

NEXT GENERATION INFORMATION SYSTEMS

*Sharing soils data regionally,
nationally and globally*

Alistair Ritchie and David Medyckyj-Scott
Informatics Team



Landcare Research
Manaaki Whenua

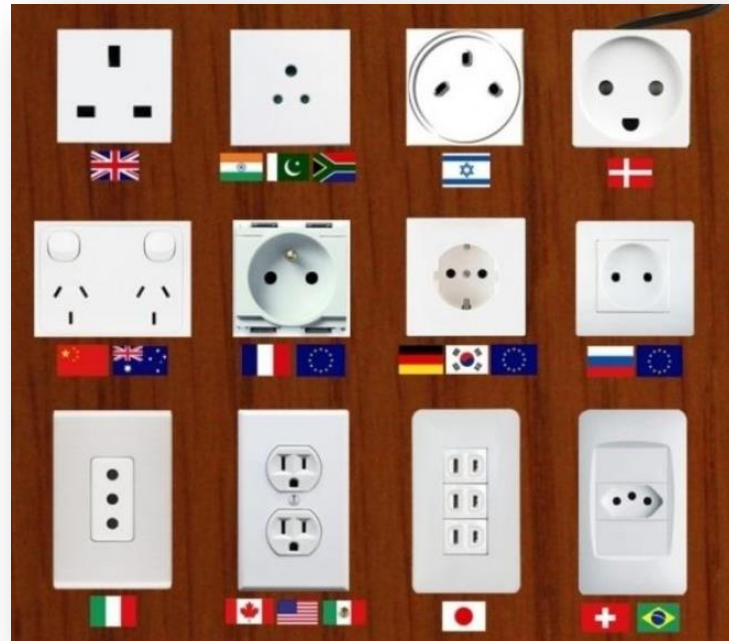
The problem space

A soil data interoperability experiment

- Background
 - An information modelling view of soils
 - Interoperability in action – demonstrations
-

The wider benefits

Lessons learnt and where next



THE PROBLEM SPACE



Healthy soils are the basis for healthy food production



Soils are the foundation for vegetation which is cultivated or managed for feed, fibre, fuel and medicinal products



Soils support our planet's biodiversity and they host a quarter of the total



Soils help to combat and adapt to climate change by playing a key role in the carbon cycle



Soils store and filter water improving our resilience to floods and droughts



Soil is a non-renewable resource, its preservation is essential for food security and our sustainable future

Broad Context

Urgent need for more information on soils at varying geographic scales

More and more data are being collected but diversity of data sources make integration and harmonisation difficult

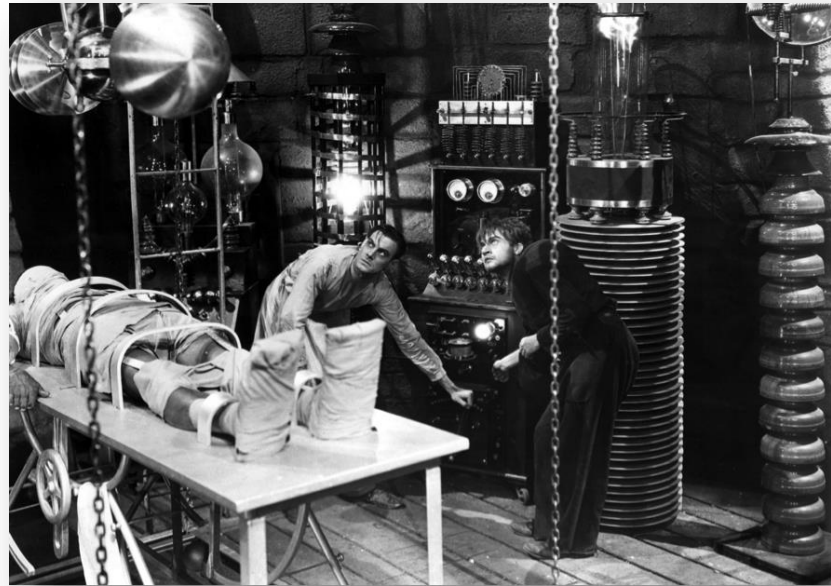
Soil data needs to be freely available and in formats that can be readily used for a wide range of purposes

There is a need for harmonization of methods

We need an internationally recognized and adopted standard for the exchange and collation of consistent harmonized soils data and information worldwide

This would allow us to...

- at a technical level: easily and efficiently transfer between databases, allowing easy access and analysis
- at local levels: easier comparison of similar soil landscapes and best-practices in similar areas
- at national levels: facilitate investigation of soil degradation or opportunities between regions for a better use of natural resources
- globally: to look at matters such as food security, biodiversity and desertification



A SOIL DATA INTEROPERABILITY EXPERIMENT

The problem and first steps to a solution

- Global/national initiatives require interoperable solutions that use an agreed standard for exchanging soils data
 - *examples include FAO's GSP, GSM, LAWA+FDN (NZ), TERN (AUS), ISRIC Harmonized world database...*
- There are existing standards but diverse and don't cover all our data exchange needs
 - *we need to reconcile them in a single, comprehensive, global standard*
- An OGC Interoperability Experiment was seen as the first step in achieving this
 - *Initiated by the International Union of Soil Sciences Working Group on Soil Information Standards*
 - *led by Landcare Research, funded in part by the MBIE IDA Programme*

Open Geospatial Consortium (OGC)

- An international industry consortium of 500+ companies, government agencies and universities participating in a consensus process to develop publicly available data exchange standards
- Different types of initiatives, e.g. Testbeds and Interoperability Experiments, used to develop standards
- Interoperability Experiments are brief, low-overhead, formally structured and approved initiatives led and executed by OGC members to achieve specific technical objectives
- An IE must focus on a single interoperability issue

Soil Data IE – Implementation

- Goal - the development and testing of a Soil Markup Language, a data encoding for soil features
- Participants: IE Initiators (CSIRO, Landcare Research, IRSIC) plus a number of other agencies e.g. USDA, USGS, Fed Uni (Aus), CRA-ABP (Italy) + Horizons Regional Council
- 6 months duration
- 4 main user stories (use cases)
- Develop a conceptual model and common exchange language (a GML-XML application schema)
- Deploy a set of demonstrators using web services (e.g. WFS, SOS) that used the schema

Use Cases

Use Case 1: **soil data integration & publication**

Publication of heterogeneous soil data from different databases at different agencies

Use Case 2: **soil sensor data**

Publication of data from sensors monitoring dynamic soil properties

Use Case 3: **soil property modelling and predictions**

Provision of high resolution estimates of functional soil properties generated using digital soil mapping techniques – e.g. the world-wide soil property predictions made by the GlobalSoilMap project

Use Case 4: **pedo-transfer functions**

Delivery of observed and interpreted soil properties (by soil type and/or by spatial distribution) in a standard format that allows the use of pedotransfer functions - algorithms that calculate additional interpreted soil properties.

Soil IE Candidate Standard

No clear candidate from existing work

Back to basics using as much existing work as possible

Minimalist approach

Don't create standards when existing standards do the job

Only extend standards when it is impossible to describe domain specific things ... soil and soil horizons

Model the thing, not the interface (soil characteristics, not portals)

Obvious, huh?

Apparently not. Egos and laziness often beat good design.

Soil IE Candidate Standard



Soil Observations

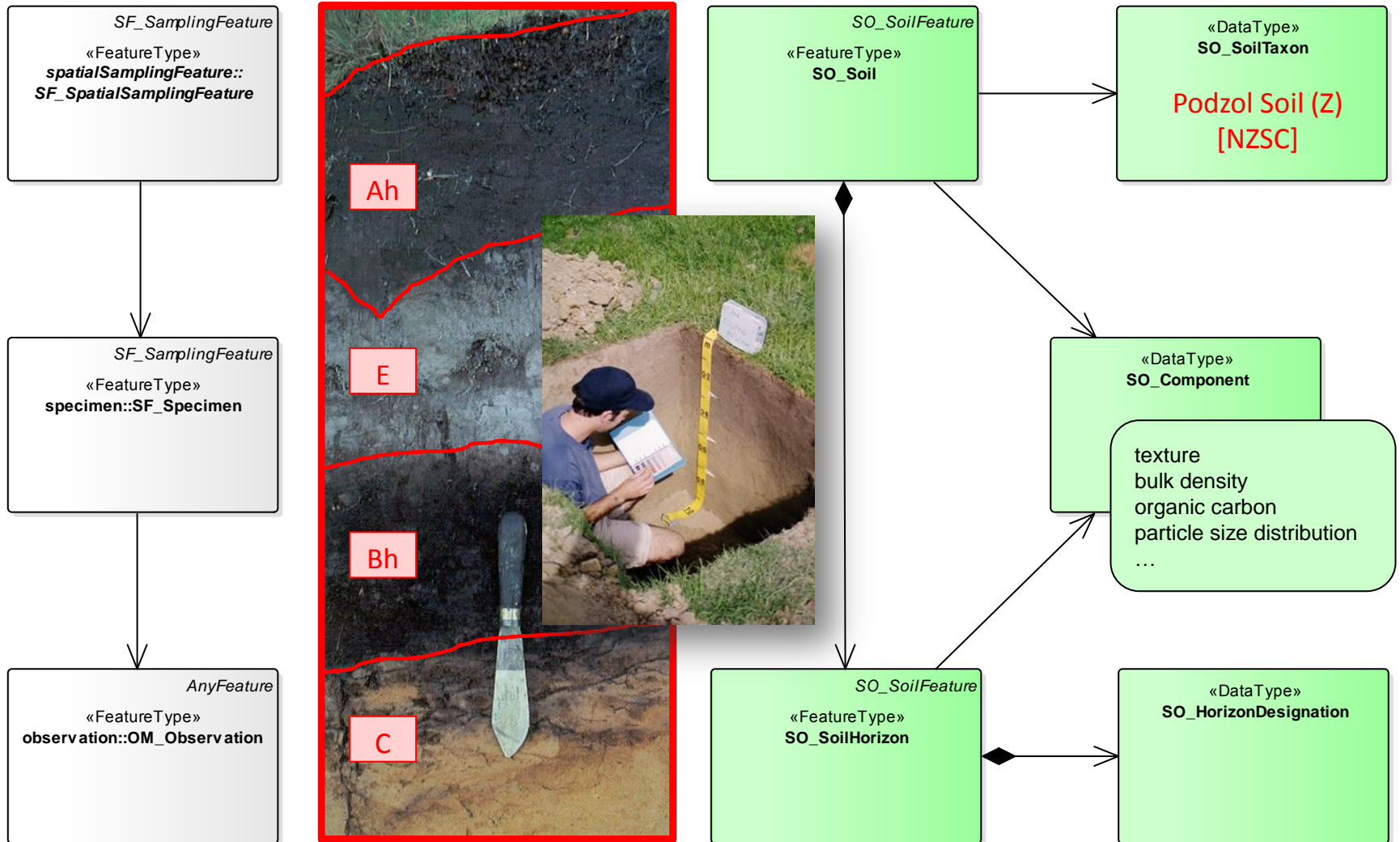
SF_SamplingFeature
«FeatureType»
spatialSamplingFeature::
SF_SpatialSamplingFeature

SF_SamplingFeature
«FeatureType»
specimen::SF_Specimen

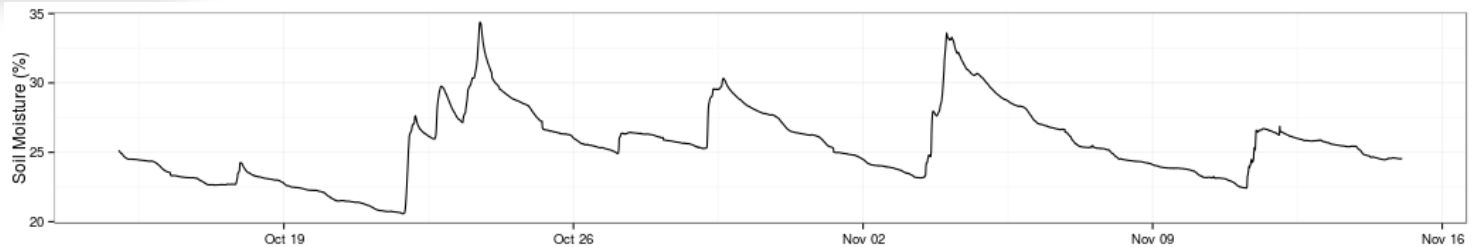
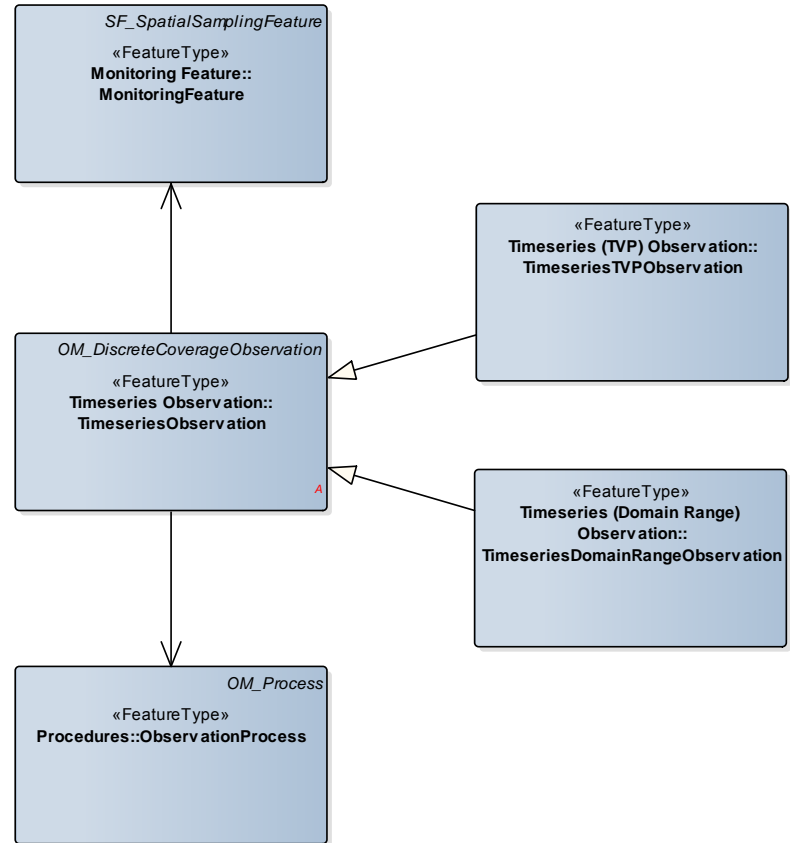
AnyFeature
«FeatureType»
observation::OM_Observation



Soil Descriptions

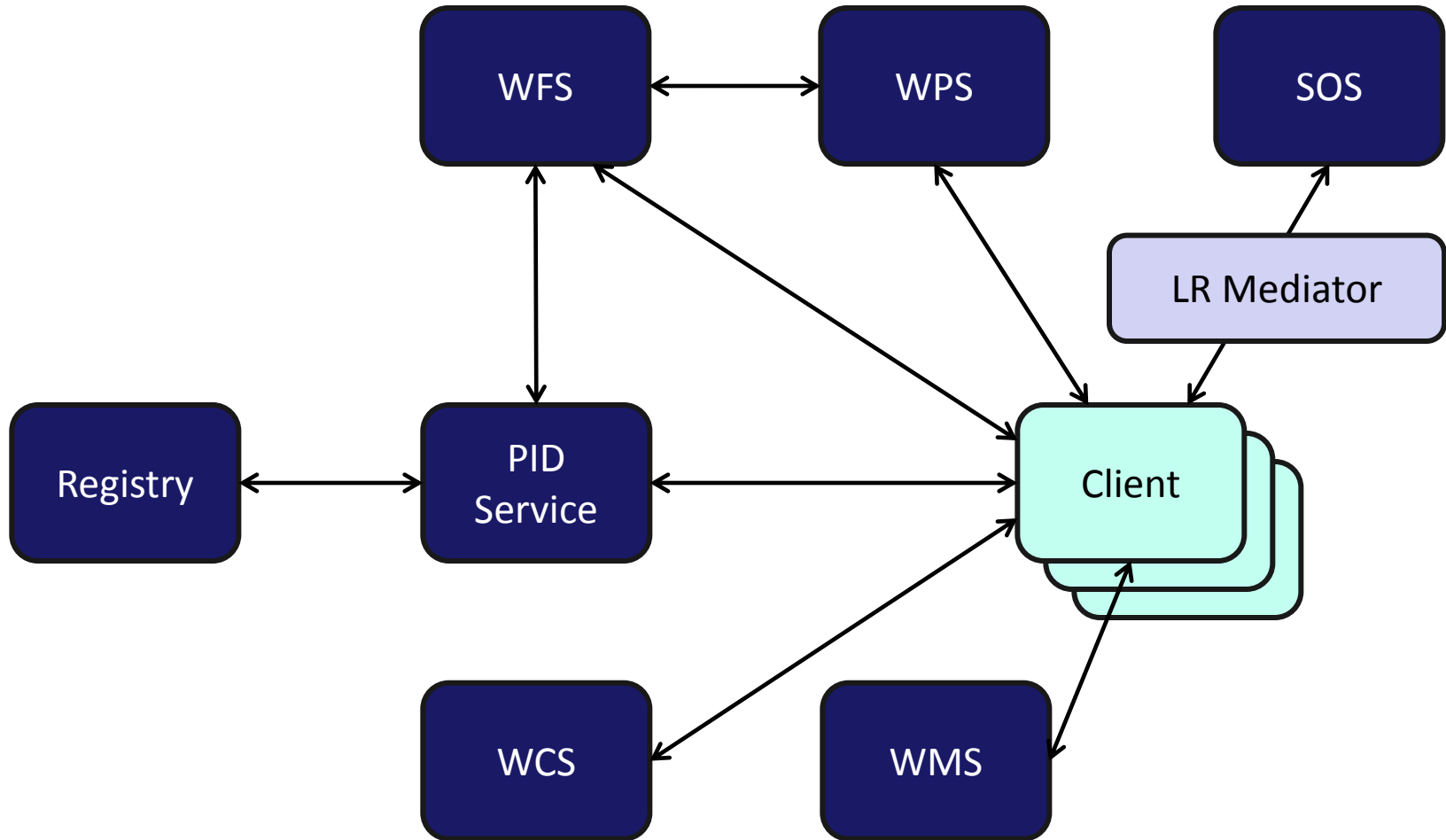


Soil Sensors



OGC15-043r3 – Timeseries Profile of Observations and Measurements

General architecture



'Simple' and 'complicated' services

Demonstration One

- Use Case Two
- Provision of soil sensor data as TimeSeriesML

Soil Moisture

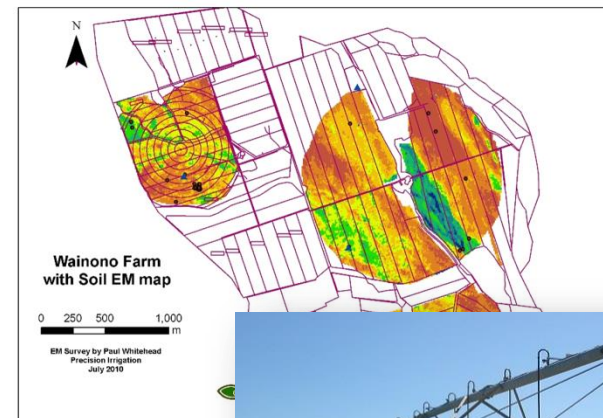
Soil Temperature

Rainfall

- Example applications

Precision agriculture

Flood hazard mapping



Demonstration One

- Contributors

Landcare Research (NZ)

Horizons Regional Council (NZ)

United States Geological Survey (US)

- Video <https://youtu.be/arMufkX6jJg>

Demonstration Two

- Use Case Three
- Model variation in soil properties in space
- Contributors
 - CSIRO Land and Water (AU)*
 - Federation University of Australia (AU)*
- Video: <https://youtu.be/xuz8ZJ1RHSM>

Demonstration Two

The image displays two overlapping software interfaces. The background interface is QGIS 2.8.3, showing a map of a coastal region with two soil data layers: 'AC18P - Organic Carbon - 0-5cm' and 'AC111 - Bulk Density - 0-5cm'. The foreground interface is the 'Soil IE Demonstrator' web application, which features a map of the same region with numbered sampling points (1-24) overlaid on a grayscale base map. The web interface includes a search bar, a 'Soil Property Query' dropdown, and a 'Base map' selector set to 'Greyscale'. The footer of the web interface states: 'This demonstration portal developed by CeRDI - Federation University Australia'.

QGIS 2.8.3 - Wien

Project Edit View Layer Settings Plugins Vector Symbio Database Web Processing Help

Layers

- ☑ SF_SpatialSamplingFeature
- ☑ AC18P - Organic Carbon - 0-5cm
- ☑ AC111 - Bulk Density - 0-5cm

0.000000
0.600000
1.200000
1.800000
2.400000
3.000000
3.600000
4.200000
4.800000
5.400000

0.000000
0.197701
0.395403
0.593104
0.790806
0.988507
1.186208
1.383909
1.520780

Browser

- Home
- Favourites
- C:/
- D:/
- G:/
- H:/
- L:/
- T:/
- U:/
- X:/
- Y:/
- MySQL
- Oracle
- PostGIS
- SpatiaLite
- OWS
- WCS
- WFS
- WMS

Getting map via WCS.

Soil IE Demonstrator

+ Add layer ⚙ Soil Property Query

Address Search

CeRDI Federation UNIVERSITY AUSTRALIA

Base map: Greyscale

This demonstration portal developed by CeRDI - Federation University Australia

OpenStreetMap contributors © CanOS

Demonstration Three

- Use Cases One and Four
- Soil sampling and observation data

Samples and archiving

Laboratory analyses

Field observations

- Example application

Provision of a comprehensive description of the physical character of a soil at a specific location



Demonstration Three

- Contributors

Landcare Research (NZ)

CSIRO Land and Water (AU)

Federation University of Australia (AU)

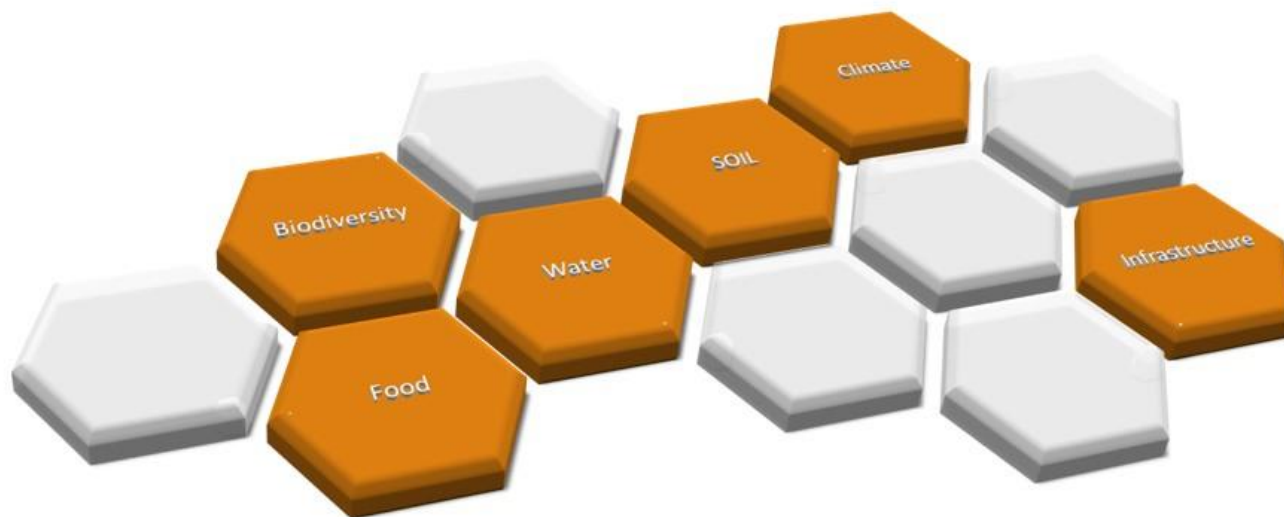
ISRIC World Soil Information (NL)

- Video https://youtu.be/0_gdq4YludU

Landcare Research client with CSIRO Registry

Federation University Australia client with ISRIC WPS

THE WIDER BENEFITS



Useful Tools

- O&M, TimeseriesML and *SoilML Services at LR

The screenshot shows the S-MAP ONLINE web application. The top navigation bar includes 'S-MAP ONLINE', 'Maps', 'Facetsheets', 'Data Provenance', 'Terms of Use', and 'Help'. A search bar is on the left. The main area displays a map of Christchurch, New Zealand, with various layers like 'Context layers', 'Soils', and 'Soil Moisture'. A 'SOIL SUMMARY' panel is visible at the bottom, showing a donut chart for 'Moderately Deep' soil (57% water, 43% clay) and a table for 'PROPERTIES'.

This is a screenshot of an 'S-map Soil Report' generated on 1-Jun-2012. It includes a title, a brief description of the report's purpose, and a table of soil properties. The report is for a location in Kaiafof, with a confidence level of Low.

The screenshot shows the OVERSEER web application login page. It features the OVERSEER logo and navigation buttons for 'Start or Enter', 'Register', and 'Login'. A 'Show help' link is present. The page prompts users to 'Login to OVERSEER Nutrient Budgets' and provides a warning about reading the latest Release Notes. There are input fields for 'Username' and 'Password', along with a 'Remember me' checkbox and a 'Login' button.

The screenshot displays the NDR Viewer interface. It features a search bar and a 'Clear Filters' button. A list of soil types is shown on the left, including 'New Zealand Soil Classification' (Organic, Arable, etc.) and 'Land Use' (Grassland, etc.). A map of New Zealand is shown on the right with several locations marked. The bottom of the interface indicates 'Total Filtered: 1411' and 'Total in map view: 1411'.

This screenshot shows the 'Site Report' interface in the NDR Viewer. It includes a 'Chemical Analysis' table with columns for 'Horizon', 'Sample Interval', and 'Value'. Below this is a 'Soil Classification' section with 'New Zealand Soil Series' and 'New Zealand Generic Classification'. A 'Soil Profile' section shows a vertical cross-section of the soil with depth markers. The 'Locality' and 'Landform' sections provide additional site context.

Useful Tools

- Vocabulary and data dictionary services

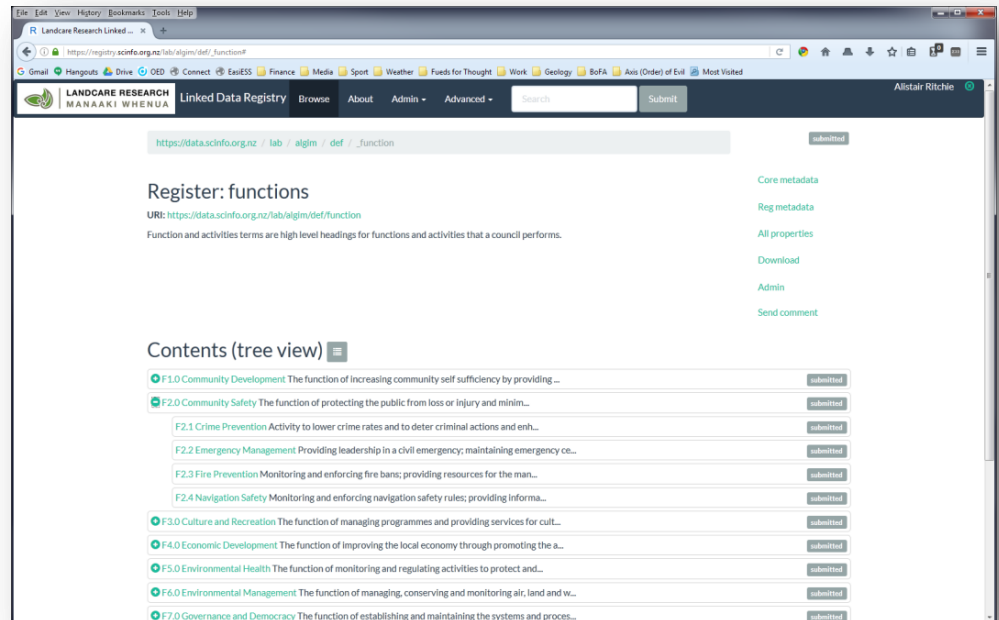
- UKGovLD Linked Data Registry

Shared vocabs and property definitions

Shared classification systems

ALGIM: Laboratory

LR: Landuse and Soil



The screenshot displays a web browser window showing the UKGovLD Linked Data Registry interface. The browser's address bar shows the URL: https://registry.scinfo.org.nz/lab/algim/def/_function. The page header includes the logo for LANDCARE RESEARCH MANAĀKI WHENUA and navigation links for Browse, About, Admin, and Advanced. The main content area displays the following information:

- Register: functions**
- URI:** <https://data.scinfo.org.nz/lab/algim/def/function>
- Description:** Function and activities terms are high level headings for functions and activities that a council performs.

On the right side of the page, there are several action links: Core metadata, Reg metadata, All properties, Download, Admin, and Send comment. Below this, there is a section titled "Contents (tree view)" which lists a series of functions, each with a "submitted" button to its right:

- F1.0 Community Development: The function of increasing community self sufficiency by providing ...
- F2.0 Community Safety: The function of protecting the public from loss or injury and minim...
- F2.1 Crime Prevention: Activity to lower crime rates and to deter criminal actions and enh...
- F2.2 Emergency Management: Providing leadership in a civil emergency; maintaining emergency ce...
- F2.3 Fire Prevention: Monitoring and enforcing fire bans; providing resources for the man...
- F2.4 Navigation Safety: Monitoring and enforcing navigation safety rules; providing Informa...
- F3.0 Culture and Recreation: The function of managing programmes and providing services for cult...
- F4.0 Economic Development: The function of improving the local economy through promoting the a...
- F5.0 Environmental Health: The function of monitoring and regulating activities to protect and...
- F6.0 Environmental Management: The function of managing, conserving and monitoring air, land and w...
- F7.0 Governance and Democracy: The function of establishing and maintaining the systems and proces...

UKGovLD Linked Data Registry

The screenshot shows a web browser window displaying the UKGovLD Linked Data Registry. The browser's address bar shows the URL: https://data.scinfo.org.nz/land/def/use/lucas-2012/_72. The page title is "Register: LUCAS LU Entry: pre-1990 planted forest". The page content includes a URI, a description of the land use classification, and a list of contents (tree view) with items like "settlements tba", "other tba", "cropland NB: not sure if it is really a par", "forest NB: not sure if it is really a paren", "natural forest Areas that since 1 Jan", "pre-1990 planted forest Areas that,", "post-1989 forest Includes post-198", "grassland NB: not sure if it is really a pa", and "wetland NB: not sure if it is really a par".

https://data.scinfo.org.nz/land/def/use/lucas-2012/_72

Register: LUCAS LU Entry: pre-1990 planted forest

URI: https://data.scinfo.org.nz/land/def/use/lucas-2012/_72

Land use classification using Kyoto-compliance regions. The data description can be found

URI: <https://data.scinfo.org.nz/land/def/use/lucas-2012/72>

Areas that, on 1 January 1990, were and presently include: 1/ radiata pine, Douglas-fir, eucalypts, or other planted species (with potential to reach ? 5 m height at maturity in situ) planted before 1 January 1990, or replanted on land which was forest la

Contents (tree view)

- settlements tba
- other tba
- cropland NB: not sure if it is really a par
- forest NB: not sure if it is really a paren
 - natural forest Areas that since 1 Jan
 - pre-1990 planted forest Areas that,
 - post-1989 forest Includes post-198
- grassland NB: not sure if it is really a pa
- wetland NB: not sure if it is really a par

Definition

broader	forest
definition	Areas that, on 1 January 1990, were and presently include: 1/ radiata pine, Douglas-fir, eucalypts, or other planted species (with potential to reach ? 5 m height at maturity in situ) planted before 1 January 1990, or replanted on land which was forest la
description	Areas that, on 1 January 1990, were and presently include: 1/ radiata pine, Douglas-fir, eucalypts, or other planted species (with potential to reach ? 5 m height at maturity in situ) planted before 1 January 1990, or replanted on land which was forest la
label	pre-1990 planted forest
notation	72
pref label	pre-1990 planted forest
type	Concept

Links

Has broader concept

- [forest](#)

LESSONS LEARNT AND WHERE NEXT



In the IE we used open standards to...

- create a simple *information model* of soils data
- *harmonise* the structure and some content of soils data between agencies
- bring data from different international soil agencies together in applications for users (*interoperability*)
- provide a way to describe and organise soil concepts, features, methods, etc (*semantics*)
- use *linked data* ideas to provide supporting information about the data inside the responses to queries
- to chain services for *processing* data
- And demonstrated cross agency *data sharing*

Issues uncovered

- Developing standards is time consuming and expensive
- We need to engage further with the global soil community to reconcile existing models
- The modelling methodology is unfamiliar to scientists
- IE participants reported difficulties deploying services
- Standards technology is not keeping up with web technology
- Assumes open access to data and good governance
- Internationally, it's unclear who should be doing this

Why did Landcare engage globally?

- To ease the use of soil data by our scientists and customers
- We learn from other people's experiences with implementations saving us significant research time and funding
- Ensures we/NZ conform with current best practices and stay aware and engaged with future developments regarding sharing soils data/standards
- Maximises opportunities to achieve research impact
 - *The lessons learnt will be valuable for other data infrastructure activities such as the New Zealand's Biological Heritage, OL&W*
- It's got wider value than just soils
 - *It's directly applicable to data integration and information sharing in infrastructures such as the EMaR/LAWA NFDS*
- Grows New Zealand's expertise and our credibility internationally

Conclusions

- Soil and land information is needed for a wide range of applications
- Data are difficult to access, data interoperability between agencies & their systems even harder
- Standards e.g. ISO 19100 & OGC provide us with the tools and guidelines to improve access, sharing and interoperability
- Linking data to semantic knowledge makes the 'data smarter'
- What we learnt for soils can be applied more widely for the benefit of environmental data sharing in NZ



Alistair

e: ritchiera@landcareresearch.co.nz

t: 03 321 9722

David

e: medyckyj-scott@landcareresearch.co.nz

t: 06 353 4979

Soil IE Engineering Report at <http://www.opengeospatial.org/docs/er>