Australia's Scalable Drone Platform – What it is, Lessons Learnt, Moving Forward

Jens Klump, Timothy Brown, Steve Quenette, Owen Kaluza, Siddeswara Guru, Samitha Amarapathy, Chris Peters, Donald Hobern, and Gerhard Weis

17 October 2023 | eResearch Australasia 2023

MINERAL RESOURCES
www.csiro.au

appf Australian Plant Phenomics Facility

AuScope

MONASH University

tern Ecosystem Research Infrastructure

Australian Research Data Commons

CSIRO

We acknowledge the Traditional Owners of the land, sea and waters, of the area that we live and work on across Australia. We acknowledge their continuing connection to their culture and we pay our respects to their Elders past and present.

Agenda

- Introduction and goals of the session (5 min)
- Introduction to the Australian Scalable Drone Cloud project (15 min)
- Use case examples (15 min)
- Discussion topics (25 min)
  - Building a better Australasian network
  - What capabilities would everyone like to see from ASDC?
    - Where are the gaps in existing tools/platforms
    - Existing workflows for new use-cases?
  - How to make use of synergies between projects in Australian and abroad
  - *Your topic here!*

Shared BoF notes

tinyurl.com/37ynu29u

Australian Scalable Drone Cloud

A national platform for drone research

Why do we need a national drone platform?

UAV's enable near-real-time monitoring to address the critical scale-gap between ground- and satellite-based observations but…

Issue 1: *Measurements* are not *data* and *data* is not *knowledge*

The Vision: *Better technology = better data = improved knowledge*

Measurements → **Data** → **Knowledge**

Reality: *Better technology = more complexity*

Farm/Field Data Capture	Photogrammetry	Raw 3D Data	Analysis (post-processing)	Publication
Mission planning	Data management	Data management	Which software?	How to publish data?
Hardware challenges	Which software?	How to view & interrogate data?	Data management	How to publish code & workflows?
Weather challenges	Which pipeline?	File formats	Which pipeline?	Long term storage
Data management	Which settings?	Collaboration & Sharing outputs	Pipeline config	Minting DOI's
	Output and settings tracking		Methods development	Data Discovery & Reuse
	Data storage		How to record analysis steps?	
	Compute requirements		How to view & interrogate data?	

For most applications, drone data is only a supporting component - Multitude of other data to integrate as well

Issue 2: *Everything about working drone data and processing pipelines is hard!*

- Drone and 3D Geospatial time-series data are:
  - Large and hard to work with (and getting larger and harder to work with)

<- 5Ha RGB flight, 750mb, 800 jpg's

<- 5Ha Multispectral flight = 15.7GB, 4,000 TIF files (20x larger)

- Processing workflows are complex; bespoke; often developed in-house
  - Expensive to develop, not easily shared / reused

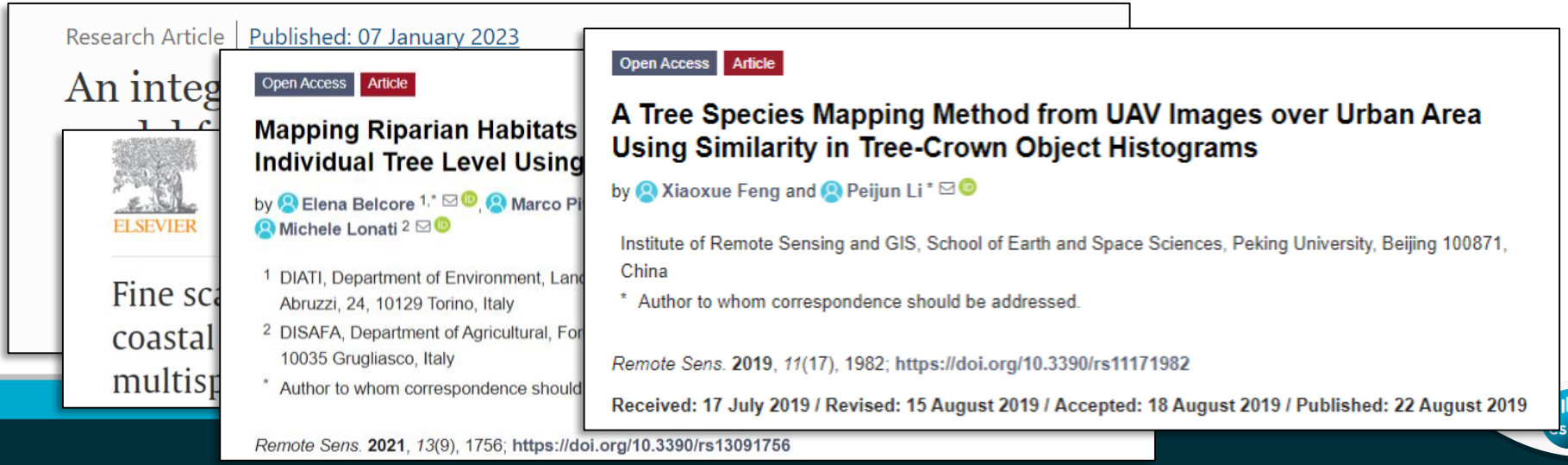
- Data not often/easily managed according to FAIR principles
  - *FAIR data = Findable, Accessible, Interoperable, and Reusable*[1]

1. nature.com/articles/sdata201618

Issue 3: *Published solutions are rarely immediately re-usable by others*

- Many pubs report an in-house “solution”; few provide that capability to the reader
- Have you actually “solved” anything if no one else can make use of the solution?
- Solve once for everyone:

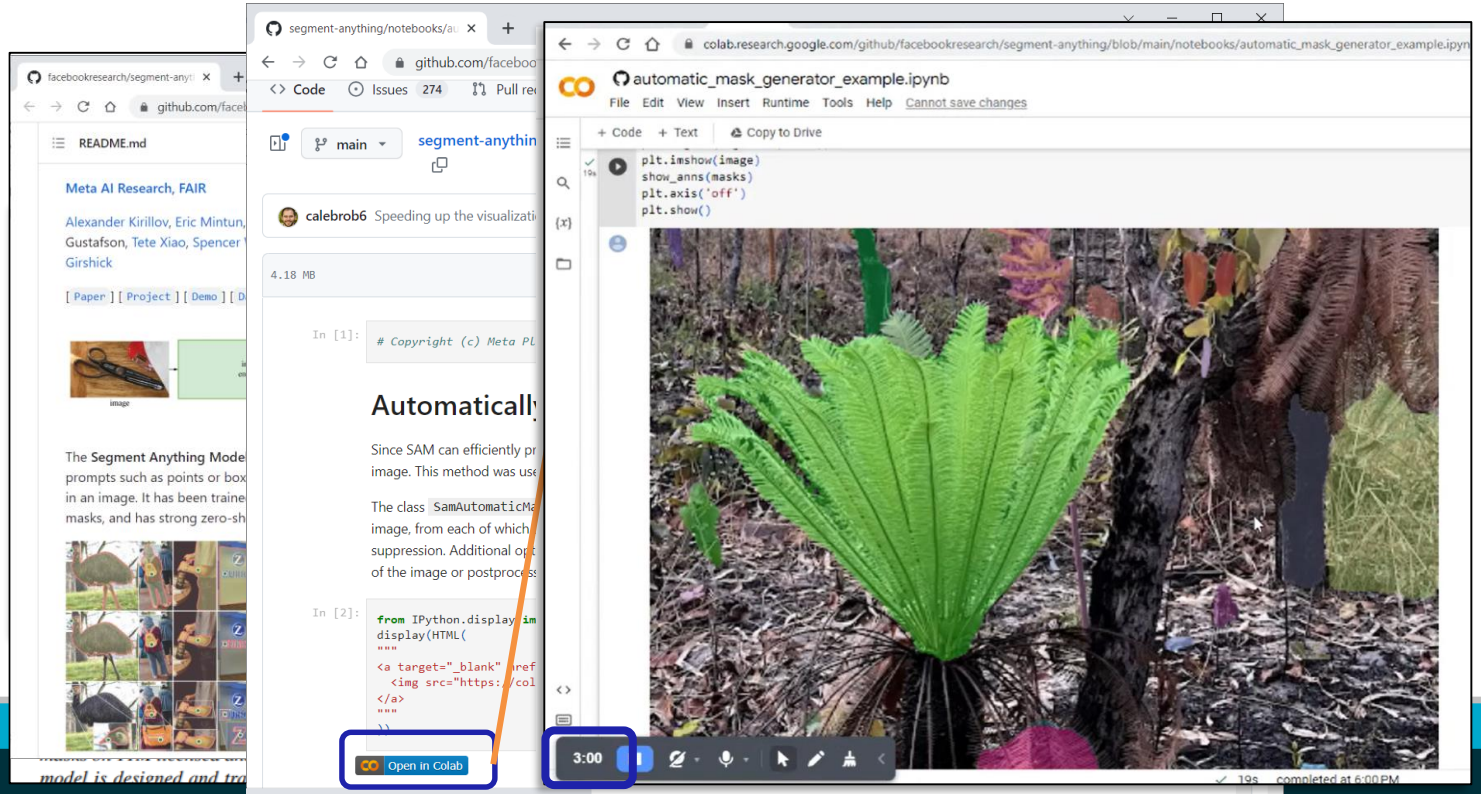
  Our work should be able to start where the last work finished; not have to re-implement the entire pipeline first
- Re-usability needs to become a community expectation



Community standards + Reusable code + Cloud infrastructure = immediate re-use of capability!

Example: Segmenting images – Long standing computer vision challenge

Meta's "Segment Anything" code: Apply an advanced ML segmentation model on your images with a $15K GPU for free in < 3 minutes from publication.

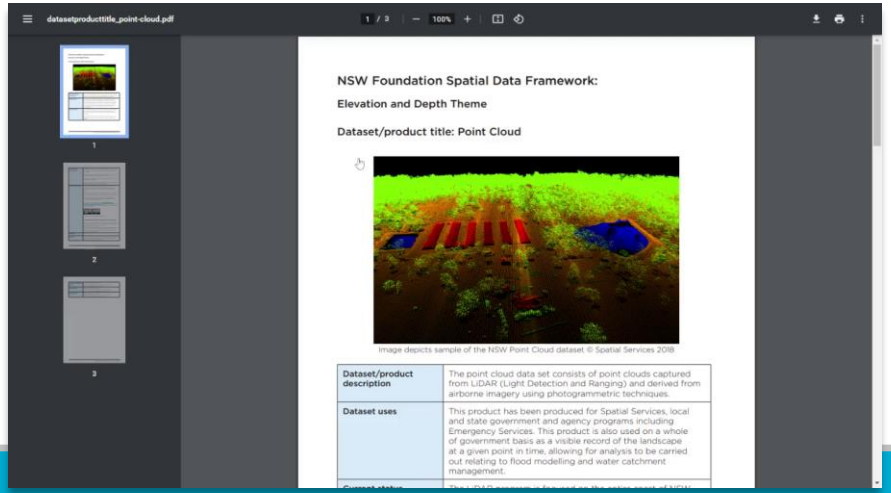


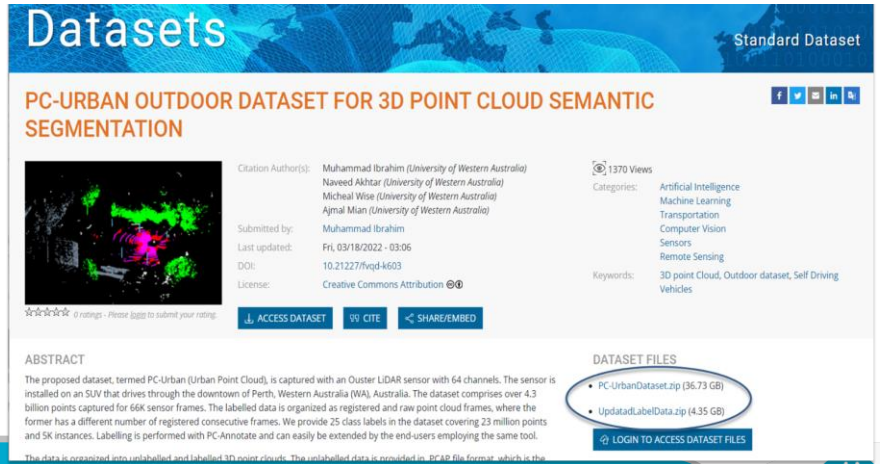
Issue 4: *Much of modern data isn't FAIR w/out platforms to support (re)use*

For large, 3D and time series data: Being able to see & interact with data is essential to access

- Not Accessible, Interoperable or Reusable w/o high end computer and expensive software
- Same issue only worse for sharing capabilities and pipelines (eg ML/AI workflows)

Data and workflows are not FAIR if they are functionally inaccessible





Australian Scalable Drone Cloud (ASDC) - A national platform for drone research

- Funding: Australian Research Data Commons (ARDC)
- 5 Project Partners from diverse research domains:
  - Lead: Monash University (Monash eResearch)
  - NCRIS: Australian Plant Phenomics Facility (APPF); Terrestrial Ecosystem Research Network (TERN); AuScope
  - CSIRO Mineral Resources (Perth); Reefs: Australian Institute of Marine Science (AIMS)

Goals

- Coalesce an integrated ecosystem for all steps in drone research: Capture -> publishing
- Lower barriers to entry and democratize access to needed capabilities
  - "FAIR-from-capture"
  - Enable a "Solve once for everyone" approach
- Cross-domain enables national-scale data interoperability

Australian Institute of Marine Science

APPF Australian Plant Phenomics Facility

AuScope

Monash University

CSIRO

ARDC Australian Research Data Commons

tern Ecosystem Research Infrastructure

‘Platforms’ approach = adapt, adopt and link existing tools

Capability

Photogrammetry Data & User Management

File manager
Upload external files (LiDAR, other software)

Analysis

Visualisation

Publish data and workflows

Enabling software

ODM (Open Drone Map)

ODM

jupyter
Drone workbench (TERN CoESRA VL)

CESIUM

terria

ARDC Research Data Australia
TERN SHaRED

Integrated Platform

Cloud Native Architecture

kubernetes

Australian Research Cloud

ARDC Nectar Research Cloud

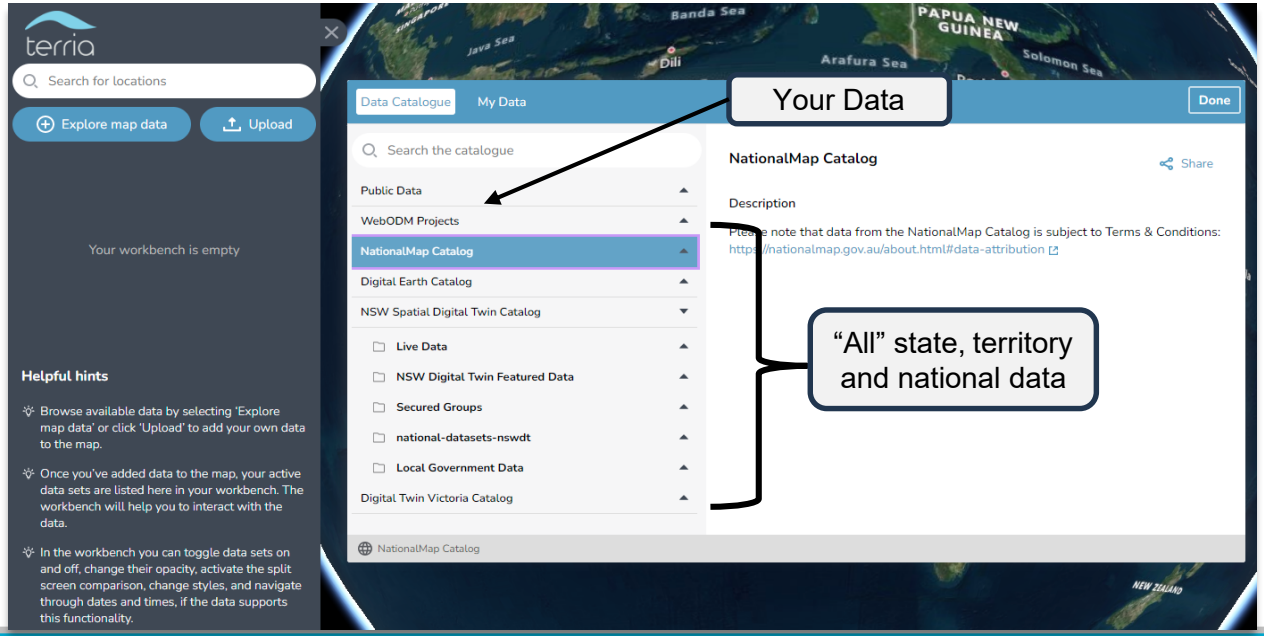
Commercial cloud (on roadmap)

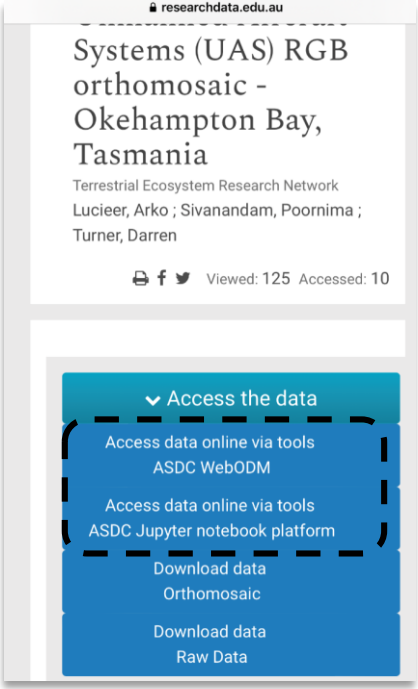
Microsoft Azure

amazon web services

Tackling Accessibility, Interoperability and Reusability

- Linking user's drone data with "cloud-native" data from <u>all</u> National and State repositories.
- Publish code, pipelines & data back to national data repositories

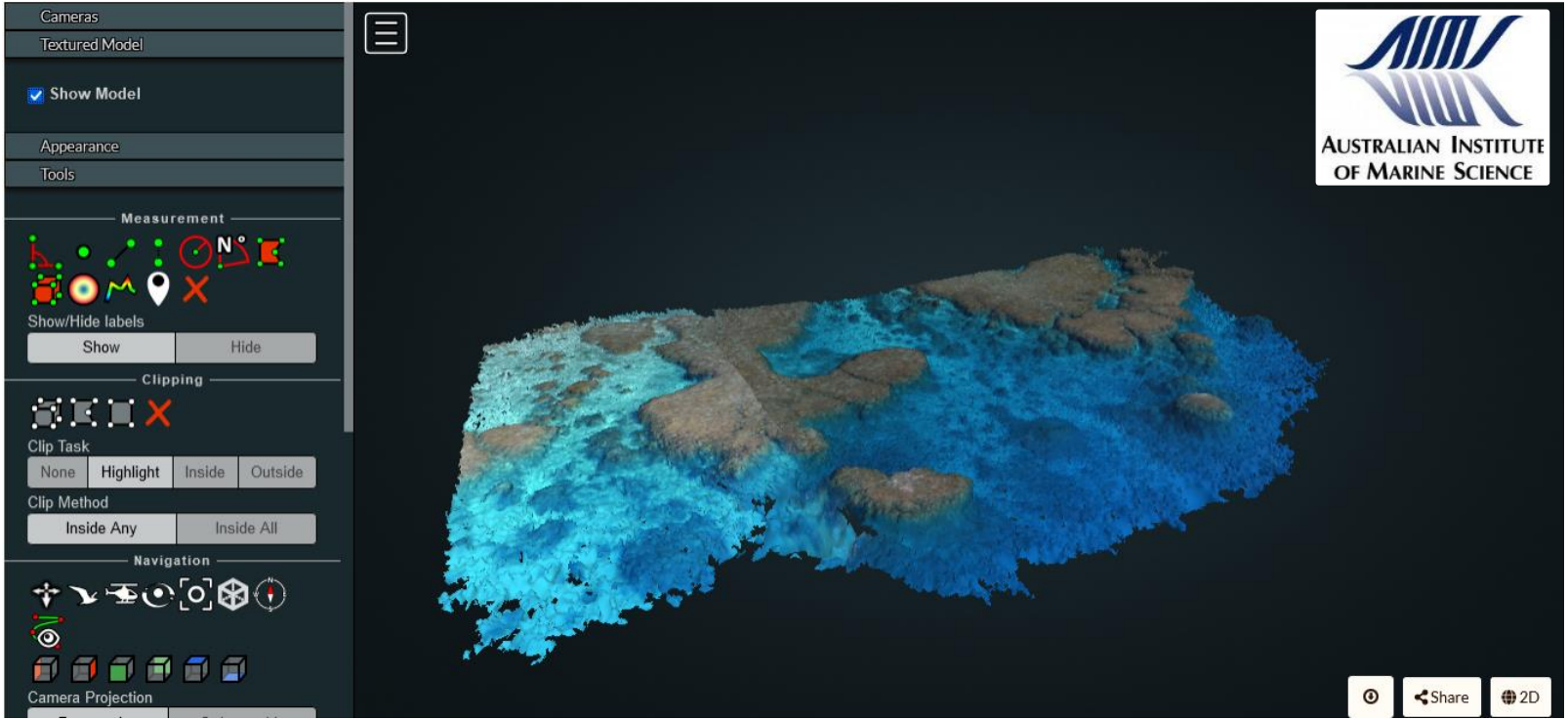




Use Case Examples

1. Reef monitoring
2. > Structural Geology
3. > Victorian Coastal Monitoring Program
4. > Weed detection
5. Wheat plot height

Reef Photogrammetry



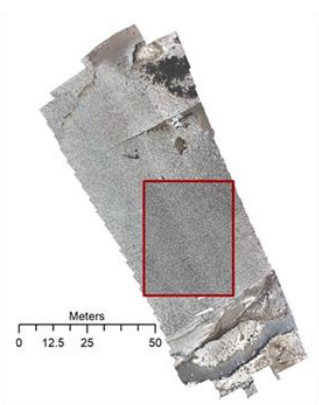
Structural Geology

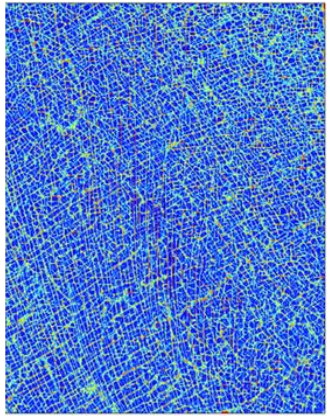
Orthophoto mosaic

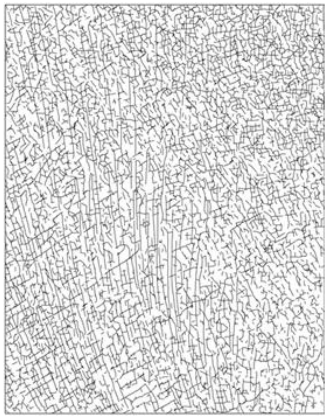
AOI

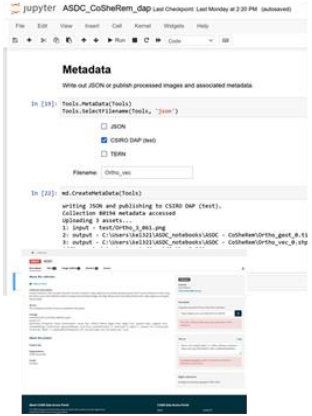
Intensity map

Vectorized features & metadata









Data Capture and Pre-processing

Data Processing, Analysis, and Visualisation

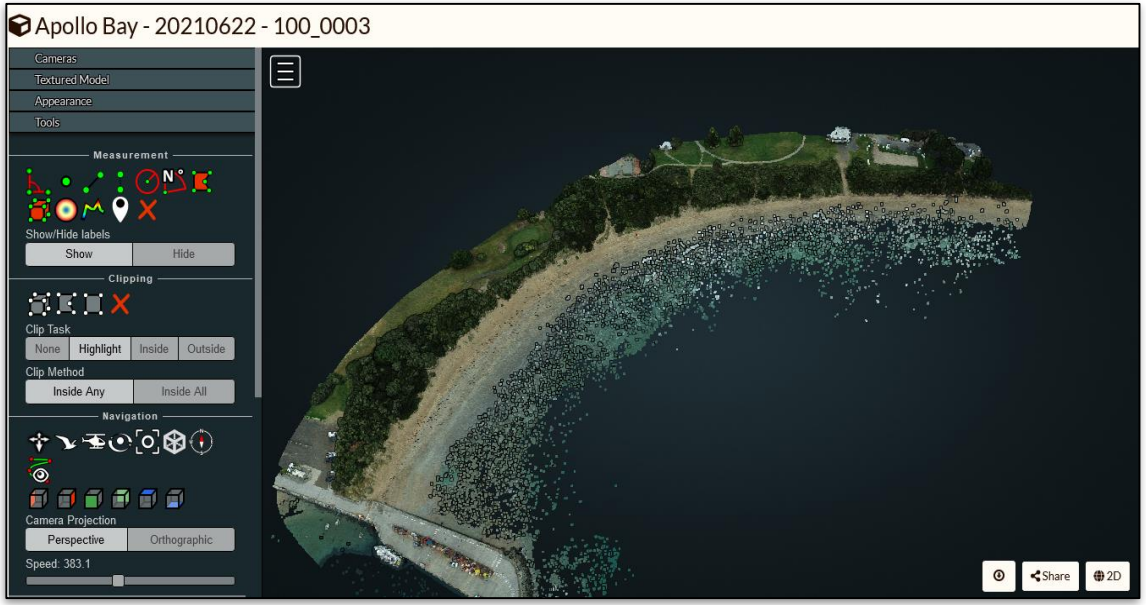
Data Archiving and Publication

Victoria Coastal Monitoring Program (VCMP)

Citizen Science project to using drones to monitor Victorian beaches

- >40 sites
- 700 surveys
- 180 citizens

- Data currently processed/hosted by propeller aero
  - Need long term, scalable, open source platform
  - *Data discovery/download*
  - *Code development and Methods development*



Learn more: *marineandcoasts.vic.gov.au/marine-and-coastal-knowledge/victorian-coastal-monitoring-program*

Machine Learning Pipelines for Weed Detection *(and everything else!)*

Creating an ecosystem of re-usable solutions

The challenge:

- 11 weeds cost Australian primary industries sector >> $400 million each year[1]
- Same species = amenable to ***"solve once for everyone"*** approach
- Many groups building solutions but no easy way to:
  - share solutions, algorithms or data, audit code, no integrated data management
  - find others working on same problem
  - build shared national collections of annotated images

Solution is both technical and social:

- Social (communities of practice):
  - People need to be onboard; supported to implement reusable solutions; aware of other's work
- Technical: National Research Infrastructure to support re-use + standards, documentation
  - ASDC platform provides user-friendly entry to using ML for aerial weed detection
  - Native integration between ASDC and AWS (for enterprise scaling)
  - Build library of reusable ML solutions

African Lovegrass

Orange Hawkweed

Siam Weed

Department of Primary Industries

Charles Sturt University

2pi Software

1. agrifutures.com.au/wp-content/uploads/2022/04/AgriFutures_Weeds-newsletter_Autumn-2022.pdf

Discussion

Discussions: Where are the gaps?

- Discussion topics (25 minutes)
  - Building a better Australasian network
  - What capabilities would everyone like to see from ASDC?
    - Where are the gaps in existing tools/platforms
    - Existing workflows for new use-cases?
  - How to make use of synergies between projects in Australian and abroad
  - *Your topic here!*

International Connections

RDA supports the proposed Small Uncrewed Aircraft and Autonomous Platforms Data WG through its TIGER programme.

The WG will develop recommendations for data and metadata best practices for interoperability with EOSC and other VREs.

RDA Small Uncrewed Aircraft Systems IG

Next meeting:

06 November 2023

12.00 - 13.30 UTC
18.00 - 19.30 AWST
21.00 - 22.30 AEDT

https://www.rd-alliance.org/new-rda-working-group-small-uncrewed-aircraft-and-autonomous-platforms-data-kick-meeting

Summary

- What's happening next?
- How to become involved?

We'd love to collaborate - how can we help solve your drone/data challenge(s)?

Jens Klump: jens.klump@csiro.au
Tim Brown: tim.brown@anu.edu.au

Thank you!

Australian Plant Phenomics Facility

AuScope

MONASH University

tern Ecosystem Research Infrastructure

Australian Research Data Commons

CSIRO

LEFTOVERS

ASDC next steps

- The ASDC project has been extended at a smaller scale until March 2024.
- The extension will be used to stabilise the technical operation of the platform and reduce technical debt.

Establishing Australia's Scalable Drone Cloud

Update with more succinct slide and modular design slide

Drones provide sensing capabilities that address the critical scale-gap between ground and satellite-based observations.

Australia's Scalable Drone Cloud (ASDC) offers a competitive advantage to researchers: lowering barriers to near-real-time & societally-relevant drone-derived information

The ASDC will:

- establish national best practice for drone data analytics, informed by
  - fundamental research applications,
  - industry applications, and
  - national applications
- coalesce novel and bespoke drone data analysis capabilities
  - co-design advanced tooling and methods (incl. point-clouds, machine-learning, best-in-breed tools etc)
  - cloud-native scale-out to society-scale applications (incl. data and campaign interoperability)
  - exemplar data-processing pipelines made reusable and accessible
- Accelerate the adoption of FAIR Principles for drone-acquired data, including FAIR-from-capture

A novel approach - a partnership between…

- three NCRIS capability areas - Australian Plant Phenomics Facility (APPF), AuScope and the Terrestrial Ecosystem Research Network (TERN),
- a university - Monash University (via the Monash Drone Discovery Platform and the Monash eResearch Centre), and
- a government research agency - CSIRO (piloted through CSIRO Mineral Resources)

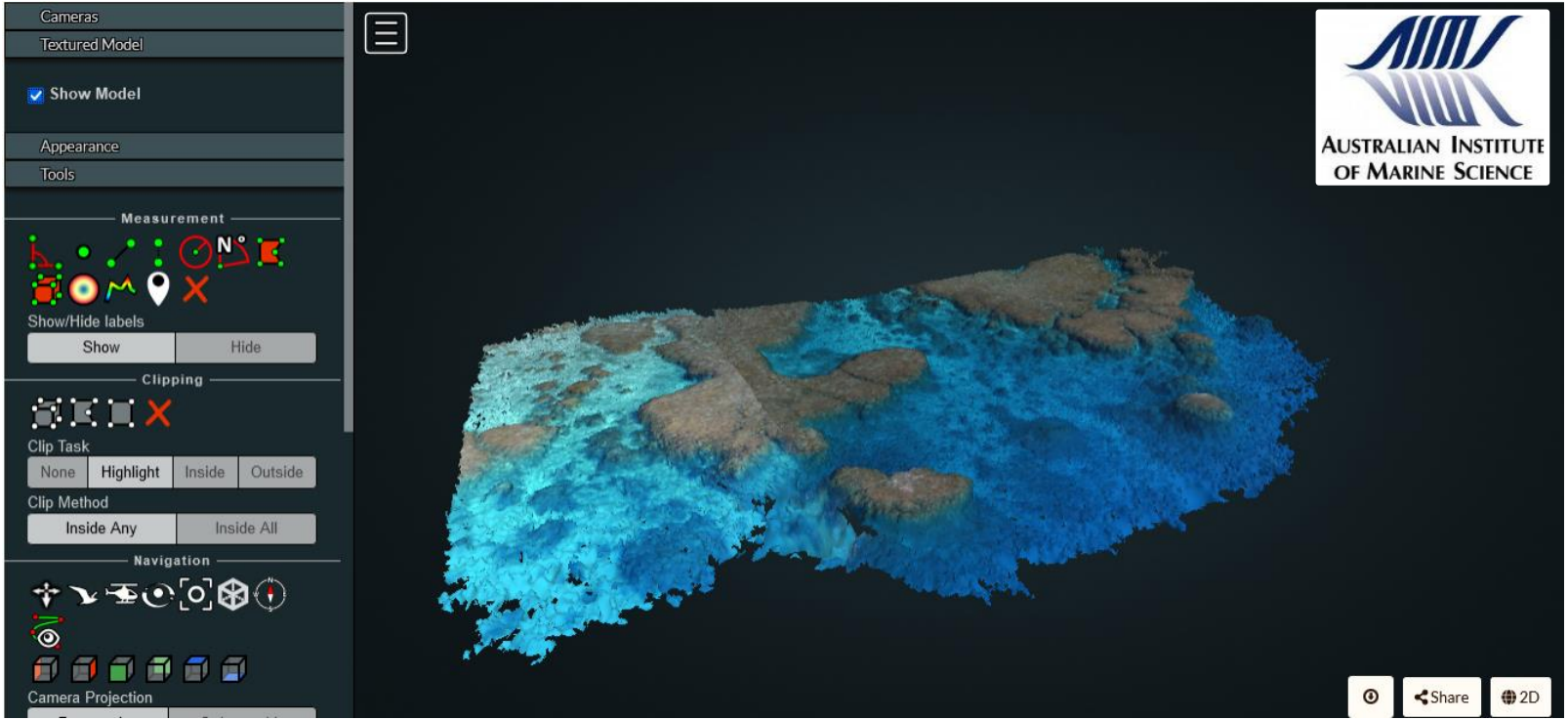
… to transform their already substantial drone data processing capabilities towards best-practice, automation and intelligent decision making, establishing a nationally accessible ecosystem for drone-related research and innovation.

Supported as a 3yr ARDC Platform project commencing 2020

http://www.asdc.org.au

APPF Australian Plant Phenomics Facility · AuScope · CSIRO · MONASH University · tern Ecosystem Research Infrastructure · ARDC Australian Research Data Commons

Reef Photogrammetry



Machine Learning Pipelines for Weed Detection *(and everything else!)*

TIM to update slide

The Challenge:

- African Lovegrass (ALG) is a widespread invasive in Australia but no quick low cost detection methods
- Conventional detection methods (visual surveys): expensive, slow, can ALG spread seed
- Machine learning (ML) methods are robust but not easy for wider usage

Solution:

- ASDC platform provides user-friendly entry to using ML for aerial weed detection
- Develop native integration between ASDC and Amazon Web services
- Reusable ML tool trained on ALG (but can be trained for any other plants)
- Goal: Any users interested in detecting a new plant can train the system, then ML new model rolled out for everyone (Conservation/landcare, local government associations, farmers, endangered plant surveys, etc)

Project Partners:

- 2PI Software, Bega; Snow Monaro, LLC, USydney

Wheat Plot Height

The Challenge:

- How can researchers and breeders measure plant biomass with semi-automated reproducible workflows?
- Existing methods are well established but none are yet implemented into a turn-key workflow

Solution:

- ASDC implemented the URAF plot height pipeline into a semi-automated pipeline with integrated FAIR data management.
- This makes it much easier for breeders to use the same standardized methods; better track and share data

Project Partners:

- Ramesh Segaran (Unmanned Research Aircraft Facility (URAF)
- WheatHub (ARC Industrial Transformation Research Hub)
- Australian Plant Phenomics Facility (APPF)