

# Soundscape Atlas

Yvan's presentation



# Data Management: Metadata

Importance of instrument metadata for data reuse/re-analysis

Processed data/data products - tight coupling between them, source data, metadata - valuable for discovery?

Linking species data to data chunks - identifying source of species identification

User centered design - for discovery and for metadata creation

Automatic metadata generation from instruments (SensorML?)



# Data Management: Infrastructure and Workflows

Understanding of OAI-PMH vs API for data access

GIS ‘endpoint’ as a valuable MERIDIAN ‘product’

Importance of clever redirect away from non-public data

CIOOS interface collaboration?

Use of existing profiles (e.g., ORCID? User profiles?) to simplify ingest

“Fit for purpose” data



# Focus Areas, Collaborations and Future opportunities I

## Focus Areas

- Training new users in acoustic data
- Provide regulations, standards and clear guidelines for acoustic data
- new/different ways of analysing and visualizing acoustic data -> Machine Learning
- Ensure sustainable curation of metadata
- Support research community/data providers addressing data management needs
- Provide Big Data solutions -> management, analysis, visualisation
- Become a mediator to inform public and decision makers



# Focus Areas, Collaborations and Future opportunities II

## Collaborations

- CIOOS
- Industry
- Government -> e.g. DFO
- FRDR (and Libraries in general)
- Research Groups, NGO's, government, industry -> in relation to use cases



# Focus Areas, Collaborations and Future opportunities III

## Future Opportunities

- Data Store (? part of CIOOS ?)
- Tackle the topics that are difficult to deal with by CIOOS -> e.g. Big Data
- Sustainable MERIDIAN
- Support decision makers -> provide tools for non-experts
- Be go-to organisation for acoustic data



# Sound Detection and Classification

- **Data**
  - Limitations in quantity and quality of annotated data
  - Data augmentation techniques might help here
    - Physically motivated augmentation using sound propagation model
    - Deep learning techniques (GANs, Style Transfer, Sequence generation with RNNs, etc)

Goal: provide an open database for training deep learning models (the ImageNet of marine acoustics)
- **Network structures**
  - Combine computer vision (CNNs) and Speech recognition (LSTMs) strategies (like human analysts do)
  - Goal: provide flexible models that can be specialized as necessary
    - Open source library for model development
    - Pre-trained models for transfer-learning/fine tuning
- **Workshop**
  - Organize a sound detection and classification workshop next year



# Acoustic Data Analysis and Tools

- **Visualization tools**

- "Data driven" (Time series, scatter plots, ...)
- Directional (long-term) data
- Take advantage of open-source libraries: **Plotly (Python)**, Shiny (R), Leaflet (JavaScript), ...
- Learn from what's already out there (Oceans 2.0, ...)

- **Analysis tools**

- Does MERIDIAN not provide data storage?!
- Analysis tools will only be useful if combined with data serving services!
  - Sandbox (for ONC data only) and CIOOS
- Focus on tools that integrate data from different sources
- Specific ideas for data-analysis products:
  - Lightweight algorithms
  - Citizen science?!
  - Apply Deep Learning to sonar data (e.g. FORCE in Bay of Fundy)
  - Glider trigger (e.g., do not resurface if a boat is right above you!)
  - Data decimation

- Do not reinvent the wheel

- Industry involvement



**DALHOUSIE**  
**UNIVERSITY**

