A SMART groundwater portal: An OGC web services orchestration framework for hydrology to improve data access and visualisation in New Zealand

Hermann Kluga, Alexander Kmochb

a University of Salzburg, Interfaculty Department of Geoinformatics (Z_GIS), Schillerstr. 30, Building 15, 3rd Floor, 5020 Salzburg, Austria
b AUT University, Geoinformatics Research Centre (GRC), 640 Great South Road, Manukau 2025, New Zealand

Article info
Article history:
Received 23 October 2013
Received in revised form 29 April 2014
Accepted 30 April 2014
Available online 9 May 2014

Keywords:
WebGIS
Interoperability
Harmonisation
Metadata
CSW
SDI

Abstract
Transboundary and cross-catchment access to hydrological data is the key to designing successful environmental policies and activities. Electronic maps based on distributed databases are fundamental for planning and decision making in all regions and for all spatial and temporal scales. Freshwater is an essential asset in New Zealand (and globally) and the availability as well as accessibility of hydrological information held by or held for public authorities and businesses are becoming a crucial management factor. Access to and visual representation of environmental information for the public is essential for attracting greater awareness of water quality and quantity matters. Detailed interdisciplinary knowledge about the environment is required to ensure that the environmental policy-making community of New Zealand considers regional and local differences of hydrological statuses, while assessing the overall national situation. However, cross-regional and inter-agency sharing of environmental spatial data is complex and challenging. In this article, we firstly provide an overview of the state of the art standard compliant techniques and methodologies for the practical implementation of simple, measurable, achievable, repeatable, and time-based (SMART) hydrological data management principles. Secondly, we contrast international state of the art data management developments with the present status for groundwater information in New Zealand. Finally, for the topics (i) data access and harmonisation, (ii) sensor web enablement and (iii) metadata, we summarise our findings, provide recommendations on future developments and highlight the specific advantages resulting from a seamless view, discovery, access, and analysis of interoperable hydrological information and metadata for decision making.

1. Introduction
Population growth and increasing land use intensity has led to growing demands and exploitation of natural resources in New Zealand. Following the almost total allocation of surface water bodies, groundwater bodies are among the most important and valuable natural resources available, but at the same time they are also the most endangered ones (White, 2001; White and Reeves, 2002; White, 2007). To understand the hydrological state of the environment and groundwater dynamics, datasets and measurements need to be made available and accessible to scientists, planners, and stakeholders to ensure proper decision making.

In New Zealand, datasets are presently archived by different institutions (e.g. Crown Research Institutes, regional councils, and Ministries). These institutions hold spatial data and metadata in various formats that use different nomenclature, storage technologies, interfaces and even languages. This setup complicates searches, discoveries, and accessibility for users. Thus, searching for, obtaining, and pre-processing of datasets consume valuable time and personnel resources on both the data provider and data user side (Beran and Piasecki, 2009; Ames et al., 2012). This has resulted in the need for new data management techniques, tools, and data models to effectively manage the vast amount of new hydrologic data being collected (Goodall et al., 2008; Wojda and Brouyère, 2013).

Data must be collected, transmitted, stored, error checked, manipulated, retrieved for analysis, and shared within the hydrologic community under commonly accepted rules and standards (Carleton et al., 2005; Ranatunga et al., 2011, Kao et al., 2011; Horsburgh et al., 2009; Yang et al., 2010). In order to plan and implement integrated water resources management practices, and to facilitate a rational exploitation and allocation of the available...
resources, in-depth and timely information on the location-based occurrence of groundwater, its particular characteristics, potential risks, and hazards is needed as a future development direction (White, 2006; Kiehle, 2006).

To fulfil the requirements of political initiatives such as the New Zealand Resource Management Act (RMA) from 1991 (Ministry for the Environment New Zealand, 2013a), or the state of the environment reporting in general (Ministry for the Environment New Zealand, 2013b), a sound understanding of groundwater occurrence, recharge, exploitation, and water level changes is required. However, data search and visual or raw data access is hampered by lacking public access to many spatial datasets and the distribution of datasets across 16 regional councils within New Zealand.

The objective of this proposed OGC (Open Geospatial Consortium) web services orchestration framework for hydrology is (1) to update the international scientific community on a holistic approach of hydrological data resource management, (2) to outline procedures of SMART (Simple, Measurable, Achievable, Repeatable, and Time-based) hydrological data management at multiple and cross regional scales, and (3) to compare the scientific state of the art methodologies and techniques with the present development status in New Zealand.

We consider free and open source products, and more importantly, publicly accessible international standards to frame the setup of a New Zealand-wide groundwater knowledge inventory portal. This aims to amass New Zealand’s groundwater information from heterogeneous (distributed) sources into one seamless view across New Zealand. In order to ensure trans-regional, interdisciplinary (hydrology, climatology, land use science, soil science or earth sciences in general), and trans-cultural usability for a wide community of different user groups (e.g. public sector bodies, private companies and citizens), the portal and related services need to consider aspects of multilingualism (e.g. Te Reo Maori support) and data interpretation needs. In this respect, this article considers aspects of data and metadata capturing, (re-)organisation and data harmonisation, as well as semantic and technical interoperability in order to produce seamless geographic information on groundwater for better decision-making support and business utilisation in future.

2. The OGC web services framework for hydrology

2.1. Interoperability criteria and data harmonisation

The interoperability of datasets and services ensures the ability of two or more autonomous (distributed) data sets, data sources or services to be used seamlessly side by side and chained in a processing workflow in a meaningful and consistent way. This needs to occur considering differences in language, content, or context. For example, Groundwater Mapping Units (GMU) as basic geographic mapping elements are defined and mapped differently, depending for instance on regional classifications, mapping rules and/or surveyor’s interpretation.

Harmonisation is considered as an agreed procedure where two individual parts are brought into same structure and semantics to be merged into a beneficial aggregate; we instigate this by creating and using common standards across groundwater related data providers and users. The harmonisation process integrates two basic types of groundwater information: (1) spatial layers either being raster or vector representing the defined GMU, and (2) point datasets on the location of groundwater samples and observations and measurements entailing groundwater properties used for the representation of groundwater thematic maps (e.g. the New Zealand National Groundwater Monitoring Programme, NGMP). The harmonisation of attributes attached to the polygons and points requires cooperation between data providers to make standards uniform and coherent. Moreover, harmonisation is required to ensure interoperability (Laniak et al., 2013). Interoperability refers to the property of a datasets or service, whose structure and interfaces are well defined in order to work with other datasets or services without any restrictions in access or implementation. Fig. 1 visually explains the harmonisation process from different perspectives (technically, logically and semantically).

International groundwater and IT experts are collaborating in making groundwater maps accessible in a seamless digital distribution. Recently, the second Groundwater Interoperability Experiment (GW2IE) chartered by Boyan Brodaric from Natural Resources Canada has been setup (Brodaric and Booth, 2011). The result of this process is an application schema for groundwater data serving as a backbone for international data interoperability.

To enable interoperability, the national or regional data schema need to be mapped to the target schema. The creation of schema mappings is a task for data experts, who are familiar with the semantics and characteristics of their own data and therefore can identify how these relate to the concepts defined in the standardised target schema (e.g. the GW2IE schema). The Humboldt Alignment Editor (HALE, http://community.esdi-humboldt.eu/projects/hale) can be used as an open source tool for creating the map and executing the resulting transformation on the data.

2.2. Data access and visualisation

Within the following OGC web services orchestration framework, we foresee data providers offering their data using web services. According to the World Wide Web Consortium (W3C), web services are self-contained and self-describing application components using XML (eXtensible Markup Language) and HTTP (Hypertext Transfer Protocol) to facilitate communication with other applications using open protocols.

The main OGC compliant web services specifications for spatial data interoperability are the Web Feature Service (WFS, OGC, 2010b) for vector and discrete feature data access, Web Coverage Service (WCS, OGC, 2010a) for continuous coverage and raster data access, Web Map Service (WMS, OGC, 2006) for pre-rendered image data access for web maps, and the Catalogue Service for Web (CSW, OGC, 2007b) for publishing and searching collections of descriptive information (metadata) about geospatial data and services. Fig. 2 shows how the different services and user activities are connected and interdependent and compliant systems are capable of accessing, displaying, and harvesting the distributed spatial data.

The mentioned services are applications that provide data access. Each instance needs to be accessed and queried independently. These services can be connected to a common mediating software component layer, called Enterprise Service Bus (ESB). Then a direct communication between the participating services is possible via the exchange of messages. These messages are typically encoded in XML, and XML schemas describe and constrain structure and content types of such XML encoded documents.

This concept is called a Service Oriented Architecture (SOA) and is a software architecture pattern widely used in business applications (Erickson and Siau, 2008; Huang et al., 2011; Tian and Huang, 2012; Fils et al., 2009; Hildebrandt and Dollner, 2010). Main characteristics of SOAs are the encapsulation of smaller applications. Data providers and/or consumers use these applications as re-usable services. The orchestration of these services to larger applications serves the overall use-case for searching, discovering, accessing and processing of information. Throughout this document we use the phrase ‘use-case’ to refer to any user interaction and use of the web portal (Klug and Kmoch, 2014).
Fig. 1. Achieving technical, logical and semantic interoperability.

Fig. 2. How OGC compliant web services are orchestrated as a framework.
For in situ observations and measurements (e.g., water level measurements), another service component is the sensor interface in the Sensor Observation Service specification (SOS, OGC, 2012). This interface delivers descriptive sensor information in a standardised format using the Sensor Model Language (SensorML, OGC, 2007e), and Observations and Measurements (OGC O&M, OGC, 2011a, ISO 19156, 2011), and gathers/receives measurements from hydro-climate field sensors (Bermudez et al., 2009). SOS is part of the OGC Sensor Web Enablement (SWE, OGC, 2011b) initiative and allows the aggregation of sensor descriptions and observation time-series (Brimoring et al., 2011). The OGC O&M standard describes how time series and observation data can be encoded within OGC compliant interoperable systems. WaterML2.0 is a specialisation of O&M and already widely used to exchange hydrological time-series data.

The GeoScience Markup Language (GeoSciML) was developed as a geology-specific GML application schema under the governance of the international interoperability working group of the Commission for the Management and Application of Geological Sciences (CGI), which is a commission of the International Union of Geological Sciences (IUGS). Recently CGI-IUGS has shifted the further GeoSciML development under the umbrella of the OGC. GeoSciML is a standards-based data format that provides a framework for the application-neutral encoding of geoscience thematic data. This includes mapped geological units, borehole data, and rock specimens, as required by the hydrologists, and allows the querying and exchange of digital, interoperable geospatial information between data providers and users. It allows the communication of data from different formats and different spatial content into one comparable content and format framework.

The Groundwater Markup Language (GWML) is a GML application schema that extends GeoSciML and provides a structural framework for the interoperable exchange of individually defined groundwater related data. The purpose of GWML is to provide a generic, i.e. top-level classification, or the conceptual schema for groundwater data. GWML is also connected to O&M and Sampling Features specifications (OGC, 2007a). It has been designed according to ISO 19103 (2005); ISO 19118 (2011); and (ISO 19136, 2007) and follows the practice of GeoSciML and GML.

The above mentioned specifications permit developments in a Service Oriented Architecture (SOA). The Unified Modelling Language (UML) is a general-purpose modelling language in software engineering and its representation of geographic information follows (ISO 19103, 2005). UML models can be converted into a World Wide Web Consortium (W3C) XML Schema Definition (XSD) document following the ISO 19118 (2011) encoding guidelines. Such XSD documents define the structure and allowed elements of a GML application schema and are used for the assessment and validation of the syntactic compliance of the GML instance documents. These documents are the main messages that are exchanged via the interoperable OGC compliant web services participating in a SOA. The GetCapabilities operation common to all OGC web services provides the service metadata i.e. available requests for a service instance, request parameters and message types for integration into SOA.

2.3. Three dimensional visualisation

All of the above mentioned standards apply to two-dimensional data visualisation. Three-dimensional (3D) visualisation of wells and boreholes is for instance provided by the Canadian groundwater portal GIN (Brodaric et al., 2011) in addition to the 2D spatial visualisation (web mapping) with OGC WMS, and access to the datasets with OGC WFS and the Groundwater Markup Language (GWML, Boisvert and Brodaric, 2012) output encoding described above. This provides additional analysis and discovery services for 3D groundwater processes.

Similarly to HTML4, the HTML5 document structure is accessible via the Document Object Model (DOM). The user can interact with and modify the graphics through a browser through JavaScript. With X3DOM (2012), the German Fraunhofer Institute IGD (Institut für Graphische Datenverarbeitung) created an experimental open source, pure JavaScript framework with extensive X3D support that harnesses the power of X3D, HTML5 canvas and WebGL in a platform independent completely browser-based environment (Behr et al., 2009). X3D is a modern 3D declarative XML-based Markup Language (ISO/IEC, 2008) that evolved from and succeeded the former non-XML language Virtual Reality Modelling Language (VRML, ISO/IEC, 1997).

Based on the combination of X3D, HTML5 canvas and WebGL, several experimental and demonstrative approaches have evolved, e.g. Schmidt and May (2012) and Zhang et al. (2011).

2.4. Metadata for datasets and services

Metadata are data or information on the data itself, and thus refers to structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use or manage datasets, collections, and services. It describes the geographical and domain context, how to interpret the data values, where and when it has been obtained and analysed, who the maintaining institution is and how to get the data. The OGC standard for the Catalogue Service for Web (CSW, OGC, 2007b) describes the web service interface specification for standardised metadata discovery, access and manipulation. The definition of a set of metadata elements is mandatory according to ISO 19139 (2007), allowing identification of the information resource for which the metadata is created. Its classification, as well as the identification of its geographic location and temporal reference, quality and validity, conformity with implementing rules on the interoperability of spatial data sets and services, constraints related to access and use, and the organisation responsible for the resource are to be captured. In addition, metadata elements related to the metadata record itself are necessary. Thus, a specific groundwater metadata schema needs to be determined for datasets and services to extend the following ISO standards.

All defined metadata need to be based on the ISO 19115 (2003) standard, whereas additional service-level metadata should conform to ISO 19119 (2005). For the implementation the metadata for datasets and services are defined in ISO 19139 (2007), which is the derived XML Schema Definition (XSD) representation. Relying on free and open source products and developments, different metadata tools such as GeoNetwork Opensource, the ESRI Geoportal Server or PyCSW can be used to register, upload, and download metadata records as well as related spatial data, documents, PDFs or any other content.

2.5. Thesaurus and gazetteer services

End users can search for hydrological information using key-words, areas, or points of interest. To yield the respective search results, a thesaurus and a gazetteer service infrastructure is required. A thesaurus is a reference work where words are grouped according to their (multilingual) similarity of meaning. It provides a uniform and consistent vocabulary for indexing metadata. A gazetteer is a dictionary or directory referencing place names with their geographical locations.

The Resource Description Framework (RDF) from the W3C provides the foundation for publishing and linking metadata through an RDF Schema (RDFS). The RDF as a data format has its
own query language - the SPARQL query language. SPARQL makes it possible to send queries and receive results through web services and supports HTTP (Huang et al., 2011; Erickson and Siau, 2008; Tian and Huang, 2012; Cruz et al., 2012).

RDFS is supported by the Simple Knowledge Organization System Reference (SKOS), which is a common data model for sharing and linking knowledge via the web (Ma et al., 2011). Thesaurus, taxonomies, classification schemes and subject heading systems share a similar structure and are used in similar applications. SKOS captures these similarities and makes them explicit in order to facilitate data and semantic information exchange across applications. The “ThManager” is a free and open source tool that facilitates the creation and visualisation of SKOS RDF vocabularies (Lacasta et al., 2007).

The Web Ontology Language (OWL) belongs to the family of knowledge representation and has its place in the very young ontology languages (Horrocks et al., 2003; Pulido et al., 2006; Buccella et al., 2011; Stock et al., 2012). The Rule Interchange Format (RIF) is part of the infrastructure for semantic web enablement and facilitates the exchange of diverse existing rules (Zhao et al., 2009).

3. The status of data access and visualisation in New Zealand

The present state of groundwater representation in New Zealand compared to the scientific state of the art is summarised in Table 1. This summary is further detailed in the following sections.

3.1. Available data access and viewing services

Many groundwater related datasets are scattered around the 16 regional and unitary councils, the seven Crown Research Institutes, and commercial companies (e.g. consultancies and energy sector). The different available datasets are partly announced on the websites of the organisations, through web forms or Excel sheets or need to be personally requested. Thus, the establishment of a comprehensive list of available groundwater related resources across New Zealand has not yet been possible.

Data exchange is presently based on personal communication, resulting in usage negotiations and physical data exchange workloads. The data exchange itself is mainly based on file based copies of the requested data repository.

However, since the enactment of the New Zealand Geospatial Strategy (Land Information New Zealand, 2007) several spatial data portals have emerged which provide search, discovery, view and download services based on OGC standards. Land Information New Zealand (LINZ), a government department responsible for land titles, survey systems, topographic and hydrographic information, started the LINZ Data Service, which provides a wealth of publicly-funded topographic and hydrographic datasets for New Zealand, accessible through OGC CSW, WMS, and WFS, as well as a custom Application Programming Interface (API). Landcare Research publishes a considerable number of soil and biodiversity-related maps with a similar solution (LRIS). The LINZ and LRIS portals are backed by a product of the New Zealand based company Koordinates, which also hosts publicly accessible datasets from some New Zealand regional and city councils (e.g. Greater Wellington Regional Council). The National Institute of Water and Atmospheric Research (NIWA) has built a publicly accessible CSW catalogue system and continually registers their datasets, which can now be discovered. However, the access to most datasets is still conducted through manual processes instead of interoperable web services. The Institute of Geological & Nuclear Sciences Ltd (GNS) publishes the New Zealand Quaternary Geological Map (QMAP) not only in print and file copy, but also through GeoSciML/WFS and WMS services. GNS also provides online access to a wealth of other datasets, but in many cases encapsulated in proprietary web applications. The data for the National Groundwater Monitoring Programme (NGMP) is hosted in the Geothermal and Groundwater Database (GGW) at GNS, but the measurements are only accessible through a web form, and not as a web service.

White and Reeves (1999); White et al. (2010), and White et al. (2011) developed three dimensional (hydro-)geological models for a number of areas in New Zealand. Models covering parts of the Bay of Plenty region are viewable online within a custom view and query application, but are not interoperable yet.

3.2. Aspects of data harmonisation

In Europe, the Infrastructure for Spatial Information in the European Community (INSPIRE, Directive, 2007/2/EC), as well as progress and domain specific endeavours facilitate interoperability experiments on data harmonisation across Europe (e.g. soil data harmonisation procedures as described in Morvan et al. (2008); Pourabdollah et al. (2012); Klug and Brectz (2012), while in New Zealand this process needs to be stimulated. Internationally, OGC, ISO, W3C, OASIS (Organisation for the Advancement of Structured Information Standards), Web3D and other organisations provide standards for distributed services for data transformation, discovery, view and best practice for download. New Zealand adopted several of these and also declared them as national standards:

- WMS 1.3; AS/NZS/ISO 19128 (2005)
- GML, OGC (2007d)) v3.2.1 is also known as ISO 19136 (2007)
- UML standardised geographical information and service models ISO 19103 (2005)

National groundwater related data collection programmes such as the New Zealand NGMP harmonise data before they are stored
Internally. However, this is mainly based on a manual acquisition of information from regional councils. Methodologies used to elaborate groundwater related information are to a large extent region and organisation specific. Most existing groundwater related data sets in New Zealand exist in heterogeneous formats, e.g. as ESRI shapefiles, Hilltop XML, geodatabase/file based, MS Excel spread sheets or other (proprietary) formats. Each dataset comes with a specific set of attributes, representing a data provider’s internal data model. Fig. 3 depicts the variety of spatial and temporal datasets and the related challenges that are to be considered for data harmonisation. Integration into the NGMP data repository is based on manual compilations.

Some procedures foresee manual measurement (e.g. water table level) with notes in field books that need to be transferred to digital format later. Some measurements are based on in situ sensors recording for instance water table levels to data loggers. Measurements are then manually transferred to update the respective data repository. Some sensors telemeter the measurements in real time, but usually to data repositories inaccessible by the public or not available through standard compliant procedures. We are not aware of any publicly accessible service using the OGC SOS and the O&M or WaterML2.0 encoding specification for real time telemetering and public access of measurements in New Zealand so far.

For some measurement properties, several closely related analysis procedures exist that differ in values and meaning, but describe the same phenomenon. For example, phosphorus can be measured in different fractions (phosphate, total organic available phosphorus, or total phosphorus). Each of the fractions can be analysed with different methodologies resulting in non-comparable results (Cade-Menun and Lavkulich, 1997).

Many quantitative values are given in classes instead of single values, considering their high spatial or temporal variability. Class definitions differ between regions and are adapted to the entire value range found within their boundaries or to the aims of the individual groundwater investigation. Numbering and naming within the datasets and between different datasets is not consistent. For example, properties regarding the different datasets about wells and springs use “name”, “label”, “wellid” and “id” to reference wells. The elevation nomenclature is also not consistent, e.g. “Z” or “elevation” is used to name the corresponding column. For the groundwater level measurements often “GWL_(MASL)” is used, but also just “Z”. The context is required to recognise if the elevation of the well or the actually measured value is meant. Furthermore it is partly unclear whether the groundwater level is characterised as metres below surface, metres below well head or metres above sea level.

3.3. Available metadata services

A couple of data providers provide standard based services for the search and discovery of their data (LINZ, LRIS, NIWA, Geodata.govtnz, Koordinates). However, many regional and city councils, governmental agencies and research institutes also offer datasets for download on their web sites; yet they either provide metadata as static content on their website, or no metadata at all. These website repositories are
not standard compliant and thus automated harvesting of this information is impossible.

3.4. Thesauri and gazetteer services

In New Zealand, thesaurus and gazetteer services are important in connection with place names since terms are available in English and Te Reo Maori. A list of scientific freshwater terms was hosted on the website of the Hydrological Society of New Zealand as a hydrological glossary, but it was abandoned and is not available any more. However, as mentioned before, LINZ provides a gazetteer service but specific groundwater related places and a groundwater related thesaurus has not been deployed so far.

4. Discussion and recommendations

4.1. Data access, harmonisation, and visualisation

4.1.1. Our findings

New Zealand’s available groundwater related datasets have not been comprehensively listed. We found datasets in different structures and not yet in an interoperable format. Datasets are neither harmonised across data providers, nor do they fulfil the interoperability criteria required for a Spatial Data Infrastructure (SDI). Specific web services based data access capabilities are barely available. Harmonisation efforts resulting in new databases (e.g. NGMP) are labour intensive and only occur in certain time intervals. Even worse, with each copied dataset for the consumer, or a transfer to another database, an out-dated, non-comprehensive, non-updated, and non-maintained instance is created, and further work on these instances is not connected to the original dataset.

Experiences from previous European SDI projects such as the GS Soil project (www.gssoil-portal.eu/), ThermoMap (www.thermomap-project.eu), and the NatureSDIplus project (www.nature-sdi.eu/) have shown that Intellectual Property Rights aspects are considered as one of the most challenging endeavours in the data providing and data exchange process. This is especially true when data is not published under the open data initiative. In New Zealand, access restrictions result from different data collectors, management and user responsibilities and interdependencies. The biggest challenge seems to be the dissimilar licence agreements across the data providers, as well as preoccupation with data quality issues and respective liabilities that might arise.

4.1.2. Our recommendations

For the harmonisation of spatial groundwater related datasets across New Zealand we recommend identifying and selecting important groundwater related datasets to establish a common core data structure (target schema), in order to achieve technical interoperability between the diverse range of existing datasets. This requires definitions of the parameters, properties, types, and coding of the parameter values, explicit disclosure of the respective data quality and possibly a minimum set of metadata elements that comprises auxiliary information for meaningful automated harmonisation procedures and groundwater data quality information. In principle, the knowledge is based on a non-extensible catalogue of objects and rules of how to define attribute parameters for these objects. Parameter definitions according to these rules include, for instance, code lists with relevant explanations for each code and class definitions of classified numerical values. Once the target schema is developed, schema mappings can be applied to ensure the automatic bilateral transformation of the corresponding source data from or into the target schema.

Data providers should be aware that with the emerging web service architecture the existing licence agreements might need to be revisited, because end users might request data repositories at any time. Thus, a user might request the identical data record many times when running a data modelling routine multiple times. On the one hand, the consumer does not want to be billed several times for the same dataset, while on the other hand, those recurrent data requests cause extra network and system load on the data provider side.

4.1.3. Advantages

Interoperable datasets across New Zealand would support the Ministries and regional councils in providing seamlessly available groundwater data across New Zealand. This would enable web mapping services, taking advantage of the latest available datasets at the data providers’ side, to provide a seamless geospatial representation across New Zealand. Thus, information transfer in regard to the resource management act and State of the Environment reporting could be done with less human interference and therefore could save valuable labour resources. Out-dated copies of datasets could be avoided. Timely automated updates of the harmonised NGMP/GGW database would continuously provide the most recent picture of the environment at any time without labour intensive workloads. Valuable further works done on the original datasets at the consumer side can be made accessible to create its full exploitation value. With the integration of Web Processing Services (WPS, OGC, 2007c), statistics and derived environmental indicators could provide near-real time decision-making support once telemetered sensor datasets are available and accessible.

4.2. Sensor Web Enablement

4.2.1. Our findings

To understand the hydrological processes from a local to regional scale, the state of the environment, and the groundwater dynamics, real-time information about the environment is required. Moreover, decisions should be based on the latest findings to avoid irrecoverable damages based on poorly informed decisions. Our observations revealed that some measurements are based on manual field data recording. This procedure is error prone and labour intensive and cannot be done for large areas. Proprietary, non-standard compliant techniques are used for in-situ measurements, causing interoperability challenges due to non-standard compliant data storage and data transmission routines. When data is migrated and converted from the proprietary data collection systems into local processing and analysis systems, second instances of dataset are created, which are not further updated and maintained.

4.2.2. Our recommendation

To ensure a near real time update of, for example, groundwater level or groundwater composition, we recommend employing more telemetered field instruments. These field instruments and telemeter possibilities have not been discussed in this paper, but there are many energy autarchic, low cost, open source and platform independent developments described in the scientific literature. More importantly, it is to be ensured that the distribution and accessibility of measurements occurs in the described OGC and ISV standard compliant ways (SWE/SOS, O&M/WaterML2.0).

4.2.3. Advantages

With the standards based distribution of field measurements, end users would have near real time access to the latest recordings. Databases such as the NGMP could be updated in real time.
(at least for some measurements) based on the automated procedures described above. Decision support could be based on the latest measurements. Integrated in WPS, the resulting indicators and statistics could be accurately provided to users at any time. The labour workload would be decreased by the automated procedures. Based on predefined indicators, decision makers or stakeholders could be automatically informed on harmful developments (e.g. by SMS or Email) when employing techniques from the SWE suite.

4.3. Metadata, thesaurus and gazetteer services

4.3.1. Our findings

Presently, there is no groundwater related thesaurus available in New Zealand and regrettably the valuable public collection of freshwater terms has been removed from the website of the Hydrological Society of New Zealand.

4.3.2. Our recommendation

Despite the underlying differences in database content and historical approaches to groundwater classification and mapping, a common approach should be initiated to ‘translate’ the available local groundwater legends. Common vocabulary for groundwater related maps and content should be developed in collaboration with Maori. This translation step deals with the semantic interoperability required for the contextual and conceptual levels within the web services oriented approach. Such a translation requires the application of common standards to groundwater property descriptions and groundwater analytical results. The property descriptions are increasingly viewed as a pathway to data interoperability between regional councils, Crown Research Institutes (CRIs) and other private and governmental organisations. This procedure should lead to a groundwater related thesaurus which could be based on the removed freshwater glossary from the website of the Hydrological Society of New Zealand. Related Maori and English terms and their adequate mutual translations should be added. To avoid extensive work load in establishing the thesaurus, it should not replace but complement the GEneral Multilingual Environmental Thesaurus (GEMET) developed by the European Environment Agency, while adhering to its syntactical and semantic interoperability. The continuously updated OpenStreetMap (OSM) Nominatim geocoding service and the OpenStreetMap, which is made available for reuse under the Open Database Licence (ODBL) share-alike licence is recommended to complement the LINZ gazetteer.

4.3.3. Advantages

A comprehensive thesaurus is expected to increase the quality of search results for specific datasets at specific places in distributed data repositories. Labour intensive searching for required datasets and information about the dataset itself could be avoided. Related Maori and English terms would support search functionalities in both languages resulting in respective datasets to be found independent from the search language. Once different distributed services are connected, a search is only necessary with one search engine, but the list of results also shows metadata records from other repositories. Searching for specific place names or regions might be more powerful with the OSM gazetteer, as it contains additional volunteered geographical information provided by the public community.

5. Outlook

The SMART (Simple, Measurable, Achievable, Repeatable and Time-based) web service orchestration (Table 2) is intended to become New Zealand’s groundwater related one-stop-shop where users achieve access to available distributed groundwater related resources with a single query. The catalogues storing the data and services are maintained by the data collectors and providers. The datasets remain updated and the respective metadata records are distributed on the fly via a Catalogue Service Web interface. In future, this SMART infrastructure should Save Money and Reduce Time for searching, discovering and accessing spatial groundwater related datasets.

Acknowledgements

The authors wish to acknowledge the six year funding (07/2011–06/2017) of the Ministry of Business, Innovation, and Employment (MBIE), New Zealand. We are also very thankful to the SMART project participants and volunteering regional councils that provide a variety of information and datasets. Works were mainly based on data from the three New Zealand regions Manawatu-Wanganui, Waikato, and Bay of Plenty. Furthermore we would like to acknowledge the on-going collaborations, e.g. through the Google Summer of Code programme, with the 52 North initiative, Germany.

References


