NOAA Open Data / Open Science Working Group: Update & Status

August 30, 2022

<u>Subcommittee</u>
Robert Grossman, U. Chicago
Chelle Gentemann, NASA
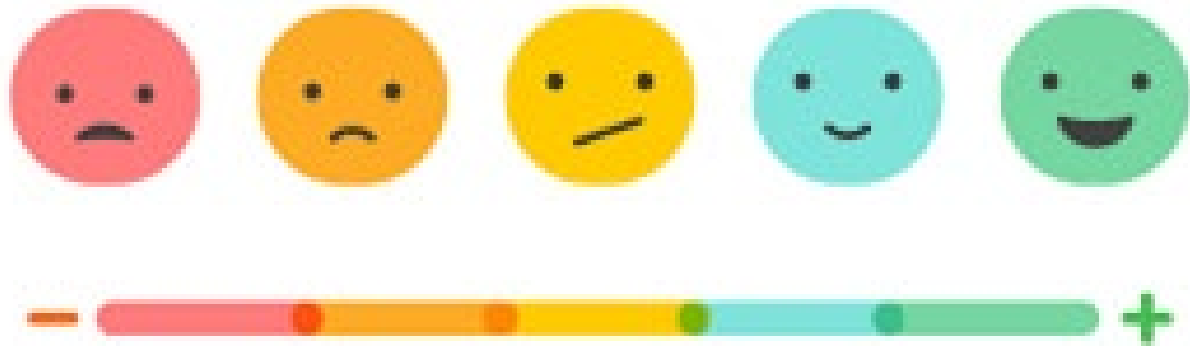Jason Hickey, Google (stepped down)
Zhaoxia Pu, U. Utah
Ilene Carpenter, HP
Chris Lendardt, UNC
Tony Wu, AeroMarine

Asked NOAA and others about open data / open science

- Sent out discussion questions around open data / open science to selected NOAA and cloud providers
- 11 returned questionnaire
- Interviewed 10 of those that responded to the questionnaire.

Interviewees

IT

Dave Fischman, Intrapreneur, SaaS PM, CIO

Kim Valentine, Geospatial Information Officer, NOS

Eugene Burger, Assoc. Director of IT, PMEL

HQ

Tony LaVoi, CDO

Derek Hanson, Attorney-Advisor

Joseph Pica, Acting Director, NCEI

Adrienne Simonson, Director, BDP

Scientists

Douglas Rao, Research Scientist, CISESS

Kevin Garrett, Physical Scientist, NESDIS

Sid Boukabara, Senior Scientist, NESDIS

NMFS

Karen Sender, Information Architect, NMFS

Nancy Majower, Chief Information Officer, NMFS

IOOS

Tiffany Vance, Ocean Technology Transition Program Manager, IOOS

Melissa Zweng, DMAC Program Coordinator, IOOS

Matt Biddle, Data Management Analyst, IOOS

Cloud

Shane Glass, Developer Advocate, Google

Rob Emanuele, Geospatial Architect, Microsoft

Ana Pinheiro Privette, Principal Program Manager, Scientific Data, Amazon Sustainability

Questions

OS Principles	Anything Unique?	Change	Challenges	Needs
Any? Describe? Useful? Improvements?	Are there any differences specific to your office/data/….	Approaches & techniques to manage adoption of OS?	Challenges meeting your mission, tradeoffs between mission and users	What would help?

Opportunities	Barriers	Anything else?
What could openness enable?	Barriers to serving open data/science to users	Open ended question about open science…

Outline of Report on Open Data / Open Science

1. Introduction
    1. Goal of report
    2. Definition of open data and open science
    3. Where does this report fit in?

1. Survey Responses
    1. Open data / open science frameworks used
    2. Common problems, opportunities and challenges identified
    3. Office/division/research area specific challenges and opportunities identified
    4. What is the impact on users of open data/open science and how is it quantified?

2. Findings

3. Recommendations

4. Summary

Report Timeline

Period	Focus of WG	Number of meetings
March - July	Discussion with NOAA scientists	9 meetings
August	Outline discussed and finalized, writing assignments, and initial prose	2 meetings
Sept	First draft; get informal feedback from selected SAB members	3 meetings
Oct	Second draft of report; draft report sent to SAB for feedback; prepare first draft of SAB presentation	2 meetings
Nov/Dec	SAB presentation Dec 1; revise report based upon SAB presentation	Report presented to SAB

On track

Early WIN

BDP → NODD

220+ NOAA datasets on the Cloud Service Providers (CSPs) platforms through the NOAA Open Data Dissemination (NODD) Program

Some of the topics discussed in the Open Data / Open Science Working Group

Discussions around open science*

NOAA's move to open science would benefit from high-level leadership of open science principles that becomes part of core messaging from headquarters. This should be followed through by announcing concrete changes to **promote and incentivize** open science across all policies and procedures, including evaluations, funding, and promotion.

*This topic was discussed in the ODOS working group.

Discussions around open science*

The distributed nature of NOAA data access points has created barriers to open science, redundant workflows, and inconsistent adoption of policies. A centralized open science office could provide guidance, checklists, and develop **dashboards to ensure adherence** to open science principles and tracking of adherence to policies and mandates.

*This topic was discussed in the ODOS working group.

Discussions around open science*

Streamlining the policy and procedures to **enable** open data/open science is a high priority.

*This topic was discussed in the ODOS working group.

Discussions around open science*

Issuing DOIs for NOAA datasets should be made the **highest priority** and bottlenecks removed.

*This topic was discussed in the ODOS working group.

Discussions around open science*

Training the workforce (both current and future) on how to do open science in an immersive way is necessary.

Discussions around open science*

External users,
especially
commercial cloud
providers, are a key
partner in increasing
accessibility of data
for unforeseen uses.

*This topic was discussed in the ODOS working group.

Summary

- This is hard
- This is complicated
- Need for coordinated approach
- More agile to avoid missed opportunities
- NOAA has open science experts who want to help

Open Science rewards

Report

How can our report best help NOAA move to open science / open data?

What do you need from us? Have we interviewed the right constituents?

Backup

Questions

1. Do you use a particular framework or set of **principles for open data/open science** in your mission and the work of your office/division? If so, can you please describe these? Are the principles helpful? Do you have any suggestions for how they can be improved?
2. What are some of the ways that applying open data / open science principles are **different** for the work in your office/division than would be for other components of NOAA, and more generally, for other disciplines or application domains? Or if there are no relevant differences, please let us know.
3. The expectations (eg. from open data to open FAIR data with license, or from open access publication to open access publication with DOI for all data and software so that results can be reproduced) for open data/open science are growing and the underlying technologies are rapidly evolving. How does this impact your mission? Have you developed any **approaches or strategies** for managing the changing expectations and evolving technologies?
4. What are some of the **challenges** that you face applying open data/open science principles to your mission? Are there any **tradeoffs** between meeting your internal mission and serving the user community with open data/open science? Are there any criteria you use to manage these tradeoffs?

Questionnaire

5. What do you **need** or what are you missing to meet these challenges?
6. What are some new or different **opportunities** that open data / open science practices could better enable or inspire? How will this help you support your mission and user community?
7. How do you evaluate whether the open data / open science approach you are using for your mission serves the needs of your user community? What are some of the **barriers and impediments** providing open data/open science to your user communities? How varied is your user community and what kinds of challenges does that create for meeting their open data / open science expectations?
8. What questions should we have asked, but **didn't**, and how would you answer them?

Interviewees

NOAA

- Dave Fischman, Intrapreneur, SaaS PM, Office Chief Information Officer
- Eugene Burger, Associate Director of IT, PMEL
- Sid Boukabara, Senior Scientist, Office of Systems Architecture & Advanced Planning, NESDIS
- Douglas Rao, Research Scientist, CISESS (NOAA Cooperative Institute)
- Joseph Pica, Acting Director, National Center for Environmental Information
- Tony LaVoi, Chief Data Officer
- Adrienne Simonson, Director, Big Data Program
- Kim Valentine, Geospatial Information Officer, NOS
- Nancy Majower, Chief Information Officer, NMFS
- Derek Hanson, Attorney-Advisor
- Karen Sender, Information Architect, NMFS
- Kevin Garrett, Physical Scientist, NESDIS
- Tiffany Vance, Ocean Technology Transition Program Manager, Integrated Ocean Observing System Office
- Matt Biddle, Data Management Analyst, Integrated Ocean Observing System Office
- Melissa Zweng, DMAC Program Coordinator, Integrated Ocean Observing System Office

Cloud providers

- Shane Glass, Developer Advocate, Google
- Ana Pinheiro Privette, Principal Program Manager, Scientific Data, Amazon Sustainability
- Rob Emanuele, Geospatial Architect, Microsoft