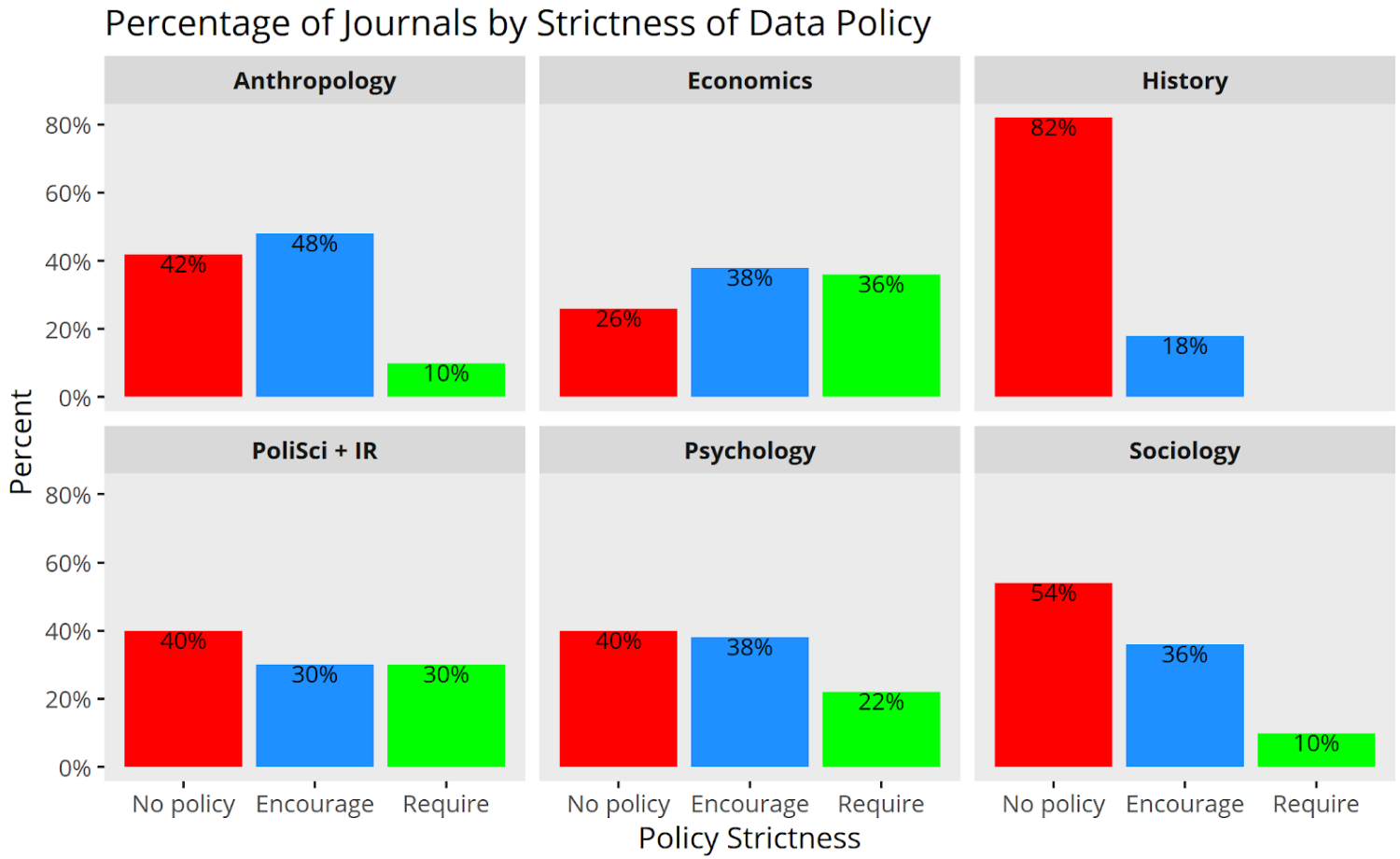
Data Policies of top 50 Social Science Journals

Crosas, Gautier, Karcher, Kirilova, Otalora, Schwartz. 2019
Data Policies of Highly-Ranked Social Science Journals,
preprint, https://osf.io/preprints/socarxiv/9h7ay

No Policy

Encourage data sharing

Require data sharing



Data Sharing is good for you

“Gain more citations and visibility by sharing data”

“Openness as a continuum of practices”

https://elifesciences.org/articles/16800

eLife

Magazine | Feature Article

Biochemistry and Chemical Biology

Point of View: How open science helps researchers succeed

Erin C McKiernan, Philip E Bourne, C Titus Brown, Stuart Buck, Amye Kenall, Jennifer Lin, Damon McDougall, Brian A Nosek, Karthik Ram, Courtney K Soderberg, Jeffrey R Spies, Kaitlin Thaney, Andrew Updegrove, Kara H Woo, Tal Yarkoni « see less

Funding	Description	URL
Shuttleworth Foundation Fellowship Program	funding for researchers working openly on diverse problems	shuttleworthfoundation.org/fellows/
Mozilla Fellowship for Science	funding for researchers interested in open data and open source	www.mozillascience.org/fellows
Leamer-Rosenthal Prizes for Open Social Science (UC Berkeley and John Templeton Foundation)	rewards social scientists for open research and education practices	www.bitss.org/prizes/leamer-rosenthal-prizes/
OpenCon Travel Scholarship (Right to Research Coalition and SPARC)	funding for students and early-career researchers to attend OpenCon, and receive training in open practices and advocacy	www.opencon2016.org/
Preregistration Challenge (Center for Open Science)	prizes for researchers who publish the results of a preregistered study	cos.io/prereg/
Open Science Prize (Wellcome Trust, NIH, and HHMI)	funding to develop services, tools, and platforms that will increase openness in biomedical research	www.openscienceprize.org/

Advances in Computational Reproducibility

Best practices for Computational Reproducibility:

- Share data and code in open trusted repositories
- Use persistent links from publication to data and code
- Citation to data and code should be a standard
- Document data, code, workflows, and computational environment
- Use open license for your code and data

(Stodden et al. 2016, Enhancing reproducibility of computational methods)

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Packaging research artefacts with RO-Crate Cite

Article type: Resource Paper

Authors: Soiland-Reyes, Stian[a; b; *] | Sefton, Peter[c] | Crosas, Mercè[d] | Castro, Leyla Jael[e] | Coppens, Frederik[f] | Fernández, José M.[g] | Garijo, Daniel[h] | Grüning, Björn[i] | La Rosa, Marco[j] | Leo, Simone[k] | Ó Carragáin, Eoghan[l] | Portier, Marc[m] | Trisovic, Ana[n] | RO-Crate Community, [o] | Groth, Paul[p] | Goble, Carole[q]

Affiliations: [a] Department of Computer Science, The University of Manchester, UK | [b] Informatics Institute, University of Amsterdam, The Netherlands | [c] Faculty of Science, University Technology Sydney, Australia | [d] Institute for Quantitative Social Science, Harvard University, Cambridge, MA, USA | [e] ZB MED Information Centre for Life Sciences, Cologne, Germany | [f] VIB-UGent Center for Plant Systems Biology, Gent, Belgium | [g] Barcelona Supercomputing Center, Barcelona, Spain | [h] Ontology

Short Paper P-RECS '20, June 23–26, 2020, Stockholm, Sweden

Advancing Computational Reproducibility in the Dataverse Data Repository Platform

Ana Trisovic
Institute for Quantitative Social Science, Harvard University
Cambridge, MA, USA
anatrisovic@g.harvard.edu

Philip Durbin
Institute for Quantitative Social Science, Harvard University
Cambridge, MA, USA

Tania Schlatter
Institute for Quantitative Social Science, Harvard University
Cambridge, MA, USA

Gustavo Durand
Institute for Quantitative Social Science, Harvard University
Cambridge, MA, USA

Sonia Barbosa
Institute for Quantitative Social Science, Harvard University
Cambridge, MA, USA

Danny Brooke
Institute for Quantitative Social Science, Harvard University
Cambridge, MA, USA

Mercè Crosas
Institute for Quantitative Social Science, Harvard University
Cambridge, MA, USA
mcrosas@g.harvard.edu

Learn from the Replication Crisis

communications psychology

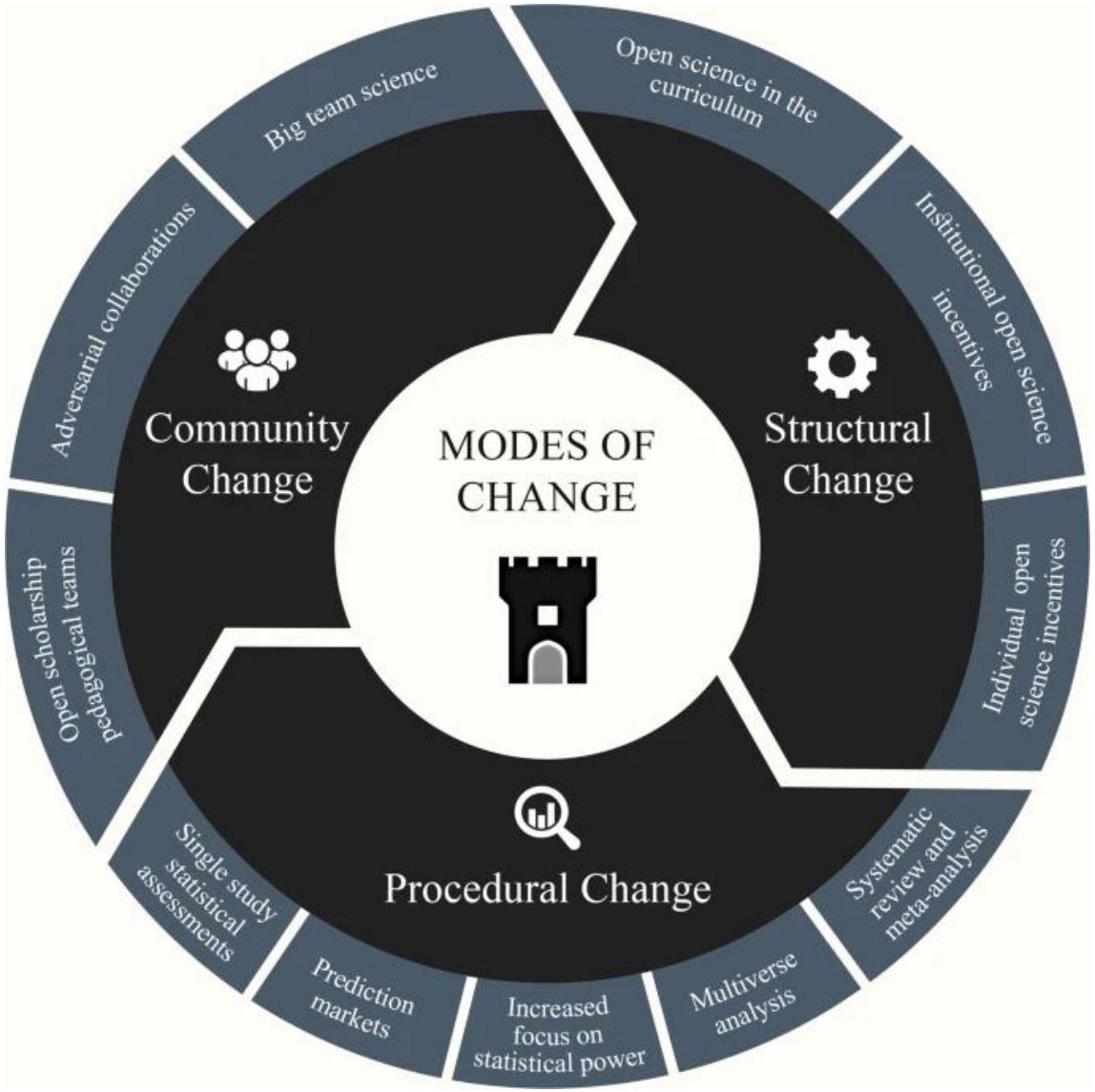
Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature › communications psychology › perspectives › article

Perspective | Open access | Published: 25 July 2023

The replication crisis has led to positive structural, procedural, and community changes

Max Korbmacher, Flavio Azevedo ✉, Charlotte R. Pennington, Helena Hartmann, Madeleine Pownall, Kathleen Schmidt, Mahmoud Elsherif, Nate Breznau, Olly Robertson, Tamara Kalandadze, Shijun Yu, Bradley J. Baker, Aoife O'Mahony, Jørgen Ø. -S. Olsnes, John J. Shaw, Biljana Gjoneska, Yuki Yamada, Jan P. Röer, Jennifer Murphy, Shilaan Alzahawi, Sandra Grinschgl, Catia M. Oliveira, Tobias Wingen, Siu Kit Yeung, ... Thomas Evans + Show authors



Pre-registration, as an option to avoid cognitive biases

PNAS

RESEARCH ARTICLE | PSYCHOLOGICAL AND COGNITIVE SCIENCES |

The preregistration revolution

Brian A. Nosek, Charles R. Ebersole, Alexander C. DeHaven, and David T. Mellor
Authors Info & Affiliations

Edited by Richard M. Shiffrin, Indiana University, Bloomington, IN, and approved August 28, 2017 (received for review June 15, 2017)

March 12, 2018 | 115 (11) 2600-2606 | https://doi.org/10.1073/pnas.1708274114

Pre-Registration Myth	Pre-Registration Reality
✗ Stifles creative or exploratory research	✓ Asks researchers to specify whether research is exploratory
✗ Guarantees quality and fully addresses questionable practices such as *p*-hacking	✓ Only improves quality and addresses questionable practices when done well
✗ Is irrelevant for certain types of research studies (e.g., qualitative research)	✓ Is relevant to many types of research studies and many study components
✗ Fully solves the file-drawer problem (i.e., publication bias)	✓ Only addresses publication bias if the pre-registered study is public and findable (i.e., located on a public platform or external registry)
✗ Is easy to do	✓ Is challenging to do well and requires collaboration within the broader research community
✗ Is time-consuming and expensive	✓ Can save time or add no additional time; may offset costs of errors

COS 10 YEARS — CENTER FOR OPEN SCIENCE — Science Works Best in the Open

About | Engage | Open Science | Research | Events | Blog | Contact

Future-proof your research.
Preregister your next study.

What is Preregistration?

When you preregister your research, you're simply specifying your research plan in advance of your study and submitting it to a registry.

Preregistration separates *hypothesis-generating* (exploratory) from *hypothesis-testing* (confirmatory) research. Both are important. But the same data cannot be used to generate *and* test a hypothesis, which can happen unintentionally and reduce the credibility of your results. Addressing this problem through planning improves the quality and transparency of your research. This helps you clearly report your study and helps others who may wish to build on it. For instructions on how to submit a preregistration on OSF, please visit our help guides.

For additional insight and context, you can read The Preregistration Revolution. (preprint)

https://www.cos.io/initiatives/prereg
https://www.acf.hhs.gov/opre/blog/2022/08/pre-registering-studies-what-it-how-do-you-do-it-and-why

Data Sharing Experiences:

1. Foundations, Principles, Implementation
2. **Open Data Across Sectors and Domains**
3. Science with Computation and AI, and Openness in mind

Open Government, Open Data

Transparency.

"**Transparency means providing the public with information about their government's activities.** It contemplates disclosure about, for example, what federal agencies have done or will do. Transparency's premise is that citizens are entitled to know what, how, and why government does what it does."

Participation.

"Citizens are entitled to more, however, than a transparent view of their government from the outside looking in. Participation emphasizes citizens' voice in public affairs, recognizing that public officials stand to benefit from the perspective of expert and nonexpert knowledge that resides outside of government. **Participation is fostered by expanding citizens' opportunities to express their views about policy alternatives, and in ways beyond voting in elections.**"

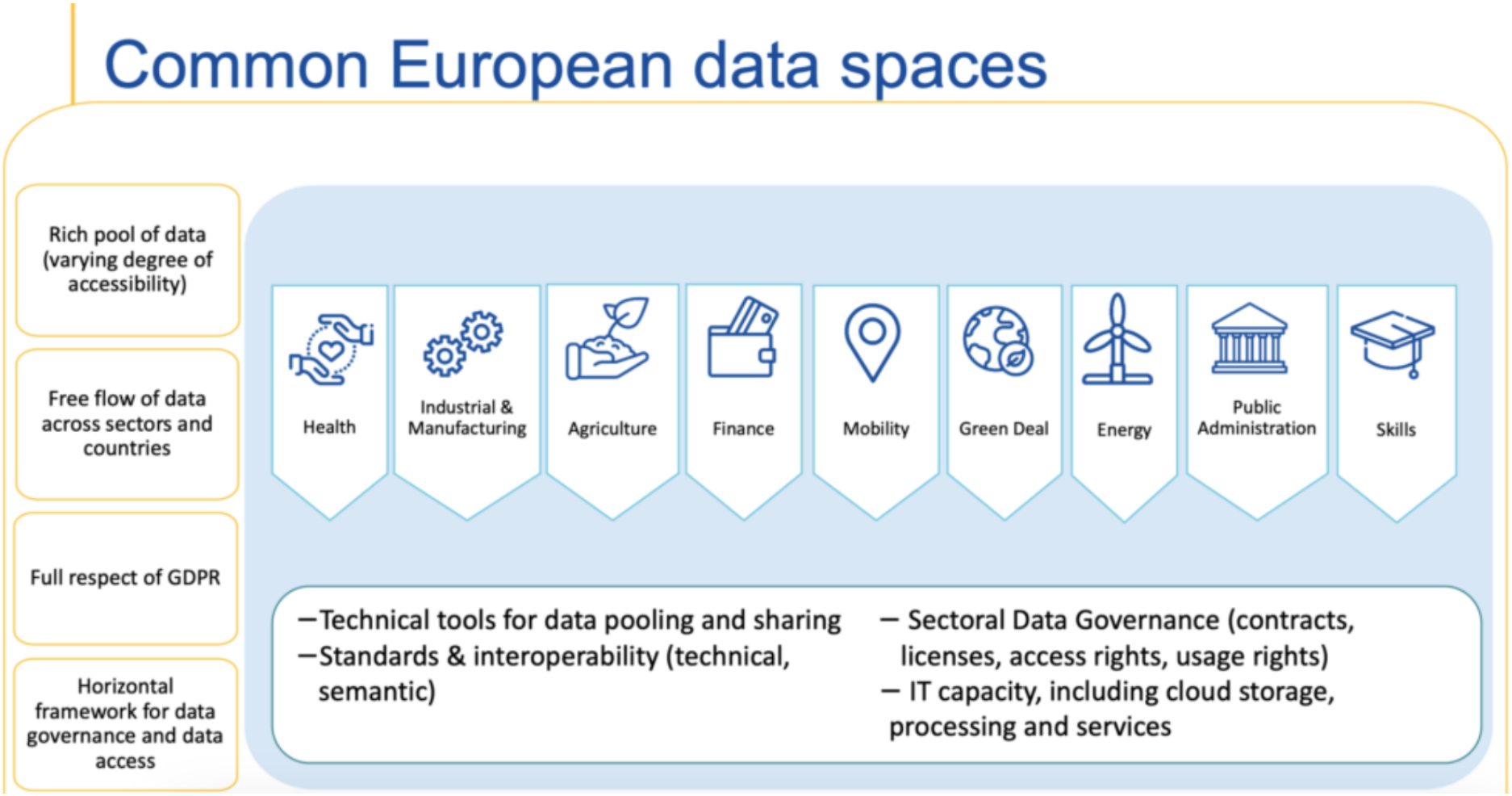
Collaboration.

"Collaboration further erodes the us-versus-them divide between citizens and government by taking participation to another level. Citizens are capable, after all, of more than simply registering their views about policy alternatives defined in advance. They can usefully help shape the government's agenda. They can also help determine even the tools and methods by which public policy goals are pursued. **Where government is collaborative, citizens become true partners with government, in both the identification and pursuit of public goals."**

Using Data for world´s global challenges

Opening Government Data

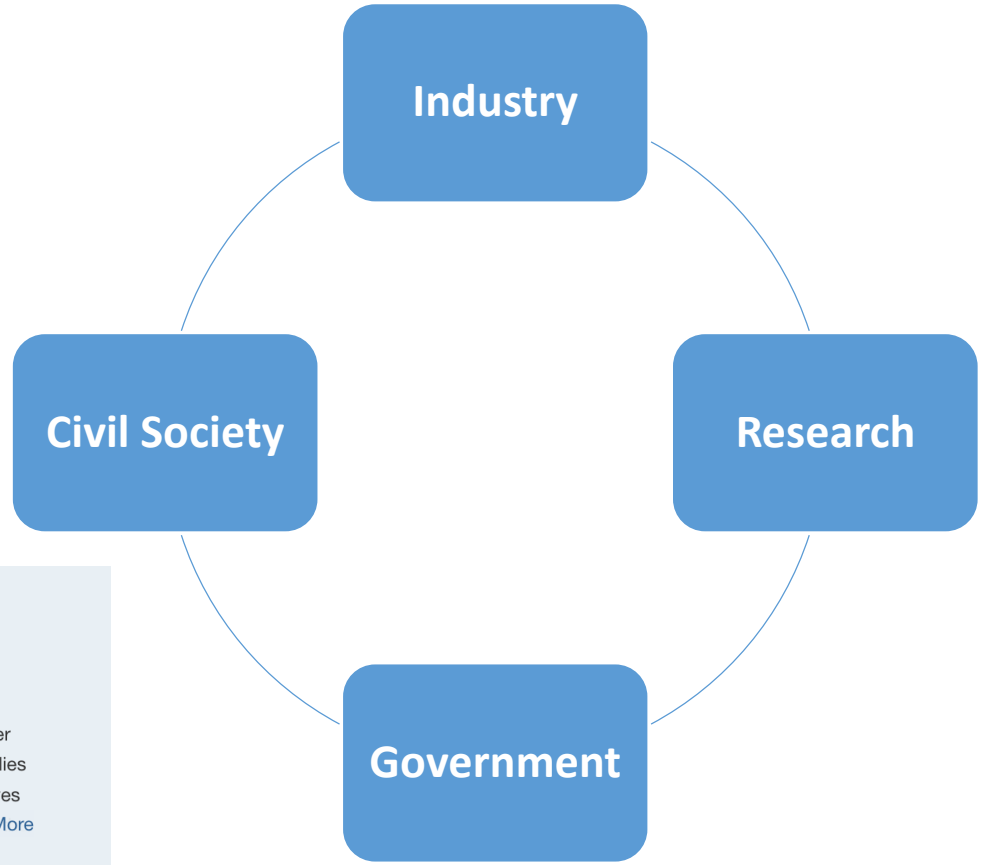
Data Spaces: Data, tools, services exchange for industry



EOSC: Data, tools, services exchange for research

Cocreation: Academia, Gov, Industry, Citizens

The joint production of innovation between combinations of industry, research, government and civil society



OECD publishing

OECD Science, Technology and Industry Policy Papers

OECD

Knowledge co-creation in the 21st century

A cross-country experience-based policy report

The importance of knowledge co-creation – the joint production of innovation between industry, research and possibly other stakeholders, such as civil society – has been increasingly acknowledged. This paper builds on 13 cross-country case studies and co-creation experiences during the COVID-19 pandemic to characterise the diversity of knowledge co-creation initiatives and identify lessons for policy. The paper identifies a strong rationale for policy to support knowledge co-creation becau... More

Published on June 16, 2021

In series: OECD Science, Technology and Industry Policy Papers (view more titles)

Read online | Download PDF | Get citation details

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

CODATA:
Committee on Data of the International Science Council

- **Mission:** Connect data and people to advance science and improve our world
- **Strategic activities** divided into four priorities:
  - Decadal Program '**Making Data Work for Cross-Domain Grand Challenges**' and the Global Open Science Cloud Initiative.
  - **Data Policy:** promoting principles, policies and practices for FAIR Data and Open Science;
  - **Data Science:** advancing the frontiers of the science of data;
  - **Data Skills:** building capacity for Open Science by improving data skills and the functions of national science systems needed to support open data.

WorldFAIR: Cross-Domain Interoperability

11 disciplinary and cross-disciplinary case studies to advance implementation of the FAIR principles and, in particular, to improve interoperability and reusability of digital research objects, including data.

- FAIR Implementation Profiles
- Cross-Domain Interoperability Framework
- FAIR Assessment

COMMITTEE ON DATA CODATA INTERNATIONAL SCIENCE COUNCIL

RDA RESEARCH DATA ALLIANCE

https://worldfair-project.eu/

BSC Barcelona Supercomputing Center Centro Nacional de Supercomputación

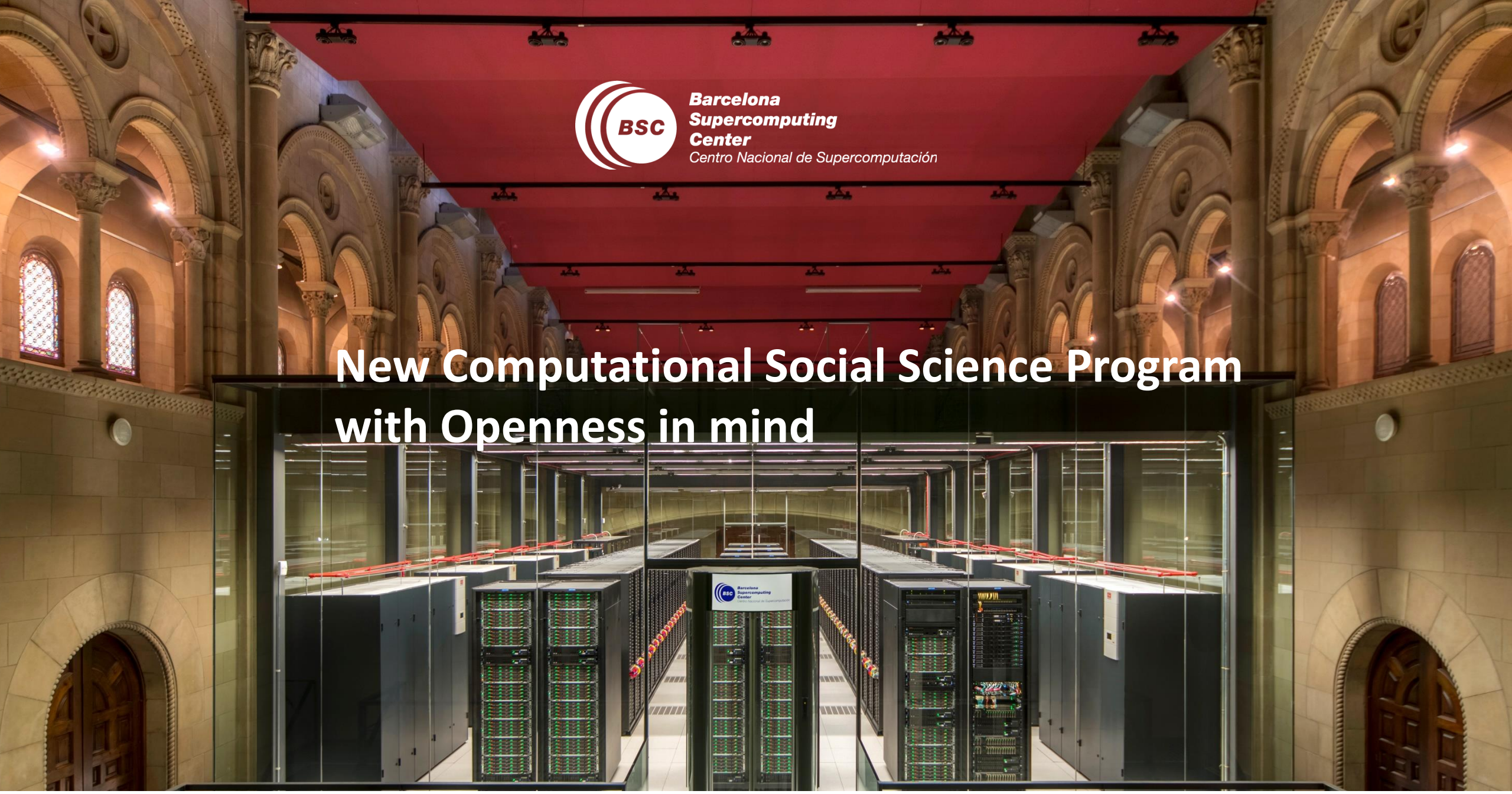
DDI-Cross Domain Integration

- The EC-funded WorldFAIR project is coordinated by CODATA, with RDA as a major partner, and will produce an initial draft of the Cross Domain Interoperability Framework (CDIF) recommendations.
- Using existing specifications: DCAT, Schema.org**, DDI-CDI**, SKOS/XKOS, SSSOM, OGC Observations & Measurements/I-PROV, …

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Data Sharing Experiences:

1. Foundations, Principles, Implementation
2. Open Data Across Sectors and Domains
3. **Science with Computation and AI, and Openness in mind**



New Computational Social Science Program with Openness in mind

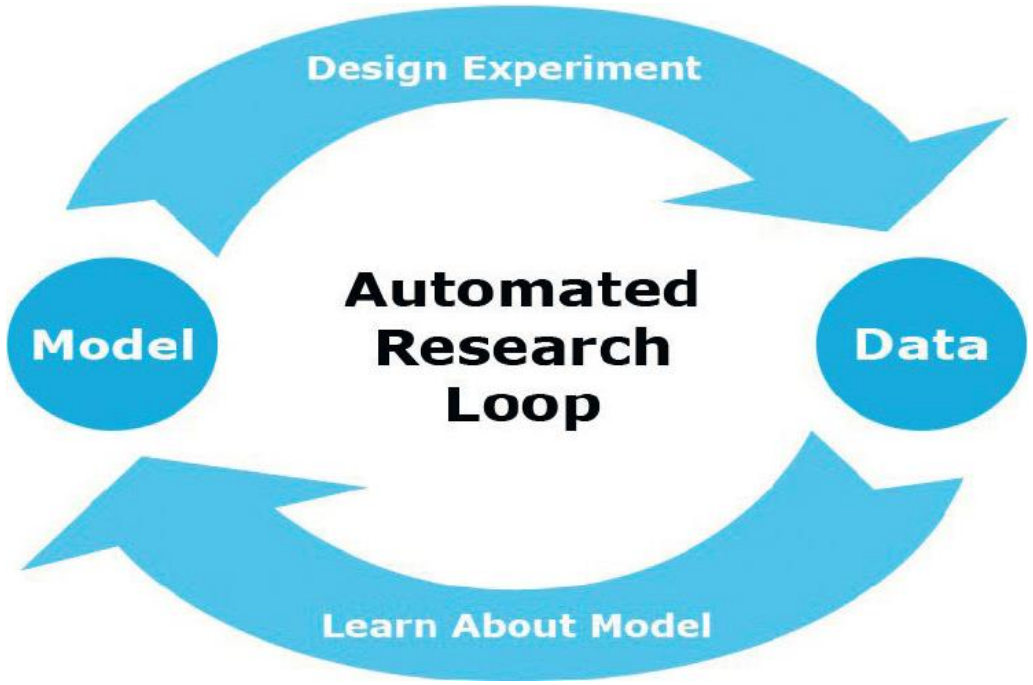
Generalitat de Catalunya
Departament de Recerca
i Universitats

GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

UNIÓN EUROPEA
Fondo Europeo de Desarrollo Regional

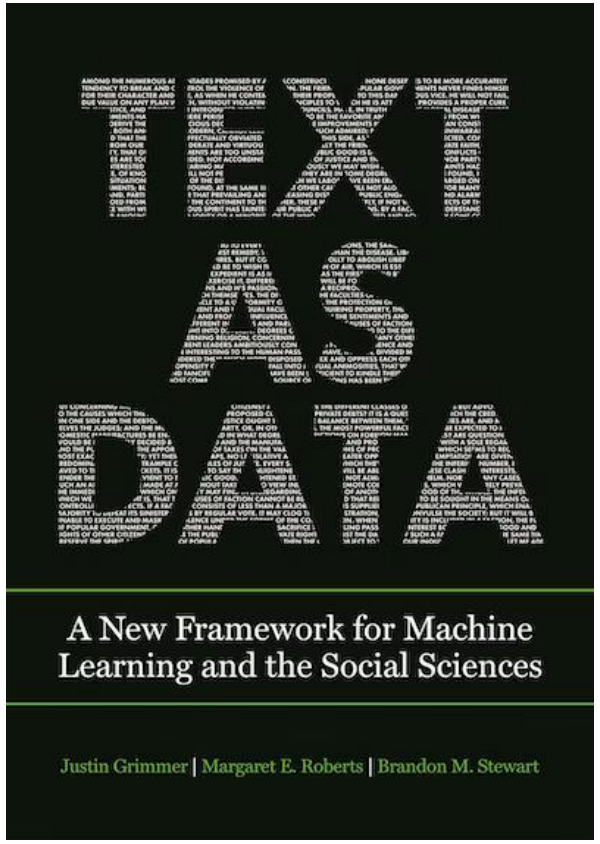
AI to accelerate to Model-Data Loop



- An iterative Data-Model loop for scientific inference
- AI-assisted research workflows for accelerated discovery
- FAIR workflows
- Team Science

National Academies of Sciences, Engineering, and Medicine. 2022. Automated Research Workflows for Accelerated Discovery: Closing the Knowledge Discovery Loop. Washington, DC: The National Academies Press. https://doi.org/10.17226/26532.

Text as Data: ML for Social Sciences



Methods to use Text as Data in three stages of the research process (not necessarily sequential):

- **Discovery:** develop a research question, conceptualization (clusters of documents)
- **Measurement:** categorize text based on the concepts (from theory, discovery phase) to provide descriptions (summaries of the data)
- **Inference (causal and prediction):** Use text to make prediction of changes in the future or understand the effect of an intervention.

Iterative and Inductive model vs Deductive Model

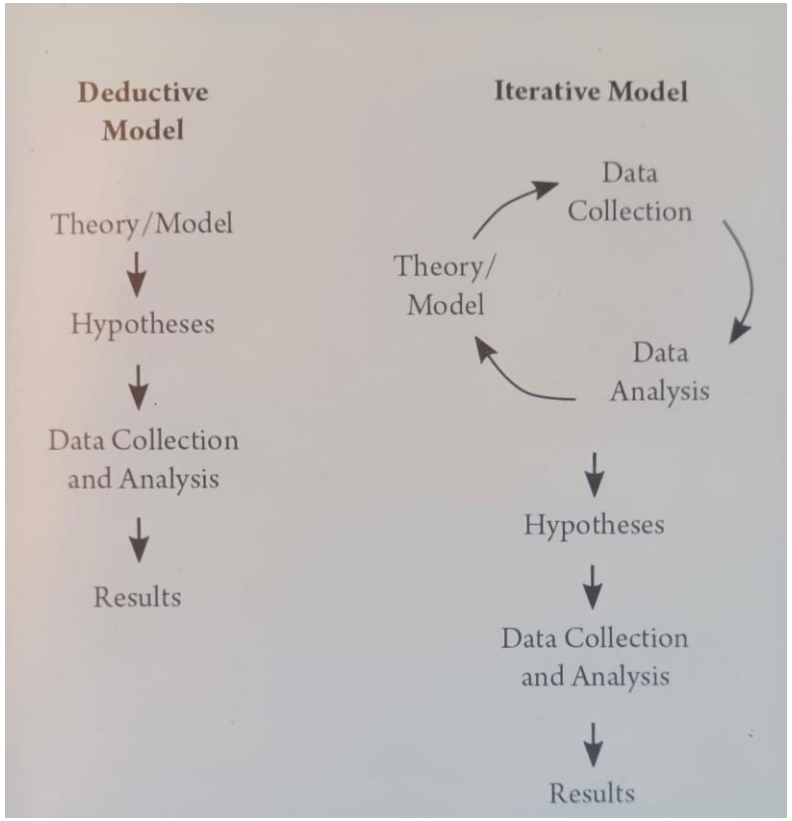
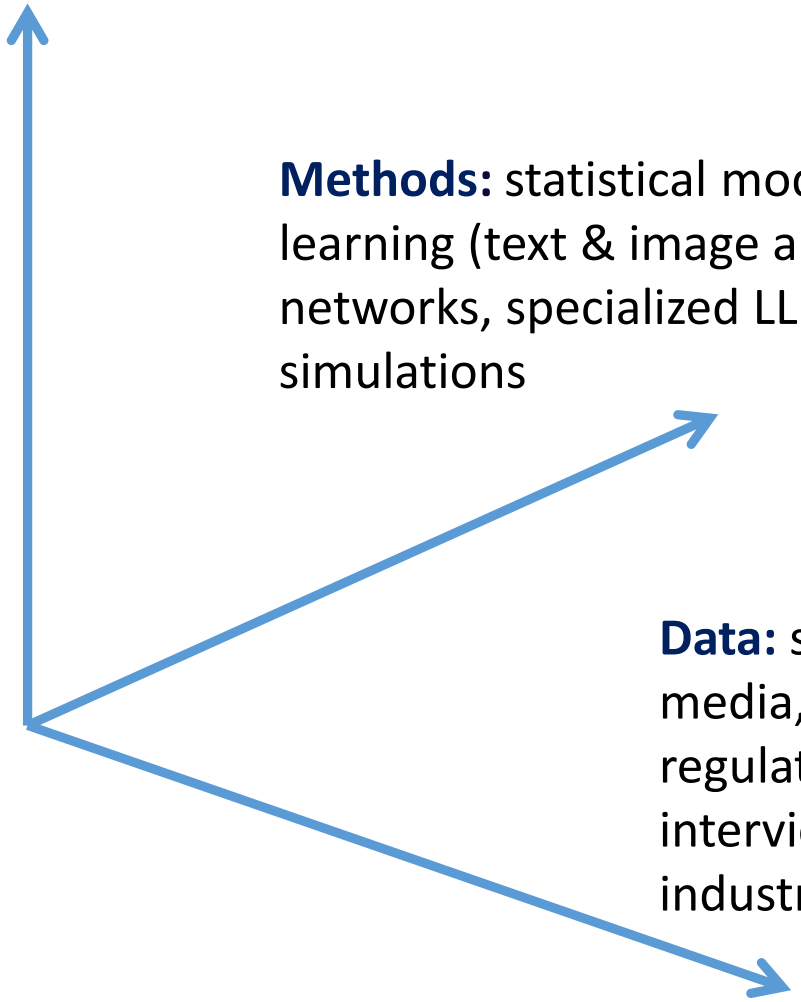


Figure 2. from Text as Data, Grimmer, Roberts, Stewart

Advancing Computational Social Science @BSC

Research areas:

- Demography
- Democracy Quality
- Socioeconomic impact of climate change
- Social Innovation
- History and cultural heritage
- …



Methods: statistical models, machine learning (text & image analysis), social networks, specialized LLMs, agent-based simulations

Data: surveys, national statistics, social media, historical archives, laws, regulations and national plans, interviews, citizens volunteered data, industry data

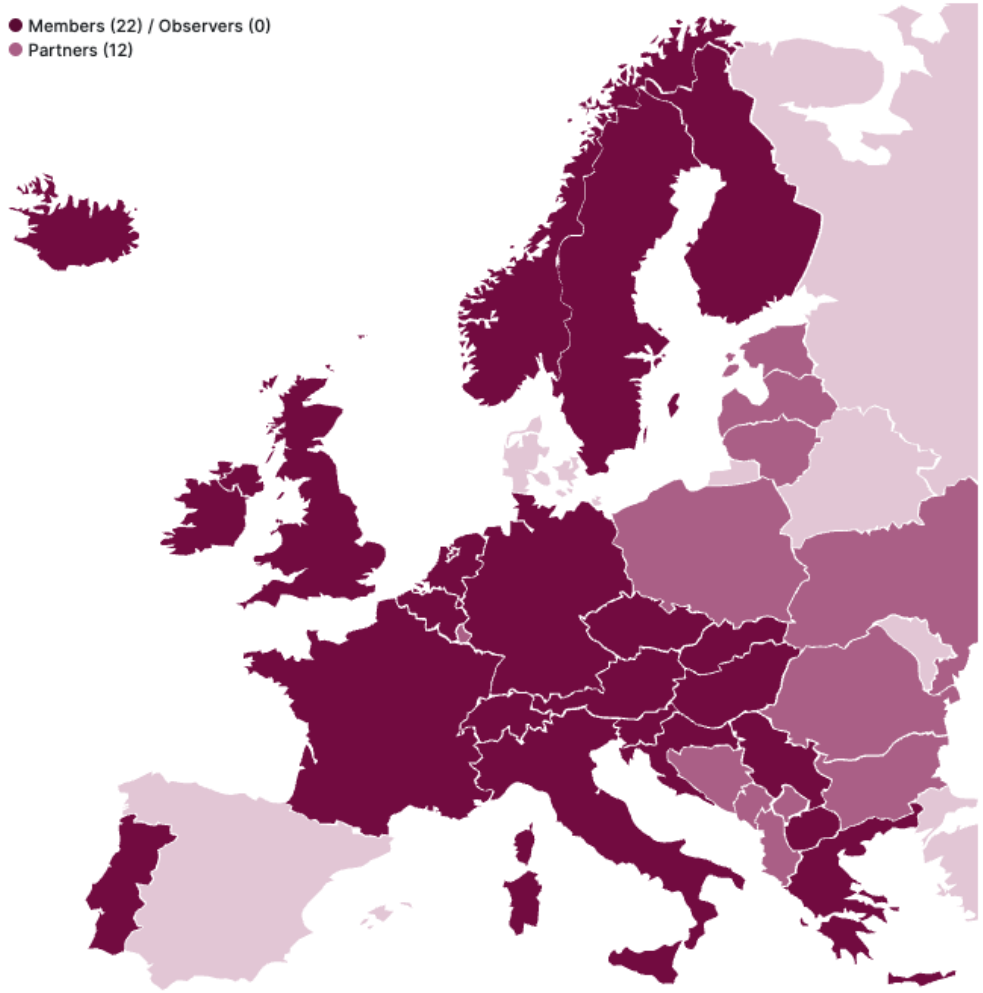
Program´s future Tools and Services

- To **build a FAIR data repository for Social Sciences (in Spain), integrate with CESSDA**
  - **Federate existing social science data repositories**
  - Do not duplicate efforts, use an open-source data sharing platform (dataverse.org)
- To **integrate datasets in the repository with supercomputing**
  - Implement **open computational workflows** to prepare dataset analysis for High-Performance Computing
- To **explore AI to assist with generation of common metadata**
  - Create metadata mappings and connect datasets across disciplines/domains in the data space
- To **test systems for sharing and analyzing sensitive data for social science research**
  - Set up data enclaves with levels of access and explore the use of differential privacy (OpenDP.org)

The need for a Social Science Data Repository in Spain

Start a process to become part of CESSDA:

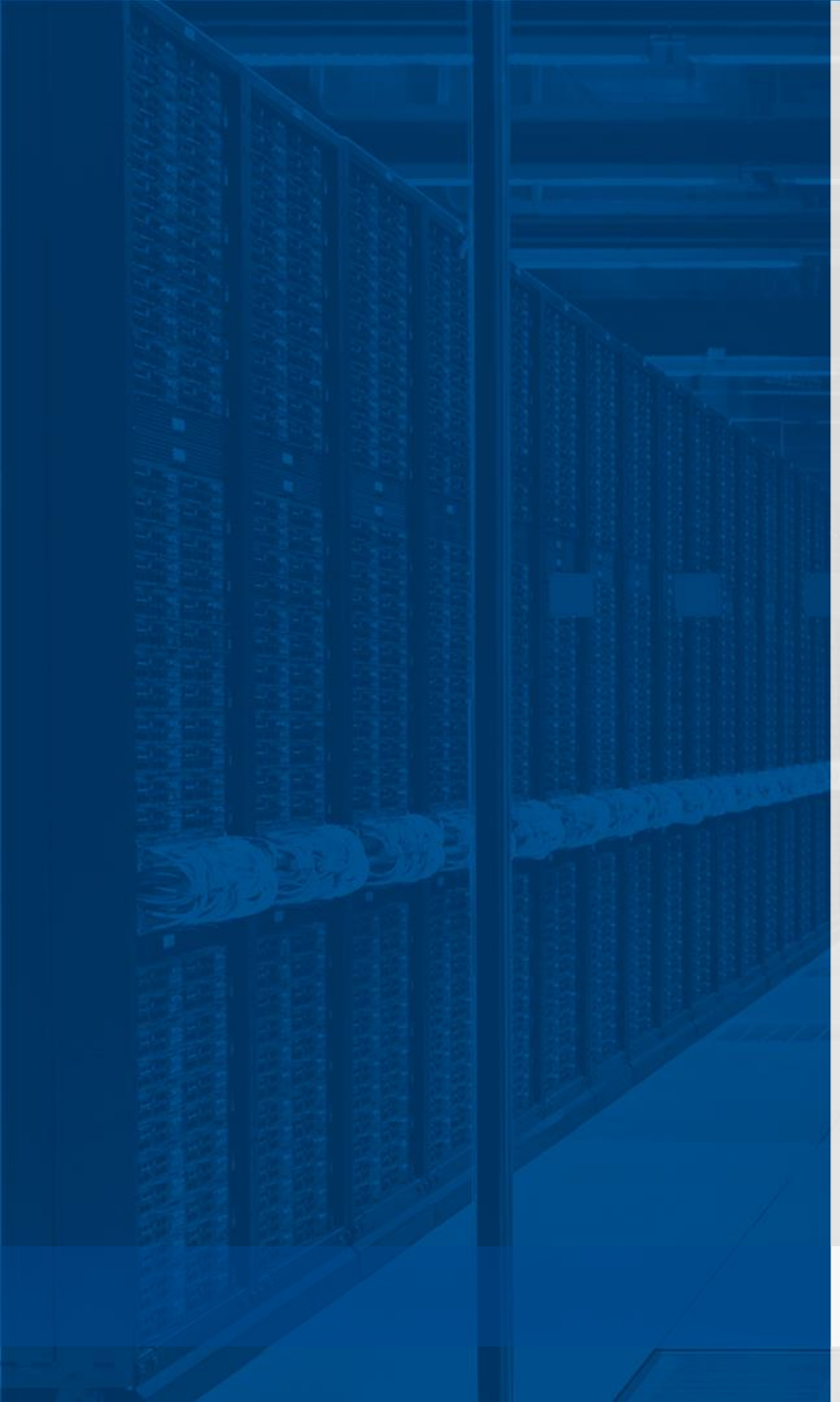
- Spain is not yet part of it
- The Computational Social Science program @BSC could contribute to:
  - FAIRification of social science data
  - Data repository based on Dataverse
  - Integration with other European (and world-wide) repositories, EOSC, and Data Spaces



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Summary

1. **Foundations, Principles, Implementation**
    1. FAIR, Data citation, Reproducibility principles and best practices
    2. Advance with data repositories & journal/funding data policies
2. **Open Data Across Sectors and Domains**
    1. Co-creation research, Industry, Governments, Civil society
    2. Build Interoperability across scientific domains
3. **Science with Computation and AI, and Openness in mind**
    1. Increase use of AI/Ml in science, more computing, more data
    2. Open automated workflows (large data-computation integration)



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Thank you

merce.crosas@bsc.es