

# The Essential Future of Open Data and How Publishers Must Help



Brooks Hanson  
Executive VP, Science  
AGU

[bhanson@agu.org](mailto:bhanson@agu.org)

Following Stall et al., *Nature* 5 June 2019

# Fact 1: The 21<sup>st</sup> century (and beyond) will be the century of convergent science for the benefit of humanity

1 habitable planet

All major sustainability issues are core to convergent/transdisciplinary science:

- Population 7-10 billion; food?
- Major health issues are convergent science: include clean air, soil, and water, geohealth, etc.
- All ecosystems and biogeochemical cycles are human dominated
- Huge climate change experiment underway
- Key limited resources
- Imminent energy transition
- Transportation, navigation, trade (e.g., economy) all science based
- Natural hazards, weather, and climate forecasting

All require massive data and integration with modeling and across science,

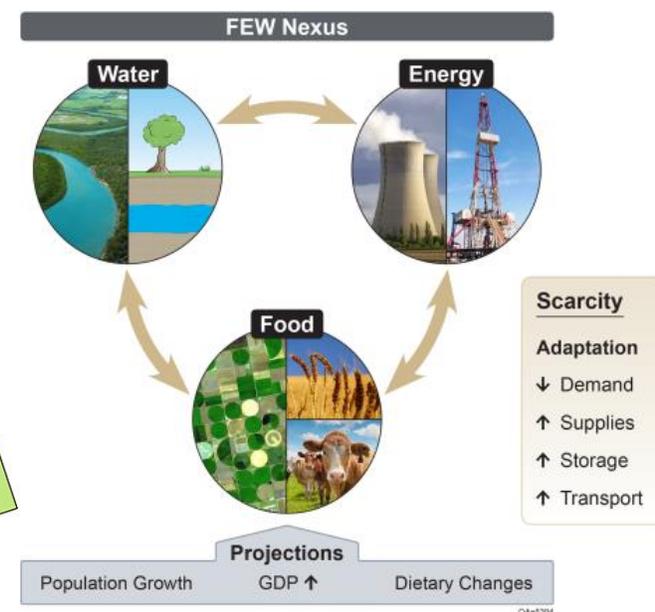
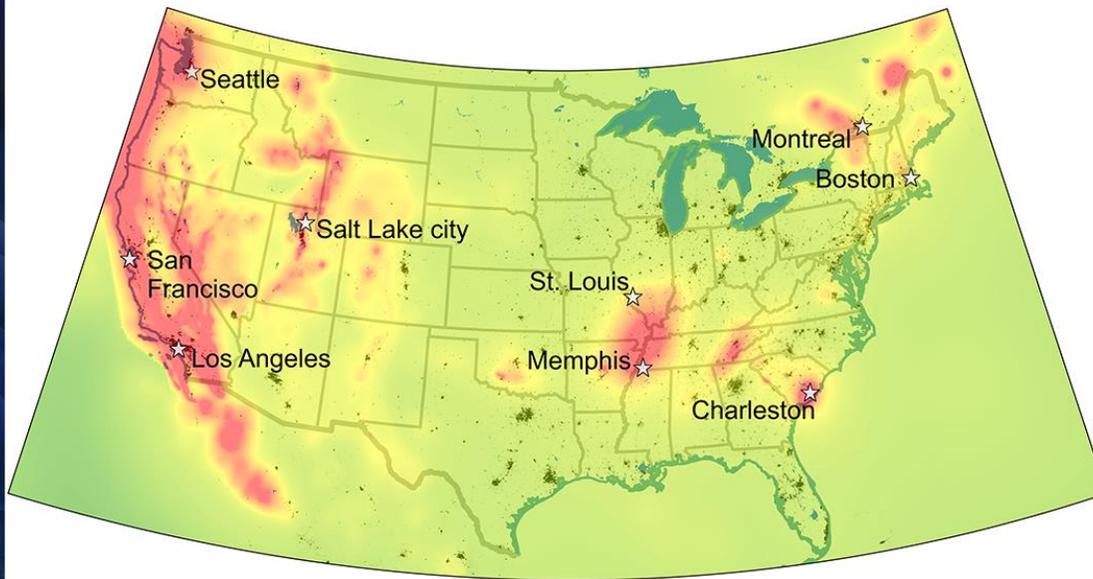
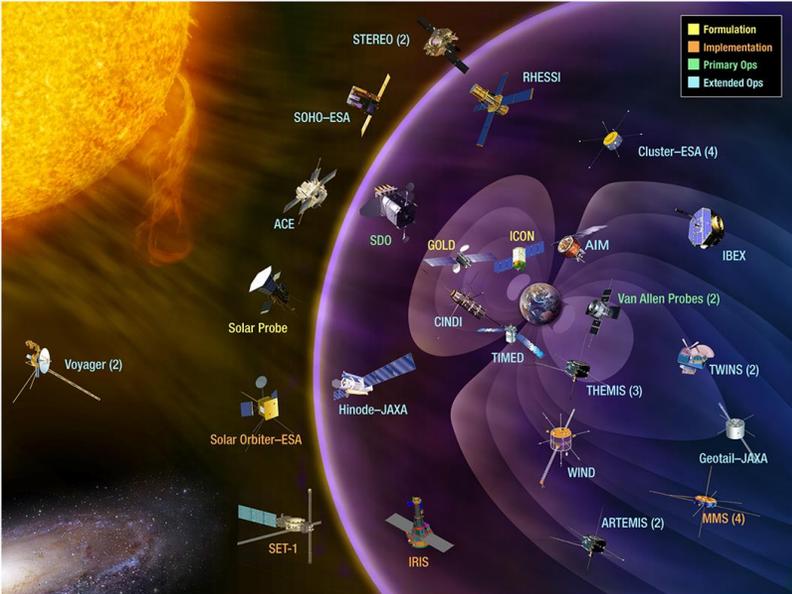


## Fact 2 (Follows Fact 1): Integrated, interoperable data across the sciences are essential to address these challenges

“We've arranged a global civilization in which most crucial elements profoundly depend on science and technology. We have also arranged things so that almost no one understands science and technology. This is a prescription for disaster. We might get away with it for a while, but sooner or later this combustible mixture of ignorance and power is going to blow up in our faces.” (1995)

We must prove Carl wrong....Data are the key.





## Earth and Space Science is Essential for Society

**Special AGU Issue, April 2017** Open Access; see link at <http://publications.agu.org>

27 Commentaries across AGU journals illuminate the deep and growing benefits of research in the Earth and space sciences for humanity.

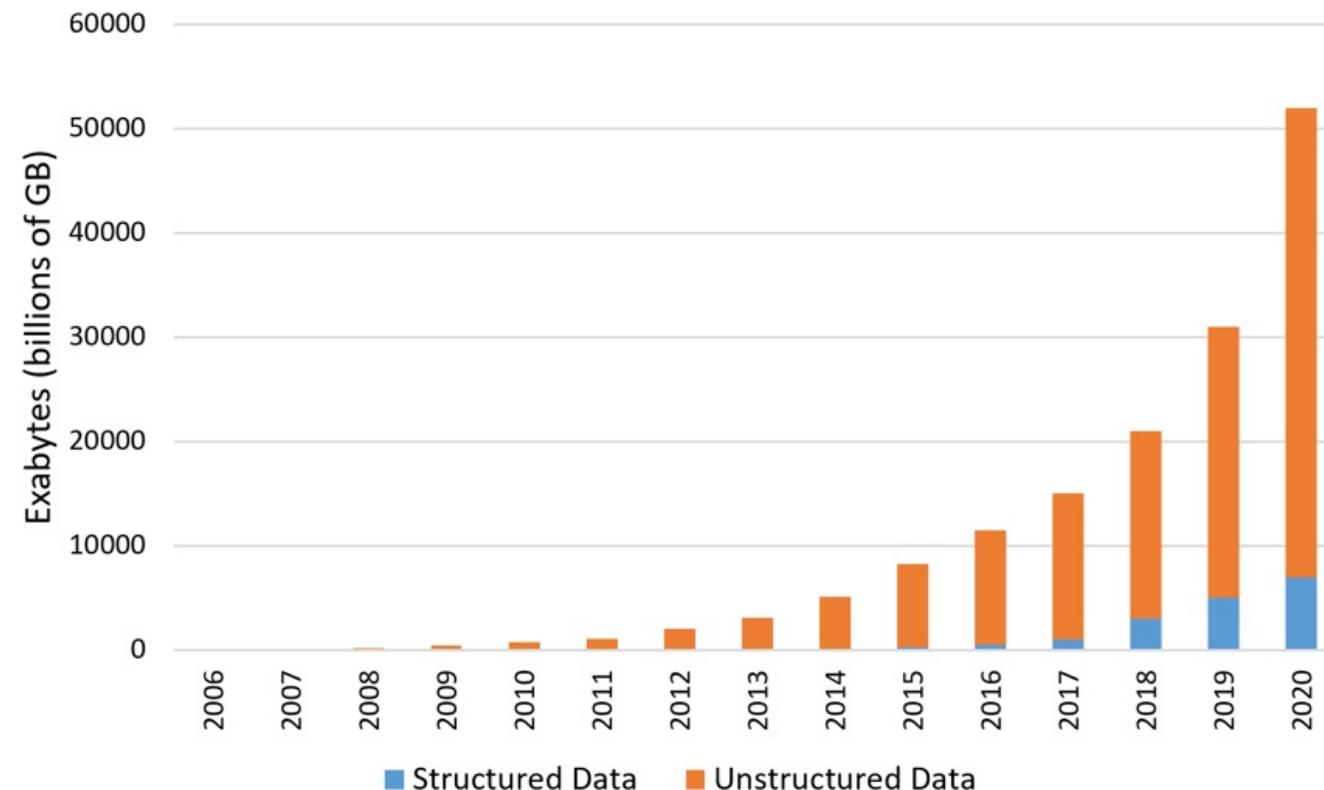
- Fact 3: All parts of ESS contribute science that is critically important to society
  - Fact 4: International shared data are critical
- Fact 5: Future depends on continued international growth of ESS and connections across other sciences

## Fact 6: We are at a tipping point for scientific data (or slightly beyond it)

- Sensor explosion starting
- Satellite and instrument explosion already here
- Citizen engagement expanding
- Critical need for data management

Digital Data Storage is Undergoing Mind-Boggling Growth  
*EE Times*, Sept 2016

### The Cambrian Explosion...of Data



# Therefore:

## Data science IS science. It is core, not an extra...

- Data are not supplemental and shouldn't be treated as supplements
- Data Management Plans are backwards
- Data repositories should be “Science Centers of actionable information and knowledge”

## Enhancing and ensuring the integrity and transparency of research is critically part of our contract for support with society—it's essential for humanity.

- This means FAIR data
- Dan Gilmore: “Transparency isn't just for journalists: Any organization that wants to be trusted should use it...”

## We are thinking about and discussing “impact” way too narrowly



# We have and know the key pieces:

- Many reports/studies (NAS, Horizon 2020, other national studies)
- FAIR data initiative, FORCE11 Guides, ORCID, CrossRef, DataCite, RDA, Make Data Count, CRediT...
- Publisher best practices; TOP Guidelines
- Funder best practices
- Repository best practices
- Discussions on research assessment--catch up with how science is done now (teams) and needs to be done, not was done 20 years ago
- Many more



# Publishers Play a Critical Role

Responsibility for Integrity, Quality, Value, Communication to the Public, Standards, Ethics...

Publications are much more than the PDF/Html file

- All the connected resources and information
- References, Data, Software
- Define how science is done in the 21<sup>st</sup> century



# AGU's position statement on data affirms that

**“Earth and space sciences data are a world heritage. Properly documented, credited, and preserved, they will help future scientists understand the Earth, planetary, and heliophysics systems.”**



Initial version  
1997  
Current version  
2015  
Update scheduled  
2019



# ENABLING FAIR DATA PROJECT

[HOME](#) / ENABLING FAIR DATA PROJECT

OVERVIEW

LEADERSHIP

COMMITMENT STATEMENT

AUTHOR GUIDELINES

RESOURCES

The [Laura and John Arnold Foundation](#) has awarded a grant to a coalition of groups representing the international Earth and space science community, convened by the [American Geophysical Union \(AGU\)](#), to develop standards that will connect researchers, publishers, and data repositories in the Earth, space, and environmental sciences to enable [FAIR](#) (findable, accessible, interoperable, and reusable) data on a large scale. **This project will accelerate scientific discovery and enhance the integrity, transparency, and reproducibility of this data.**

## Commitment to Enabling FAIR Data in the Earth, Space, and Environmental Sciences

[READ MORE](#)

Publication of scholarly articles in the Earth, space, and environmental sciences community is conditional upon the concurrent

# Publisher Signatories Include...

- American Geophysical Union
- Copernicus Publications
- Ubiquity Press
- California Digital Library – CDL
- Wiley
- PANGAEA, Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research (AWI), Center for Marine Environmental Sciences, University of Bremen (MARUM)
- WDC Climate, Deutsches Klimarechenzentrum (DKRZ)
- *Science*
- *Science Advances*
- PLOS
- Elsevier
- F1000
- *Nature*
- *Scientific Data*
- Taylor & Francis Group
- Hindawi Ltd.

Over 100 signatories to date!

# Author, Reviewer, Editor – Open and FAIR – New Data Guidelines – One-Page Reference



1. Summary of objectives for open and FAIR data
2. Resources for Editors, Reviewers, and Authors
3. Information and description of the tools for editors, reviewers, and researchers
4. Data Citation Training, Guidelines, and Examples
5. Software Citation Guidelines and Examples
6. Enabling FAIR Data Project - Overview
7. FAIR Guiding Principles

<http://bit.ly/FAIROnePage>

ADVERTISEMENT



COLLABORATE. INNOVATE.  
COMPETE. TALK TO US.

FIND OUT MORE

---

# Making Magic Happen: Implementing and Contributing Data Citations in Support of Today's Scholarship

By **SCHOLARLY KITCHEN** | DEC 12, 2018 | **9 COMMENTS**

# SCIENTIFIC DATA

OPEN

## A data citation roadmap for scientific publishers

Helena Cousijn<sup>1,†,\*</sup>, Amye Kenall<sup>2,\*</sup>, Emma Ganley<sup>3</sup>, Melissa Harrison<sup>4</sup>, David Kernohan<sup>5</sup>, Thomas Lemberger<sup>6</sup>, Fiona Murphy<sup>7</sup>, Patrick Polischuk<sup>8</sup>, Simone Taylor<sup>9</sup>, Maryann Martone<sup>10</sup> & Tim Clark<sup>11,12</sup>

Received: 18 September 2018

Accepted: 4 October 2018

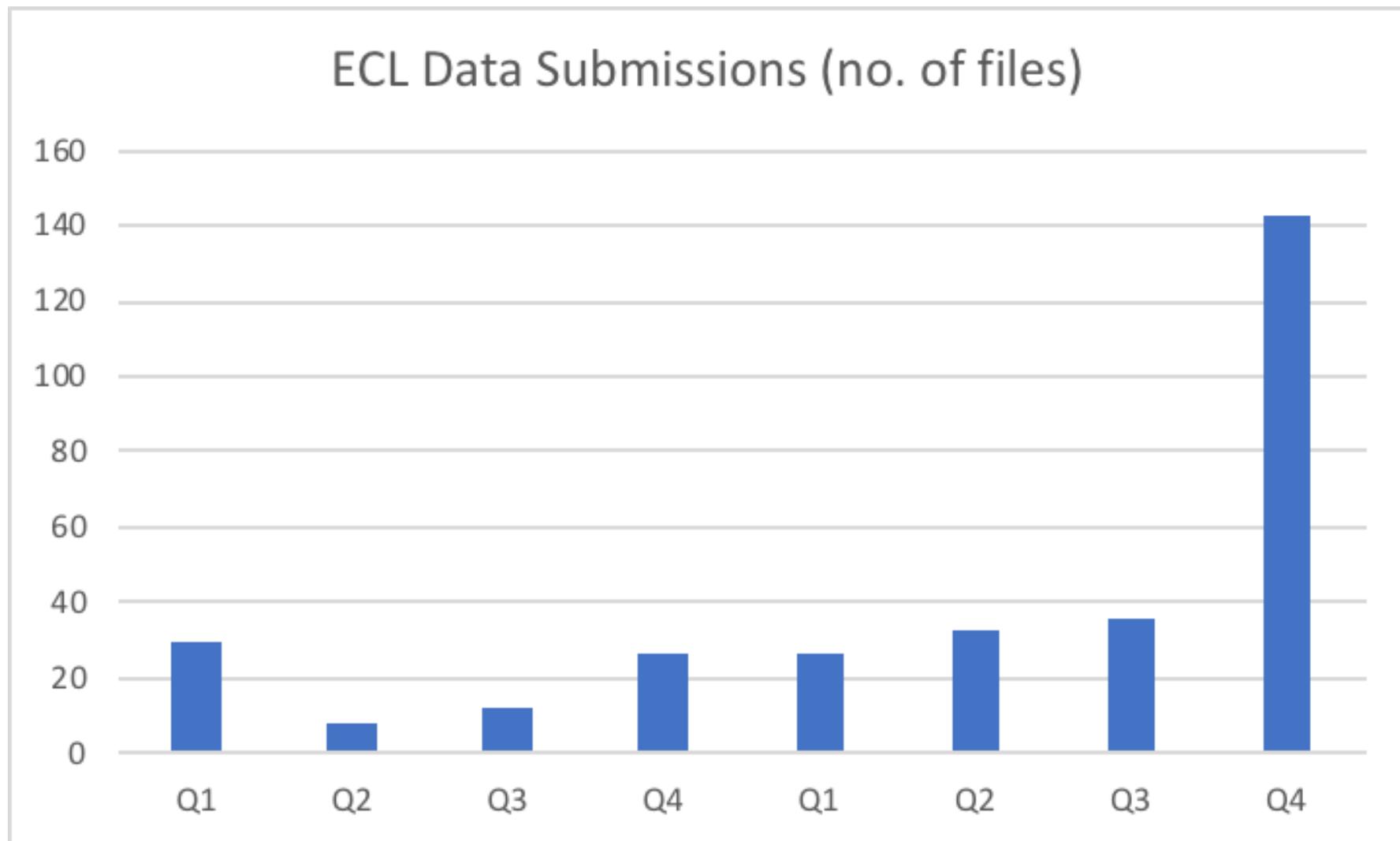
Published: 20 November 2018

This article presents a practical roadmap for scholarly publishers to implement data citation in accordance with the Joint Declaration of Data Citation Principles (JDDCP), a synopsis and harmonization of the recommendations of major science policy bodies. It was developed by the Publishers Early Adopters Expert Group as part of the Data Citation Implementation Pilot (DCIP) project, an initiative of FORCE11.org and the NIH BioCADDIE program. The structure of the roadmap presented here follows the “life of a paper” workflow and includes the categories Pre-submission, Submission, Production, and Publication. The roadmap is intended to be publisher-agnostic so that all publishers can use this as a starting point when implementing JDDCP-compliant data citation. Authors reading this roadmap will also better know what to expect from publishers and how to enable their own data citations to gain maximum impact, as well as complying with what will become increasingly common funder mandates on data transparency.

# Leading Practices for Publishers

1. Use Identifiers: DOI for references/datasets, ORCID and CRediT for Authors, Institutional identifiers (in progress), Funder Registry
2. Use repositories for data and software; work with domain repositories in your field
3. Place data citations in references, and ideally in metadata
4. Tag data citations correctly (so they can be identified by Crossref and linked)
5. Open up your references; include references from any supporting information
6. Talk to and explain process to and help researchers
  - Lots of excuses not to act
  - Lots of good actors too
    - They need help and support
7. Raising new challenges also...

## Submissions in last two years to EarthChem Repository



# A Real Story from Just a Few Days Ago:

Dear Editors,

I have read with great interest the paper entitled "...and have also fully replicated the uncertainty analysis by using the algorithm available at <https://...> I noticed a typo in the line 371 of the file "...", where the ...anomaly is normalized [incorrectly]. I sent an email to the correspondent author but so far he did not answer me .

The 371 line of the file "" reads: ...where it should read ... Note the change in the denominator.

The above correction affects almost all figures; see attached document.

I am looking forward to hearing from you.

**Reply:** We can confirm that the coding error discovered by Dr. ... was present in the code used to generate the results presented in .... We have corrected the error and provide a summary of the resulting errata in the text and figures of the paper. In addition, we will be correcting the codebase and output on the public .... website and providing full documentation of the changes to the results.



# We need a large, coordinated effort across science stakeholders to catalyze broad culture change and assemble these pieces:

- Societies, Publishers, NGO's, Academies, Funders, Science-partners (everyone) are key, and we need common, collective voice and urgency.
- Need to implement support at all levels and TOGETHER.
- Implement demonstrable collective actions (awards, support, practices, policies, messages, credit..)
- Broad engagement to support infrastructure and culture change internationally
- Talk about societal impacts, not personal ones



We need to go BIG because we have to live here



not here