Toitū Te Whenua
Land Information
New Zealand

Enabling Integrated Marine Management

Part 1: Technical Proof of Concept

Appendix 2:

Use Case Summary: Te Papa Atawhai Department of Conservation

1 September 2022

Te Kāwanatanga o Aotearoa
New Zealand Government

Prepared by

Role	Name	Date
Science Advisor, Marine, Te Papa Atawhai Department of Conservation	Enrique Pardo	1-09-2022
Project Manager, Te Papa Atawhai Department of Conservation	Amelie Auge	1-09-2022

Contents

Te Papa Atawhai Department of Conservation: Mapping Marine Key Ecological Areas

Use case overview

Aotearoa New Zealand is nationally and internationally committed to manage its marine biodiversity; however, there is currently inconsistency on how different government bodies apply regulatory frameworks for managing marine and coastal biodiversity. One of the fundamental aspects to solve is how to effectively share existing data and allow potential users (e.g., regional councils, government agencies, etc) to analyse it in consistent ways. Te Papa Atawhai Department of Conservation (DOC) has been developing a set of criteria for identifying Key Ecological Areas (KEAs) that can inform decisions related to the management of marine biodiversity.

The DOC use case developed an online tool to delineate candidate KEAs within the marine environment using a set of criteria thresholds through the Datamesh. The use case explored how the Datamesh technology can support the development of an open-access analytical tool to delineate KEAs at national and regional scales. DOC, the National Institute of Water and Atmospheric Research (NIWA) and Oceanum collaborated during the use case to: (1) identify a series of test KEA criteria and sample of datasets to generate analyses; (2) link all required input datasets for the analyses to the Datamesh; (3) develop the analytical tool; and (4) conduct user testing internally and with regional councils. A full report of the use case where each step was described is available in Appendix 3 – DOC final report.

The main output of the use case was a prototype online application, the 'KEA App'. The KEA App is Datamesh-enabled and fully performs the mapping of KEAs. The KEA App brings together data from multiple source servers (including the partner agencies' data portals) via the Datamesh without requiring downloading the data. The KEA App and Datamesh were extensively tested with key stakeholders such as regional councils, which showed strong interest and support for future development of these tools.

Key Findings

The Datamesh provides an intuitive and innovative solution to access standardised data from different organisations and enabled the DOC use case purpose of mapping KEAs.

The DOC use case showed how a Datamesh can be transformational in how marine spatial data are accessed in Aotearoa New Zealand providing a single access point to a range of data held on different data servers. The flexible and functional data architecture provided by the Datamesh could save time and resources if fully operationalised, which ultimately could help create efficiencies in marine management.

The Datamesh also provided a platform for the development of the prototype KEA App which represents an innovative way of running analyses on spatial data, while maintaining the data on source servers. The KEA App can improve coastal planning initiatives with a user-friendly interface to map KEAs.

The two significant limitations identified during the use case were: (1) The Datamesh only partially incorporated metadata from the linked datasets, and (2) The owners of datasets need good data management discipline, including having complete metadata, in source portals. The DOC use case noted that these can be overcome with proper planning and governance in place to drive the implementation of the Datamesh.

Overall, the DOC use case provided an excellent test of the functionality of the Datamesh. The use case goals were achieved, as all data were successfully linked and the mapping of KEAs was successful, with the additional achievement of building the prototype KEA App. The use case also demonstrated that the Datamesh or a similar solution can provide a range of benefits to government departments, including regional councils, and improve marine management in Aotearoa New Zealand.

Project Outputs

The DOC use case achieved all set objectives for the overall Datamesh proof of concept. Below, each objective is assessed against the DOC's use case findings. Key outputs achieved are listed, along with some key lessons learnt.

Output 1: Connections between existing GIS web services and/or datasets (without replicating the data across to this system) from the four partner organisations supporting faster and lower-cost computation of bulk data than existing systems.	**Achieved**

Findings:

- All data needed for KEA analyses linked successfully to the Datamesh.
- The Datamesh allowed KEA analyses to run directly from the data source via the Datamesh, providing lower-cost computation of large datasets.

Key outputs:

- Linked 1168 datasets hosted on the DOC Marine Portal (total size: 9 GB)
- Linked 46 datasets hosted on other data portals (total size: 2 GB)
- A Datamesh would therefore provide a lower-cost computation because organisations would not have to store third-party data on their own networks (only their own data). Data custodianship and version control for data is maintained.
- The Datamesh provided more flexibility for data architecture, allowing better searching and downloading capabilities, therefore saving time.

Lessons Learnt:

- The success of a Datamesh will rely on good data management from agencies, including having all data in online servers (e.g., data portals) and comprehensive metadata, to allow efficient links in a Datamesh (particularly for searching). A level of curation by data custodians will be required for a Datamesh to remain functional in the long-term.
- A Datamesh will require agencies to have data on an API, and preferably publicly available to improve the digital capacity across government agencies for marine management.

Output 2: Interoperability of various data formats, standards, and scales. **Achieved**

Finding: All types of data successfully integrated (from various platforms, in several formats)

Key outputs:

- The Datamesh provided the ability to create data packages to be exported into common operating platforms. These data packages were successfully created in the Datamesh and exported to personal workspaces.
- Linked datasets came from different data portals
- Linked various data formats (raster data, shapefiles, "virtual" datasets)
- The Datamesh supported the use of the KEA App at national, regional, and user-defined spatial scales
- The Datamesh provided user-friendly options for spatial data formats (shapefiles, geoTIFF and kml). This function greatly improved the user experience.

Output 3: A web service (conforming to OGC API standards) enabling users to query and stream data into their own GIS systems and analysis platforms. **Achieved**

Finding: The KEA App developed in the use case is Datamesh-based and queries, streams, and uses datasets through the Datamesh.

Key outputs:

- The cloud based Datamesh enabled users to stream datasets directly into the online analytical tool created as part of the use case: the prototype KEA App. This analytical tool used the datasets directly from their original sources via the Datamesh to perform complex analyses.
- Outputs from the analytical tool were then available for download in various formats into users own GIS systems for subsequent analysis, or they could be exported back in the Datamesh to make them accessible to other users via the search or shared directly with others.

Output 4: Safe and secure access to sensitive data repositories. **Achieved**

Finding: This was not applicable directly to the DOC's use case as the fundamental principle was to create a fully open-access analytical tool. The function was, however, successfully tested during development.

Key outputs:

- All datasets are set automatically as restricted when first linked in the Datamesh. Only users with an authenticated email from the organisation can access the data in this setting. When a project team member used the Datamesh with an email other than from DOC, the datasets did not display when searching relevant keywords and could not be accessed. When settings were changed and the data were set to be shared with other organisations or all users, they became available to be searched and downloadable.
- Sensitive data could, therefore, be secured with restricted access in the Datamesh.

Lessons Learnt:

- Ideally a search should still return details of restricted data in the Datamesh so that users know they exist and provide contact details for requesting the data access. This would improve marine management by ensuring users know what data exist, even if they do not have access to it. This functionality could easily be enabled.

Output 5: A spatial catalogue viewer with query functionality such as searching and filtering. **Achieved**

Finding: The Datamesh interface provided a useful and user-friendly spatial catalogue with good search, filter and download functions.

Key outputs:

- Easy search function using keywords from titles, descriptions, and tags
- Good filtering function to search by time periods and area, including customised areas.
- The viewer is user-friendly and intuitive and provides display of the data for vector-based datasets, but only extent for grid-based datasets. This was not an issue for the DOC use case and searching for data in the Datamesh with these settings was efficient.
- If metadata are not well curated in the source servers, a manual step to edit metadata and tags in the Datamesh will be required.

Lessons Learnt:

- The key challenge noted is that the efficiency of the search function in a Datamesh will rely on metadata automatically imported with the dataset from its source server. There were some limitations with this function during the use case due to limited amount of metadata being imported.
- Multiple copies of the same data could be linked in the Datamesh if data custodians for each dataset are not clearly defined.

Output 6: Access to analytical tools (source code) using open formats. **Achieved**

Finding: Successful development of a prototype KEA analytical tool (the KEA App). Open-access and web-based, using data through the Datamesh.

Key outputs:

- NIWA developed R scripts for the KEA analysis that was translated into Python by Oceanum to create the App. Data are called through the Datamesh.
- KEA analysis scripts were easily linked to the Datamesh via the KEA App, so users do not have to download or manipulate the data before running the analytical tool.

Lessons Learnt:

- Input data for the analyses run by the KEA App will remain in the original data sources and are accessible via the Datamesh. This is a significant benefit of a Datamesh, allowing the data linked in the KEA App to remain automatically up to date.

Output 7: Insights/models from analytical queries run across multiple datasets. **Achieved**

Finding: KEA analyses use multiple large datasets from different sources, that are run and visualised using the KEA App.

Key outputs:

- The mapping of the three KEA criteria required access to between 11 and 51 datasets for each criterion.
- The mapping of KEA criteria is spatially constrained by reference to either a bounded analytical area informed by 1 of 4 other datasets (e.g., regional council boundaries), or user-selected datasets linked in the Datamesh.
- The datasets used in the KEA analyses were sourced from four different online servers.

Output 8: Performance of the system meets partners' and users' needs. **Achieved**

Finding: The Datamesh provides an excellent platform to improve marine and coastal management, using data from different sources through a single portal. The Datamesh enables the use of data from their original source, maintaining version control by data custodians.

Key outputs:

- The Datamesh provided a successful platform for DOC to test the application of the KEA criteria using the KEA App.
- DOC found that overall, the Datamesh and KEA App achieved our purposes.
- User testing feedback:
  - regional councils highlighted the need for a tool like the KEA App for regional planning and management
  - overall, strong interests in the technology, with some identified improvements needed.

Recommendations

High Level Recommendation

Overall, DOC recommends progressing the use of a Datamesh amongst government agencies beyond this Proof of Concept. The Marine Geospatial and Information Steering Group would be best placed to develop the roadmap on how this can be achieved. A multiagency business case could provide the necessary resources to undertake this process.

If the Datamesh is progressed beyond a Proof of Concept its success in informing marine management will be reliant upon its uptake. This includes ensuring that both data custodians make data available to the Datamesh and end users are aware of the Datamesh and its functionality. We therefore recommend investment in developing a management plan for the Datamesh and promoting it amongst data providers and end users to maximise its ability to inform marine management. A key component of this will be highlighting its dual functionality – that of providing centralised access to spatial datasets from multiple providers, and as a platform for developing apps to run standardised spatial analyses.

Technical Requirements Recommendations

- Better acquisition of metadata from source servers
- Better description of the functionality in the Datamesh
- Additional filtering options, e.g., 'by source repository', 'by owner'
- Better metadata associated with exported datasets from the Datamesh

- Additional option for the user to select the projection for the export function
- Additional clear descriptions of Owner/Authority/Sharing license/Attributions
- Regular automatic refreshing of data links
- Automated notifications/alert when source datasets used in analyses or selected in a package are updated