Cloud-based national on-line services to annotate and analyse underwater imagery

Roger Proctor (IMOS-AODN), Tim Langlois (UWA), Ariell Friedman (Greybits), Sebastien Mancini, Xavier Hoenner (IMOS-AODN) & Brendan Davey (TPAC)
• IMOS is Australia’s Integrated Marine Observing System

• Australia has the third largest marine jurisdiction of any nation on Earth —13.86 million km$^2$

• Maritime activity is predicted to generate € 63 billion/year by 2020

• IMOS undertakes systematic and sustained multi-disciplinary observations across this jurisdiction

• It turns these observations into data, products and analyses that can be freely used and reused (abiding by FAIR principles)

• IMOS is now seen as delivering ‘operational’ information

• It does this through the Australian Ocean Data Network (AODN)

What is IMOS?

Sebastien Mancini talk earlier today … and Tuesday 1145!
IMOS Autonomous Underwater Vehicle

30,000 Stereo images of benthos in typically 50m x 50m square
AUV image annotation carried out with Squidle on desktop

Get started:
ANNOTATING
collection using chosen annotation scheme

...OR choose from:
"Magical Suggestions"!
from automated machine learning algorithms

3 annotation schemes to choose from
Or machine learning algorithms
Another kind of imagery is video ...

one particular instrument is a

BRUV – Baited Remote Underwater Video

Australia is currently the world leader in remote underwater and diver stereo-video sampling. (e.g. stereo-BRUVs, stereo-DOVs, stereo-TV)

IMOS is looking to support this data collection platform
Why do this?

Body-size information is essential for:
- many fish/ecological metrics
- standardise sample unit area

Most stereo-image annotation (in Oz) is done with:
• EventMeasure (.EMObs)

Mono annotation
• FinPrint Annotator
• FISH_ROCK
• VMS
Analysis … also currently a desktop, process

Here’s Waisea from Wildlife Conservation Society Fiji about to get started

Each video records about one hour of footage, and typically takes many, many hours to analyse and extract the annotations

Courtesy: Jordan Goetze
Why is imagery important?

- Imagery fast becoming a tool of choice of State of Environment Reporting
- Enables fast, ‘cheap’ repeat sampling to assess health of area ($5-10k/BRUV)
- World-wide interest in its use
- Video == BIG DATA, laborious workflow, mostly desktop
- Because mostly desktop, sharing is difficult
- Cloud can address this and improve efficiency and uptake
Australian Research Data Cloud

National Research Infrastructure Roadmap


Marine Research Data Cloud

National service to annotate and analyse underwater imagery by leveraging existing software initiatives:

- **Squidle+**: exploration, management and annotation of georeferenced images & video
- **GlobalArchive**: exploration, sharing and querying of annotation data
- And establishing cloud online imagery repository
Key features of SQUIDLE+

**Flexible data storage:**
Sync with existing data storage infrastructure (i.e.: data linked from AODN). Avoids needing to copy and duplicate data. Takes minutes instead of days to import data into the system.

**Flexible, translatable annotation schemes:**
Users can define their own annotation schemes or select from existing ones, and can translate between them meaning all annotations can be viewed in a unified consistent framework.

**Collaborative / automated labeling**
Data can be annotated by different users with different skill levels and automated algorithms can be called upon to speed up the annotation process.

"**Media object**" annotation
Images, videos, mosaics, etc can all be annotated using the same consistent framework.
For synthesis and interpretation of annotation datasets there is a need for:

- Centralised agnostic data archive
- Standardised formats
  - Platform/Sampling information
- Future ready
- Quality control
- Ease of use and access

GlobalArchive offers

- Project sharing or open data
- Solving problems of data storage
- Added value through synthesis
- Ensuring time series data will be available
GlobalArchive

Flexible import of historical and modern annotation data

Direct import from leading stereo annotation software (SeaGIS - EventMeasure)

Working with other annotation initiatives to ensure data transferability & interoperability (GlobalFinPrint)

FinPrint

22,000 BRUVs worldwide, https://globalfinprint.org/
A National Service for Underwater Imagery

Imagery stored here
- Still imagery
- BRUVs
- DOVs
- Towed video
- Etc.

Annotation done here

EventMeasure - fish annotation

Squidle+ - Habitat annotation

Annotations stored here

GlobalArchive

Analytics done here

SoE Reporting

Marine Research Data Cloud 2018
TPAC Repository >500Tb for imagery and annotation files

Global Archive overview

Metadata

Optional Validation

Mandatory Validation

22-step validation e.g. species/region

Data sync/upload tool

Annotation files

Annotation software (Event/Measure)

Imagery

Users

Summaries

Reporting app
Data wrangling
  ○ Cleaning the data and ironing out inconsistencies is often laborious and manual

Data upload
  ○ Handles all the syncing and upload of annotation & imagery data

Access to original imagery
  ○ No standardised methods for upload and retrieval
  ○ Links to annotations are often not easily maintained
  ○ Often ends up archived on drives in boxes under someone's desk

Standardisation of annotation data
  ○ Usually difficult to reuse / validate between organisations
  ○ With links to original imagery
  ○ Future hooks to Squidle+
Datasync Tool

Organise data
select files on local computer / drive

Create metadata
using included tools to ensure consistency

Validate
cross check data files and imagery

Sync annotations
upload metadata and files to GlobalArchive

Upload imagery
to cloud storage repository
An example of the value of GlobalArchive
The Australian National BRUV synthesis
Started at a week-long workshop in early Feb 2018

• 32 BRUV researchers

• 6 Government institutions and 6 Universities

➢ 80% of available annotation analyses loaded into the cloud

➢ Represents an investment of ~$10M

Curtin, Deakin, Flinders, JCU, Utas, UWA, AIMS, CSIRO,
NSW-DPI_F, SA-DEWNR, WA-DBCA, WA-DPIRD_F
The Australian National BRUV synthesis

- 20,022 BRUV deployments
- 1,888 species
- 2,693,906 individual fish
- 660,481 length measurements
Initial synthesis: information of value to

- State of Environment reporting
- Potential for improved fisheries management
- Conservation from broad scale assessments
Workflow for State of Environment reporting app

- Users
- Global Archive
- Summaries
- Life history information
- Marine Science Cloud
- PostgreSQL
- State of Environment Reporting

Diagram showing the integration of users, global archive, summaries, life history information, and marine science cloud with PostgreSQL database and R programming language.
Using the AODN portal for State of Environment reporting

Click on a ‘blob’
Using the AODN portal for State of the Art reporting

AODN Open Access to Ocean Data

1. Select a Data Collection
2. Create a Subset
3. Download

Step 2: Create a Subset

Global Archive SoE Reporting - Metrics

Temporal
From Min Reset
To Max Reset

Management area
Port Stephens Great Lakes Marine Park

Helper Layers
- Australian Marine Parks
- Bathymetry Base Layer

Simple Base Layer
Port Stephens Great Lakes Marine Park
Number of Species - Fished (red); No Take (green)
Port Stephens Great Lakes Marine Park
Total abundance - Fished (red); No Take (green)
A National Service for Underwater Imagery

Repository >500Tb

- Imagery stored here
  - Still imagery
  - BRUVs
  - DOVs
  - Towed video
  - Etc.

User Upload Interface

Annotation done here

Annotations stored here

Analytics done here

SoE Reporting

Squidle+ - Habitat annotation

EventMeasure – fish annotation

GlobalArchive

Marine Research Data Cloud 2018