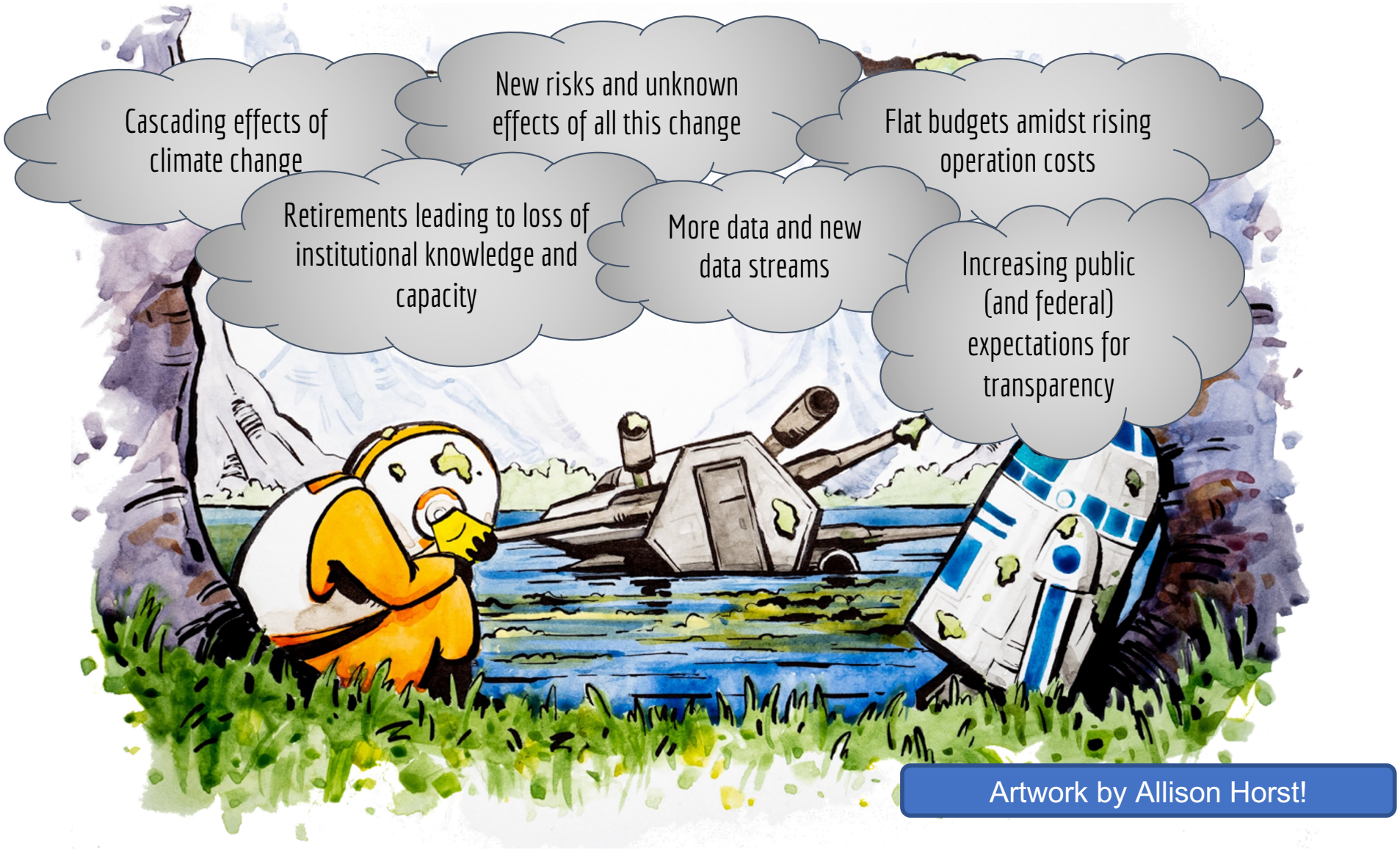
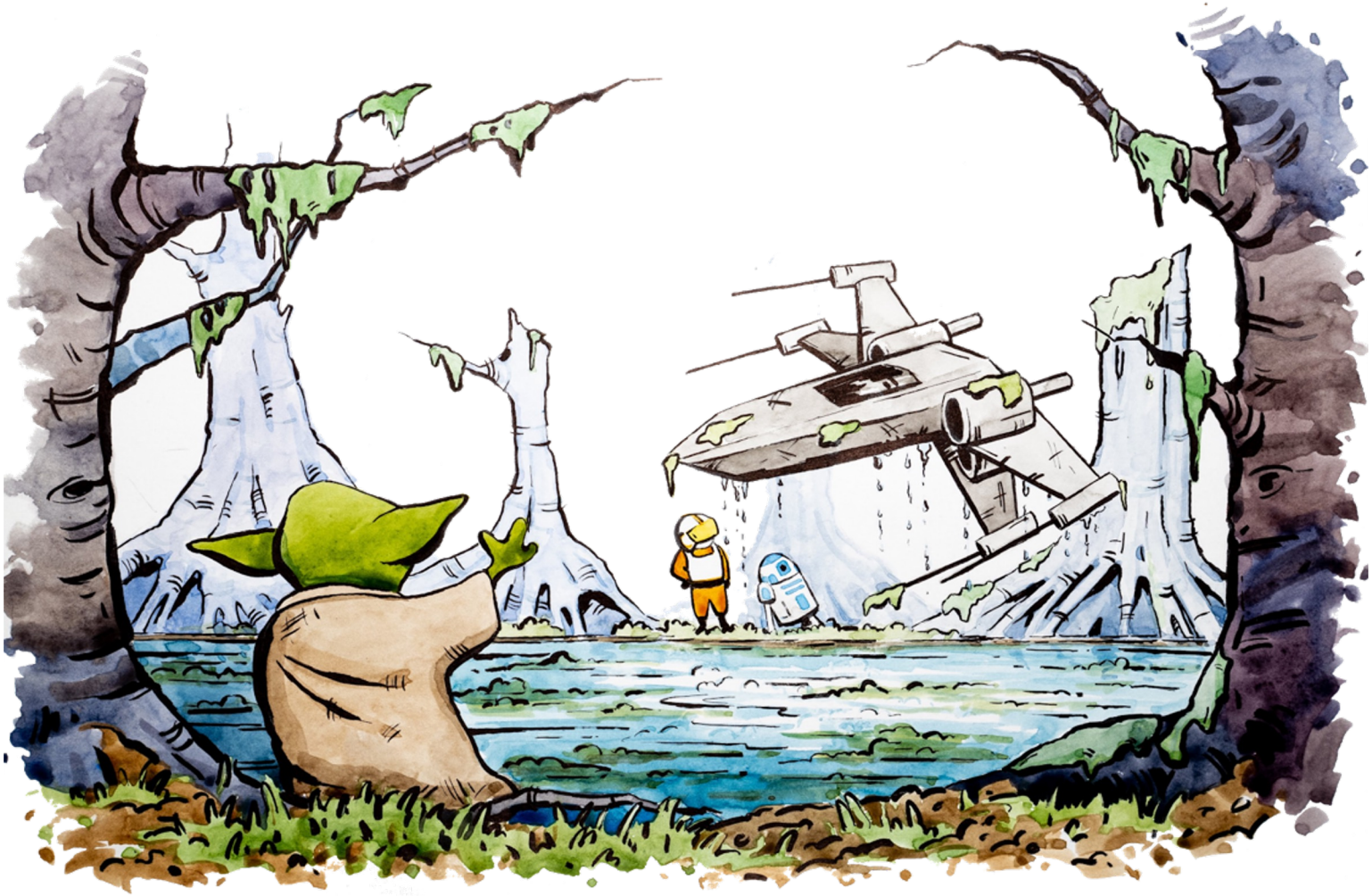
NOAA Fisheries Open Science and the 2023 Year of Open Science

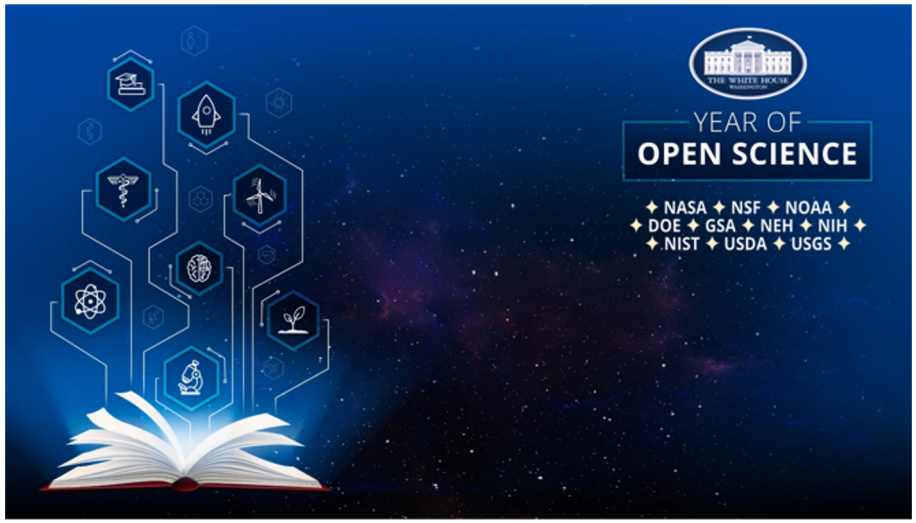
Eli Holmes, Ph.D
Northwest Fisheries Science Center
NMFS Openscapes, Co-PI
NMFS Open Science, Lead

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


Cascading effects of climate change
New risks and unknown effects of all this change
Flat budgets amidst rising operation costs
Retirements leading to loss of institutional knowledge and capacity
More data and new data streams
Increasing public (and federal) expectations for transparency
Artwork by Allison Horst!



NMFS Open Science & Year of Open Science 2023



2020-2022 Openscapes program

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

What is Open Science?

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

NOAA FISHERIES

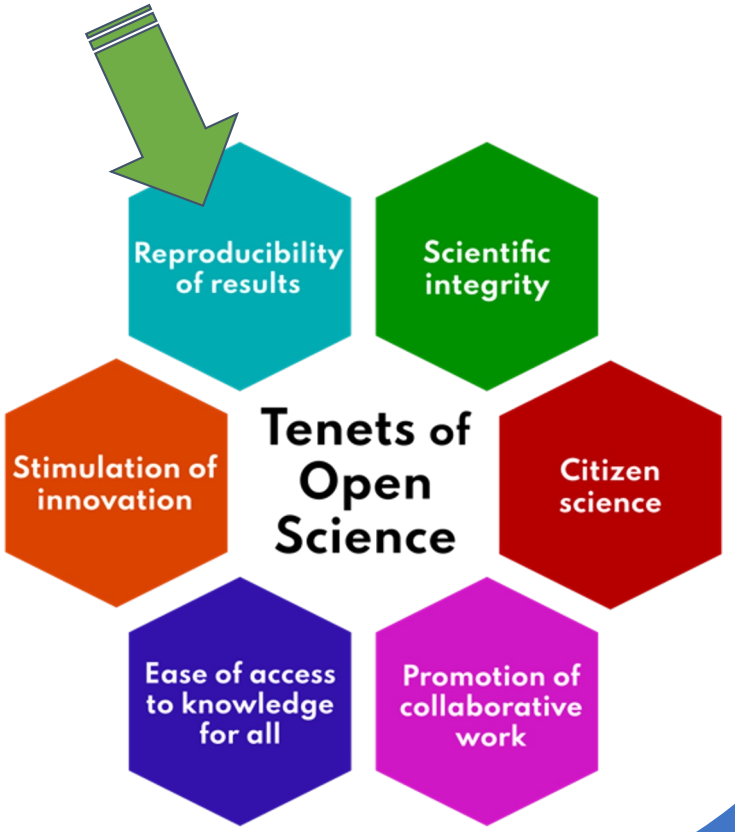
The Reproducibility Crisis in Science

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"



NOAA FISHERIES

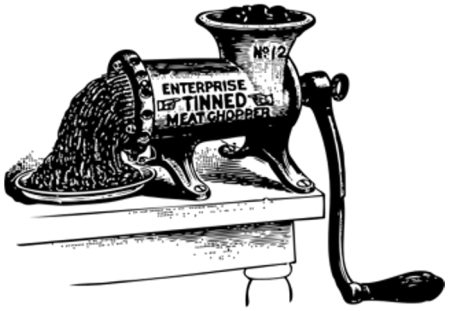
Data

- Analyses, plots, tables with no documentation (just the final product)
- Manual undocumented manipulations
- Many data file in different formats
- Scripts of various analyses
- Emails, emails, emails
- Lots of Google docs
- Files on individual folders
- Data of unknown provenance

Unreproducible product:

- Paper
- Decision
- Report

Decisions that impact protected species, human communities, fishing, land use

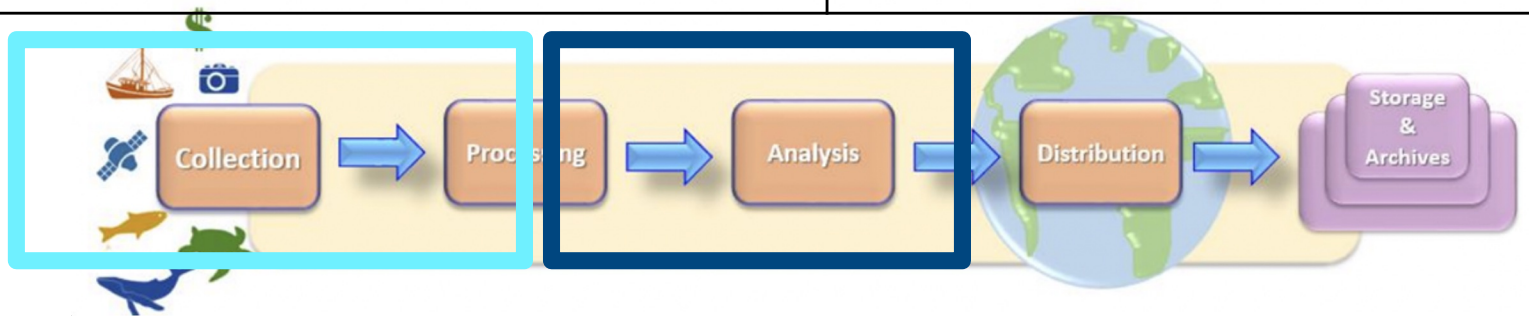


NOAA FISHERIES

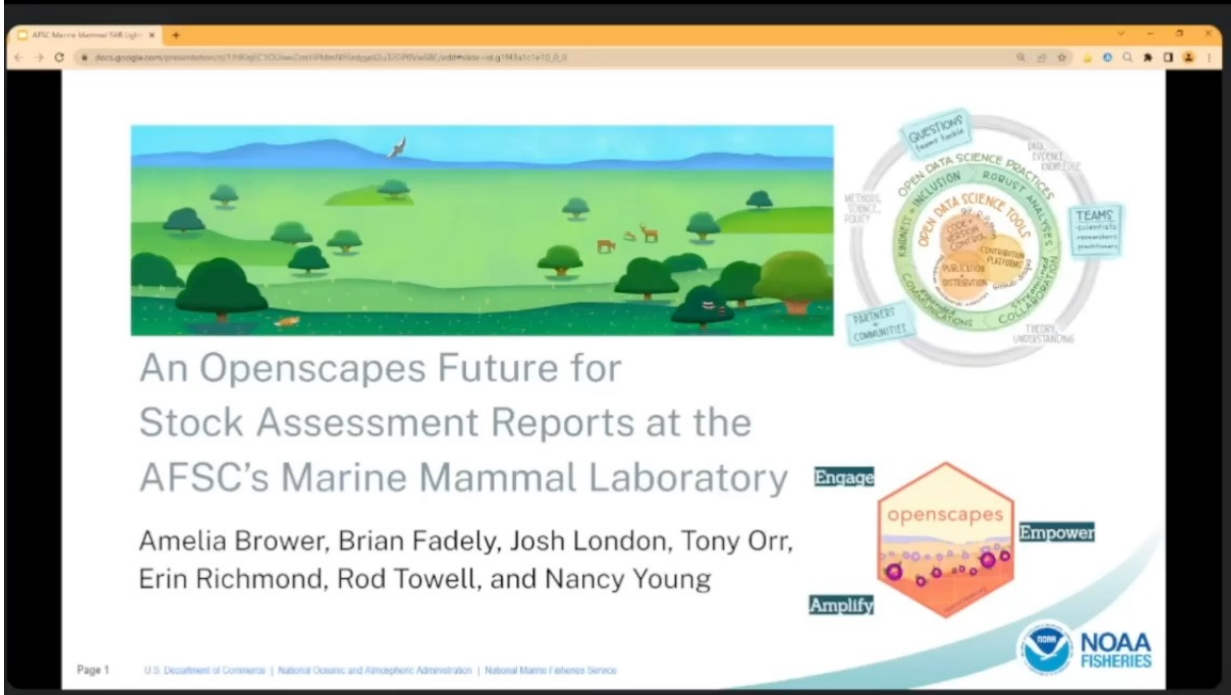
How does one create a “reproducible scientific pipeline”?

- **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)


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
Engage
openscapes
Empower
Amplify
NOAA
FISHERIES
Page 1
U.S. Department of Commerce | National Oceanic and Atmospheric Administration | National Marine Fisheries Service


Joshua London - NOAA F...


Amelia Brower - NOAA F...


Tony Orr - NOAA Federal

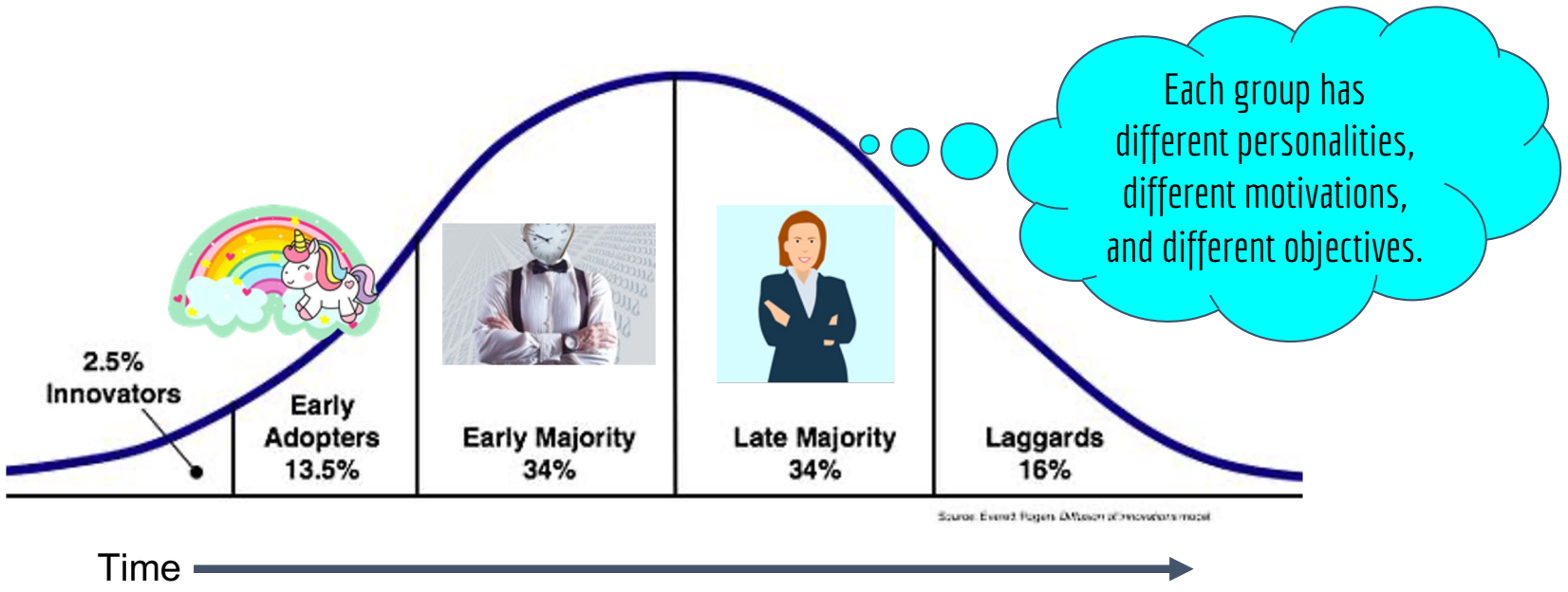

Rod Towell - NOAA Federal


You

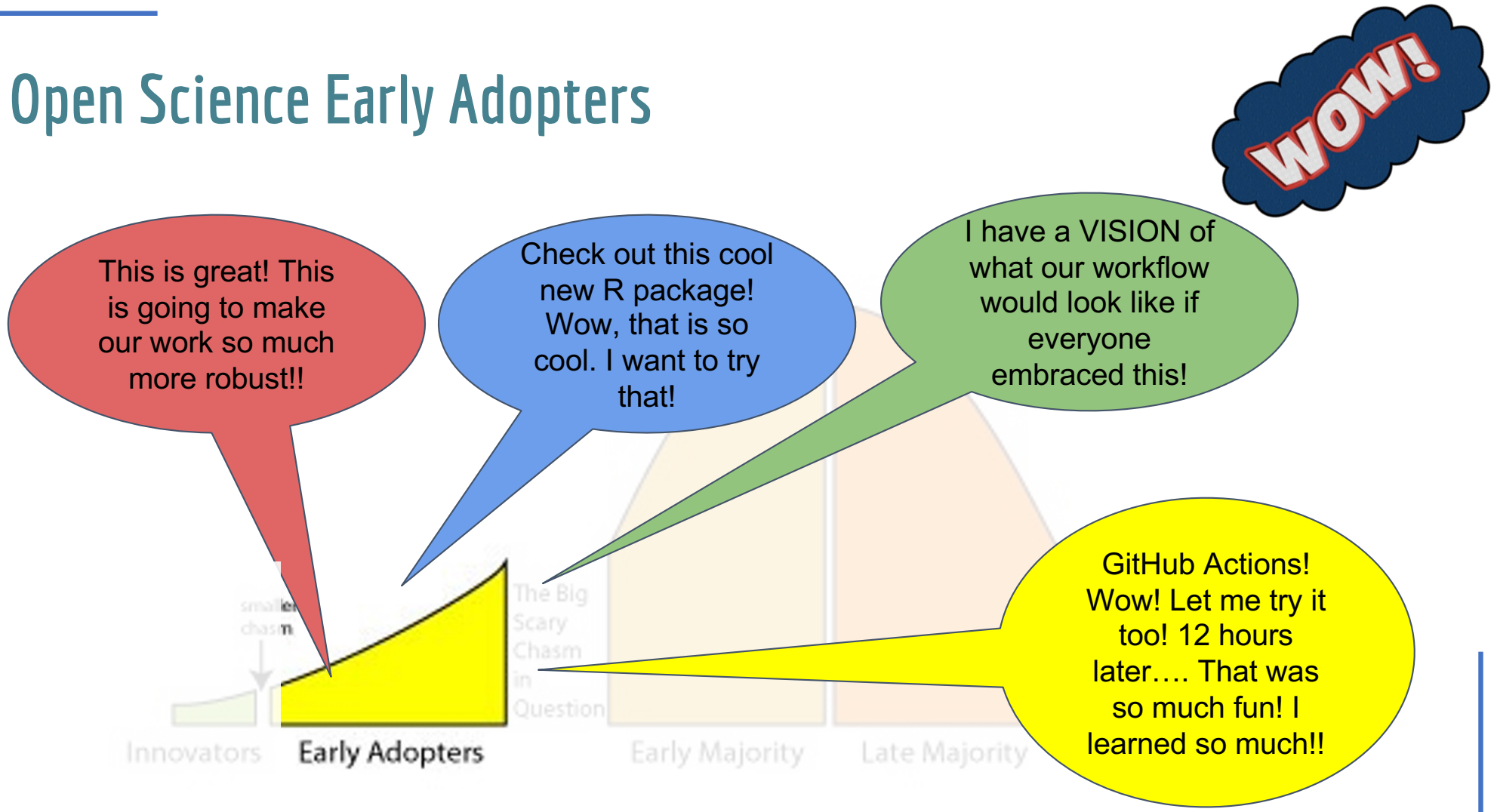

Fine for one team, but how to we spread new ways of working throughout an organization??

EM Rogers (1962) "Diffusion of Innovation" theory

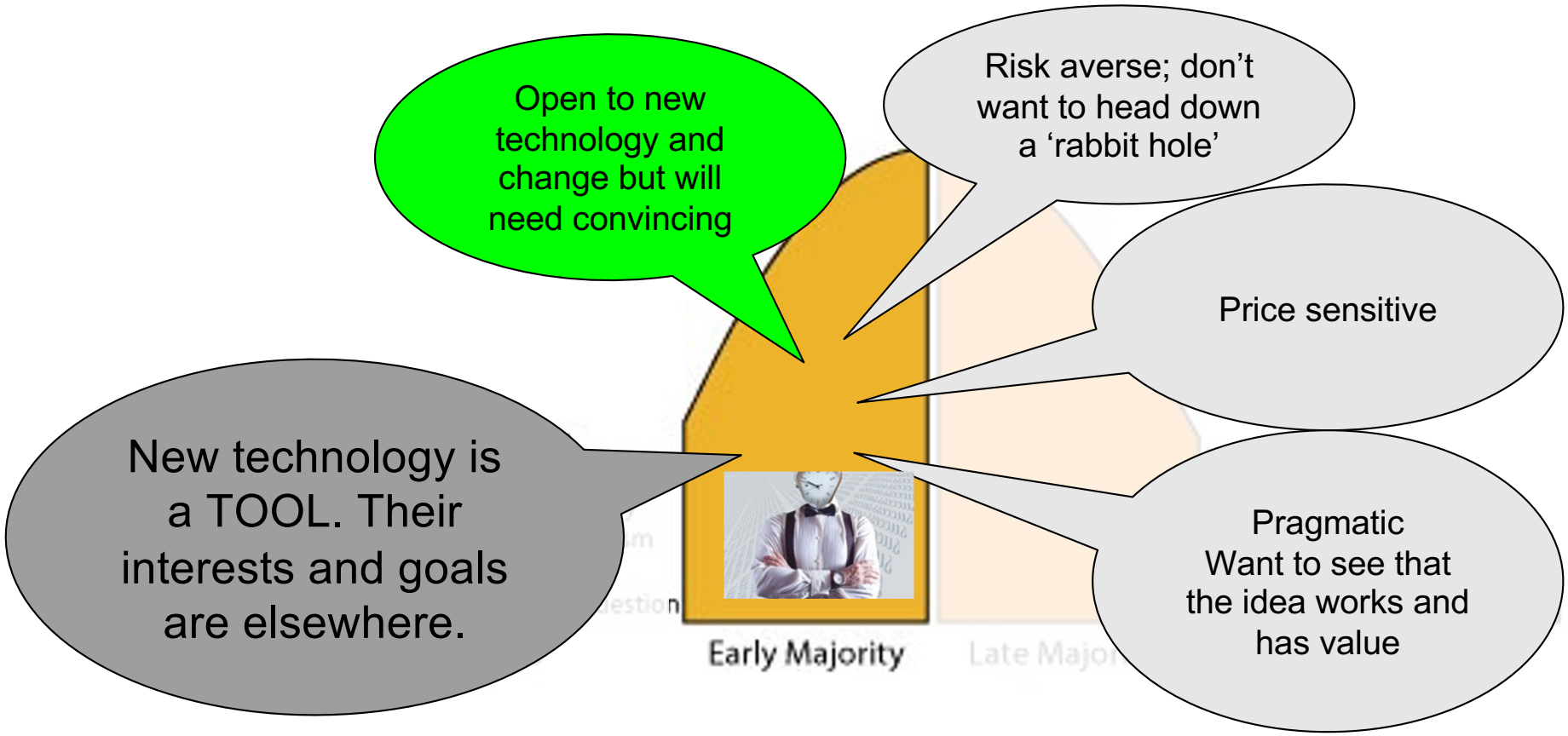
Predictable progression of stages as idea diffuses through a population



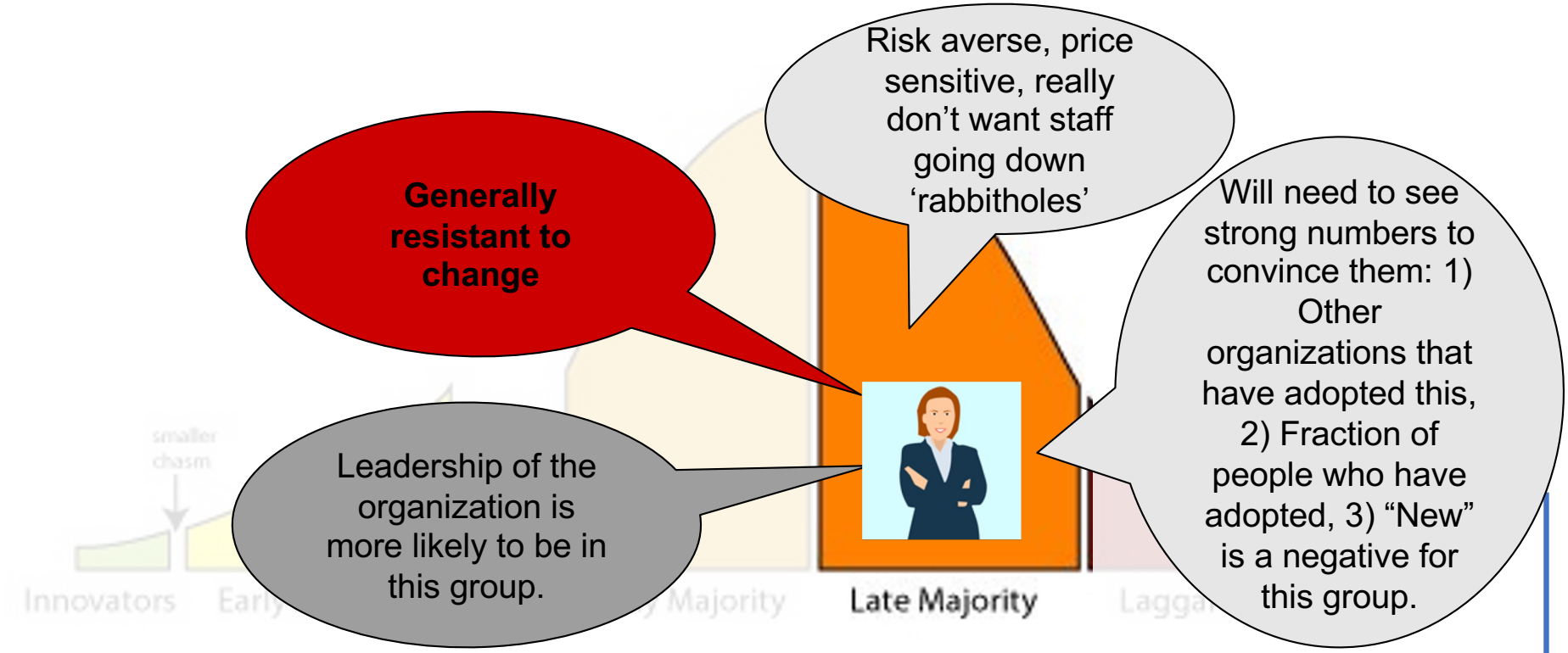
Open Science Early Adopters



Early Majority: Open to innovation but risk adverse

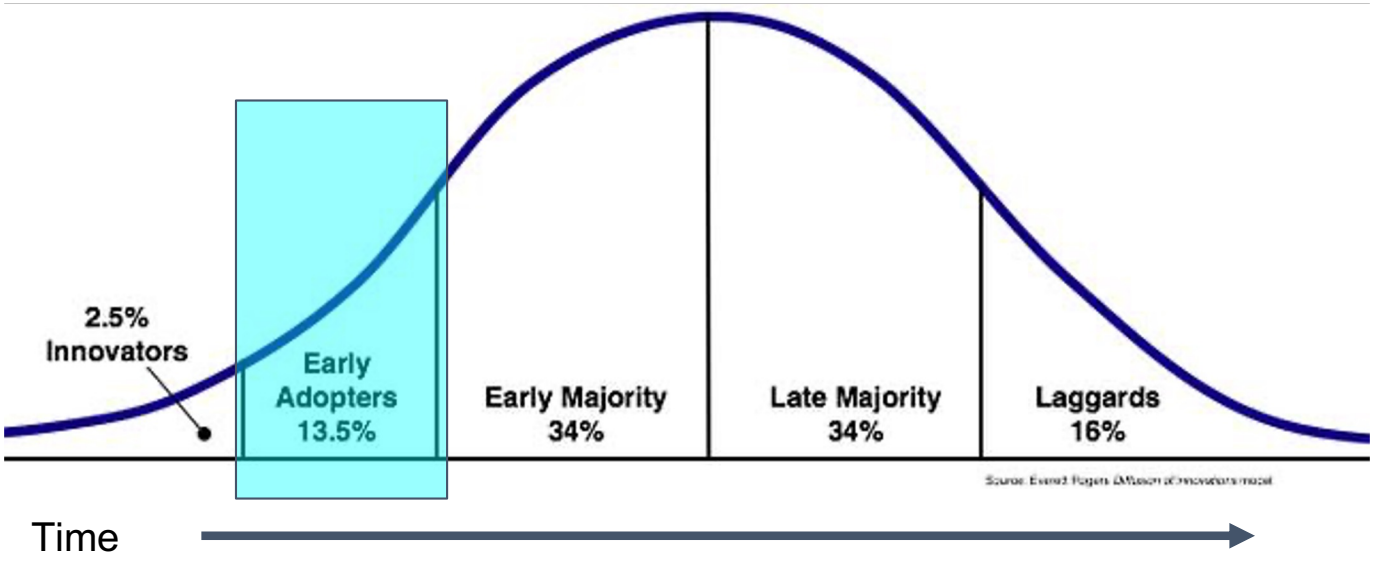


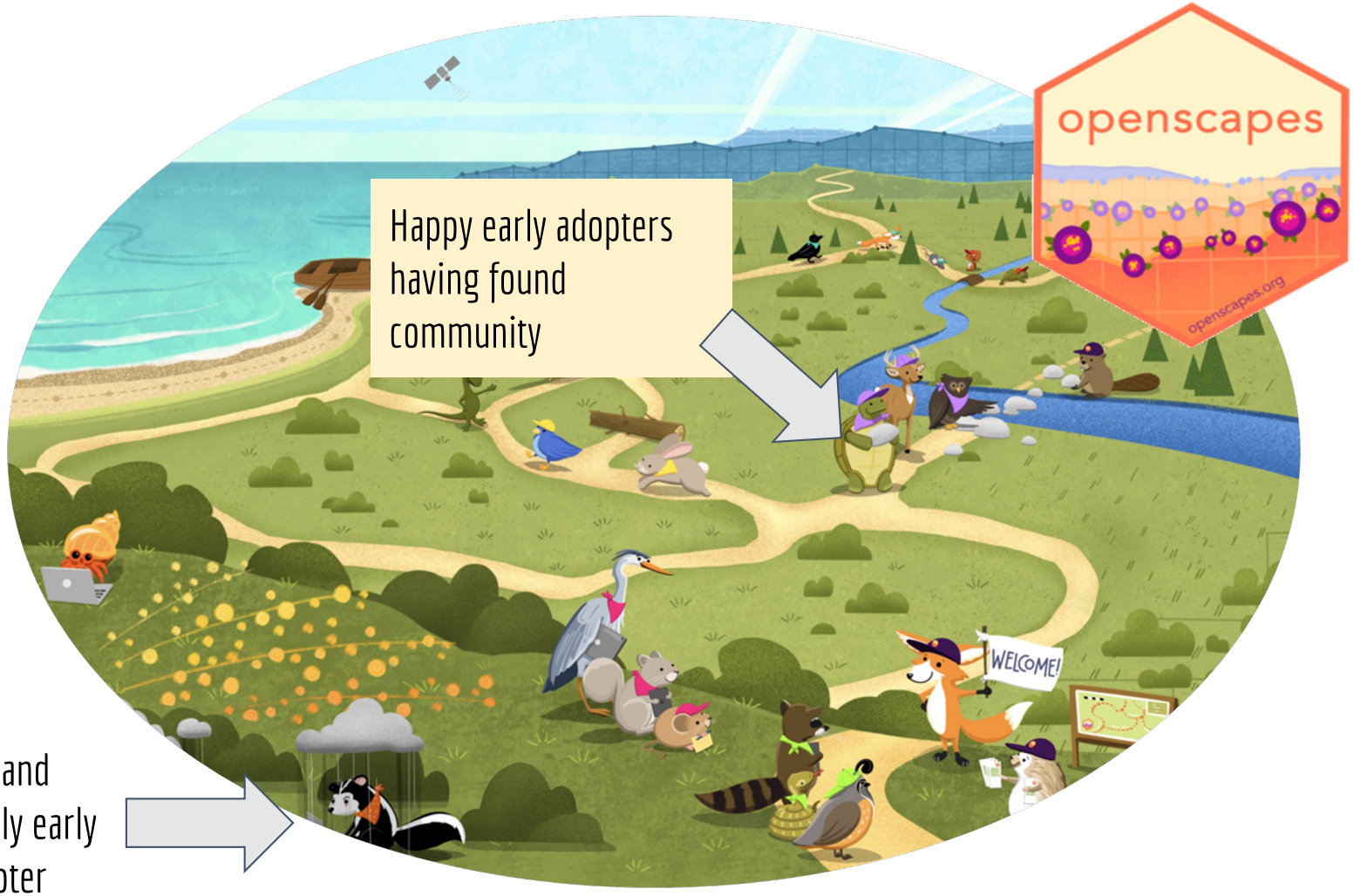
Late Majority often includes organizational leadership



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.




openscapes
openscapes.org
Happy early adopters having found community
Sad and lonely early adopter
WELCOME!

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.

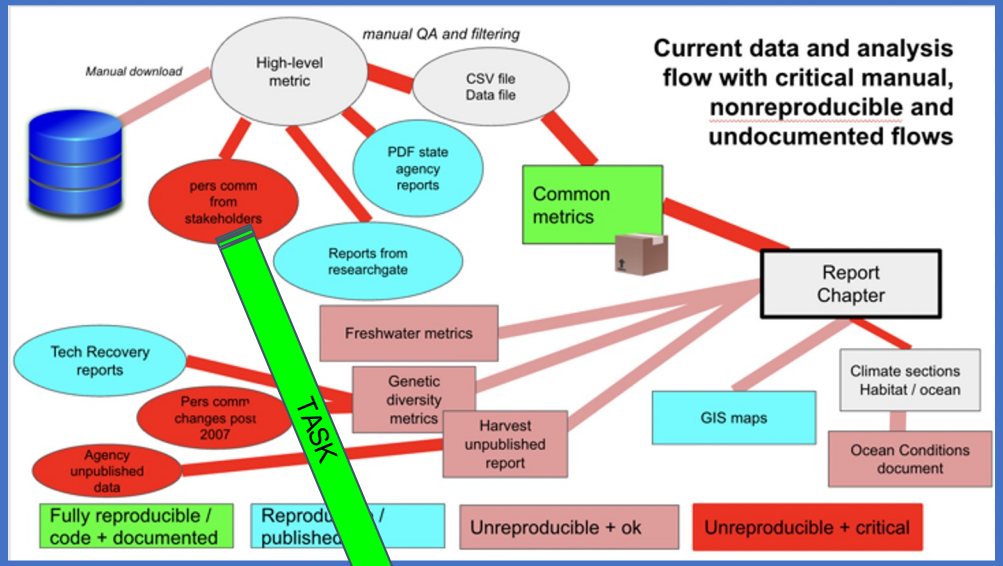
No coding or software skills required

Openscapes works with many environmental orgs

https://openscapes.org/

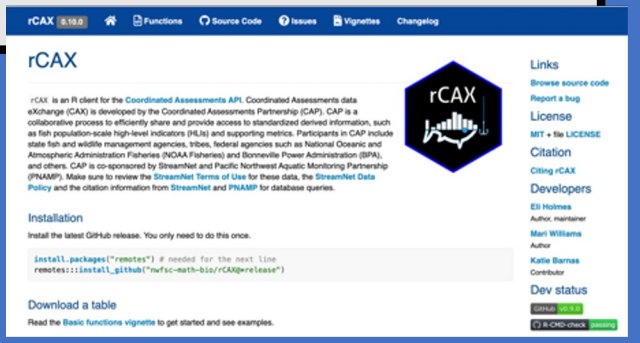
NOAA FISHERIES

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

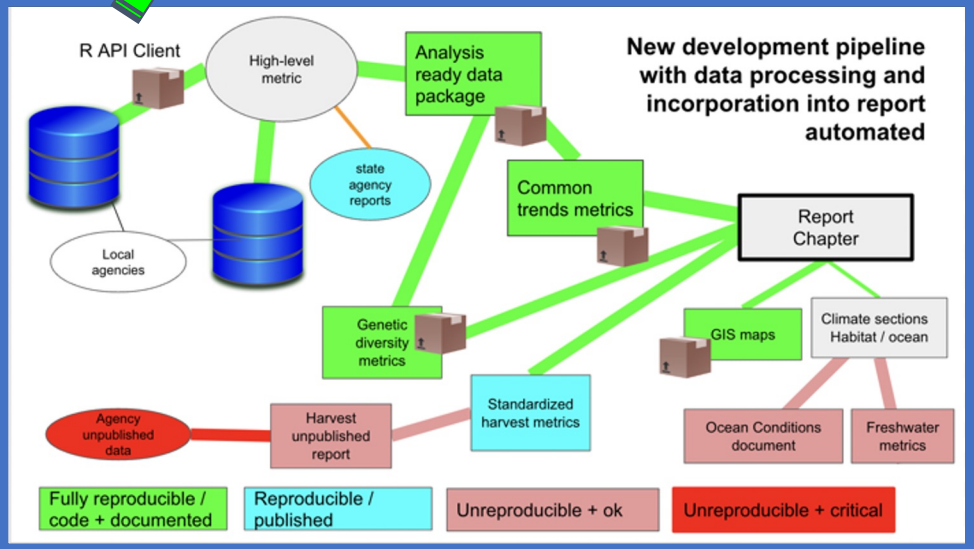


TASK

Set up single email address for pers comms re data requests

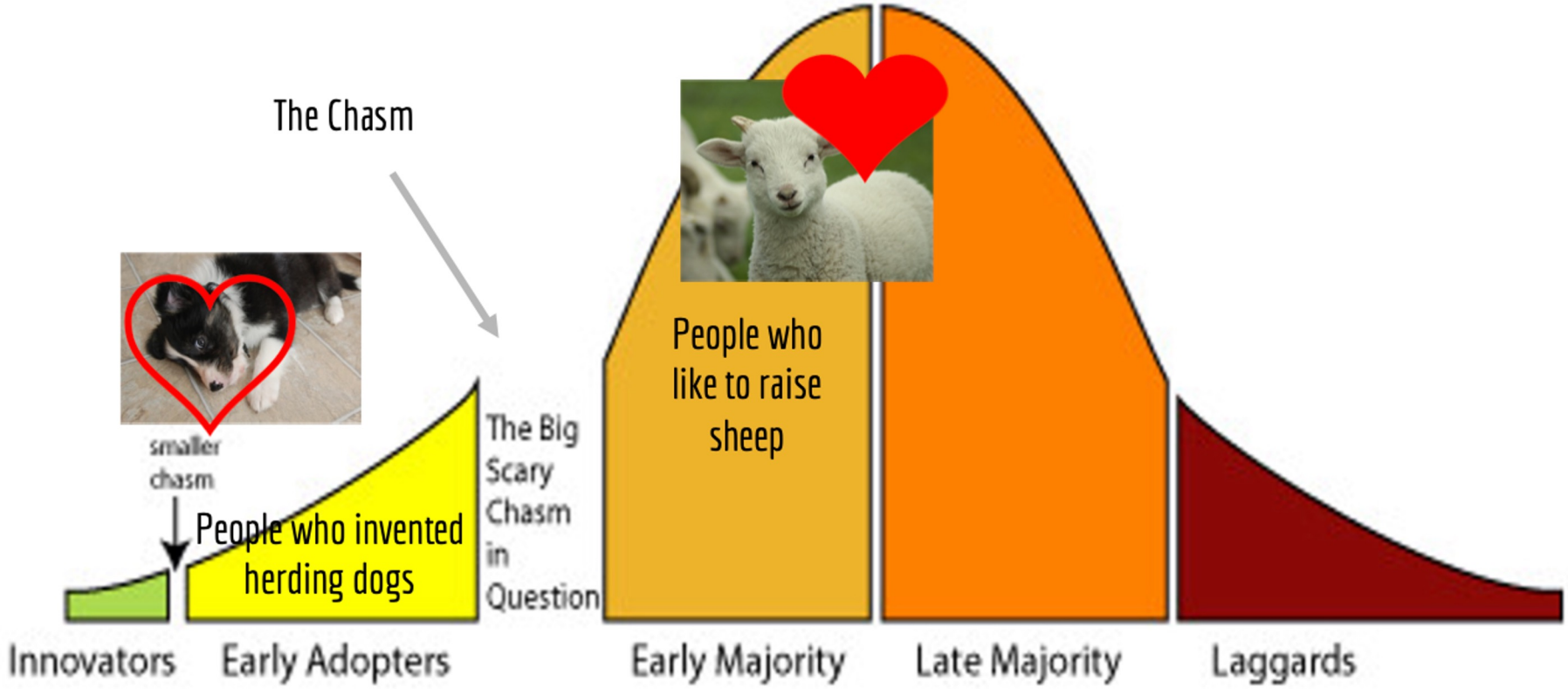


TASK



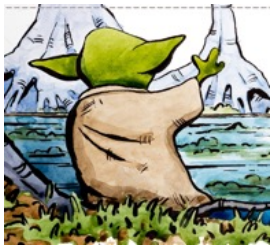

"Crossing the Chasm"

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.

How do you cross the Chasm?

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

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

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

		YEAR 1				YEAR 2				YEAR 3			
		2023		2024				2025				2026	
TASK NO.	TASK TITLE	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2
1	Develop Mentor Community												
1.1	Outreach to new NMFS mentors												
1.2	Openscapes Mentors Cohort Activities												
1.3	Identify and Address Cross-Center Roadblocks												
2	Empower Research Teams												
2.1	Engage NMFS teams												
2.2	Openscapes Champions Cohorts												
2.3	Community Skills Workshops (R/Python)												
2.4	Pathways to Open Science (DEI)												
3	Amplify Open Science Leaders												
3.1	Community engagement & webinars												
3.2	Coordinate, comms, tag-ups												

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

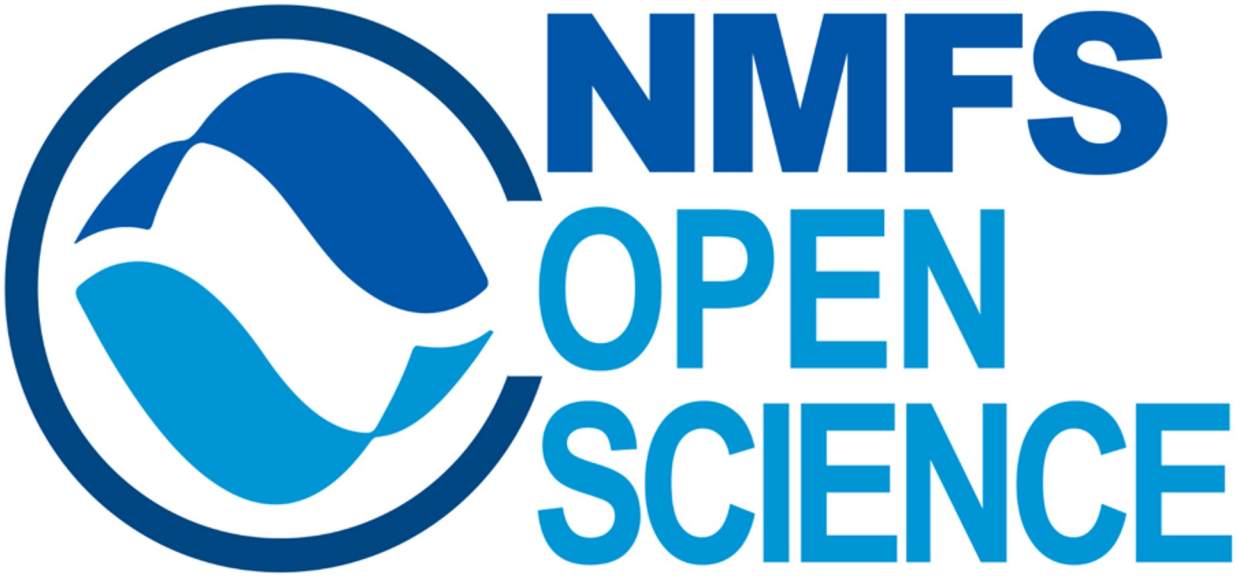
Supporting the infrastructure for Open Science

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

posit™ Connect


NMFS
OPEN
SCIENCE