About me
http://lorenabarba.com

- SC19 Reproducibility Chair; JupyterCon 2020 General Chair

- NASEM committee "Reproducibility and Replicability in Science"

- NASEM committee "Open Source Software Policy Options for NASA"

- NumFOCUS Board of Directors, 2014–2021

- Founding editor and past AEiC of The Journal of Open Source Software

- Editor-in-Chief of IEEE *Computing in Science and Engineering*
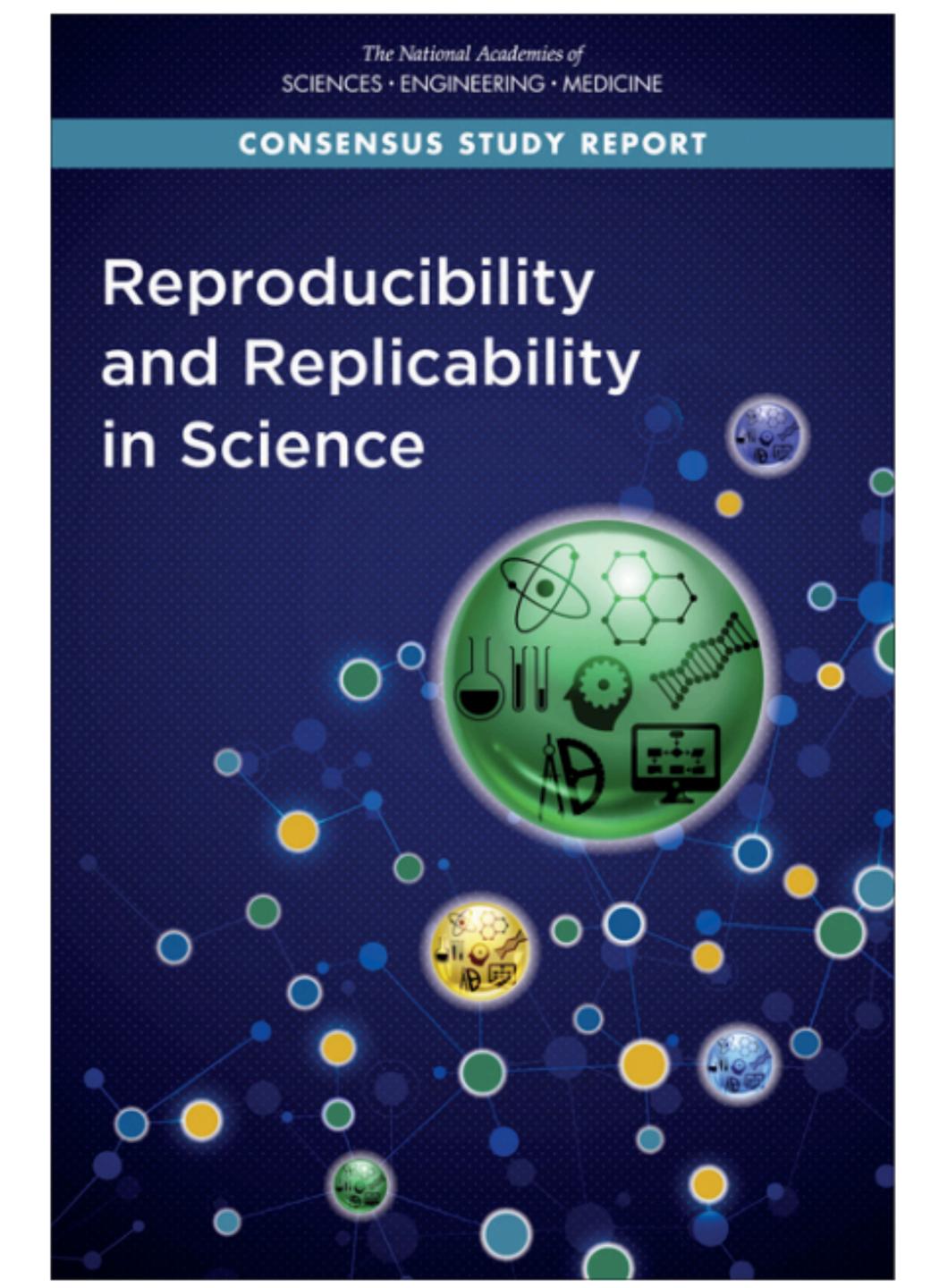
- Author "Reproducibility PI Manifesto"

SC16 panel on Reproducibility and Repeatability

Reproducibility PI Manifesto
2012

- I teach my graduate students about reproducibility

- All our research code (and writing) is under version control

- We always carry out verification & validation (and make them public)

- For main results, we share data, plotting script & figure under CC-BY

- We upload preprint to arXiv at the time of submission to a journal

- We release code at the time of submission of a paper to a journal

- We add a "Reproducibility" declaration at the end of each paper

- I develop a consistent open-science policy & keep an up-to-date web presence

The National Academies of
SCIENCE · ENGINEERING · MEDICINE

CONSENSUS STUDY REPORT

Reproducibility
and Replicability
in Science

- Study mandated by public law 114-329 (Jan. 2017)
- commissioned by NSF to The National Academies of Sciences, Engineering and Medicine (NASEM)
- 15 experts convened
- 18 months of in-person meetings, teleconferences, commissioned papers, deliberations, writing
- report released 7 May 2019

http://doi.org/c5jp

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

CONSENSUS STUDY REPORT

Open Source Software
Policy Options for NASA Earth
and Space Sciences

- ad-hoc committee to investigate and recommend best practices for NASA as it considers whether to establish an open code and open models policy, to complement its open data policy
- Convened Oct. 2017, report delivered Sep. 21, 2018
- "Accelerating the pace of science is a key goal."

https://numfocus.org

NumFOCUS was founded in 2012 to provide a fiscal umbrella for many open-source software projects that have become essential for science and research. Our sponsored projects benefit from a range of services: fiscal, legal, operational, and more.

NUMF[ ]CUS

OPEN CODE = BETTER SCIENCE

ANNUAL REPORT

2019

LORENA BARBA
NumFOCUS Co-Chairperson

https://numfocus.org

84

Sponsored &
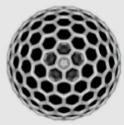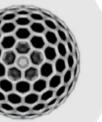Affiliated Projects

$5M

Total revenue in
2020
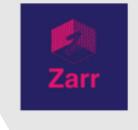
NUMFOCUS
OPEN CODE = BETTER SCIENCE

Top challenges of reproducible research

▸ creation, curation, usage and publication of research software

▸ acceptance, adoption and standardization of **open-science** practices;

▸ misalignment with academic incentive structures and institutional processes for career progression

When our world changed…
January 8, 2020

STAT

HEALTH

WHO says mysterious illness in China likely being caused by new virus

By Helen Branswell Jan. 8, 2020

Reprints

Nucleotide | Nucleotide ▼ | [   ]

Advanced

GenBank ▼   Send to: ▼

Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1, complete genome

GenBank: MN908947.1

⚠️ **This sequence has been updated. See current version.**

FASTA   Graphics

```
LOCUS       MN908947                30473 bp ss-RNA     linear   VRL 12-JAN-2020
DEFINITION  Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1, complete
            genome.
ACCESSION   MN908947
VERSION     MN908947.1
KEYWORDS    .
SOURCE      Wuhan seafood market pneumonia virus
  ORGANISM  Wuhan seafood market pneumonia virus
            Viruses; Riboviria; Nidovirales; Cornidovirineae; Coronaviridae;
            Orthocoronavirinae; Betacoronavirus; unclassified Betacoronavirus.
```
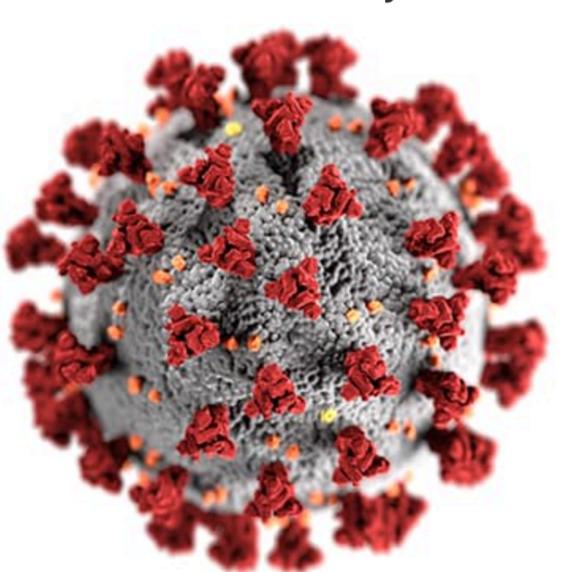
CORONAVIRUS

COVID-19 Datasets

The Lens has assembled free and open datasets of patent documents, scholarly research works metadata and biological sequences from patents, and deposited them in a machine-readable and explorable form.

Human Coronaviruses Data Initiative
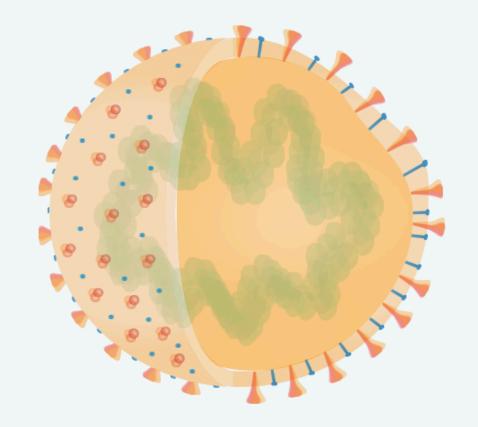
https://about.lens.org/covid-19/

COVID-19 Molecular Structure and Therapeutics Hub

Aggregating critical information to accelerate COVID-19 drug discovery for the molecular modeling and simulation community.

https://covid.molssi.org   https://github.com/MolSSI/covid
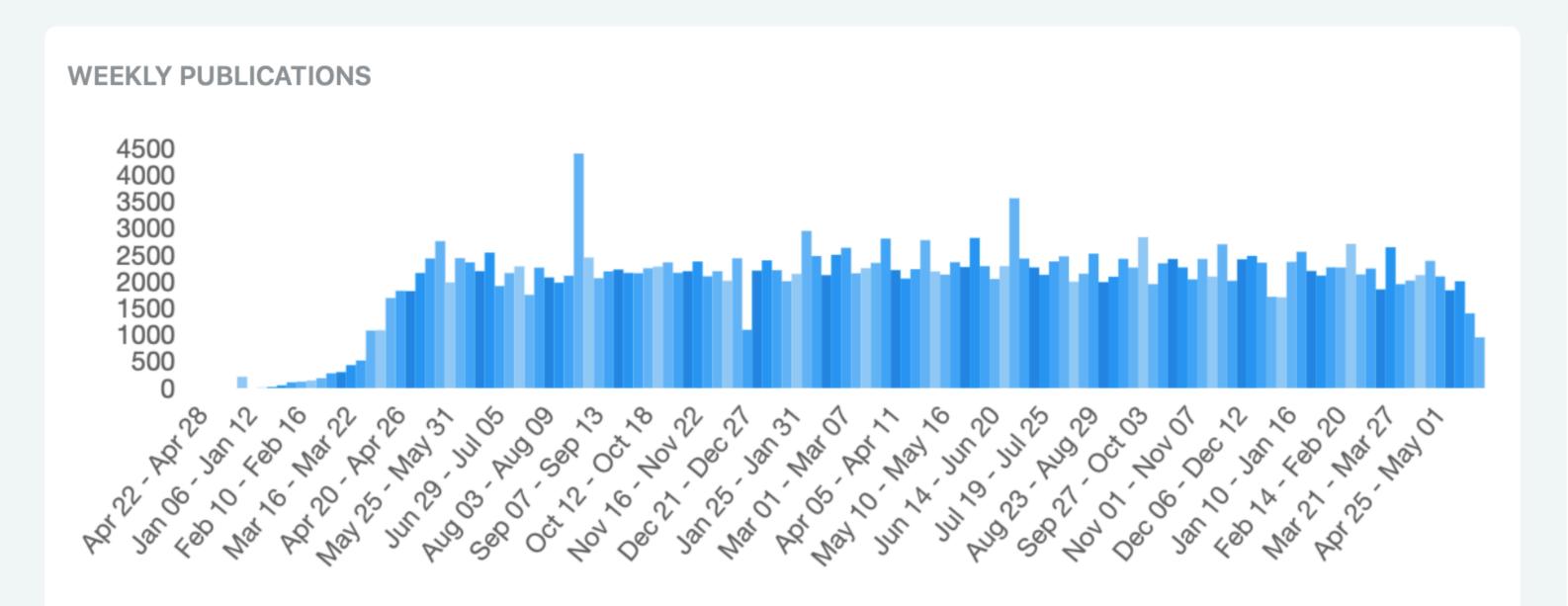
LitCovid is a curated literature hub for tracking up-to-date scientific information about the 2019 novel Coronavirus. It is the most comprehensive resource on the subject, providing a central access to 254031 (and growing) relevant articles in PubMed. The articles are updated daily and are further categorized by different research topics (e.g. transmission) and geographic locations.

CITE   FAQ   DOWNLOAD   LONG COVID

WEEKLY PUBLICATIONS

LATEST PUBLICATIONS

TREATMENT

Three doses of mRNA COVID-19 vaccine protects from SARS-CoV-2 infections in Japan.

Hotta, Katsuyuki et al. • J Intern Med

MECHANISM

Angiotensin-Converting Enzyme 2 and Furin Expression in the Appendix of Children with COVID-19.

https://doi.org/gmh9wh
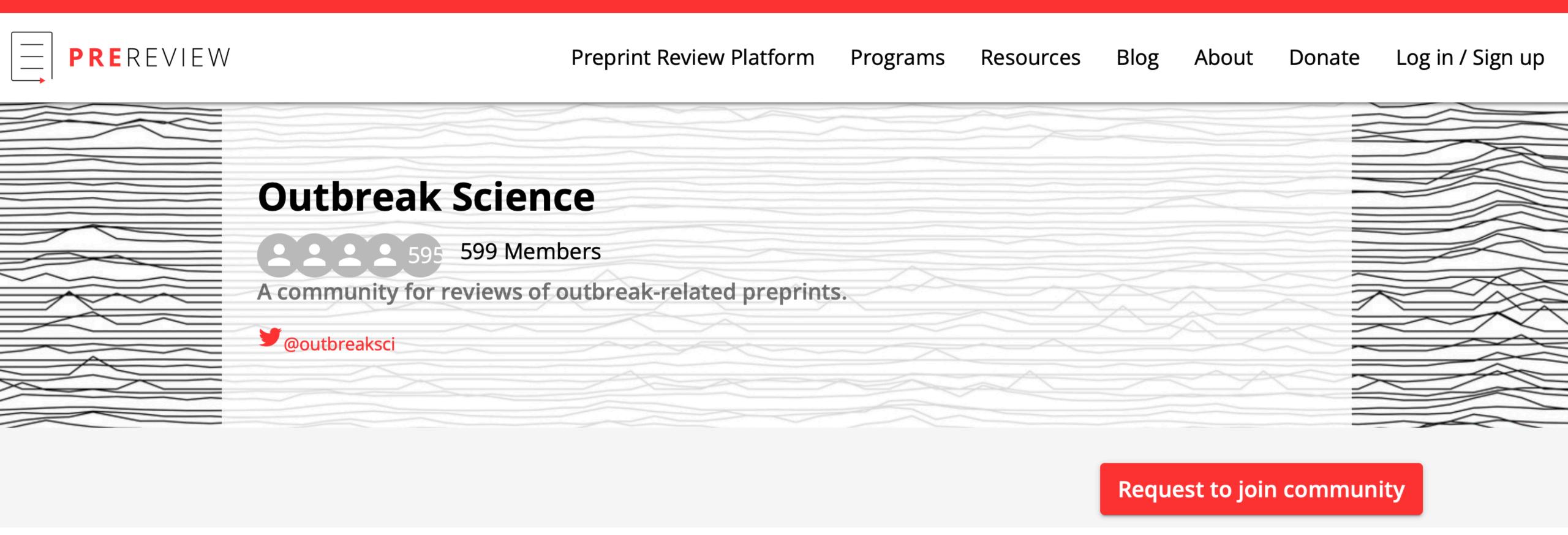
COVID-19 SARS-CoV-2 preprints from medRxiv and bioRxiv

23,592 Articles (17,973 medRxiv, 5,619 bioRxiv)

Most recent first   Page 1: Articles 1-10 | Next ➡

http://connect.medrxiv.org/relate/content/181

PREREVIEW

Preprint Review Platform   Programs   Resources   Blog   About   Donate   Log in / Sign up

Outbreak Science

595   599 Members

A community for reviews of outbreak-related preprints.

🐦 @outbreaksci

Request to join community

https://outbreaksci.prereview.org

"Our work is a direct response to the flawed way scientific research is evaluated. Behind closed doors, a handful of unpaid reviewers —selected opaquely and mainly through personal connections— use subjective criteria to decide the fate of a research article."

PREreview founders: Daniela Saderi, Samantha Hindle

Open Access to ACM Digital Library During Coronavirus Pandemic

March 30, 2020

Dear ACM Members:

As the coronavirus/COVID-19 pandemic continues, we at ACM would like to do what we can to help support the computing community. Many computing researchers and practitioners are now working remotely. In addition, teaching and learning have also moved online as more and more campuses close.

We believe that ACM can help support research, discovery and learning during this time of crisis by opening the ACM Digital Library to all. For the next three months, there will be no fees assessed for accessing or downloading work published by ACM. We hope this will help researchers, practitioners and students maintain access to our publications as well as increasing visibility and awareness of ACM's journals, proceedings and magazines. Please be sure to inform your colleagues that the ACM DL is now open, and will continue that way through June 30, 2020.

This global health crisis is a unique challenge that has impacted many ACM members. We would like to express our concern and support for all who are affected by this outbreak.

Home > About > Press Releases

> Elsevier gives full access to its content on its COVID-19 Information Center for PubMed Central and other public health databases to accelerate …

Elsevier gives full access to its content on its COVID-19 Information Center for PubMed Central and other public health databases to accelerate fight against coronavirus

New York, March 13, 2020

From today, Elsevier, a global leader in research publishing and information analytics specializing in science and health, is making all its research and data content on its COVID-19 Information Center available to PubMed Central, the archive of biomedical and lifescience at the US. National Institutes of Health's National Library of Medicine, and other publicly funded repositories globally, such as the WHO COVID database, for as long as needed while the public health emergency is ongoing. This additional access allows researchers to use artificial intelligence to keep up with the rapidly growing body of literature and identify trends as countries around the world address this global health crisis.

April 15, 2020

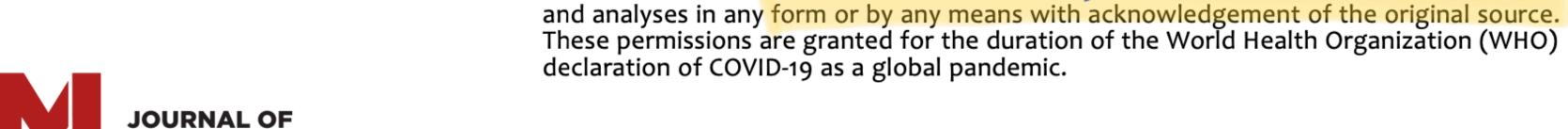# Expanded access to JSTOR during COVID-19 crisis

We are continuing to expand access to JSTOR for students, faculty, and institutions impacted by COVID-19. Access to all these resources will be available through **June 30, 2020**:
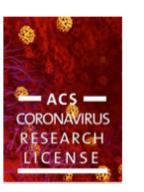
Expanded JSTOR resources for institutions

- **Expanded access to journals and primary sources**: For participating academic institutions that currently license some, but not all, JSTOR Archive and Primary Source collections, we will turn on access to all unlicensed collections at no cost.

- **Expanded access to books:** With support from dozens of publisher partners, more than 35,000 books are available at no charge for JSTOR participating academic institutions.The number of titles available through this effort is growing as more publishers opt in.

During a pandemic crisis

Another exponential spread: open science

JCIM

JOURNAL OF
CHEMICAL INFORMATION
AND MODELING

Viewpoint

A Community Letter Regarding Sharing Biomolecular Simulation Data for COVID-19
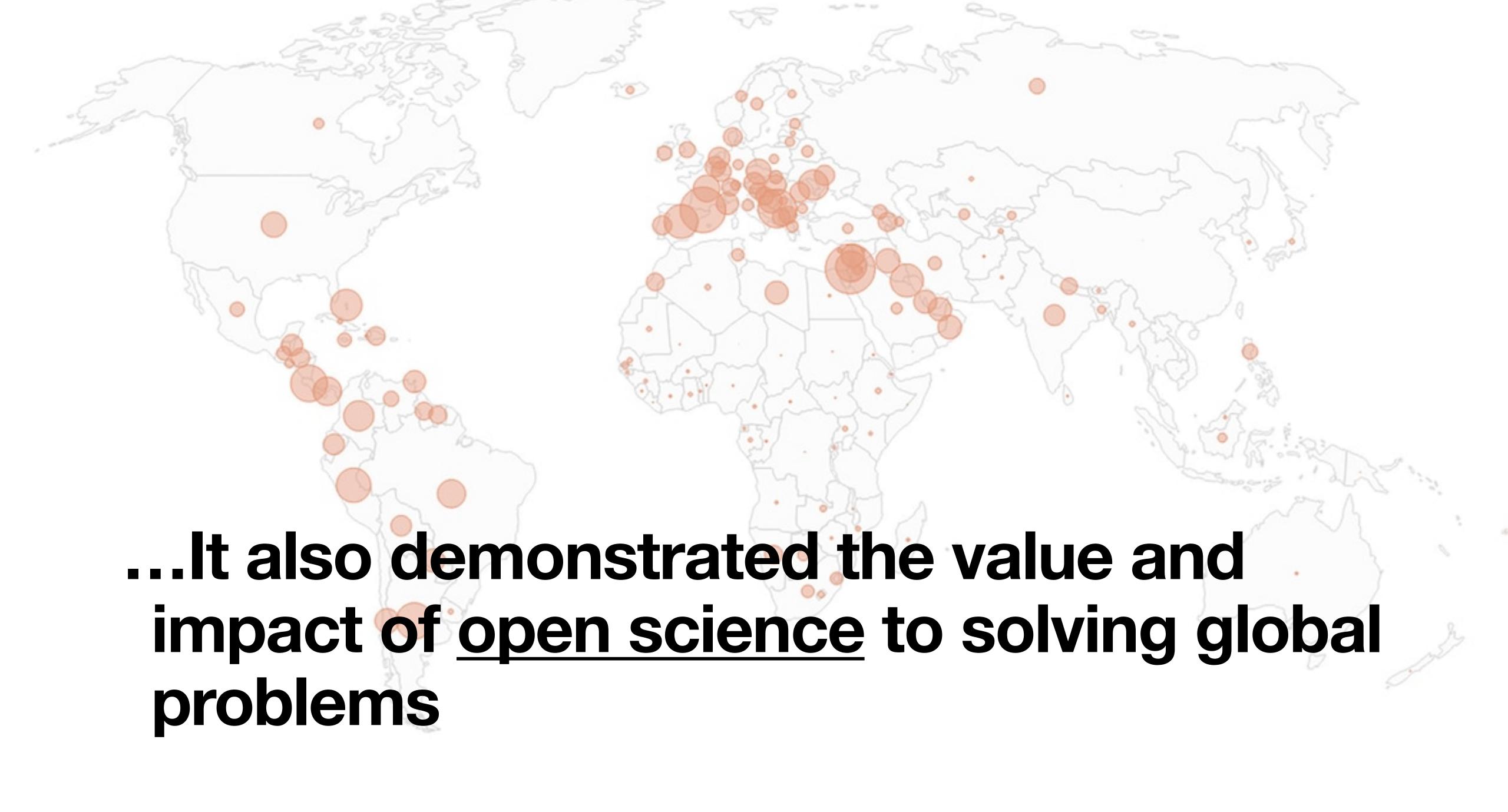
Rommie E. Amaro* and Adrian J. Mulholland*

Read Online

During a pandemic crisis
Another exponential spread: open science

Sharing Biomolecular Simulation Data for COVID-19
Commitments

▸ to making results available quickly via pre-prints: arXiv, bioRxiv, ChemRxiv…

▸ to make available input files, model-building and analysis scripts (e.g., Jupyter notebooks), and data necessary to reproduce the results;

▸ to use open data-sharing platforms to make available results as quickly as possible;

▸ to share algorithms and methods in order to accelerate reuse and innovation;

▸ to apply permissive open-source licensing strategies.

Ahem… my Reproducibility PI Manifesto, 2012

It took a major global health crisis!

…It also demonstrated the value and impact of <u>open science</u> to solving global problems

Global problems
Existential threats

- Global health

- Food security

- Planetary ecosystem sustainability

- Economic & social equity

- Energy & climate

- Water

Open Science

What is it?

Open science "aims to ensure the free availability and usability of scholarly publications, the data that result from scholarly research, and the methodologies, including code or algorithms that were used to generate those data"

NASEM (National Academies of Sciences, Engineering, and Medicine). 2018. *Open Science by Design: Realizing a Vision for 21st Century Research.* https://doi.org/gfxzc4

FIGURE 1-1 Open science alignment.
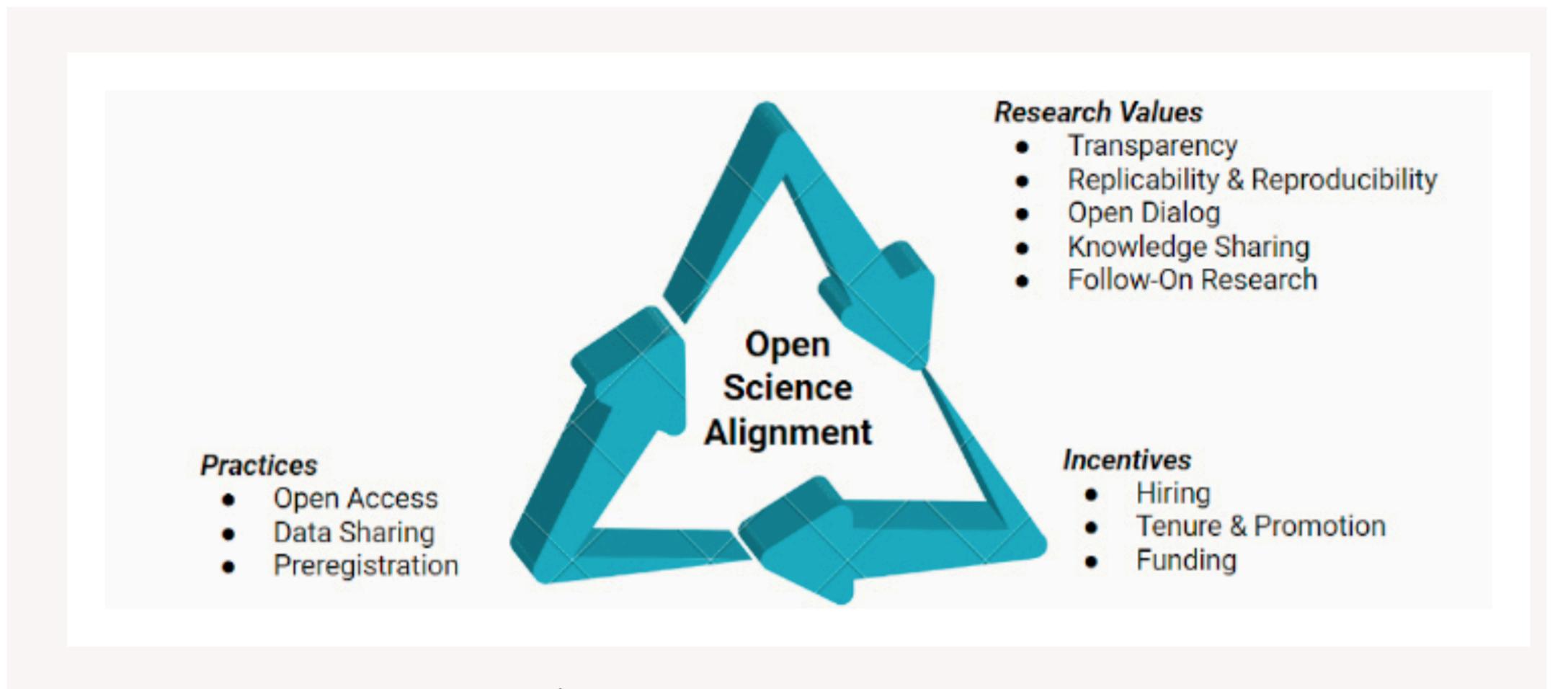
Vision for EU
2016

"Open Science represents a new approach to the scientific process based on cooperative work and new ways of diffusing knowledge by using digital technologies and new collaborative tools."

https://doi.org/gk7tw3

European Open Science Cloud, EOSC
A long-term initiative, launched 2018

- Ecosystem federating research data infrastructure, interlinking FAIR data and interoperable services to support open science

- Trusted digital space for hosting and processing research data

- offers distributed and cloud computing resources for researchers

EUROPEAN OPEN SCIENCE CLOUD

🏠 › Resources

All Resources   **248**

CATEGORIES

Access physical & eInfrastructures	63
Aggregators & Integrators	22
Processing & Analysis	115
Security & Operations	14
Sharing & Discovery	65
Training & Support	22

"Open Science is transparent and accessible knowledge that is shared and developed through collaborative networks"

Vicente-Saez, R. and Martinez-Fuentes, C., 2018. Open Science now: A systematic literature review for an integrated definition. *Journal of Business Research*, 88, pp.428–436. https://doi.org/gc5sjb

Computer Science > Computers and Society

Defining the role of open source software in research reproducibility

Lorena A. Barba

Reproducibility is inseparable from transparency, as sharing data, code and computational environment is a pre-requisite for being able to retrace the steps of producing the research results. Others have made the case that this artifact sharing should adopt appropriate licensing schemes that permit reuse, modification and redistribution. I make a new proposal for the role of open source software, stemming from the lessons it teaches about distributed collaboration and a commitment-based culture. Reviewing the defining features of open source software (licensing, development, communities), I look for explanation of its success from the perspectives of connectivism -- a learning theory for the digital age -- and the language-action framework of Winograd and Flores. I contend that reproducibility engenders trust, which we routinely build in community via conversations, and the practices of open source software help us to learn how to be more effective learning (discovering) together, contributing to the same goal.
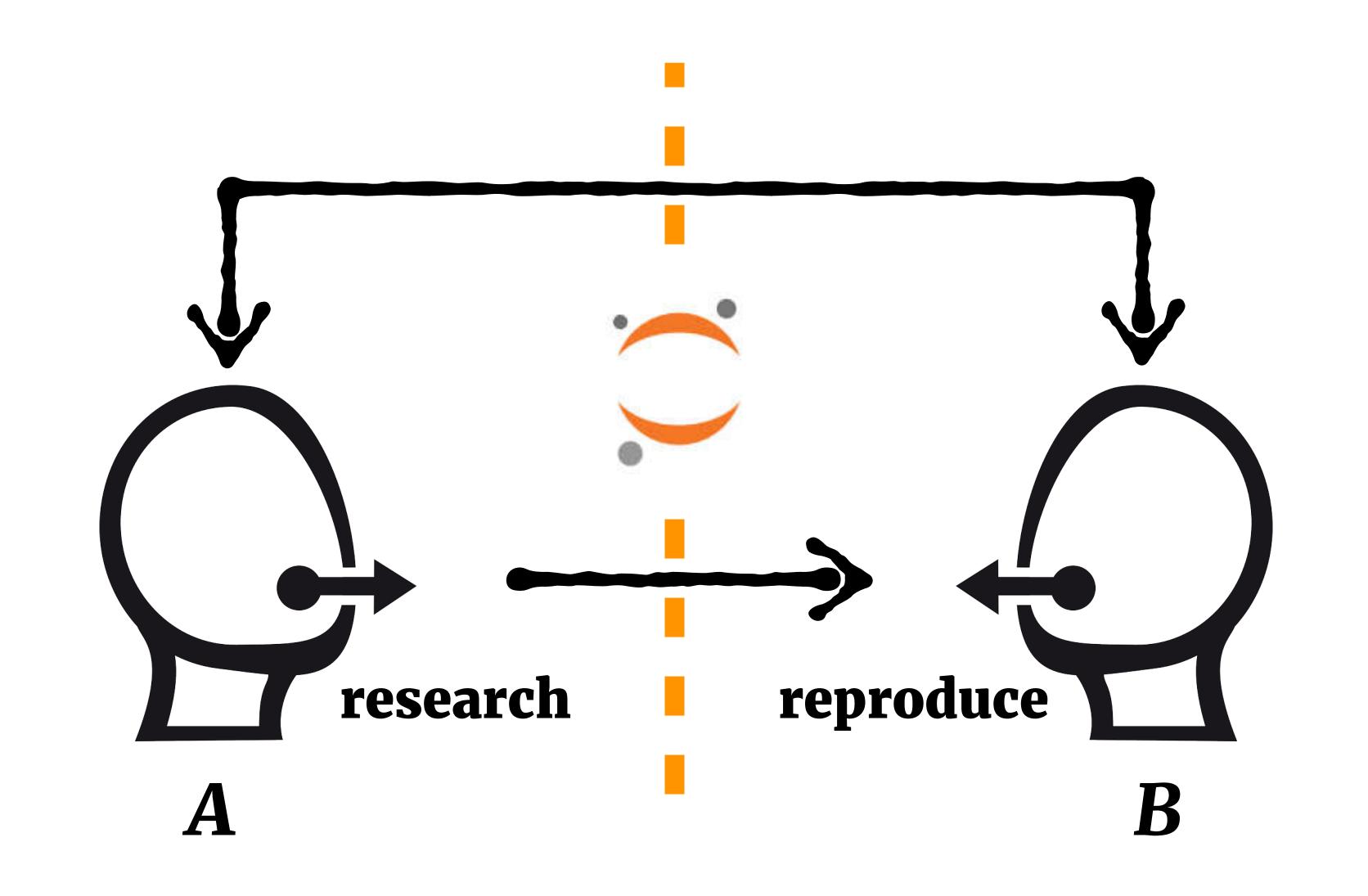
"Science is a conversation"

—Stephen Downes ("connectivism")

‣ a conversation between scientists and their body of knowledge
‣ a conversation among scientists
‣ a conversation between scientists and machines…

Conversation builds trust

research   reproduce

A   B
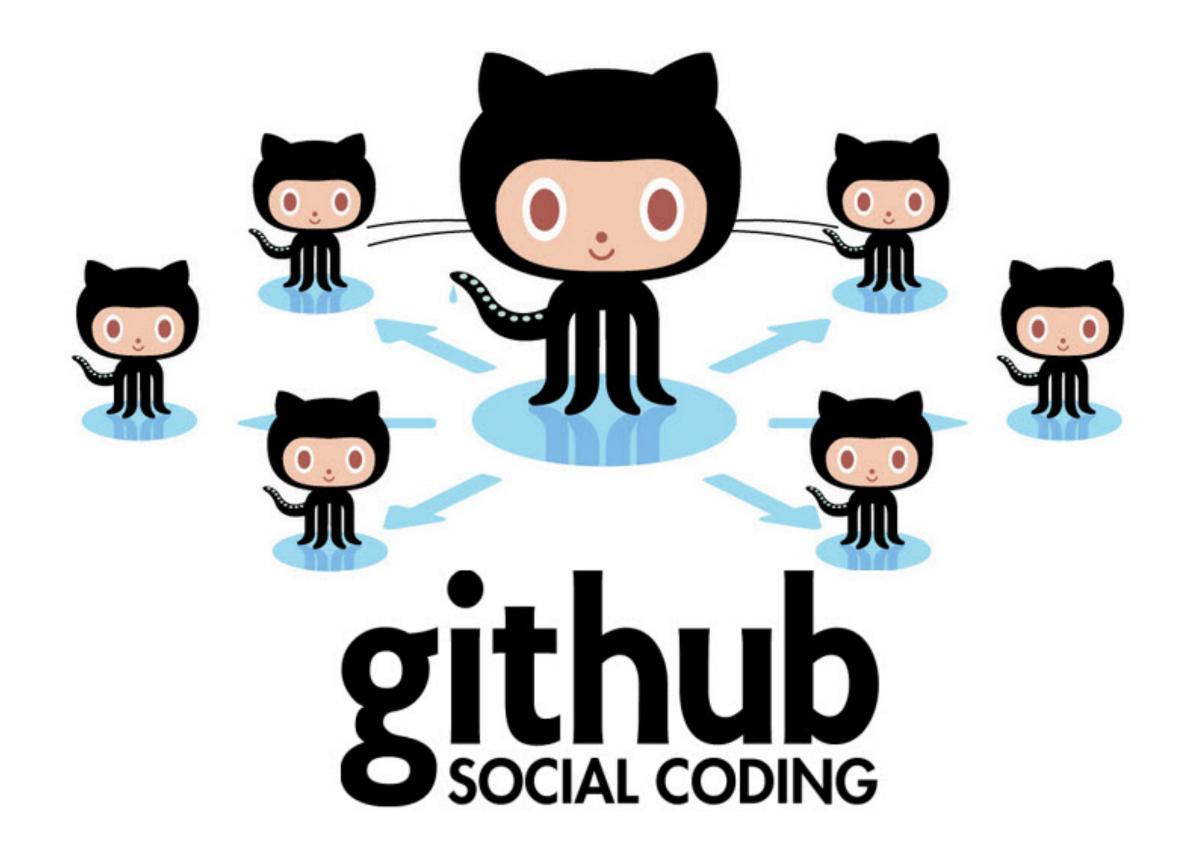
"I have a button here. I push the button. That's not a conversation."
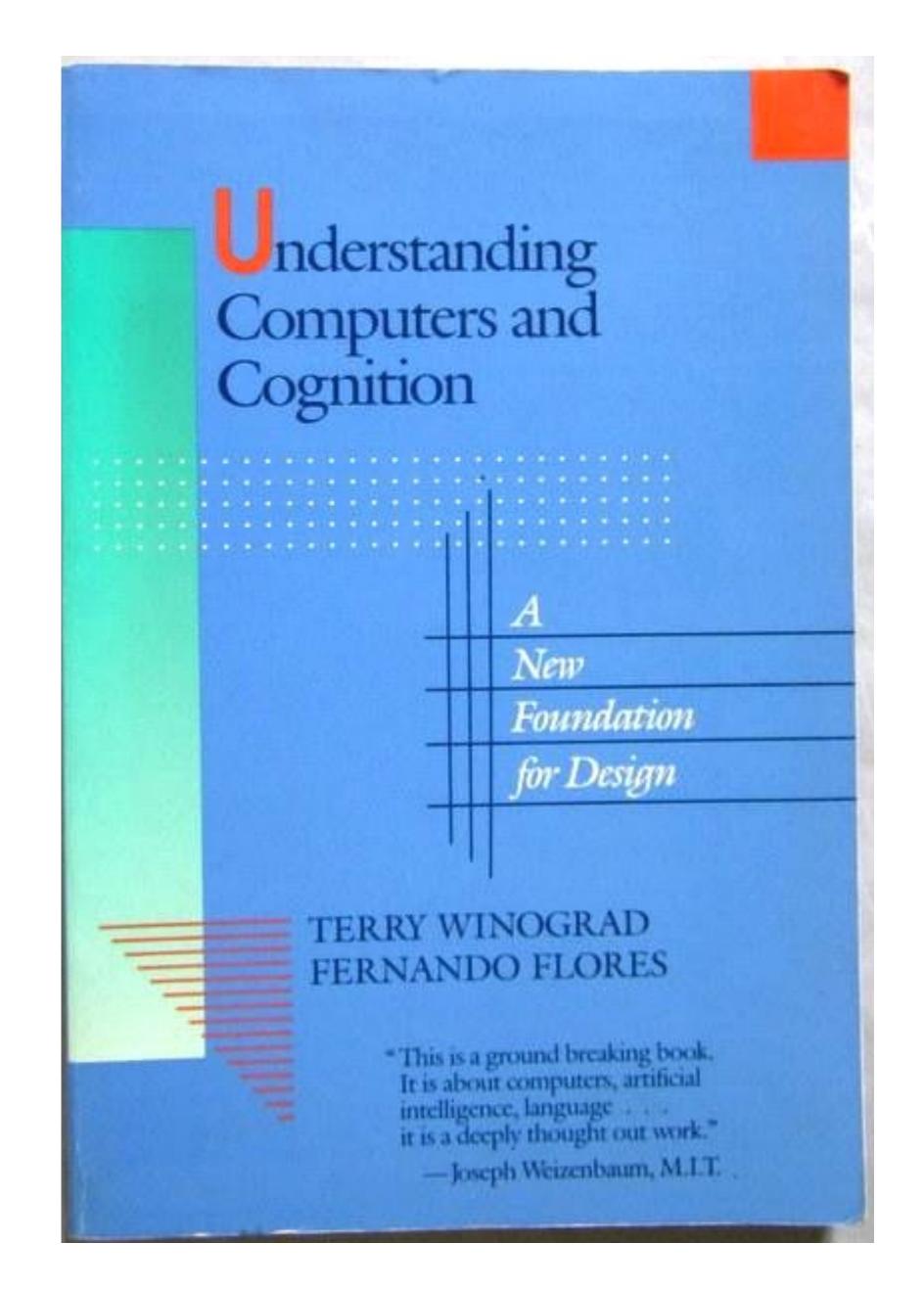
— Paul Pangaro (cybernetics)
Rethinking Design Thinking, PICNIC Festival Amsterdam (2010)

github
SOCIAL CODING

Understanding Computers and Cognition

A New Foundation for Design

TERRY WINOGRAD
FERNANDO FLORES

"This is a ground breaking book. It is about computers, artificial intelligence, language ... it is a deeply thought out work."
— Joseph Weizenbaum, M.I.T.

Winograd & Flores, 1986

NASA Transform to Open Science
Mission: 2022 to 2027

- 2023 is the Year of Open Science

- Deploy a learning platform and develop a core curriculum in open science

- Train 20,000 scientists from the Earth and space science community

- $130 million investment through FY2027

https://science.nasa.gov/open-science/transform-to-open-science
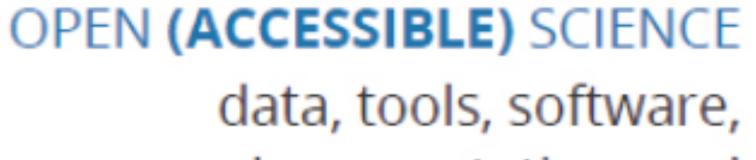
The Four Meanings of "Open" in Open-source Science

OPEN **(TRANSPARENT)** SCIENCE
scientific process and results should be visible, accessible, and understandable

OPEN **(ACCESSIBLE)** SCIENCE
data, tools, software, documentation, and publications should be accessible to all (FAIR)
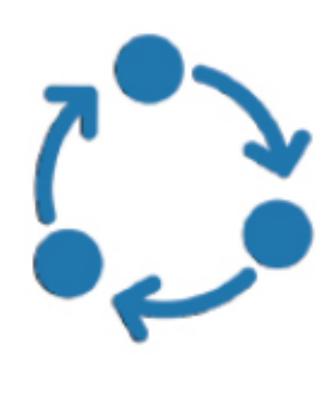
OPEN **(INCLUSIVE)** SCIENCE
process and participants should welcome participation by and collaboration with diverse people and organizations
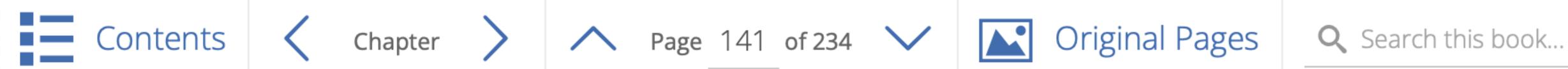
OPEN **(REPRODUCIBLE)** SCIENCE
scientific process and results should be open such that they are reproducible by members of the community

National Academies of Sciences, Engineering, and Medicine. 2022. Open Scholarship Priorities and Next Steps: Proceedings of a Workshop—in Brief. https://doi.org/10.17226/26557

RECOMMENDATION 6-9: Funders should require a thoughtful discussion in grant applications of how uncertainties will be evaluated, along with any relevant issues regarding replicability and computational reproducibility. Funders should introduce review of reproducibility and replicability guidelines and activities into their merit-review criteria, as a low-cost way to enhance both.

National Academies of Sciences, Engineering, and Medicine.
Reproducibility and Replicability in Science, 2019

http://doi.org/c5jp

DESIRABLE CHARACTERISTICS OF DATA REPOSITORIES FOR FEDERALLY FUNDED RESEARCH

Guidance by the
SUBCOMMITTEE ON OPEN SCIENCE

of the
NATIONAL SCIENCE AND TECHNOLOGY COUNCIL

May 2022

https://doi.org/hwrm

Table 1. Desirable Characteristics of Repositories for Managing and Sharing Data Resulting from Federally Funded or Supported Research

	Free and Easy Access	The repository provides broad, equitable, and maximally open access to datasets and their metadata free of charge in a timely manner after submission, consistent with legal and policy requirements related to maintaining privacy and confidentiality, Tribal and national data sovereignty, and protection of sensitive data.
	Broad and Measured Reuse	The repository ensures datasets are accompanied by metadata that describe terms of reuse and provides the ability to measure attribution, citation, and reuse of data (e.g., through assignment of adequate and openly accessible metadata and unique PIDs).

Page 126

RECOMMENDATION 6-5: In order to facilitate the transparent sharing and availability of digital artifacts, such as data and code, for its studies, the National Science Foundation (NSF) should

- **develop a set of criteria for trusted open repositories to be used by the scientific community for objects of the scholarly record;**
- **seek to harmonize with other funding agencies the repository criteria and data management plans for scholarly objects;**

National Academies of Sciences, Engineering, and Medicine.
Reproducibility and Replicability in Science, 2019

http://doi.org/c5jp

"Scientific process and results should be visible, accessible, and understandable."

NASA TOPS open science definition

SC19 Reproducibility Chair

- Artifact Description appendix required
  - Standard form asks about software, data, or other digital artifacts
- Reviewed, innovative double-open model
- Reproducibility Challenge

Open peer review of Appendices
Artifact Description and Article Evaluation

- Constructive: conversation with authors typical

- Artifact availability: many authors provide URLs to GitHub or lab website

- GitHub does not provide guarantees of persistence

- Researchers need advice and technical support!

Open peer review
A cultural change

- peer review can suffer from implicit bias that disadvantages women, minorities and people from less-prestigious institutions

- educate reviewers about bias in the review process, and equip them with tools to interrupt this bias

- Double-blind peer review does not help build awareness of bias, is clunky, and easy to "guess" author identity

- Example of open review: The Journal of Open Source Software

Closed ⊘ **40 tasks done** [REVIEW]: ExaFMM: a high-performance fast multipole method library with C++ and Python interfaces #3145

whedon opened this issue on Apr 3, 2021 · 52 comments

tingyu66 commented on Apr 22, 2021 · edited ▾

> On `Ubuntu 18.04.5 LTS (GNU/Linux 4.15.0-118-generic x86_64)` the `./configure`
> `CXX=g++` stops with
>
> ```
> checking for fftw3.h... yes
> checking for fftwf_malloc in user specified location... no
> checking for fftw_malloc in user specified location... no
> configure: error: cannot find fftwf library.
> ```

Exafmm-t requires the FFTW3 shared library in both single and double precision (`libfftw3f.so` and `libfftw3.so`), so we suggest to install `libfftw3-dev` package, as described in our documentation:

```
apt-get update
apt-get -y install libfftw3-dev
```

If you want to build FFTW3 library from source, you probably need to compile separately for each precision if I remember correctly. Then add the directory of the shared libraries to `LIBRARY_PATH` and `LD_LIBRARY_PATH` environment variables.

pitsianis commented on May 16, 2021

I am done with review. I like the Jupyter notebook.

☺ 👍 1

https://joss.theoj.org

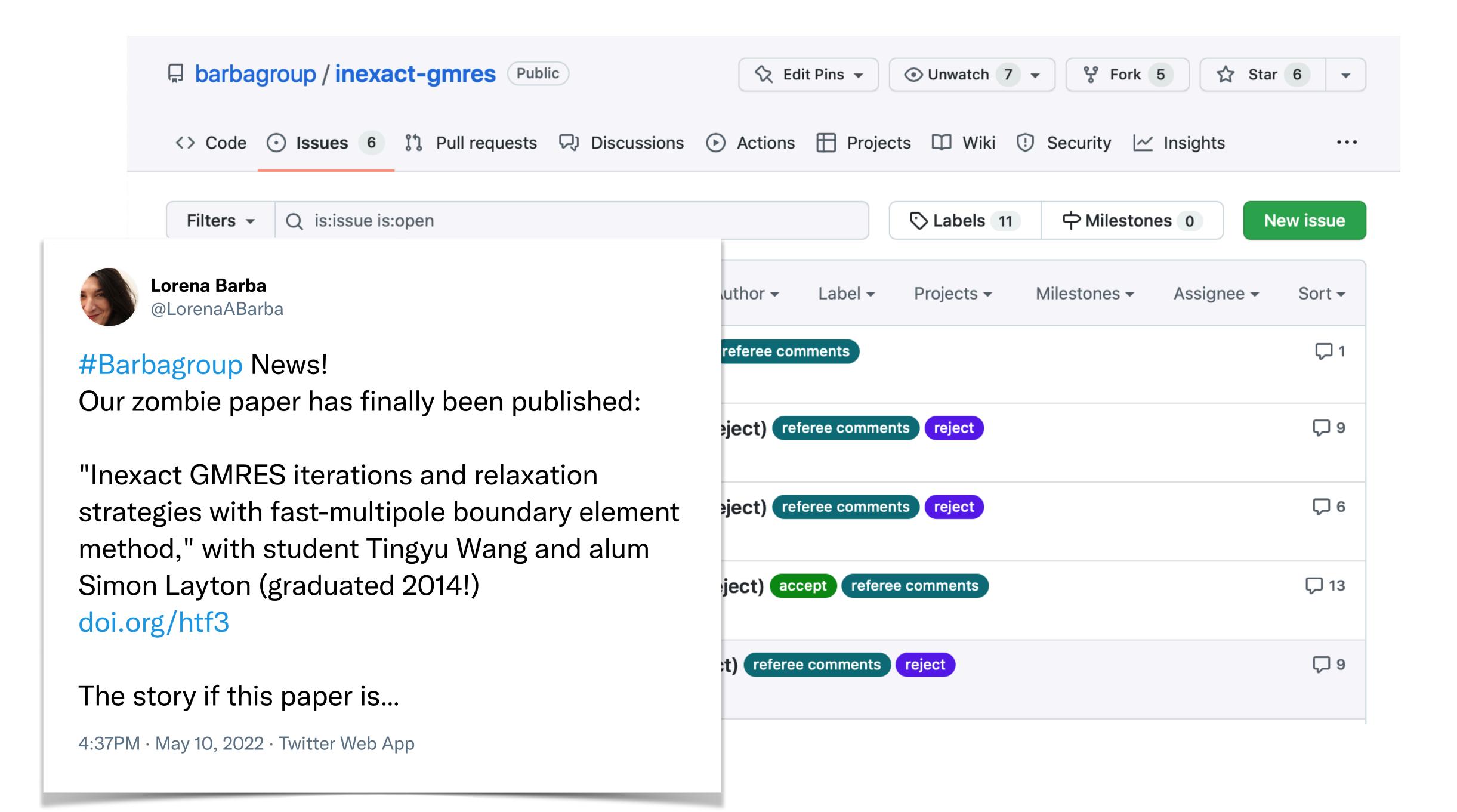Article | Open Access | Published: 18 January 2019

The effect of publishing peer review reports on referee behavior in five scholarly journals

Giangiacomo Bravo, Francisco Grimaldo, Emilia López-Iñesta, Bahar Mehmani & Flaminio Squazzoni ✉

https://doi.org/gfvpwd

barbagroup / **inexact-gmres** — Public

Edit Pins | Unwatch 7 | Fork 5 | Star 6

<> Code | ⊙ Issues 6 | ⋔ Pull requests | 💬 Discussions | ▶ Actions | ⊞ Projects | 📖 Wiki | ⊘ Security | 📈 Insights | ⋯

Filters | is:issue is:open | 🏷 Labels 11 | 🚩 Milestones 0 | New issue

uthor ▾ | Label ▾ | Projects ▾ | Milestones ▾ | Assignee ▾ | Sort ▾

referee comments | 💬 1

eject) referee comments reject | 💬 9

eject) referee comments reject | 💬 6

eject) accept referee comments | 💬 13

ct) referee comments reject | 💬 9

Lorena Barba
@LorenaABarba

#Barbagroup News!
Our zombie paper has finally been published:

"Inexact GMRES iterations and relaxation strategies with fast-multipole boundary element method," with student Tingyu Wang and alum Simon Layton (graduated 2014!)
doi.org/htf3

The story if this paper is...

4:37PM · May 10, 2022 · Twitter Web App

"It is time to improve the quality, transparency, and accountability of the peer review system."

Haffar, S., Bazerbachi, F. and Murad, M.H., 2019, April. Peer review bias: a critical review. In Mayo Clinic Proceedings (Vol. 94, No. 4, pp. 670-676). https://doi.org/gnjhwk