

# Data Management for MBONs

A case study for data management in multi-agency federally funded projects

# Biological data are messy

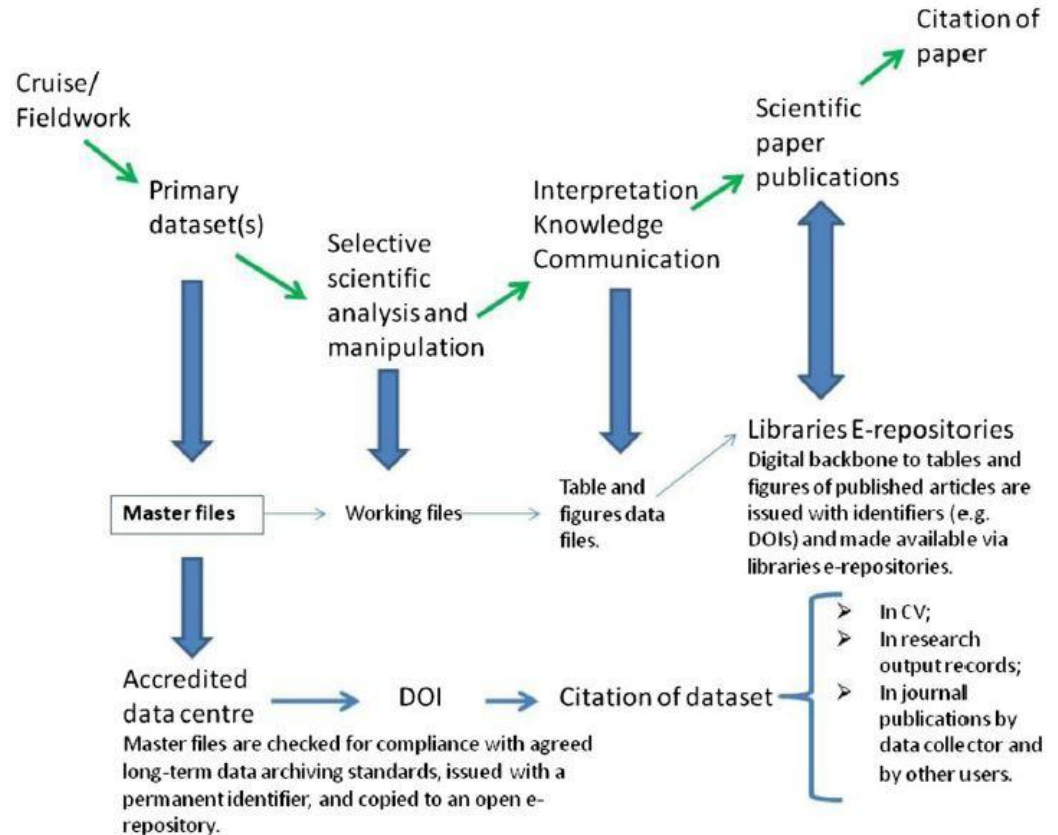


# Networked Science

Neilsen (2011) pointed out that the benefits of Open Data are not only an increase in data availability, but also a cultural change resulting in an interesting new approach to the conduct of science: “The reinvention of discovery is one of the great changes of our time. To historians looking back a hundred years from now, there will be two eras of science: prenetwork science, and networked science. We are living in a time of transition to the second era of science. But it’s going to be a bumpy transition, and there is a possibility it will fail or fall short of its potential.” (Gallagher et al. 2015)



**Fig. 1** End-to-end flow of data and information going from collection to publishing of data. SCOR/MBLWHOI/IODE (2014)

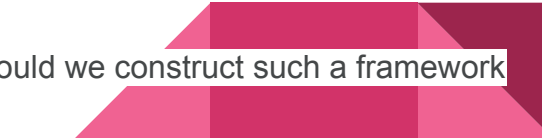


# Problem Statement

How do we provide guidance to multi-agency federally funded projects collecting biology data about adequate data management?



# Questions to Consider

- How can these data be best managed so that they are available for the long term?
  - Is managing the data through a commercial partner's data management system sufficient or appropriate for long-term preservation and availability of federally funded data assets?
  - What is the role of federal data centers like the National Centers for Environmental Information at NOAA or one or more NASA DAACs in these projects?
  - What guidance should federal program managers for the grants/funding opportunities be providing to the project PIs at this stage of projects already in motion?
  - What guidance should we be putting into future funding opportunities to ensure open, public access for federally funded data production?
  - If federal data centers are to take on a role for these and similar data, are they fully equipped to do so? Do they require additional funding from the project or funding programs to successfully take on the additional workload?
  - If federal grant-based projects are collaborative across agencies with multiple groups putting in funding, who gets responsibility for the data?
  - What might we learn from the developing ideas of the NIH Commons<sup>1</sup> idea? Could we construct such a framework across earth science and not just within one agency?
- 

# ad hoc Marine Biodata Task Team

- Representatives from federal and associated agencies collecting marine data (NASA, NOAA, BOEM, Navy, IOOS, Marine Mammal Commission, NSF, Smithsonian)
- Best Practices Document

