ARDC Platforms Project Progress Report

Project Information

This section is so you can provide an update on progress of the project against plan, as well as to flag changes that will need to be made to the project plan as a result of changing circumstances

1. Project Details:

Project title: GMRT-AusSeabed

Lead organisation: Geoscience Australia

Other organisations involved in project: Bureau of Meteorology, Deakin University, James Cook University, Lamont-Doherty Earth Observatory, CSIRO, Australian Antarctic Division

Contact person name: Kim Picard

Reporting period: 14 Aug to 1 Dec

2. Project Deliverables/Goals:

This platform will provide seamless, high-quality elevation and bathymetry data to the consistent standards required for oceanographic models and predictions. This work will accelerate research by reducing manual effort, avoiding duplication and removing the barrier of specialised skills needed for accurate bathymetry data.

The outputs, suitable for high-performance computing or desktop environments, ensure the service is scalable and accessible to the broader research community. The consistency, resolution, and processing capability delivered by this platform improve the accuracy, validity and reliability of modelling. As such, this platform will benefit coastal communities nationally by informing environmental management and risk mitigation.

3. Progress Against Project Milestones and Deliverables:

Update progress of project milestones and deliverables in the table below **or in a spreadsheet** if preferred.

Use the following criteria and colours to assign a status to each work package:

[Full spreadsheet here](#)

Status Key

On Track	On track and will be completed on time.
Minor Delay	Minor delays but will not impact the project completion date or other work packages.
Significant Delay	Planned completion date will impact the delivery of other work packages or the project completion date.
Completed	Completed.

Milestone / Deliverable / Work Package	Details of progress including explanation of any variation (include links to relevant documentation)	Agreed Due Date	Actual Completion Date or Expected Completion Date	Status (highlight colour as per table above)
MS1. End-User Platform Requirements Survey (Component 1)	**End user survey to assess the business requirements and potential application of the platform.**	05/03/2021	05/03/2021	
MS2. End-User Platform Requirements Report (Component 1)	**Summary report of the user survey findings.**	16/07/2021	15/08/2021	
MS3 – Platform Data Requirements and design report (Component 3)	**Multiple workshops and publication of associated reports.**	27/07/2021	15/08/2021	
MS5 – Datasets (Component 2)	Dataset delivery is in progress. At this stage, no impact on the project. Considering that publication of the datasets to the AusSeabed portal isn't critical to the success of the platform, we will delay this to next calendar year due to other priorities for publication	26/10/2021		
MS-4 Point Cloud conversion services (delivered through the design documents - MS7, task 53)	Milestone as a technical deliverable is no longer relevant with the technology choices made. Milestone instead has become a design milestone that formed part of the technical design document.	26/11/2021		
MS7 – Future state architecture and workflow process (Component 3)	Draft document has been provided to the ASB Steering Committee (for information) and provided to GMRT AusSeabed SC for review and feedback.	26/11/2021		
Milestone 8 – "MS8 – Cloud infrastructure set up locally for data ingestion into GMRT and subset datasets ingested (Component 3)" has been divided into 3 separate milestones reflecting the stages of development for the platform working with the selected technology.				
MS8a - Metadata harvesting development (row 40)	Initial harvesting scripts have been developed and are being refined. Expectation that the metadata harvesting will be a continuing evolving process as additional data is delivered.	29/11/2021	N/A	

Milestone / Deliverable / Work Package	Details of progress including explanation of any variation (include links to relevant documentation)	Agreed Due Date	Actual Completion Date or Expected Completion Date	Status (highlight colour as per table above)
MS8b - GA API development	Scoping has commenced.	17/12/2021	N/A	
MS8c - Configuration of tileDB ingestion service	Initial ingestion process has been trialed and is continuing to be refined to meet the desired specification and incoming data requirements.	17/12/2021	N/A	
MS6 - Connecting to Jupyter Hub	No change since last update	21/01/2022		
MS9 – GMRT grid composer modified to enable user-control setting (Component 3)	No change since last update	25/03/2022		
MS10 – Connection to Ocean Model and delivery of GMRT-AusSeabed Platform PoC	No change since last update	27/05/2022		
MS11 – Additional capabilities identified added to the Grid Composer (Component 3)	No change since last update	10/06/2022		

Progress against our KPIs are below:

KPI 1: Subset test datasets delivered to GA in a timely fashion for model development

As per variation of MS 5, dataset delivery and publication has been combined into 1 delivery mid-September.

KPI 2: Projects artefacts published on AusSeabed in a timely fashion

Our first three reports (User need analysis, workshop #1 and #2) have been published on AusSeabed website on 15 Aug. These have been delayed by 1 month from original plan due to a late decision to publish through the GA process.

KPI 3: Platform tested by key users returning positive and constructive feedbacks

n.a.

KPI 4: Platform code published to GitHub in a timely fashion

GitHub repositories are sitting within the AusSeabed organisation:

- https://github.com/ausseabed/reap-gsf (reading the GSF)
- https://github.com/ausseabed/bathy-datasets (dataset construction, metadata, geometry, tileDB schema outputs)

4. (Inter)dependencies

Have there been any changes to the project dependencies during the reporting period? Please list below.

None identified at this time

5. Risks

What risks arose during the reporting period and are there any changes to the level of risks or risk mitigation strategies identified in the project plan as a result? Use the table below or **link to a risk register**.

Risk	Controls/Mitigation Strategy	Residual Risk Rating (after controls are in place)	Risk Owner
Data Delivery and Publication delayed	Minor restructure of the development activities to allow more time for the data delivery to progress. As stated in MS5 of the progress report table, publication is not critical to the outcome of the project, thus delaying action is deemed reasonable.	Low	NL
Lack of technical design clarity	Technical design document is being circulated and presented through to the AusSeabed and ARDC Steering committee for oversight and review.	Low	NL

6. Issues/Challenges

Are there any issues or challenges that you have encountered during the reporting period?

Receiving the data from the data partners took longer than expected.
- Some delay was caused by GA not providing clear expectations on how datasets should be provided and data specifications, and miscommunications within the instruction on how to provide metadata to GA.
- Time has been spent producing templates and resolving inconsistencies in the metadata templates received.
- Getting data specifications right has been important and knowing how to communicate these specifications to data providers has been challenging.
- Some delay also due to receiving some datasets on hard drives, and minimal access to physical office space (COVID19 restrictions) to collect and upload this data.

6. Anything else you would like to tell us about the project:

Impact information

This section is provided so you can report on the impact of the project over time. The Government wishes to demonstrate the impact on researchers (and industry and the general public) of the NCRIS investment, and the ARDC Board is particularly concerned with being able to demonstrate the value of the investment in Research Platforms.

ARDC realises that a number of possible impact metrics are lag indicators, but wants to provide guidance about what metrics you should be collecting now to assist with telling a good impact story later.
Note: we understand you may not be able to provide all of these metrics initially.

7. Platform impact metrics

Impact metric (*add relevant details/links in next section)	Progress report 1	Progress report 2	Progress report 3	Progress report 4
Number of datasets published (preferably with record in Research Data Australia)*	0	o		
Number of publications using data generated by platform*	0	0		
Number of publications that cite /reference the platform*	0	2 ASB reports		
Social media mentions of platform*	7	0		
Blog posts/news articles that reference platform*	1	1 website page 1 promo video 1 ASB newsletter		
Number of users*	0	0		
Number of active users (over 90-day period)	0	0		
Number of jobs run (i.e. number of processes submitted through the Platform)	0	0		
Instruments supported	0	0		
Number of engagement/outreach activities (e.g. seminars, presentations, demos, meetings with new stakeholders)*	2+	3 presentations, 1 poster		
Number of researchers trained (online or face-to-face)*	0	0		
Availability of the platform (days in operation vs days offline)	0	0		

8. Links to publications and other communications

Published datasets generated by platform

Date	Link to a metadata record *or* Citation with DOI
Example: 20/10/2021	https://researchdata.ands.org.au/parkes-observations-project-semester-2010aprs/445704
	N/A

Publications using data generated by platform

Date	Citation with DOI
Example: July 1999	N. Paskin, "Toward unique identifiers," in *Proceedings of the IEEE*, vol. 87, no. 7, July 1999. https://doi.org/10.1109/5.771073
October 2021	Greenslade, D., Davies, G., Picard, K. 2021. GMRT-AusSeabed PL-019 User Needs Survey Summary: Setting User Requirements. Geoscience Australia, Canberra. http://pid.geoscience.gov.au/dataset/ga/145798
October 2021	Lennard, N., Sixsmith, J. 2021. GMRT AusSeabed PL019 Workshop-1: Attribute Values and Definitions. Geoscience Australia, Canberra. http://pid.geoscience.gov.au/dataset/ga/145794
October 2021	Lennard, N., Sixsmith, J. 2021. GMRT AusSeabed PL019 Workshop-2: Data Structure and Formats; Backend and Delivery. Geoscience Australia, Canberra. http://pid.geoscience.gov.au/dataset/ga/145795

Publications that cite/reference the platform

Date	Citation with DOI
Example: July 1999	*Example:* N. Paskin, "Toward unique identifiers," in *Proceedings of the IEEE*, vol. 87, no. 7, July 1999. https://doi.org/10.1109/5.771073
October 2021	Baldry, K., Picard, K., Leplastrier, A., Arnold, M., Lennard, N., Crossman, D., Evans, N., Mackay, K., Townsend, N., Yule, C. 2021. AusSeabed Annual Highlights Report 2020/21. Geoscience Australia, Canberra. https://dx.doi.org/10.26186/145959
October 2021	Leplastrier, A., Picard, K., Arnold, M. 2021. AusSeabed Annual Progress Report 2020/21. Geoscience Australia, Canberra. http://pid.geoscience.gov.au/dataset/ga/146010
October 2021	AusSeabed Steering Committee 2021. AusSeabed 2021/22 Work Plan. Geoscience Australia, Canberra. http://pid.geoscience.gov.au/dataset/ga/146009

Social media mentions of the platform

Link
Example: https://twitter.com/AusBiocommons/status/1251724453173538817?s=20
https://twitter.com/GeoscienceAus/status/1392615635155312643
https://twitter.com/SchmidtOcean/status/1394688223872491521
https://twitter.com/KimPicard8/status/1332038765351632896
https://twitter.com/ARDC_AU/status/1390131719349706752
https://www.linkedin.com/posts/australian-research-data-commons_as-a-professional-coastal-engineer-im-activity-6795897433063854080-0qd4
https://www.linkedin.com/posts/geoscience-australia_seabed-data-bathymetry-activity-6798382531285786624-Ehyi
https://www.linkedin.com/posts/geoscience-australia_last-week-our-team-from-ausseabed-were-found-activity-6740455331530346496-QiNA

Blog posts/news articles that reference the platform

Link
Example: https://twitter.com/AusBiocommons/status/1251724453173538817?s=20
The project page has been established by the ARDC (https://ardc.edu.au/project/gmrt-ausseabed/).
New webpage on AusSeabed: http://www.ausseabed.gov.au/gmrt
AusSeabed newsletter http://www.ausseabed.gov.au/__data/assets/pdf_file/0020/111746/AusSeabed-Newsletter-Oct-2021.pdf
Promotional video: https://www.youtube.com/watch?v=hUXjs-8tqgo&t=3s
Released a project launch article with the ARDC (https://ardc.edu.au/news/piecing-together-the-puzzle-of-australian-seabed-data/) and circulated through social media.

9. Details on engagement, outreach and training activities

Activity	Activity description	No. of participants	Date of activity
AODN Technical Advisory Group quarterly meeting		Unknown	May 21
Geohab 2021	An overview of the project was provided at the end of a general AusSeabed talk	Unknown	May 21
ASB Steering Committee presentation	Presented update to SC	14	04/11/2021
GMRT AusSeabed Steering Committee	Presented update to SC	16	17/11/2021
ASB Executive Board	Presented update to EB	7	24/11/2021
AODN TAG	Presented overview to AODN TAG	19	30/11/2021
ASB Quarterly Showcases	Presentation on the project in the AusSeabed Qtrly Showcase	50+	March, July, September
Public GA seminar https://www.youtube.com/watch?v=CCYkyYp8Uyw&list=PL0jP_ahe-BFmyBvlLgrRIPrMamu-oEpMP&index=11	Presented as part of an overview seminar on the progress of AusSeabed	220	25/08/21
Spotlight Presentation at the Forum for Operational Oceanography	Poster presentation and discussion to engage end-users for testing and feedback	TBC	22/11/2021
Present at GEBCO Map the Gaps symposium	Present as part of overall presentation on AusSeabed	TBC	29/22/2021
Abstract in preparation for Hydrospatial 22			Feb 22
Abstract in preparation for Locate 22			May 22

10. Implementing FAIR Platforms and Outputs

FAIR project outputs

Describe the project's progress in making the **project** outputs (publications, code, etc) FAIR.

Code repository for the project has been established within GitHub at:

- https://github.com/ausseabed/reap-gsf (reading the GSF)
- https://github.com/ausseabed/bathy-datasets (dataset construction, metadata, geometry, tileDB schema outputs)

FAIR platform outputs

The AusSeabed data portal and the GMRT synthesis, which serves FAIR and Open data nationally and globally. The GMRT AusSeabed project is leveraging these existing publication processes in these two established services to publish all provided data through this pathway, while also handing the data within the developing platform.

- To date, all datasets that have been provided have been added to the AusSeabed Publication Schedule and will be published in the new year
- The following technical decisions have been made with respect to the platform's technology choices:
    - o Metadata will be
        - delivered to meet ISO19115-3 catalogue standards
        - Stored within a JSON file format, according to the Spatio Temporal Asset Catalog (STAC), which is an open standard that is well adopted within the spatial community
        - The Open Data Cube is being used to manage the metadata, an open source initiative led by Geoscience Australia.
    - o Data is being harvested into a universal storage architecture that is fully open source. (www.tiledb.com)
    - o All bespoke development is being managed within public GitHub repositories
        - https://github.com/ausseabed/reap-gsf (reading the GSF)
        - https://github.com/ausseabed/bathy-datasets (dataset construction, metadata, geometry, tileDB schema outputs)

11. Details on users of the platform

If possible, provide a brief breakdown of where these users come from; for example, which institutions the core user groups are from, or how many users from each of the following broad categories: .edu.au, .gov.au, .org.au, .edu (international academic institutions), .com (gmails etc.), other.

We expect that we will engage 2900-3600 users per annum, based on AusSeabed and GMRT usage statistics. We expect a range of users from the academic, government and industry sectors.

We cannot provide a specific breakdown on where our users will come from as the platform is not live. However according to the 95 respondents to the end-user survey completed in Feb, users from all around Australia responded with highest percentage from WA, NSW, Vic and Qld, as well as internationally. All sectors were represented, with the most associating with private industry, Federal government, university and Research Institute (85%).

12. Impact stories

Please detail any research impacts or highlights here. Other impacts could include outcomes of industry collaborations, platform community activities leading to new policy or practice, etc.

The adoption of a universal storage architecture with tileDB represents a significant step forward for the community in the following ways:

1. Efficiencies to be gained by the user community
    a. Engaging with the data at an earlier stage of processing reduces the lead time and potentially the skill set required to access data ready for use in further processes.
    b. Interacting with data that is in a cloud performant storage format will enable more efficient work practices.
2. Data unlocked that was previously unavailable/inefficient to use
    a. Delivery of historical data that has traditionally been difficult and time consuming to work with can be moved to a performant storage architecture that is not tied to a commercial service provider.
3. Potential reduction in the cost of doing business
    a. Elevation data is a fundamental dataset and represents one of many contributing data inputs to many fields of research. Often, the end user of this data is a science domain expert, but may not be an elevation/seafloor data expert. They work with the data in a sporadic manner, for a specific purpose and likely would rely on expertise to achieve a valid outcome. There is currently no cost effective service available to service this type of user (who are the majority).