

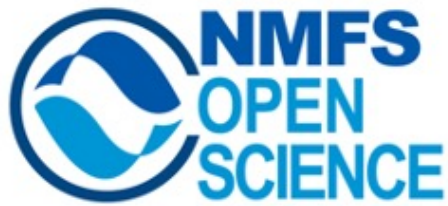
NOAA Fisheries Open Science and the 2023 Year of Open Science

Eli Holmes, Ph.D

Northwest Fisheries Science Center

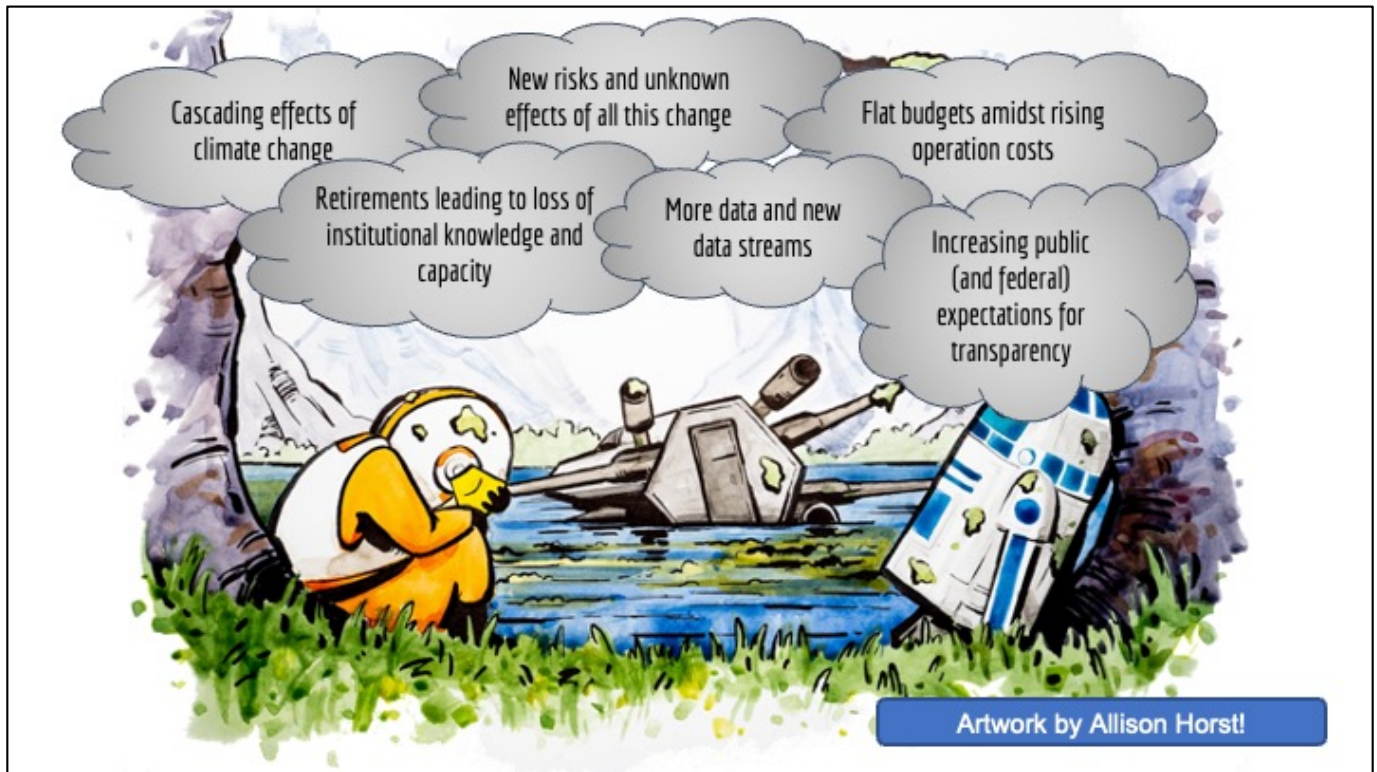
NMFS Openscapes, Co-PI

NMFS Open Science, Lead

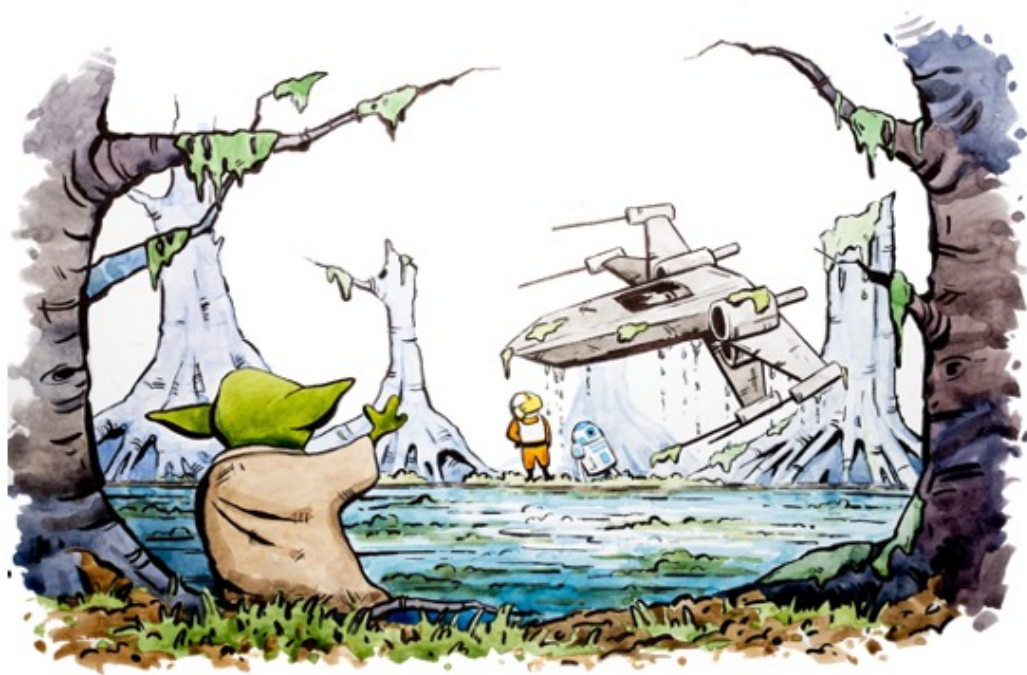


<https://nmfs-opensci.github.io/>





Art: <https://allisonhorst.com/>





**NOAA
FISHERIES**



2020-2022
Openscapes
program

NMFS Open Science &
Year of Open Science 2023



nfms-openscapes.github.io
openscapes.org

The White House announces 2023: A Year of Open Science

A multi-agency initiative across the US Federal Government to spark change and inspire open science engagement through events and activities that will advance adoption of open science.

- ✦ NASA
- ✦ National Oceanic and Atmospheric Administration
- ✦ National Science Foundation
- ✦ Department of Energy
- ✦ General Services Administration
- ✦ National Endowment for the Humanities
- ✦ National Institutes of Health
- ✦ National Institute of Standards and Technology
- ✦ US Department of Agriculture
- ✦ US Geological Survey



Gentemann, Chelle L., Shrestha, Sudhir, Ivey, Yvonne, & Hall, Cynthia. (2023, February 9). TOPS February 9 Community Forum. Zenodo.
<https://doi.org/10.5281/zenodo.7626005>

<https://open.science.gov/>

<https://www.whitehouse.gov/ostp/news-updates/2023/01/11/fact-sheet-biden-harris-administration-announces-new-actions-to-advance-open-and-equitable-research/>

<https://nasa.github.io/Transform-to-Open-Science-Book/index.html>

<https://www.earthdata.nasa.gov/news/year-of-open-science>

What is Open Science?



White House [Office of Science and Technology Policy](#) (OSTP) official definition in 2023 Year of Open Science



Slide from: Gentemann, Chelle L., Shrestha, Sudhir, Ivey, Yvonne, & Hall, Cynthia. (2023, February 9). TOPS February 9 Community Forum. Zenodo. <https://doi.org/10.5281/zenodo.7626005>

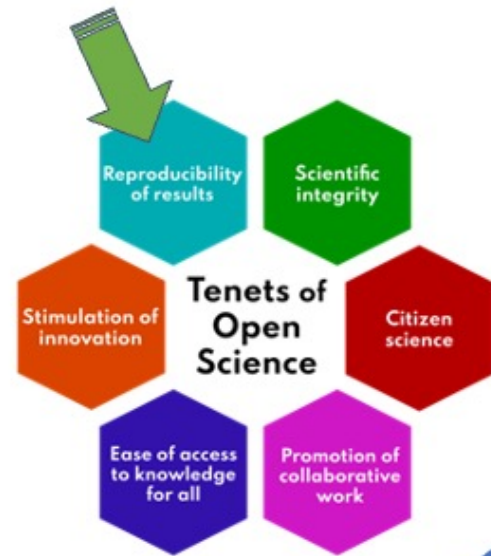
The Reproducibility Crisis in Science

Scientific fields have been rocked by the “reproducibility crisis” that has been building for the last 10 year or so, although really came to fore around 2015.

Journals begin requiring authors share the raw data and code

Recently scientific studies have shown that significant (over half) of studies cannot be replicated – even with the raw data and written methods.

Journals are moving toward requiring that authors share the “data to paper pipeline”



Do a lit search of “Reproducibility Crisis” post-2020 to find reams of publications on this. These are a few that are particularly relevant to my talk.

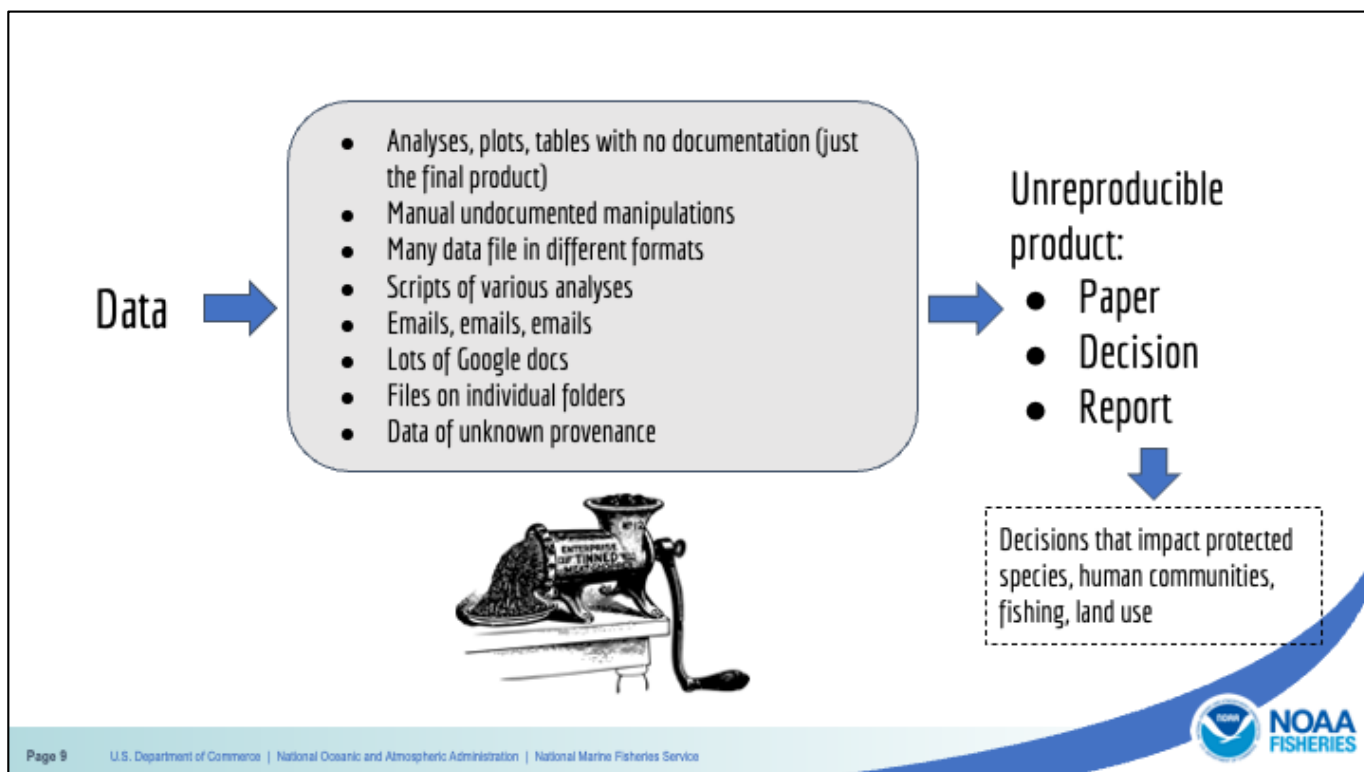
Treves, A. (2022), “Best available science” and the reproducibility crisis. *Front Ecol Environ*, 20: 495-495. <https://doi.org/10.1002/fee.2568> Especially read the Supplemental information: <https://esajournals.onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Ffee.2568&file=fee2568-sup-0001-Supinfo.pdf>

Discussing the controversy around the 2018 EPA Proposed rule. “Strengthening Transparency in Regulatory Science” Daniel J. Hicks (2023) Open science, the replication crisis, and environmental public health, *Accountability in Research*, 30:1, 34-62, DOI: [10.1080/08989621.2021.1962713](https://doi.org/10.1080/08989621.2021.1962713)

Stagge JH, Rosenberg DE, Abdallah AM, Akbar H, Attallah NA, James R (February 2019). [“Assessing data availability and research reproducibility in hydrology and water resources”](#). *Scientific Data*. 6: 190030. Bibcode:2019NatSD...690030S. doi:10.1038/sdata.2019.30. “results might be reproduced for only 0.6% to 6.8% of all 1,989 articles” largely due to missing data, methods or code.

Stodden, V., Seiler, J., & Ma, Z. (2018). An empirical analysis of journal policy effectiveness for computational reproducibility. *Proceedings of the National Academy*

of Sciences, 115(11), 2584-2589. Cannot replicate even with the raw data.
<https://doi.org/10.1073/pnas.1708290115>



Summarizing discussions during 3 years of Openscapes cohorts with NMFS scientists.

How does one create a “reproducible scientific pipeline”?

- **Data:** Data management and documentation
- **Data wrangling:** Eliminating manual manipulation of data
- **Analysis:** Adopting a documented pipeline rather than a patchwork of poorly documented analyses

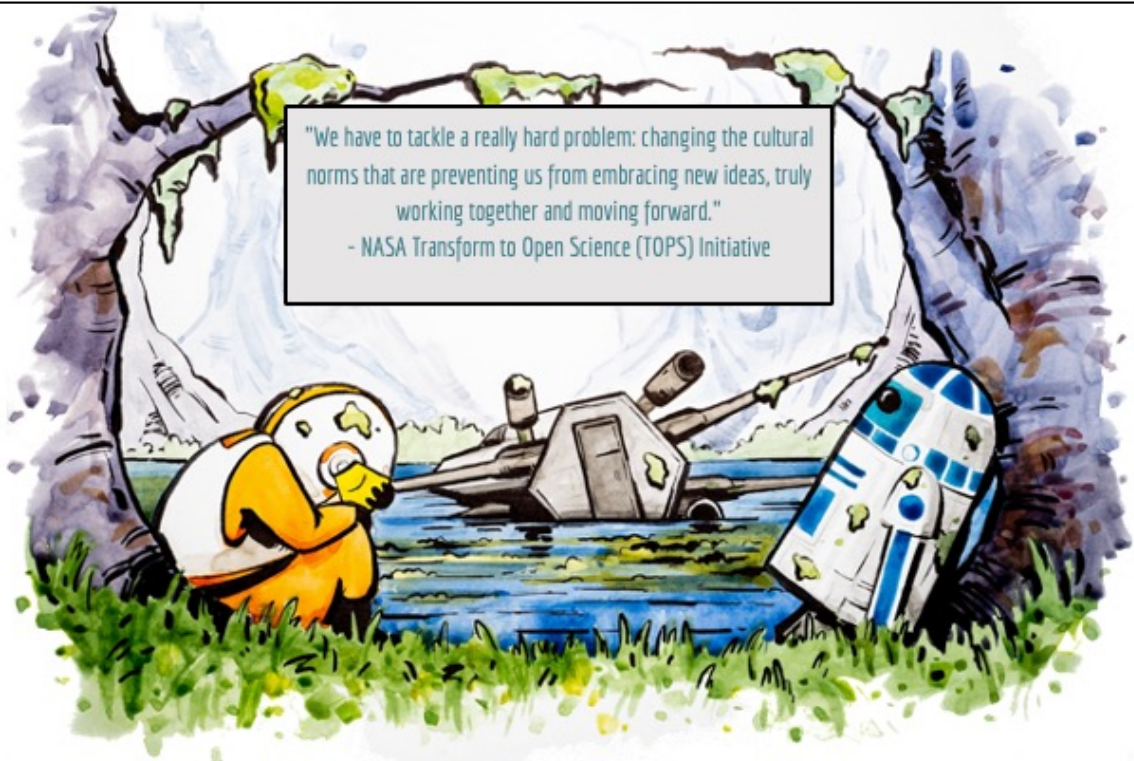
- **Version-control:** all changes and decision documented
- **Text and code integrated**
- Include a “repository” with a “make” file that reproduces the final product
- A “devcontainer” of the environment
- New skills, new tools, new ways of working



[Fisheries Information Management Modernization Workshop 2020, Tech Memo](#) September 17-19, 2019, NMFS Office of Science and Technology (OST)

[Fisheries Information Management Modernization Workshop 2020, Tech Memo](#) September 17-19, 2019, Review and evaluate practical and tangible actions to modernize the data and information system of NMFS

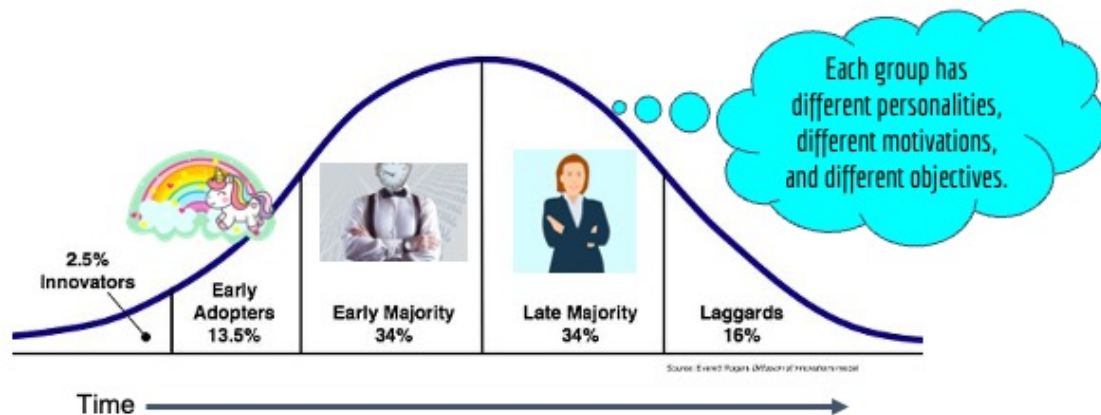
"We have to tackle a really hard problem: changing the cultural norms that are preventing us from embracing new ideas, truly working together and moving forward."
- NASA Transform to Open Science (TOPS) Initiative





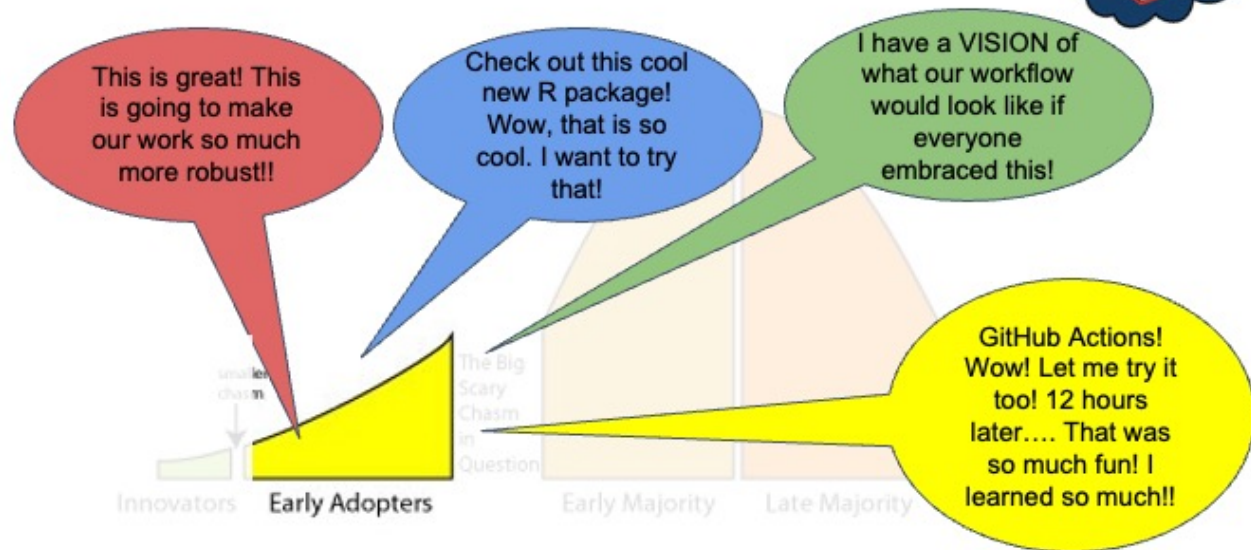
EM Rogers (1962) “Diffusion of Innovation” theory

Predictable progression of stages as idea diffuses through a population

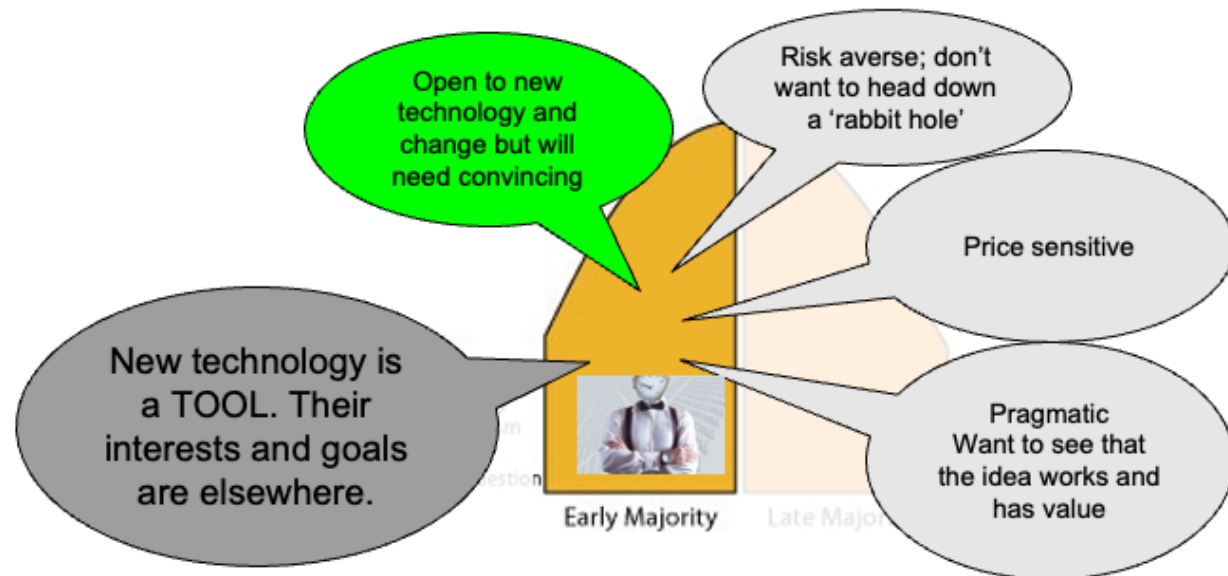


Rogers, E. M. (2003). *Diffusion of innovations*. New York, NY [u.a.]: Free Press. ISBN: 0-7432-2209-1, 978-0-7432-2209-9

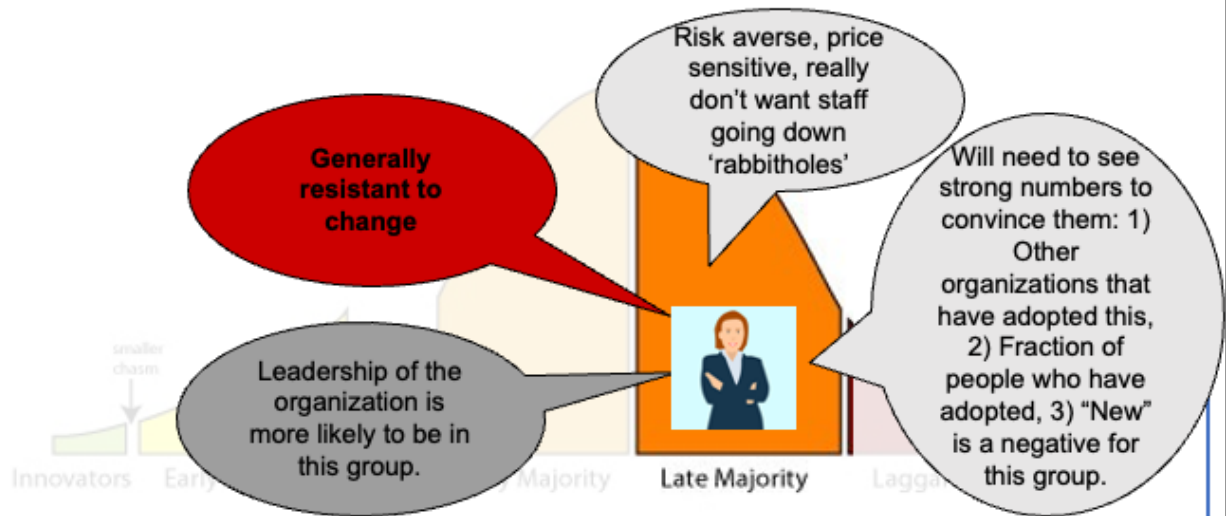
Open Science Early Adopters



Early Majority: Open to innovation but risk adverse

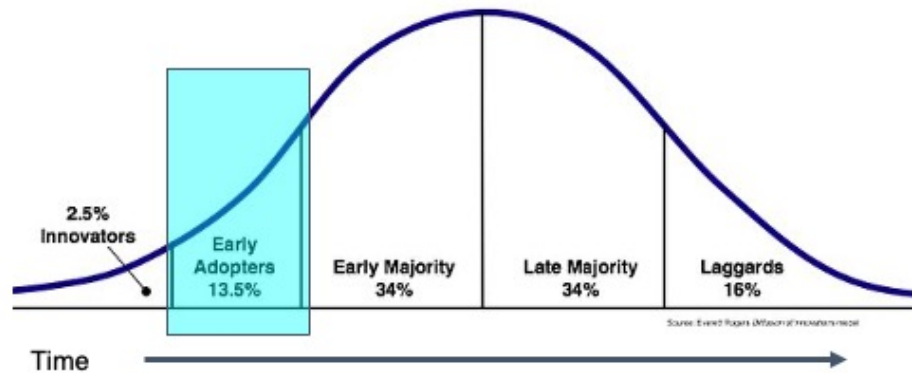


Late Majority often includes organizational leadership



The Early Adopters are critical to diffusion of innovation

1. Early Adopters develop the innovation into something of value
2. Their **energy and effort** is what drives the initial diffusion process, but that is a hard and slow process.





Sad and
lonely early
adopter

Happy early adopters
having found
community

openscapes

openscapes.org

WEIGHME!

NMFS Openscapes training in Open Science



At NMFS, a grassroots effort due to desire from staff for training in Open Science

9 NMFS Champions Cohorts (40 staff ea)

- 2020: Winter NEFSC
- 2021: Spring NWFSC
- 2021: Fall NWFSC, AFSC, SEFSC, NEFSC
- 2022: Winter AFSC
- 2022: Summer SEFSC/SERO
- 2022 Fall 4 cohorts 6 science ctrs, WCRO

<https://nmfs-openscapes.github.io/>



What is Openscapes?

Not your traditional training/workshop

- Cohort-based remote sessions for teams: introduce concepts and workflows; facilitate teams to talk about problems then go and solve them, with accountability and support.
- It's about getting stuff done. It's about identifying and making progress on barriers
- "A process to help you build better lanes of communication" -Laura Waters, SE Regional Office

Sustainability built-in

- Strengthening a teaching & learning culture within teams & orgs. Not just for scientists; admin, IT staff, etc, welcomed. Equitable.

No coding or software skills required

Openscapes works with many environmental
orgs



<https://openscapes.org/>



What is Openscapes? This is not your traditional training or workshop.

These are cohort-based remote sessions for teams, where Openscapes introduces concepts and workflows; facilitate teams to talk about problems then go and solve them, with accountability and support.

It's about getting stuff done. It's about identifying and making progress on barriers
Laura Waters at the SERO described it as "A process to help you build better lanes of communication"

What's important too is that there is sustainability and scalability built-in. It's about strengthening a teaching & learning culture within teams & orgs. Not just for scientists: supervisors, admin, IT staff, etc, welcomed. Equitable.

Openscapes has led 10 Champions Cohorts so far, half of which have been with NMFS! At NMFS it's been a grassroots effort

During Openscapes sessions, scientists talk about issues that are hindering their work. Big issues: reproducibility and tracking, team awareness, on- and off-boarding, too much duplication of effort across the center, agency and year. Last need for team training to collaborate openly and effectively.

Next slides

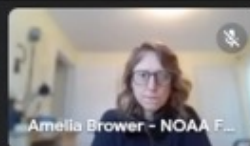


An Openscapes Future for Stock Assessment Reports at the AFSC's Marine Mammal Laboratory

Amelia Brower, Brian Fadely, Josh London, Tony Orr,
Erin Richmond, Rod Towell, and Nancy Young



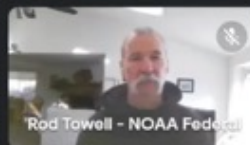
Joshua London - NOAA F...



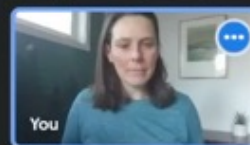
Amelia Brower - NOAA F...



Tony Orr - NOAA Federal

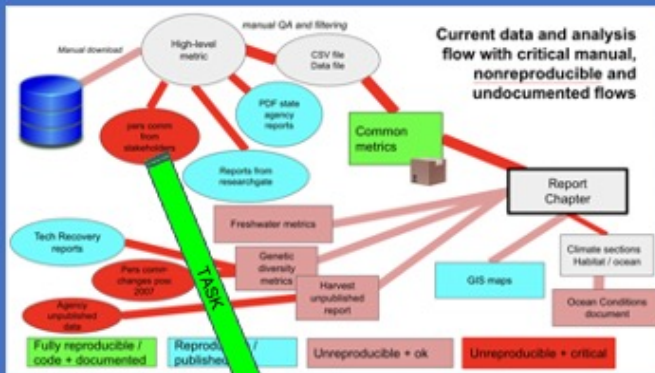


Rod Towell - NOAA Federal



You

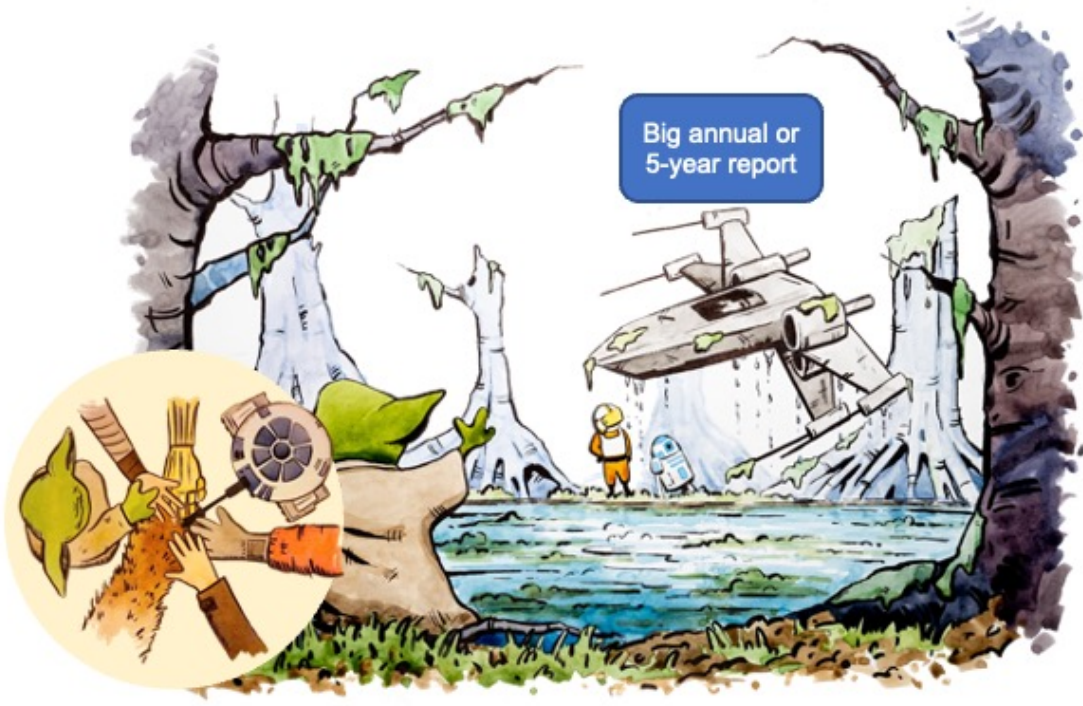
PNW Salmonid Viability Report (NWFSC) + Status Reviews (WCRO) Team



Set up single email address for pers comms re data requests



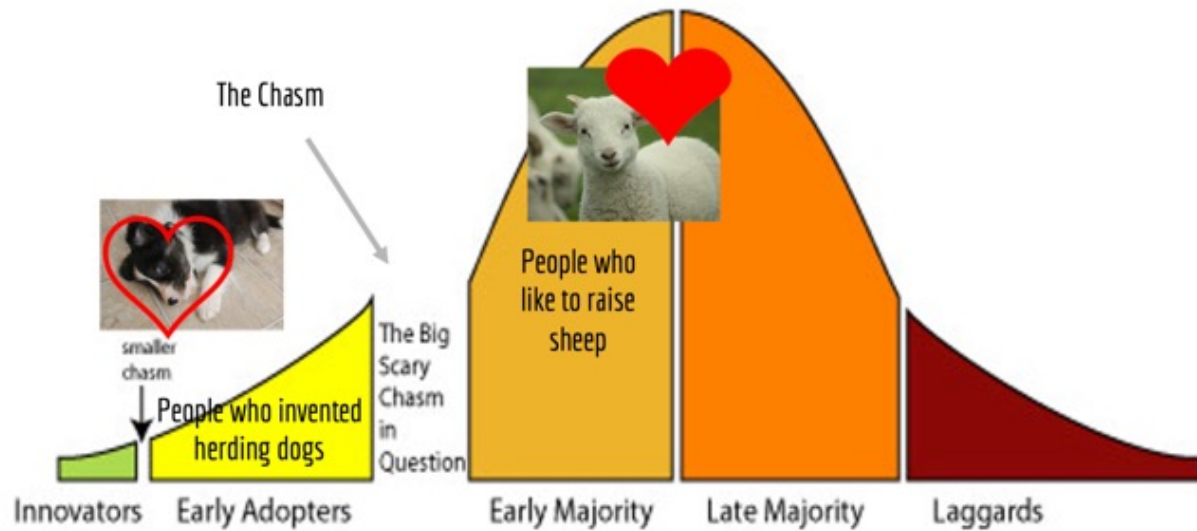
Big annual or
5-year report





Moore, G. A. (2014). *Crossing the Chasm, 3rd Edition: Marketing and Selling Disruptive Products to Mainstream Customers*. Harper Business.

The invention of the herding dog analogy



How do you cross the Chasm?

Option 1. A charismatic communicator
“salesperson” who is has deep connections with the
“majority” but also understands the innovation



Hmm, that's kind of hard and not obvious how to do.

From: Moore, G. A. (2014, January 28). *Crossing the Chasm, 3rd Edition: Marketing and Selling Disruptive Products to Mainstream Customers*. Harper Business.

How do you cross the Chasm?

Option 2. Judiciously choose a single market for the crossing. Put all your effort there.



From: Moore, G. A. (2014, January 28). *Crossing the Chasm, 3rd Edition: Marketing and Selling Disruptive Products to Mainstream Customers*. Harper Business.

Choose a single market for the crossing

Create many use cases. Pick the one where you can reduce a major and clear pain point and there isn't a good alternative.

2022 -- Big Government Reports

- Big time savings
- Savings in staff time can be quantified
- Staff eager to automate soul-crushingly tedious work
- Solves a transparency and documentation problem

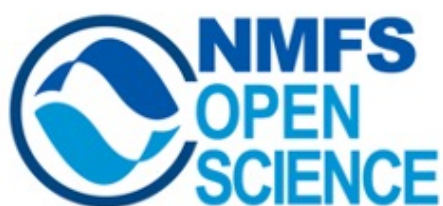
From: Moore, G. A. (2014, January 28). *Crossing the Chasm, 3rd Edition: Marketing and Selling Disruptive Products to Mainstream Customers*. Harper Business.



2023 Year of Open Science and Beyond

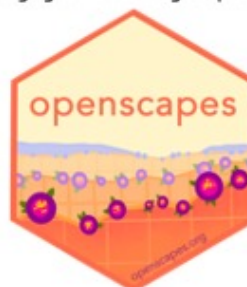
NMFS Open Science

The overarching vision of NMFS Open Science is to support scientists, developers, and policy analysts within NOAA Fisheries (NMFS) in fulfilling NOAA's Open Science mandates: NOAA Data Strategy, DOC Open Source Code Policy, Federal Data Strategy, and the Federal Open Access Memo.



NMFS Openscapes

is concerned Open Science training in workflow and technical skills needed at the individual and team level. We focus on helping all staff engaged in data-driven science and decision-making at NMFS. Support an active and engaged mentor group across NMFS.

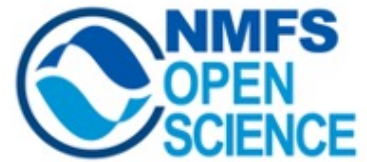


NMFS Openscapes Plan 2023-2026

[illegible]

NMFS Open Science is a strategic group

Triage the most pressing needs for scientists, developers, and policy analysts within all of NOAA Fisheries and take leadership roles to find solutions.



NMFS Openscapes	NOAA Fisheries Integrated Toolbox	NOAA Fisheries Integrated Modeling System	NMFS R User Group
The NOAA Fisheries logo and the Openscapes logo, which is a hexagon with a blue background and the word "openscapes" in white.	The NOAA Fisheries Integrated Toolbox logo, which is a hexagon with a blue background and the text "NOAA Fisheries Integrated Toolbox" in white.	The FIMS logo, which is a hexagon with a blue background and the text "FIMS" in white.	The NOAA Fisheries Users R logo, which is a hexagon with a blue background and the text "NOAA FISHERIES" in white, with a large "R" and "USERS" in blue.

<https://nmfs-opensci.github.io/>

Supporting the infrastructure for Open Science

Support for scientific software, package development, templates, utilities



Data science is highly dependent on soft infrastructure: development platforms, cloud virtual machines, and product delivery systems for data-science products. Support governance teams for these platforms.



GitHub Governance Team (live April 3, 2023. <https://sites.google.com/noaa.gov/nmfs-st-github-governance-team/home>)

