Manaaki Whenua
Landcare Research

The MBIE Innovative Data Analysis Programme – a plan for realising enduring value

FINAL Report

Prepared for: Ministry for Business, Innovation and Employment

December 2018

The MBIE Innovative Data Analysis Programme – a plan for realising enduring value

Final report

Contract Report: LC3339

David Medyckyj-Scott

Manaaki Whenua – Landcare Research

Revision History

Date	Version	Authors/Editors	Modification
18.10.18	1.0	David Medyckyj-Scott	First version submitted to MBIE
15.12.18	1.1	Anne-Gaelle Ausseil, David Medyckyj-Scott	Revisions to structure and content following meeting with Kylie Hood, Senior Investment Manager, Contestable Investments, MBIE on 24 Oct 2018.

Reviewed by:	*Approved for release by:*
Anne-Gaelle Ausseil IDA Programme Leader Manaaki Whenua – Landcare Research	Sam Carrick Portfolio Leader – Characterising Land Resources Manaaki Whenua – Landcare Research

Contents

Glossary

ALA: Atlas of Living Australia, Australia's national biodiversity database.

DOC: Department of Conservation

e-IDI: Environmental Integrated Data Infrastructure. MfE.

ELFIE: Environmental Linked Features Interoperability Experiment. An Open Geospatial Consortium (OGC) interoperability experimental.

EMAR: Environmental Monitoring and Reporting Project – MfE and Regional Councils

GBIF: (Global Biodiversity Information Facility). GBIF is an international organisation that focuses on making scientific data on biodiversity available via the internet using web services. New Zealand is a signatory and responsible for a national GBIF node (which currently does not formally exist).

HPC: high performance computing.

HTC: high throughput computing.

JSON: JavaScript Object Notation, is a lightweight data-interchange format.

JSON-LD: JSON for Linked Data. A method of encoding Linked Data using JSON.

LAWA: Land, Air, Water Aotearoa (LAWA), an environmental data and information repository. Initially a collaboration between New Zealand's 16 regional councils and unitary authorities.

LINZ: Land Information New Zealand.

LRIS: The Land Resource Information System is NZ's spatial representation of land and soil. It comprises a suite of national-scale spatial data (observed and derived) layers, including: NZ Resource Inventory (NZLRI) and Land Use Capability (LUC), the National Soils Data Repository and the National Soil Archive (physical specimens). A research-oriented data infrastructure platform provides the framework for managing and accessing digital data. Data and information products are then published through a suite of web applications and online services.

LRIS Portal: The Land Resource Information Systems portal. A repository of New Zealand land-related science datasets and information. One service within the LRIS data infrastructure.

LUCAS: Land Use and Carbon Analysis System. Used to support New Zealand's international climate change reporting to the United Nations Framework Convention on Climate Change.

LUMASS: Land-Use Management Support System.

MBIE: Ministry of Business, Innovation and Employment.

MfE: Ministry for the Environment.

MPI: Ministry for Primary Industries.

MLVT: Māori Land Visualisation Tool, also known as WhenuaVis. A web application developed and support by Manaaki Whenua.

NEMS: National Environmental Monitoring Standards. Goal is to establish best practice for the ongoing measurement of NZ's environment. The work is funded by regional councils, MfE, major power generators and the MBIE.

NSDR: The National Soils Data Repository. A component of the LRIS data infrastructure, the NSDR is a versatile observation database that hosts the original National Soil Database (NSD) and other full and partial soil profile descriptions, with associated laboratory analyses. Because of IP and data privacy with respect to the soil data collected on private land, i.e. the data have commercial value for the land owner, much of the data cannot be released to the public.

OGC: Open Geospatial Consortium. An international industry consortium of over 521 companies, government agencies and universities participating in a consensus process to develop publicly available interface standards. Works closely with the International Organization for Standardization (ISO).

OWL: Web Ontology Language – a family of knowledge representation languages for authoring ontologies.

QGIS: Open-source cross-platform GIS desktop software, formerly known as Quantum GIS.

RDF: Resource Description Framework - a standard model for data interchange on the Web.

SDM: Species Distribution Modelling.

SKOS: Simple Knowledge Organization System – a W3C recommendation designed for representation of thesauri, classification schemes, taxonomies, subject-heading systems, or any other type of structured controlled vocabulary (Wikipedia).

Stats NZ: Statistics New Zealand.

SSIF: MBIE's Strategic Science Investment Fund.

UN GGIM: United Nations Committee of Experts on Global Geospatial Information Management.

1 Introduction

The Innovative Data Analysis (IDA) programme was an MBIE-funded research project led by Manaaki Whenua – Landcare Research that ran for 4 years (2014–2018).

The aim of the IDA programme was to research and develop processes to integrate and harmonise high priority heterogeneous land resource and biodiversity datasets to support a step change in the quality of environmental reporting. The programme aimed to deliver impact by supporting central and regional government to report on the state of the New Zealand environment in a standardised, statistically robust and transparent way. The programme was aligned with key initiatives such as the State of the Environment (SOE) reporting, Environmental Monitoring and Reporting (EMaR), and the National Science Challenges. IDA used next-generation data analysis techniques and worked with data custodians and end users to develop statistical indicators for soil health, land use, and species occupancy. The programme focused on extracting knowledge and value from existing environmental data sets. Its focus was not research outcomes but technical and social infrastructure outcomes.

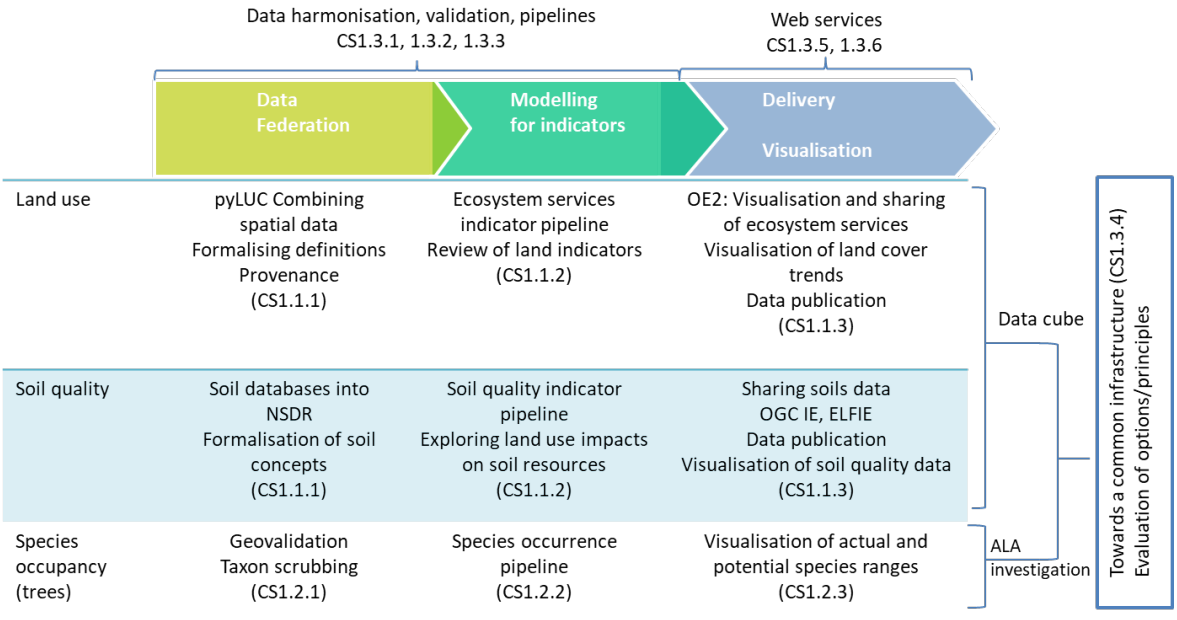

Figure 1: Diagram summarising the components of the IDA Programme.

As shown in Figure 1, the IDA programme had three domain components: land use, soil quality, and species occupancy. For each of these domains, the key components of the programme included the following.

- Data federation: bringing together heterogeneous spatial data from multiple sources to produce a suite of higher value information products.
- Modelling indicators: developing indicators to respond to the pressure, state, and impacts framework for environmental reporting.

- Visualisation and delivery: enabling use by central and regional government and other agencies, via existing portals (e.g. the LRIS Portal, Landcare Research's national land information portal) and via open standards-based web data services.

Underpinning these components, characterisation of provenance, quality, uncertainties, and workflows of the data and information were also explored to enable auditable processes.

1.1 MBIE targeted appraisal outcome

In May 2018, MBIE requested a targeted appraisal from an external panel to provide constructive comments on the outcomes from the IDA programme. During this appraisal, it was noted and recognised that:

- although the application areas of the IDA were important, the greatest enduring benefit of the IDA programme was the development and demonstration of best practices in handling, transforming, and linking spatial data, with an emphasis on the development of standards and data provenance. This was noted by all end users who were contacted during the MBIE targeted appraisal.
- the technical know-how and tools developed by the IDA programme were scientifically and technically significantly ahead of what end-user organisations were currently using and what those organisations looked to the IDA team to provide direction for their current and future work.
- the feedback from end users who had been involved in the programme was very positive, there was clear and strong support for the work that had been done and a recognised need for the work the programme had undertaken. However, both the MBIE reviewers and the end users consulted were concerned about the continuation of the research and technical development that was initiated by IDA now the programme was nearing the end.

The IDA programme leveraged existing data and existing information systems, and quite a large portion of the technical work was highly dependent on, and in some cases tightly integrated into, existing Manaaki Whenua services and internal data infrastructure:

- The data processing pipelines specifically deal with problems with certain environmental data managed by Manaaki Whenua and use components of Manaaki Whenua's data infrastructure. This means that much of the technical work is not shareable.
- Some of the work undertaken involved the creation of proof of concepts rather than full implementations. Further work would be required to make the proof of concepts operational and not all may be worth making operational anyway because, for example, the technology has moved on since the research was done and better ways now exist.
- Likewise, much of the data underpinning the indicators cannot be made accessible for reuse by others. Reasons for this include: the data are derived from commercial input data; there are data privacy constraints; in some cases, it is unclear who is the custodian of the data and thus whether Manaaki Whenua has the right to publish it.

Despite the issues listed above, the IDA programme has generated a large number of outputs. These include white papers, academic papers, and reports (32 were submitted as part of the targeted appraisal process and another three written in the final 2 months of the programme). These will help ensure that other researchers and stakeholders are aware of the outcomes of the programme going forward.

Following the conclusions from the targeted appraisal, both MBIE and Manaaki Whenua were concerned about ensuring the 'enduring value' of the IDA programme. The need to do this was recognised in the funding proposal and a critical step (1.3.7) was included to develop a strategy for enduring value of the programme. This document sets out that strategy but in the form of a plan rather than a strategy. It is worth noting that as the IDA programme has now ended, new funding sources need to be found to implement the plan proposed in this document.

2 Scope of this document

In this plan, we have organised the outputs of the IDA programme into three categories:

- know-how (section 3)
- data/information (section 4)
- technology (section 5).

For each, we describe the highlights from the IDA programme focusing in particular on highlights that have potential for enduring value, where IDA was able to leverage other funding to take the research and development further or generate enduring value,[1] and finally we list opportunities for further work through applications for new funding or partnership.

In the appendix more detail on each category is presented in tabular form under the following column headings.

Column 1. Short description of IDA output. As there were a large number of outputs, we have focused on key output or outputs where we can see a potential path to enduring value.

Column 2. How we leveraged additional value during the IDA Programme. During the lifetime of the programme we have worked to ensure that the programme has aligned with other projects and initiatives and that the knowledge gained and the outputs created in the IDA programme have been shared

[1] As specified in the MBIE contract, 3 core funded programmes (now SSIF funding) were directly aligned with IDA (the LRIS Programme that includes the National Soils Data and the SCENZGrid Capability Fund) or indirectly (Ecosystem Services Core). The work that was being undertaken in these programmes was essential to the IDA programme as they provided basic data sources on soil and ecosystem indicators (CS 1.1.1) and supporting technical infrastructure (relevant to RA3). This associated SSIF support was essential to complement the IDA work and ensure the success of the programme.

with stakeholders. Often the IDA output has been an important input to other projects.

Column 3. Lists what Manaaki Whenua has committed to doing this financial year (2018/19) to ensure longer term benefit from the programme. This is being achieved through the alignment of some of Manaaki Whenua's MBIE Strategic Science Investment funding (SSIF) and other funding.

Column 4. Opportunities for broader uptake of the outputs by stakeholders and other agencies within New Zealand and/or internationally or activities that could be possible through applications for new funding. In some cases, we have also described where there are barriers to uptake or to gaining enduring value.

In section 6 we provide more detail on those activities that we are continuing to progress during 2018–19 to generate greater enduing value.

Section 7 explores the value of IDA for Maori.

3 Know-how

3.1 Highlights

The IDA programme has produced over the life time of the programme a number of outputs that have been used to transfer knowledge, expertise and learning from programme staff to end users and others (Appendix – Table 1). The 'know how' was mainly delivered through ongoing consultation with end users, presentations to stakeholder groups and the research community, workshops, and a variety of documents (reviews, think-pieces, technical reports and scientific publications).

- Developing environmental indicators**.** We successfully engaged with key end users and stakeholders (incl. MfE and EMaR) relying on land, soil and species occupancy information for reporting purposes. The IDA programme provided input into discussions on indicator development and methods to maximise the reuse of disparate and variable data in biodiversity and land domains. We were invited to contribute to the Technical Advisory Group for the Land Domain report 2018. The added value provided by IDA staff led to a formal secondment to MfE to work closely with the Land Domain report editor to refine and finalise the report and associated products. This secondment helped with the overall consistency and cohesiveness of the report, providing guidance to analysts for data and information interpretation.
- Recording data provenance**.** The work on the NZ Land Use (LU) classification regenerator, pyLUC, (see section 5), led us to investigate data provenance, and as result we extended pyLUC to demonstrate how provenance data could be generated by a data modelling pipeline and published. Provenance capture was also implemented in LUMASS (Land-Use Management Support System) and tested on an application of SedNetNZ. Through presentations and papers this work has enabled researchers and end users to become aware of the importance of increasing transparency, openness, and reproducibility of environmental data modelling and how

provenance data can be captured, published and visualised. This is an essential piece of work that will be increasingly core to reporting requirements for central government and for publicly defensible analysis in the National Science Challenges.

- Developing reproducible data pipeline**.** An important aspect of the bio-data component of the IDA programme considered approaches to species distribution modelling consistent with the IDA data pipeline and the generation of species occupancy indices. We investigated a broad range of approaches that could be adopted and evaluated some specific online SDM toolboxes. A strong preference was expressed that the focus should be the data pipeline and the application programme interface (API) of that pipeline to widely available programming environments, specifically 'R', which is the dominant language used by biodiversity researchers (McCarthy et al. 2018).
- Assessing spatial representativeness. Our assessment of the coverage and representativeness of current soil quality monitoring sites (Cavanagh et al. 2017; Manaaki Whenua 2017) was referenced in the Land Domain report as an indication of the necessary direction of future soil sampling strategies in NZ (Ministry for the Environment & Stats NZ 2018).
- Improving data exchange through standards**.** IDA actively engaged in national initiatives looking at improving environmental data sharing. For example, we worked with the LAWA technical team regarding architectures and data exchange standards to achieve data federation. Through such engagement, along with events such as hosting an Environmental Data Summit during the Open Geospatial Consortium (OGC) Technical and Planning Committee meeting that took place in Palmerston North in December 2017, IDA has increased awareness in New Zealand about the fundamental needs for robust data management principles, and optimal design of components of environmental data infrastructures.

These know-how activities had a range of flow-on benefits (see Figure 2):

- **Capability building (external):** Knowledge transfer through working with regional councils e.g. Horizons Regional Council, Environment Southland and land domain technical advisory group (TAG) member for Ministry for the Environment. The Environmental Data summit and the white paper on POC for land and soil indicators.
- **Capability building (internal):** A review of environmental indicators for NZ, soil quality data management, structuring and barriers, species distribution modelling review. (The review of the current global status of species distribution modelling is enhancing our internal capability and capacity, with a view to integrating best-practice into an NZ data-integration platform such as the ALA).
- **Capability building (internal and external):** Data provenance review of best practices and presentations at conferences.
- **Connecting with the international community:** Throughout the IDA programme, members of the team have engaged with international peers to collaborate and leverage from IDA related international initiatives. Such interaction brought a number of benefits: accelerated technical development; opportunities to leverage other participants' know-how; and adopting international standards and related technologies to shorten the development process in the programme and thus costs. In addition, through this activity we increased New Zealand's international research

and technical profile in the areas of environmental data infrastructure, data standards and interoperability. IDA supported the involvement of Manaaki Whenua staff in an Open Geospatial Consortium (OGC) Interoperability Experiments (IE). For example, to test the exchange of soil description and observation data between agencies in Oceania, North America, and Europe an IE was conducted over 6 months. This IE successfully reconciled a set of existing soil data exchange models into a single draft standard that was then implemented as a set of interoperable prototype data services and demonstration clients. The IDA team, in partnership with some aligned SSIF funding, was able to ensure the experiment could subsequently be adapted for handling soil quality data in the IDA programme.

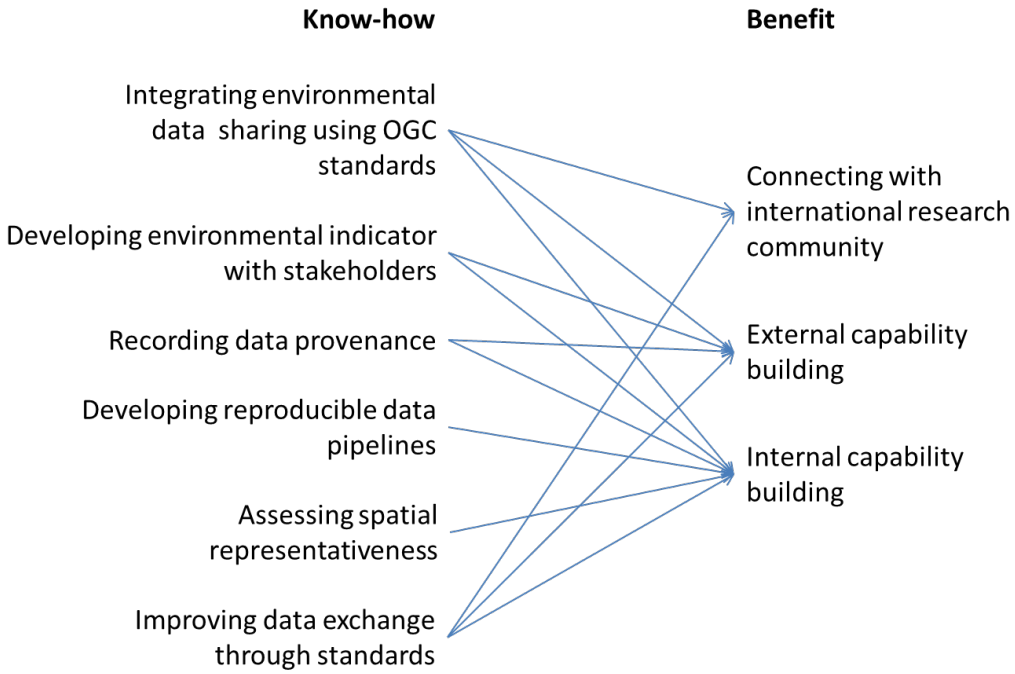

Figure 2: Know-How benefits from IDA.

3.2 Uptake supported by other funding sources

- **National Science Challenge:** commissioned a data think piece for the Our Land and Water National Science Challenge (Medyckyj-Scott et al. 2016) and a Nexus 'think piece' report on the development of environmental/ sustainability indicators (Garrett et al. 2016).
- **MfE**: consulting for national environmental reporting providing scientific leadership in land domain indicators and supporting the writing of the Land Domain report. Research undertaken in IDA used in new projects: on land-use intensity indicators; improving the grassland layer in LUCAS; river water quality changes in response to land-use intensity.
- **Envirolink**: research undertaken in IDA used in new projects: a review of soil quality and trace elements; best practices for naming soil site surveyed; a review of land-use classification for Environment Southland, and creation of guidelines for reporting land fragmentation.

- **SSIF**: we contributed expertise to IPBES Asia-Pacific Regional Assessment and to central government initiatives such as the Living Standards developed by the Treasury and the System of Environmental Economics Accounting by Statistics New Zealand. Based on the learnings from the IDA programme, additional internal investment has been aligned toward enhancing specific SDM datasets and methods by improving impacts of sampling bias and data uncertainty.

3.3 Opportunities for future work

- Ensuring knowledge transfer through presentations and papers on the research and development undertaken in the IDA Programme to specific audiences to ensure the learning and outputs from IDA are relayed to future research projects and technical development work.
- The EMaR land group has expressed interest in transferring IDA knowledge on data provenance to their protocol for creating a land fragmentation data layer.
- Continuation of the indicator work via the NSC OLW on soil natural capital and via Living Standards for Treasury.
- Continued engagement with MfE, particularly for Environmental Aotearoa 2019.
- Working with international bodies and initiatives such as the Union of Soil Sciences (IUSS) Working Group on Soil Information Standards and the FAO Global Soil Partnership to further the work on soil data harmonisation and technical standards for soil data exchange.
- Technical enhancements to proprietary and open source GIS applications so that they are better able to consume more complex types of environmental data provided by web services (APIs) that use the ISO/OGC service interface specifications.

4 Data and Information

4.1 Highlights

The IDA programme delivered a number of datasets and summary statistics tailored to the end users working in the environmental reporting space, as specified in the scope of the programme (Appendix –- Table 2). Much of these data continue to be available through a variety of data access points provided by Manaaki Whenua or other data services. However, because of data privacy, ownership or IP issues and/or because data were derived from other data that included commercial data, we are unable to make some data publicly available, e.g. the soil quality data.

- **Soil data and information.** Using a number of extant data sources, IDA created a well-documented, nationally meaningful soil quality dataset of clear provenance (see Know-How section). This dataset has been fundamental to the development of State of Environment reporting for the Land Domain, as legislated under the 2015 Environmental Reporting Act. It is widely accepted that consistent, reliable and meaningful information on the Land Domain had critical data gaps for New Zealand, as identified through the Statistics NZ / MfE-led national Environmental Domain Plan

(Statistics New Zealand 2012). With the inclusion into the IDA programme of the '500 Soils Database' of national soil quality monitoring, we undertook a major data validation exercise manually checking each data entry, cross-checking with lab sheet information, and dealing with legacy data collection and management issues such as inconsistent coding of properties across records. This data will be added into the Manaaki Whenua National Soil Data Repository.

- **Land-use data and information.** Land-use information has been identified as a major gap in data in the last Statistics NZ/ MfE-led Land domain report. While there are several GIS databases referring to how the land is used, none provide a full picture, and they need to be combined to improve our understanding of how intensively the land is used and how that use is changing. We successfully reconstructed three national land use classifications (namely for PLUTs, LUNRZ, and LUNZ,[2] see Manderson et al. (2018) for details). The methodology was successfully automated in pyLUC (see section 5). Our work on land use generated new projects with MfE, including an improvement of the grassland information in the LUCAS project, and a workshop on definition of land-use intensity indicators. This could support the Water Reform policy development through modelling of land-use pressures on fresh water (MfE/MPI Water Directorate). Subsequently, we successfully transferred ecosystem services models that were developed in the past (Ausseil et al. 2013) into our LUMASS (the Land Use Management Support System (Herzig 2017)) software and updated the models to use the latest land-use map developed by Manderson et al. (2018), LCDB version 4 and Agribase version 2015. The resulting outputs were used to create a set of ecosystem services spatial layers (greenhouse gas emission, nitrate leaching, water yield, sediment lost and sediment retained).
- **Biodiversity information.** The basis for the national-scale bio-data repository harvested from heterogeneous data-sources was created through mechanisms such as the GBIF network. New Zealand has been a signatory to GBIF since its inception in 2001. Part of the responsibility of membership is the provision of a national GBIF Node – a national clearing-house for biodiversity data that provides support to potential data providers to mobilise their content. IDA mobilised the relevant and open data associated with our Nationally Significant Collections and Databases (around 1.5 million records from a total of 6 million supplied to GBIF for New Zealand) for the development of biodiversity-domain indicators. In addition, we incorporated key data not available in GBIF from the National Vegetation Survey Databank and the BioWeb threatened species database maintained by DOC to further extend the coverage for development of biodiversity-domain indicators.

4.2 Uptake supported by other funding sources

- **MfE**: summary soil quality data submitted to MfE to support State of the Environment reporting for Land Domain. Research undertaken in IDA used in a review of a process

[2] Primary Land Use Types (PLUTs) – a classification developed to assist with the design of soil quality monitoring frameworks, LUNRZ Motu's Land Use Rural NZ (Hill & Sparling 2009), Manaaki Whenua's LUNZ – Land Use of New Zealand.

to create a land-use map and the spatial nitrate leaching update for environmental reporting (Ausseil & Manderson, 2018).

- **Horizons RC/Dairy NZ**: provided with nutrient budgeting information using land-use data from IDA (Manderson et al. 2016 (Unpublished, see Singh et al. 2017).
- **Lincoln AgriTech**: provided advice on collating data about artificial drainage using land-use data from IDA (Manderson 2018).
- **Stats NZ:** are using some of the ecosystem services layers created in the IDA programme for System of Environmental-Economic Accounting.
- **DOC**: in response to the sudden incursion of myrtle rust in May 2017, the Department of Conservation (DOC) contracted Manaaki Whenua to produce improved distribution and abundance maps of native Myrtaceae. This was an excellent opportunity to build on IDA for its bio-data harvesting, data validation, and cleaning process. As a result, a rapid response with a first draft of such maps was produced within a week. The information generated by IDA was used by DOC to target collection for seed-banking of key species that might be devastated by the rust. DOC, together with NIWA, used the data as part of their spread risk assessment process. This is a good example of the enduring direct benefit of IDA development to delivering benefits to New Zealand (Wiser et al. 2017).

4.3 Opportunities for future work

Stats NZ are using some of the ecosystem service/conditions data as examples for their ecosystem accounting work.

There are many opportunities for greater use of soil and land-use data, but the use of commercial data as input layers and effect of reuse, data ownership and privacy issues need to be resolved. Looking ahead there is going to be a growing tension between the need for research data to be citable in scientific papers and available for reuse by other scientists and the fact that, for data privacy and IP reasons, certain data cannot be open data.

Future development of a national biodata platform is contingent on the current MBIE review of nationally Significant Databases and Collections leading to implementation of the recommendations of the extant review documents. The funding of national services that provide environmental data into global science data systems such as GBIF (biodiversity data) or the proposed FAO Global Soil Data System (GLOSIS) also needs resolving.

5 Technology

5.1 Highlights

Consultation during the programme with major stakeholders (MfE, DOC, regional councils, and others) indicated that a 'front-end' (i.e. user interface such as a website) was not a high priority compared with the more significant challenges of sourcing credible and

representative data for reporting. Therefore, the focus of IDA was on back-end infrastructure technical development and thus much of what has been developed is not intended for use by end users or shareable.

Many of the components, e.g. pipelines, services, are highly dependent, and in some cases tightly integrated, with Manaaki Whenua's Nationally Significant databases and other databases, the related infrastructure and other services. Other components are scripts and code specific to deal with harmonising data (and had to resolve associated data cleaning, validation and integration challenges). Nevertheless, we see opportunities for ongoing use or re-use of some of the technology developed in the programme (Appendix –Table 3).

- **Harmonising data for integration.** IDA defined a process for aggregating, integrating and standardising relevant species occurrence and soil quality data. As an example, for species occurrence data, we implemented procedures for extracting data from GBIF, Naturewatch, NVS and NZVH. A set of web-services were associated with the New Zealand Organisms register to validate the spelling of organism names and to provide a set of Globally Unique Identifiers with which to track name-related data through subsequent data-manipulation pipelines. IDA then expand this list of names to produce a list of all equivalent synonyms, infra-specific taxa, and misapplied names harvested through the NZOR network together with the currently accepted name. As an example of what was achieved, an original set of 216 tree names were expanded to a possible 940 names which may be then linked to relevant observation data.
- **Validating data for analysis.** IDA developed and implemented a generalised model for capturing taxon concept changes over time. The model/data can be used to disambiguate historical observation/survey records (such as NVS/GBIF) to provide improved information on species distribution ranges, which in turn supports ecological analysis and improved decision-making, specifically on threat status (and potentially risk status in a biosecurity context). IDA also developed an NZ-specific geo-validation service that is now a component of the NZOR data integration and quality assurance pipeline.
- **Providing a platform for creating workflows and data pipelines**. We produced a NZ Land Use (LU) classification regenerator, pyLUC, that enables automatic generation of LU classifications using the latest source data, to a high standard of reliability, and on demand according to end-user needs. pyLUC is platform-independent and allows one to define and construct LU classifications that are easily reproducible. It uses versioned input data available from the LRIS Portal, increasing the transparency of the process, but is extensible to access data from other data source end points, e.g. web services.
- **Publishing data for sharing knowledge.** An open source registry system was deployed to test its suitability for the publication and management of land-use classification systems. This Linked Data Registry software used semantic web standards and technology to store, organise, search, and publish classification data on the web. The use of the Simple Knowledge Organisation System (SKOS) proved to be convenient, light-weight ontology for describing and organising the land-use classes. Increasingly, web services (APIs) will be the way other scientists, business, and government will access environmental data – either data collected from sensors, surveys or as a result of analysis and modelling. IDA investigated and successfully set up a test data publication service using the forthcoming WFS 3.0 specification. WFS

3.00 is the next version of the OGC Web Feature Service standard. It will be a modernised service that aligns with the current architecture of the Web, has a focus on improving the developer experience and will likely be a blueprint for the next versions of other OGC service types.

- **Integrating environmental data.** IDA looked at new ways to integrate and publish environmental data. We looked at the use of linked data, common domain conceptual models, ontologies, e.g. SKOS, and vocabularies as a way to achieve cross agency data sharing and integration. We proved this concept by applying Linked Data API and WFS API practices and semantics to the work undertaken in the OGC Environmental Linked Feature Interoperability Experiment (ELFIE) project (Blodgett et al. 2018). The concept was also demonstrated in the MfE e-IDI project. Building on the ELFIE project, the e-IDI project and other technical and data learnings, a proof of concept of a multi-indicator infrastructure for soil quality and land-use data was developed by the IDA programme (Ritchie et al. 2018). Together, the results showed one way by which environmental data from Manaaki Whenua and other NZ data providers could be formally described, shared, and integrated.
- **Towards a common multi-indicator infrastructure**. A significant component of a multi-indicator infrastructure, providing centralised data storage as well as possible compute services, could be a 'data cube'. Data cubes commonly hold spatio-temporal information such as sequences of satellite images or model output for a series of time steps, though single snapshots (or iterations of snapshots) of non-temporal data can also be stored. We investigated readily available data cube technologies to see if such technology could provide a service that would integrate data from various sources (model and observational) in a central location where it would be easily accessible to researchers. We found that current open-source tools had their shortcomings, and none was an 'ideal' candidate, though the Open Data Cube (ODC) came close. Our conclusion was that this technology was not yet at the point where it could be easily implemented and used operationally for large-scale data (Jolly 2018).

  Beyond the simpler data cube concepts, the IDA programme recognised that a more advanced framework was required that could link large multi-resolution and multi-domain datasets together. These multi-resolution data cubes could enable the next generation of analytic processes to be developed. The Open Geospatial Consortium's Discrete Global Grid System (DGGS) could be part of the solution (DGGS, 2017). IDA supported co-authorship of the DGGS Specification – Topic 21 [OGC 15-104r5] (Purss & Gibb 2017). The OGC DGGS will become a standard in the ISO 19100 series of geospatial standards. Organisations that are building (or have built) systems using the DGGS standard include UK Defence Science and Technology Laboratory, Google (S2), Uber (H3), GeoScience Australia and Statistics Australia (AusPIX).
- **Visualising data.** We investigated data visualisation techniques for effective communication between scientists and stakeholders on the state and trend of the environment. This included novel ways to visualise soil quality, land cover and species occurrence data as infographics; dynamic, map-based visualisations; and interactive visualisations that allow exploration of spatio-temporal changes for land resource data at a variety of scales and soil data uncertainty (Cowie et al. 2018). We also created a R-shiny web app which allows the exploration of variability in selected species distributions according to spatial granularity size (10, 25, 50, and 100 km

grids), time-slices of primary occurrence data, and taxonomic resolution. The visualisation tool presents the underlying primary occurrence data as well as range maps.

5.2 Uptake supported by other funding sources

SSIF: MBIE SSIF Platform and SSIF infrastructure supported the publication of data created in the IDA programme both during year 4 of the programme and going forward (National Soil Data Repository, LRIS Portal, OurEnvironment website, NZOR). SSIF funding also assisted in the incorporation of technical components developed in the IDA programme into the Manaaki Whenua data pipelines, internal data infrastructure and publication services.

5.3 Opportunities for future work

- Evaluation of data cube technology in the MBIE Advanced Remote Sensing of Aotearoa for Next Generation Land Cover Mapping Programme (see section 6.3).
- Pilot projects in New Zealand to explore the potential of DGGS, e.g. the use of DGGS for National Environmental Monitoring sites. Such projects could possibly be combined with further investigations into the use of data cubes (Nativi et al. 2017) but in a New Zealand context.
- Establishment of environmental vocabulary services to support an NZ wide environmental data infrastructure and specific applications such as LAWA. Such services would support the consistent definition of environmental data properties throughout New Zealand with relationship mappings used to define the semantic relationship between properties (e.g. 'equivalent to', 'broader than', 'related to').
- Investigate the demand for and opportunities arising from providing open access to API based services developed in IDA, including those based on new standards such WFS 3.0, by which third parties, e.g. government, business, could access Manaaki Whenua's environmental data. Establish a set of agreed data encodings to be used in the publication of environmental data via web services (APIs) to meet the needs of different end users – application developers, geospatial experts, scientists. Manaaki Whenua could also work with stakeholders to provide environmental data using the next generation of Open Web Services with data published as linked data (this requires data providers to publish URIs for their environmental features) thus allowing the linking of environmental data over the web.
- Extending pyLUC into other land related areas and further developing the functionality provided by pyLUC, e.g. broadening the types of data sources that can be consumed by pyLUC.
- More research into appropriate ways to visualise environmental data uncertainty for different audiences, e.g. citizen, policy maker, scientist.

6 Continued activities in 2018–19 after the end of the IDA Programme

The following is a list of activities Manaaki Whenua have committed to undertaking post the end of the IDA programme.

6.1 Know-how activities (see Table 1 in Appendix)

- **A workshop with key stakeholders**, including Māori representatives, to discuss this plan and other opportunities for uptake of outputs from the IDA programme and follow on projects.
- **A workshop to promote better data management practices** of environmental data (e.g. standards for data collection and management, data citation and provenance, etc.) using the lessons learnt from IDA to ensure the learning and outputs from IDA will be relayed to future projects. This workshop should provide an opportunity to expand the user base of the research and what has been learnt. This will be an activity of Manaaki Whenua's Characterising Land Resources Portfolio and run jointly with NIWA.
- **Website**: Creation of an area on the Manaaki Whenua website which will summarise the MBIE IDA Programme and provides links to papers, reports, software and tools.
- **Engagement with GBIF secretariat** on biodata management in NZ and the Pacific (as GBIF Node Manager). Engagement with iNaturalist, the largest Citizens Science platform for capturing biodata, through board membership of the NZ governance of iNat.
- **Journal publications**: Paper submitted to Environmental Modelling & Software journal on provenance and IDA. Paper to be submitted on an enhanced LUMASS modelling platform (Herzig & Rutledge n.d.).
- **Publication of the OGC ELFIE** Technical Engineering Report (approved by OGC members in December 2018 (Blodgett et al. 2018).

6.2 Data activities (see Table 2 in Appendix)

- **Soil quality data** continue to be managed by Manaaki Whenua as part of the work it does with respect to its role as a custodian of nationally significant databases (SSIF Infrastructure funding). We will be working on the inclusion into the NSDR of the additional soil quality information that was processed by Manaaki Whenua for the land domain report 2018.
- **Ecosystem services data** continues to be managed by Manaaki Whenua as part of the work it does with respect to its role as a custodian of nationally significant databases (SSIF Infrastructure funding).
- **Species Distribution Model**: On-going commitment by Manaaki Whenua to species distribution modelling through multiple strategic hires. Further Manaaki Whenua investment in a species distribution modelling platform. Subject to funding for licensing of R-Shiny server, the R-Shiny app prototype for viewing the species distribution of Myrtaceous species will continue to be accessible (https://shiny-

data.landcareresearch.co.nz/IDAmapperMyrtle/ username: MBIE Password: review2018).

6.3 Technology (see Table 3 in Appendix)

- **Soil**: the ideas and technical know-how will be applied to begin work to create OWL/RDF ontologies from ANZSoilML and the SoilDataIEML (soil data exchange standards) and then to publish SKOS versions of key vocabularies for Manaaki Whenua managed data.
- **Land use**: We plan to make pyLUC open source and broaden provenance to other LUMASS components. We will undertake a usability evaluation of the land cover change viewer tool and the tool will be made available online as part of the OurEnvironment website.
- **Biodiversity**: Continued Manaaki Whenua support for the taxon scrubbing service in NZOR.
- **Cross-domain**: Evaluation of the use of data cube technology for storing and managing access to Sentinel-2 remotely sensed imagery. Collaboration with Centre for Space Science Technology funded by MBIE Advanced remote sensing of Aotearoa for next generation land cover mapping Programme.

7 Value to Māori

The generic goal of environmental data harmonisation and integration from IDA is not specific to the Māori world. The tools, data, and principles apply across both Māori and western views. The harmonising of existing datasets will however ultimately improve access to information for Maori-land owners.

The integration of traditional knowledge/mātauranga Maori into models and data structures is in progress in other research programmes[3], so was thus outside the scope of IDA.

During the programme, engagement with Māori was done at the research level through several avenues:

- First, two Manaaki Whenua staff were part of the land TAG and ensured that te Ao Māori was included in the land domain report by including some new Māori land indicators (share of land cover), with concurring media coverage on Radio Waatea to promote the benefits of the land domain report for Tangata Whenua. The share of land cover was created using a robust pipeline intersecting a cleaned-up layer of Maori land block 2017 and the latest version of LCDB (version 4.1).

[3] To the author's knowledge, this includes the Soil Health MBIE research programme led by Manaaki Whenua and Te Tāhū o te Pātaka Whakairinga Kōrero - Next Generation Indigenous Knowledge led by University of Waikato.

- Second, there was interaction with researchers through the MBIE Soil Health programme. Their focus is on the development of concepts of soil health from a Māori perspective. While the research is still in progress, the integrated dataset on soil quality will be invaluable as a fundamental data infrastructure to support the harmonisation of soil quality physical measurements to infer Māori indicators.
- The role of IDA in the myrtle rust response was a big success story and proof of value and benefits to NZ including Māori, given the importance of several of the myrtle family (pohutukawa, mānuka, kānuka, rātā). The harmonisation of the kānuka (*Kunzea ericoides*) data during the first year of IDA was also highly relevant as it is a taonga species. The updated dataset could help Māori landowners identify potential areas for tea-tree oil production.

The WhenuaVis tool[4] (Māori Land Visualisation Tool) was intended to be our avenue for targeted data delivery services but lack of funding for Whenua Vis has prevented the release of Māori-relevant information through this platform. However, IDA staff are now involved in the evolving Māori Land Online Service (led by Te Puni Kōkiri, the Ministry of Māori Development) and we will be using the expertise gained from IDA and some of the data in that initiative.

8 Progress towards achieving post-contract outcomes

The IDA post-contract outcomes highlighted that by 2020 *"central and regional government, iwi and business sectors are using a terrestrial data landscape, populated with consistent, authoritative and defensible data to support a step-change in the accuracy and acceptability of environmental reporting and improved land management decision-making. Relevant Science Challenges both add to the data landscape and use the analytical infrastructure from this project to improve data access, sharing and re-use, as well as support data-driven science discovery. By 2023, this coherent terrestrial data landscape becomes the authoritative evidence base on NZ's natural capital, enhancing policy and management decisions across the Natural Resources Sector (NRS) and enabling NZ to better respond to emerging risks and opportunities. Relevant Science Challenges continue to benefit from re-use and repurposing of data, saving millions in avoided costs and delivering world leading science."*

Our IDA programme has made progress towards these bigger goals, by virtue of the outputs generated and by allowing us to draw out the key principles that will be essential to achieving these outcomes, as well as the barriers and important challenges that need to be addressed.

A consistent, authoritative and defensible terrestrial data landscape necessitates:

- That we start by maturing New Zealand's environmental data ecosystem beginning with improved data management practices and data curation by all the actors in the

[4] https://whenuaviz.landcareresearch.co.nz/

ecosystem (Medyckyj-Scott et al. 2016). To quote a speaker at a recent Australian Data Science conference – we need to 'retain, curate and federate'.

- Improved investment in fundamental data-collections (including adequate, stable and future-proofed financial support for the Nationally Significant Databases and Collections) to allow re-use and re-purposing of legacy data. What can be achieved just using existing data has been well demonstrated in the IDA Programme. The addition of new environmental data will bring additional advantages and understanding with respect to state, trend, impact and opportunities of the land and biodiversity resources of New Zealand.
- That data are fit for purpose (sufficient resolution and spatial and temporal data coverage, with the required properties being measured or modelled) and openly available to ensure accurate, acceptable and defensible reporting on indicators.
- In addition, there need to be improved incentives and support for local and regional government to improve data management and data accessibility. Short-term investments such as IDA, eIDI or even TFBIS will not deliver the post-contract outcome. Further progress will not be possible without national investment in both new and extant infrastructure such as the ALA/GBIF for bio-data and similar for the land resources.
- Good progress is being made in some sectors to improve access to environmental data, but care needs to be taken to ensure that once-common organisational data silos are not replaced with sector data silos as a result of funding being only available to the players in that sector.

An analytical infrastructure will require substantial cross-agency efforts and strong governance.

- There are several technology frameworks and possible architectures that we have investigated, and decisions need to be made on which is the most appropriate.
- Adopting standards and international best practices across agencies is critical, it reduces the risk we will reinvent the wheel and we can build on the lessons gained by overseas institutions.
- National standards should be adopted for data collection within a domain, e.g. soil quality, so that harmonisation work can focus on the more difficult challenge of harmonising across data domains – through using semantics, ontologies and linked data.
- For bio-data to fully realise the goals expressed within IDA and formerly by TFBIS, the New Zealand government needs to make a sustained investment in data accumulation and the physical and data infrastructure at national scale such as GBIF, the ALA, BSS, and NZOR.

9 Improving data access, sharing, and interoperability

One of the recommendations of the MBIE targeted appraisal was looking to the IDA team to investigate options for recognition of nationwide platform for integrated data analysis, including environmental reporting.

Various initiatives are happening in different sectors with the aim of improving access to and integration and harmonisation of environmental data. The challenges to be solved are less technical and more about the social architecture including more joined up thinking across the sectors. Manaaki Whenua believe this is a job for MBIE and that they should facilitate the investigation of options for recognition of nationwide platform for integrated environmental data analysis, through the hosting of an initial stakeholder workshop. This should be done in consultation with MfE, Statistics NZ, regional councils, DOC, MPI, business, LINZ, Māori agencies (e.g. Te Puni Kōrkiri) as well CRIs and universities. Manaaki Whenua will support this.

With reference to species occurrence data, the widely acknowledged solution to a national platform for data mobilisation, pipe-lining and analysis, would be the adoption and support for a New Zealand instance of the Atlas of Living Australia infrastructure. Adoption of this digital infrastructure formed part of the recommendations of the recent Royal Society review of the National Taxonomic Collections (Royal Society of New Zealand, 2015) and the Decadal Plan for Taxonomy and Biosystematics in Australia and New Zealand 2018–2027 (Australian Academy of Science & Royal Society 2018). These reports, and their recommendations, have been based on numerous stakeholder workshops over a period of three years.

The IDA Programme has laid the ground for further conversations based on what has been learnt. Three examples illustrate this:

1 As the IDA programme has progressed it has become clear that users have different needs when it comes to how they discover and access data and the way the data are presented. This means implementing, operating, and sustaining multiple access routes to data. This requires having people with the necessary skills to implement the access routes and the financial resources to operate and maintain the different delivery mechanisms. In the IDA programme, we were often dealing with volumes of environmental data that could be analysed on the desktop. Increasingly, this is not going to be the case: in data science, the data can be very large and require access to institutional or commercial data centres and HPC and HTP systems. NESI provides the HPC, but support for assisting scientists to use commercial cloud facilities, such as Google Earth Engine,[5] is absent.

2 Interoperability of cross-domain data requires improved uptake of data standards, and common data service architectures to allow distributed data to be located and integrated for analysis and visualisation. The technology aspect of data interoperability is not a high research priority. But, to ensure we can respond to the demands of new techniques in data delivery and analytics, interoperability becomes a responsive task – ensuring that state-of-the-art technology is used in a consistent and disciplined way. Being responsive also means infrastructure providers must be prepared, and willing, to update the technology they deploy as the community of users adopts new tools.

[5] https://earthengine.google.com/

3 Information interoperability, the use of ontologies and vocabularies to consistently organise content, is of high operational importance and needs to be supported by ongoing research. The research should be focussed in two areas: the ongoing engagement with domain scientists to define ontologies that discretely describe the phenomena they study; and the development of techniques that build on this expert knowledge to help machines suggest or extend ontologies from available collections of structured data. This work requires deep engagement with the global community of data and information scientists who are working on these problems and must cover both applied and more theoretical communities. The applied communities (e.g. OGC domain and standards working groups) provide robust testing and development of work by experiment and implementation while the research community (e.g. collaborations with institutions like CSIRO or Australian Cooperative Research Centres) provide the domain knowledge and signal the possible direction of future work.

10 Next step for this plan

This plan will be discussed with key end users. This will be done in a half-day workshop early in 2019. The goal of the workshop will be to find ways to bring about the broader uptake and to identify where there may be opportunities to secure funding to further develop the outputs from the IDA Programme or undertake related activities. Ideas for other possible ways of securing enduring value will be sought.

11 References

Ausseil AG, Dymond JR, Kirschbaum MUF, Andrew R, Parfitt R 2013. Assessment of multiple ecosystem services in New Zealand at the catchment scale. Environmental Modelling & Software 43: 37–48. doi: http://dx.doi.org/10.1016/j.envsoft.2013.01.006

Ausseil A-GE, Manderson A 2018. Spatial nitrate leaching extent: an update for environmental reporting. Manaaki Whenua – Landcare Research Contract Report LC3245 prepared for Ministry of the Environment.

Australian Academy of Science & Royal Society 2018. Discovering Biodiversity: A decadal plan for taxonomy and biosystematics in Australia and New Zealand 2018–2027. Australian Academy of Science and Royal Society Te Apārangi: Canberra and Wellington. https://www.science.org.au/files/userfiles/support/reports-and-plans/2018/taxonomy-decadal-plan-lo-res-v200618.pdf

Blodgett D et al. eds 2018. OGC Environmental Linked Features Interoperability Experiment Engineering Report. Open Geospatial Consortium 18-097r1.

Cavanagh J, Munir K, McNeill S, Stevenson B 2017. Review of soil quality and trace element State of the Environment monitoring programmes. Envirolink Report 1757-HBRC256.

Cowie A, Medyckyj-Scott D, Heuer T 2018. Spatio-temporal web mapping of scientific data - approaches for effective user experiences. In: Cartographies of change: then, now and tomorrow. Proceedings of the 9th National Cartographic Conference Geocart' 2018, CartoPress Occasional Publication No. 4. Pp. 121–127.

Discrete Global Grid Systems Abstract Specification version 1.0. 2017. Open Geospatial Consortium. 15-104r5. http://docs.opengeospatial.org/as/15-104r5/15-104r5.html

Garrett L et al. 2016. Co-Innovation leads to high impact indicators, Our Land & Water National Science Challenge Indicators Think-Piece. http://www.ourlandandwater.nz/assets/Uploads/Co-innovation-Leads-to-High-Impact-Indicators-Think-Piece-for-Our-Land-and-Water.pdf

Herzig A 2017. LUMASS – user guide. Vers. 0.3. Lincoln, Manaaki Whenua – Landcare Research. https://bitbucket.org/landcareresearch/lumass/downloads/LUMASS_UserGuide_v0.3.pdf

Hill RB, Sparling GP 2009. Soil quality monitoring. Land Monitoring Forum. Land and Soil Monitoring: A guide for SoE and Regional Council reporting. Hamilton: Land Monitoring Forum (2009). Pp. 27–88.

IPBES 2018. Summary for policymakers of the regional assessment report on biodiversity and ecosystem services for Asia and the Pacific of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. Bonn, Germany, IPBES.

Jolly B 2018. A review of the current state of freely available data cube software. Manaaki Whenua – Landcare Research Contract Report LC3377.

Jolly B, Medyckyj-Scott D, Spiekermann R, Ausseil, AGE 2017. Automated generation of technical documentation and provenance for reproducible research. Hydrological Society Conference: Data Access and Visualisation Workshop, Napier, New Zealand.

Julian JP, de Beurs KM, Owsley B, Davies-Colley RJ, Ausseil AGE 2017. River water quality changes in New Zealand over 26 years: response to land use intensity. Hydrology & Earth System Sciences 21: 1149–1171. https://doi.org/10.5194/hess-21-1149-2017.

Kmoch A, Klug H, Ritchie A, Schmidt J, White P 2015. A spatial data infrastructure approach for the characterization of New Zealand's groundwater systems. Transactions in GIS. Early View. 10.1111/tgis.12171.

Manaaki Whenua 2017. Beyond today – Soil Science into the future. Soil Horizons Newsletter.

Manderson A 2018. Mapping the extent of artificial drainage in New Zealand. Landcare Research Report LC3325 prepared for Lincoln AgriTech as part of the MBIE-funded research Transfers Pathway Programme.

Manderson A, Lilburne L, Vetrova V, 2016. Spatial Overseer N-loss modelling for the Rangitikei Catchment. Unpublished Landcare Research report for DairyNZ. Palmerston North. (Also reported in Singh et al 2017.)

Manderson A, Jolly B, Ausseil A-G 2018. The NZ Land Use Classifier. Manaaki Whenua – Landcare Research Contract Report LC3335. (In press)

McCarthy J, Cooper J, Wiser S 2018. Evaluation of online toolboxes for implementing Species Distribution Modelling. Manaaki Whenua – Landcare Research Contract Report LC3251.

Medyckyj-Scott D, Stock K, Gibb R, Gehagan M, Dzierzon H, Schmidt J, Collins A 2016. Our Land and Water National Science Challenge: a data ecosystem for land and water

data to achieve the challenge mission 59. Palmerston North, Manaaki Whenua – Landcare Research.

Ministry for the Environment & Stats NZ 2018. New Zealand's Environmental Reporting Series: Our land 2018. Retrieved from www.mfe.govt.nz and www.stats.govt.nz. Publication number: ME 1350.

Nativi S, Mazzetti P, Craglia M 2017. A view-based model of data-cube to support big earth data systems interoperability. Big Earth Data 1 (1-2): 75–99, DOI: 10.1080/20964471.2017.1404232

Purss M, Gibb R 2017. Discrete Global Grid Systems. In: ISO/TC 45th Plenary Meeting. Standards in Action Outreach Seminar, Wellington, New Zealand.

Ritchie A 2017. OGC Environmental Data Interoperability Experiments. In: ISO/TC 45th Plenary Meeting. Standards in Action Outreach Seminar2, Wellington, New Zealand.

Ritchie A. et al. 2016. OGC Soil Data Interoperability Experiment Technical Engineering Report. Open Geospatial Consortium. 16-088r1.

Ritchie A, Osorio-Jaramillo J 2017. National Environmental Monitoring Site Identification System Envirolink Grant: 1729 – HZLC137.

Ritchie A, Osorio-Jaramillo J, Medyckyj-Scott D and Blaschek, M 2018. A proof of concept (PoC) multi-indicator environmental data infrastructure. A MBIE Innovative Data Analysis Programme White Paper. Manaaki Whenua - Landcare Research Contract Report LC3396.

Royal Society of New Zealand, 2015. National Taxonomic Collections in New Zealand. December 2015. https://royalsociety.org.nz/assets/Uploads/Report-National-Taxonomic-Collections-in-New-Zealand-2015.pdf

Rutledge DT, Price R, Hart G 2015. National guidelines for monitoring and reporting effects of land fragmentation. Landcare Research Contract Report LC2144 for Regional Council Land Monitoring Forum.

Rutledge DT, Manderson A, Lilburne L, Ausseil A-GE, Belliss S, Price R 2016. Methodology for a GIS-based Land-use Map for Southland: A Review Envirolink Advice Grant: ESRC274.

Singh R, Elwan A, Horne D, Manderson A., Patterson M, Roygard J 2017. Predicting land-based nitrogen loads and attenuation in the Rangitikei River catchment – the model development. In: Currie LDS, Hedley MJ eds Science and policy: nutrient management challenges for the next generation. http://flrc.massey.ac.nz/publications.html. Occasional Report No. 30. Fertilizer and Lime Research Centre, Massey University, Palmerston North, New Zealand.

Spiekermann et al. 2017. Innovative data analysis for land resources – the role of provenance. Manaaki Whenua – Landcare Research Contract Report LC2838.

Spiekermann R, Herzig A, Dymond J, Basher L, Betts H, Jolly B, Medyckyj-Scott D, Ausseil A-G, Phillips C 2018. SedNetNZ and the challenges of science-driven erosion control in policy and practice. AAG Annual Meeting, 10–14 April 2018, New Orleans, USA.

Statistics New Zealand 2012. Stocktake for the Environment Domain Plan. Wellington, Statistics NZ.

Wiser SK, Cooper JA, Arnst EA, Richardson SJ 2017. Mapping of native Myrtaceae species in New Zealand. Landcare Research Contract Report LC3065 for the Department of Conservation.

Appendix

Column 1. Short description of IDA output. As there were a large number of outputs, we have focused on key outputs or those where we can see a path to enduring value.

Column 2. How we leveraged additional value during the IDA Programme. During the lifetime of the programme we have worked to ensure it has aligned with other projects and initiatives and that the knowledge gained and the outputs created in the IDA programme have been shared with stakeholders. Often the IDA output has been an important input to other projects.

Column 3. Lists what Manaaki Whenua (MW) has committed to doing this financial year (2018/19) to ensure longer term benefit from the programme. This is being achieved through the alignment of some of Manaaki Whenua's MBIE Strategic Science Investment funding (SSIF) and other funding.

Column 4. Opportunities for broader uptake of the outputs by stakeholders and other agencies within New Zealand and/or internationally OR activities that could be possible through applications for new funding. In some cases, we have also described where there are barriers to uptake or to gaining enduring value.

Note that there is not necessarily an association between the cells in each row.

Table 1. 'Know how' output from IDA and enduring value activity

Output from IDA	Enduring value activity		
	During programme [funder if not IDA or IDA part funded]	**Committed 2018/19 FY**	**Broader uptake or activities possible with additional funding**
General knowledge sharing	Knowledge transfer through working with regional councils e.g. Horizons Regional Council, Environment Southland. Specific examples given below.	A workshop with key stakeholders, including Māori representatives, to discuss this plan and other opportunities for uptake of outputs from the IDA programme and follow on projects.	Manaaki Whenua (MW) LINK seminar summarising the IDA programme for government and science (in Wellington).
	Knowledge transfer with government, e.g. StatsNZ, MfE, DOC, LINZ. This included two Manaaki Whenua LINK seminars (short seminars and discussions for environmental policy-makers in Wellington) including Next Generation Information Systems (Dec 2016), Data, data everywhere.... Innovative Data Analysis (October 2016). Other specific examples given below.	A workshop to promote better data management practices for environmental data (e.g. standards for data collection and management, data citation and provenance, etc.) using the lessons learnt from IDA to ensure the learning and outputs from IDA will be relayed to future projects. This workshop should provide an opportunity to expand the user base of the research, and invitations should be extended to researchers and practitioners in universities, other CRIs and private research institutions. This will be an activity of MW's Characterising Land Resources Portfolio and run jointly with NIWA.	

Output from IDA	Enduring value activity		
	During programme [funder if not IDA or IDA part funded]	**Committed 2018/19 FY**	**Broader uptake or activities possible with additional funding**
	Knowledge transfer to science and research. For example, the Commissioned Nexus 'think piece' report for the Our Land and Water National Science Challenge on data structures (Medyckyj-Scott et al. 2016), which set out best practice with respect to data management and the technical environment required to enable a frictionless land and water data ecosystem, has been picked up by government and research initiatives. [NSC OLW funded]	Creation of an area on the Manaaki Whenua website which will summarise the IDA programme and provide links to papers, reports, software and tools.	
	Working with Maori – see 'Value to Māori' above.		
	IDA ran an Environmental Data Summit during the Open Geospatial Consortium (OGC) Technical and Planning Committee meeting that took place in Palmerston North in December 2017. The Summit was designed for providers of environmental data in both the private and public sector. Nearly 80 people attended the event from a mixture of research, government and business organisations. The research and technical development undertaken by staff working on IDA was praised by the OGC speakers.		

Output from IDA	Enduring value activity		
	During programme [funder if not IDA or IDA part funded]	**Committed 2018/19 FY**	**Broader uptake or activities possible with additional funding**
Research into environmental performance indicators for land	Commissioned Nexus 'think piece' report for the Our Land and Water National Science Challenge on the development of environmental/ sustainability indicators, and the importance of co-innovation in this process. [NSC OLW funded]		A research paper from aligned co-funding from Plant & Food Research is underway that will conceptualise the various soil functions and processes to be considered for potential land uses (Lilburne et al. n.d.).
	Contributed expertise through MfE's Technical Advisory Group to the review of land domain environmental indicators during the production of the Land Domain report ('Our Land') 2018.		Staff from the IDA team is contributing to the Our Land and Water National Science Challenge to define soil natural capital functions and services that underpin the land-use suitability concept.
	Following the expertise provided through the TAG, a secondment was put in place for Dr Ausseil, IDA Programme Leader, to provide scientific leadership in land domain indicators and writing support of the Land Domain report (Ministry for the Environment and Stats NZ 2018). [MfE funded]		Continued engagement of MW staff with MfE; in particular, for the Environmental Aotearoa 2019 and the natural capital assessment led by MfE for the Living Standards Framework of Treasury.
	Contributed expertise to the review of environmental indicators used in the Asia-Pacific region, and the use of an ecosystem services framework to better assess relationships between pressure and state of the land (IPBES 2018). [SSIF funded]		
	Worked with Treasury investigating an ecosystem services framework for the Natural Capital of the Living Standards Framework. [SSIF funded]		

Output from IDA	Enduring value activity		
	During programme [funder if not IDA or IDA part funded]	**Committed 2018/19 FY**	**Broader uptake or activities possible with additional funding**
Research into environmental performance indicators for biodiversity	Engagement in governance, reviews and stakeholder workshops for improved national-scale biodata management.	Proactive engagement with Royal Society review of National Taxonomic Collections in New Zealand, the Australasian Decadal Plan for Taxonomy and Biosystematics. Engagement with GBIF secretariat on biodata management in NZ and the Pacific (as GBIF Node Manager) Engagement with iNaturalist, the largest Citizens Science platform for capturing biodata, through board membership of the NZ governance of iNat.	All future development of a national biodata platform is contingent on the current MBIE review of nationally Significant Databases and Collections leading to implementation of the recommendations of the extant review documents.
Review of soil quality data	A review of soil quality and trace elements was undertaken to create a stock take of available information for environmental reporting (Cavanagh et al. 2017). [Envirolink and IDA part funded]		
	Recommending best practices for site naming for future soil quality collection and tracking over time which led to an Envirolink project with Horizons Regional Council defining a National Environmental Monitoring Site Identification System (Ritchie & Osorio-Jaramillo 2017). [Envirolink funded]		Possible follow on Envirolink project with Horizons Regional Council to investigate governance and policies for a Site Identification System.

Output from IDA	Enduring value activity		
	During programme [funder if not IDA or IDA part funded]	**Committed 2018/19 FY**	**Broader uptake or activities possible with additional funding**
Research into data provenance	IDA report on the value of creating provenance data, how it might be managed, and described how it might be used in several Manaaki Whenua systems (Spiekermann et al. 2017).	Report will be made available on MW website.	
	Research presented at several national and international conferences, e.g. to regional councils at the NZ Hydrology Society Conference workshop: Data Access and Visualisation in November 2017 (Jolly et al. 2017), at the American Geophysical Union Fall Meeting (AGU 2017) and American Association of Geographers (AAG 2018), both in New Orleans (Jolly et al. 2017; Spiekermann et al. 2018).		
	Provenance capture implemented in PyLUC (see below).		
	Paper submitted to Environmental Modelling & Software journal on provenance and IDA programme.		
	Provenance capture implemented in LUMASS (Land-Use Management Support System). Going forward, provenance data will be available along with model outputs from LUMASS.		

Output from IDA	Enduring value activity		
	During programme [funder if not IDA or IDA part funded]	**Committed 2018/19 FY**	**Broader uptake or activities possible with additional funding**
Creation of ecosystem services layers and land-use data layers	Research used in new projects: • reviewing land-use intensity indicators (on-going till June 2019) [MfE funded] • with MfE on improving grassland layer in LUCAS (ongoing till June 2019) [MfE funded] • River water quality changes in New Zealand over 26 years and responses to land-use intensity (Julian et al. 2017)	A paper to be submitted on an enhanced LUMASS modelling platform (Herzig & Rutledge, n.d.).	
	The IDA team reviewed a land-use classification for Environment Southland and created guidelines for land fragmentation that were then used by EMaR as a baseline for their reporting framework on land fragmentation (Rutledge et al. 2015, 2016). [Envirolink funded]		
Investigated approaches to Species Distribution Modelling	IDA Report 'Evaluation of online toolboxes for implementing Species Distribution Modelling' (McCarthy et al. 2018). The report will be made available on the MW website.	Based on the learnings from the IDA programme, additional internal investment has been aligned toward enhancing specific SDM datasets and methods by improving impacts of sampling bias and data uncertainty.	
Led the Open Geospatial Consortium (OGC) Soil Data Interoperability Experiment – working with overseas partners to create a soil data exchange	OGC technical engineering report (Ritchie et al 2016). [Part funded by IDA. SIFF Infrastructure also funded some staff time.]		The soil data exchange standard is being considered as one of the possible standards to be used in the FAO Global Soil Partnership's (GSP) GLOSIS – the Global Soil Information System.

Output from IDA	Enduring value activity		
	During programme [funder if not IDA or IDA part funded]	**Committed 2018/19 FY**	**Broader uptake or activities possible with additional funding**
standard and implemented 4 demonstration clients using the standard			Knowledge of soil data harmonisation and publication gained in the IDA programme is contributing to the better sharing of soil data globally through engagement in GSP (supported by MPI) and the development of a new version (3.0) of the joint NZ/Australian soil data exchange standards ANZML (supported by MBIE SSIF Platform and Australian funding).
Participation and technical lead of the OGC Environmental Linked Feature Interoperability Experiment (ELFIE) with US Geological Survey, LINZ (NZ), BRGM (France), INSPIRE (EU), Natural Resources Canada (CA), Horizons Regional Council (NZ)	Knowledge shared with NZ partners. (See OGC Environmental Data Summit – above). The IDA team were involved in the ISO Technical Committee 45th Plenary Meeting as invited members of the New Zealand delegation. Hosted by LINZ, this was the first time ISO TC 211 had met in New Zealand. A member of the IDA team spoke at an associated public event, the ISO/TC211 Outreach Seminar, describing the work conducted in IDA on Environmental Data Interoperability Experiments (Ritchie 2017).	ELFIE OGC Technical Engineering Report (approved for release by OGC members in December 2018 (Blodgett et al. 2018)).	

Output from IDA	Enduring value activity		
	During programme [funder if not IDA or IDA part funded]	**Committed 2018/19 FY**	**Broader uptake or activities possible with additional funding**
	Staff from IDA provided technical expertise, gained from the research undertaken in IDA to the e-IDI POC project. (http://www.mfe.govt.nz/more/science-and-data/e-idi---real-time-environmental-data-one-spot). The project, funded by the Better Public Services Seed Fund, was led by MfE, and involved Horizons Regional Council, Swirling World, and a number of other NZ agencies, e.g. LINZ. [IDA supported MW staff time.]		
Proof of concept of a multi-indicator infrastructure for soil quality and land-use data	An IDA White Paper focused on summarising the requirements, and the technical and social architecture to support the delivery and analysis of land and soil health indicators. (Ritchie et al 2018).		We identified the need for desktop GIS like ArcGIS or QGIS to be extended to handle complex data provided (complex feature type GML) by the WFS standard[6] before soils data could be used by such applications.

[6] MW did an experiment with a QGIS plug-in in the IDA programme. The plugin was developed by BRGM, the French Geological Survey, but we found it needed a considerable amount of further development work before it would work with our soil data.

Table 2. Data and information provided by IDA and enduring value activity

Output	Enduring value activity		
	During programme	**Committed 2018/19 FY**	**Broader uptake or activities possible with additional funding**
Well-documented, nationally meaningful soil quality dataset	Summary data submitted to MfE to support State of the Environment reporting for Land Domain – Environment Aotearoa 2015 report (Ministry for the Environment & Stats NZ 2018). [MfE funded]	LRIS Programme (SIFF Infrastructure funding) - Soil quality data continues to be managed by MW as part of the work it does with respect to its role as a custodian of nationally significant databases.	A barrier to wider use is the question of who owns the data, whether it includes information that a land owner would not want released publicly and who is the custodian of the national dataset, i.e. regional councils or MW. MW might be considered the custodian of a soil quality database for scientific use. Another challenge is the tension between research data being citable in scientific papers and available for reuse by other scientists and the fact, for reasons cited above, that the soil quality data cannot be open data.
	The assessment of the coverage and representativeness of current soil quality monitoring sites (Cavanagh et al. 2017; Manaaki Whenua – Landcare Research 2017) was referenced in the Land Domain report as an indication of the necessary direction of future soil sampling strategies in NZ (Ministry for the Environment & Stats NZ 2018).	Investigating the inclusion into the NSDR of the additional soil quality information that was processed by Manaaki Whenua for the land domain report 2018. This will provide a link between site data captured under the 500 soils project, and sites currently used by councils for SOE monitoring. This dataset would be extremely valuable to ensure that soil quality data collected until 2017 are securely recorded into the NSDR and ensure the provision of an authoritative and clean legacy for that data.	Further progress requires agreement with RCs that soil quality data can be aggregated into a single database (NSDR) and used for research by scientists in New Zealand.
	Two datasets were assessed for inclusion into NSDR: '500 Soils Database' of national soil quality monitoring and Land Use and Carbon Analysis System (LUCAS) soil data used to support NZ's international climate change reporting to the UN Framework Convention on Climate Change.		

Output	Enduring value activity		
	During programme	**Committed 2018/19 FY**	**Broader uptake or activities possible with additional funding**
Land-use domain data	Land-use map developed by Manderson et al. (2018) used in various internal and external projects, e.g. mapping the extent of artificial drainage in New Zealand, in the Our Land and Water sources and flows programme, an MPI project on producing a framework for pasture quality (2018).		Motu and other agencies are asking to use the layer. However, the layer is in part derived from Agribase. For the data to be used more widely, there needs to be an all of government approach to AsureQuality to open up Agribase so it can be shared when it is part of other data or when other data is derived from it.
	Partnered with Horizons Regional Council to provide nutrient budgeting information in the Rangitikei catchment (Manderson et al. 2016 see Singh et al. 2017) and advice around collating information on artificial drainage (Manderson et al. 2018). [Horizons/Dairy NZ and Lincoln AgriTech funded]		
Set of ecosystem services spatial layers. This includes data for greenhouse gas emission, nitrate leaching, water yield, sediment lost and sediment retained	The data are available for download from the LRIS Portal (lris.scinfo.org.nz) and through related web services (APIs).[7]	LRIS Programme (SIFF Infrastructure funding) – ecosystem services data continues to be managed by MW as part of the work it does with respect to its role as a custodian of nationally significant databases.	Stats NZ are using some of the ecosystem service/conditions data as examples for their ecosystem accounting work.
	Report for MfE on spatial nitrate leaching extent for national environmental reporting (Ausseil & Manderson 2018). [MfE funded]		

[7] At the time of writing this report, the data layers have been viewed 1,400 times on the LRIS Portal, with over 60 downloads of the data.

Output	Enduring value activity		
	During programme	**Committed 2018/19 FY**	**Broader uptake or activities possible with additional funding**
Biodiversity domain data	Use of the IDA validation, data aggregation pipeline and visualisation tools to support DOC/MPI myrtle rust response by mapping of NZ species in the myrtaceae family. [IDA funded the pipeline, DOC/MPI funded the mapping]	MW has an on-going commitment to species distribution modelling through multiple strategic hires. MW investment in a species distribution modelling platform.	Reports and papers prepared on myrtle species mapping. Current MPI funding to support predictive species niche modelling, in support of the myrtle rust response.
	Using IDA data pipeline, key data not previously available in GBIF from the National Vegetation Survey Databank and the BioWeb threatened species database added to Manaaki Whenua proto NZ GBIF Node to extend the coverage for development of biodiversity-domain indicators.		All future development of a national biodata platform is contingent on the current MBIE review of nationally Significant Databases and Collections leading to implementation of the recommendations of the extant review documents.
	Biodiversity data integrated into the Australasian Virtual herbarium, part of the Atlas of Living Australia, and continue to be made available via the Atlas of Living Australia platform.		New Zealand has been a signatory to GBIF since its inception in 2001. However, the formal New Zealand GBIF Node does not currently exist in any formal sense. Funding required to support its development or maintenance.
	Used pipeline to produce improved distribution and abundance maps of native Myrtaceae (Wiser et al. 2017) at request of DOC. First draft of maps was produced within a week, and refined maps in August 2017. [DOC funded]	The R-Shiny app prototype for viewing the species distribution of Myrtaceous species will continue to be accessible (https://shiny-data.landcareresearch.co.nz/IDAmapperMyrtle/ username: MBIE Password: review2018).	

Table 3. Technology developed during the IDA programme and enduring value activity

Output	Enduring value activity		
	During programme	**Committed 2018/19 FY**	**Broader uptake or activities possible with additional funding**
Domain-specific logic for harmonising and cleaning soil quality, land use and species occupancy data	Workflows and scripting have been re-used for integration and validation of soil data as part of the process for importing legacy soil data into the NSDR.	See Table 2.	
	Re-usable pipeline and related packages to support data integration, validation, and analysis of biodiversity data re-used by Manaaki Whenua.	Continued MW support for NZOR with new functionality added that was developed in the IDA programme.	All future development of a national biodata platform is contingent on the current MBIE review of nationally Significant Databases and Collections leading to implementation of the recommendations of the extant review documents.
	Biodiversity indicators pipeline extension to integrate with species distribution modelling (SDM) and visualisation outputs re-used by Manaaki Whenua.	The review of the current global status of species distribution modelling is enhancing our internal capability and capacity, with a view to integrating best-practice into an NZ data-integration platform such as the ALA.	
The IDA Linked Data Registry.[8]	The registry was used to create and store vocabularies for the e-IDI PoC project. ALGIM/Horizons Regional Council then looked to establish their own instance using the IDA Registry as a template.		We have know-how to establish vocabulary services and registers that can support environmental data systems such as Land Air Water Aotearoa (LAWA).

[8] The Linked Data Registry, a test implementation of the CSIRO Linked Data Registry (itself based on the UKGovLD design for a Linked Data registry), was trialled to test its suitability for the management and public/private access to codes, code lists, vocabularies, e.g. land-use classifications, ontologies and other reference resources.

Output	Enduring value activity		
	During programme	**Committed 2018/19 FY**	**Broader uptake or activities possible with additional funding**
Open Geospatial Consortium (OGC) Soil Data Interoperability Experiment – soil data exchange standard	Underpins the National Soil Data Repository (NSDR) developed and operated by Manaaki Whenua. [SSIF Infrastructure funded]		Informing the work of Pillar 4 and Pillar 5 of the FAO Global Soil Partnership.
	Contributed to the thinking underpinning the development of the OGC standards-based NZ Environmental Observation Data Profile (EODP) for the discovery of environmental observation data and time-series data (Kmoch et al. 2015). [IDA part funded]		
The OGC Environmental Linked Feature Interoperability Experiment (ELFIE) showed how to achieve cross-domain integration of data using standards in a lightweight, web-searchable, web-developer-friendly form		LRIS Programme (SIFF Infrastructure funding) – The ideas and technical know-how will be applied to begin work to create OWL/RDF Ontologies from ANZSoilML and the SoilDataIEML (soil data exchange standards) and then to publish SKOS versions of key vocabularies for MW managed data.	Where appropriate all MW services should provide EFLIE JSON-LD[9] representations. This is something we could work towards, funding permitting.
			With funding, we would look to: i) work with NZ agencies who have the capability and willingness to implement the ELFIE approach; ii) work with NIWA/GNS to get consistent earth science observation data delivery from CRIs; iii) then promote these end points with other agencies to encourage adoption.

[9] JSON-LD (JavaScript Object Notation for Linked Data), is a method of encoding Linked Data using JSON. Linked Data is a method of publishing structured data so that they can be interlinked and available using semantic queries.

Output	Enduring value activity		
	During programme	**Committed 2018/19 FY**	**Broader uptake or activities possible with additional funding**
			Work with the OGC to have ELFIE JSON-LD (and its subsequent versions) as a defined and widely adopted output format for the next generation of Open Web Services (OWS) (this assumes WFS 3.0 (see below) proves successful and the basis of OGC services going forward). We would then like to work with OGC partners (CSIRO, USGS, BRGM, NRCan) to further develop coherent earth science data delivery using JSON-LD (assumes well defined domain models).
Experimental web services, e.g. an implementation of Web Feature Services 3.0 standard	In 2014/15 IDA staff hosted two workshops with representatives from NIWA, GNS, LINZ, MfE, HRC regarding data exchange standards and services for NEMS and the design of a National Environmental Information Infrastructure.		MW will be investigating demand for and opportunities arising from providing WFS 3.0 API by which third parties, e.g. government, business, can access MW's environmental data. Development of a QGIS plug-in to access data provided using the WFS 3.0 standard.
Nationally focussed taxon scrubbing service – improved 'fuzzy string matching' and classification of species names provided by end-users linked to standardised entries within NZOR	Improved taxon-matching service developed and integrated into NZOR. [Integration work SSIF Infrastructure funded]	Continued MW support for taxon scrubbing service in NZOR.	All future development of a national biodata platform is contingent on the current MBIE review of nationally Significant Databases and Collections leading to implementation of the recommendations of the extant review documents.

Output	Enduring value activity		
	During programme	**Committed 2018/19 FY**	**Broader uptake or activities possible with additional funding**
Geo-coding and validation services – i) converts coordinates New Zealand specific geo-reference standards; ii) validates location of species, e.g. that the recorded location of a land-based species is not in a lake or a river	NZ-specific geo-validation services now a component of the NZOR data integration and quality assurance pipeline. [Integration into NZOR SSIF Infrastructure funded]		May be made public as part of the work to develop a suite of API services for use by science, government and business. Market analysis of need for such service will be undertaken by MW during 2018/19 FY. If need is demonstrated, funding exists to evolve the validation services to open, production APIs services.
Created pyLUC – a platform-independent technology that provides a framework for defining and constructing easily reproducible land-use changes workflows with provenance.		Make pyLUC open source and downloadable from a suitable publicly accessible code repository.	There is the potential for extending pyLUC into other land related areas. For example, a meeting was held with representatives of the EMAR land group regarding the use of pyLUC to create a land fragmentation data layer.
			There is further development work that could be done to pyLUC to extend its capabilities including: increasing the amount of metadata about the input data by automatically ingesting metadata from the Koordinates Portals, broadening the types of data sources e.g. consuming web services, make the tool more generic so it can be used for more than land-use classifications, provide support for processing entirely in vector-space.
Provenance capture implemented in LUMASS (Land-Use Management Support System) and tested on an application of SedNetNZ	Provenance generation components part of LUMASS open source software. (https://bitbucket.org/landcareresearch/lumass)	Broaden provenance to other LUMASS components.	

Output	Enduring value activity		
	During programme	**Committed 2018/19 FY**	**Broader uptake or activities possible with additional funding**
Investigated readily available data cube technologies to see if such technology could provide a service that would integrate data from various sources (model and observational)	IDA Report (Jolly 2018). Report will be made available on MW website.	Evaluation of the use of data cube technology for storing and managing access to Sentinel-2 remotely sensed imagery. Collaboration with Centre for Space Science Technology funded by MBIE Advanced Remote Sensing of Aotearoa for Next Generation Land Cover Mapping Programme.	
Contributed technical expertise to the development of an advanced coordinate reference framework that can link large multi-resolution and multi-domain environmental datasets together[10]	Co-author of Open Geospatial Consortium's Discrete Global Grid System (DGGS) Abstract Specification – Topic 21 [OGC 15-104r5] (See also Purss & Gibb 2017). [Part IDA funded]		The governments of Canada, Australia, China, and UK are investing very significantly in DGGS. The UN's GGIM has recommended DGGS as a common global geography for sharing national statistics. Organisations building (or have built) systems using the DGGS standard include the UK Defence Science and Technology Laboratory, GeoScience Australia and Statistics Australia (AusPIX).
	Google and Uber are the most well-known private companies that have chosen to use DGGS at the core of their business systems. Earlier this year, 2018, Uber released its DGGS software code (called H3) as an Open Source project.[11]		Horizons Regional Council propose to develop an internal pilot and also to separately explore the use of DGGS for National Environmental Monitoring sites.

[10] This framework, Discrete Global Grid Systems, is an alternative to the system of Latitude and Longitude and country specific flat earth map projections, such as NZ Transverse Mercator (NZTM). DGGS are designed from the ground up for meeting the needs for location and spatial data in Big Data analysis.

[11] Uber's H3:

Output	Enduring value activity		
	During programme	**Committed 2018/19 FY**	**Broader uptake or activities possible with additional funding**
	StatsNZ and LINZ, assisted by a member of the IDA programme, held a workshop on DGGS. [SSIF funded]		StatsNZ Content Standards work proposing use of DGGS for specifying location for NZ government data. StatsNZ and LINZ, helped by MW, are establishing a DGGS pilot. NZTA has signalled enthusiasm for their innovation unit to be involved. NZ Defence's GeoINT group are expressing interest and starting to think about the value of DGGS.
Created a set of data visualisations and infographics	Demo visualisations available for viewing at vizdemo.landcareresearch.co.nz	Land cover change viewer will be extended to national coverage and made available online.	Tool visualising uncertainty and variability of soil data may be added to S-map Online or MW's Soils Portal as an educational tool.
	Presentation on the land cover change viewer at the 9th National Cartographic Conference Geocart'2018 with associated paper in the Proceedings.	Usability evaluation of one or more of the tools.	The novel way of visualising NZLRI data, developed by IDA, at a level not normally possible or to be used to compare data for different locations could be implemented as a standalone web tool.

---

a. Uber announce H3's release, 16 January 2018 https://twitter.com/ubereng/status/953261851072135169, and

b. Open source project repository https://github.com/uber/h3

Output	Enduring value activity		
	During programme	**Committed 2018/19 FY**	**Broader uptake or activities possible with additional funding**
Two R-Shiny apps – soil health (part of POC of multi indicator platform) and species occurrence			Additional and on-going investment for an operational platform including a commercial licence for R-Shiny will be required to support external access where access to sensitive data is required.