Announcement of Workshop

International Salmon Data Laboratory (ISDL)  
2019-01-25

Pacific Salmon Commission Boardroom  
600 - 1155 Robson Street, Vancouver, Canada

# Background

This workshop immediately follows a separate workshop, also sponsored by the International Year of the Salmon (IYS), wherein ecologists will investigate the status and trends in salmon around the world. They will identify a series of data sets to be assembled and integrated, then discuss using them to relate spatial and temporal patterns in salmon dynamics to habitat changes from anthropic global warming. Predicting salmon populations will requires mobilizing extensive ecological data and applying new analyses. Better decisions, based on that better knowledge, will require better decision-support products and practices.

WHAT the ecologists need exceeds the capabilities of the technologies presently accessible[[1]](#footnote-1) to them. We will identify HOW those needs can be met with modern technology for data processing, analysis, and presentation.

The problems involve the standardization and integration of new and historical datasets, plus metadata about practices and precision, plus myriad habitat indicators. Analyses of the integrated data will build understanding of the mechanisms underlying salmon dynamics, but implementation of these new practices will required automation and user-friendly tools. The intention is that more complete, up-to-date, effectively communicated ecological knowledge about salmon will mean better decisions about salmon habitats and fisheries, thereby improving the resilience of salmon populations in a rapidly changing world.

# Goals of THIS Workshop

* Consider examples, provided by salmon ecologists, of critical issues, historical datasets, proposed analyses, desired products, emerging practices, and future requirements. Determine what these imply for data processing.
* Specify the next generation of technology, tools, and practices for collecting, integrating, analyzing, and communicating salmon information. This design will be valuable throughout the environmental sciences.
* Identify means to lower barriers to implementing these new practices and striking new collaborations throughout the international salmon community.
* Decide on the strategy, plan, and budget required to effect this transformation.
* Identify experiments by this data laboratory that will produce the irresistible examples required for wide adoption.

# Longer-Term goals of ISDL – Preliminary

## Breakthrough

* 1. Mobilization. Deliver technologies to support the identification, assembly, standardization, and integration of diverse and complicated historical datasets into the Salmon Knowledge Graph.
  2. Ontology. Assist the salmon *community of practice* with standard names for parts, practices, ideas,.
  3. Schema. Design and implement a graph schema sufficient to encompass ecological complexity. May involve multiple graphs (layers).
  4. Habitats. Mobilization and integration of fine-scale habitat indicators to examine the mechanisms causing events and trends in salmon survival, distribution, and traits.
  5. Analyses. Demonstrate statistical analysis (typically R) in context of a knowledge graph.
  6. Communication. Effective communication of data and results to diverse stakeholders and decision-makers. Interactive graphics and decision support products from salmon knowledge graph.

## Collaboration

* 1. Knowledge management and collaboration support. For an international community of practice.
  2. On-line workgroups. In the context of a salmon knowledge graph (2.1).
  3. Data-sharing arrangements. Safe and rewarding.
  4. Best appropriate practices. Evolving via shared experiences from around the world, possibly captured as annotated workflow diagrams (3.4).

## Operationalization

* 1. Predictive. Meet data management requirements for projections about the future of salmon, e.g. recovery potential analysis (RPA).
  2. Efficiency. Maintenance of data streams feeding data lakes (3.4).
  3. Facilitate new models. Preparation for population viability analysis, individual-based models, salmon evolution, colonization,.
  4. Automation. Workflows and GUIs for data processing that bounces between R and neo4j.
  5. Field data. Immediately mobilized via cloud computing with attendant automation of QC and rewards.
  6. Collection platforms. Voluminous collections from genomics and genetics, new sensors, drones and robots (AUVs, sailbots), citizen science, etc.
  7. Priorities. Projects and workgroups.

## Adoption

* 1. Attractive. New efficiency in data collection, processing, and presentation.
  2. Easy. Reduce barriers to adoption, support participation by ecologists, facilitate ease of use: automation, interactive workflows, webpages, documentation,.
  3. Accessible. International participation: cloud services, localization,.
  4. Rewarding. Modern analyses that are easy-to-use ,yet safe, and levered by integrated data. Interactive visualization. Effective decision support products and practices.

# For more information

Scott Akenhead

[Scott@S4S.com](mailto:Scott@S4S.com)

1.250.210.4410

# See also

[Presentation of ISDL at GraphConnect 2018, New York](https://www.youtube.com/watch?v=ZNWEBoduACk&feature=youtu.be)

<salmondatalab.slack.com>

<github.com/int-salmon-data-lab>

[international year of salmon workshop on salmon status and-trends](https://yearofthesalmon.org/salmon_event/international-year-of-salmon-iys-workshop-on-salmon-status-and-trends/)

# Agenda – Penultimate.

Morning agenda: fish-heads meet tech-heads

Short presentations, each with generous discussion, for each of five steps in information flow for salmon, from field to decisions. See following table.

* Questions and discussions:
* Unfettered ideas about the future of salmon data processing.
* What are the current practices, what are the trends, what is possible?
* Transition to what, exactly, and how, exactly?
* What are the impediments, requirements, show-stoppers, friction, politics?
* What are the rewards? E.g., efficiency, better “decision support products,” ability to address critical problems identified in preceding IYS Salmon Status and Trends workshop.

Afternoon agenda: The Experiments as befits a “data laboratory.”

* Create a 4-year work plan (2019-2022) with projects, deliverables, budget, benefits, and budget.
* Define projects and workgroups for immediate “experiments” that will encourage adoption of new practices.
* Define the success stories and proof-of-concept required by funders.

# Presentations (Proposed)

|  |  |  |
| --- | --- | --- |
| Information Step | Person | Topic |
| Introduction | Mark Saunders  Director, International Year of the Salmon NPAFC, Vancouver BC | IYS Objectives, Participation, Highlights, and Time-Line. [LINK](https://yearofthesalmon.org/) |
| Kelly Chapman Vancouver Island University Powell River BC | A Précis of the Preceding IYS Workshop: *Salmon Status and Trends* |
| Scott Akenhead Pacific Biological Station Nanaimo BC | “A vision for ISDL” |
| Step 1. Data Capture | Matt Deniston Sitka Technology Portland OR | “Go Directly to the Cloud: GeoOptix™ as an Example of Better Data Capture and Better Information Flow” |
| Step 2. Data Assembly | Bruce Patten Fishery & Assessment Data, DFO Pacific Region Pacific Biological Station, Nanaimo BC | “An overview of salmon data processing in DFO Pacific Region.” |
| Step 3. Data Integration | Jeff Morris Neo4j, Inc. San Mateo CA | “Neo4j Uniquely Enables Data Integration and Modern Analyses.” [LINK](https://neo4j.com/blog/graph-algorithms-neo4j-connections-drive-discoveries) |
| John Song,  Systum Inc. Los Angeles CA | “Data: Current Trends and Tools for the Transformation, Storage, and Analyses of Disparate Data Sets.” |
| Step 4. Analysis | Sue Grant Status of Salmon Program, DFO Pacific Region Annacis Island, Richmond BC | “An Introduction to DFO’s State of the Salmon (SOS) Program.” |
| Step 5. Communication | Kelly Chapman Vancouver Island University Powell River BC | “Effective Communication Between Ecologists and Decision-Makers.” |

# Attendees – So Far

|  |  |  |
| --- | --- | --- |
| Person | Interest | Contact |
| Mark Saunders  Director, International Year of the Salmon | International Year of the Salmon [LINK](https://yearofthesalmon.org/) | msaunders@yearofthesalmon.org |
| Jeff Morris Neo4j, Inc., San Mateo CA | Graph database technologies and tools. | jeff@neo4j.com |
| John Song,  Systum Inc., Los Angeles CA | Data systems architecture and user interfaces. | jsong222@gmail.com |
| Matt Deniston CEO Sitka Technology, Portland OR | Technology and services support for environmental programs. | Matt@sitkatech.com |
| Keith Steele Systems Architect Sitka Technologies Group, Portland OR | Data systems architecture and user interfaces for environmental data. | Keith@sitkatech.com |
| Bruce Hecht Cambridge Innovation Center Cambridge MA | Corporate innovation and high-tech startups. | Bruce.Allen.Hecht@gmail.com [LINK](https://www.prnewswire.com/news-releases/seaahead-and-cic-partner-to-open-bostons-first-ocean-innovation-hub-300775371.html) ; [LINK](https://cic.com/) |
| Kelly Chapman Vancouver Island University, Powell River, BC | Knowledge transfer. | Kelly.Chapman@gmail.com |
| Scott Akenhead DFO, Pacific Biological Station, Nanaimo, BC | a vision for ISDL; data integration to enable new statistical models. | Scott@s4s.com |
| Bruce Patten Head, Fishery & Assessment Data  DFO, Pacific Biological Station, Nanaimo, BC | Fisheries and ecology data processing in DFO. | Bruce.Patten@dfo-mpo.gc.ca |
| Jim Irvine, Research Scientist DFO, Pacific Biological Station, Nanaimo, BC | IYS Salmon Status and Trends workshops. [LINK](https://www.researchgate.net/profile/James_Irvine5) | James.Irvine@dfo-mpo.gc.ca |
| Kim Hyatt  Research Scientist DFO, Pacific Biological Station, Nanaimo, BC | Extrapolating from intensively studied salmon populations to many salmon populations [LINK](https://www.researchgate.net/profile/Kim_Hyatt) | Kim.Hyatt@dfo-mpo.gc.ca |
| Vladimir Radchenko, Executive Director North Pacific Anadromous Fish Commission Vancouver BC | Assembling international data and knowledge regarding Pacific salmon [LINK](https://npafc.org/) | VLRad@napfc.org |
| Gérald Chaput Research Scientist Gulf Fisheries Centre, Moncton, NB | Canadian Atlantic Salmon [LINK](https://www.researchgate.net/profile/Gerald_Chaput) | Gerald.Chaput@dfo-mpo.gc.ca |
| Sue Grant Head, Status of Salmon Program DFO, Annacis Island, Richmond, BC | New tools for interactive analysis and visualization of salmon data | Sue.Grant@dfo-mpo.gc.ca |
| Catherine Michielsens Director of Modeling and Data Management Pacific Salmon Commission, Vancouver, BC | Advanced statistical models for salmon population dynamics and fisheries forecasting. | Michielsens@psc.org |
| Merran Hague Quantitative Fisheries Biologist Pacific Salmon Commission, Vancouver, BC | Salmon data collection and processing for salmon fisheries management decisions | Hague@psc.org |
| Mark McMillan Database manager Pacific Salmon Commission, Vancouver, BC | Salmon database designs, including biological observations | McMillan@psc.org |
| Andrew Munro  Fisheries Scientist  Alaska Dept. Fish & Game, Anchorage, AK | “Southern Alaska Salmon and People:” an experiment in salmon data assembly. | [Andrew.Munro@Alaska.gov](mailto:Andrew.Munro@Alaska.gov) 1.907.267.2260 |
| Nathan Bendriem North Pacific Anadromous Fish Commission Vancouver BC | International Year of the Salmon | NBendriem@yearofthesalmon.org |
| Stephanie Taylor North Pacific Anadromous Fish Commission Vancouver BC | International Year of the Salmon | STaylor@npafc.org |
| Bruce Baxter  Fishery & Assessment Data  DFO, Pacific Biological Station, Nanaimo, BC | Fisheries databases in DFO. | Bruce.Baxter@dfo-mpo.gc.ca |
| Shelee Hamilton Fishery & Assessment Data  DFO, Pacific Biological Station, Nanaimo, BC | Fisheries databases in DFO. | Shelee.Hamilton@dfo-mpo.gc.ca |
| Jason Parsley  Fishery & Assessment Data  DFO, Pacific Biological Station, Nanaimo, BC | Fisheries databases in DFO. | Jason.Parsley@dfo-mpo.gc.ca |
| Steve Schut Fishery & Assessment Data  DFO, Pacific Biological Station, Nanaimo, BC | Fisheries databases in DFO. | Steve.Schut@dfo-mpo.gc.ca |
|  |  |  |

1. available and useable without unwelcome training. [↑](#footnote-ref-1)