Image Data Work Group Meeting #1

March 31, 2025 | Meeting Summary

The Image Data Work Group is one of two Work Groups supporting ROSA's Data Governance Program in 2025. The Image Data Work Group is focused on all image data (video and still) from any source.

Work Group Goals

- Where should ROSA recommend people publish fisheries-related data?
- How should data be organized to maximize discovery and reusability?
- What else can ROSA do to support data governance for this data type?

Potential topics for recommendations

- Documentation (including metadata)
- Metadata standards and recommended repositories
- How to organize and structure datasets within a 'project'
- Any other specific info to be included (i.e. special fields, tags)
- Catalog of experimental designs used
- How to be good data managers for our future selves
- How to preserve and publish data to answer regional or cumulative impact questions

We had eight participants, in addition to staff from ROSA and Intertidal Agency.

We started by mapping the landscape of image data sources. We had a detailed discussion of the device capabilities, data products, uses, and destinations for habitat cams ("hab cams") and drop cams, including SPI/PV cameras. We were not able to dig as deeply into data outputs from ROVs and BRUV but captured some notes on what might be similar across all four data sources. Some overarching topics:

- Project design may be very similar between offshore energy habitat assessment and stock assessment data collection, but there may also be important differences in experimental design and other survey elements.
- A lot of species and other information in images isn't analyzed or reported to the client if it's out of scope for the project. There may be information in the original image data that could be extracted later for other research purposes.
- Even for well-documented datasets, if you don't know the data well it's hard to know how to use it.
    - An image quality flag, automated or entered manually, could be useful for reuse.
- Review processes range from fully automated (AI processing) to manual review and annotation by human reviewers, and projects may apply a mix of automation and human review.
- Video files are very large and currently stored on local hard drives (cheaper) and/or cloud storage (can be expensive). Clients generally only ask for derived data products and selected stills or video clips; they don't provide funding for long-term storage or extra processing.
- Technology is changing quickly so standards need to be flexible and adaptive to avoid locking in obsolete tech.

Potential opportunities for ROSA to support:

- Create a data product / metadata catalog that points back to raw data and contacts to request access. This would require minimum metadata standards for all data to be included in the catalog.
- Design the framework for a regional / cumulative impacts analysis using image data and map out data needs from that structured report.
- A pooled video library that improves automation for all partners (including AI providers) and increases the amount of information that can be captured from any image file.