

Welcome to...

# Co-creating a shared framework for ocean data management

*Finding common ground on terminology*

27 Oct, 2024

WELCOME!



Data is not information,  
information is not knowledge,  
knowledge is not understanding,  
understanding is not wisdom

*~Clifford Stoll*

A pyramid diagram with four horizontal levels, each containing a word in white, bold, uppercase letters. The pyramid is light blue and is set against a dark blue background with a faint, abstract pattern of overlapping lines and shapes. The levels from top to bottom are: WISDOM, KNOWLEDGE, INFORMATION, and DATA.

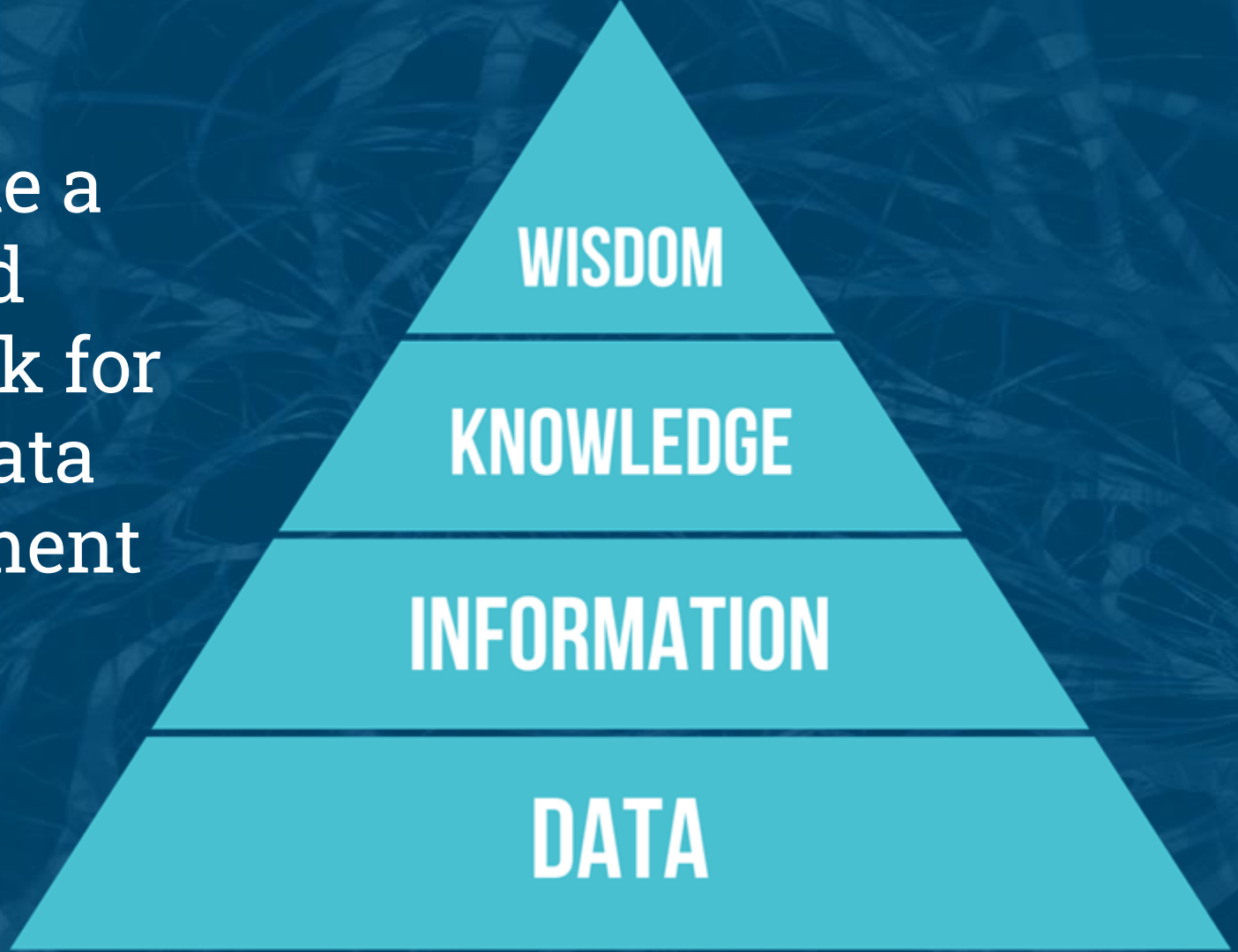
**WISDOM**

**KNOWLEDGE**

**INFORMATION**

**DATA**

Co-create a  
shared  
framework for  
ocean data  
management



# Session conveners



**Erin Satterthwaite**, Scripps Institution of Oceanography, UCSD



**Jeanette Gann**, NOAA, Alaska Fisheries Science Center, Juneau



**Naomi Boon**, Tula Foundation



**Tim Van Der Stap**, Hakai Institute

# INTRODUCTIONS



# Introduction & Polls

Let's get to know each other a little better!

What do you consider 'data'?



Join at [slido.com #2316 866](https://slido.com/join/2316866)

What data type do you most work with?

How do you call the data lifecycle process?





slido



What science domain or data type do you have the most experience with? Or if you don't have any experience please write “none”

① Click **Present with Slido** or install our [Chrome extension](#) to activate this poll while presenting.

slido



# What do you consider data?

① Click **Present with Slido** or install our [Chrome extension](#) to activate this poll while presenting.

slido



What do you call the process of data collection to use/action (e.g., data lifecycle)?

① Click **Present with Slido** or install our [Chrome extension](#) to activate this poll while presenting.



# BACKGROUND & RATIONALE

---

# Goal

---

Bring together ocean professionals to develop a **shared framework & language for effective ocean data management**

# Workshop Context

---

- To our knowledge, there is not one consistent conceptual model for the data lifecycle in oceanography/ocean sciences
- Today, we will start with a basic framework and build up to a conceptual framework with key phrases, definitions, steps, important considerations and outputs

# Workshop outcomes & outputs

---

- Outcomes

- An agreed upon conceptual framework that will:
  - make the data lifecycle more concrete, tangible and practical
  - help people understand what is expected of them
  - be useful for different audiences and data types

- Outputs

- Distill outcomes from workshop into a tangible output e.g. summary report
- Publish the conceptual framework
- Follow up workshop? Online or PICES 2025

# A Basic Shared Framework

STAGES	OUTPUTS
Purpose	Goal
Data & sample collection acquisition	Raw data/samples
Data & sample processing/QA/QC	Processed data
Data analysis/modeling	Summarized/derived data
Data serving/archiving	Archived/served data
Data standardization	Standardized data
Data visualization/use	Data products/models
	Metadata?



# Agenda

---

9:30 AM - 10:10 AM | Invited Presentations

Steve Diggs, University of California: Office of The President (UCOP) –  
Overview of climate-related data issues across the UC system

Prof. Chunhua Han, National Marine Data and Information Service-China  
(NMDIS)– Methods and practices of ocean data lifecycle management in China

10:10 AM - 10:40 AM | Conceptual Model Part 1: Discussion & Activity

10:40 AM - 11:00 AM | Break

11:00 AM - 12:15 PM | Conceptual Model Part 2: Discussion & Activity

12:15 PM - 12:30 PM | Closing & Next Steps



*Agenda also on Whova!*

# INVITED PRESENTATIONS



# Invited presentations

---



**Steve Diggs**, University of California: Office of The President (UCOP)

*Overview of climate-related data issues across the UC system*



**Chunhua Han**, National Marine Data and Information Service-China (NMDIS)

*Methods and practices of ocean data lifecycle management in China*



## **Steve Diggs:**

*California Digital Library (CDL), CA, USA*

Steve Diggs is a Research Data Specialist with the University of California Office of the President (UCOP), serving as the Senior Product Manager for the data publishing portfolio at the University of California Curation Center (UC3). He previously was the Technical Director for the Hydrographic Data Office at Scripps Institution of Oceanography, UC San Diego.

With decades of experience in designing and operating full-stack information systems, Steve focuses on making data more valuable, discoverable, and accessible to the research community. He is also active in international data and science teams, serving on the Board on Research Data and Information (BRDI) under the National Academies of Sciences, Engineering, and Medicine (NASEM) and the UN Ocean Decade Data Coordination Committee. Steve's extensive work with researchers in the field and lab gives him a deep understanding of the challenges and opportunities in managing scholarly data.



**Chunhua Han:**

*National Marine Data and Information Service (NMDIS)*

Prof. Chunhua Han is a senior research fellow at National Marine Data and Information Service, Ministry of Natural Resources of China. She is mainly engaged in scientific research and operational work in the fields of national marine information resources management, and marine mineral resources big data application technology research and development. She has undertaken more than 10 provincial and ministerial projects, presided over the development of 2 marine industry standards, written more than 20 departmental rules, regulations and technical specifications, edited 2 monographs, published over 30 papers, and compiled more than 30 important marine data management technical reports. She is now taking the leading role in innovating concepts and technical methods of classifying marine big data resources for the purposes of building and operating the marine big data resources system in China.

# Conceptual Model Interactive Activity & Discussion

---

# Conceptual Model Activity

---

- Conceptualizing the process
- Defining the parts of the process (“Stages”)
  - What different components to a data lifecycle or workflow are there / do you record?
  - What are some examples of outputs from each stage?
  - ‘Vote’ for the term(s) that resonate with your approach
- What are some cross-cutting themes to be considered across (all) the stages?
- Reflections & discussion
  - Ideas Box - what stages or outputs are we missing?
  - Are all recorded stages and outputs applicable to all data types? If not, how do they differ? Any discrepancies?

*Aim: Have a general outline of the data process, including the main parts of the process (‘stages’) as well as the outputs from each stage.*



# Conceptual Model Activity

- Conceptual diagram example:

## Part of the process ("Stage"): **Data Collection**

*Do you use other terminology for this stage of the process? If so, what are other words you use for this stage of the process - please stick them here, e.g.:*

Data acquisition

## Output: **raw data**

*Are there any other terms or outputs for this stage of the process? If so, stick them here, e.g.:*

*Unprocessed data, source data, raw observations, original data, input*





A serene landscape featuring a clear blue sky with scattered white clouds and a calm blue ocean. The horizon line is visible in the lower third of the image. The word "Break" is centered in the sky area, with a small blue horizontal line underneath it.

Break

# Conceptual Model: Parts of the Process

---

*Discussion*

# QUESTIONS, ERRORS, ADDITIONS

- Is anything missing?



- Are there any strong feelings about the stages?



- Does this general flow capture the intricacies of your data type?



# Co-developing Definitions & Important considerations of the Conceptual Model

---

*Group time*

Final shared model



*Discussion*

# Closing & next steps

---

## *Next steps*

---

- A **summary document** detailing the shared conceptual model (e.g., published in PICES Press)
- **Publish the conceptual model** with a url & doi to continue to iterate on with our communities
- Follow-up **PICES session/workshop proposal?** - *due Nov 15*
- **Community & continued dialogue** around this shared oceanographic data lifecycle conceptual model
  - Make the stages and steps more concrete
  - Provide a shared vocabulary
- Other thoughts?

*Please let us know if you are especially interested in helping with any of these*



Thank  
you!