OGC Environmental Data Interoperability Experiments

Alistair Ritchie
Research Data Architect/Engineer
Manaaki Whenua – Landcare Research
OVERVIEW

• Context
• Soil Data Interoperability Experiment (Soil Data IE)
• Environmental Linked Feature Interoperability Experiment (ELFIE)
• Where next?
Context

- SOIL is essential to ALL life
- The most complex biological material on the planet
- We need to better understand and manage our global soil resources
- We just don’t know enough!
- Urgent need to exchange data and information on our soils
- Need a structured, flexible and long lived global soil information system
  - information architectures, information models and information transfer systems - interoperability
  - web semantics, web services and web information implementation – access and use
- Well defined standards are essential to this system

- Cross out soil add any other environmental feature
- Can’t think in isolation ... no pure, domain-specific information systems
The Open Geospatial Consortium and Interoperability Experiments

• ‘The Open Geospatial Consortium (OGC) is an international industry consortium of over 529 companies, government agencies and universities participating in a consensus process to develop publicly available interface standards.’
  From: http://www.opengeospatial.org/ogc

• New Zealand members:
  - Hawkes Bay Regional Council, Horizons RC, Land Information NZ, Manaaki Whenua, Ministry for the Environment, NIWA

• Interoperability Experiments – standardization by doing

• ‘Brief, low-overhead, formally structured and approved initiatives led and executed by OGC members to achieve specific technical objectives’
  From: http://www.opengeospatial.org/ogc/programs/ip

• Should lead to the formation of a Standards Working Group that moves the IE results to a formal specification
Soil Data Interoperability Experiment

• Started by the IUSS Working Group on Soil Information Standards

• OGC Initiators
  - CSIRO (AU)
  - Manaaki Whenua (NZ - Initiative Manager and Technical Lead)
  - ISRIC World Soil Information (NL)

• Active Participants
  - Federation University of Australia (AU)
  - USDA Natural Resource Conservation Service (US)
  - Agribiology and Pedology Research Centre (IT)
  - USGS (US)
  - Horizons Regional Council (NZ)
  - Tumbling Walls (US)

• Reconcile five existing standards ...

... into a single standard ...
• Not quite ... point to prove ... can use existing standards
Soil Data IE 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. GlobalSoilMap project soil property predictions*

• Use Case 4: pedo-transfer functions
  *Process observed and interpreted soil properties using of pedo-transfer functions - algorithms that calculate additional interpreted soil properties*
Soil IE Exchange Standard

• Reviewed five existing standards
  • Australia and New Zealand Soil Mark-up Language
  • e-SOTER Soil and Terrain Mark-up Language
  • INSPIRE D2.8.III.3 Data Specification on Soil
  • ISO 28258:2013 Soil quality – Digital exchange of soil-related data
  • IUSS/ISO ‘Wageningen Proposal’ (effort to reconcile 1. and 4.)

• No clear candidate from this work

• Back to basics using as much existing work as possible
  • Don’t bind the information model too tightly to technology
    – Abstract and implementation specifications
Soil Descriptions

Podzol Soil (Z) [NZSC]

Ah
E
Bh
C

texture
bulk density
organic carbon
particle size distribution
…
Soil IE Implementation

Vocabularies
Soil Taxonomies
Observable Properties

Registry

WFS
Soil Data

WPS
Execute PTFs

SOS
Sensor Data

PID
HTTP URIs (conneg)

WCS
Gridded Data

WMS
Maps

Client
Demonstration – Soil time series data

OGC Soil Interoperability Experiment

- Properties:
  - Soil Moisture
  - Soil Temperature
  - Rainfall

- Contributors:
  - Manaaki Whenua (NZ)
  - Horizons RC (NZ)
  - USGS (US)
Demonstration – Soil property surfaces

• Contributors
  - CSIRO Land and Water (AU)
  - Federation University of Australia (AU)
Demonstration - Soil descriptions

• Use Cases One and Four
  - Field observations
  - Sampling
  - Laboratory analyses
  - Pedo-transfer functions

• Contributors
  - Manaaki Whenua (NZ)
  - CSIRO Land and Water (AU)
  - Federation University of Australia (AU)
  - ISRIC World Soil Information (NL)
Entry: moderately well drained


A. Soils that have a horizon between 60 and 90 cm of the mineral soil surface with 50% or more low chroma mottles on cut faces or ped faces, or B. Soils that have a horizon between 30 and 90 cm of the mineral soil surface with 2% or more redox segregations.

Definition

description: A. Soils that have a horizon between 60 and 90 cm of the mineral soil surface with 50% or more low chroma mottles on cut faces or ped faces. Or B. Soils that have a horizon between 30 and 90 cm of the mineral soil surface with 2% or more redox segregations.

label: moderately well drained
notation: M
notation: mw

type: Concept
Results

• Created a simple *information model* of soils data
• *Harmonised* the structure and some content of soils data between agencies
• Brought data from different soil agencies together in applications for users (*interoperability*)
• Provided a way to describe and organise soil concepts, features, methods, etc (*semantics*)

• Started a conversation with Dave Blodgett at the USGS

How do we link together all of our environmental data?
How do we do it in a way that is ‘the HTML for environmental data’?
What does that even mean?

- First question ...
- Cross-domain science
- Integrating data from multiple domains
- Lower the data integration overhead
What does that even mean?

- Yeah, but ... ‘the HTML for environmental data’?
- HTML is pervasive web technology and used without thought or controversy.
- OGC X(G)ML and web services are widely regarded as a Bad Thing
- REST and JSON are a Good Thing
- (Depending on your perspective.)
- It’s a pointless argument ... there’s no one way of doing things ... but imagine ...
Environmental Data on the (Google-able) Web

• ... being able to search for a monitoring station by name (location?)
• Get a summary in a Google Knowledge Panel (or equivalent)

• Find out about related features
  - monitoring station -> reach -> river -> catchment -> ...
  - moisture sensor -> paddock -> irrigator -> aquifer -> ...

• Data may be distributed across multiple data stores/agencies

• Standard interoperability problem, but ...
Environmental Data on the Web

• ... needs to be solved in a way that is friendly to web developers

• No strange formats, protocols or bewildering (and numerous) data structures
  - They like ReST and JSON and get scared and angry when dealing with anything else

• But also need to maintain discipline
  - If you want to seamlessly link data across servers the data need to be consistent – standardized
Environmental Linked Feature Interoperability Experiment

- ‘Demonstrate the use of existing and pending OGC standards for the encoding of environmental observation data in an integrated dataset of features linked according to ReSTful and Linked Data principles.’

- Initiators:
  - U.S. Geological Survey (US)
  - Land Information New Zealand (NZ)
  - BRGM (FR)

- Participants
  - Horizons Regional Council (NZ)
  - Manaaki Whenua - Landcare Research (NZ)
  - Natural Resources Canada (CA)
  - ESRI (US)
  - Tumbling Walls (US)
  - Meta-linkage (AU)
  - INSPIRE (EU)

From: [https://github.com/opengeospatial/ELFIE](https://github.com/opengeospatial/ELFIE)
Applying OGC/W3C Best Practices

Spatial Data on the Web Best Practices

SDW-BP provides a refinement of DWBP for spatial data. We use globally unique persistent HTTP URIs for Spatial Things.

- **Best Practice 1:** Use globally unique persistent HTTP URIs for Spatial Things
- **Best Practice 3:** Link resources together using ID and the link
- **Best Practice 4:** Use spatial data encodings that match your target audience
- **Best Practice 9:** Describe relative positioning
- **Best Practice 10:** Use appropriate relations to link Spatial Things
- **Best Practice 11:** Provide information on spatial things
- **Best Practice 12:** Expose spatial data through 'convenience APIs'

'KEEP IT SIMPLE!'

- HTML landing pages with structured data mark-up
- Don’t rely on specialist registries or catalogue services

---

**Note - Other SDW-BP not cited**

- **Best Practice 2:** Make your spatial data indexable by search engines

---

The engine ... the ID *and* the link
https://data.example.org/id/some-thing
'ReST with discipline'

GeoJSON, Well Known Text, ...
JSON-LD – JSON for the Semantic Web

• Need a disciplined approach – some way of standardizing content
  – Properties
  – Definitions
  – Vocabularies
  – Links (aka relations)

• JSON-LD
  – Looks like JSON
  – Linked/embedded ‘context’
  – Binds to RDF Schema/OWL ontology
  – Objects have @ids (URIs)
  – Can be converted to RDF

https://www.w3.org/TR/json-ld/
Use Cases

• NZ: Stream Monitoring Quantity / Quality (impacts of water allocation decisions):
  - Monitoring sites
  - Existing water allocations
  - Diversions / discharges
  - Water quality
  - Blue line network

• NZ: Agriculture (soil moisture state for irrigation application decisions)
  - Soil moisture time series
  - Allowed to pump / consent data.
  - Water allocation

• US: Watershed Boundary Dataset Monitoring Summary
• US and Canada: Monitoring Index
• France: Surface/ground water interaction
• France: Ground water level forecast
• And a bunch more ...
Demonstrations

- Not finished, nothing to be demonstrated (could show you but would have to kill you)
(Preliminary) Results

- Can use JSON (as JSON-LD) and ReST (as Linked Data)
- Data can be linked across servers (well, it is the web)
- Data can be indexed for search engines
- The are useful ‘core’ ontologies to standardize basic content
  - SKOS, schema.org, SOSA, OWL-Time ...
- There are useful OGC application schema for environmental data
  - WaterML 2.0 Parts 1 to 3; GroundwaterML 2; HY_Features; GeoSciML, ...
  - But ... very few of them are RDF-ready (now OWL-RDF ontologies)
- Solves a ‘mass-market’ problem but doesn’t meet all needs
  - Advanced geospatial querying
  - Semantic web interencing/reasons
  - Good entry point though
Where next?

• Honestly not sure

• Soil standards have ground to a halt in international politics (UN FAO etc)
  – No technical problems, all social

• ELFIE ... first we need to finish it
  – Some very interesting follow up projects
  – Gateway drug ... get hooked on standards with the simple stuff, introduce advanced tech when needed

• Can be confident the Semantic Web is part of the future, and here it is ...
QUESTIONS?

Soil IE stuff ...

Engineering Report **OGC16-088r1**
http://www.opengeospatial.org/docs/er

GitHub (model, schema, examples)
https://github.com/opengeospatial/SoilDataIE

Sydney Demonstration – YouTube: https://www.youtube.com/user/ogcvideo