Fishing for Clarity: Knowledge Modeling to Support Cross-Organizational Collaboration and Data Sharing About Salmon Escapement

To be held at the AFS Washington-British Columbia Chapter Meeting, Vancouver BC March 9, 2025 (all day)

Workshop Description

In this day-long workshop (sponsored by Fisheries and Oceans Canada, PSMFC StreamNet, and PNAMP), participants will learn about using a knowledge modeling approach to integrate data across multiple organizations engaged in monitoring salmon spawner abundance. Focusing on the challenge of being able to integrate salmon escapement observations across diverse entities, workshop attendees will develop a shared ontology describing core data requirements and concepts that underlie salmon escapement observations. The workshop will equip participants with the skills to build scalable, interoperable data models, an important component of enabling integrated data across organizations for better analysis and decision-making. By the end, attendees will leave with practical tools and a clear roadmap for advancing data sharing and collaboration in salmon management.

Target Audience: This workshop is ideal for professionals working in salmon monitoring, analysis and management.

Workshop Objectives: By the end of the workshop, attendees will be able to:

- Understand the key principles and methodologies of knowledge modeling.
- Learn how to build and apply an ontology to structure data for cross-organizational integration.
- Identify and address challenges in integrating heterogeneous data from different systems and organizations.
- Gain hands-on experience in creating and refining ontologies that support data interoperability and semantic clarity.
- Develop strategies for ensuring the scalability and sustainability of data models in collaborative environments.

Scope:

- Introduction to knowledge modeling and ontology concepts, with a focus on data integration challenges.
- Step-by-step guidance on designing a knowledge model tailored to cross-organizational data needs.
- Practical exercises on defining and structuring an ontology that reflects real-world data scenarios.
- Group discussions and case studies drilling into specifics of salmon escapement.
- Develop shared approach for continuing development and maintenance of a shared ontology

Participants will leave with the skills to:

- Build foundational knowledge models that can integrate disparate data sources.
- Create ontologies that provide a common framework for understanding data across organizations.
- Apply knowledge modeling to real-world scenarios, enhancing interoperability and reducing data silos.
- Contribute to the long-term success of collaborative data initiatives through a sustainable and adaptable knowledge framework.

DRAFT

PROPOSED Workshop Schedule (as of January 27, 2025)

8:00 AM - 9:00 AM – Coffee/Meet & Greet

9:00 AM - 9:10 AM – Opening Remarks
Facilitators: Jen Bayer (USGS/PNAMP) & Tom Bird (Fisheries and Oceans Canada)

- **Problem Statement**: Salmon management and research are often fragmented across different domains and stakeholders. Data are defined, scoped, and measured differently in local and broad-scale research contexts. This workshop aims to pilot a knowledge modeling approach to address these challenges.
- **Workshop Objectives:**
    - Introduce the general approach
    - Start working on a specific context (cross-border escapement data)
    - Develop a commitment to follow up and complete the work following the workshop

Outcome: Understand the workshop's context and objectives.

9:10 AM - 9:30 AM – Framing the Focus of the Workshop
Facilitators: Matt Jones (NCEAS) & Shirley Stephen (NCEAS)

- **Content:**
    - Introduction to ontologies: What they are and how they can help solve data integration challenges.
    - Overview of how to establish competency questions (CQs) and model controlled vocabularies (CVs).
    - Introduce the specific work we'll do during the workshop.

Outcome: Attendees understand the concept of knowledge modeling and the overall goal for the workshop.

9:30 AM - 10:15 AM – SESSION 1: Overview of the Salmon Research and Monitoring Domain
Facilitators: Salmon Domain Experts

- **Goals**:
    - Define the domain and focus on data comparisons (e.g., escapement data between BC and US salmon streams).
    - Identify competency questions: What do we want to learn from integrating data?
    - Define the scope: Focus on raw data related to escapement estimates as they offer the most potential for broader insights.
- **Discussion**: What data will we work with? Examples include NuSEDS data and others.

Outcome: Establish a shared understanding of the domain and the questions we aim to address.

10:15 AM - 10:30 AM – Break

10:30 AM - 12:00 PM – SESSION 2: Key Concepts and Terms Brainstorming
Format: Interactive Brainstorming and Group Discussion
Preparation: bring your controlled vocabularies and data dictionaries

- **Challenge**: Different datasets use different terms for similar concepts, so we need to align key terms and concepts to link data effectively.
- **Goal**: Identify and define vocabulary hierarchies and mappings between them.
- **Breakout Groups:**

DRAFT

    - Discuss and identify relevant concepts.
    - Define terms and relationships needed to link data to competency questions.
    - Use existing controlled vocabularies to inform discussion.

Outcome: Refine organization of mappings and terms, get user feedback on whether the mapping is sufficient to answer the competency questions

---

12:00 PM - 1:00 PM – Lunch

1:00 PM - 2:30 PM – SESSION 3: Formalizing ideas into formal structure with Ontology Software
Format: Hands-On Workshop

- **Challenge**: The concepts identified so far need a formal structure and computer language to show how they are related and to ensure clarity.
- **Goal**: Formalize the structure of the ontology, organizing terms into hierarchical relationships and aligning overlapping concepts. Concretize into a formal ontology language.
- **Facilitators**: Ontology Engineers
    - Part 1: Introduction to basic ontology design principles.
        - Use tools like Protégé or spreadsheets to organize terms.
    - Part 2: Group activity: Refine concepts into classes, subclasses, and relationships.

Outcome: Participants become familiar with how to use patterns designed in previous session to build re-usable ontologies for their own query needs.

2:30 PM - 2:45 PM – Break

2:45 PM - 4:15 PM – SESSION 4: Continue and share ontology building results
Format: Group work and presentation
Goal: Have participants walk away with confidence that we are going to have a scalable method for comparing and integrate terms across datasets

- **Challenge**: We have seen the concepts and methods, what do we need to do to make them FAIR, how do we improve adoption?
- **Goal**: Discuss strategies implementation
    - Approaches to hosting terms and definitions (e.g., using PURLs or w3id).
    - Plan a scalable method to compare and integrate terms across datasets.
    - Consider how to create data dictionaries and conduct semantic similarity comparisons.
    - Demonstrate how to ground external vocabularies in the workshop output ontology
- **Discussion**: Determine the next steps for development and identify key actions for follow-up.

Outcome: A roadmap for continuing the work

4:15 PM - 4:30 PM – Closing Remarks and Next Steps

- **Facilitators**: Tom Bird and Jen Bayer
- Review the workshop's key accomplishments and next steps for ongoing collaboration.

Outcome: Confirmed action plan for follow-up and next steps to advance the knowledge modeling approach.

Summary of Expected Workshop Outcomes

By the end of the workshop, attendees will:

1. Have a shared understanding of the salmon research and monitoring domain, focusing on escapement data.
2. Have collaboratively developed a rough ontology structure, identifying key terms and relationships.

DRAFT

3. Have a clear plan for continuing the knowledge modeling work and harmonizing data across domains.
4. Commit to supporting the ongoing development of this work and its application to broader research and management objectives.

DRAFT