

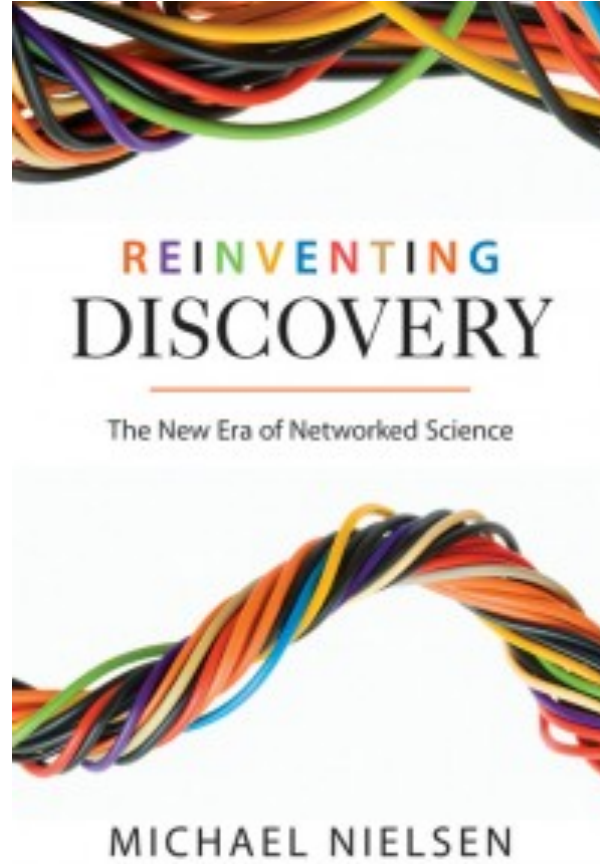
# Open and Reproducible Science

*Data Science for Coral Reefs*

*CRESCYNT - Data Rescue Workshop – 2018*

Adapted by Julien Brun from Mark Schildhauer, SNAPP training 2015

crowdsourcing



collective intelligence

micro-expertise

transdisciplinary

# Why Open Science now?

- Technology is available (World Wide Web)
- Growing politicization of science:  
need for transparency
- Importance of large-scale/interdisciplinary science
- Efficiencies in re-using or sharing available data,  
code

*A return to fundamental premise of science:  
objective, repeatable, replicable, “general”*

# Need for Open Science

*"It is essential that the scientific community work urgently to make standards for analyzing, reporting, providing access to, and stewardship of research data operational..."*

*Failure to make research data and related information accessible not only impedes science, it also breeds conflicts."*

Ralph J. Cicerone, President of U.S. National Academy of Sciences ("Ensuring Integrity in Science", Science 5Feb2010, p. 624.)

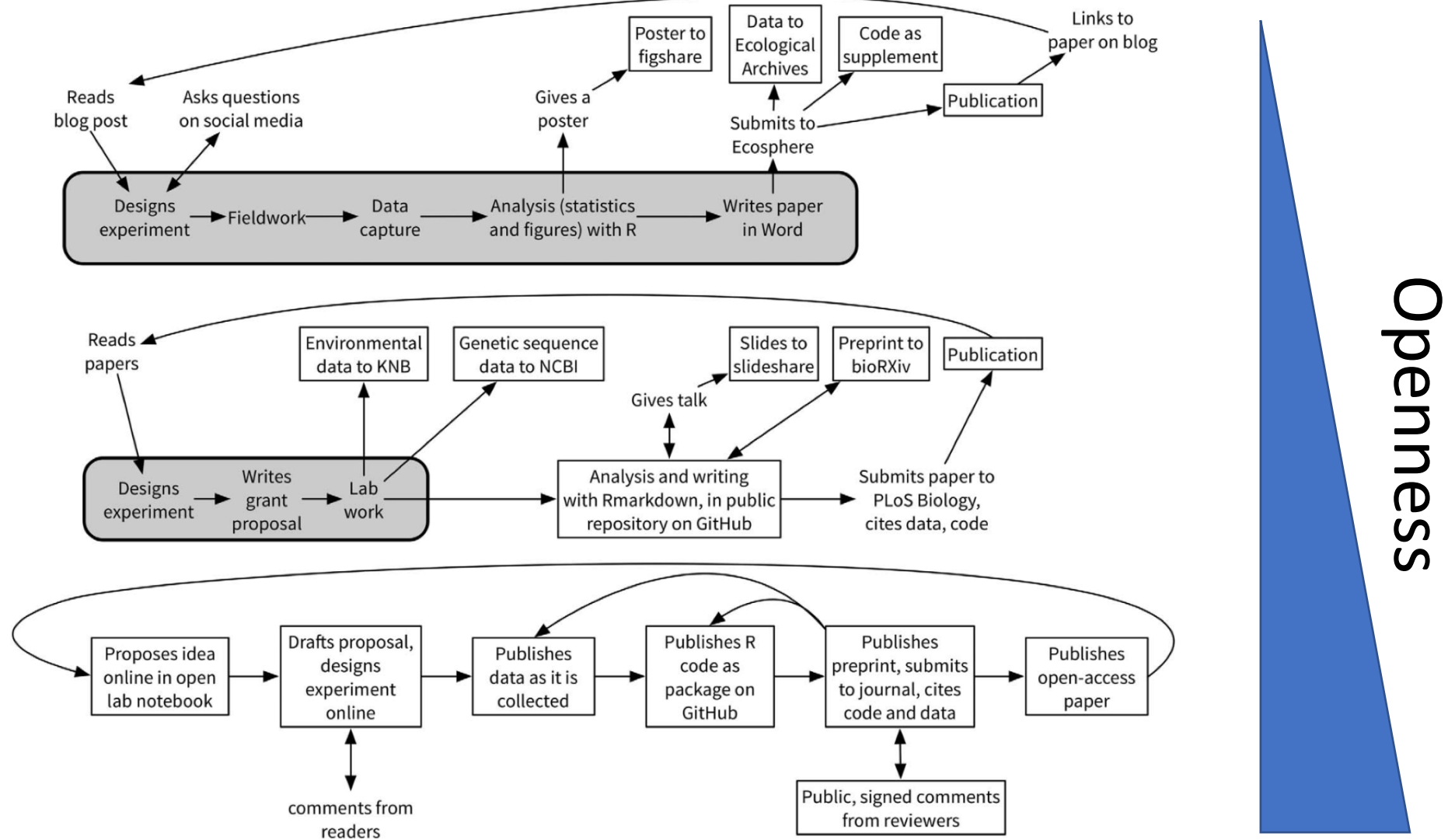
# Open Science

- ✧ Open Access
- ✧ Open Data
- ✧ Open Source
- ✧ Open Notebook

# Open Science

- ✧ Open Access
- ✧ Open Data **and Products**
- ✧ Open Source
- ✧ Open Notebook

# Open Science Workflows



Openness

# Open Access

Rapid, highly affordable or free  
access to the latest scientific findings

--

for everyone/anyone





# Open Data

Rapid, highly affordable or free access to ALL the data supporting scientific findings

Open data is data that are:

- Properly licensed for re-use
- Accessible w/o gates (e.g., paywall, login)
- Use open formats (formats you can work with)

# FAIR Principles

- Findability
- Accessibility
- Interoperability
- Reusability

# Open Notebook

Rapid, highly affordable or free access to the design details, logistical considerations, analytical decision points and justifications, anecdotes/marginalia, etc. that support scientific findings

# Open Source

Rapid, highly affordable or free access to the algorithms and other code used in an analysis, to enable examination and verification of appropriateness, and ideally suitable for re-use

# Obstacles to Open Science?

- Lack of time and funds
- Lack of rewards: money, status, promotion
- Breaking from traditional publication model
- Other concerns: scooping? poaching?

# Some Resources

Nielsen, Michael. 2011. *Reinventing Discovery: The New Era of Networked Science*. Princeton University Press. 208pp.

Hampton et al. 2015. *The Tao of open science for ecology*. *Ecosphere* 6(7), article 120.

Wilkinson et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship, *Sci. Data*.  
<https://doi.org/10.1038/sdata.2016.18>