



**NOAA**  
**SCIENCE**  
**ADVISORY**  
**BOARD**

# **NOAA SAB WORKING GROUP ON OPEN DATA & OPEN SCIENCE**

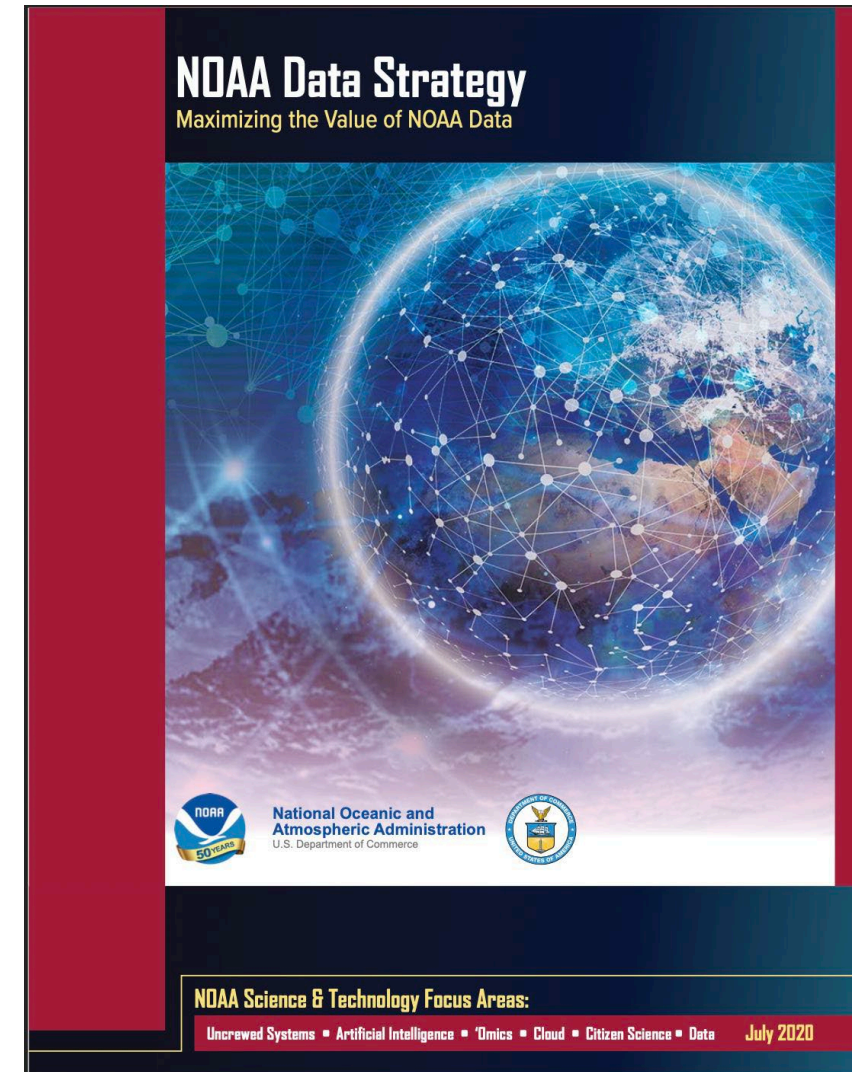
**Dec 1, 2022**

Working Group Members: Chelle Gentemann (Co-chair),  
Robert L. Grossman (Co-Chair), Jason Hickey, Ilene  
Carpenter, Chris Lendardt, Zhaoxia Pu, and Anthony Wu



# Open Data & Open Science

- NOAA has a well-defined Data Strategy and Cloud Strategy that clearly identifies goals, objectives and mechanisms to support open data, and NOAA has made substantial progress implementing these strategies.
- Our process:
  - 9 surveys from NOAA scientists and leaders
  - 7 virtual meetings between the committee and those that filled out the surveys.
  - Collected input from approx. 18 NOAA scientists, leaders, and outside experts
- The Working Group report highlights some issues, challenges and opportunities that arose from this process.



# Scoping Open Science

## Open Data Scope and Definition

- There is a growing consensus on the definition and scope of open data.
- Open Data is a public data asset that is machine-readable; available (or could be made available) in an open format; not encumbered by restrictions, other than intellectual property rights, that would impede the use or reuse of such asset; and based on an underlying open standard that is maintained by a standards organization.

\* NOAA Data Strategy, July 2020, derived from Evidence Act (II.a17)]

\*\*National Academy of Sciences study on Open Source Software, 2018.

## Open Science Scope

1. In general, software developed by NOAA and NOAA supported projects should be **open-source** with a permissive license that encourages engagement and reuse.
2. In general, research by NOAA and NOAA supported projects should be **accessible and reproducible**.

## Open Science Related Definitions

- **Open-source software\*** Software whose source code is under an open-source license, by which the copyright holder grants to anyone the rights to inspect, modify, and distribute the source.
- **Reproducible research\*** Research published with all the necessary data, source code, and configurations to run the analysis again, re-creating the results and data products.

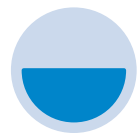
# Timeline



**Feb–Mar 2022**

Discussion of  
WG Scope

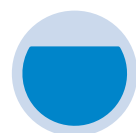
Formulation of  
Surveys



**Mar–Jul 2022**

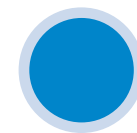
Surveys and  
Interviews

Scoping  
document  
presented to  
SAB in April



**Aug–Oct 2022**

First Draft of  
Report



**Nov–Dec 2022**

Second Draft  
of Report

Report  
presented to  
NOAA SAB



**NOAA SCIENCE ADVISORY BOARD**

# RECOMMENDATIONS

---

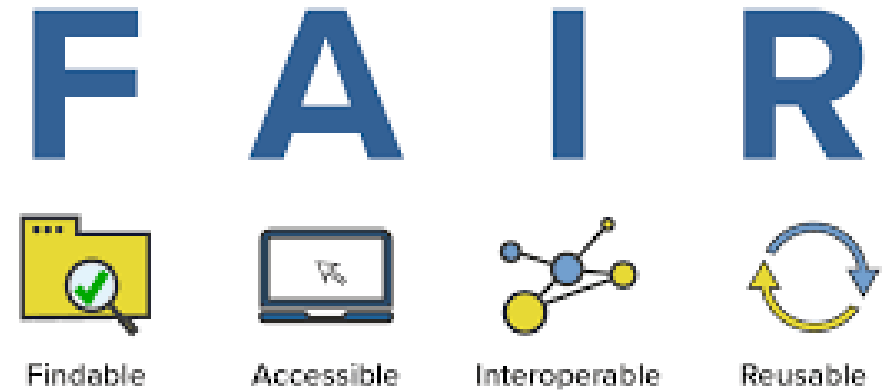
# Recommendation 1

**Recommendation 1. NOAA should follow the principles of FAIR open data and, whenever possible, these principles should be prioritized over other mission requirements.**

1.1 Specifically, since FAIR is open to interpretation, NOAA data should all be required to have persistent identifiers ([PIDs](#)), metadata, open-access APIs, and a standard open license (eg. cc-0 or cc-by).

1.2 Issuing PIDs for NOAA datasets should be made the [highest priority](#) and bottlenecks removed.

1.3 NOAA should consider the [impact](#) of any use restrictions on purchased commercial data on reproducibility and scientific impact and strive to minimize the use of non-open data whenever practical, as well as negotiating contracts that transition this data to open data after an appropriate time period.



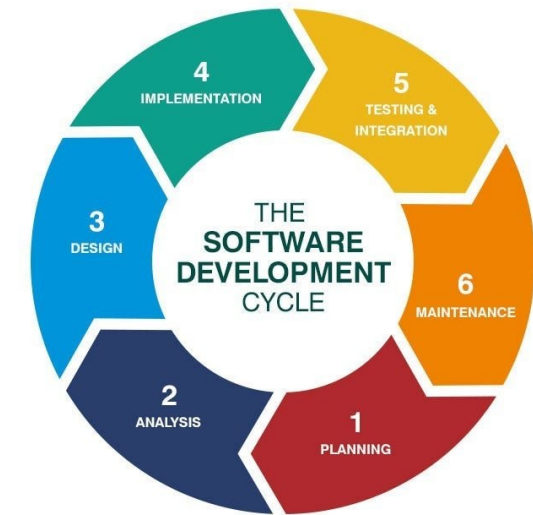
# Recommendation 2

## Recommendation 2: Encourage and support the use of open-source software as a key component of open science.

2.1 NOAA should develop agency wide guidance recommending the use of [permissive open-source software licenses](#) for most projects, unless there are compelling reasons otherwise.

2.2 Any new projects should agree [at the project start](#) to use a widely accepted permissive open-source software license and terms requiring this should be included in NOAA contracts and partnership agreements.

2.3 NOAA developed open-source software will require maintenance over time and NOAA should develop [agency wide guidance](#) on how this will be supported if the project that develops the software is ended, while other NOAA projects that rely on the software continue.





# Recommendation 3

**Recommendation 3: NOAA should engage with the open science community around open reproducible research and support workforce training on how to do open, collaborative, and reproducible science in support of the NOAA mission. Specifically:**

- 3.1 When NOAA scientists publish scientific papers, the software and configurations used for figures, tables, and **core results should be made available at time of publication.**
- 3.2 NOAA should invest in **workforce development** in broad support of open-source software, make research results reproducible at the time of publication, and more generally open science.
- 3.2 NOAA should sponsor or leverage an annual conference or other annual event, such as a session at a larger scientific conference, with an accompanying report to **engage consistently with the external scientific community** around open data, reproducible research, and more generally open science.





# Recommendation 4

**Recommendation 4: Consider providing consistent guidance across the agency for best practices, checklists, and dashboards to track adherence to open science principles, policies and mandates across the enterprise, while still supporting NOAA's distributed culture of data and science.**



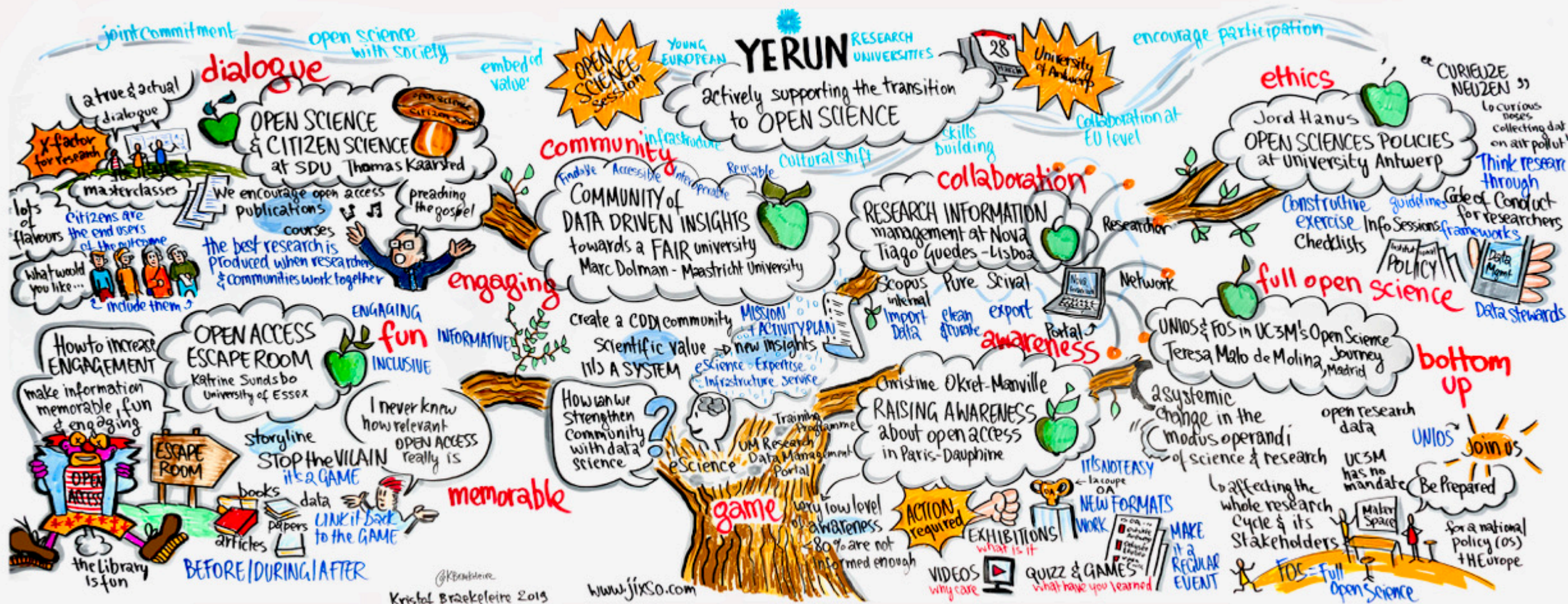
# Special Thanks To:

We want to give special thanks for all those that completed the surveys and participated in the interviews that followed.





# DISCUSSION



# References

[NASEM 2018] National Academies of Sciences, Engineering, and Medicine. Open Source Software Policy Options for NASA Earth and Space Sciences. National Academies Press, 2018

[NOAA 2020] NOAA Cloud Strategy: Maximizing the Value of NOAA's Cloud Services, July, 2020, retrieved from <https://sciencecouncil.noaa.gov/Portals/0/2020%20Cloud%20Strategy.pdf> on November 1, 2022.

[NOAA 2020] NOAA Data Strategy, Maximizing the Value of NOAA's Data, July, 2020, retrieved from <https://sciencecouncil.noaa.gov/Portals/0/2020%20Data%20Strategy.pdf> on November 1, 2022.

[Wilkinson 2016] Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg et al. "The FAIR Guiding Principles for scientific data management and stewardship." Scientific data 3, no. 1 (2016): 1-9.