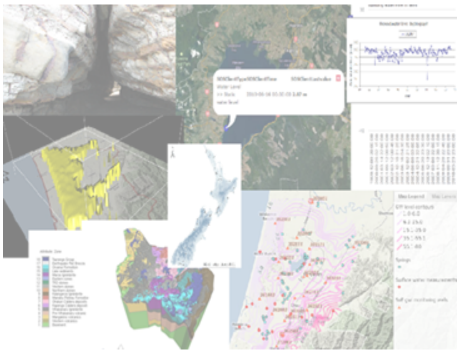
A CONTEXT-BASED GROUNDWATER DATA INFRASTRUCTURE

ALEXANDER KMOCH



A dissertation submitted to
Auckland University of Technology
in fulfilment of the requirements for the degree of
Doctor of Philosophy (PhD)

School of Engineering, Computer and Mathematical Sciences

Auckland– 2017

SUPERVISORS:
Assoc Prof Jacqueline Whalley
Assoc Prof Hermann Klug
Prof Philip Sallis

ABSTRACT

Groundwater bodies are important and valuable natural resources. To better understand the hydrological state of the environment and groundwater dynamics, data sets and measurements need to be made available and accessible to scientists, planners, and stakeholders to allow for proper decision making support. A common challenge in hydrogeological modelling is to discover available data and fit these correctly into required input formats for specific modelling tools. The results need be made available again as inputs for analyses, presentation, and for subsequent use in dependent modelling routines. Information exchange via the internet has become faster, but data sets remain scattered both in location and formats. Present research in hydrogeology and freshwater resources management can be significantly supported and accelerated by relating, reusing and combining existing data sets, models and simulations in a streamlined, computer-aided and networked fashion.

In this thesis Design Science Research (DSR), Grounded Theory (GT) and Case Studies are triangulated in a Geographical Information Science (GIScience) research framework in order to design a Spatial Data Infrastructure (SDI) concept that addresses the full data life cycle in the context of hydrogeology in New Zealand. The 'Hydrogeology Infrastructure' was designed as a distributed system of platform- and location-independent services. It describes which data formats, interfaces and services are required in order to integrate inter-organisational data management, processing, hydrogeological modelling, and visualisation. A series of networked and open standards-based prototypes of the components of the 'Hydrogeology Infrastructure' were implemented, tested, evaluated and discussed. A web-based user interface was developed that demonstrates the access to the distributed functions and services of the infrastructure.

Formerly disconnected and distributed data sets can now be used for hydrogeological data analysis, visualisation and modelling from within one user-facing application. Enabling interoperability of environmental data management tools, scientific modelling routines and data visualisation processes will improve natural resources management and produce better and reproducible environmental knowledge.

CONTENTS

i INTRODUCTION 1
1 INTRODUCTION 2
1.1 Rationale and Motivation 2
1.2 Aim and Objectives 6
1.3 Research Design 7
1.4 Contributions and Publications 11
1.5 Thesis Structure 15
2 LITERATURE REVIEW 18
2.1 Hydrogeology and Open Geodata in New Zealand 19
2.2 Spatial Data Infrastructures 23
2.3 Semantic Web and Linked Data Principles 29
2.4 Sensor Web Enablement and Environmental Sensors Integration 31
2.5 Geoprocessing in the Web and Linked Models 34
2.6 3D Visualisation in the Web 41

ii METHODS AND MATERIALS 45
3 METHODOLOGY 46
3.1 Context and Perspective 46
3.2 Grounded Theory 49
3.2.1 Hypothesis Testing versus Emergence 49
3.2.2 Constant Comparative Method 51
3.2.3 Research Situation, Minor Literature Review 51
3.2.4 Coding and Memoing 52
3.2.5 Sampling and Saturation 53
3.2.6 Sorting and Writing-Up 54
3.3 Design Science Research 54
3.3.1 Awareness 54
3.3.2 Suggestion 56
3.3.3 Development 57
3.3.4 Evaluation 57
3.3.5 Conclusion 57
3.4 Case Studies for Data Collection and Evaluation 57
3.5 Summary 61
4 CASE STUDIES 65

4.1 Horowhenua Local Case Study and Data Sets 66
4.2 New Zealand Regional and National Data Sources 77
4.3 SMART Aquifer characterisation Programme (SAC) 83

iii RESULTS 86

5 A GROUNDED DESIGN THEORY FOR A HYDROGEOLOGY INFRASTRUCTURE 87
5.1 Data Sources 88
5.1.1 Services and Encoding Standards for Data Sources . . . 92
5.1.2 Sensor Web Integration 96
5.1.3 Summary of Selected Types for Data Sources Integration 100
5.2 Metadata 100
5.2.1 Catalogue Service for Metadata Management 102
5.2.2 Metadata Provenance 103
5.2.3 Vocabulary Services and Semantic Convergence 105
5.3 Geoprocessing 108
5.4 Geovisualisation 113
5.5 Summary 118

6 IMPLEMENTATION AND EVALUATION OF THE DEVELOPED DESIGN 120
6.1 Data Sources Integration 120
6.1.1 Groundwater SDI 120
6.1.2 Sensor Web Integration 124
6.2 Metadata Management 132
6.3 Enabling the Model Web through Geoprocessing 139
6.4 Development of 3D Scene Generation as WPS Process 142
6.5 Framework Evaluation 148
6.6 Summary 155

iv DISCUSSION 156

7 DISCUSSION 157
7.1 Improving Groundwater Data Interoperability 157
7.2 Data Sources 159
7.3 Metadata 162
7.4 The Groundwater SDI 166
7.5 Geoprocessing 170
7.6 Geovisualisation 174
7.7 The Hydrogeology Infrastructure 176
7.8 Grounded Design with Case Studies 179

8 CONCLUSION AND OUTLOOK 182

8.1 Answering the Research Questions 183
8.2 Research Contributions and Implications 186
8.3 Outlook and Future Work 187

BIBLIOGRAPHY 189

ACRONYMS 214

V APPENDIX 219

A SWE FOR OBSERVATION DATABASES LISTINGS 220
A.1 SOS Interfaces for CLIDB Database 220
A.2 SOS Interfaces for NGMP Database 221

B SDI INSTANCE DOCUMENTS 223
B.1 GroundWaterML2 Instance Documents 223
B.2 Sensor Web Instance Documents 229

C FRAMEWORK METADATA LISTINGS 236

D GEOVISUALISATION PROCESS 246

E EARTHVISION XYZ EXPORT TO GEOTIFF TRANSFORM 252

LIST OF FIGURES

Figure 1	Conceptual Water Information Cycle	3
Figure 2	GIScience, holistic connecting discipline	8
Figure 3	Triangulation of Methods	10
Figure 4	Graphical Thesis Index	16
Figure 5	Earth Sciences in focus	47
Figure 6	Discrepancies in research	48
Figure 7	Grounded Theory Process	50
Figure 8	Constant Comparative Method	52
Figure 9	Design Science Research Process Model and Cognition	55
Figure 10	DSR Knowledge Contribution Framework	56
Figure 11	Theoretical rooting of GT in Data from Case Studies	59
Figure 12	Design Science Research extended process flow	61
Figure 13	New Zealand regions	65
Figure 14	Lake Horowhenua Conceptual Model	67
Figure 15	Geological Model of Horowhenua	70
Figure 16	Geological Cross Section	78
Figure 17	SMART Aquifer Characterisation Programme	84
Figure 18	Framework Interaction Workflow	88
Figure 19	WPS X3D Process	117
Figure 20	Framework Integration and Interaction	118
Figure 21	Groundwater Portal Frontpage	121
Figure 22	SDI and Portal	121
Figure 23	Groundwater Portal Aquifers Web Map	124
Figure 24	Upper Rangitaiki sensor station	125
Figure 25	Plotting sensor data on the website	126
Figure 26	UML class diagram of NGMP Data Access for 52N SOS	128
Figure 27	NGMP-SOS Access through the Groundwater Portal	129
Figure 28	MQTT SOS/SPS Mediation	131
Figure 29	MQTT SOS/SPS Reconfiguration Message Passing	131
Figure 30	Portal Metadata Distributed Search	134
Figure 31	Portal NZHS Glossary	137
Figure 32	Portal NGMP Observed Property URI Vocab	138
Figure 33	QGIS layered model grid	140
Figure 34	Plots of MODFLOW model	140

Figure 35 Visualising data from different OGC data sources 141
Figure 36 X3D scene generation as WPS process, activity diagram 143
Figure 37 Groundwater Portal 3D Viewer, Terrain 145
Figure 38 Groundwater Portal 3D Viewer, Terrain and Wells . . . 146
Figure 39 Groundwater Portal 3D Viewer, Geological Layers . . . 146
Figure 40 Groundwater Portal 3D Viewer, Subsurface Waterlevel . 147
Figure 41 Implementation Coverage Matrix 150
Figure 42 Evaluation results for the Implementation 152
Figure 43 UML class diagram of X3D scene generation as 52N WPS process 251

LIST OF TABLES

Table 1 Hydrological data management software in New Zealand 21
Table 2 Evaluation Criteria 60
Table 3 Horowhenua case study data coding for geological units, structures and surface 73
Table 4 Horowhenua case study data coding for wells, borelogs, lithology observations 74
Table 5 Horowhenua case study data coding for catchments and management zones 74
Table 6 Horowhenua case study data coding for hydrological and hydrogeological monitoring sites and related measurements 75
Table 7 Horowhenua case study data coding for hydrological and hydrogeological numerical model results/outputs 76
Table 8 Horowhenua case study data coding for map visualisation and unstructured data 76
Table 9 Emerged categories and concepts for data formats and access 101
Table 10 Emerged categories and concepts for metadata formats and access 108
Table 11 MODFLOW Web Processing Service (WPS) 111
Table 12 Emerged categories and concepts for processing 112
Table 13 OGC to X3D mapping 116

Table 14 Mapping NGMP and CLIDB 127

LISTINGS

Listing 1 Service linkage via `parameter` property from WFS to SOS 129
Listing 2 Example SPS Tasking JSON on MQTT from Jenkins . . 132
Listing 3 SensorML Capabilities of 'tasked' procedure from LOW to HIGH 132
Listing 4 NGMP Observed Property 1679 Standing Water Level as RDF/SKOS encoded 138
Listing 5 Excerpts from CLIDB-SOS `CLIDBHelper` class 220
Listing 6 Excerpts from CLIDB-SOS `DescribeSensorDAO` class . . 220
Listing 7 Excerpts from CLIDB-SOS `GetFeatureOfInterestDAO` class 220
Listing 8 Excerpts from CLIDB-SOS `GetObservationDAO` class . . 220
Listing 9 Excerpts from CLIDB-SOS `SosCacheFeederDAO` class . . 220
Listing 10 Excerpts from NGMP-SOS `GGWHelper` class 221
Listing 11 Excerpts from NGMP-SOS `DescribeSensorDAO` class . . 222
Listing 12 Excerpts from NGMP-SOS `GetFeatureOfInterestDAO` class 222
Listing 13 Excerpts from NGMP-SOS `GetObservationDAO` class . . 222
Listing 14 Excerpts from NGMP-SOS `SosCacheFeederDAO` class . . 222
Listing 15 Instance encoding example for catchments/management zones as `gwml2:GW_ManagementArea` 223
Listing 16 Instance encoding example for NGMP monitoring sites as `gwml2:GW_MonitoringSite` which is also a specialisation of `sams:SF_SpatialSamplingFeature` 225
Listing 17 Instance encoding example for Horowwhenua springs data sets as `gwml2:GW_Spring` 225
Listing 18 Instance encoding example for wells with bore lithology as `gwml2:GW_Well` 226
Listing 19 Instance encoding example for a basic EODP sampling location of an NGMP site 229

Listing 20 Instance encoding example for a `SOS/OM_Procedure` sensor description of a soil moisture sensor at the Upper Rangitaiki case study site 231
Listing 21 Instance encoding example for a `SOS/OM_Procedure` process description for groundwater sampling protocol for the NGMP database as generated through the adopted 52N SOS software 233
Listing 22 Instance example of SMART Rainfall Recharge dataset ANZLIC/ISO metadata record 236
Listing 23 OWC Context example New Zealand overview 238
Listing 24 Instance example of JoHNZ ANZLIC/ISO metadata record 240
Listing 25 RDF/SKOS encoded NZHS Freshwater Glossary 244
Listing 26 RDF/SKOS encoded NGMP Observed Properties from NGMP Parameters 244
Listing 27 Excerpts from `WmcSupervisionAlgorithm` class 246
Listing 28 Excerpts from `X3DUtil` class 246
Listing 29 Excerpts from `WmcParserFactory` class 247
Listing 30 Excerpts from a Horowhenua layer context document . 247
Listing 31 Excerpts from X3D scene encoding for the Horowhenua scene 248
Listing 32 Excerpts from `WmcTest` 249
Listing 33 Excerpts from portal Scala views Scala template 250
Listing 34 Excerpts from EarthVision to GeoTIFF transform program 252

DECLARATION

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.

Auckland, 2017

Alexander Kmoch

ACKNOWLEDGMENTS

I would like to express my profound gratitude to my supervisors, Associate Professors Jacqueline Whalley and Hermann Klug, and Professor Philip Sallis, for their guidance and advice, for letting me work freely, yet keeping me from straying too far. The combination of you as my supervisory team strongly reflects the interdisciplinary fields I am aiming to link through this work. At the junction of your academic convictions and conventions it was where the fertile grounds were created on which I had to reshape my thinking over and over again.

I am very grateful to my family, who always believed in me, supported me and never let me down. Even though I was disagreeable, distracted and unavailable oh so often, I love you very much.

I am very thankful to Horizons Regional Council, Brent Watson, Sean Hodges, Abby Matthews and Jeff Watson for the rich data sets from the Horowhenua reports.

Thanks to my fellow colleagues at GNS Science, Paul White, Magali Moreau, Dr Stewart Cameron and Dr Chris Daughney, for their assistance in all matters hydrogeology, liaising with regional councils and their imperturbable belief in the advancements and innovation resulting from my work.

I am forever thankful to Dr Evelyn Uuemaa for help on proof-reading, structure and graphics in the dissertation and her continuous demand to simplify explanations in the text.

Thanks also to Akbar Ghobakhlou and the late GRC colleagues, Dr Subana Shanmuganathan and Kitty Ko for taking me in and their on-going support.

Thank you, Steffen Reichel for your on-going critique and mentorship of my software development skills.

Thanks to Dr Kevin Lee for the refreshing philosophical discussions about science, knowledge and ontologies.

Thanks to fellow geo data folks at Landcare, NIWA and LINZ, especially Alistair Ritchie, Dr Jochen Schmidt, Robert Gibb, Andrew Watkins and David Medyckyj-Scott for inspiring conversations, idea bouncing and joint pursuit towards a standards-based New Zealand environmental geodata infrastructure.

I am also very thankful to the SMART project participants and volunteering regional councils that provided a variety of information and data sets. Furthermore, I would like to acknowledge the collaborations on the SOS server developments with the 52°North initiative, Germany.

I am very grateful to the members of the OGC GroundwaterSWG and GW2IE for allowing me to observe, participate in, and contribute to the development of GWML2.

Finally, I wish to acknowledge the six year funding (07/2011-06/2017) of the Ministry of Business, Innovation, and Employment (MBIE), New Zealand, contract number C05X1102, for the SMART Aquifer characterisation (SAC) programme which sponsored my scholarship.

Part I

INTRODUCTION

1 INTRODUCTION

1.1 RATIONALE AND MOTIVATION

Groundwater resources in New Zealand are under pressure, but not adequately understood (Lowry et al., 2003; White & Reeves, 2002). Large economic value comes from water supplies to agriculture, industry and domestic users (White, 2001; White, Sharp, & Kerr, 2001). The many environmental benefits of water resources also include the maintenance of aquatic ecosystems and provision of recreation opportunities (Harding, Mosley, Pearson, & Sorrell, 2004). Currently, both surface water and groundwater resources are under pressure from development, e.g., as demonstrated by an approximate 50 percent increase in allocation between 1999 and 2006 (Ministry for the Environment, 2006; White, 2006). 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, 2007; White & Reeves, 2002).

To understand the hydrological state of the environment and groundwater dynamics, data sets and measurements need to be made available and accessible to scientists, planners, and stakeholders to allow for proper decision making support. Data must be collected, transmitted, stored, error checked, manipulated, retrieved for analysis, and shared within the hydrological community under commonly accepted rules and standards (Carleton, Dahlgren, & Tate, 2005; Horsburgh et al., 2009; Kao, Ranatunga, Squire, Pratt, & Dee, 2011; Ranatunga, Walker, & Sheahan, 2011; Zhang, Meratnia, & Havinga, 2010). Fig. 1 illustrates that many agencies and organisations are involved in this data management process and depend on each other's domain expertise, data and information. Thus, the entirety of services they provide can be viewed as a system, and this system needs to have multiple components that run on different machines (in different locations, networks, organisations) in order to complete its tasks and be fully functional. 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 occurrence and spatial properties of groundwater bodies, their particular characteristics, potential risks, and hazards is required (Kiehle, 2006; White, 2006). This indicates the need for future development of integrated management practices and information systems.

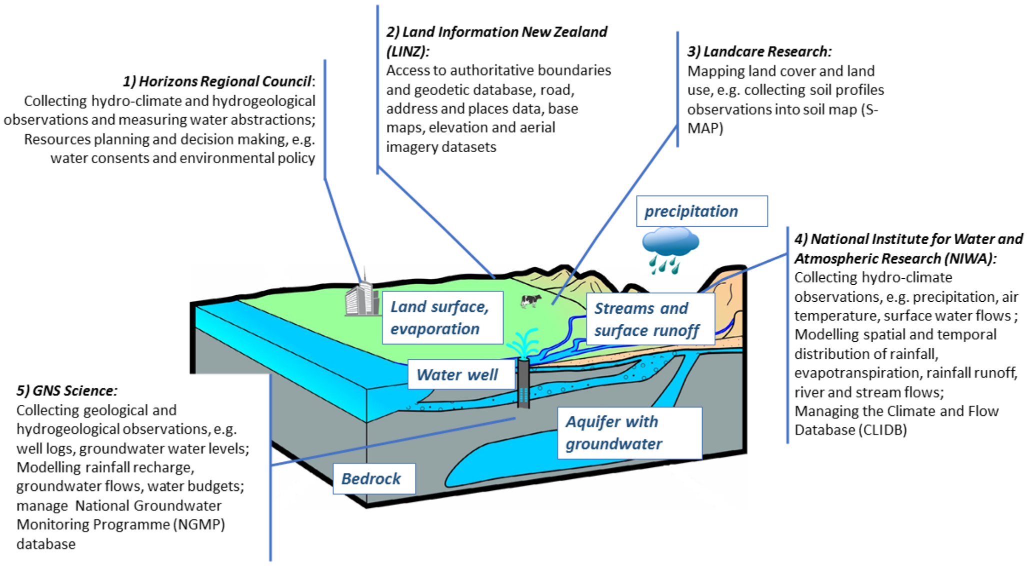


Figure 1: Conceptual presentation of the distributed nature of water information (observations, datasets, models) collectors, providers and users participating in planning and implementing integrated water resources management. In order for Horizons RC (1) to make informed decisions on groundwater pumping allocation, data transfer and modelling needs to be followed through from 1) to 5) in order for GNS Science to develop a useful groundwater flow model.

Up to fifty per cent of the time and cost investments in projects involving spatio-temporal data are related to searching, discovering, accessing, understanding, transforming and preprocessing of relevant data sets, before the data can actually be used and the project progress (Bandaragoda, Tarboton, & Maidment, 2006). Searching for, obtaining, and pre-processing of data sets consumes valuable time and personnel resources for both the data provider and data user (Ames et al., 2012; Beran & Piasecki, 2009).

In New Zealand this need has been discussed by Lowry et al. (2003). Besides reports and research articles, the actual collected and derived data as well as the software model codes have often not been published or otherwise made accessible and are being archived in disparate internal computer systems of the independent organisations (William, Milke, & Raffensperger, 2009). To improve this situation an integrated environmental modelling approach has become a vision or roadmap for the future (Granell, Díaz, Schade, Ostländer, & Huerta, 2013; Laniak et al., 2013).

Spatially continuous modelling exercises across administrative, state and/or organisational boundaries are currently impossible because interoperable - technically and semantically harmonised and standardised - data sets and modelling interfaces are not available (Horsburgh, Tarboton, Hooper, & Zaslavsky, 2014). Furthermore, software routines for environmental modelling often are either not publicly or freely accessible or they don't they support standardised data interfaces for their integration, which hampers repeatability, transferability and automated interchanges of model results (Granell et al., 2013).

The SDI paradigm is well described and in general a well understood and documented methodology that has been applied globally, governmental and private organisations, agencies and communities. New Zealand's geospatial strategy is published and CRIs and regional councils are increasingly building on that foundation and contribute data. Regarding the particular domain of groundwater there is not an information gap itself, but the discovery, accessibility and in particular web-based visualisation is limited. As described earlier, initiatives in North America, Europe and Australia have implemented and scientifically documented methodologies, data models and technologies.

Environmental telemetry is applied and used throughout New Zealand. However, these telemetry solutions and sensor networks are often proprietary, commercial black box systems. Data collection and aggregation happens via closed system architectures or data loggers, which are collected manually and therefore, are not available in real-time relative to their original sampling frequency. Under such data management practices the data sets are eventually consolidated within the disparate data infrastructures of their custodians. Nevertheless, as described earlier regions and research institutes start to open up such data sets to make them accessible through an SDI. Therefore the SDI paradigm not only suitable, but a necessary foundation to be considered.

Gap 1: Groundwater information in New Zealand is scattered across regions, agencies and organisations. Data sets, measurements and processing results exist, but there is no consistent way to discover and access different groundwater related data sets across administrative or hydrological boundaries in New Zealand.

Web-based processing has been documented. Open Geospatial Consortium (OGC) WPS as web service interface and several open source software packages have been evaluated to enable the operational use of geospatial and hydrological algorithms in the web. Also the use of OGC-based input data sources for WPS-based models has been described. Interfaces and frameworks like Open

Modelling Interface (OpenMI)2 and the open modelling system Object Modelling System (OMS)3 have been demonstrated to enable inter-connection of models. However the flow of input data into such environmental modelling frameworks and its attached models, and then the return of the output data into the SDI as a new data set, is not well described yet.

Gap 2: An automated web-based integration with groundwater flow and water balance models has partially been described in few international scientific publications. Yet input data transformation has always been constrained to support one fixed modelling configuration for the respective case study. Furthermore, the results are mainly used for presentation and visualisation purposes, not for dynamically coupled re-use in another dependent model.

Gap 3: The reusability of processed data within the SDI is limited, and final processing results are only used for presentation and reporting. Models that require other models as functional input, like groundwater flow and water balance models have not yet been described to run within a dynamic web-based on-demand processing environment.

Several large scale 'multi-system' environmental or geoscientific web- and web services-based geospatial data processing and visualisation frameworks have been described. For example, Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI)-HIS system is very sophisticated and provides a lot of processing and visualisation functionality within an accompanying desktop program. However, the CUAHSI-HIS system is limited when used for processing and visualising hydrogeological rather than surface hydrology data. The Canadian Groundwater Information Network targets national hydrological data management but has less focus on scientific data visualisation and no inherent modelling capabilities.

As a result of the literature review it becomes obvious that many of the independent components required for hydrogeological data storage, processing or visualisation already exist but that these components were neither developed to work together in a single harmonised data infrastructure nor in a distributed networked fashion. No overarching frameworks have applied all available techniques, methods and standards consistently in a distributed and completely web-based hydrogeology infrastructure.

OGC standards are well established building blocks in numerous research publications in the last decade and constantly improved in an iterative participatory public process. The integration of the disparate and diverse data sources with groundwater flow and water balance models in a networked on-demand fashion is possible with OGC standards in order to improve automation and reproducability and consequently, support water management decision making (Latre et al., 2013).

1.2 AIM AND OBJECTIVES

The primary objective of this research is to enable a holistic and integrated view of groundwater and groundwater-related data, models and hydrogeological research. The study aims to develop a novel research method for designing SDIs in order to create a new data enriched perspective on New Zealand's groundwater resources. Subsequently, it will be reviewed which standards need to be integrated and how these standards can be implemented in an SDI software framework. Furthermore, this thesis aims to identify and solve gaps, issues and ambiguities with the existing and to be developed/refined suites of (inter-)national standards, models and techniques that might inhibit or exacerbate their application for use in a comprehensive hydrogeological data framework.

Objectives:

1. Review methods, standards, formats and encodings in the literature and in relevant applied technologies and software systems for integrated groundwater data management, analytics, simulation and visualisation in regards to gaps revealed from the literature and compare to New Zealand groundwater data management practices and analysis tools.

2. Identify, select and discuss an appropriate set for a distributed/networked groundwater data and modelling infrastructure and design a reference implementation.

3. Advance theory for designing interoperable data and analytics infrastructures based on the selected methods, standards, formats and technologies and address current challenges and short comings in existing methods.

It is not an objective of this thesis to create or model new standards or data models. However, participating in the community process around the development of a standard data model through the Groundwater Interoperability Experiment 2 (GW2IE) and the New Zealand National Environmental Monitoring Standards (NEMS) / Environmental Observation Data Profile (EODP) working group has provided insight into the practice of domain experts.

Pragmatic software engineering practices are applied to develop a functional web-based prototype system. The design and interoperability of the software implementation is evaluated on the case studies in order to ascertain the credibility, transferability, dependability and confirmability of the underlying theory.

1.3 RESEARCH DESIGN

One of the big goals of GIScience is to improve the use and reuse of (geographical) data in scientific analysis, visualisation and modelling, and the formal methods thereof, which are required to become more interdisciplinary and interconnected to solve the challenges of our time (Goodchild, 1992; Mark, 2002). While GIScience has been disputed for being an independent scientific discipline for a period around its emergence (Pickles, 1997), its interdisciplinary nature is of more relevance than ever (Goodchild, 2006, 2009; Yang, Raskin, Goodchild, & Gahegan, 2010). This study is intended to contribute to the field of GIScience, in particular on the interdisciplinary information science aspects, where methods and techniques around a Digital Earth, spatial and semantic data interoperability for visualisation, distributed geoprocessing and knowledge engineering are essential (Blaschke, Strobl, & Donert, 2011; Goodchild, 2009). Figure 2 shows how GIScience can lead interdisciplinary research that encompasses knowledge from the Earth Sciences and Computer Science.

In summary, groundwater information in New Zealand is scattered across regions, agencies and organisations. Data sets, measurements and processing results exist, but there is no consistent way to discover and access different groundwater related data sets across administrative or hydrological boundaries in New Zealand. Thus, the capability for reuse of the processed data is limited, as final processing results are only used in presentation or external reporting. Models that require other models as functional input cannot run within a dynamic web-based on-demand processing environment to ensure up-to-date requested information on environmental conditions. However, application of (predominantly OGC/ISO) standards are established building blocks in numerous research publications in recent years and constantly improved in an iterative participatory public process. But for overarching frameworks, none or only few have applied available techniques, methods and standards consistently as a distributed web-based framework theory. Putting that into perspective with the foundations and research aims of GIScience, consequently the following Research Questions are formulated

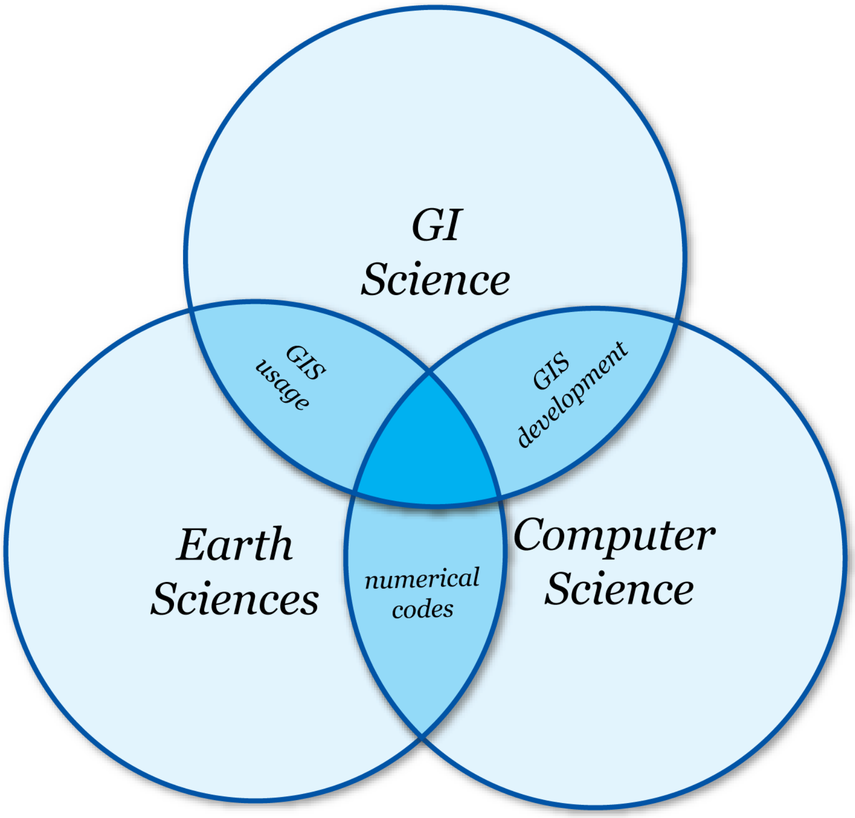


Figure 2: GIScience, as an holistic connecting discipline lays claim to the central cross section

for this study:

Question 1: What are the required interfaces and data formats to comprehensively capture describe and share hydrogeological and groundwater related data sets in New Zealand to enable data interoperability for hydrogeology ('A Groundwater SDI')?

Question 2: Can existing scientific groundwater flow and water balance simulation models and software modules be linked into a networked environment to enable interoperable models ('A Model Web')?

Question 3: Can such models be provisioned with available interoperable data sets ('Context-based Data Provisioning')?

Question 4: Can the visualisation of interoperable data and models be realised with the same design principles ('Visualisation as an interoperable Model')?

Question 5: Does the new (theory of the) framework design overcome the challenges and closes the gaps from the literature?

The starting point to address the objectives of this study and answer the research questions outlined above is 1) the ever increasing collection of specifically hydrogeological as well as environmental data in general, in a wide variety of formats, 2) the evolving data formats, data and database models and formalised standards and 3) their documented application to data collection, storage, harmonisation, discovery, access and transfer in the scientific literature. Thus, the body of data analysed in this study is scientific literature, formal standards developed by expert groups and the data sets and technologies from the case studies.

All of those are in fact human artefacts, for which the author needs to adopt a research paradigm to address the subjectivity involved with handling a possible inherent bias (Creswell, 2003; Holliday, 2010; Saunders, Lewis, & Thornhill, 2012). But as compared to unfiltered statements like from interviews or questionnaires, research articles and formal standards have been prepared by experts in their field and already passed through a process of peer review and refinement. Here, critical realism will act as a bridge between positivist and interpretivist stances and guide the process of theorising throughout this thesis under the assumption that knowledge is not value free and bias should be articulated accordingly (Bhaskar, 1998; Dobson, 2002; Mingers, 2004). The modes of inquiry that are going to be used are qualitative in nature and comprise GT, DSR and case studies (Beck, Weber, & Gregory, 2013; Glaser & Strauss, 1967; Yin, 1994). These three methods are applied jointly in triangulation, as show in Figure 3.

GT will serve as the overall guiding process for exploring, comparing and explaining observations and distil emerging patterns, as well as for the critical evaluation of the design process and the generated artefacts. The researcher attempts to derive a general theory to solve the challenging interdisciplinary and cross-domain research objectives, grounded in the views and formalisations expressed in the body of data for this study. This will involve iterative stages of data collection and refinement of interrelationship of categories of information (Strauss & Corbin, 1994, 1998). This process of developing theory is going to be the prime informant for the iterative design stages of the envisioned integration framework (Gregor, 2006; Gregory, 2010). Emerging categories and theoretical sampling of the similarities and the differences of information (implemented standards, addressed problems, orchestrated au-

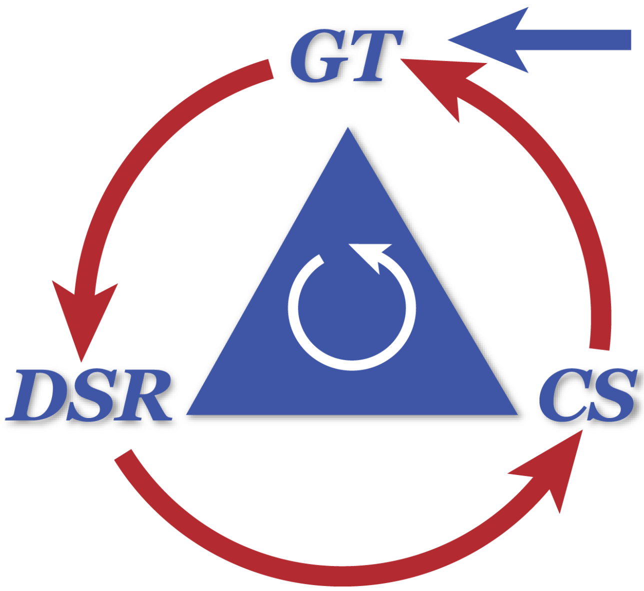


Figure 3: The research methods Grounded Theory (GT), Design Science Research (DSR), and Case Studies (CS) are triangulated resulting in a unique research framework (coined as Grounded Design).

tomated services) will be constantly compared with the framework design and thereof generated artefacts to solve the articulated research questions for the domains of GIScience and Hydrogeology.

While GT provides the means to reason over all data as they become available, this study is bound to the notion of Case Studies twofold: 1) As a practical consideration this research is going to be applied to GIScience for Hydrogeology in New Zealand in geographical case study areas on a local scale, regional scale and national scale. 2) As qualitative strategy of inquiry, in which the researcher explores in depth the events, activities and processes that generate and consume data sets in hydrogeological research. The case studies are constrained by availability of distributed data sets and scientific models and codes for the geographical case study areas over the time frame of this study, which are the data from specific groundwater research projects from the Institute of Geological and Nuclear Sciences GNS and Horizons Regional Council, climate data from the National Institute for Water and Atmospheric Research NIWA and the various emerging geoportals in New Zealand.

Research methodology and theoretical framework are described in depth in Chapter 3. Here the rationale, background and development of the new 'Grounded Design' method is explained. This method is then applied over the literature review (Chapter 2) and the case studies (Chapter 4) in order derive the knowledge necessary in order to design an SDI concept that will provide answers to the aforementioned research questions. The resulting theory and design decision are extensively elaborated on in Chapter 5. Eventually, the instantiation of a real world implementation of the developed design is described and tested in Chapter 6.

1.4 contributions and publications

The time horizon of three years is indicating a constant change within the domain of research of this study. The field evolves quickly, and becomes self-reliant on its former state. I believe Grounded Theory was advantageous over a purely empirical positivist approach where a quantitative analysis of the literature, standards and technologies might not have provided a reliable way of developing the theory and inform an adaptive dynamic design process.

Facts, which I concluded along the way of this study, are now partially happening with and without my interference in the domain, e.g. expert communities developing and/or adopting new data and service integration models that were not envisioned three years ago. My contributions to the literature along the way since the inception of this dissertation can be seen as current additions to theory at that time and do not represent timeless and authoritative parts of this dissertation. In the following paragraphs I describe my contributions to the literature, how much I contributed to the single research articles and how the status and content is relevant in the context of this thesis. This will be of particular interest to the reader: Although these articles and their described technical framework artefacts were accepted as valid contribution to the overall body of knowledge in the field, some results were subsequently advanced and partly superseded in the final dissertation through the Grounded Theory process which feeds back into the course of the design process.

1. Ghobakhlou, A., **Kmoch, A.**, & Sallis, P. (2013). Integration of Wireless Sensor Network and Web Services. In J. Piantadosi, Anderssen, R. S., & J. Boland (Eds.), *MODSIM2013, 20th International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December 2013, ISBN: 978-0-9872143-3-1.* doi: 10.13140/2.1.3508.0001

   This paper describes integration of Wireless Sensor Networks (WSNs) into a Service Oriented Architecture (SOA) by proposing a web service proxy linkage from the low level sensor network Message Queue Telemetry Transport (MQTT) to the high level the OGC Sensor Web Enablement (SWE) semantics to treat sensors in an interoperable, platform-independent and uniform way. My specific contribution proposes the mapping OGC SWE semantics into the low-level MQTT protocol for interoperable, bi-directional interface specifications and (meta) data encodings for the real time integration of sensors and sensor networks into a web services architecture. This has been refined in 5) Klug, Kmoch, and Reichel (2015) were the needs of bi-directional sensor web integration were found crucial to implement operationalisation as described in 4) Klug and Kmoch (2015).

2. Klug, H., & **Kmoch, A.** (2014). A SMART groundwater portal: An OGC web services orchestration framework for hydrology to improve data access and visualisation in New Zealand. *Computers & Geosciences, 69(0), 78-86.* doi: 10.1016/j.cageo.2014.04.016

   This research article comprises a fundamental review of the existing literature at the time. The authors 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, international best practices in data management developments are contrasted with the current situation for groundwater information in New Zealand. Finally, for the topics (i) data access and harmonisation, (ii) sensor web enablement and (iii) metadata, the findings are summarised and recommendations on future developments are formulated. My main contributions are large parts of literature review and the comparison of technological details, as well as informing parts of the conclusions and recommendations to reflect a logical and coherent use of standards.

3. **Kmoch, A.**, & Klug, H. (2014). Visualization of 3D Hydrogeological Data in the Web. *GI_Forum 2014 - Geospatial Innovation for Society, (2014), 16-24.* doi: 10.1553/giscience-2014s16.Visualization

   In this research article the authors demonstrate a framework to enable a web-based (platform independent) retrieval and visualisation of three-dimensional hydrogeological information via the web browser. This research article represents my vision on distributed data and processing services to prepare an on-demand 3D visualisation of geological and hydrological data. Here I firstly mention and demonstrate the 'context-based' encapsulation of referenced data sources as means of 'transport-by-reference' to geoprocessing services via OGC WPS. Furthermore, I explicitly treat the generation of visual output as a specialised form of geoprocessing.

4. Klug, H., & **Kmoch, A.** (2015). Operationalizing environmental indicators for real time multi-purpose decision
making and action support. *Ecological Modelling, 295(0), 66-74.*
doi: 10.1016/j.ecolmodel.2014.04.009

 In this research article the authors make the claim how numerous indicators have been proposed and operationalised using computing techniques, yet many of the approaches are hampered due to incompatible data formats, data availability limitations, and/or unavailable modelling routines. Klug proposes the major frame from a landscape ecology perspective and provides a heuristic conceptual basis for driving the next generation of real-time multi-purpose data assembling, evaluating, modelling, and visualisation towards the operationalisation of decisions. My distinct contribution here is advancing integrative theory on how publicly available and standardised environmental information as OGC compliant Sensor Observation Services (SOSs) with its data formats Observations & Measurements and Water Markup Language 2.0 automatically feed into WPSs for timely information delivery, discovery and access of the spatially explicit environmental conditions.

5. **Kmoch, A.**, Klug, H., Ritchie, A. B. H., Schmidt, J., & White, P. A. (2015). A Spatial Data Infrastructure Approach for the Characterization of New Zealand's Groundwater Systems. *Transactions in GIS, Early View.*
doi: 10.1111/tgis.12171

 In this research article the authors propose how the future information needs of stakeholders for hydrogeological and hydro-climate data management and assessment in New Zealand may be met with an OGC compliant publicly accessible web services framework, which aims to provide integrated use of groundwater information and environmental observation data in general. The authors jointly document stakeholder and community interaction processes and the development of a draft standard for an EODP for consideration for New Zealand NEMS. Besides managing and writing the paper, my distinct contribution with essential relevance to this thesis is the conceptual mapping (and experimental implementation) of the Institute of Geological and Nuclear Sciences (GNS) National Groundwater Monitoring Programme (NGMP) database and the National Institue for Water and Atmospheric Research (NIWA)

Climate and Flow Database (CLIDB) from their legacy data schemas into interoperable OGC data services and thus, the integration into the groundwater data framework.

6. Klug, H., **Kmoch, A.**, & Reichel, S. (2015). Adjusting the Frequency of Automated Phosphorus Measurements to Environmental Conditions. *GI_Forum 2015 - Journal for Geographic Information Science - Geospatial Minds for Society, 1, 590-599.* doi: 10.1553/giscience2015s590

   In this paper the authors provide a framework where low frequency phosphorus measurements in the Mondsee catchment can be adapted to high frequency measurements during storm events. When heavy rainfall is observed, a threshold event triggers a reconfiguration task for the phosphorus measurement device. While Klug's main interest here lies in the application 'Operationalizing environmental indicators' (Klug & Kmoch, 2015), the framework described here also continues and refines the OGC SWE (SOS/SPS) and MQTT 'Integration of Wireless Sensor Network and Web Services' (Ghobakhlou, Kmoch, & Sallis, 2013). My particular contribution is the advancing of the 'SOS/SPS-on-MQTT' design and the automated integration of geoprocessing via OGC WPS for the realisation of the indicator threshold process.

7. **Kmoch, A.**, Klug, H., White, P., Whalley, J., & Sallis, P. (2015). Bridging Domains: A web-enabled interoperable Groundwater Research Framework. *In 9th Symposium of the International Society for Digital Earth (ISDE). Halifax.*

   This poster presentation illustrated a networked and open standards-based integration of storing, accessing and transformation of distributed hydrogeological data (OGC-based groundwater SDI) with domain specific software modules and processes (USGS MODFLOW exposed as a WPS process) in a case study in New Zealand. My contribution comprises article text as well as the underlying conceptual and technical work. The co-authors provided editorial support and advice from their environmental domain perspective. The particular importance of this poster was the prototyped WPS integration of widely used hydrogeological simulation or modelling codes for predictions and decision support with the data from an SDI within a web-based an open standards context.

The advancements of theory through my contributions in these publications are snapshots in time, additional artefacts, and are acted on, discussed and advanced throughout the course of this dissertation. They are inputs to the Grounded Theory process and the final comprehensive framework design presented in the following chapters.

1.5 THESIS STRUCTURE

Figure 4 shows a graphical index of the contents and structure of this thesis.

Chapter 1 - Introduction: This chapter establishes the context of this study and provides an overview of the research contribution, the methodology and the structure of this dissertation.

Chapter 2 - Literature Review: This chapter presents an exhaustive literature review of the current state in the relevant fields. In the light of the guiding methodological principle, which is Grounded Theory, an extensive literature upfront is typically not recommended. However, for the practical structure and academic requirements regarding this dissertation the complete literature review is presented here. Literature regarding the context and rationale of this thesis which has been described in the introduction (Chapter 1), as well as methodological literature referring to the research design (Chapter 3) is not reproduced here.

Chapter 3 - Methodology: This chapter elaborates on the overall research methodology employed to undertake this study. The interplay of Grounded Theory, Design Science and Case Studies is explained in detail.

Chapter 4 - Case Studies: This chapter describes the case studies used in this thesis. The conceptualisation strategy is demonstrated from sampling from case study data and literature towards generating categories and emerging patterns.

Chapter 5 - Framework Design: This chapter represents the developed theory and how this theory informed the framework design. In Grounded Theory it would represent the write-up, which is the first and foremost result of a pure Grounded Theory study. However, in this thesis the aim is not only to develop theory to advance understanding of current approaches, but also to advance theory to point to solutions as they emerge. Through the transdisciplinary nature of this study it will become apparent that through the data

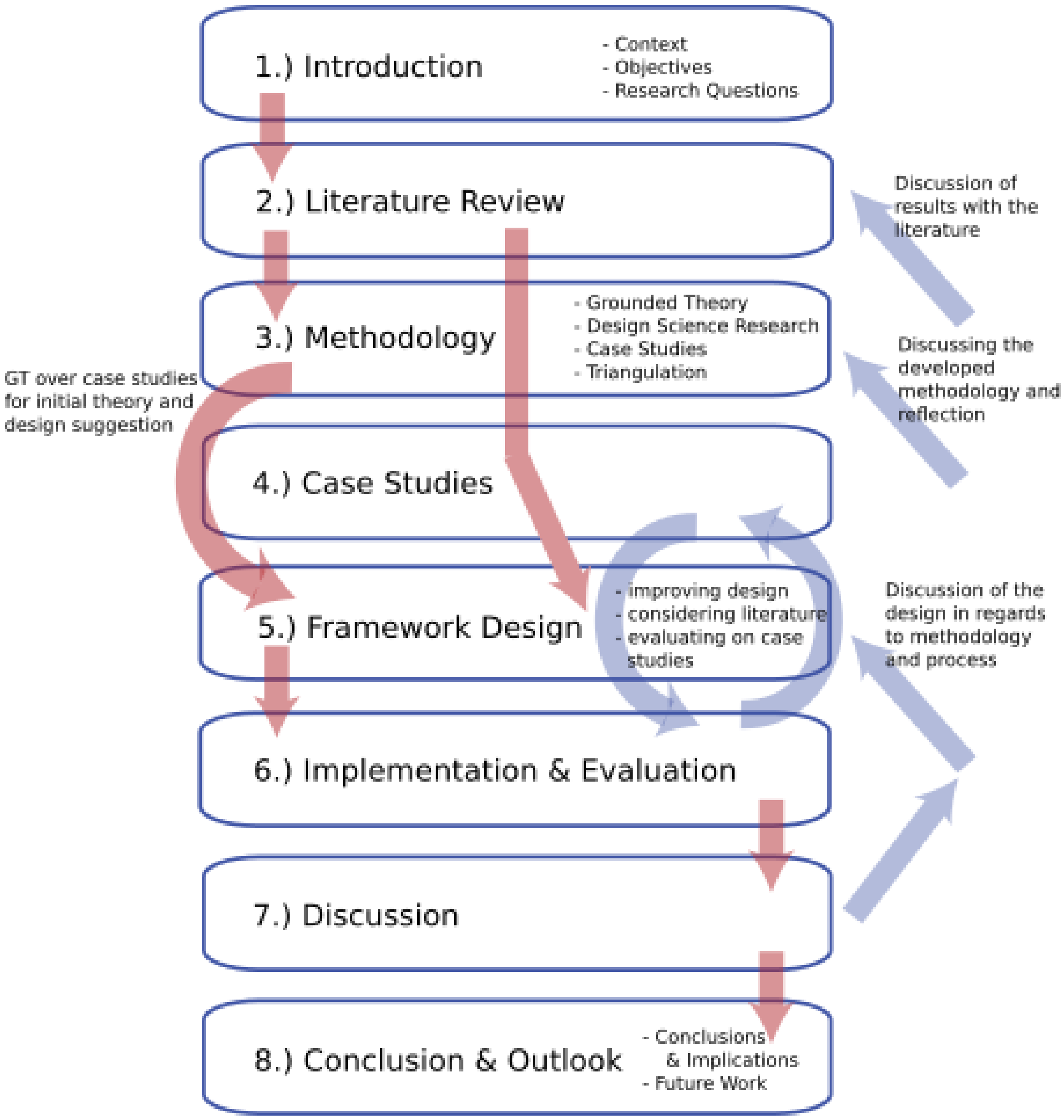


Figure 4: Graphical Thesis Index

challenges and solutions emerge, yet on a meta-level these are not reconciled sufficiently as of now. This is captured in the framework design which has been iteratively improved as the theory emerged from the data.

Chapter 6 - Implementation & Evaluation: This chapter summarises the prototypical implementation of the conceived framework design to evaluate technical applicability in respect to the theory. This pathway was necessary to ensure the iterative feedback in the design science process. Additionally, it provided the author with further valuable 'data' to assess fit, work, relevance, and modifiability of the framework theory.

Chapter 7 - Discussion: In this chapter the theory, framework design and implementation results are discussed and compared to the current state-of-art as presented in the literature review. Furthermore, the author's contributions to the body of knowledge are critically examined as per Grounded Theory as well as limitations of this work.

Chapter 8 - Conclusion: This chapter concludes over the research, summarises its implications for practice and outlines future work.

2 LITERATURE REVIEW

This chapter provides the review of the current and past literature in the immediate fields of interest for this study. In the first section (2.1) the current state of geodata management in New Zealand is reviewed within the context of hydrogeological research. Subsequently, the relevant international literature for the related GIScience disciplines is reviewed. This chapter contributes to the Research Objectives 1 ('Review methods, standards, formats') and 2 ('Identify, select and discuss an appropriate set thereof') and provides the immediate foundations to answer the Research Question 1 ('What are the required interfaces and data formats?').

Furthermore, the sections 'Spatial Data Infrastructures' (2.2), 'Semantic Web and Linked Data Principles' (2.3), 'Sensor Web Enablement and Environmental Sensors Integration' (2.4) are also providing the necessary literature background to investigate into the Research Questions 2 ('Can existing models and software modules be linked to enable interoperable models?') and 3 ('How can such models be provisioned with available interoperable data sets'?). The section 'Geoprocessing in the Web and Linked Models' (2.5) focusses on the specific requirements for groundwater modelling software and interoperable web processing in relation to Research Question 2 and 3.

To answer Research Question 4 ('Can the visualisation of interoperable data and models be realised with the same design principles?') the section '3D Visualisation in the Web' (2.6) reviews the state-of-the-art of web visualisation and visualisation of geological, hydrological and hydrogeological data.

Finally, in each section the current limitations in methods are outlined to inform Research Question 5: ('Overcome challenges and gaps from the literature?'). Thus, through Research Objective 3, theory for designing interoperable data and analytics infrastructures can be advanced and current challenges and short comings in existing methods can be addressed in the course of this study.

2.1 HYDROGEOLOGY AND OPEN GEODATA IN NEW ZEALAND

New Zealand's freshwater resources are extremely valuable for the economic and environmental benefits that they provide. Large economic value comes from water supplies to agriculture, industry and domestic users (White, 2001; White et al., 2001). The many environmental benefits of water resources also include the maintenance of aquatic ecosystems and provision of recreation opportunities (Harding et al., 2004). Currently, both surface water and groundwater resources are under pressure from development, e.g., as demonstrated by an approximate 50% increase in allocation between 1999 and 2006 (Ministry for the Environment, 2006). Increasingly, the effects of land use on water quality are being recognised by communities. Our current understanding of New Zealand's groundwater resources is generally less than required to meet current and future water resources management challenges (White, 2006).

Freshwater management is an important task for all regional and unitary councils in New Zealand and is legislated by the Resource Management Act (Ministry for the Environment New Zealand, 2013). In the last decade regional councils have faced an increasing demand on freshwater resources from agriculture. While most surface waters are already allocated, the actual volumes of available groundwater resources are not well understood (Lowry et al., 2003; White & Reeves, 2002).

Stakeholder information requirements for the characterisation of New Zealand's groundwater resources were assessed by Lowry et al. (2003) and William et al. (2009) and since the publication of 'Groundwaters of New Zealand' (Rosen & White, 2001), much resource characterisation at the subregional scale has been completed by regional councils as the responsible authorities. These assessments have been incorporated into regional policies on groundwater quantity and quality. Reports and research articles are made public in many instances, but the underlying data is mostly archived in internal databases. This essential environmental data remains 'at rest' and is not immediately available for national or cross-regional water resource management. Large scale analyses e.g. from Daughney and Reeves (2006) are unique efforts at a certain point in time with limited decision support value, because they rapidly become outdated.

Research organisations and regional councils, as well as private companies, undertake drilling, pump tests, water sample analysis, flow recording with manual data entry or sensor-based systems. There is a great variety of tools, applications and databases related to the integration of live telemetry data and environmental sensor networks and to manage hydrogeological and environmental data in general. Interoperability issues between commercial software products for data management and analysis are due to a lack of standardised data interchange formats. Thus, data managers evaluate either commercial or free and open source software for geospatial data to publish publicly available parts of their data via OGC web services. Kmoch, Klug, Ritchie, Schmidt, and White (2015, tbl. 2, p. 9) examined available tools and databases for hydrological and hydrogeological data in use in New Zealand (Table 1) and found that increasing support for OGC standards in these tools was becoming apparent.

Population growth and increasing land use intensity has led to growing demands and exploitation of natural resources in New Zealand.

To understand the hydrological state of the environment and groundwater dynamics, data sets and measurements need to be made consistently available and accessible to scientists, planners, and stakeholders to ensure proper decision making also across regional councils.

The Oxford Dictionary states that 'Interoperability' is 'The ability of computer systems or software to exchange and make use of information'[1]. In a geospatial context interoperability includes the use of standards for transfer and exchange of geographical information and the integration of different data types. For the implementation of SDI seven interoperability levels are suggested: technical, syntactic, semantic, pragmatic, dynamic, conceptual and organisational levels (Manso & Wachowicz, 2009; Manso, Wachowicz, & Bernabé, 2009). The formalisation of interoperability models for the implementation of Spatial Data Infrastructures need to specify the applicable levels of data interoperability (Sheth, 1999).

Klug and Kmoch (2014) summarise interoperability as that property of a data set or service, whose technical structure or software interfaces are well defined in order to work with other data sets or services without restrictions on implementation. They reviewed the state-of-the-art of data interoperability from a GIScience point of view and compare it with their observations regarding groundwater data management in New Zealand. They suggest that open and easy access to hydrogeological data and information would be a key for an increasingly timely and efficient assessment and management of groundwater resources for New Zealand. This is following up on the findings of Lowry et al. (2003, p. 579) that *'most gaps appear to be predominantly information and implementation issues.'*

1 http://www.oxforddictionaries.com/definition/english/interoperability

Table 1: Overview of main hydro(geo)logical data management systems in use in New Zealand, adopted from Kmoch et al. (2015)

Tool	Description
Tideda™	DOS/Windows-based database and reporting application for hydrology-related time series data, uses special file formats, prequel to Hilltop, now maintained by NIWA (NIWA, 2013).
Hilltop Data Tamer™	Windows-based database and reporting application suite for hydrology-related time series data, uses special file formats, server system which provides REST-style XML access and can import HydroTel™data. OGC support has been demonstrated with basic WFS and SOS implementation in EODP testbed (Rodgers, 2015).
Hydstra™	Database and reporting application for hydrologic time series data. This Australian-based company was acquired by German-based company Kisters around 2003 that promotes Wiski which provides data migration paths (Kisters, 2015a)
Kisters Wiski™	Fully-fledged data management and reporting system for hydrological data and time series, Kisters is currently working on WaterML 2.0 and took part in the OGC surface water interoperability experiment (Kisters, 2015b, 2015c).
AQUARIUS Time-Series™	AQUARIUS Informatics' software has recently been introduced to New Zealand and provides tools for water data management and supports the latest global standards set by the USGS, ISO, WMO, and OGC (Informatics, 2015).
HydroTel™	Telemetry / sensor system from New Zealand-based iQuest company, which was acquired by Kisters in 2007 (IQUEST, 2011). Kisters and HydroTel interoperability tests are described in IQUEST (2011).
Oracle™and spatial/Locator™extensions	Many regional councils and agencies use an Oracle database with and without its spatial extensions and implement independent data models to store hydrogeological and hydrological data including information related to bores, wells, springs and hydroclimate data, e.g. the GGW including the NGMP, EW bore database (Waikato Regional Council) and NIWA CLIDB.

Increasingly, data is made available and accessible via downloads on websites. However, download possibilities range from Portable Document Format (PDF) reports, time series in spread sheets or comma-separated text files or other proprietary data formats and Application Programming Interfaces (APIs) of (non-) commercial software products. Examples include current web portals such as the GNS GGW and NGMP databases or NIWA's CLIDB. New Zealand Regional Councils increasingly display regional hydro-climate and water quality information for the citizens on their websites through custom data viewers or download sites, again through incongruent custom or proprietary formats. Thus, constraints on the standardised and structured access of hydrogeological and related data and information are a significant impediment to immediate easy reuse for national and cross- regional projects. Constraints include inconsistencies across organisations such as data format and standardised APIs as well as parameter naming and descriptions. Observations at data access workshops which were documented by Kmoch et al. (2015) confirm that there is a need and desire for improved interoperability for environmental data in New Zealand.

The SMART Aquifer Characterisation programme (SAC) suggested a linked geoportal for groundwater data and information, which was intended to interlink its research outputs with the data residing at other data custodians, like regional councils and research institutes, to provide a seamless view over New Zealand's groundwater resources (Klug, Daughney, Verhagen, Westerhoff, & Dudley Ward, 2011). The main challenges in implementing such a geoportal include interdisciplinary data harmonisation and interoperability of services across New Zealand. Klug and Kmoch (2014) and Kmoch et al. (2015) argue that technical interoperability could be improved with a complimentary set of international standards, like Geographic Markup Language (GML), Web Map Service (WMS), Web Feature Service (WFS) and Catalogue Service for the Web (CSW), because multiple initiatives around the globe have also started to employ such standards in attempts to build geoportals and SDI.

- Most gaps appear to be predominantly information and implementation issues.
- Essential environmental data remains 'at rest'. It is either not immediately available or is made accessible via websites through custom interfaces or proprietary formats.
- The main challenges of implementation include interdisciplinary data harmonisation and interoperability of services across New Zealand.

- Interoperability issues between commercial software products for data management and analysis are due to a lack of standardised data interchange formats. However, increasing support for OGC standards in these tools was shown.

2.2 SPATIAL DATA INFRASTRUCTURES

Data must be collected, transmitted, stored, error checked, manipulated, retrieved for analysis, and shared within the hydrological community under commonly accepted rules and standards in order to plan and implement integrated water resources management practices (Carleton et al., 2005; Horsburgh et al., 2009; Iwanaga, El Sawah, & Jakeman, n.d.; Kao et al., 2011; Ranatunga et al., 2011). 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 to facilitate a rational exploitation and allocation of the available resources (Kiehle, 2006; White, 2006; Yang et al., 2010).

At the end of the last century Albrecht (1999) discussed 'offline' geospatial information standards. Since then these standards transcended to 'online' web service technologies with an increasing amount of available, web and cloud based, geospatial resources and modelling functions (Bailey & Chen, 2011; Ballagh et al., 2011; Zhao, Foerster, & Yue, 2012). While offline geospatial content still has its value, the diffusion of this information available as hard copy maps, digital images or PDF files is limited. The internet, as a fast, efficient, and effective information distribution medium, offers sufficient capabilities to provide continuously updated and 'live' information. Consequently, information and data infrastructures were developed and the notion of SDI became prevalent for environmental and spatial data and information systems.

Thus the need for information systems arose and techniques like Enterprise Architecture (EA) formalised the design of information systems, in particular within companies (Zachman, 1987). In EA the architect creates a descriptive framework, and then specifies the information systems architecture based upon that neutral, objective framework. Booch (2010) explains Technical Architecture (TA) and its relation to EA. TA focuses on the architecture of software-intensive systems that are used by a business to achieve its mission. EA in comparison focuses on the architecture of a business that uses software-intensive systems (Booch, 2010).

Zachman (2015) also argues that EA is not about building IT models, but about solving the general problems of the integration of all of the components of the system. However, EA is not so much a strategic planning methodology, but a way to build the enterprise information systems iteratively and incrementally. Eventually, he agrees that there is a perception that Enterprise Architecture is too costly and time consuming.

Hinkelmann et al. (2016) present a new 'metamodelling' paradigm for next generation enterprise information systems, which moves beyond the classical EA model-driven engineering approach and towards a more continuous alignment of business and IT. They explain that agility is an essential quality for the engineering of enterprises and enterprise information systems, and only through changing classical EA methods businesses can react to continuous and unexpected change in a rapidly evolving environment.

To fulfil the requirements of political initiatives such as the New Zealand Resource Management Act (RMA) from 1991 (Ministry for the Environment New Zealand, 2013), or the state of the environment reporting in general (Ministry for the Environment New Zealand, 2013) 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 data sets and the distribution of data sets across the regions within New Zealand. In New Zealand, data sets are maintained by different custodians, e.g. Crown Research Institutes (CRIs), regional councils, and Ministries. These institutions hold spatial data and metadata (data about data (ISO 19115, 2003; ISO 19139, 2007)) in various formats that use different nomenclature, storage technologies, interfaces and even languages. This situation complicates search, discovery, and accessibility for users. The search, retrieval, and pre-processing of data sets consume valuable time and personnel resources on both the data provider and data user side (Ames et al., 2012; Beran & Piasecki, 2009). Bandaragoda et al. (2006, fig. 1) identified that 'inconsistent data formats and existence and consistency of metadata are the biggest concerns'. Furthermore, she explains that hydrologists use from at least 10% to more than 50% of their time for data preparation. This has resulted in the need for new data management techniques, tools, and data models to effectively manage the vast amount of new hydrological data being collected (Goodall, Horsburgh, Whiteaker, Maidment, & Zaslavsky, 2008; Wojda & Brouyère, 2013).

Resource management decisions are based on environmental data with spatial and temporal properties. Klug and Kmoch (2015) conclude that, the larger the number of data sets and the less structured the data sets are, the worse the situation will become. Since we are not only looking at single data users but also considering multi-vendor architectures and multi-user applications, a considerable loss of economic and production power due to inefficiency and ineffectiveness would accrue. Thus, networked, web-based Geographic Information System (GIS) provide means to process and analyse spatio-temporal data from distributed sources and derive valuable information to inform policy development (Crompvoets, Vancauwenberghe, Bouckaert, & Vandenbroucke, 2011; Latre et al., 2013). At the same time, information retrieval has become faster, but data sets remain scattered both in location and formats.

Land Information New Zealand has published the New Zealand Geospatial Office (NZGO) in 2007 (Land Information New Zealand, 2007). Also, New Zealand recommends and relies on OGC standards for the foundation of an NZ SDI. Since then governmental agencies and research institutes are starting to develop strategies to make data accessible via OGC standards. The OGC was founded in 1994 and 'is an international industry consortium of more than 500 companies, government agencies and universities participating in a consensus process to develop publicly available interface standards' (OGC, 2015).

In OGC terms are there two main types of implementation standards, i.e. service interfaces and data encodings, which are backed by conceptual and logical models. Besides the mentioned service interfaces WMS (OGC & Beaujardiere, 2006), WFS (OGC, Peter, & Vretanos, 2010; OGC & Vretanos, 2005b) and CSW (OGC, Nebert, Whiteside, Vretanos, & Editors, 2007), following service interface are also important standard components within an SDI: The Web Coverage Service (WCS) (OGC & Baumann, 2006, 2010) and the Sensor Observation Service (OGC, Bröring, Stasch, & Echterhoff, 2012), can deliver multi-dimensional continuous or discrete spatial and spatio-temporal data sets, e.g. from sensor networks (i.e. SOS server) or satellite remote sensing. Secondly the data encoding standards provide the conceptual data model as well as a corresponding standard encoding description for the data, where GML (OGC, 2007) provides the core elements. GML is an XML schema that defines elements to describe geographic entities and their properties. Therefore, encodings, i.e. GML application schemas, which inherit from GML, are inherently interoperable.

There are many SDI and geoportal initiatives and research projects that describe in the literature how they implemented and advanced OGC standards. The OGC undertook several so called Interoperability Experiments to drive the maturing of the developed standards, amongst them the first groundwater interoperability experiment (GWIE1). Boisvert and Brodaric (2012); Brodaric, Booth, Boisvert, and Lucido (2015) illustrate the Groundwater Information Network (GIN) for Canada which, through mediation services, allows for interoperable hydrogeological data exchange between Canada and the US. GIN also supports WFS and SOS as OGC services and Water Markup Language (WaterML) 2.0 for hydrological time series as well as the first version of the also GroundWaterML, Groundwater Markup Language (GWML) to describe groundwater features like wells, springs or aquifers. GWML1.1 is currently being reviewed and advanced towards an OGC standard in the second groundwater interoperability experiment (GW2IE) (Brodaric et al., 2016).

For the hydrological domain WaterML 2.0 (OGC & Taylor, 2012) was published 2012 as a community effort of the OGC Hydrology Domain working group (HydroDWG), which is a joint collaboration with the WMO commission on Hydrology (WMO CHy). WaterML 2.0 is a specialisation of Observations & Measurements (ISO 19156, 2011; OGC & Cox, 2011), which is a fundamental GML application schema used to describe and deliver observations, measurements or specimen of all kinds as well as telemetered data and time series in a standardised manner. WaterML 2.0 and Observations & Measurements (O&M) encoded hydrological data are typically served through OGC SOS, however, (Bermudez et al., 2009) discussed serving observations also via WFS. As groundwater is inherently dependent on the containing geological bodies, feature models for geological descriptions have been developed. Bore logs or geological specimen, sample description and observations are typically served via WFS (GeoSciML, 2011; Sen & Duffy, 2005). Besides the WaterML 2.0 conceptual model for hydrological data, other so called feature models for groundwater have been developed Boisvert and Brodaric (2012).

Dahlhaus, MacLeod, and Thompson (2011) presented UBSpatial, an interoperable federated groundwater database that employs GWML2 and further OGC standards and provides a web mapping data portal for the visualisation of Victoria's groundwater resources in Australia. GWML2 is also based on GML. GWML2 links elements from the Geoscientific Markup Language (GeoSciML). GeoSciML is an OGC GML complex application schema, too, and is mainly used to describe geology, e.g. bore logs, rock samples, geological units and formations.GWML2 as a GML application schema extends GeoSciML and provides a structural framework for the interoperable exchange of individually defined groundwater related data. The purpose of GWML2 is to provide a generic, i.e. top-level classification, or the conceptual schema for groundwater data. GWML2 is also connected to O&M and Sampling Features specifications (OGC & Cox, 2007). It has been designed according to (ISO 19103, 2005; ISO 19118, 2011; ISO 19136, 2007) and follows the practice of GeoSciML and GML.

The OneGeology project is an international initiative which aims to provide a geological map of the planet (Sen & Duffy, 2005). It sources the geological maps in the GeoSciML format from each participating country via OGC WFS and WMS. 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 geoscientific 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. Web Map Context (WMC) (OGC & Vretanos, 2005a) documents were used to keep track of the state of the currently active web map with its active layers, zoom, bounding box and styling properties, as well as to deliver predefined collated maps. WMC documents provide a structured and standards compliant way to define a map context holding references and styling declarations for spatial data sources. However, its main limitation of being designed only to collate WMS layers have been superseded by the OpenGIS®Web Services Context (OWC) specification, which now caters for more generic OGC services and data linkage (OGC, Brackin, & Gonçalves, 2014b). An Atom-based XML (OGC, Brackin, & Gonçalves, 2014a) encoding is defined to provide full contextual representation capabilities for situational awareness, catalogue search collections or analytical processing results. A JavaScript Object Notation (JSON) encoding based on Geographic JSON (GeoJSON) has also been suggested, which is currently being developed in an interoperability test-bed.

Wojda and Brouyère (2013) proposed object-oriented hydrogeological data model, named Hg2o, for groundwater projects, which contributes to the Infrastructure for Spatial Information in Europe (INSPIRE) (EU Directive, 2007) process in Europe and follows recommendations of ISO/TC211, OGC and the European Geospatial Information Working Group. Among other large-scale

EU-wide implementations building the European SDI, like NatureSDIplus for nature conservation or GS Soil for soil data across the EU (Klug & Bretz, 2012), the GIGAS project issued a series of technical reports on interoperability, relying on OGC standards (Portele et al., 2009).

The Australian Bureau of Meteorology (BOM) developed the Water Data Transfer Format (WDTF), which is used for the transfer of hydrological and meteorological observations between the federal states/territories of Australia. BOM also applied it to groundwater data (Kao et al., 2011; Ranatunga et al., 2011). WDTF is also an XML data format and experiences of WDTF development have contributed to maturing WaterML 2.0.

The CUAHSI, a research organization representing more than 100 U.S. universities and international water science-related organizations, and the United States Geological Survey (USGS) started to develop a water information service network for the US. Those water information services and technologies were reviewed through further international collaboration, which advanced components like WaterML towards OGC interoperability. Besides the proprietary binary format WaterOneFlow, CUAHSI developed the first version of WaterML and also supported the development of WaterML 2.0. The integrated 'CUAHSI Hydrological Information System' (CUAHSI-HIS) comprises a hydrological SDI and a desktop-based processing and visualisation tool. With a focus on hydrological time series and an integrated desktop software package, however, CUAHSI does not allow a holistic visualisation of aquifers and groundwater properties (Horsburgh et al., 2014; Peckham & Goodall, 2013).

Finally, data sources including their respective metadata sets should be discoverable (e.g., by a web-based keyword search). OGC CSW provides these capabilities. de Andrade, Baptista, and Leite Jr (2011) and Yue et al. (2011) describe how a federation of catalogues through the CSW service interface improves overall access to distributed metadata records. Standards, tools, and interfaces are being developed by the OGC for data transfer, processing and visualisation and comprise crucial building blocks in a distributed Spatial Data Infrastructure (Latre et al., 2013; Maguire & Longley, 2005; Yang et al., 2010).

Common data exchange formats and service interfaces are required to achieve technical and syntactical interoperability and to transfer hydrogeological feature and observation data between users. GML provides the foundation for application schemas for feature encoding, e.g., GeoSciML and GWML, and WaterML 2.0 to exchange hydrological and geological feature and observation data. Initiatives on national and international scale build on such standards include AuScope (Australia), the Groundwater Information Network (Canada) and CUAHSI (North America), as well as the activities in the INSPIRE. These standards have gained attention in the literature and WaterML 2.0 has also been adopted by the World Meteorological Organisation (Atkinson, Dornblut, & Smith, 2012). Klug and Kmoch (2014, pp. 3.2) review these developments and standards and argue how they contribute to and correlate with the efforts of the NZGO to establish a nation-wide SDI including approved international (ISO) and Australian and New Zealand (AS/NZS) standards.

- Networked, web-based GIS provide means to process and analyse spatio-temporal data from distributed sources, and to provide continuously updated and 'live' information.
- Inconsistent data formats and existence and consistency of metadata are the biggest concerns, but OGC standards were described and adopted as solution in many studies.
- EA formalised the design of information systems, in particular within companies. But there is a perception that EA is too costly and time consuming.
- There is need for new 'metamodelling' paradigms for next generation (inter-organisational, cross-sectoral) enterprise information systems.

2.3 SEMANTIC WEB AND LINKED DATA PRINCIPLES

For the meaningful integration of geographic data sets GIScience and the Geosciences research topics have been continuously focussing on semantic methodologies using ontologies and their machine-readable encoding (Cruz, Monteiro, & Santos, 2012; Evangelidis, Ntouros, Makridis, & Papatheodorou, 2014; Gahegan, Luo, Weaver, Pike, & Banchuen, 2009; Lutz et al., 2009; Stock et al., 2012). End users can search for (hydrological) information using keywords, areas, or points of interest. Those frameworks are not discrete components by themselves, but techniques and methodologies to integrate generic resources in a (web-based) distributed environment. To yield the requested 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. Thus, a thesaurus is a collection of concepts - terms of reference in a particular community or domain with, collated and described with their attributes and properties and inherent relationships. It provides a uniform and consistent vocabulary for indexing metadata (Ma, Carranza, Wu, van der Meer, & Liu, 2011). A gazetteer is a dictionary or directory referencing place names with their geographical locations. In current GIScience and semantic literature, thesauri and gazetteer services are discussed as implementations of special profiles of so called vocabulary or ontology services accessible through Hyper Text Transfer Protocol (HTTP) protocols WFS and SPARQL Protocol And RDF Query Language (SPARQL) (S. J. D. Cox, Simons, & Yu, 2014; Perry & Herring, 2012; West, 2015).

The semantic heterogeneity of information and naming of phenomena, natural spatial entities and scientific parameters imposes a general challenge to relate similar meanings of information in data sets and services. Research in semantics and linked data is investigating in how to overcome such challenges with the use of ontologies. In that aspect, instantiations of thesauri and gazetteers can be based on similar description systems, e.g. Resource Description Framework (RDF), Simple Knowledge Organization System (SKOS), or Web Ontology Language (OWL), and can be made accessible through web services APIs like the SPARQL or Geographic SPARQL (GEOSPARQL) (with spatial indexing and search capabilities) (Antoine Isaac & Ed Summers, 2008; Baker et al., 2013; Batcheller & Reitsma, 2010; Choor alil & Gopi nathan, 2015; S. J. D. Cox et al., 2014; Glimm, Horridge, Parsia, & Patel-Schneider, n.d.; Horrocks, Patel-Schneider, & van Harmelen, 2003; Ma, Carranza, Wu, & van der Meer, 2012; Perry & Herring, 2012; Vitolo, Elkhatib, Reusser, Macleod, & Buytaert, 2015).

RDF from the world Wide Web Consortium (W3C) provides the foundation for publishing and linking metadata through an RDF Schema (RDFS). RDFS is supported by the SKOS, which is a common data model for sharing and linking knowledge via the web. Thesauri, 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, Nogueras-Iso, López-Pellicer, Muro-Medrano, & Zarazaga-Soria, 2007).

OWL belongs to the family of knowledge representation and has its place in the very young ontology languages (Batcheller & Reitsma, 2010; Buccella et al., 2011; Horrocks et al., 2003; Pulido et al., 2006; Stock et al., 2012). The Rule Interchange Format (RIF) is part of the infrastructure for semantic web enablement (Ramos, 2015) and facilitates the exchange of diverse existing rules(Devaraju, Kuhn, & Renschler, 2015; Zhao et al., 2009).

The RDF/SKOS and RDF/OWL ontologies can be exposed and queried through a web-based query language - the SPARQL query language. SPARQL works over web services and supports HTTP and various Linked Data encodings like JSON Linked Data (JSON-LD) and triple notations to better integrate spatial and semantic data in the web.

Linked Data principles build on top of the semantic web. Identifiers of concepts in ontologies are in Uniform Resource Identifier (URI) form, which could even be resolved like web links. S. J. D. Cox et al. (2014) is against that notion, and argues the Domain Name System (DNS) system is not robust enough for permanent identifiers on domains. However, the host part, the path and parameter elements of the URI can precisely define and delineate a resource, its origin and identity properties. This relates to the Representational State Transfer (REST) paradigm, where web resources should also be navigated with the correct HTTP verb, host, path and parameters (Fielding, 2000). Thus, observed properties and feature identifiers could be used in combinations with Linked Data vocabularies to add semantic glue for groundwater data interoperability (Car, Cox, Stenson, Atkinson, & Fitch, 2015; de Andrade et al., 2011; Janowicz, Scheider, Pehle, & Hart, 2012; Tian & Huang, 2012).

- The semantic heterogeneity of information and naming of phenomena and entities causes a challenge to relate similar meanings of information in data sets and services.
- The semantic web is not an assembly of discrete components, but techniques and methodologies to identify and describe generic resources in a (web-based) distributed environment. Identifiers of concepts in ontologies can thus be interlinked and related via their machine-readable encoding.

2.4 SENSOR WEB ENABLEMENT AND ENVIRONMENTAL SENSORS INTEGRATION

Our natural environment is multi-dimensional, multi-functional and highly dynamic (de Groot, 2006; Grabaum & Meyer, 1998). Its inherently complex and transdisciplinary nature demands higher-order systems thinking (Tress, Tress, van Apeldoorn, & Fry, 2003). The interwoven nature of real-world problems requires holistic integrated approaches to apply state-of-the-art knowledge to explain, explore, and predict environmental phenomena to ensure proper resource management strategies (Klug, 2012; Potschin & Klug, 2010). Most importantly, geographically referenced environmental information is required in near real time to ensure early warning and a timely adaptation. Example stressors are climatological extreme events or overexploitation of water resources.

To acquire near-real time environmental data, wireless sensor networks have been used since decades and have been regularly reviewed and continuously improved (Akyildiz, Su, Sankarasubramaniam, & Cayirci, 2002; Baronti et al., 2007; Schimak, Rizzoli, & Watson, 2010; Yick, Mukherjee, & Ghosal, 2008). However, such technical sensor network developments did not consider the standardised data distribution beyond their local technical use cases.

In GIScience the 'feature' term is used for the abstraction of a real world phenomenon into a 'thing' for the model depiction of the real world phenomenon (ISO 19136, 2007). An (OGC/ISO defined) observation is an abstract feature entity that consists of the procedure (the measurement process, sensor, or algorithm) that generated an observation result, its phenomenon time (when did the phenomenon occur that was measured), the result time (when was the result of the measurement available), the observed property (what was being observed, which phenomenon) the feature of interest (basically the location feature or entity at which the observation was conducted) and the actual result (ISO 19156, 2011). Furthermore, metadata elements like data quality and additional dependent parameters can be added too.

For in situ observations and measurements (e.g. water level measurements), another service component is the sensor interface in the SOS. This interface delivers descriptive sensor information in a standardised format using the Sensor Model Language (SensorML) (S. J. D. Cox & Botts, 2007), and O&M, and gathers/receives measurements from hydro-climate field sensors (Bröring et al., 2011). Currently SensorML 1.0.1 is the standard format for sensor and process metadata for the SOS interface, but version 2.0 has been recently published, too. However, it has not been prescribed in current interface specifications nor been demonstrated in implementations of SOS. However, SensorML has been proposed as a possible alternative data encodings for Sensor Planning Service (SPS) and WPS payloads (N. Chen, Hu, Chen, Wang, & Gong, 2012; N. Chen, Wang, & Yang, 2013).

There is a discussion in the field how the `FeatureOfInterest` of an Observation should be handled. In SOS this is usually encoded as a `SpatialSamplingFeature` which therefore really only binds the sampling location. However, it can be a real world feature, like a well or weather station or an abstract feature that just provides a handle of a place based on coordinates. Additionally, separation is managed with an additional `SampledFeature` element, which could refer to the lake, river or aquifer, even the atmosphere, as a greater real world feature that was sampled at one particular location (Botts, Percivall, Reed, & Davidson, 2008; Leighton et al., 2015; Liang, Croitoru, & Tao, 2005; Ritchie, Hodges, et al., 2014).

In comparison to the WFS interface, the SOS specification directly describes the necessary semantics to access and deliver time series (Bermudez et al., 2009). SOS is part of the OGC Sensor Web Enablement (SWE) (Bröring et al., 2011; OGC & Robin, 2011) initiative and allows the aggregation of sensor descriptions and observation time series. The OGC O&M standard describes how time series and observation data can be encoded within OGC compliant interoperable systems. WaterML 2.0 as a specialisation of O&M is already widely used to exchange hydrological time series data via SOS. This has resulted in a SOS configuration profile for hydrology, which basically constrains certain characteristics and behaviour to align better with the needs of users of hydrological data (Leighton et al., 2015). Jirka, Bröring, Kjeld, Maidens, and Wytzisk (2012) also propose a lightweight station profile for SOS-based observation data delivery for fixed sensor stations. Sensor descriptions, configurations, and capabilities can be expressed in SensorML. Main interfaces for data transmission and interaction with sensors and actors are the SOS and the SPS (OGC, Bröring, Stasch, & Echterhoff, 2011).

However, OGC SWE semantics heavily rely on XML-based web services, which induce data volume overhead through the XML markup and subsequently impacts data transfer on low bandwidth unreliable network connections. Alternative sensor uplink methods are presented as 'Intelligent, Manageable, Power Efficient and Reliable Internetworking Architecture (IMPERIA)', a centrally controlled architecture that vertically integrates a WSN network stack with the publish/subscribe messaging middleware MQTT-S (Hunkeler, Lombriser, Truong, & Weiss, 2013). MQTT is assessed as suitable bi-directional transport protocol for wireless sensor networks following the message queue pattern (J. Chen et al., 2013). This has been picked up and an MQTT integration with OGC SWE semantics has been proposed and conceptualised by (Ghobakhlou, Kmoch, & Sallis, 2013). Klug, Kmoch, and Reichel (2015) continued research and presented a prototype system, which sticks to the respective operation names for SOS and SPS. The payload data encodings are implemented as light-weight JavaScript Object Notation aligning with the conceptual data model of O&M and a parameter list for SPS Tasking. Another approach has been published through the OGC 'SensorThings API Part 1: Sensing' (OGC, Liang, Huang, Khalafbeigi, & Kim, 2015). While still relating to the ISO observation concept, the SensorThings API standard changes for example the O&M 'feature' pattern to accommodate for generic sensors. This is in favour of a data stream centric and industry-focussed conceptual model and weakens the geographic aspect of the observation data model. Furthermore, 'Part 1' only describes data collection and distribution, a tasking profile is planned as a future work activity and is announced to be defined in a separate document as 'Part 2' of the SensorThings API.

While these are comparatively recent activities which have not found a broader echo in the literature, it demonstrates that MQTT is a preferred medium for low-level sensor connectivity in combination with OGC/ISO observation semantics, and another indicator for coalescence of sensor web and classic SDI paradigms (Jazayeri, Liang, & Huang, 2015).

- Wireless sensor networks have been used since decades but did not consider the standardised data distribution beyond their local technical use cases.
- SOS specification directly describes the necessary semantics to access and deliver time series and SPS for tasking/configuring sensors.
- OGC SWE semantics heavily rely on XML-based web services, which induces data volume overhead and negatively impacts data transfer on low bandwidth unreliable network connections.
- MQTT is a preferred medium for low-level sensor connectivity but has no default alignment with OGC SOS or SPS semantics.

2.5 GEOPROCESSING IN THE WEB AND LINKED MODELS

From the general goal of interoperable data discovery and sharing, SDIs and geoportals evolved to integrated systems of systems, not only providing data, but also processing routines and visualisations of the processed geospatial data to support science and education as well as policy and decision making for particular environmental domains.

GIS-based geoprocessing for Earth Sciences is typically applied in surface hydrology, soils and basic geology as specialisations of physical geography. Modelling approaches need a thorough data basis as a foundation, but spatial data for integrated environmental modelling is scattered and difficult to obtain, publicly unavailable and data access is hampered. In consequence a lot of time resources and costs need to be invested for state-of-the-art data acquisition. Subsequently, updates of previous indicator modelling results are almost unavailable on a real-time basis since real-time data are rarely accessible. Furthermore, software routines for environmental modelling often are neither publicly accessible nor do they support standardised data interfaces for their integration. Both modelling and monitoring efforts are considered as the key for sustainable environmental planning (Jorgensen, Refsgaard, & Hojberg, 2007) resulting in environmental decision support systems rapidly progressing since the beginning of this century (Matthies, Giupponi, & Ostendorf, 2007). Interdisciplinary and multi-purpose integrated models are becoming more important but also more complex (Pahl-Wostl, 2007; Voinov & Bousquet, 2010). Multi-purpose examples are 'Coupling a hydrological water quality model and an economic optimisation model to set up a cost-effective emission reduction countermeasure scenario for nitrogen' (Cools et al., 2011) or 'An integrated approach to linking economic valuation and catchment modelling' (Kragt, Newham, Bennett, & Jakeman, 2011). However, both examples did not consider real time measurements for modelling, but Hart and Martinez (2006) highlight the importance of real-time environmental information from sensor networks for better process-understanding and informed decision making. Thus, linking integrated environmental modelling with real-time environmental sensor data has been proposed as a necessary roadmap for the future (Granell et al., 2013; Laniak et al., 2013).

To analyse and describe conditions such as extreme precipitation and overuse of groundwater resources e.g. for irrigation purposes, many ecological modelling routines resulting in indicator proposals have been developed. Among modelling routines described in the literature are many hydrological modelling examples dedicated to irrigation needs (Blaney & Criddle, 1950; Minacapilli, Iovino, & D'Urso, 2008; Wriedt, Van der Velde, Aloe, & Bouraoui, 2009) or minimum flow in rivers (Gao, Gao, Zhao, & Hörmann, 2010; Thomas, Steidl, Dietrich, & Lischeid, 2011).

Once diverse data sets for hydro-climatological time series, soil coverage, land use, elevation and geology data are available through the SDI, data processing can consequently be migrated into and automated within the web. Thus, connecting SDI data provisioning and web processing services to solve information needs allows non-GIS-experts the use of such tools with recent data (Heavner, Fatland, Hood, & Connor, 2011; Peckham & Goodall, 2013; Smith & Lakshmanan, 2011).

As groundwater and surface waters are often intrinsically connected, the modelling of groundwater-surface water interaction is challenging. However available models could be linked, which means, while running in geographically related areas and sharing boundary conditions, water exchange (e.g. seepage, recharge, discharge, flow) between a groundwater system and a surface water system can be calculated at the same time steps within the simulation, or conceptually for time periods. Groundwater and groundwater-surface water interaction models and numerical simulations are available, documented and reviewed for New Zealand. However, these are implemented in either proprietary and/or desktop-based software codes, like FEFLOW or ArcGIS. Free and open source tools for such tasks are fit for purpose and increasingly applied, e.g. GRASS-GIS (D. Chen, Shams, Carmona-Moreno, & Leone, 2010; Jasiewicz & Metz, 2011; Li, Di, Han, Zhao, & Dadi, 2010).

The finite differences flow model MODFLOW from USGS is considered a defacto standard in expert groundwater modelling. MODFLOW is open source and can be integrated into general purpose programming environment with a Python library (Harbaugh, 2005; Hughes et al., 2015). However, MODFLOW has rarely been described in web-enabled scenarios due to the very specific model calibration for the catchments/groundwater sheds of interest (Horak, Orlik, & Stromsky, 2007; Jones, Jones, Greer, & Nelson, 2015; Matthies et al., 2007).

Usländer, Jacques, Simonis, and Watson (2010) and Díaz, Bröring, McInerney, Libertá, and Foerster (2013) provide insight in designing environmental software applications based upon an open sensor service architecture. However, without proper measures for validation, model (auto-)calibration is hardly achieved (Green & van Griensven, 2008; Raj Shrestha & Rode, 2008).

Streamlining inputs and outputs of data from and to models and the implications of complexity and uncertainty for integrated modelling and impact assessments are a big scientific and technical challenge and need to be considered as equally important as uncertainty propagation (Krysanova, Hattermann, & Wechsung, 2007; Refsgaard, van der Sluijs, Højberg, & Vanrolleghem, 2007). Thus, with increasing complexity the uncertainty of model results increases and there is a strong need for emulation techniques for the reduction and sensitivity analysis of complex environmental models (Andrews, Croke, & Jakeman, 2011; Crout, Tarsitano, & Wood, 2009; Makler-Pick, Gal, Gorfine, Hipsey, & Carmel, 2011; Ratto, Castelletti, & Pagano, 2012).

Klug and Kmoch (2015) highlight that repeatability and transferability of the discussed approaches are hampered, because modelling tools and software often designed to work stand-alone and not within a 'network of models and data'. Thus, automated interchanges and digital representations of model results are presently specialised solutions. As a consequence, the lack of near real-time data access in a ready to use data format and immediate usability in an interoperable workflow for scenarios such as aquifer depletion, well contamination through seepage in contribution zone or flooding (e.g. as in liquefaction events) are rather post-processed than forecast. This is severely impacting resource managers' agility and reaction time and putting people, infrastructure and the environment at risk. Consequently, a workflow from initial data capturing to the real-time provision of conditioned information delivery to end users for prepared decision making needs to be automated and operationalized.

Spatial interpolation of environmental properties and the assignment and communication of related uncertainty is of particular interest in hydrogeology, as property distributions underground can only be validated at certain locations, e.g. wells, bores or springs. A special case in using MODFLOW for predictive uncertainty analysis implements a 'data worth' algorithm for improved decision support on data acquisition priorities (Wallis et al., 2014). UncertWeb and Intamap are two European projects that employ OGC web services and formats to get data, run models and return interpolated data and calculated uncertainty in a web-based fashion (Bastin et al., 2013; Pebesma et al., 2011). Arheimer, Wallman, Donnelly, Nystrom, and Pers (2011) present E-HypeWeb, an interdisciplinary and participatory web-based modelling framework, with a focus of integrated modelling and web presentation of the results and predictions. The demonstrated integrated modelling approaches however have been used for analysis, interpolation and prediction and visualisation of air quality, precipitation, surface hydrology and nitrate transport in European countries (Lindström, Pers, Rosberg, Strömqvist, & Arheimer, 2010). Giuliani, Ray, and Lehmann (2011) examine how a 'grid-enabled' SDI could support running distributed models for environmental sciences with online data and concludes that a standardised linkage of models and data through a networked grid infrastructure is challenging.

Running disparate model software modules in parallel does not provide any means of control, when to exchange data. The OpenMI specification (The OpenMI Association Technical Committee, Bröring, Stasch, & Echterhoff, 2010) is an interface specification which has been developed within a pan-European initiative to link models/simulations that have spatial and temporal overlap.

OpenMI has also been advanced into an OGC standard and can be considered for model interaction/linking (Knapen et al., 2013).

But an important consideration is that the connection of the models to the SDI should be based on agreed data formats and protocols. The prior or observation data for processing, but also the results to be returned shall be in such formats again, to related raw and processed data with same technologies within the SDI. This has not been described nor intended with OpenMI. CUAHSI-HIS uses its stand-alone desktop software as a control components to download data from US water services and run OpenMI-enabled models locally on the computer (Castronova, Goodall, & Ercan, 2013).

The OMS as another generic environmental modelling framework. It has been selected in a collaboration of three New Zealand CRIs (NIWA, Landcare Research, AgResearch) for interoperable freshwater modelling purposes over other integrated methods (Elliott, Ritchie, & Snow, 2012; Rutledge, Elliott, Snow, Ritchie, & Turek, 2012). Similar to OpenMI, OMS is a development and runtime environment to run models from reusable components. OMS presents interfaces which need to be implemented by the model software codes that need to be linked. In comparison with OpenMI, it presents more flexibility in programming languages and execution control. While standardised and well described interface descriptions are used as abstraction to hide complexity of the underlying modelling codes software linkage like OpenMI and OMS are also called 'tightly' coupled.

The OGC WPS (OGC, Mueller, & Pross, 2015; OGC & Schut, 2007) is a service interface description that allows distributed and chained processing in the web, computing power can be extended into the cloud, i.e. multiple instances run in parallel on different servers in the internet, and can be managed synchronously or asynchronously. This concept has been also addressed as SOA (Goodall, Robinson, & Castronova, 2011; Granell, Díaz, & Gould, 2010). In a SOA available services, which could be data provider or processing algorithms or something else are so called 'loosely coupled' services. That means by using Simple Object Access Protocol (SOAP) and its Web Service Description Language (WSDL) the service descriptions/capabilities are available at runtime and can be evaluated on demand (de Jesus, Walker, Grant, & Groom, 2012). The systems do not need to know all members initially.

Main characteristics of SOAs are the encapsulation of smaller applications. Data providers and/or consumers use these applications as reusable services. The orchestration of these services to larger application workflows serves the overall use-case for searching, discovering, accessing and processing of information.

By relying on a predefined set of standards and service interface descriptions, the system becomes flexible to add new services to the orchestration. This concept comes from the computer science field of information systems, is widely used in business applications, and has been successfully applied in GIScience studies, too (Erickson & Siau, 2008; Meek, Jackson, & Leibovici, 2016).

Process chaining is described within the WPS standard (OGC & Schut, 2007) that recommends that it should be accomplished in one of three ways:

1. Creating a process within a WPS that calls other WPS processes in a sequence.
2. Cascading service calls within an execute request.
3. A Business Process Execution Language (BPEL) engine.

Many achievements in web-based geoprocessing focus on logically chaining Open Geospatial Consortium (OGC) web services, via BPEL and analogously Business Process Markup Notation (BPMN) systems that are based on SOAP and WSDL technologies (N. Chen et al., 2012). BPEL has become a popular choice for orchestrating and executing workflows in the web environment. For example, BPELPower was designed as a generic BPEL execution engine with enhancements for executing geospatial workflows (Meek et al., 2016; Yu et al., 2012).

The OGC also discussed and proposed REST as an alternative approach for web services architecture for the geospatial community (C. Reed, 2009; Fielding, 2000). Several OGC prototype implementations demonstrate REST interfaces, however it is not yet standardised.

Alternative software architecture pattern have been successfully applied in distributed environmental geoprocessing, e.g. via exposing hydrological models as WPS and then coupling such models via OpenMI (Castronova, Goodall, & Elag, 2013). A WPS mediation approach was proposed to dispatch models as services on different computing back-ends (Giuliani, Nativi, Lehmann, & Ray, 2012). These architectural patterns represent a cascading rather then chaining approach of orchestrating geospatial workflows.

Many scientific models need to run repeatedly to calculate solutions for different scenarios. Software implementations with Message Passing Interface (MPI) cluster spreading have been demonstrated and are effective on specialised High Performance Computing (HPC) clusters (Dongarra, 1994). However, the MPI cluster approach does not consider the SDI paradigm or flexible internet HTTP-data layer linkage. The recent trend of 'Big Data' and 'Cloud Computing' has also been addressed in the literature, and OGC web services for SDI and WPS for processing and modelling in the web is a recurring topic across scientific disciplines to tackle the large amounts and variety of data and to unify currently still very disconnected environmental modelling routines (Evangelidis et al., 2014; Granell et al., 2016; Vitolo et al., 2015).

The seemingly independently designed process model with its limited input and output definitions in the WPS 1.0.0 specification has led to various approaches to improve semantic annotation of the parameters fields (Henson, Pschorr, Sheth, & Thirunarayan, 2009; Janowicz et al., 2012). Such approaches included development of ontologies, semantic web services and semantic mediation layers, in order to be able to express additional domain meaning for input and output data type definition (beyond the technical data format) that could be evaluated by computer algorithms (Farnaghi & Mansourian, 2013). This has been addressed by the latest WPS 2.0 standard, which has decoupled the process model from the core WPS specification. On technical suggestion is to use the SensorML process model as an alternative method to describe parameters, inputs and outputs (OGC, Mueller, & Pross, 2015). N. Chen et al. (2012) have experimented with SensorML to construct a geoprocessing workflow system. This current situation indicates that SensorML is a potent medium for a convergence towards a unified encoding for the currently independent, disjunct process encodings for WPS (process as algorithm), SPS (process as sensor or actor) and SOS (process as sensor or algorithm). Consequently, if algorithms are processes which generate results, a WPS process is naturally the procedure for a SOS observation data stream. This fits well with the OGC/ISO observation concept.

Surface hydrological modelling seems to be well understood and a variety of models and software implementations are available. CUAHSI-HIS is a very good example of such an integrated infrastructure with linkage of WaterML time series data into hydrological models. Linking several of those hydrological models has been described with OpenMI. OMS3 provides similar type model coupling to OpenMI. However OMS3's architectural concept is based around composing modular geoprocesses into a workflow, whereas OpenMI provides interfaces to exchange data at same time step between (hydrological/hydrogeological) models. Both OMS3 and OpenMI need to run within a local model execution environment and thus can not be decoupled in a distributed computing setting.

Besides the 'Interoperable Freshwater Models' project (Elliott et al., 2012; Rutledge et al., 2012) described above there are no advancements in integrated modelling approaches documented for New Zealand in the literature. Current interorganisational modelling exercises stick to separate and independent model runs with occasional file-based (manual file transfer) data exchange. Furthermore, the metadata (data quality, provenance) for the data different source elements should be propagated and transported together with the process outputs to be able to track metadata lineage throughput the modelling exercises. This has been partially described with UncertWeb (Bastin et al., 2013) and Intamap (Pebesma et al., 2011) based on interpolated point air quality observation data with a focus on communicating uncertainty for the interpolated values. But has neither generalised on overall data quality, provenance or metadata lineage, nor has it been examined on other environmental, i.e. hydrogeological use cases.

- Data for integrated environmental modelling is scattered and difficult to obtain. State-of-the-art data acquisition and processing into the right formats costs a lot of time and resources.
- Updates of previous modelling results are almost unavailable on a on-demand basis and examples did not consider real time measurements for modelling.
- A workflow from initial data capturing to the real-time provision of conditioned information delivery to end users for prepared decision making needs to be automated and operationalized.
- The modelling workflow needs to allow linking of models via distributed and chained processing in the web. OpenMI and OMS can be considered for model interaction/linking, but need to run within a local model execution environment and thus can not be decoupled in a distributed computing setting.

2.6 3D VISUALISATION IN THE WEB

For web-based visualisation of 3D geospatial data and models, several technologies, standards and methodologies have recently been described in the literature (M. E. Cox, James, Hawke, & Raiber, 2013; Evans, Romeo, Bahrehmand, Agenjo, & Blat, 2014; Morozov, Chubak, & Blyth, 2009; Schmidt & May, 2012). As current web browsers support native graphics acceleration technologies, with WebGL as state-of-the-art, many 3D/4D visualisations have been developed for the web browser (Resch, Wohlfahrt, & Wosniok, 2014; Stein, Limper, & Kuijper, 2014; WebGL, 2012). However, these applications lack an explicit declarative approach to the mapping of the geoscientific data into the visual domain.

Three-dimensional (3D) visualisation of geological data like wells and boreholes has been demonstrated by Beard (2006) and the Canadian Groundwater Information Network (GIN, 2011) portal with the ISO standard Virtual Reality Modeling Language (VRML) (ISO/IEC, 1997) (ISO/IEC, 1997). In addition to 2D spatial visualisation (web mapping) with OGC WMS, access to data sets with OGC WFS and the GWML output encoding has provided additional analysis and discovery services for 3D groundwater processes (Dahlhaus et al., 2015). Typically, web map viewing applications display 2D pictorial representations of geospatial groundwater data layers from WFS and WCS layers via OGC compliant WMS and graphs for hydrological time series data from SOS. Data from SOS servers are often represented and grouped by their sampling locations, i.e. sensor stations. Locations are plotted on a web map and the time series data for available phenomena are then requested and graphed (Brodaric et al., 2011; Dahlhaus et al., 2011).

Alternative developments enable 3D (hydro-)geological, or hydro-climatological data visualisation in the web using specific 'portrayal' methods. M. E. Cox et al. (2013) describes a web sockets link to a scientific visualisation server where the prepared data sets reside. The OGC 3D draft standards Web Perspective View Service (WPVS) and Web 3D service (OGC, Schilling, & Kolbe, 2010) propose a service interfaces for server-side rendered images. A few developments have been advanced for urban landscape visualisation in conjunction with OGC City Geographic Markup Language (CityGML) (Gröger, Kolbe, Nagel, & Häfele, 2012; Gröger & Plümer, 2012). However, research seems to have stalled or developments progressed into closed, inaccessible, commercial space. The described OGC 3D view services drafts and proposals do not address the origin of data in-depth and Commercial CityGML examples are not reproducible, because they only visualise feature sourced from a local database. The challenge is that, when assembled from different sources, the metadata for the different source elements should be communicated and transported together with the aggregated scene elements. Only then licensing and provenance information can be coalesced into metadata lineage for the visualisation products.

From an academic point of view declarative approaches are advantageous because they are reproducible and transformation is based on formal patterns. OGC WMS provides such a declarative visualisation for simple features. The OGC Simple Features specification (van den Brink, Portele, & Vretanos, 2012) describes their structure as a flat attribute table and one allowed geometry element. The OGC specifications Styled Layer Descriptor (SLD) and Symbol Encoding (SE) provide rule-based means to style and decorate web map rendering based on the feature attributes and by its geometry type (i.e. point, line, and polygon). In contrast to flat attributes of simple features, complex features can have nested properties, which themselves can be complex features again. WMS cannot select rendering rules for complex features, as SLD selectors are based on the flat property type structure. GeoSciML as a complex application schema faces the same challenge. The GeoSciML community developed a portrayal service to enable a unified web mapping method [2].

Traditionally, detailed 3D visualisations had to be rendered on powerful desktop computers and workstations with high-performance graphics acceleration. For web browsers, plug-ins were necessary to show advanced 3D content. Java applets, 3D viewers plug-ins, Adobe Flash and Microsoft Silverlight became popular that could harness the computer's graphics capabilities for the web content. Now Hyper Text Markup Language (HTML) in version 5 allows developers exposing graphics acceleration via WebGL, even on mobile phones providing improved real-time rendering capabilities. A range of JavaScript libraries now provide native and direct access to modern web HTML5 and 3D acceleration capabilities.

Extensible 3D - A 3D Scene Graph Markup Language (X3D) (ISO/IEC, 2008) is the successor of the ISO standard VRML97 and has been demonstrated in the web browser particularly with x3dom (Behr, Eschler, Jung, & Zöllner, 2009; Behr, Jung, Franke, & Sturm, 2012; Mouton et al., 2014). X3dom is a JavaScript library that makes extensive use of the HTML5 and WebGL integration of modern browsers. It distinguishes itself from other visualisation frameworks through its focus on integrating X3D scenes directly into the HTML Document Object Model (DOM), where it can be manipulated like other declarative HTML content.

While all these approaches render static data sets into a 3D scene, none considered a dynamic integration with on-the fly generated data sets. To

2 GeoSciML and Portrayal documentation http://www.geosciml.org/

visualise the spatio-temporal distribution of phenomena geostatistical algorithms can be exposed via OGC WPS; they can use near-real-time observation data from SOS services, and thus, can be integrated into web based visualisation systems (Castronova, Goodall, & Elag, 2013; Pebesma et al., 2011). WPS allows distributed and chained processing in the web, computing power can be extended into the cloud, i.e. multiple instances run in parallel on different servers in the internet to extend visualisation as a service in the SOA (Giuliani et al., 2012; Hildebrandt & Döllner, 2010) to connect web based spatial data provisioning and web processing services.

- Many current plugin-free 3D/4D visualisations have been developed for the web browser. But they lack an explicit declarative approach of mapping geoscientific data into the visual domain.
- When assembled from different sources, the metadata for the different source elements should be communicated and transported together with the aggregated scene elements, because licensing and provenance information should be coalesced into metadata for the visualisation products.
- While current approaches render static data sets into a 3D scene, none have described a dynamic integration with on-the fly generated data sets.

Part II

METHODS AND MATERIALS

3 METHODOLOGY

3.1 CONTEXT AND PERSPECTIVE

GIScience, also Geographical Information Science and Systems, is a multi-faceted research discipline and comprises a wide variety of topics (Goodchild, 1992; Mark, 2002). The investigation into the data management and interoperability of geographical data sets as well as environmental data sets with spatio-temporal context with tools and processes for scientific analysis, visualisation and modelling is a particular challenge of the Information Science aspect of GIScience (Yang et al., 2010). Methods and techniques around semantic web and linked data, data visualisation, distributed computing or knowledge engineering are common research topics in pure Information Science and Systems research. While such research is mainly conducted in the domain of business economics and pure Computer Science studies, the interdisciplinary Information Science aspects are an essential research aspect to integrate Information Science not only with the wider field of Geography, but even further with the domain of Geosciences or Earth Sciences (Blaschke et al., 2011; Gahegan et al., 2009).

To advance theory in GIScience, existing gaps need to be revealed and the current understanding of phenomena needs to be improved. The context of this research is 1) the ever increasing collection of specifically geoscientific as well as environmental data in general, in a wide variety of formats, 2) the evolving data formats, data and database models and formalised standards and 3) their application in data collection, storage, harmonisation, discovery, access and transfer, in the Earth Sciences. For this thesis the immediate focus is to digitally better enable Hydrogeology and the related fields where necessary (Figure 5). A central assumption in this research framework is that methodology and techniques of geoscientists, GIS practitioners, information modellers and computer scientists are not sufficiently coalesced to improve efficiency and effectiveness of data handling, for example in hydrogeological assessments (Bandaragoda et al., 2006). Figure 2 (in Chapter 1) demonstrates how GIScience as an inherently holistic discipline can be the connecting link between the Earth Sciences and Computer Science. Since there are also non-

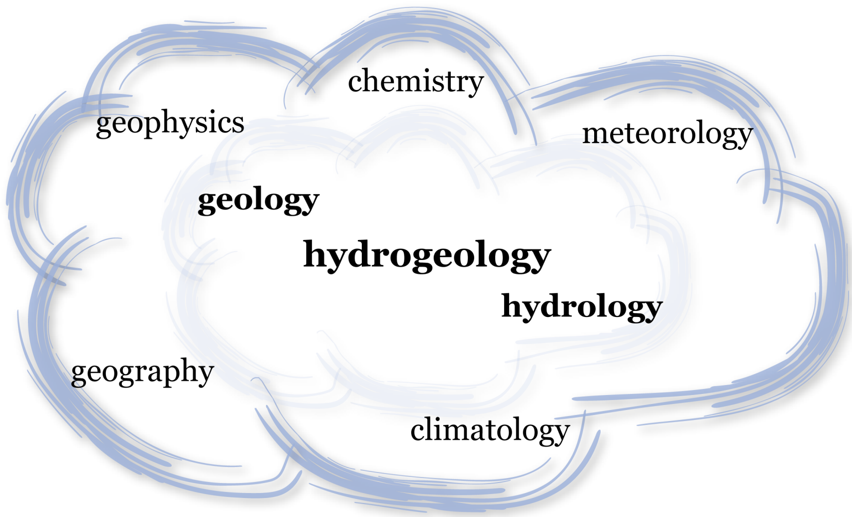


Figure 5: Earth Sciences in immediate and indirect focus for digital integration in this study

academic professions involved, such studies are not only interdisciplinary but of transdisciplinary nature (Tress et al., 2003).

The researcher needs to explain his research perspective to clarify the application of methods and modes of scientific inquiry. This is particularly important, if qualitative methods are employed and possible bias is involved (Creswell, 2003; Holliday, 2010; Saunders et al., 2012).

Figure 6 shows the challenging irony of research required to be conducted in a strictly positivistic community. GIScience as a research discipline comprises a large variety of research topics, with some of them being classically of quantitative and empirical nature, by for example determining statistical and spatial relationships of geographical features. Other topics are not always as clearly objective or empirical in nature, e.g. the GIS software design and development as part of GIScience theory development and knowledge contribution. Thus, to avoid disguising the conduct of this study under the cloak of classic scientific method, an explicit approach has been chosen that better represents the needs of framework development in the GIScience and Systems discipline.

Classical philosophy of science differentiates between objective and subjective epistemological stances, which define the researcher's relationship with the notion of what is knowledge and truth. An objective epistemology with a positivist, empirical paradigm embodies the Scientific Method and assumes an external reality. Thus, there is only one truth and the generated knowledge is independent of the researcher (Goles & Hirschheim, 2000; Gray, 2014).

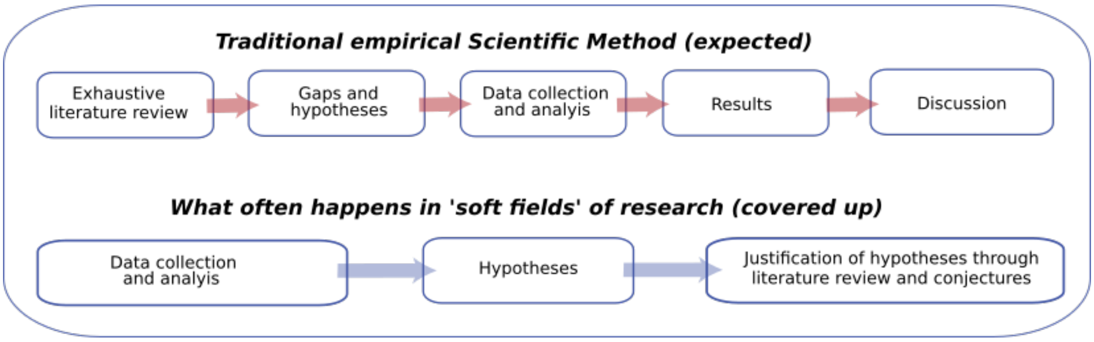


Figure 6: Discrepancies in modes of research typically adopted in GIScience, adopted from Walsh (2014)

A subjective epistemological stance assumes that knowledge is not value free and thus the reality is dependent on the researcher, human perceptions and the social context. Therefore, bias should be articulated accordingly (Bhaskar, 1998; Dobson, 2002; Mingers, 2004). Interpretivist and constructivist paradigms are commonly described as subjective perspectives (Feast & Melles, 2010; Hirschheim, 1985). In contrast to traditional approaches, philosophical pragmatists deny the correspondence notion of truth completely and propose that truth essentially is what works in practice, thus in Information Science and Systems research progress is achieved when existing technologies are replaced by more effective ones (Goles & Hirschheim, 2000; Porra, Hirschheim, & Parks, 2014). Critical Realism is another, more mediating and increasingly accepted philosophy of science (Bhaskar, 1998; Dobson, 2002). It can be seen as a bridge between positivist and constructivist/interpretivist stances and guides the researcher in addressing bias in qualitative and mixed methods research (Scott, 2007; Zachariadis, Scott, & Barrett, 2010). A researcher with a Critical Realism point of view reflects on the underlying knowledge generating process, the conflicts in society and takes on an emancipatory role (Easterbrook, Singer, Storey, & Damian, 2008).

Although most researchers focus on one particular research method, some suggest combining one or more research methods via means of triangulation (Couclelis, 2009; Gray, 2014; Sikolia et al., 2013). The research stance taken in this research is a mixed methods approach that integrates several qualitative methods of inquiry, i.e. DSR, GT and Case Studies (Beck et al., 2013; Glaser & Strauss, 1967). GT acknowledges that the researcher's background has an impact on their ability to let patterns emerge from the data, and that preconceived ideas can influence their ability thereof. Here the reflecting perspective of Critical Realism is supporting the researcher's discipline in handling bias and avoiding preconceived ideas. The constant critical reflection is in tune with letting categories emerge while constantly comparing theory with new data (Birks, Fernandez, Levina, & Nasirin, 2013; Thornberg, 2012). From the DSR perspective it is important to address perceived challenges that a new design should overcome (Beck et al., 2013; Gregor, 2002; Gregory, 2010). Here the researcher needs to take an emancipatory stance how the emerging theory can inform the design iterations. The research methods and how they are interwoven and applied will be explained in the following sections.

3.2 GROUNDED THEORY

GT was first developed by the two sociologists Barney Glaser and Anselm Strauss in 1967 and was first published in 'The discovery of grounded theory' (Glaser & Strauss, 1967). Glaser describes GT as being a full research methodology. Its core emphasis is on discovery of theory from data. Since then GT has been used in a variety of disciplines including Information Systems and Software Engineering (Urquhart, Lehmann, & Myers, 2010). As opposed to an empirical positivist approach of testing existing theory, GT presents a method for generating theory from data. This data can be collected from interviews, documents, observation and from other sources as soon as the data may become available. Data analysis involves the identification of concepts, subcategories, and categories and how they relate to each other. The emerging categories and their relationships are then checked with existing literature in the field to explain how they relate to each other. While in Information Systems research GT is often used for coding their data, the prime result of this methodology is the development of an underlying theory (Urquhart et al., 2010). Figure 7 illustrates the stages, processes and in- and outputs of GT. The following sections describe these stages in detail.

3.2.1 *Hypothesis Testing versus Emergence*

One particular aspect that differentiates GT from many other research methods and methodologies is that it is explicitly emergent and does not aim to prove or falsify a hypothesis (Glaser, 2008; Glaser & Strauss, 1967). The aim is to understand the research situation, as Glaser states it, to discover the theory implicit in the data (compare with Figure 6). GT provides reconciliation where the field of geographical and environmental spatio-temporal data and software integration for earth sciences is complex and scattered. The in-depth

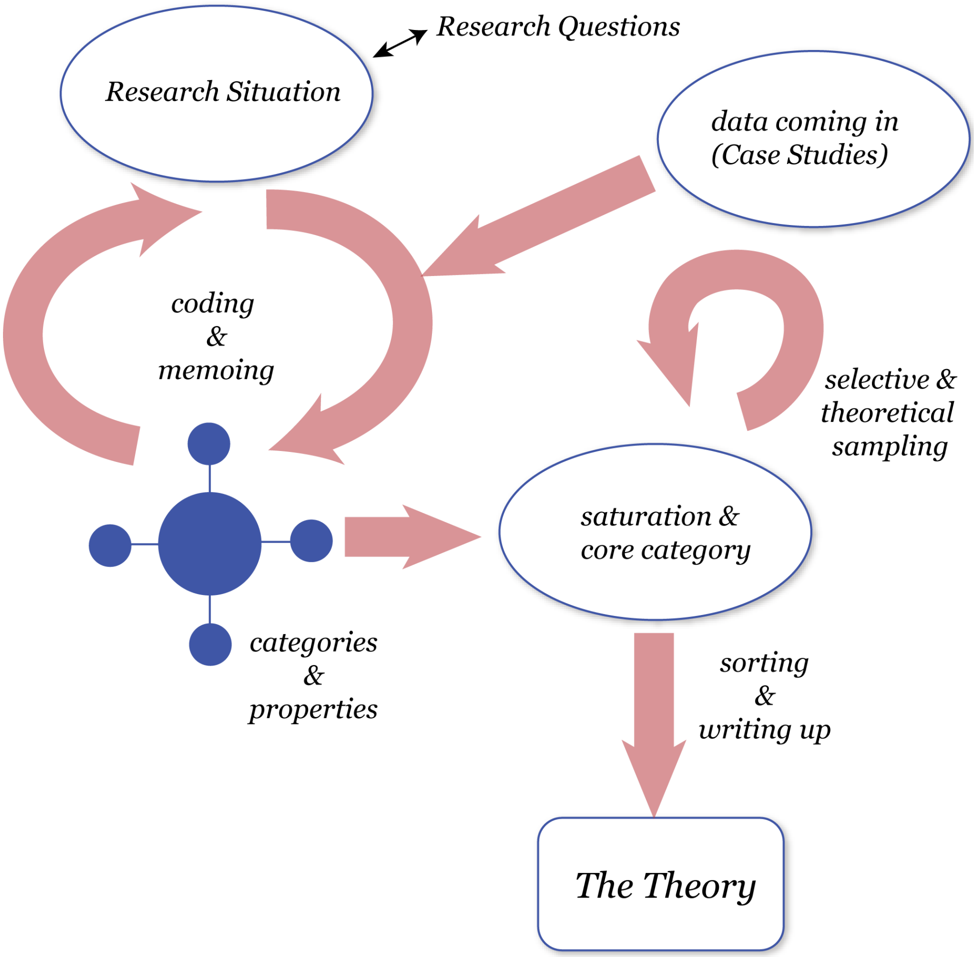


Figure 7: Grounded Theory Process, from Research Situation (Section 3.2.3), via Coding and Memoing (Section 3.2.4) and employing the Constant Comparison Method (Section 3.2.2) in order to distil categories and properties (Section 3.2.4) from data, i.e. the Case Studies; eventually achieve saturation (Section 3.2.5) and find the core category; then writing up (Section 3.2.6) the resulting theory;

analysis of the situation based on GT will reveal required patterns to inform a holistic design approach subsequently (Beck et al., 2013).

3.2.2 *Constant Comparative Method*

The single most important and powerful tool in GT is the Constant Comparative Method (Glaser, 2008; Sikolia et al., 2013; Strauss & Corbin, 1994). It is a process by which codes, as well as later emerging categories arising out of each piece of data are constantly compared against the codes and categories from the same data, other observations and with the emerging theory to produce higher levels of abstraction. In the onset the research situation is observed: What is the current situation, which forces and constraints work and what are consequences? This will be then transcribed into patterns and constantly compared, what categories are suggested by that data and subsequently with the emerging theory in mind. That is the constant comparative method - initially comparing data with data, and later comparing data with theory.

3.2.3 *Research Situation, Minor Literature Review*

In GT initial research questions are to be avoided and that an in-depth gap analysis of the field under investigation might blind the researcher's ability to let theory emerge. But to frame the area of interest and based on first impressions from the field as well as to comply with academic guidelines for dissertations, the researcher can start off with a minor literature review to be able to understand technical jargon and the current research situation within the data. Glaser instead suggests background reading in adjunct fields to provide the understanding to make sense of the data, while avoiding the most closely related literature. Consequently, the increasingly comprehensive literature review can be progressed as the literature becomes relevant to the emerging categories and theory, as otherwise early reading may constrain coding and memoing too much (Glaser & Strauss, 1967; Strauss & Corbin, 1994; Thornberg, 2012).

An important concern is how the research treats disagreement between the emerging theory and the literature. The researcher's ambition throughout to fit the emerging theory to the data and to make sense of the actual situation. A means to tackle the assumption that the theory could be wrong is to stick critically with the constant comparative method and also to seek disconfirming evidence. A study that applies Grounded Theory seeks to extend the theory so that it makes sense of both the data from the case study and the data from the literature (Birks et al., 2013; Sikolia et al., 2013).

3.2.4 *Coding and Memoing*

The analysis of the body of data starts with coding, i.e. excerpting key points which are then assigned a code. Initially the code might be a phrase that summaries the key point in two or three words. Over time similar code patterns emerge. Through constant comparison codes that share meaning can be abstracted to categories. Figure 8 shows an example of how applying the constant comparative method can result in absorbing the data into an existing category, or creating a new because of sufficient distinction. Glaser lists several abstract theoretical coding structures which can be used as a framework to describe how the categories relate to each other. This is called Theoretical Coding (Glaser, 2008; Glaser & Strauss, 1967).

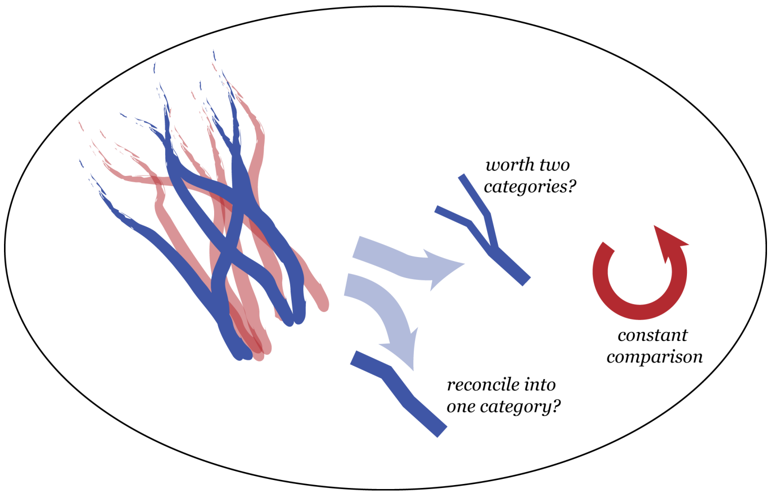


Figure 8: Processes in Constant Comparative Method on data: absorbing data into an existing category or creating a new category

Categories and properties

A category is a theme or distinct context which embodies certain characteristics. It is interpreted in the light of the emerging theory. Properties in this case are basically sub-categories that split off finer distinctions in the categories where appropriate.

Core category

Eventually several categories emerge as a result of data analysis and one category will be found to emerge with a higher frequency. This higher frequency category is able to account for most variations in the data and also connects meaningfully with many other emerging categories. This is likely to be the core category. Guidelines advise that the coder should not be too eager to choose a core category or to choose too early in the data collection. However, when it becomes clear that a category is mentioned with high frequency and is well connected to other categories, it should be adopted as the core category.

Memos

Memoing is an on-going process of writing conceptual notes throughout the GT process whenever ideas occur to the researcher. The memos are intended to guide emergence of conceptual links between categories the researcher notes down their relations to different categories. In GT it is assumed that the theory is concealed in the data and coding makes some of its components visible. Building on that memoing captures the relationships which link the categories to each other.

3.2.5 *Sampling and Saturation*

As categories emerge from data, the researcher seeks to add samples in a way that it further increases the diversity for the purpose of strengthening the emerging theory. This is done by defining properties of the categories and how those properties effectuate from category to category. Once the core category is established, the researcher then only codes for the core category and those categories that are closely related to the core. Collecting and interpreting more data about a particular category over time at some point does not add any new insight about that category, its properties, and its relationships. Then the category is said to have reached Theoretical Saturation and the researcher can then stop collecting data and cease coding for that category.

3.2.6 *Sorting and Writing-Up*

Once the data collection and coding are finished and categories saturated, the theoretical memos can be arranged on a conceptual level, which is called Sorting. Sorting allows forming an outline of the theory which aims to explain how the categories relate to the core category. As the theory starts to emerge, the researcher can conduct extensive literature review to see how the literature in the field relates to their emerging theory. Finally, the resulting theory is written up. In this study the process of coding, sampling and memoing in fact informs a design process. Every iteration presents a new refined blueprint and is included into the theory and thus becomes subject to the constant comparison method. The generic design development principles are outlined in the following section.

3.3 DESIGN SCIENCE RESEARCH

DSR is a particular 'lens' of research methods in Information Systems (IS) research which fosters creation of new knowledge through the designing and implementation of innovative artefacts (Gregor & Hevner, 2013; Hevner & Chatterjee, 2010, p. 27). Figure 9 demonstrates how DSR proposes a scientific research framework for innovative software development and the creation of new artefacts based on following criteria: 1) awareness, 2) suggestion, 3) development, 4) evaluation and 5) conclusion. These criteria are then executed along iterative stages to be improved, where the first stage merely represents the starting point for DSR theory development (Beck et al., 2013; March & Smith, 1995; Owen, 1998; Takeda, Veerkamp, Tomiyama, & Yoshikawa, 1990; Vaishnavi & Kuechler, 2004).

3.3.1 *Awareness*

The starting point in DSR is awareness - the recognition of a problem which can be solved by using or developing new artefacts. On one side, GT and Case Study frame the awareness. On the other side, the researcher's experience in Computer science and GIScience and the perceived discrepancies at the first encounters with geodata management in the Earth Sciences raised awareness for the need of improvement. Based on the 'Solution Space vs Problem Space Maturity' (Figure 10) matrix of Gregor and Hevner (2013) the proposed method acts around the interface between 'Exaptation' (non-trivial

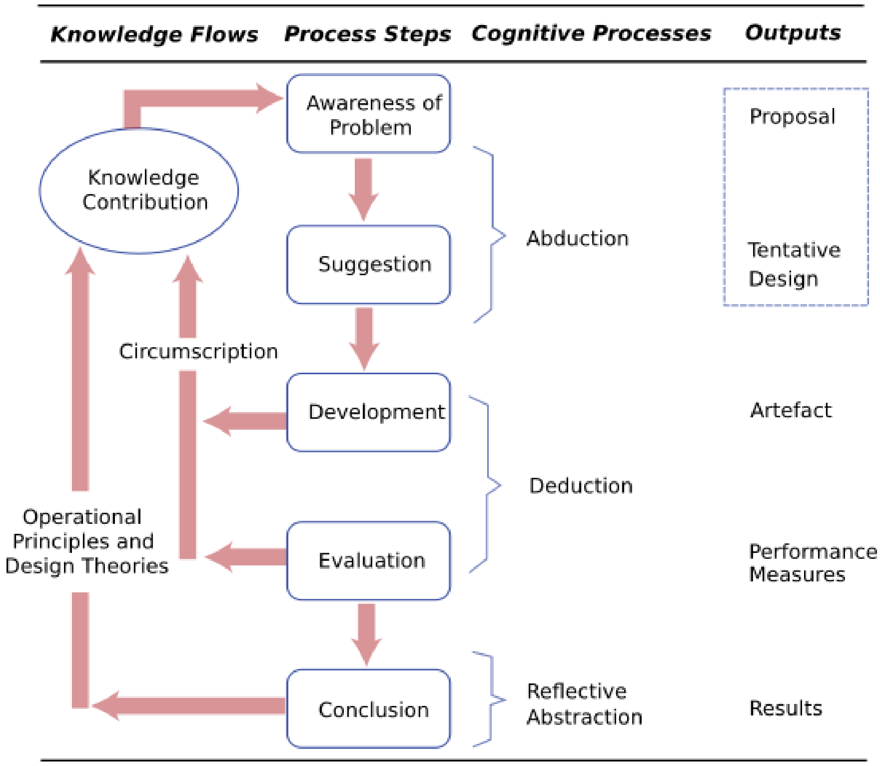


Figure 9: Design Science Research Process Model and Cognition throughout the cycle, adopted from Gregor and Hevner (2013); Takeda et al. (1990); Vaishnavi and Kuechler (2004)

extension of known solutions for new problems) and 'Invention' (new solutions for new problems) and critically distances itself from 'Routine Design' (applying known solutions to known problems). Further considerations in the complexity of the problem are the different levels of conceptualisations:

- real world trying to be captured in data: e.g. geographical feature abstraction and data formats designed by humans to capture real world objects/features and besides standardised ways, i.e. OGC/ISO, there are different ways used to do that
- different types of problem analysis of how networked (geographical) information systems can use those data: a) how effective/efficient they are, and b) being used by humans with data in data formats, which are human artefacts themselves
- which data types and systems orchestrations are more comprehensive than others, and how well information systems that use those data can handle those data

- mentioned challenges are addressed in partially, but rarely as a holistically for GIScience and Systems and Earth Sciences

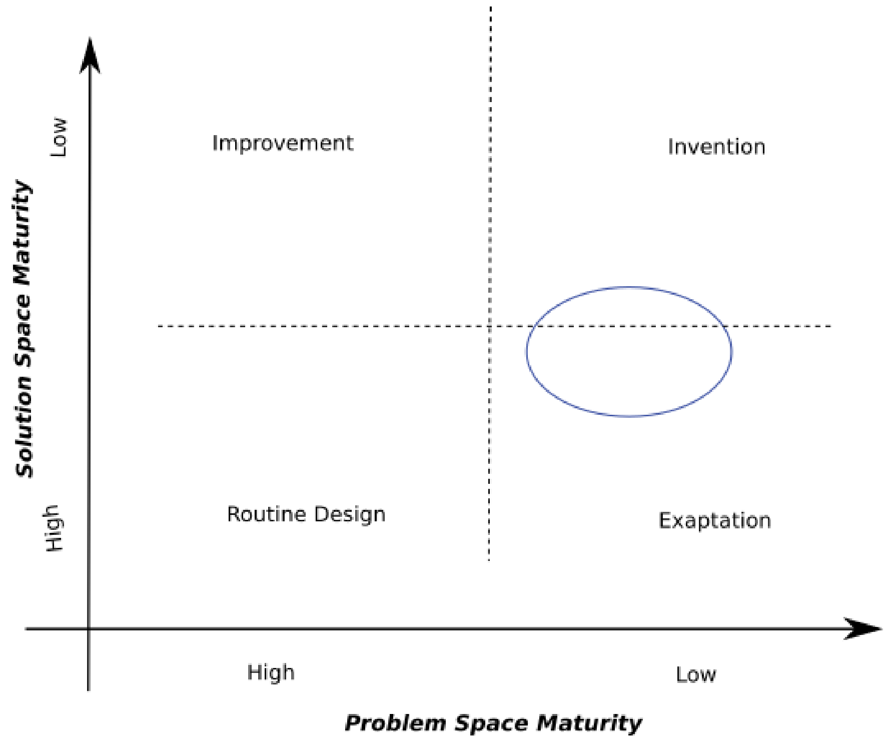


Figure 10: DSR Knowledge Contribution Framework and where this study's outputs reside, from Gregor and Hevner (2013)

3.3.2 *Suggestion*

Consequently from the awareness of the problem, suggestions are made, discussing what kind of artefact might solve the problem. Here GT informs design suggestion process. That is a particularly important aspect, because GT alone would replicate the current fuzzy state-of-the-art which is not a succinct design but a confused landscape of patchy realisations, addressing only subsets - with GIScience theory and integrated geoscientific (or environmental) modelling as the emerging desire, but no proven theory how to achieve that. GT informs the design process to improve the design theory, based on the case study data and literature and the revealed gaps in the existing theories - which results in design suggestions (Beck et al., 2013; Thornberg, 2012). Furthermore, consultation with domain experts (e.g. hydrogeologists) is necessary to ensure applicability in the target domain.

3.3.3 *Development*

Based on the suggestions, an artefact is designed and implemented in the development phase. The implementation can be pedestrian and prototypical, depending on the anticipated goal. Furthermore, the design as well its implemented artefact is supposed to be improved over multiple iterations.

3.3.4 *Evaluation*

In the evaluation phase of the design process it is checked whether the artefact solves the problem, and its strengths and weaknesses are analysed. Software interfaces and data models can be tested experimental through manual implementation if no reference benchmark is available or if the goal is not a quantitative performance improvement, but the general testing of applicability of a new design. Here again, domain experts can be included into the process in order to evaluate the applicability of the developed artefacts in the domain of interest.

3.3.5 *Conclusion*

In the conclusion stage results and future aspects such as open questions or plans for further development are compiled and discussed. This can well be the input for the next iteration or if the developed theory for the design becomes clear and eminent it should be written up and documented accordingly. The GT process of developing theory is the main guiding input for the iterative improvement stages of the framework design (Gregor, 2006; Gregory, 2010). Emerging categories and theoretical sampling of the similarities and the differences of information (implemented standards, addressed problems, orchestrated automated services) are consulted and reviewed with the framework design and thus the next stage of implementations (Owen, 1998; Takeda et al., 1990). Thus, the grounded theory is objectively refined into a novel design theory.

3.4 case studies for data collection and evaluation

GT provides the instruments to reason over all data as they become available. This creates a theory which is grounded in the data and which is the foundation of the design blueprints. The data is predominantly drawn from the context of the Earth Sciences - as the geographical case study areas on various scales, from local to global scale.

In Case Study research as qualitative strategy of inquiry the researcher explores events, processes and forces that generate and consume data. It is investigating phenomena in real-life settings looking for supporting empirical evidence for a well-formulated theoretical model (Yin, 1994). This implies that the a priori testing defined hypothesis is constructed from the existing literature. Also clear research questions, case selecting criteria, information about data collection methods, and clear explanation of the data analysis process need to be provided for rigorous research (Dubé & Paré, 2003).

Some of the presented characteristics of Case Study research in Information Systems contradict GT methodology, in particular the a priori paradigm. However, case studies can be used to generalise theory from. Thus Case Study research is appropriate to frame the method triangulation for this research (Couclelis, 2009; Creswell, 2003; Creswell & Miller, 2000; Scott, 2007).

For GT the case studies are one of main sources of data, as well as emerging technological and theoretical developments regarding environmental data management. Figure 11 illustrates how theory is drawn from, but also constantly reconciled (compared) with the data. An exploratory literature review can also serve as data source, particular when sampling for emerging categories. The aim is to compare literature to the emerging theory in the same way that data is compared to the emerging theory. In an emergent study, it is likely to be unknown at the beginning which literature will turn out to be relevant later. Thus, the constant comparison of GT remains an essential core process.

In DSR, progress is achieved when existing technologies are replaced by more effective ones. DSR suggests substantive tests in the sense of natural science research. Not only must an artefact be evaluated, but the evaluation criteria themselves must be determined for the artefact in a particular environment. The literature review revealed the state-of-the-art (best practices, important considerations) and the respective gaps in the interdisciplinary field of this study. From the state-of-the-art best practices and the current gaps identified in the literature a set of data infrastructure evaluation criteria were established. Each criterion is either a 'shall have' consideration from the state-of-the-art best practices or an 'overcoming the gap' translation from an identified gap in the literature review. Table 2 lists the written out evaluation criteria. Each criterion is identified by a unique identifier (ID). The prefix given to each ID indicates the domain from which the criterion was

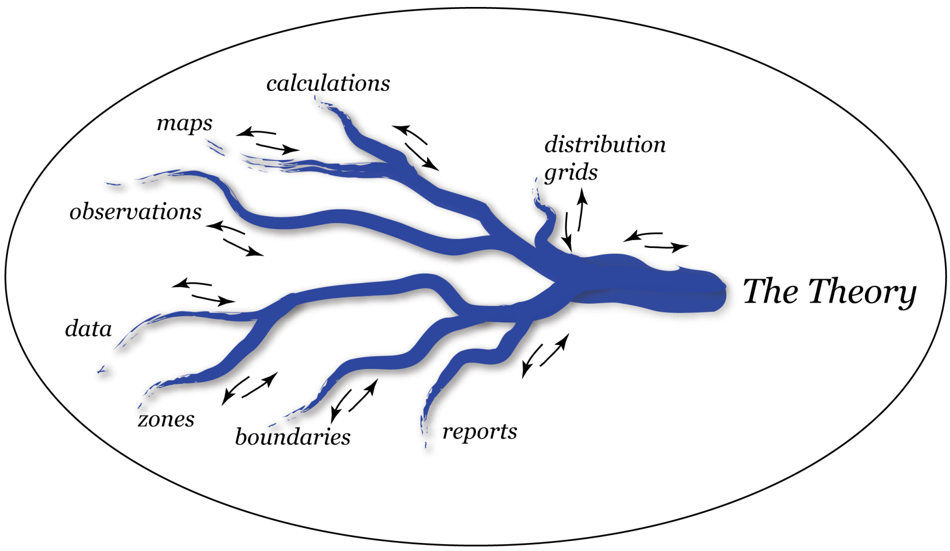


Figure 11: Theoretical rooting of GT in Data from Case Studies

derived - i.e. HYNZ (hydrogeological data situation in New Zealand), SDI (classic Spatial Data Infrastructure), SWE (Sensor web enablement and live telemetry integration), MOD (hydrogeological modelling and model web realisation), VIZ (the visualisation component). The implemented components will be evaluated with respect to their application to the case studies.

By their nature, single case studies and instantiations of a design do not meet the requirement of 'generality' that is defined for research. Typically, in qualitative research some type of 'opinion per person' data is collected, e.g. from interviews or surveys. The body of data analysed in this study are data sets and technologies from the case studies, formal standards developed by national and international expert groups correlated with the scientific, peer-reviewed literature. Those data are human artefacts. However, all those data are generated in a either consensus oriented way or peer-reviewed embody expert knowledge on different level. The major case study is the set of data sets in the groundwater domain and the different available data encodings.

Table 2: Resulting evaluation criteria from the literature review

ID	Criterion
HYNZ1	environmental data is immediately available and accessible in standardised data interchange formats
HYNZ2	implementation addresses interdisciplinary data harmonisation and interoperability of services
SDI1	supports networked web-based GIS access
SDI2	formats are consistent
SDI3	metadata exists in consistent formats and can be used for provenance, licensing and lineage of data sets
SDI4	the design paradigm surpasses classic information systems design approaches
SDI5	phenomena, entities and parameters can be formally described and accessed in machine-readable formats
SDI6	identifiers and concepts can be inter-linked, related and thus, descriptions re-used
SWE1	wireless sensor network data collection uses standards to improve access to the data and sensor configuration
SWE2	data protocols are appropriate for low bandwidth unreliable network connections
MOD1	user is concerned with data pre- and post-processing when modelling
MOD2	user can easily run updates with same model configuration but with updated source data
MOD3	modelling can include real-time measurements
MOD4	(modelling) workflow can be automated and operationalised from data capture via modelling and prepared information (visualisation) for decision making
MOD5	(modelling) workflow can be distributed over the web and link data and modelling resources across multiple organisations
VIZ1	works plugin-free in web browser
VIZ2	can generate transportable model 3D scene
VIZ3	provides declarative translation from geoscience data to visual model
VIZ4	support transport and communicate metadata
VIZ5	support dynamic on-the fly generated data sets/sources

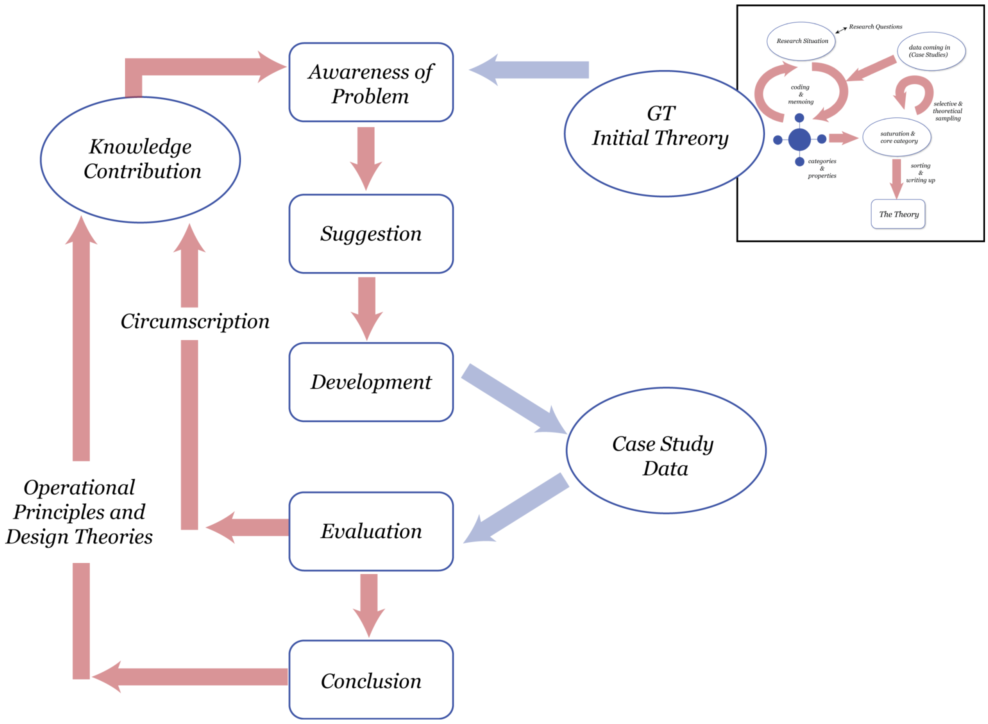


Figure 12: The 'Grounded Design' method - Design Science Research extended process flow with Grounded Theory and Case Studies

3.5 SUMMARY

In this section the research methodology as well as the interplay and triangulation of the selected methods are discussed. Figure 12 illustrates how the Design Science Research process flow has been extended with Grounded Theory and Case Studies. This results in the explicit triangulated research framework, hereby referenced as 'Grounded Design' method, that will be applied throughout this dissertation. The literature review (Chapter 2) presents the current state in the relevant fields. Sampling from the New Zealand case study data (Chapter 4) and the presented literature will generate categories and find emerging patterns. The resulting framework design (Chapter 5) will be implemented and its technical applicability evaluated in respect to the theory and in relation to the New Zealand case studies (Chapter 6).

Rigorous principles of good qualitative research are a reproducible and transparent explanation of methods, reliable submission to the chosen methods as well making appropriate claims. These criteria allow judging the quality of the subjective research as compared with the positivists' labels: credibility (internal validity), transferability (external validity), dependability or consistency (reliability) and confirmability (objectivity) (Creswell, 2003; Creswell & Miller, 2000; Sikolia et al., 2013).

Why Grounded Theory?

The core features of GT, i.e. the Constant Comparison method and the coding from all available data to retrieve emerging patterns that solely stem from the body of data, make it a formidable mode of inquiry for the researcher to examine the outlined transdisciplinary problem domain and develop relevant, applicable and generalizable theory. Literature review is conducted for new codes and categories as they emerge and theory is constantly compared with new data and amended or re-articulated. GT provides the main pathway along the Objectives 1 ('Review methods, standards, formats') and 2 ('Identify, select and discuss an appropriate set thereof') towards answering the Research Question 1 ('What are the required interfaces and data formats?') in chapters Chapter 4 and Chapter 5.

Why not only Grounded Theory?

Originally developed in and for the social sciences, most literature about GT refers to the data as the interviews with participants. However, in order to develop a theory for GIScience, Information Systems, Computer Science and the Geosciences this study wants to draw from existing geodata, software systems, standards/best-practices and the literature. Hence this theory claims to improve GIScience methods for interdisciplinary geodata management, processing and visualisation in distributed computer systems. Thus, GT is triangulated with DSR to provide the theory for the systems design and its iterations and with Case Study research as source of data and artefacts.

Why Design Science Research?

The first and foremost feature of DSR is to provide a rigorous and structured methodology to design an artefact, i.e. create and innovative piece of software that improves or supersedes existing implementations. DSR maintains that only through implementations, i.e. realisations of a theory in a real-world artefact, can reveal, prove or falsify properties, behaviour and relationships between components that might or might not have been theorised beforehand. Furthermore, drawing from the experiences of the implementations the design will be improved through iterations of implementations and exploring its properties, behaviours and relationships. DSR guides the work on Objective 3 to advance the theory for designing interoperable data and analytics infrastructures. Thus, Research Questions 2 ('Can existing models and software modules be linked to enable interoperable models?'), 3 ('How can such models be provisioned with available interoperable data sets'?) and 4 ('Can the visualisation of interoperable data and models be realised with the same design principles?') will be answered through the exploratory and iterative implementations, improvements and design evaluations in Chapter 5 and Chapter 6.

Why not only Design Science Research?

Design Science Research is a relatively young research discipline and the literature is not as abundant. The arrangement of a multitude of interfaces and standards and formats (i.e. ISO/W3C/OGC orchestration) cannot easily be measured in quantitative metrics, in particular if there are no gold benchmarks available to compare with. The new framework design implements a methodology which has not been explicitly described like this, thus a comparison is only possible on a descriptive and explanatory basis. Also, the 'magical' spark of creativity, the idea generating process to arrive at the design blueprints and their improvement is not well understood. Here, GT provides both: qualitative means of evaluation; and a documented, abductive/inductive approach that is grounded in the actual (Case Study) data.

Why the Case Study method?

The case study notion espouses the mentioned two methods with a limited set of domain (the various Earth Sciences disciplines) data constrained in time and space. Case Study results and generated theory shall be generalizable and applicable to other case studies. Case Study in GIScience has a particular nuance of geographical case study areas and thus serves as an explicit body of data for GT and DSR. Finally, through the triangulation of the described research methodologies, the Research Question 5: ('Can the gaps from the literature be overcome?') can be answered in Chapter 8.

Why not only the Case Study method?

Case Study research in a rigorous scientific inquiry requires a fixed priori defined research questions and quantitative measures for validation to prove or falsify the a priori hypothesis. However, in this study through the triangulation with GT and DSR with both having a posteriori theory generating paradigm, the Case Study method is not executed on its own. It serves as an explicit field of inquiry for GT and DSR, and by its nature, the generated theory from within the constraints of this Case Study are sought to be generalizable to other (in particular geographical) case studies or even globally.

Finally, in an analogy, the GT processes of sampling and constant comparison could be depicted in a cognitive conceptual space. After Tobler's First Law of Geography (Tobler, 1970), closer things are more related to each other than more distant things. Logically, similar concepts are closer to each other in our mentioned cognitive space. GT aims to produce theory through letting patterns emerge from the data. In GIScience alike, patterns arise from the data through geographic analysis which characterises the spatial relationships in the data. Therefore, an analogy could be made between the constant comparison method of GT and spatial analysis. In conclusion, GT is a fitting method for spatial thinking in conceptual space. This explicit research methodology is now subsequently applied to develop an advanced groundwater data infrastructure design theory on the premise that the various Earth Sciences disciplines' data and models need to be integrated through explicit GIScience principles.

4 CASE STUDIES

This research is fundamentally based in a New Zealand context (Figure 13). This chapter provides a description of the case studies, data sets and online data sources that were provided for the purposes of this research, or publicly available, and which were used as sources in this study for developing the groundwater infrastructure design. Three case studies served as local (Horowhenua - Section 4.1), regional (NZ Regional - Section 4.2) and national level data sources (NZ National - Section 4.2 and SAC - Section 4.3) and application targets alike. As described in the Methodology (Chapter 3.2), these case studies were used as one of the main sources of data for coding the data using a grounded theory approach (Figure 11, p. 59).

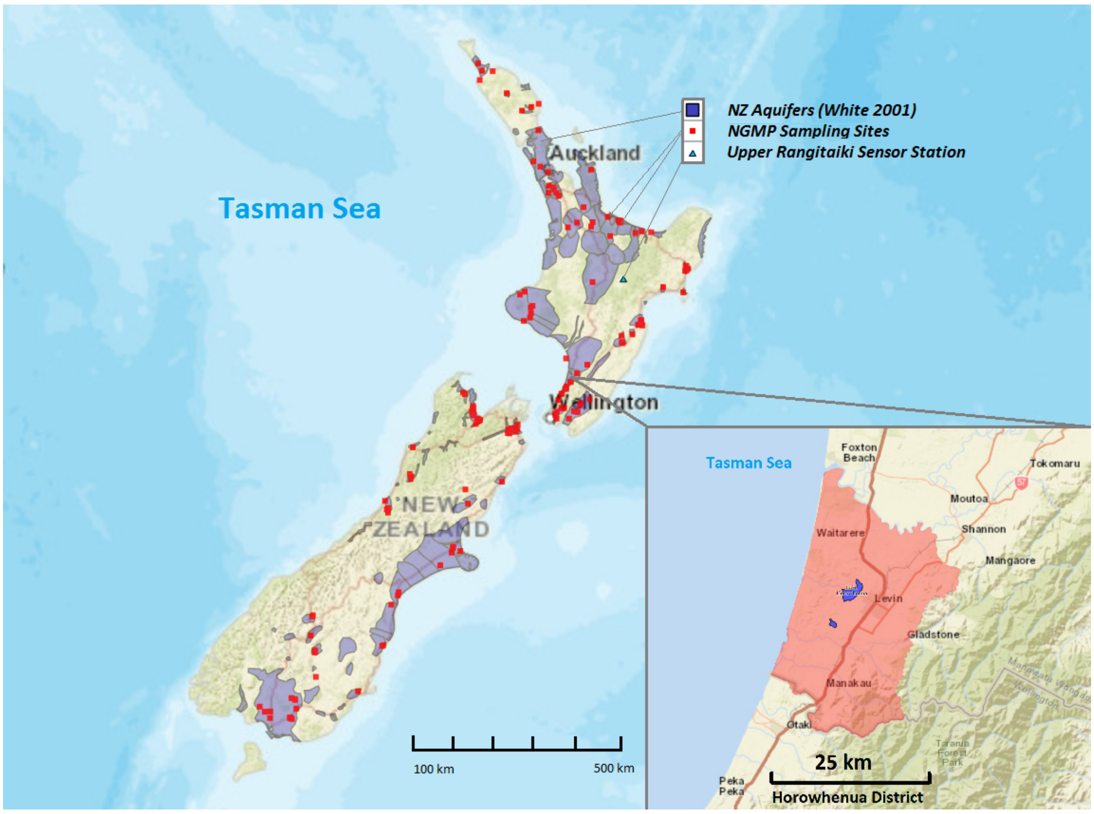


Figure 13: New Zealand, with case studies: Horowhenua with Lake Horowhenua, NGMP sampling sites, aquifer delineations (White, 2001), and the SMART Upper Rangitaiki Sensor station

The theory emerged from exploration of these data repositories and served as starting point for the design process (Figure 12, p. 61). For each case study the data sets, metadata, data access and their corresponding domain properties were analysed, abstracted and conceptualised through the GT approach as described in Chapter 3.2. Their emerging categories where sorted and integrated into a framework design in reconciliation with the state-of-the-art literature, as discussed in Chapter 5, under critical reflection through the GT methods Selective Coding and Constant Comparison (Section 3.2). The categories were then coded into tabular form as discussed in the following sections. These tables represent the condensed conceptualisation of the groundwater data domain and serve as the foundation for the framework design (Chapter 5). In Chapter 6 the implementation is then evaluated against the case studies that are described here.

4.1 HOROWHENUA LOCAL CASE STUDY AND DATA SETS

The local scale case study area is the Horowhenua municipal district located on the west coast of the lower North Island and is part of the Manawatu-Wanganui Region (Horizons, Figure 13). Lake Horowhenua is the defining natural feature in this rather flat landscape, which only towards the south east becomes slightly more elevated and hilly (Lake Horowhenua and Conceptual Model, Figure 14).

Two GNS science reports, which report on geological and hydrogeological assessments, conducted in 2010, cover geology and hydrogeology in-depth (White, Raiber, Della Pasqua, Zarour, & Meilhac, 2010; White, Raiber, Zarour, et al., 2010). The data repository for these two studies was analysed in this research and the results of this analysis are described in this section. The subsequently following tables illustrate the abstraction and conceptualisation from the data in preparation for the design process.

- Table 3 (p. 73): Data sets describing geological units/structures, elevation and surfaces
- Table 4 (p. 74): Data sets containing well descriptions, borelogs, lithology observations
- Table 5 (p. 74): Data sets about catchments and management zones
- Table 6 (p. 75): Data sets from hydrological and hydrogeological monitoring sites and related measurements

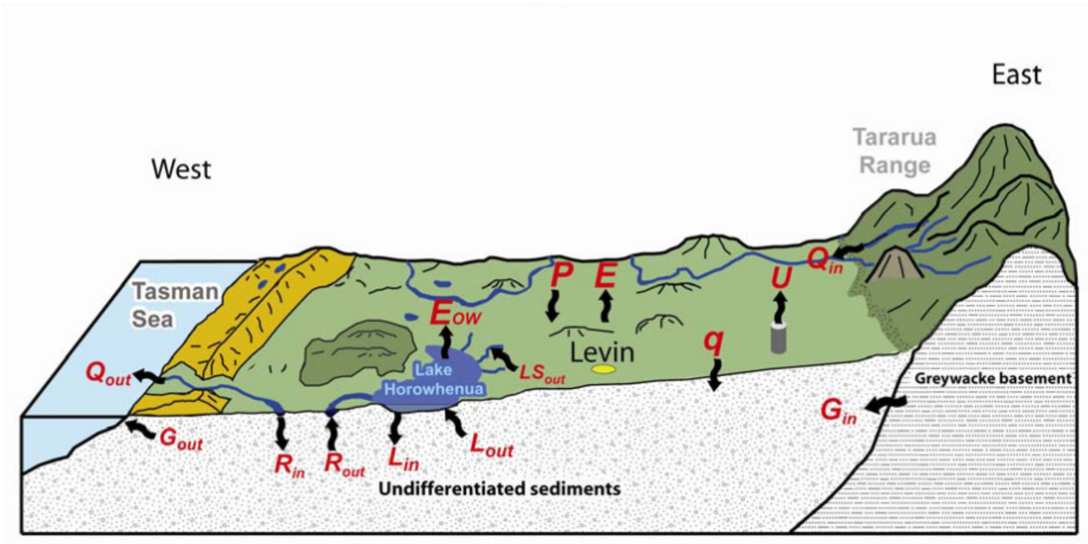


Figure 14: Conceptual depiction of water budget components in the Horowhenua area (cf. Figure 13), from East to West approx. 20km. The model is not to scale. Water inflows: P = Precipitation, Q_{in} = Surface water inflow, G_{in} = Groundwater inflow (including R_{in} and L_{in} flows seeping from rivers and lakes into groundwater), q = Groundwater recharge; Water ouflows: E = Evapotranspiration, Q_{out} = Surface water outflow, G_{out} = Groundwater outflow (including R_{out} and L_{out} flows from groundwater into rivers and lakes), U - abstraction and consumptive use. Source: White, Raiber, Zarour, et al. (2010, fig. 3)

- Table 7 (p. 76): Data sets from hydrological and hydrogeological numerical model results/outputs
- Table 8 (p. 76): Data sets solely for map visualisation and unstructured data set like reports and images

The data repository for the reports from White, Raiber, Della Pasqua, et al. (2010) and White, Raiber, Zarour, et al. (2010) consists of mainly files of different types that were either used as input data for the assessment of geology and water budget, or have been produced in the course of those studies. The data sets are related to bore logs and lithology, soil properties, climate and physical and chemical river/lake hydrology observations. The geospatial vector data is stored in shapefiles. Many shapefiles consist of polygon layers that serve to outline occurrences of geological or hydrological features, like aquifers, geological units, basins, catchments or water management zones.

Another group of shapefiles contain point layers for the geolocations of wells, springs, and generally measurement locations, and therefore identify where samples of well lithology and groundwater related measurements are located. In some cases shapefiles are used to present the characteristics or occurrences of features or properties of those in different 'depths' or along lengths (e.g. for well logs).

The actual measurement series or full sample description reside in Microsoft Excel spreadsheets (XLS) or comma separated values (CSV) files and are more thoroughly interpreted in the reports. In the reports of White, Raiber, Della Pasqua, et al. (2010) and White, Raiber, Zarour, et al. (2010) the source data sets are not directly referenced. For the current study, the reports were used to understand how the data sets in the repository were used to create products that were subsequently presented in the reports as descriptive results. However, full meta-data information was not consistently available for all data sets, neither in the reports of White, Raiber, Della Pasqua, et al. (2010) and White, Raiber, Zarour, et al. (2010), nor in adjunct descriptions. In many cases additional interpretation, inference or personal inquiry was necessary to judge lineage, data quality and extended metadata like units references, controlled vocabularies or used coordinate reference systems.

Many output data sets in the Horowhenua data repository were originally created to produce maps in ArcGIS, which were subsequently added as 2D visual aids in the reports of White, Raiber, Della Pasqua, et al. (2010) and White, Raiber, Zarour, et al. (2010). Thus, many of the shapefiles only contain small slices of information, rather than a complete representation, in order to allow for the visualisation or mapping of certain relevant aspects as a figure in the report. They do not serve as part of the common data exchange information model.

Some data sets are the results of raw sampling or measurement data, while some are preprocessed (as in data scrubbing, quality assurance) and some are the outputs of final processing steps. These post-processing outputs may have been generated by a model, algorithm or simulation, or a visualisation such as ArcGIS map production. In some cases, the shapefiles contain metadata for the presented features, i.e. information about a contact person, or a responsible authority.

The GNS assessments used grid files of geological unit surfaces, which were largely produced by exporting them from EarthVision[1]. EarthVision is a geological modelling tool that is used to model (based on well logs input, interpolation and a geologists experience) the 3D properties of and distribution of earth materials in the ground. For simplicity and data size reasons, these exports, in most cases, were created as 'top of layer' 2.5D curved planes in space, or as shapefiles delineating the occurrence of a material or geological unit via a 2D polygon for map visualisation purposes. When using such grid files and map layers it is impotant to note that the Coordinate Reference Systems employed are not always consistent. The New Zealand Map Grid (NZMG), New Zealand Transverse Mercator 2000 (NZTM2000) and World Geodetic System 1984 (WGS84) coordinate systems are widely used in New Zealand. All modern GIS software packages provide a means of reprojecting and geo-rectifying layers, but in some cases (such as data stored and circulated in XLS/CSV files) the coordinate system is implicitly assumed. This meant that the reference system was not always known. Through personal inquiry with the data curator it became clear that the NZMG system was widely used as the implicit standard reference system. At this stage it is important to state, that NZMG has been declared obsolete by Land Information New Zealand (LINZ) and its use should be replaced by the current NZTM2000 projection, but in reality NZMG is still widely used.

As part of the assessments for the Horizons Regional Council there was a 3D geological model created as shown in Figure 15. The refined data sets for this area tat were used in this 3D model include groundwater level time series from monitoring wells and four geological layers (Greywacke basement, quaternary layers Q6, Q5, Q2Q3Q4, and Holocene).

National groundwater related data collection programmes such as the New Zealand NGMP harmonise data before they are stored in a database. This

1 EarthVision, developed by Dynamic Graphics, website: http://www.dgi.com/earthvision/evmain.html

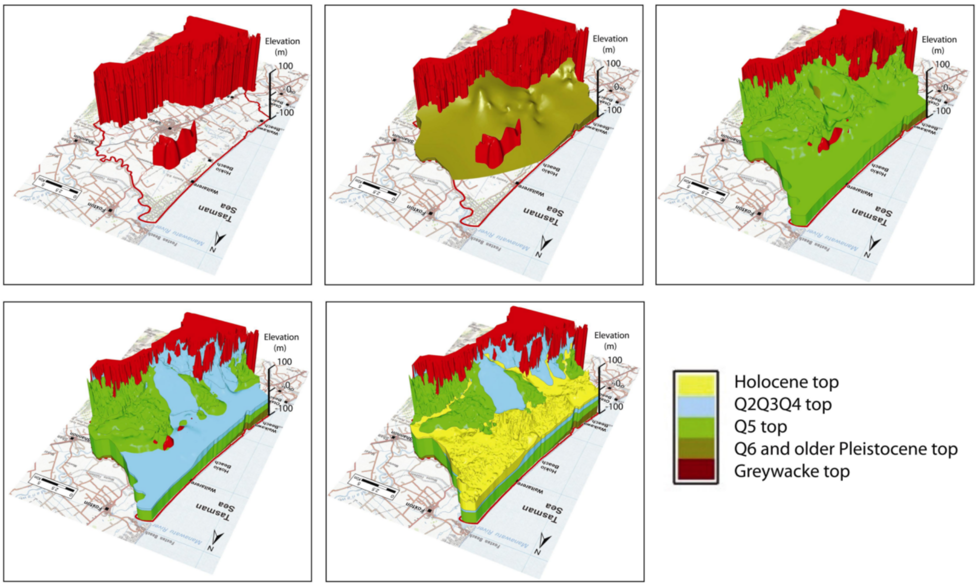


Figure 15: Geological Model of the Horowhenua area (cf. Figure 13) showing five model layers. The view is from North (lower edge) towards North (upper edge), with the coast to the Tasman Sea on the left side. North to South approx. 25km. Source: White, Raiber, Della Pasqua, et al. (2010)

data harmonisation is mainly based on the manual process after acquisition of the 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 data set comes with a specific set of attributes, representing a data provider's internal data model. Integration into the NGMP data repository is based on manual compilations.

Some procedures stipulate 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 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.

For some measurement properties, several closely related analysis procedures exist which 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.

Numbering and naming within the data sets and between different data sets is not consistent. For example, properties regarding the different data sets 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 often unclear whether the groundwater level is characterised as metres below surface, metres below well head or metres above sea level. New Zealand's available groundwater related data sets have not been comprehensively listed. Consequently, there are data sets 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 SDI. Harmonisation efforts resulting in new databases (e.g. NGMP) are labour intensive and only occur in certain time intervals. Even worse, with each copied data set for the consumer, or a transfer to another database, an outdated, non-comprehensive, non-updated, and non-maintained instance is created, and further work on these instances is not connected to the original data set.

In conclusion it can be said that the Horowhenua reports case study revealed important issues. The hydrogeological data itself from a Geoscience domain knowledge point of view seems comprehensive and appropriate to cover all required domain elements. The analysed data sets contain immediately usable data about geological units/structures, elevation and surfaces (Table 3, p. 73), well descriptions, borelogs, lithology observations (Table 4, p. 74) and information about catchments and management zones (Table 5, p. 74). Furthermore, the Horowhenua repository contained data sets from hydrological and hydrogeological monitoring sites and related measurements (Table 6, p. 75) and processed data, like from hydrological and hydrogeological numerical model results/outputs (Table 7, p. 76) and data sets which were solely used for map visualisation as well as unstructured data set like referenced reports and images (Table 8, p. 76). But in conjunction with the Horowhenua reports from White, Raiber, Della Pasqua, et al. (2010) and White, Raiber, Zarour, et al. (2010) the information from the analysed data sets was appropriate and effective to provide a comprehensive view on necessary data type requirement for the resulting framework design (Chapter 5).

The technical data aspects resulting patterns from this analysis only deliver basic information for file and folder data management practices. From that perspective this case is not sufficient as it reflects only that file-based data management is being used. However, the other two case studies, 'New Zealand Regional and National Data Sources' (Section 4.2) and 'SAC' (Section 4.3), reveal additional patterns.

Table 3: Horowhenua case study data with coding from GT, type mapping and concepts for geological units/structures, elevation and surfaces

Data source	Type / Domain	Concepts / Categories
model_boundary.shp	Shape file, extent of study area	zone boundary
Makorokio_fan.shp	Shape file, extent of geological unit	boundary, geological unit identity
Q6_extent.shp	Shape file, extent of geological unit	boundary, geological unit identity
Q2Q3Q4_extent.shp	Shape file, extent of geological unit	boundary, geological unit identity
poro_grey.xls (=poro_grey.ply)	extent of geological unit	boundary, geological unit identity
Pleistocene.shp	Shape file, extent of geological unit	boundary, geological unit identity
Greywacke_extent.shp	Shape file, extent of geological unit	boundary, geological unit identity
organic_10m_thickness.shp	Shape file, sorted bore logs	observations, borelog
organic_5m_thickness.shp	Shape file, sorted bore logs	observations, borelog
Waikawa_fan.shp	Shape file, extent of geological unit	boundary, unit identity
well_sorted_floodplain_ deposits.shp	Shape file, extent of geological unit with metadata	boundary, unit identity, geological metadata
Ohau_gravel_fan.shp	Shape file, extent of geological unit	boundary, geological unit identity
Holocene_line.shp	Shape file, extent of geological unit	boundary, geological unit identity
Holocene_base.shp	Shape file, extent of geological unit	boundary, geological unit identity
Holocene.shp	Shape file, extent of geological unit	boundary, geological unit identity
QMAP_clip.shp	Shape file, extent of geological unit with metadata	boundary, unit identity, geological metadata
Greywacke_top_250.dat	top surface of geological layer/unit	geological unit identity, gridded surface
Holocene_top_250.dat	top surface of geological layer/unit	geological unit identity, gridded surface
Q2Q3Q4_top_250.dat	top surface of geological layer/unit	geological unit identity, gridded surface
Q5_top_250.dat	top surface of geological layer/unit	geological unit identity, gridded surface
Q6_top_250.dat	top surface of geological layer/unit	geological unit identity, gridded surface
DTM250.dat	land surface elevation	elevation, gridded surface
DEMgrid.rrd	surface elevation grid	elevation, DEM
XYZ.shp	shapefile, extent of geological layer/unit	boundary, geological unit identity
Horo_contours20m.shp	Shape file, elevation as contour lines	elevation, DEM, DTM
Lake Horowhenua bathymetryedited SG1-130.xls	spreadsheet, elevations	elevations
structure_shannonAnticline .shp	Shape file, fold	boundary/line, geological structure identity
structure_otehrfaults.shp	Shape file, fault line	boundary/line, geological structure identity
structure_levinFault.shp	Shape file, fault line	boundary/line, geological structure identity
structure_levinAnticline.shp	Shape file, fold	boundary/line, geological structure identity
structure_KoputaroaSyncline .shp	Shape file, fold	boundary/line, geological structure identity
Sewell1991_sections_ all.shp	Shape file, cross section with basic information	geological structure identity, metadata
Sewell1991_sections.shp	Shape file, cross section with basic information	geological structure identity, metadata
HughesCrossSection.shp	Shape file, location of referenced cross section	geo-referenced report

Table 4: Horowhenua case study data with coding from GT, type mapping and concepts for wells, borelogs, lithology observations

Data source	Type / Domain	Concepts / Categories
Wells_1106_ withlogsinmodel.shp	Shape file, well locations with lithology observations	observations, borelog
wells_geologicalmodel_ 21_March2010.shp	Shape file, well locations	location, well identity
Wells_within_ boundary.shp	Shape file, wells with diverse collated information	well identity, well metadata, observations
Wells_withLithologicalInfo .shp	Shape file, well locations	location, well identity
Well_locations.shp	Shape file, wells with diverse collated information	well identity, well metadata, observations
DrillHoleControlSites.xls	spreadsheet, locations for drill hole IDs	borehole / well identity
map_rock_types.shp	shapefile, geological observations	observations, geological metadata
Lithology_description_ processed.xls	spreadsheet, geological observations	observations, geological metadata
Wells-with-lithology depths.xls	spreadsheet, geological observations	observations, geological metadata
gravel_pseudo_logs.xls	shapefile, geological observations	observations, geological metadata
Wells_1106_withlogs.shp	Shape file, well locations with lithology observations	observations, borelog
Sands_greater150.shp	Shape file, wells with diverse collated information	well identity, well metadata, observations
sand_greater150.xls	spreadsheets, sorted bore logs	observations, borelog

Table 5: Horowhenua case study data with coding from GT, type mapping and concepts for catchments and management zones

Data source	Type / Domain	Concepts / Categories
Study_area_new.shp	Shape file, extent of study area, catchment information	boundary, catchment / management zone feature
NZTopoMap250k	New Zealand topographic map	background map, visualisation
Groundwater_catchments_ with_names.shp	Shape file, extent of catchment and add. information	boundary, catchment / management zone feature and metadata
Groundwater catchments_ polygonsNZMG.shp	Shape file, extent of catchment and add. information	boundary, catchment / management zone feature and metadata
Groundwater catchment.shp	Shape file, extent of catchment and add. information	boundary, catchment / management zone feature and metadata
New_management_ subzones.shp	Shape file, extent of study area, catchment information	boundary, catchment / management zone feature
Kopaturoa.shp	Shape file, extent of catchment and add. information	boundary, catchment / management zone feature and metadata

Table 6: Horowhenua case study data with coding from GT, type mapping and concepts for hydrological and hydrogeological monitoring sites and related measurements

Data source	Type / Domain	Concepts / Categories
Horowhenua_springs.shp	Shape file, springs with seasonal flow characteristic	spring identity, spring metadata
Telemetered_Groundwater_ Sites.shp	Shape file, locations for observation wells	well identity, metadata
Horizons Horowhenua groundwater data 2008-11-27.xls	groundwater level observations, locations and summary statistics	observations and sampling location basic metadata
Ohau_monitoring_sites.shp	flow gaugings observations, locations and summary statistics	observations and sampling location basic metadata
export_output.dbf	database file, wells with diverse collated information, observations and aquifer properties	well identity, well metadata, observations and aquifer properties
Monitoring_sites_used_for_ water_budget.shp	flow gaugings observations, locations and summary statistics	observations and sampling location basic metadata
Water_allocations.shp	Shape file, locations of water allocations, consents for water abstractions (surface and groundwater) with metadata	consent metadata, observations
water_budget_sites.shp	flow gaugings observations, locations and summary statistics	observations and sampling location basic metadata
Consented_wells_ horowhenua.shp	Shape file, well locations with allocation metadata	observations, well identity, metadata
3720_Flow_gauging_data_ with_site_nos_coordinates.xls	flow gaugings observations, locations and summary statistics	observations and sampling location basic metadata
gaugingexport_Horowhenua.xls	flow gaugings observations, locations and summary statistics	observations and sampling location basic metadata
hydrometric_archive.xls	flow gaugings observations, locations and summary statistics	observations and sampling location basic metadata
Water allocation by catchment.xls	spreadsheet, water allocation metadata for catchments	observations, catchment identity, metadata
Rain Normals 1961-1990.xls	rainfall observations, locations and summary statistics	observations and sampling location basic metadata
Rainfall_july2009.xls	rainfall observations, locations and summary statistics	observations and sampling location basic metadata
water_quality.shp	water quality sampling locations	sampling location basic metadata
chemistry.xls	water quality observations, locations and summary statistics	observations and sampling location basic metadata
Radiocarbon.shp	Shape file, coordinates and addresses	name indicates point locations of radio carbon sampling sites
land use cover and soil units and properties		
9soils_clipped.shp	Shape file, extent of soil units with metadata and soil properties	boundary, unit identity, soil metadata
Soil_horowhenua.shp	Shape file, extent of soil units with metadata and soil properties	boundary, unit identity, soil metadata
Soil_codes_fund.dbf	soil units metadata	soil units classification and metadata
landuse-assessment-zones.shp	Shape file, extent of landuse cover units with metadata	boundary, unit identity, landuse classification
landcover-assessment-zones.xls	landuse cover metadata	landuse classification and metadata

Table 7: Horowhenua case study data with coding from GT, type mapping and concepts for hydrological and hydrogeological numerical model results/outputs

Data source	Type / Domain	Concepts / Categories
Horowhenua yields .aux/.adf for the subcatchments	numerical model yield output	gridded model output
Horowhenua_groundwater_ yields_corrected.pmf	numerical model yield output	gridded model output, groundwater yield
Horowhenua_surface_water_ yields_corrected.pmf	numerical model yield output	gridded model output, surface water yield
hor_yield_1	arc binary grid, yield, groundwater model output	yield, groundwater flow model, grid
hor_yield_2	arc binary grid, yield, groundwater model output	yield, groundwater flow model, grid
hor_yield_fwe	arc binary grid, yield, groundwater model output	yield, groundwater flow model, grid
rainfall *.adf	pre-processed from NIWA station data into arc binary raster grid	gridded model output, rainfall distribution
Surface_Water_Expert_ Interpolation.dbf	pre-processed from rainfall, soils etc. into summary database	point model output surface water flows
evapotranspiration	pre-processed from rainfall, soils etc. data into raster grid	distributions
rain_median_1978-2007_ horo.img	pre-processed from rainfall into raster grid	distributions
m4_5km_v2_filled_ NIWArunoff.asc	binary grid, rainfall run-off model output	rainfall run-off model, grid
Soil recharge August 2009_ new.xls	pre-processed from soils properties into summary spreadsheet	soil recharge potential
Rainfall recharge summary1.xls	pre-processed from rainfall, soils etc. into summary spreadsheet	rainfall recharge

Table 8: Horowhenua case study data with coding from GT, type mapping and concepts for map visualisation and unstructured data

Data source	Type / Domain	Concepts / Categories
derived data sets solely for map visualisation		
Horowhenua gwl contours.shp	shapefile, contour lines indicating flow direction	flow direction, visualisation
ContinuousFlow_ALL.shp	shapefile, arrows indicating flow direction	flow direction, visualisation
Estimated_inflow_outflow .shp	shapefile, arrows indicating flow direction	flow direction, visualisation
Rivers_Horowhenua.shp	Shape file, stream and river reaches	metadata, only to show rivers on map
Ohau_River.shp	Shape file, river reaches with gaining or loosing flow characteristic	no further metadata, only to visualise river parts on map
lakes.shp	Shape file, lakes	no metadata, only to show water bodies on map
gaining_loosing_2010.shp	Shape file, river reaches with gaining or loosing flow characteristic	metadata of how inferred
Arrows groundwater flow direction.shp	poly-lines of calculated groundwater flow direction	groundwater flow direction, visualisation
Contour_Aet_Jan1.shp	contour lines of calculated evapotranspiration distribution	evapotranspiration distribution, visualisation
rainfall_contours_pline.shp	contour lines of mean calculated rainfall distribution	mean rainfall distribution, visualisation
Mean_1978.shp	mean rainfall distribution contours	rainfall distribution, visualisation
unstructured (images, written text, reports) information used		
TIDEDA plots.doc	plots from flow archives	images, graphs
LoggedBoresfromLiterature .shp	Shape file, locations of referenced bore record	geo-referenced report
Summary of previous work in the Manawatu-Wanganui Region, including the Horowhenua.doc	report review	metadata summary available reports
NIWA rainfall report FINAL.pdf	report about rainfall characteristics	report, text
Bibliography - Manawatu-Wanganui Region.doc	review available literature	metadata summary available literature / journal articles
PETLAB Database maps S24 S25 R25.doc	maps indicating locations of rock specimen samples	images, maps
Rainfall recharge summary.doc	report	summary reports

4.2 NEW ZEALAND REGIONAL AND NATIONAL DATA SOURCES

Many groundwater related data sets are scattered around the regional and unitary councils, the Crown Research Institutes and universities, as well as commercial companies and NGOs (e.g. consultancies, energy sector, utilities). The different available data sets are partly announced on the websites of the organisations, through web forms or Excel sheets, or need to be personally requested. Data exchange is presently based on personal communication, resulting in usage negotiations and physical data exchange workloads. Data exchange is primarily based on file copies of the requested data repository. Increasingly, data access is improving in New Zealand, and this was notable throughout the course of this research.

Prominent examples of improved data access are the now publically available New Zealand geospatial data catalogues and data portals such as LINZ, Land Resource Information Systems (LRIS), NIWA Environmental Information Browser (EIB) and Department of Conservation (DoC) Geoportal, Ministry for the Environment (MfE) Data Service, GNS data sets catalogue. Additionally, the Land, Air, Water, Aotearoa (LAWA) initiative's push to display water quality indicators from federated council databases via WaterML 2.0 has also increased the availability of data. 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. 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 data sets for New Zealand, accessible through OGC CSW, WMS, and WFS. New Zealand adopted several of these standards and 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)
- Unified Modeling Language (UML) standardised geographic information and service models ISO 19103 (2005)
- Australia New Zealand Land Information Council (ANZLIC) Metadata Profile v1.1 ISO 19115 (2003)

Landcare Research publishes a considerable number of soil and biodiversity-related maps with a similar solution (LRIS). The LINZ, LRIS and MfE data por-

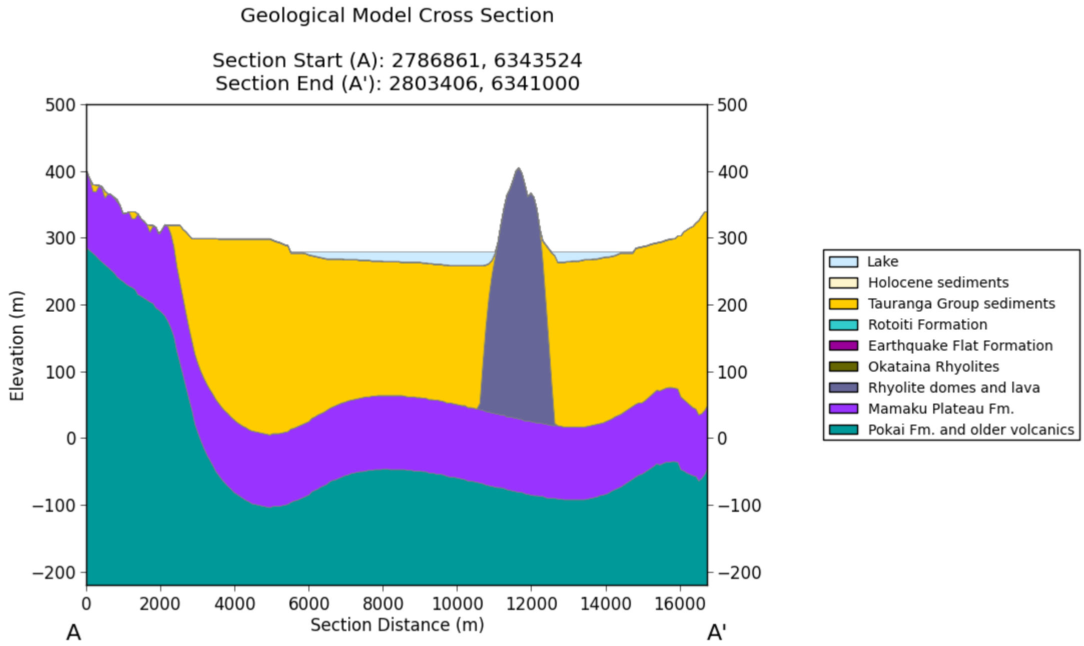


Figure 16: Geological Cross Section across Lake Rotorua from West to East (from left to right). Source: GNS Earth Beneath Our Feet (EBOF) website (http://data.gns.cri.nz/ebof/)

tals are backed by a product of the New Zealand based company Koordinates, which also hosts publicly accessible data sets from some New Zealand regional and city councils (e.g. Greater Wellington Regional Council), which beside the standard interfaces CSW, WMS and WFS also provides a custom API for data access. NIWA has built a publicly accessible CSW catalogue system and continually registers their data sets. However, access to most data sets required for hydrogeological assessments is still conducted through multiple manual search processes instead of interoperable web services. GNS publishes the New Zealand Geological Map, not only in print and electronic copy, but also through GeoSciML/WFS and WMS services. GNS also provides online access to a wealth of other data sets, but in many cases encapsulated in proprietary web applications. The data for NGMP is hosted in the GGW at GNS, but the measurements are only accessible through a web form, and not as a web service.

GNS Science is developing 3D hydrogeological 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 in terms of the originating data source. Figure 16 shows a resulting image of a cross-section retrieved from the web viewer application via a HTTP request.

Publicly accessible service using the OGC SOS and the O&M or WaterML 2.0 encoding specification for real time telemetering and public access of measurements in New Zealand are becoming more pervasive, e.g. LAWA. With the developments of the Hilltop Data Tamer software in conjunction with the rapidly improving LAWA water quality data portal it seems that the barriers to adoption are decreasing.

The approach to data conceptualisation for the case studies was shown in Tables 3 to 8. In this section's analysis about New Zealand's emerging spatial data and environmental information services and related technical data aspects, the results reveal that there is a common, recurring infrastructure pattern. It is becoming apparent that there is momentum to provide data sets and metadata about existing data sets through automated internet services, that can be accessed through user-friendly websites and standardised API at the same time. However, in summary for this case study it can be said that the establishment of a comprehensive list of available groundwater related resources across New Zealand is not yet possible in an automated fashion.

The following list enumerates the coding from GT for New Zealand regional and national data sources. The concepts and their respective type mapping are described with i) Data source, ii) Type, Domain, iii) Concepts / Categories:

1. LINZ Data Service[2]
   - i CSW, WFS, WMS, metadata, file download
   - ii data repository for public access and download of data, catalogue, general and govt geographic features, hydrography, topography, topology, public transport, gazetteer, cadastre, aerial imagery
   - iii catalogue & metadata records, CSW & ISO metadata and XML encoding, WMS visualisation, simple features WFS and file access
2. Landcare LRIS data portal[3]
   - i CSW, WFS, WMS, metadata, file download
   - ii data repository for public access and download of data, catalogue, soil, land use and cover, biodiversity, conservation areas
   - iii catalogue & metadata records, CSW & ISO metadata and XML encoding, WMS visualisation, simple features WFS and file access
3. NIWA Environmental Information Browser (EIB)[4]

2 LINZ Data Service: https://data.linz.govt.nz
3 Landcare LRIS data portal: https://lris.scinfo.org.nz
4 NIWA EIB: http://ei.niwa.co.nz/

i CSW, metadata, links or contact information

ii data repository for public access to metadata, biodiversity, marine and ocean science catalogue, hydrometric stations catalogue

iii catalogue & metadata records, CSW & ISO metadata and XML encoding

4. LINZ public geodata catalogue[5]

i CSW, metadata, links or direct data download

ii data repository for public access of upload and download, catalogue, free upload

iii catalogue & metadata records, CSW & ISO metadata and XML encoding

5. DoC Geoportal[6]

i ESRI Geoportal, with DoC public data, CSW, WFS, WMS, metadata, shapefile download

ii data repository for public access and download of data, catalogue, camping sites, reserves, hiking tracks, conservation areas

iii catalogue & metadata records, CSW & ISO metadata and XML encoding, WMS visualisation, simple features WFS and shapefile access

6. NIWA CLIDB SOS Service (Alexander Kmoch, Jochen Schmidt and Andrew Watkins)

i SOS, WaterML 2.0, O&M 2.0, SensorML, CLIDB climate data time series service as alternative interface for CLIFLO[7]

ii middleware data service for access to station metadata, hydro-climate time series data

iii SOS, WaterML, O&M 2.0, SensorML, hydro-climate time series

7. NIWA Hydro and Flows archive Kisters SOS (Jochen Schmidt and Andrew Watkins)

i Kisters software, SOS, WaterML 2.0, O&M 2.0, SensorML, TIDEDA Hydro and Flows archive, time series data service as API interface for old TIDEDA archives

5 LINZ public geodata catalogue: http://geodata.govt.nz
6 DoC Geoportal: http://geoportal.doc.govt.nz/geoportal
7 NIWA CLIFLO http://clifo.niwa.co.nz

ii middleware data service for access to station metadata, hydro-climate time series data

iii SOS, WaterML 2.0, O&M 2.0, SensorML, hydro-climate time series

8. Waikato Regional Council hydro Kisters SOS (John Hadfield and Bevan Jenkins)

    i Kisters software, SOS, WaterML 2.0, O&M 2.0, SensorML, time series data service to manage Waikato Regional Councils hydro-climate data, weather and river flow and telemetered water pumping stations, groundwater levels

    ii middleware data service for access to station metadata, hydro-climate time series data

    iii SOS, WaterML 2.0, O&M 2.0, SensorML, hydro-climate time series, groundwater levels

9. Horizons Regional Council Hilltop SOS Testbed (Brent Watson and Sean Hodges)

    i Hilltop software, WFS, SOS, WaterML 2.0, time series data service to manage Horizons Regional Councils hydro-climate data, weather and river flow and telemetered water pumping stations, groundwater levels

    ii middleware data service for access to station metadata, hydro-climate time series data

    iii SOS, WaterML 2.0, simple features WFS and file access, hydro-climate time series, groundwater levels

10. GNS Geoserver with Geological Map

    i Geoserver software, publishing a variety of geological data sets used in GNS web maps

    ii middleware data service for access to geological features and metadata, and web map visualisation

    iii simple features WFS and file access, WFS/GeoSciML and WMS

11. GNS Earth Beneath Our Feet (EBOF)

    i 'Three-dimensional geological models have been developed at GNS Sciences over the past 10 years with the long-term aims of understanding the geometry, fluid flows and layer properties of geological units in the Taupo Volcanic Zone'[8]. This website allows to enquire vertical profiles and cross-sections.

ii middleware data service for access to geological models with an image-based visualisation

iii cross-sections and virtual borelog images, HTTP API

12. LAWA

 i LAWA has been established by like-minded organisations with a view to helping local communities find the balance between using natural resources and maintaining their quality and availability. Initially a collaboration between New Zealand's 16 regional and unitary councils, LAWA is now a partnership between the councils, Cawthron Institute, Ministry for the Environment and Massey University and has been supported by the Tindall Foundation'[9]. This website allows to enquire a variety of water quality parameters, measurements and derived indicators.

 ii website for presentation of predominantly water quality data accessible through web maps and tables.

 iii water quality, web map

13. Journal of Hydrology (New Zealand)

 i The Journal of Hydrology (New Zealand) is published by the New Zealand Hydrological Society with the mission 'to further the science of hydrology and its application to the understanding and management of New Zealand's water resources' [10]. This website allows to enquire all abstracts up to the latest Volume 54 (2015) and papers for free viewing up to Volume 46 as at 12 October 2011.

 ii website for discovery of journal articles, with a free text search over title, authors and abstract, no keywords, articles are available in PDF format

 iii journal articles, unstructured data, reports

8 GNS Earth Beneath Our Feet (EBOF): http://data.gns.cri.nz/ebof/

9 Land Air Water Aotearoa (LAWA): http://www.lawa.org.nz/

10 The Journal of Hydrology (New Zealand): http://www.hydrologynz.org.nz/index.php/nzhs-publications/nzhs-journal

4.3 SMART AQUIFER CHARACTERISATION PROGRAMME (SAC)

The research undertaken for this thesis is aligned with SAC (Klug et al., 2011). SAC is a collaborative NZ-EU research programme focused on the characterisation of New Zealand's aquifers and is led by GNS. The project aims to develop a suite of mostly novel methods which will be used to passively collect data from in situ and remote sensors and satellites to infer new insights about aquifer properties and groundwater quantities. The main stakeholders are the regional and unitary councils of New Zealand as well governmental bodies like LINZ, MfE and other research institutions such as NIWA and Landcare Research. The programme has clearly defined research aims, i.e. sub projects, shown in Figure 17. The data sets and services collected and produced throughout the course of the SAC programme are also considered in this research, because this dissertation research is embedded in the overarching SAC research aim of 'Data Synthesis and Visualisation' (sample data sets available at time of this study are shown in the following section).

This case study, the SAC case study, reveals similar patterns to the afore mentioned Horowhenua case study, i.e. file-based data sets without a ubiquitous underlying domain model, but rich in specific local domain knowledge and partial online access of data. An additional advantage of this case study is the availability of an online data storage link for sensor-based live-telemetry.

SMART case study data with coding from GT, type mapping and concepts, with i) Data source, ii) Type, Domain, iii) Concepts / Categories:

1. Equilibrium Water Table (Westerhoff et al.)
   - i model output water table, advanced mean climate water table
   - ii processed from rainfall, recharge, soil and geological input data into raster grid
   - iii gridded surface, distributions
2. NZ Aquifers (White 2011)
   - i exported data sets from the book 'Groundwaters of New Zealand' (Rosen & White, 2001)
   - ii Shape file, extent of geological unit
   - iii boundary, hydrogeological unit identity, basic unit metadata
3. NZ Groundwater Information (Lovett & Cameron, 2014)
   - i extended aquifers and management zones reviewed

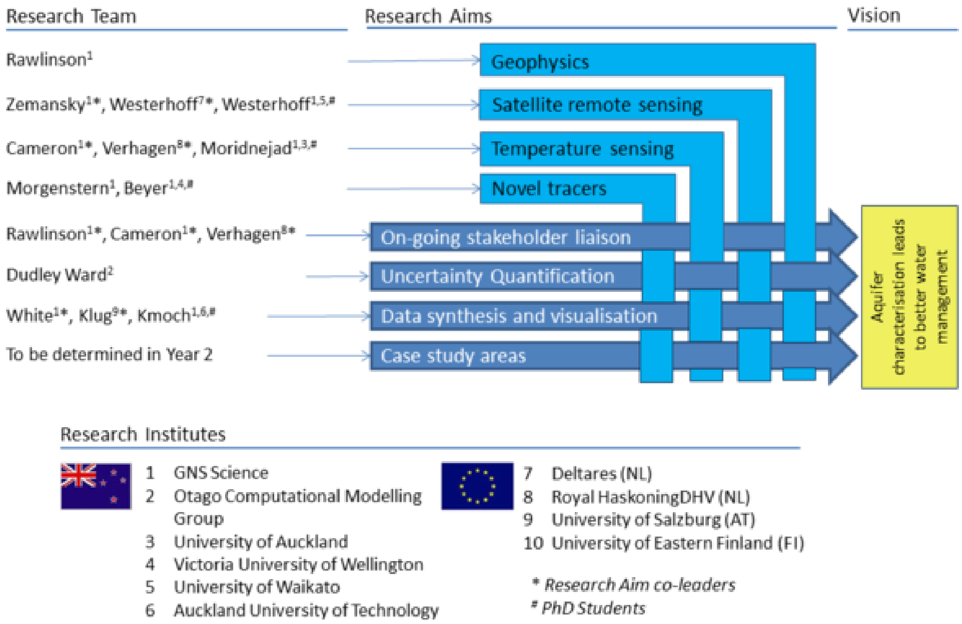


Figure 17: New Zealand-EU collaborative SMART Aquifer Characterisation Programme (SAC), source: SMART (2016) website

- ii Shape file, extent of geological unit, catchment information
- iii boundary, hydrogeological unit identity, boundary, catchment / management zone feature, basic unit metadata

4. Halon Sampling
   - i sampling locations of water samples for Halon analysis and age tracer validation
   - ii observations, locations and summary statistics
   - iii observations and sampling location basic metadata

5. NGMP SOS Service (Alexander Kmoch)
   - i SOS, WaterML 2.0, O&M 2.0, SensorML, NGMP groundwater chemistry and physical properties, data time series service as alternative interface for GGW[11]
   - ii middleware data service for access to feature/sampling location metadata, groundwater chemistry, aquifer physical properties, observations

11 GNS GGW database: http://ggw.gns.cri.nz/ggwdata

iii SOS, WaterML 2.0, O&M 2.0, SensorML, groundwater observations/measurements, sampling locations, aquifer properties

6. SMART Upper Rangitaiki Sensor station SOS server (Alexander Kmoch, Hermann Klug, Steffen Reichel)

    i SOS, WaterML 2.0, O&M 2.0, SensorML, hydro-climate demo weather station, time series, open source design

    ii middleware data service for access to the field data logger, observations

    iii SOS, WaterML 2.0, O&M 2.0, SensorML, groundwater and hydro-climate observations/measurements

Part III

RESULTS

5 A GROUNDED DESIGN THEORY FOR A HYDROGEOLOGY INFRASTRUCTURE

This chapter presents the theory and design developed for an interoperable distributed groundwater infrastructure for New Zealand by application of the 'Grounded Design' method - hence the resulting Grounded Design Theory. The following section is comprised of the presentation and discussion of the results of the GT analysis of the case study areas and the impact on the DSR framework development of a 'Hydrogeology Infrastructure' in the context of this thesis.

I decided to call this framework a 'Hydrogeology Infrastructure' to encompass the fact that this type of system projects beyond the traditional notion of SDI. While this system employs the capabilities of a SDI for Hydrogeology in particular regarding the integration of data sources, its integrated modelling and visualisation components make a more generalised 'Spatial Computing Infrastructure'. Consequently, this means that either the notion 'Spatial Data Infrastructure' needs to be redefined or that the system, which is a novel conception, requires its own terminology or name. It is for this reason that the term/name 'Hydrogeology Infrastructure' was coined.

The Hydrogeology Infrastructure design in this chapter describes which methodologies from GIScience, technical interfaces and formats, and implementation standards can be integrated to create a new holistic data enriched perspective on New Zealand's groundwater resources, without designing an isolated data framework. Figure 18 shows the interfaces and components of the infrastructure design in terms of a workflow. These components, their sub-components and modules will be described in the following sections together with their immediate and indirect interaction patterns from a conceptual perspective. This has been refined from the current state of practice of hydrogeology and geoscience data management in New Zealand as well as through several iterations of literature review, reflection and circumscription. The subsequent chapter (Chapter 6) will discuss the actual software artefacts that were generated in this study and which represent a real world implementation of the hydrogeology infrastructure presented in this chapter. For each section, the specific infrastructure components and their relevance to the

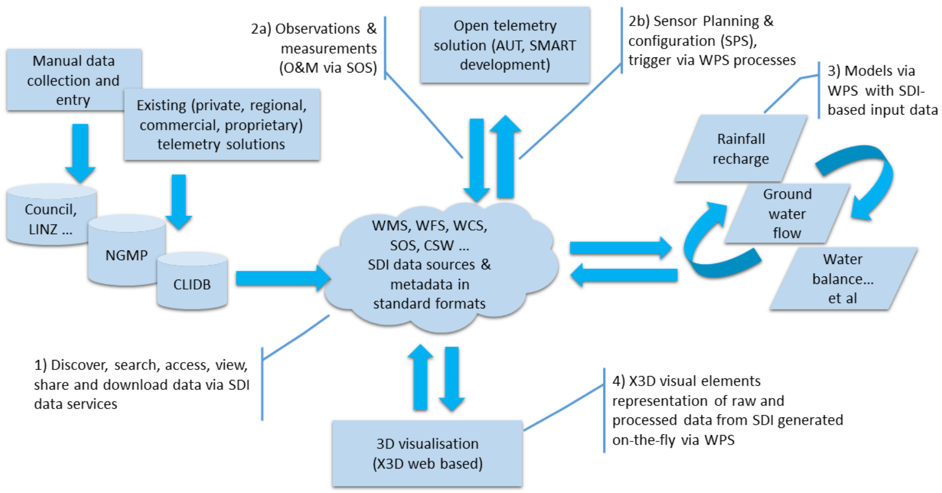


Figure 18: A schematic of the workflow, and how the data sources, geoprocessing and visualisation functions are connected via the Hydrogeology Infrastructure

overall work is presented in detail in relation to the objectives and research questions (Section 1.3) of this research.

5.1 DATA SOURCES

The following section explains how the framework design and consequently its implementation were derived from the case studies source data, and how exemplary transformation from source data in their current formats into the described reference data formats can be achieved. For the reader it is important to distinguish how the data was sourced (case studies, Chapter 4), and how the framework deals with it s a technical transformation (without changing the domain data substance) that makes the source data usable by the framework. This transformative process and hence the framework, aims to improve overall data access and usage for groundwater-related data in New Zealand.

SDI as a foundational concept for an interoperable and harmonised data exchange framework is well described in various contexts (Section 2.2 of the literature review). SDIs are implemented on many scales for a variety of purposes. The status quo and current challenges (i.e. gaps) in hydrological data management and modelling in New Zealand were reviewed in Section 2.1 of the Literature Review. To answer Research Question 1, existing SDI methods will be examined in relation to New Zealand and to the domain of hydrogeology and will be tested on the case studies (Chapter 4).

To examine, if and how the hydrogeological data sets can be transformed into standardised and widely accepted technical formats, the core concepts 'Interoperability' and 'Harmonisation' need to be clarified at first. Subsequently, viable encoding and interface standards are discussed and selected for the framework in reconciliation with the Case Studies' GT data coding categories.

These data sets comprise digital information which is based on an 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. These atoms of information modelling are used to develop domain specific data schemata. The creation of such schema is a task for data experts, who are familiar with the semantics and characteristics of the target domain and therefore can identify how these concepts need to be defined in the target schema.

An important tool used for conceptually modelling the target domain is UML, a general-purpose modelling language in software engineering. UML models can then be encoded into an XML Schema Definition (XSD) document following the ISO 19118 encoding guidelines. Such XSD documents define the structure and allowed elements of a XML application schema and are used for the assessment and validation of the syntactic compliance of the XML instance documents. But, XML is only one form of a physical data encoding format. Another increasingly adopted encoding is JSON.

Internationally, OGC, ISO, W3C, Organization for the Advancement of Structured Information Standards (OASIS), Web3D and other organisations provide standards for distributed services for data transformation, discovery, view and best practice for download. According to the W3C, web services are self-contained and self-describing application components using XML and HTTP to facilitate communication with other applications using open protocols.

International groundwater and IT experts are currently collaborating on initiatives intended to make groundwater information accessible in a standardised digital formats. In the international OGC GW2IE a GML application schema, called GWML2, for groundwater data is being developed for international groundwater data interoperability.

Data Harmonisation is considered to be an agreed procedure where two individual parts are brought into same structure and semantics are merged into a beneficial aggregate, which is the mapping into a target schema. Harmonisation is required to enable interoperability. To ensure interoperability on national or regional level, data that exists in a variety of formats and different schemata need to be mapped to an agreed on, standardised target schema. The emphasise here lies on the inherent ability to transform data losslessly between schemata. Not all data needs to exist in the interoperable exchange schema.

In New Zealand this process is stimulated in particular for environmental data. There is a targeted progress in New Zealand in terms of making hydrological measurements available via standards through initiatives as State of the Environment (SOE), NEMS and EODP. However, these initiatives only refer to a subset of groundwater related data sets. For the harmonisation of spatial groundwater related data sets across New Zealand a common core data structure (target schema) for groundwater related data sets would need to be established. In comparison with the levels of interoperability from Manso et al. (2009) and Sheth (1999) the aim of such a target schema would be to achieve technical, syntactic and semantic interoperability between the diverse range of case study data sets. Technical interoperability is established when lower-level computer systems guarantee to adhere to standards in order to be able to communicate with each other reliably and to transfer and exchange data. Syntactic interoperability defines that the format of the data that is transferred between different systems is in compliance with agreed on spatial data standards. Semantic interoperability relies on and builds on top of syntactic interoperability. It prescribes that the meaning and interpretation of the transferred information is accurately preserved.

For this study the process of data harmonisation for spatial groundwater related data sets would need to integrate different types of data - on a technical level (i.e. a computer-based data structure), syntactic (i.e. data formats are compatible), and semantic (i.e. naming and meaning of fields can be directly translated into each data format) - into the target schema as described in the Case Studies (Chapter 4):

- Horowhenua (Section 4.1, p. 73), Table 3: Data sets describing geological units/structures, elevation and surfaces
- Horowhenua (Section 4.1, p. 74), Table 4: Data sets containing well descriptions, borelogs, lithology observations
- Horowhenua (Section 4.1, p. 74), Table 5: Data sets about catchments and management zones

- Horowhenua (Section 4.1, p. 75), Table 6: Data sets from hydrological and hydrogeological monitoring sites and related measurements
- Horowhenua (Section 4.1, p. 76), Table 7: Data sets from hydrological and hydrogeological numerical model results/outputs
- Horowhenua (Section 4.1, p. 76), Table 8: Data sets solely for map visualisation and unstructured data set like reports and images
- Exemplary New Zealand Environmental Information Services (Section 4.2, p. 79)
- SAC Programme (Section 4.3, p. 83)

From the analysis of the case studies data it is apparent, with a few exceptions like GGW/NGMP, the geological map of New Zealand, soil data and hydrological time series data; that there is no real source data schema for many of the commonly used data sets. Furthermore, the breadth and variety of data sets needed for groundwater resources assessment and modelling makes it illogical to try to capture them in a single data model. However, the GT approach employed over the case studies revealed that there are general patterns of data formats (encodings) for categories of domain specific data. Thus, the approach towards data harmonisation and interoperability chosen in this study is not a single explicit semantic transformation of source data into a single target schema. Instead as few as possible, but as many as necessary existing data encodings (formats) with related data schemas were selected to expose data in an interoperable fashion. Table 9 (p. 101) lists the selection results. The general categories from the case studies are related to a common data standard. Furthermore, those selected data formats have preferred modes of data transfer and exchange, which are standard interfaces.

The process of exposing data sets through such target schemas and interfaces still requires the mapping of parameters, properties, types, and coding of the parameter values and a minimum set of metadata elements that comprise auxiliary information of the respective data source for meaningful data exchange procedures. This was done in this study for the two nationally significant databases NGMP and CLIDB, as well as on an exemplary selection of wells, bores and aquifer data sets from the case studies.

Through the classification from the case studies and in reconciliation with the review from the literature the mapping Table 9 (p. 101) was created, to define by which means data sources should deliver data in the groundwater SDI.

For each generic data type domain the table attributes the concept category with a 'Target Design Element' for the SDI.

In the following subsections the selected standards will be described regarding their capability to achieve certain levels of interoperability for the different and disparate data sources, models and visualisation processes in a groundwater SDI. Chapter 6 explains the real world application of the developed design and the technical details of the implementations that were produced in the course of this study.

5.1.1 *Services and Encoding Standards for Data Sources*

The main OGC compliant web services specifications for interoperable data exchange are the WFS for vector and discrete feature data access, WCS for continuous coverage and raster data access, WMS for rendered image data access for web maps from vector and raster data sources, and the CSW for publishing and searching collections of descriptive information (metadata) about geospatial data and services. Following the model of Manso et al. (2009) then these standardised interface and data exchange descriptions allow for all levels of interoperability. Within the described web services orchestration framework for this study, data providers offer their data using compatible and standardised (OGC) web services. 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. The specific linkage of observation data via Sensor Observation Service (SOS) into the SDI is described in more detail in Section 5.1.2.

Web Feature Service (WFS)

WFS services store and deliver discrete vector features and associated attribute data. The actual resource is a FeatureType, a layer of discrete geometries with attributes. The rendering/visualisation of discrete vector features is independent from their 'identity', styling can be applied outside of the WFS service based on the features' attributes for example. In OGC GML there are simple and complex features, where simple features have a `FeatureType` or schema that only consists of simple data types (i.e. XML literals, numbers, text, geometry), whereas complex features have elements and properties that are `FeatureType` themselves. All higher order GML application schemas are complex features.

Most importantly, feature types in GML application schemas like GWML1 and upcoming GWML2, GeoSciML3 or GeoSciML4, O&M 1.0/2.0 and WaterML 2.0 are complex feature application schemas. Additionally, they capture domain logic in their type and thus immediately improve logical and partially semantic interoperability. WFS service versions 1.0.0/1.1.0 are widely in use for simple features, recent developments promote WFS 2.0.0 particularly for better support of GML 3.2.1 which most current complex feature types / application schemas are built upon now.

In the New Zealand case study (Section 4.2), several New Zealand governmental agencies CRIs provided feature type access to their data sets, as well as means of searching and discovering metadata about these data sets. LINZ Data Service, Landcare LRIS data portal, partially NIWA Environmental Information Browser (EIB), DoC Geoportal, GNS Geological Map and MfE Data Service contributed to the 'New Zealand Environmental Information Services', which showed all the basic characteristics of an SDI as explicated in Section 2.2.

Table 9 (p. 101) lists which categories of data should be served through the WFS interface to support the groundwater SDI: Catchments and management zones (case studies: Table 5); well descriptions, borelogs, lithology observations (case studies: Table 4); geological units/structures (case studies: Table 3); and hydrological and hydrogeological monitoring sites (Case studies: Table 6).

Web Coverage Service (WCS)

A WCS service can store and deliver spatial, temporal and spatio-temporal coverages predominantly understood as (but not limited to) raster data, or regular grids, DEM/Digital Terrain Model (DTM), continuous spatial property distributions in formats like GeoTIFF / ArcGRID, but also stacked spatio-temporal grids of climate properties, which are often encoded in NetCDF. The addressable resource is a coverage and WCS supports multi-dimensional coverages of spatial, temporal and additional dimensionalities, depending on the number of different properties. Examples are physical property distributions like groundwater level and air temperature in space and time, as well as distributions and transport of chemical constituents. A slice (based on a time frame; or one of the chemical/physical properties; or for delimited spatial context) or other combinations are possible to retrieve through queries through the WCS interface. More complex queries are supported via the SQL like query language Web Coverage Processing Service (WCPS) extension which enables more selective slicing from often very large data sets, where a download of the full data set over the internet would want to be avoided. The result is a coverage again, the addressable resource is a combination of an addressable coverage and a 'select/process' request on that coverage to extract a slice.

Table 9 (p. 101) also lists which categories of data from the case studies can be handled through the WCS interface. The most relevant application of WCS for hydrogeology are elevation grids and geological surfaces (case studies: Table 3); water tables and numerical model outputs for rain fall distributions from monitoring sites (case studies: Table 6) and computed numerical model results/outputs like rain fall recharge grids (case studies: Table 7)

Web Map Service (WMS)

WMS is an interface standard for cartographic web maps and delivers pictures for specified layers. A WMS service uses spatial data layers as a resource, applies a style and renders standard image formats. Version as in OGC 1.3.0 is a desired baseline (WMS 1.3: AS/NZS/ISO 19128:2005), but 1.0.0 and 1.1.0 can also be widely found. WMS mainly serves as a visualisation / portrayal of feature or coverage data sets (layers) in the background. Through the `GetFeatureInfo` operation, the WMS service can also serve basic information from the underlying layer, i.e. feature or coverage. WMS serves as an easy visual orientation in the groundwater SDI. In the New Zealand Environmental Information Services case study (Section 4.2), LINZ Data Service, Landcare LRIS data portal, partially NIWA Environmental Information Browser (EIB), Department of Conservation (DoC) Geoportal, GNS Geological Map and GNS Earth Beneath Our Feet (EBOF) use WMS with underlying simple features to visualise available data sets on a map, as well as display the actual data meaningfully on a web map.

A variation of WMS is Web Map Tile Service (WMTS): While being a map image providing service like WMS, the distinction is that the map tiles are pre-rendered (for different zoom-levels and resolutions). Whereas in WMS layers still make sense as separate resources that either get rendered individually and combined in the client (transparency in the images provided), a WMTS is one resource in one service as all the images from that service address make up the 'layer' resource. WMTS are primarily used as so called base maps. In the case studies, LINZ Data Service provides the 'TOPO50' base topographic map of New Zealand, which is used as underlying base map in the visualisations of spatial data sets in the two Horowhenua GNS Science reports.

Data sets from hydrological and hydrogeological numerical model results/outputs (case studies: Table 7); and elevation and surfaces (case studies: Table 3) can be rendered by a WMS, displayed on a web map and inquired with the `GetFeatureInfo` operation.

A limitation of most WMS implementations is that they only handle simple feature data sets. Consequently, WMS will not immediately render the complex feature types from O&M, GWML or GeoSciML from Table 9 (p. 101). There are two solutions: The first one is to generate/provide a simple features representation of the complex feature type. This is reproducing the case study data sets from Table 8, which were solely created to allow for map visualisations for inclusion into reports. The GeoSciML community is maintaining a simple features class for map portrayal of geological units, which is used for the GNS Geological Web Map. This simple feature was driven by a user desire for a more compact structure, though not due to the restriction of current WMS implementations. This has not been done for GWML, and it is not in the focus of this study to develop a specialised simple features representation for spatial feature classes in GWML, which could immediately be mapped through the WMS interface.

The second solution is to recognise visualisation as a process, a function that is applied to the data to produce corresponding visual elements, which is in fact the underlying concept in WMS implementations. While this by itself not a new concept, the innovative contribution towards a standardised, declarative and generalisable process for 3D visualisation through this study is described in Section 5.4.

The mentioned services are applications that provide data access. Each instance needs to be accessed and queried independently to inquire data points, feature or gridded/raster data sets, or slices thereof. 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 encoded in XML, and XML schemata describe and constrain structure and content types of such XML encoded documents. The above mentioned specifications permit developments in a SOA. However, ESB implementations regarding New Zealand geodata processing are not commonly used. Instead, deferred processing is encapsulated in WPS processes, or custom scientific codes. The designated processing scheme is explained in more detail later in this chapter in Section 5.3 and Section 5.5.

5.1.2 *Sensor Web Integration*

To understand the hydrological processes on and across different scales, the state of the environment and the groundwater dynamics, real-time information about the environment, climate and landscape is required. Moreover, resource management decisions, such as allowed water allocation for an aquifer or limits for nutrient loads, should be based on the latest information in order to avoid irrecoverable damages due to poorly informed decisions.

Automated telemetry for physical hydro-climate parameters like rainfall and water-levels is broadly available. However, it causes interoperability challenges when proprietary non-standard compliant techniques are used for in situ measurements, data storage and data transmission routines. Subsequently, when data needs to be loaded from the proprietary data collection systems into processing and analysis systems, it often needs to be preprocessed and converted into a compatible format. Consequently, instances of data set are created, which are hard to update and maintain. A download via a website, also interchangeably referred to as web interface, does not provide the same benefits as compared to a programmatic web service access via an API with a standardised interface. Through the interface definition a software program can access the web service like a database,where the data is loaded from the database into the program and can immediately be put to use. An access through a website or web interface is a manual process that is conducted by a human and the download comprises typically a file-based archive which subsequently needs to be loaded into the software program to put it to use. This is often further aggravated by the fact that the file archive is not immediately digestible by the software program, because the contained data is not self-describing and needs to be manually adapted so the software program understands the meaning of the included data.

Thus, for the integration into a broader, inter-agency analytics and reporting infrastructure it is to be ensured that the distribution and accessibility of measurements occurs in publicly accessible and well described standard compliant ways. 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, CLIDB and regional councils' river flow and hydro-climate data sets could be accessible in near real-time based through automated procedures. Decision support could be based on the latest measurements. The resulting statistics could be provided to users at any time. The labour workload would be decreased by automated data pre-and post-processing. With the integration of WPS, statistics and derived environmental indicators could provide near-real time decision making support once the telemetered sensor data are becoming available and accessible (more details in Section 5.3).

In 2014 a working group was formed to formalise the emerging standards for hydro-climate and general observation data federation in New Zealand and develop the New Zealand EODP standardisation document, as a technical contribution to the MfE-endorsed NEMS. This profile refines existing information and web service specifications published by the OGC and vocabulary data standards defined by the W3C. The only scenario supported directly by this document is data discovery where the services and their responses act as indexes, providing the minimum required information to locate observation data for a phenomenon in space and time. Here also the integration of basic SDI principles and near-real-time sensor web approaches converged into a commonly integrated framework.

Table 6 outlines hydrological and hydrogeological monitoring sites and related measurements from the Horowhenua case study as important categories of data that need to be incorporated in a groundwater SDI.

The New Zealand Regional and National Data Sources case study (Section 4.2), and in further reconciliation with the literature review (Hydrogeology New Zealand, Section 2.1, Table 1), shows that such data sets from hydrological and hydrogeological monitoring sites and their related measurements are managed in a variety of software systems across agencies. For example, the NIWA CLIDB hydro-climate time series and station data can be accessed via the CLIFLO web interface[1].

NIWA is maintaining the TIDEDA Hydro and Flows archive Hydro and Flows archive and is implementing a Kisters Wiski solution to serve the TIDEDA archive through Kisters SOS interface, which supports WaterML 2.0 and O&M 2.0 as time series data encoding. Also Waikato Regional Council uses Kisters software to hydro-climate data, weather and river flow measurements; its telemetered water pumping stations; groundwater levels and station metadata. This data is accessible through a SOS interface as a web service.

Horizons Regional Council for example uses Hilltop software. In the course of the NEMS EODP initiative Hilltop participated and developed an experimental WFS and SOS interface to the Hilltop server software, which provided station metadata and locations through a Simple Features WFS interface. Hydro-climate data sets, weather and river flows measurements; telemetered water

1 NIWA CLIFLO http://clifo.niwa.co.nz

pumping stations and groundwater levels time series data were accessible through the SOS interface encoded in WaterML 2.0.

LAWA is a website for presentation of water data accessible through web maps and infographics. The LAWA database incorporates distributed data sets from the participating regional councils through WaterML 2.0. However, exact details of the employed procedure are not public.

The GNS GGW Database is publicly accessible through a web interface[2], which provides access to feature/sampling location metadata, groundwater chemistry, aquifer physical properties, observations collected and maintained through the National Groundwater Monitoring Programme. In the SMART programme sampling locations of water samples for Halon analysis and age tracer validation, distributed temperature sensing time series data sets were collected. Furthermore, a sensor station has been established in the Upper Rangitaiki catchment to test a low cost open sourced hard- and software platform with standards-based public data access.

In particular for the widely used hydro-climate time series data the OGC SOS interface and WaterML 2.0 as the transfer encoding are apparent. Thus, Table 9 (p. 101) summarises which interfaces and data formats were selected for groundwater SDI based on the described case studies information.

Sensor Observation Service (SOS)

The SOS specification describes a web service to store and access in situ observations and measurements (e.g. water level measurements) and related measurement process and sensor information. A SOS service delivers descriptive sensor information in a standardised format using the SensorML, and gathers/provides measurements from hydro-climate field sensors in Observations and Measurements (OGC O&M, ISO 19156) formats. In comparison to the WFS interface, the SOS specification directly describes the necessary semantics to access and deliver time series. SOS is part of the OGC SWE initiative and allows the aggregation of sensor descriptions and observation time series. The OGC O&M standard describes how time series and observation data can be encoded within OGC compliant interoperable systems. WaterML 2.0 is a specialisation of O&M and already widely used to exchange hydrological time series data and is an endorsed transfer format by the WMO since 2015.

SOS resources are also sorted in collections referred to as Offerings. SOS 2.0 specification allows only one procedure for an offering, with the explanation to reduce complexity and confusion with different collections harbouring the

2 GNS GGW Database: http://ggw.gns.cri.nz/ggwdata

same observations. Therefore, an offering depends mainly on the procedure to observed property mapping. The implication is that this needs special attention when addressing resources (the observations) in the SOS services. Although generically able to handle 'observations' SOS is often used for time series observations. Result format is predominantly O&M and WaterML 2.0. In SOS there is not a strict resource separation as in the former services. This makes SOS a little special. While SOS 1.0.0 is still in use with its default format of O&M 1.0, WaterML 2.0 Collections can be served with SOS 1.0.0. However, the current active and recommended standard would be SOS 2.0 with default formats O&M 2.0 and WaterML 2.0. There are also profiles in development to better cater for the hydro-climate community. These profiles, i.e. SOS 2.0 profile for Hydrology basically limit some flexibility in output to provide a more reliable data structure. Similarly, the NEMS EODP is developed as a profile of WFS and SOS to make it easier for clients to expect and interpret delivered data and service behaviours.

Sensor Planning Service (SPS)

SPS is also described in the SWE suite of the OGC. The one important operation of SPS in this groundwater infrastructure is the 'tasking' or (re-)configuration sensors or actuators to make observation, and with particular respect to hydro-climate sensor networks, reacting on previously defined measured environmental parameters. This reactive behaviour depends on well-defined interoperable input data available in near-real time which needs to be automatically processed based on user-defined threshold and indicators. Here the requirement of an immediate interoperable linking of geoprocessing capabilities into the groundwater infrastructure design becomes apparent. Geoprocessing with groundwater SDI compliant data sources will be addressed in Section 5.3 in more details.

In the course of this study these particular time series data harmonisation and interoperability principles were further applied to the case studies:

As an alternative to the CLIFLO web interface a SOS middleware was developed to access station metadata and hydro-climate time series data via SOS. Supported data encodings were WaterML 2.0, O&M 2.0 and SensorML. In a similar fashion, a SOS time series data service for GGW was developed as alternative access to feature/sampling location metadata, groundwater chemistry, aquifer physical properties, observations from the NGMP data set. The harmonisation procedure, the implementation details and evaluation are described in the next chapter (Adopting SWE for Observation Databases in Section 6.1.2).

Additionally, for the SMART Upper Rangitaiki sensor station, a middleware for the field data logger was developed to store and transmit hydro-climate time series data to an online SOS server. From there it was available and publicly accessible via the SOS interface and supported WaterML 2.0, O&M 2.0 and SensorML as output formats. Subsequently, for this sensor station a simple SPS integration with WPS was developed to further demonstrate the reactive integration between environmental measurement and threshold indicators. Section 6.1.2 explains the application and implementations of the described design.

5.1.3 *Summary of Selected Types for Data Sources Integration*

Table 9 (p. 101) summarises the data type and domain categories and concepts as emerged from case study data grounded theory coding, and how the resulting interface standards and data formats (as selected from the case studies and in reconciliation with the literature) fulfil the case studies requirements for the groundwater SDI part of the hydrogeology infrastructure.

5.2 METADATA

Metadata are data or information about the data itself. Metadata refers to structured information that describes, explains, locates, or otherwise makes it easier to discover, access or use data sets, collections, and services. For example, metadata elements 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 or where to get the data. From the analysis of the case study New Zealand Regional and National Data Sources (Section 4.2) and in reconciliation with the literature review section about SDIs (Section 2.2) it was concluded that although different metadata standards and methods to management of metadata are available and have been described in the literature, the ANZLIC Metadata Profile (version 1.1) would be the recommended geospatial metadata standard for use by New Zealand government agencies. This choice is further reinforced by the many data services in New Zealand which maintain online data catalogues that can be searched through a standards-compliant web service interface and provide metadata in the described format. In the following discussion the role

Table 9: Resulting categories and concepts for data formats and access, as emerged from case study data grounded theory coding

Type / Domain	Concept / Category	Target Design Element
boundaries or extents of features or mapping units		
	catchment	WFS/GWML2
	geological unit	WFS/GeoSciML, GWML2
	management zone	WFS/GWML2
	model boundary	WFS/GML SF, GWML2
	soil type units	WFS/GML SF, Soil Markup Language
	land use cover	WFS/GML SF
	cadastre / section	WFS/GML SF
identity or inherent domain types of features, with and without according metadata		
	hydrogeological feature types	WFS/GWML2
	geological feature types	WFS/GeoSciML
observations, measurements, samples of (environmental) phenomena		
	sampling location	WFS/GML as O&M, SOS/O&M
	well logs	WFS/GeoSciML, WFS/GWML2
	rock specimen	WFS/GeoSciML, WFS/GML
	water level	SOS/O&M SOS/WaterML
	river gauging	SOS/O&M SOS/WaterML
	rainfall	SOS/O&M SOS/WaterML
	wind speed and direction	SOS/O&M SOS/WaterML
	soil moisture	SOS/O&M SOS/WaterML
	hydro-chemistry	SOS/O&M SOS/WaterML
	geo-chemistry	WFS/GeoSciML, WFS/GWML2, O&M SOS/O&M
	aquifer properties (hydraulic conductivity, porosity, yield, flow etc.)	WFS/GWML2, O&M SOS/O&M WaterML
places, where things were measured, observed, sampled		
	sampling locations	WFS/GML, O&M
property distributions, gridded data, calculated ot otherwise derived		
	rainfall	WCS/GeoTIFF, NetCDF, ArcGRID
	evapotranspiration	WCS/GeoTIFF, NetCDF, ArcGRID
	recharge	WCS/GeoTIFF, NetCDF, ArcGRID
	probability of occurrences of sediments, e.g. gravel, sand	WCS/GeoTIFF, NetCDF, ArcGRID
	land use cover	WFS/GML
	elevation of surfaces, geological layers	WCS/GeoTIFF, NetCDF, ArcGRID

of metadata catalogues and metadata records in regards to the groundwater SDI will be explained.

Several major data providers (LINZ, LRIS, NIWA, GNS, DoC, MfE, see Section 4.2) are maintaining standards-based services for the search and discovery of their data. The main metadata formats which are used throughout New Zealand data providers besides ISO/ANZLIC metadata are CSW Record and DublinCore (*Dublin Core: DCMI Metadata Terms*, 2010), which are compatible subsets of ISO/ANZLIC metadata elements.

5.2.1 *Catalogue Service for Metadata Management*

The OGC standard for the CSW 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, 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. Metadata records in the groundwater SDI should be based on the ANZLIC Metadata Profile v1.1 (based on ISO 19115 (2003)). For additional service-level metadata the related ISO 19119 standard provides required elements. For the encoding, i.e. the data format, of metadata records for data sets and services, in which such metadata records can be delivered through the CSW interface, the ISO 19139 standard (ISO 19139, 2007) provides a standardised XSD representation. Table 10 (p. 108) summarises which categories and concepts relate to metadata in relation to the previously described CSW/`gmd:MD_Metadata` (ISO 19139) behaviour. This defined behaviour correlates well with the other OGC web service, with CSW being the interface and the metadata records being the resources that can be inquired through the service.

A CSW can be searched by many different fields, most usable through keywords and topical term based text search, spatial scale Bounding Box and temporal (e.g. inclusive time period or intersection), if a data set has a time domain mentioned. Also, geoscientific models specifically developed and calibrated for distinct catchments would be geo-referenced in a catalogue, too. Free and open source products and developments, different metadata tools such as GeoNetwork Opensource, the ESRI Geoportal Server or PyCSW can be used, too, to register, upload, and download metadata records as well as related spatial data, documents, PDFs or any other content. Details on the application and evaluation of these tools and example instance documents of metadata records are described in detail in the next chapter.

Metadata records in the CSW catalogues are underpinning and supporting the real data sets, services and processes through their anchoring capability. To understand their broader relevance within the described groundwater SDI, such metadata records were perceived as roots. Metadata records for data sets and services provide already fields like `Lineage`, `ResponsibleParty`, `LicenseConstraints`, and `DataQuality` elements to document data quality and provenance. It is of importance that within the infrastructure the framework keeps track of the metadata records, in particular when in more complex geoprocessing scenarios multiple data sets were amalgamated. The metadata records fulfil an anchoring/referencing function for their data sets and services throughout the processing steps and their lifetime within the SDI. The propagation of lineage, licence and data quality information throughout WPS Processes will be described subsequently in Section 5.3.

5.2.2 *Metadata Provenance*

Dealing with data quality, lineage and intellectual property rights aspects are another challenge in data exchange and data use processes. This is especially true when data is not published under a clear open data initiative, like the New Zealand Government Open Access and Licensing (NZGOAL) framework. In New Zealand, access restrictions result from different data collectors, management and user responsibilities and interdependencies. If data sets are improved and refined at the data provider or in a curator's database, new added value data products are created and are always published as free of charge open data sets again. While it was not the focus of this study and could not be verified throughout the course of this study, a big issue seems to be perceived from data users in the various agencies, that liabilities for data users and providers alike might arise through dissimilar licence agreements across the data providers, as well as preoccupation with data quality issues. Protection of intellectual property is typically enforced through non-disclosure of data. But where data was collected through governmental agencies, a conflict for the data provider stems from the perceived responsibility to have all data that is released also have them reliably quality checked. This is an enormous effort. As it is also not the focus of this study to develop an automated quality checking system, this Hydrogeology Infrastructure Design needs to be able to cope with these issues transparently and support future efforts, without encumbering the use of data. Thus, metadata provenance, reference to lineage and data quality, as well as licensing and access constraints (e.g. disclaimers) need to be immediately accessible and transparently propagated through all stages of the framework design.

Licenses, particularly well known license and copy right models like NZGOAL, including the application of an appropriate Creative Commons (CC)[3] license, have well defined 'degrees of freedom' and distinguish between 1) Attribution of source, 2) Derivatives allowed, 3) Commercial use of the data (or derivatives, if allowed) and 4) the ShareAlike constraint - to only re-distribute under the same license; as well as combinations thereof. In particular, when using different input data sets from various data providers in complex computations, it needs to be possible to properly discern between licensing conditions, access constraints, public or closed access and fees for a coalesced resulting metadata lineage for the new data sets what would be generated through the computation.

OGC OWC document specification describes how a so-called context document can reference OGC services, data sets; and in fact, any 'addressable' resource, and provide full contextual representation capabilities for situational awareness, catalogue search collections or analytical processing results (OGC et al., 2014b). An Atom-based XML (OGC et al., 2014a) encoding is also defined. The selection of the OWC standard and its relevance to this study is justified as follows: Metadata records describe either a service or a data set or a collection of such. A catalogue service could be queried and will possibly yield more than one metadata record for the provided search criteria. The metadata records reference an actual real data set, or process, or other addressable resource that the user would want to put to use in a software transaction, e.g. download of several data sets, visualising them or using them as inputs for geoprocessing/modelling exercise. These references, for the actual data and the corresponding metadata record need to be held together in a message format. Only through these means the complete metadata is accessible at any stage where the actual data is used, too. Below in Section 5.3 (Geoprocessing) the aspect of propagation of data quality, lineage and license information throughout complex geoprocessing scenarios with the help of OWC documents is explored in detail.

3 https://creativecommons.org/licenses/

Table 10 (p. 108) also summarises which categories and concepts relate to metadata and addressable resources in relation to the described requirement of keeping track of metadata records and their corresponding data sets or services. Linkage and reference of resources is a recurring pattern that reinforced the distributed infrastructure aspect of the developed Hydrogeology Infrastructure in this study. Example instances of OWC documents that were created and they were used in this framework are described in Chapter 6.

5.2.3 *Vocabulary Services and Semantic Convergence*

Above, section 'Catalogue Service for Metadata Management' introduced the concept of how metadata records in the CSW catalogues are underpinning and supporting the real data sets, services and processes through their anchoring capability. It was outlined how a CSW can be searched using a variety of different fields, in particular through the use of keywords, topical terms based on text; and by spatial extent. This elucidates the need to further support these basic search capabilities in the design framework for the groundwater SDI. A controlled list of keyword terms is expected to increase the quality of search results for specific data sets at specific places in distributed data repositories. Related Māori and English terms would support search functionalities in both languages resulting in respective data sets to be found independent from the search language. From the information services case studies (Section 4.2) a list of scientific freshwater terms was found to be hosted on the website of the Hydrological Society of New Zealand as a hydrological glossary. In the course of this study it was abandoned and is now available in similar form on the LAWA website. A specific groundwater related thesaurus has not been developed so far in New Zealand. Also, there was no further indication to be found in the case studies relating to intentions or activities towards the development of a controlled list of terms.

The New Zealand Hydrological Society publishes the Journal of Hydrology (New Zealand), which is considered as an important medium for the communication of scientific and operational results around water resources and their management in New Zealand. The journal articles can be accessed through a website, which provides a 'free-text' search over title, authors and abstract of the journal articles. It does neither provide a list of keywords nor does it require authors to tag their articles with keywords at all.

Such an important source of expert knowledge still should be discoverable through the groundwater SDI. Therefore, the minimal set of ANZLIC metadata elements was selected to create metadata records for journal articles. While the CSW protocol also allows for 'free-text' search over all elements of metadata records, in a distributed setting with a joint search over all possible catalogues in New Zealand a sensible approach to reduce the amount of results is needed. A distinctive advantage of the CSW protocol as compared to a plain text search are spatial and temporal search constraints, and furthermore, limiting search to selected keywords from controlled lists. Thus, to generate basic metadata elements for unstructured text documents, like the journal articles, location information and topic information needed to be extracted from the articles.

Presently, there is no groundwater related thesaurus available in New Zealand and the freshwater glossary is not accessible through a web services API form. Here it was concluded that the groundwater SDI needs to provide technical means to allow web services based access to spatial references based on place names and to a keyword list.

The RDF from the W3C provides the foundation for publishing and linking metadata elements through an RDFS. RDF as a data format has its own query language - the SPARQL. SPARQL makes it possible to send queries and receive results through web services over the HTTP protocol. The terms and identifiers of the data sets in their service location as well as seemingly free elements, like names of named parameters/observed properties which have no other identity, can be harboured as 'concepts' in systems like the SKOS. SKOS is encoded in RDF as a graph of triplets. RDF/SKOS collections/schemas can be made accessible via SPARQL servers that make those RDF graphs queryable via the graph aware SQL like query language (i.e. SPARQL). The technical details of the SPARQL and Linked Data implementation and the RDF/SKOS encoded glossary are described in Section 6.2.

This set-up added following capabilities to the groundwater SDI:

1. a controlled list of keywords (with their definition) is ubiquitously available, through web services form as well as human readable website;
2. this controlled list of keywords can be used to tag metadata records for more efficient search results;
3. unstructured text documents (like the journal articles from the case studies) could be automatically tagged with relevant keywords, based on occurrence of the words in the text.;

The other crucial issue with the inclusion of journal articles into the groundwater SDI is their lack of explicit spatially reference metadata. From the case study data sources (Section 4.2) and in reconciliation with the literature review about Spatial Data Infrastructures (Section 2.2) it was concluded that the use of a gazetteer service would provide the required capabilities. LINZ (Section 4.2) provides a gazetteer through WFS profile. Similar to the ontological aspect of the described thesaurus or glossary based keyword search for CSW queries, a gazetteer provides two very important properties of the place names concept: that are the name(s) and a corresponding spatial reference. In the literature there are also examples with a spatial SPARQL profile for a list of place names with coordinates, but the differentiation between the use of WFS or SPARQL for a geo-coding approach was not considered a distinctive design aspect, because both can be resolved via HTTP web services and exposed via Linked Data principles. Based on this design aspect locations could be found for journal articles which contained place names. The technical details of the geo-coding implementation for the journal articles based on the LINZ gazetteer are described and discussed in Section 6.2.

Once the above described vocabulary services were connected through SPARQL and semantic Linked Data principles, a metadata search could query all the different distributed catalogues and the list of results also showed metadata records from other repositories, including the journal articles.

Consequently, at the stage of creating a metadata record, thesauri and code lists (ontologies, vocabularies) need to be available to pick keywords, topics and place names. These could then be used to tag and geo-reference metadata properly with verbs/terms that have a representation in existing controlled vocabularies.

The OWC ’context’-based data handling was then used to bundle different data sets, addressed via their web service interfaces, and could can be used as referenced data packages to be sent to geoprocessing, modelling and visualisation services, which are explained in the following sections.

Table 10 (p. 108) also lists the data type and domain categories and concepts which relate to overall metadata provenance. The resulting interface standards and data formats (as selected from the case studies and in reconciliation with the literature) fulfil the case studies requirements for the groundwater SDI part of the hydrogeology infrastructure and thus contribute to research objectives 1 and 2 and add to the answer to research question 1.

Table 10: Resulting categories and concepts for metadata formats and access, as emerged from case study data grounded theory coding and the literature

Type / Domain	Concept / Category	Target Design Element
collections of metadata elements that identity and characterise spatial data sets and services for features		
	discovery through metadata records	CSW/gmd:MD_Metadata (ISO 19139)
	lineage and provenance information for data sets	CSW/gmd:MD_Metadata (ISO 19139)
collections of metadata elements that identity concepts or properties which are not captured as features		
	definition of terms in glossary or thesaurus	RDF/SKOS RDF/OWL
	definition of environmental phenomena	RDF/SKOS RDF/OWL
	references of resources	CSW/gmd:MD_Metadata (ISO 19139), HTTP/(OWC, LD), SPARQL/RDF
	collections of referenced resources	HTTP/OWC, SPARQL/RDF

5.3 GEOPROCESSING

The characterisation of a hydrogeological setting is a multi-faceted complex task and the availability, access and knowledge about provenance and quality of relevant data sets are important considerations for the hydrogeologist who is undertaking respective research. In the two sections above, 'Data Sources' (Section 5.1) and 'Metadata' (Section 5.2), Research Question 1 has been already answered to a large degree, and tables 9 (p. 101) and 10 (p. 108) summarise 'the required interfaces and data formats to comprehensively capture describe and share hydrogeological and groundwater related data sets in New Zealand to enable data interoperability for hydrogeology'.

To answer the Research Questions 2 ('Can existing scientific groundwater flow and water balance simulation models and software modules be made interoperable?') and 3 ('Can such models be provisioned with available interoperable data sets?') the envisioned 'Hydrogeology Infrastructure' advances theory for designing interoperable data and analytics infrastructures. The new design is based on the selected methods, standards, formats and technologies and addresses short-comings from existing methods.

In this section, specific attention is given to gaps revealed as a result of the literature review and to the distributed/networked groundwater analytics infrastructure design aspects. Analysis and visual exploration of the data sets requires practical support by computer applications. Although a variety of software for this purpose is freely available, they require a good command of the technology or programming language for their application in a hydrogeological assessment.

Thus, integrated proprietary software products are often used to analyse and particularly provide high-quality visualisation of the system. However, these software tools are typically desktop programs with strict licensing arrangements, limited extensibility and lack of interoperability with other applications.

One trend revealed from the literature is that most systems have to be installed and run on a desktop, which decouples concerns of data processing and data storage through a local workflow, and at the same time integrates distributed data from web resources on-demand. The advantage is that users do not need to pre-process and format data into the required data formats for the available processing packages (e.g. CUAHSI-HIS). Software packages for modelling, simulation and visualisation are integrated into the framework and are executed on the local computer. While this solves the data access and pre-processing steps, it implies two limitations: 1) The outputs of successful program runs are merely available as local (on the desktop computer) end products and it is not immediately intended to make results available by similar means into the a networked data sharing environment from where the inputs were gathered in the first place. 2) Local desktop or small working group networks do not provide the computing power of a high performance computing cluster, nor a company-wide or cloud data centre. As all the computation for all linked/integrated models have to run locally, scaling up and distributing computation is neither trivial nor flexible. To advance GIScience theory in that space, this section explains how the encapsulation and linkage of groundwater-related software modules into standardised web-enabled geoprocesses allows for flexible distribution of the computation.

The geological and water budget assessments for the Horowhenua area in the southern Manawatu-Wanganui region of the lower North Island of New Zealand served as case studies to build a groundwater model based on MODFLOW. However, to create an internet-enabled interactive MODFLOW execution and modelling environment, it needed to be integrated with the 'Hydrogeology Infrastructure'. The flexible open standards toolbox design of the OGC Web Processing Service (WPS) enables extensibility and interoperability with other web-enabled hydrology-related and GIS processing tools and conventional workflow design and execution systems. The WPS-enabled MODFLOW software module could run pre-configured and calibrated groundwater models with latest available observation data from the groundwater SDI. Furthermore, the interoperable groundwater model was linked with other WPS processes, and thus created the foundation of a so called 'model web'.

WPS Web Processing Service

The WPS is not immediately a spatial or temporal data service itself, but a generic processing service. The executable algorithms or processes of a WPS server are the resources it provides; and these resources are addressable via their process identifier. Processes require specific inputs and provide specific outputs, which are described in a standard notation in its WPS process description. In general WPS processes are encapsulating typical geoprocessing capabilities and therefore have geodata as input and provide geodata as outputs. They provide a process description, listing the input fields with their possible input data and types / formats as well which output is to be expected, type and format. The inputs and outputs can also be addressed via reference and do need to be copied to the process directly, but can then be fetched by the processing service or the subsequently dispatched process and the client can access process output which also can be retrieved as reference.

In 2015, the WPS 2.0 specification has been released and technically allows describing WPS processes through the SensorML process model. That would provide for more detailed and explicit formalised input and output description, in particular with the rich data annotation possibilities provided through 'SWE Commons' elements. However, there are no references to be found in the literature, nor are there publicly available implementations of WPS 2.0 in combination with SensorML 2.0 yet. In this study there has not been made any attempt to implement a new WPS service with SensorML. In Chapter 7, the possible implications will be discussed in more detail.

WPS processes can also be 'catalogued' with their respective metadata, i.e. identifier, long name and description, input and output parameters, and where and how they can be accessed as HTTP resources, alike other OGC services and resources under HTTP paradigm. This makes it possible to bundle WPS processes like a toolbox contexts via HTTP references, too, as described earlier (semantic linkage and context handling in Section 5.2) through OWC context documents.

MODFLOW as WPS process

The data sets and the water budget and geological model assessment reports for the Horowhenua case study (Section 4.1) were compared to the literature findings in order to expose the necessary data sets via OGC web services.

Table 11: Input data type mapping, criteria and resulting OGC/web-enabled data source conversion for MODFLOW

Originating GIS source	Criteria, Domain	OGC / Web data source target
DEM GeoTIFF, ArcGRID raster	Physical surface elevation	OGC WCS
excel spreadsheet, report, shapefile	measured / derived aquifer properties	WFS GWML2 `Aquifer` properties, `Fluidbody` or `Void` properties
Excel spreadsheet	Vertical bore log observation data	GWML2 `GeologyLog`
Shape file, boundary of study area/-catchment	Management area or aquifer/basin boundary	GWML2 `ManagementArea` (for water budget reasons)

Table 11 shows exemplary input data type, criteria and resulting OGC/web-enabled data source conversion mappings.

The original USGS MODFLOW program is a Fortran and C programming language command line utility which works solely on local files with their own specific formats. Thus, MODFLOW needed to be either extended to be able to include web data sources or be wrapped into a networked execution environment which takes care of running the MODFLOW program as well as data pre- and post-processing.

MODFLOW was integrated through a layered approach into a WPS environment, which on the higher level interacts with OGC geodata sources and are then translated into the data types, which in turn could be accessed by MODFLOW. To integrate MODFLOW with the WPS execution environment, the groundwater SDI data sets needed to be translated into MODFLOW native (i.e. Fortan matrices) data grids. Policy informing data elements for groundwater flows and budgets can be updated in the WFS GWML2 `ManagementArea` on the fly. Although it would be possible to accomplish this via the WFS transactional profile, the decision needs to be made after review by the researcher and resource officer. In typical groundwater modelling exercises MODFLOW is run many times to create many realisations of the same hydrogeological setting. Uncertainty is assessed and the most likely and reasonable representation in the eye of the hydrogeologist is taken into account for further processing and discussion.

The described procedure enables a web based data and information sharing framework for near real time access to environmental models utilising common data interfaces provides on-demand re-calculating water budgets for advanced decision support. Thus, it is possible to drive modelling routines on request and provide messages on present and future environmental conditions to those who need them. The technical implementation details and

Table 12: Resulting categories and concepts for data processing, as emerged from case study data grounded theory coding

Type / Domain	Concept / Category	Target Design Element
processes, algorithm, codes that calculate or otherwise generate values or property distributions		
	rainfall recharge, custom Fortran model with XYZ / ArcGRID output	WPS with WCS by-ref/by-value or compatible coverage output[4]
	groundwater flow/yield data from MODFLOW	WPS with SOS/O&M compatible or WCS by-ref/by-value compatible coverage output
	Equilibrium Water Table custom MatLAB with GeoTIFF, NetCDF, ArcGRID output	WPS with WCS by-ref/by-value or compatible coverage output
	GIS spatial analysis and geoprocessing (raster algebra grid preparation, clipping, shapefile merge and attribute manipulation)	WPS GIS geoprocessing with WFS or WCS by-ref/by-value compatible
sensor network, live telemetry		
	data logger and sensor information/metadata	SOS/O&M and SensorML
	sensor and sensor network management and configuration	SPS/Tasking and SensorML

current application are illustrated in Section 6.3. Table 12 summarises which categories of data for WPS modelling inputs and outputs were derived in particular from the Horowhenua case study (Section 4.1 and Table 7) and reconciled with the envisioned hydrogeology model-web design. The presented methodologies, technologies, and international standards provide for such a framework of distributed data and processing. Consequently, they can be interlinked with advanced visualisation and linked community thesauri and vocabularies.

Model Web Execution Linkage

Based on the examined hydro(geo)logical data sets from the case studies (in particular from Horowhenua, Section 4.1), and in reconciliation with international advances in the development of hydrogeological data models and open encoding standards, the 'Hydrogeology Infrastructure' design now provides necessary means of storing, accessing, discovering and downloading required data sets to provision the networked on-demand model execution environment through the developed procedure of making groundwater modelling software interoperable and link it with the groundwater SDI.

A complete WPS orchestration of a groundwater flow modelling exercise would involve a number of subsequent execution steps with WPS processes with different input data sets. Such orchestration can be achieved, for example, with conventional SOAP-based workflow systems. Eventually, MODFLOW output data sets would be saved back to SDI (e.g. as a WCS coverage grid, a number of `OM_Observations` or complex feature types, which should be identifiable via URI reference identifiers). Via the OWC such URI reference identifiers can be tracked also for transient data sets throughout the SDI. Furthermore, they could also to be linked to a 'project' or 'workspace', to allow bundled sharing for a subsequent review process.

In conclusion, encapsulating Environmental Models as WPS processes enables their use as web (HTTP) addressable bundled resources in OWC documents. This adressing and bundling scheme can now be advanced into a hierarchical recursive orchestration of context-based WPS geoprocessing, where OWC documents provide a list of bundled input resources for WPS processes. This concept will be further elaborated on in the next Section 5.4 ('Geovisualisation').

5.4 GEOVISUALISATION

In the 'Geoprocessing' section above, Research Questions 2 and 3 have been addressed. In this section the design aspects developed as part of this research will be illustrated in order to answer Research Question 4: Can visualisation be realised as interoperable model and being provisioned with data from the groundwater SDI, and thus complete the envisioned 'Hydrogeology Infrastructure'?

Apart from WMS/WMTS all the mentioned standards so far apply to data discovery, exchange and processing. Besides pictorial mapping and graphing of time series there is no common direction for visualisation. However, there are fundamental advancements in the literature and in the availability of 3D technology for the web. The following paragraph recalls the essential elements from the literature review.

HTML is the standard encoding for representation of websites in the internet. HTML features a so called DOM which holds the internal structure of the website when it is rendered in a web browser. In HTML4 as well as in the new HTML5 the DOM is programmatically accessible via an open standardised API. The user, or a software module therefore, can interact with and modify the DOM through JavaScript. With x3dom, the German Fraunhofer Institute fuer Graphische Datenverarbeitung (IGD) develops an 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. X3D is a modern 3D declarative XML-based Markup Language (ISO/IEC, 2008) that evolved from and succeeded the former non-XML language VRML (ISO/IEC, 1997). Based on the combination of X3D, HTML5 canvas and WebGL, several approaches to visualise geodata in the browser have evolved and were described in Chapter 2.6. However, none of the documented methods were immediately integrating SDI data sources with a dynamic 3D visualisation. In summary, two modes have become apparent:

1. Visual models of 3D geological model data are generated from physical data sets exported from the various modelling tools and the result are the declarative 3D scenes and textures, which are subsequently loaded onto a web space from where they are rendered in the web browser. The viewing usability greatly depends on performance of the browser's 3D rendering capabilities.

2. A second approach is a portrayal proxy, which does not render the 3D scene in the web browser, but on a dedicated rendering machine. A web browser application transmits requests containing specific view parameters like angle or zoom, to the rendering server, and gets an image in return. Here the usability greatly depends on network latency and rendering performance of the rendering machine as each change in perspective needs to resend a newly rendered image via the network.

The reason to detach model generation and model rendering from each other in case 1) is, that the visual model generation process, the translation from the data sets into the scene and textures, is complex, error prone and full of manual exception handling.

In the seconds case the rendering software understands the native geodata sets, but it's highly optimised software and a computation intensive process that does not run in the web browser.

Now, this study so far has developed comprehensive theory about standardised distributed (geo)data provisioning ('Groundwater SDI') as well as standardised distributed processing with interoperable data ('Model Web'). In reconciliation with the current state in web-based 3D visualisation, in this study it was inferred via abductive reasoning to understand the advanced visualisation as a specialisation of the 'model web'. In comparison with WMS and WMTS, the image rendering also stems from an underlying geographic data source. WMTS is statically rendered and deposited at an HTTP-accessible location. From there the images are served for respective query parameters like coordinate references and zoom level. WMS Implementation render underlying vector (e.g. shape files, database) or raster data sets (e.g. GeoTIFF) into images on demand at the time of the WMS query, also based on a variety of allowed WMS query parameters.

In the following argument the method is explained, which employed the interoperable data sources, metadata and geoprocessing theory to dynamically generate X3D which subsequently could be immediately loaded into and rendered in the web browser, and thus, advancing theory for 3D web visualisation of geographical and 3D geological and hydrogeological data sets.

The 3D geological model layers were only available as plain text grid files export from the EarthVision software and represented the top of geological layer boundaries. Based on the conceptualisation of the data sources from the case studies, these files were treated as spatial coverages with height as their attribute values. This allowed publishing such data sets as standardised OGC WCS accessible coverage layer. As explained previously, WMS implementations render from underlying data sources, and so the WCS service implementations typically provide WMS rendering capabilities. Therefore, 2D web mapping application could show the actual spatial context of the geological layers together with other data sets from catchment boundaries and sampling locations of the case study area. Here, the geological underground model was visualised as an area in the 2D view from the WMS source. The uppermost elevations of all the 3D geological model layers represent the actual land surface and should be concurrent with a DEM/DTM.

In Section 5.2 the specific employment of OWC documents through this study to keep data sets and their respective metadata records via CSW together was described. The OWC standard was designed to create contexts from different OGC data and processing resources, like WMS, WFS, WCS, SOS, and WPS.

An architecture that aggregates the results of the distributed services into the scene graph through the described WPS and OWC orchestration resembles a context-based, functional composition. Table 13 demonstrates how the OGC service and data elements are mapped and transformed to X3D elements. The mapping was derived in particular from the Horowhenua case study (Section 4.1) categories and reconciled with the groundwater SDI and model-web design. In the model output is the X3D visualisation. In Chapter 6, Section 6.4 implementation and evaluation of the described method will be examined in detail.

As an example, through a WPS algorithm spatially distributed groundwater level observations could be interpolated into a groundwater surface. The implicit spatial property interpolation from the spatio-temporal sampling data

Table 13: Mapping of data source type to visual type in X3D

OGC data source	X3D element
WMS layer	image draping nodes for (top) surface
WFS, complex features (e.g. wells)	The features will be rendered with their borehole geometry based on their designated elevation attribute
WFS, simple features	The features will be rendered with their (flat) geometry on the designated surface grid
3D geological model layer, WCS elevation data set	`ElevationGrid` as an evenly spaced grid
SOS (e.g. groundwater level), WPS process returning a coverage	The result coverage layer will be treated as elevation grid
SOS, sampling feature locations	The features will be rendered with their (flat) geometry on the designated surface grid

was defined in the OWC context document. Thus, the WPS X3D generation process dispatched another WPS process that interpolated the groundwater level surface. This dispatched process itself again pulled its referenced groundwater level observations input data from a SOS service. All this distributed services interaction implied prior knowledge of the interface descriptions of the participating services in such loosely-coupled SOA. Because they are all standardised interfaces the dispatch method could rely on the interface descriptions at runtime. In Figure 19 the process of scene generation with an additional WPS groundwater surface interpolation is described from the initial OWC to the generated X3D scene.

Three-dimensional web visualisation capabilities for environmental data in New Zealand are presently limited and established 3D models are based on cost intensive proprietary desktop software solutions. This 3D visualisation method based on distributed data sources and scene creation framework services could enable more flexible insight into the subsurface and thus a better understanding of available (ground) water resources.

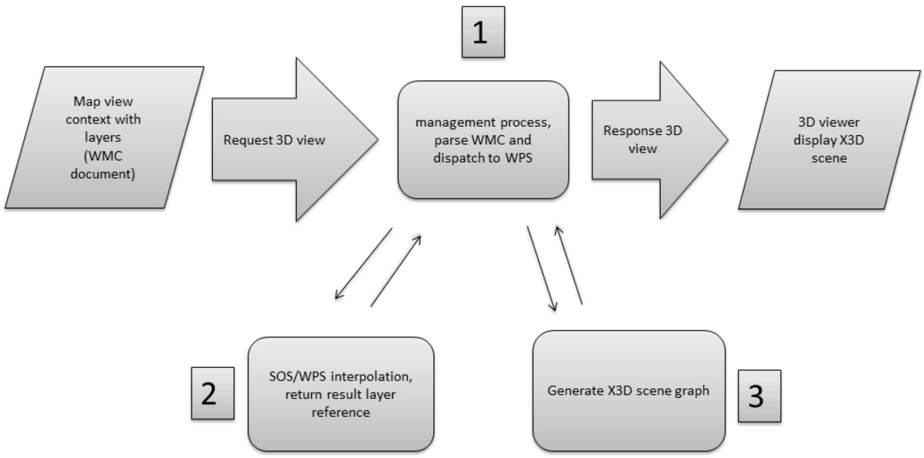


Figure 19: WPS Process for X3D scene generation, source: own illustration from Kmoch and Klug (2014)

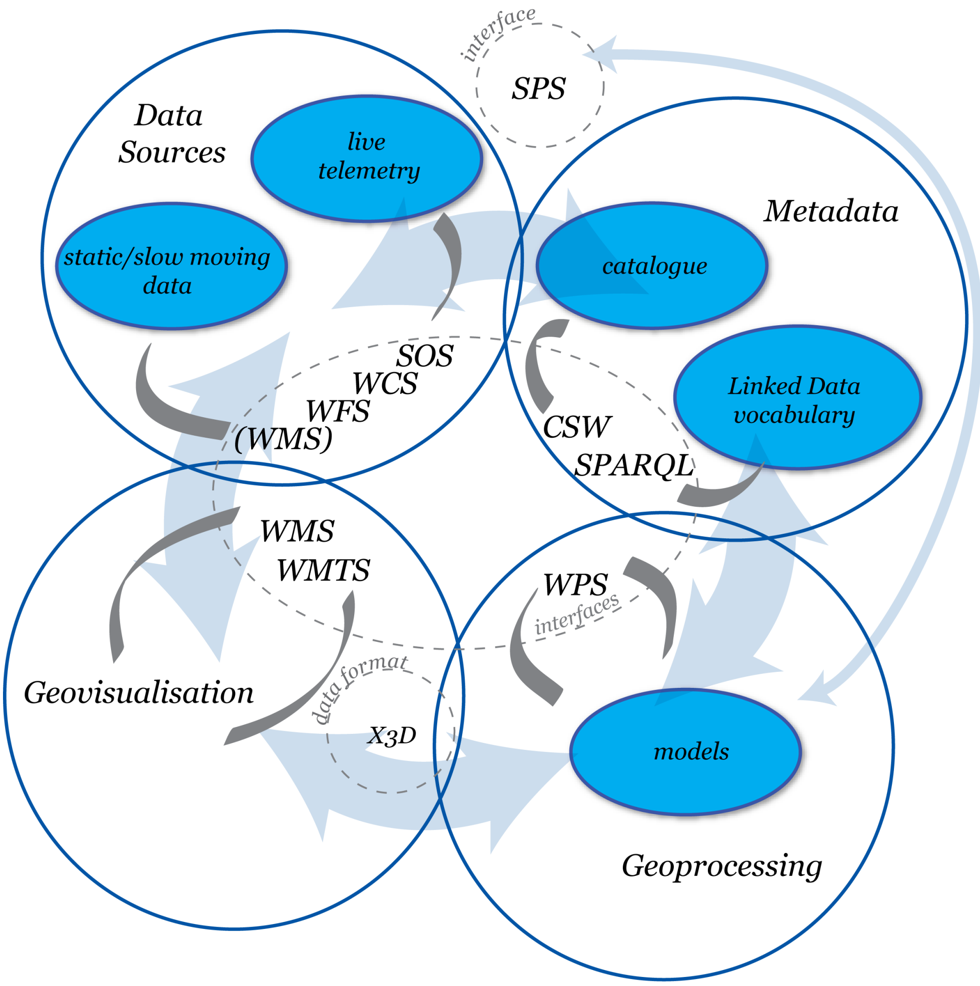


Figure 20: A schematic of how the disparate data sources and functions are unified in the Hydrogeology Infrastructure

5.5 SUMMARY

Based on the latest available international and national standards and the identified existing gaps in New Zealand, a service-oriented modular based concept was developed as shown in Figure 20. Multiple interfaces provide the means to deliver (Data Sources), use (Geoprocessing) and visualise (Geovisualisation) hydro(geo)logical data while keeping track of the data sets provenance and semantic linkage (Metadata).

The above sections described the theory for a new explicit Hydrogeology Infrastructure design. Through the Literature Review (Chapter 2), the application of the specifically developed research method (Chapter 3) to analyse and conceptualise the case studies (4), this study has reviewed methods, standards, formats and encodings in the literature and in relevant applied technologies and software systems for integrated groundwater data management, processing and visualisation. Gaps revealed from the literature were compared to New Zealand groundwater data management practices and analysis tools. Consequently, an appropriate set of such standards for a distributed/networked groundwater data and analytics infrastructure were identified and incorporated into the design for a comprehensive 'Hydrogeology Infrastructure'.

A reference implementation of the Hydrogeology Infrastructure has been developed and applied to the case studies. The required interfaces and data formats to comprehensively capture describe and share hydrogeological and groundwater related data sets in New Zealand enabled data interoperability for hydrogeology ('Groundwater SDI') and answered Research Question 1. Current challenges and short comings in existing methods were addressed, in particular by reshaping the understanding of visualisation of groundwater data as a specialisation of the model web ('Visualisation as an interoperable Model'). Also sharpening the perspective of how standardised data formats and their accessibility via standardised interfaces enabled interoperability also for existing scientific groundwater flow and water balance simulation software modules ('Model Web' and 'Context-based Data Provisioning'). Thus, Research Question 2, 3 and 4 have also been answered and this study advanced theory for designing interoperable data and analytics infrastructures based on the selected methods, standards, formats and technologies.

The envisioned 'Hydrogeology Infrastructure' application and implementation will be evaluated in the following Chapter 6, to be able to critically discuss in Chapter 7, if the new (theory of the) framework design overcomes the challenges and closes the gaps from the literature.

6 IMPLEMENTATION AND EVALUATION OF THE DEVELOPED DESIGN

The focus of this chapter is an implementation of the design theory presented in Chapter 5. The purpose of the prototype is to enable an exploration of interoperability behaviour and to act as a proof of concept. Thus, the structure of this chapter is closely aligned with the previous chapter's structure (Figure 20, p. 118). The implementation of the prototype is examined in relation to the research method from the perspectives of GT and DSR. Consequently, the development must be grounded in the data and implementation artefacts created that represent an essence of the improved design theory. Therefore, this chapter describes an evaluation of the design theory's credibility, transferability, dependability and confirmability.

6.1 DATA SOURCES INTEGRATION

To interact with the infrastructure designed in this research, a web portal was developed as a reference client, with a user facing front end and a service oriented back end. The front page of the user front end was implemented as a webpage (Figure 21), which for demonstration purposes provided immediate access to the Hydrogeology Infrastructure components described in Chapter 5.

The general architectural structure of the components is illustrated in Figure 22. The main data processing, transport and interaction with the infrastructure happens within the 'Data Interface and Services Layer'. Existing data providers (e.g. from the New Zealand case studies, Section 4.1) can be registered and connected to the infrastructure via standardised OGC web services to the exposed service interfaces.

6.1.1 *Groundwater SDI*

Table 9 (p. 101) lists the required interfaces and data formats to enable a groundwater SDI in New Zealand. These were applied to and tested with the

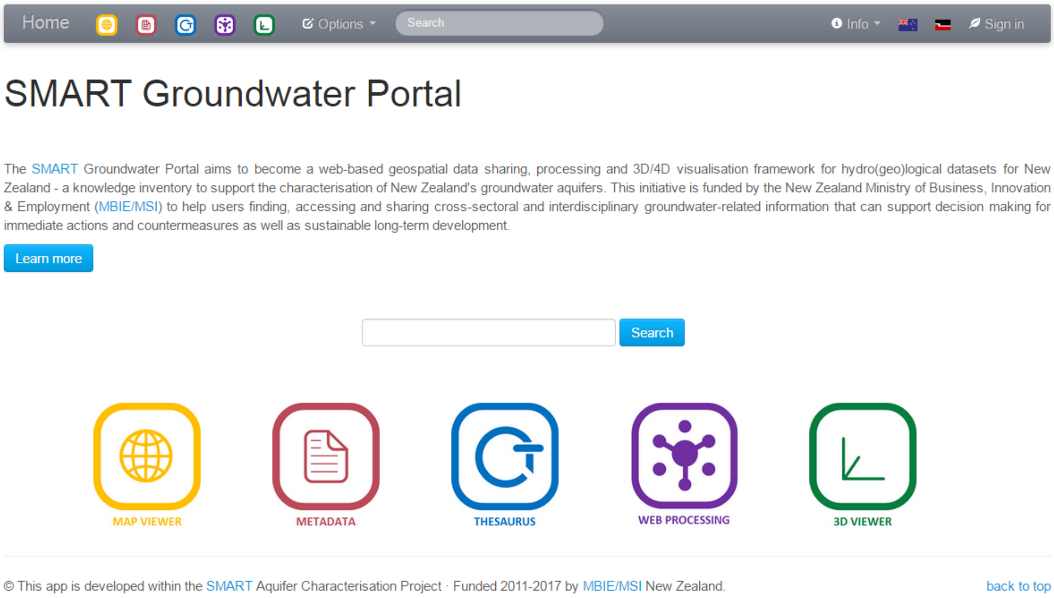


Figure 21: Groundwater Portal Frontpage, with five buttons representing the main tools, Mapviewer to view map contexts, Metadata management, Thesaurus to explore and manage vocabularies and glossaries, Geoprocessing with the examplary goundwater WPS and 3D Viewer to view 3D scenes of the map contexts

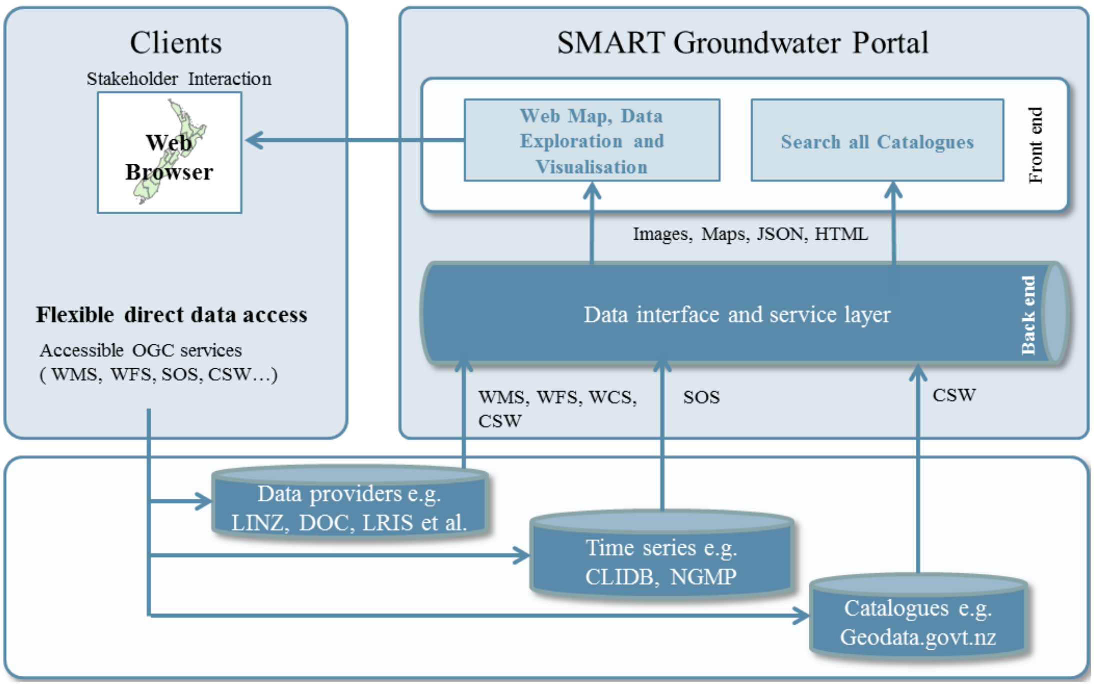


Figure 22: The groundwater portal as infrastructure client interacting with services, source: own illustration from Kmoch et al. (2015)

case studies. The main data sources should support the OGC WFS, WCS and WMS standards to serve data sets in the designated encodings.

For the purpose of sharing geospatial data in a standardised way there is a variety of open source, off-the-shelf server software readily available, for example Geoserver[1], Mapserver[2], Deegree Web Services[3], Constellation SDI[4], THREDDS Data Server[5], tinyWFS[6], ncWMS[7] or QGIS[8]. The OSGeo Foundation[9] hosts a large collection of free and open source GIS related software packages. Whichever component is selected for the prototype, it must be able to serve GML complex features. This requirement reduces the number of choices. Geoserver is a very popular interoperable spatial data server application that is widely deployed through the OneGeology initiative to serve GeoSciML. Geoserver is OGC compliant and supports several required formats including WFS/GML for simple features and complex features through the 'app-schema' extension.

Thus, Geoserver was selected as the component for managing the transformation and harmonisation processes. Technical transformations from source data sets into structured XML type formats (e.g. GML applications) were performed through publishing, thereby simply exposing the data, using the Geoserver software.

Through the New Zealand Environmental Information Services case study (Section 4.2) it was apparent that many New Zealand organisations already provide data sets via WFS/GML simple features access. However, data access to geological units/structures, well descriptions, borelogs, lithology observations (Table 4, p. 74), elevation and surface data sets (Table 3, p. 73), catchments and management zones (Table 5, p. 74) were only to be found in GNS's WFS/GeoSciML Geoserver implementation of the New Zealand Geological Map.

Data sets from hydrological and hydrogeological numerical model results (i.e. outputs) like rainfall recharge, evapotranspiration or the equilibrium water table are predominantly raster data sets (Table 7, p. 76). Thus, these were published through Geoserver's WCS/GeoTIFF capabilities. The NetCDF data sets were successfully published with THREDDS.

1 Geoserver website: http://geoserver.org/
2 Mapserver website: http://mapserver.org/
3 Deegree website: http://www.deegree.org/
4 Constellation website: http://www.constellation-sdi.org/en/
5 Thredds website: http://www.unidata.ucar.edu/software/thredds/current/tds/
6 tinyWFS website: http://www.mapserver.org/tinyows/
7 ncWMS website: http://www.resc.rdg.ac.uk/trac/ncWMS/
8 QGIS website: http://www.qgis.org/en/site/
9 OSGeo website: http://www.osgeo.org/

Two specific harmonisation challenges were related two the EarthVision geological layers data files and the complex features data provisioning. As explained in Section 4.1 the EarthVision data export provided simple XYZ surfaces. The file format for the XYZ data layer files is plain text and holds a single coordinate pair per line with respective Z value (i.e. in this case elevation). These files needed to be translated into a WCS/GeoTIFF file, because GeoTIFF is a compatible raster format that can be handled by Geoserver. GeoTIFF was also a common format for elevation raster files in the Horowhenua case study. A Java program was implemented that calculated the bounding box and regular spacing from each one of these XYZ files and built a GeoTIFF raster (code listing in Appendix 34). The point distribution in the XYZ surfaces was regularly spaced, but only described the existing points where the surface existed. There was no particular notion of a NODATA field, as non-existence of values was not encoded in the XYZ file. A Geotiff file needs to have values for all the spatial context is codes for. Thus, the 'empty' positions were explicitly filled with a standard NODATA value. Subsequently, these layers could be published via Geoserver.

- catchment, management zone, model boundary:
  `GW_ManagementArea` (Appendix B.1, Listing 15)
- monitoring sites: `GW_MonitoringSite` (Appendix B.1, Listing 16)
- springs: `GW_Spring` (Appendix B.1, Listing 17)
- well descriptions:
  `GW_Well` (link lithological logs via gwWellGeology property, Appendix 18)
- well logs or borelogs, lithology observations:
  `GW_GeologyLog` (Appendix B.1, Listing 18)
- hydrogeological units and aquifers:
  `GW_Aquifer` (Appendix B.1)

The designated complex features from the Horowhenua data sets were mapped via Geoserver app-schema configuration into WFS/GWML2 elements (full listings in Appendix B.1).

A graphical map viewer was implemented to retrieve the SDI data and to visualise them on a map. Figure 23 shows an image of the web map with aquifer and geological models boundaries and NGMP monitoring sites. Geoserver (and THREDDS) provided the capabilities to render WMS images

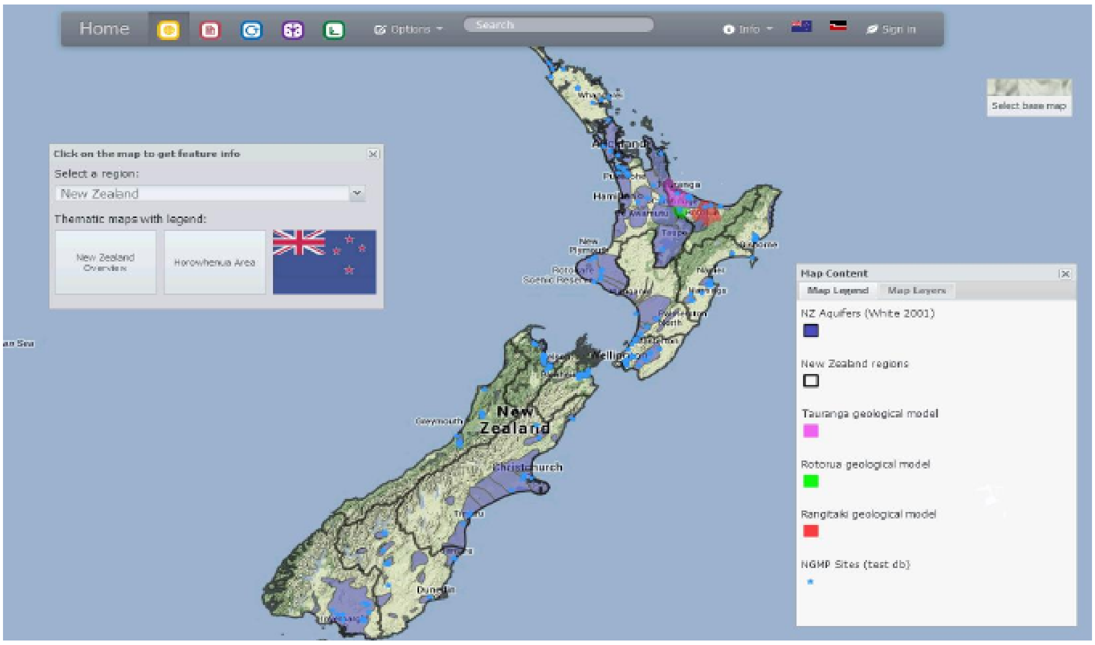


Figure 23: Groundwater portal with web map view of New Zealand, with boundaries of aquifer systems and geological models

from the simple feature data sets and WCS coverages (GeoTIFF, NetCDF). Complex features where not immediately displayed on the map, instead, structured data retrieval was triggered through a two step approach linked through the simple feature map access.

6.1.2 *Sensor Web Integration*

The integration of live telemetry measurement data into the Groundwater SDI as a data source was tested and evaluated in several iterations using the SMART Aquifer characterisation Programme (case study, in particular with the sensor station in the Upper Rangitaiki catchment, and the nationally significant databases GNS NGMP and NIWA CLIDB.

Upper Rangitaiki Sensor Station SMART Case Study

In the first stage a SMART case study site was established in the Upper Rangitaiki catchment, east of Lake Taupo. Data was acquired from a monitoring station deployed by the researcher. Figure 24 shows the setup and location of the sensor station, which is still in operation. It is comprised of a field computer (Raspberry Pi) with a direct internet link and a sensor board (Waspmote) that has 12 typical meteorological, hydrological and pedological sensors attached (i.e., wind speed, wind direction, rainfall, 1x groundwater

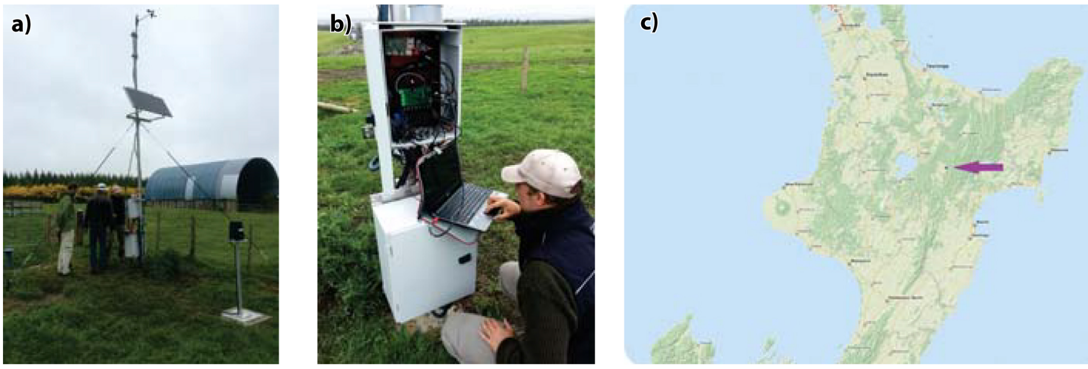


Figure 24: (a) Upper Rangitaiki sensor station set-up, with rain collector, soil moisture in the ground; (b) the researcher with field computer and sensor boards using a live internet connection via 3G; (c) location of the site, east of Lake Taupo, North Island

probe, 5x temperature and 3x soil moisture). The Raspberry Pi and Waspmote can be monitored and reprogrammed from a web browser. This sensor station setup allows scaling up to a multitude of low cost, low energy sensor stations throughout a catchment, with only one field computer that serves as data logger for backup.

The station is running a Linux operating system. A Java process collected the ground-truth measurements from the locally connected environmental sensors. A Java software library was developed ('wasp-uplink') that added upload capabilities to the data logger process. This software library encapsulated each measurement into an O&M `OM_Measurement` encoding and submitted the measurements via SOS `InsertObservation`. The observation encoding needs to specify an `ObservedProperty`, a Procedure, a Feature Of Interest and an Offering. The SOS 2.0 specification prescribes that the Procedure needs to exist at the time of `InsertObservation`. The library checks if the Procedure exists, and if it does not, then a Procedure is is created based on the locally available sensor station and sensor metadata from the data logger. The generated SensorML for the soil moisture probe is provided, as an exemplar, in Appendix 20. The data logger was equipped with a GPRS/3G modem connection which provided the necessary wireless internet access via cellular network.

The sensor and circuit board instrumentation collects data and forwards them to the field computer in 10-minute intervals to the field computer. The field computer transfers observation data in 10-minute intervals via a 3G mobile data connection to an online SOS server. From the service the observations are available in a standardised open format O&M 2.0/ WaterML 2.0.

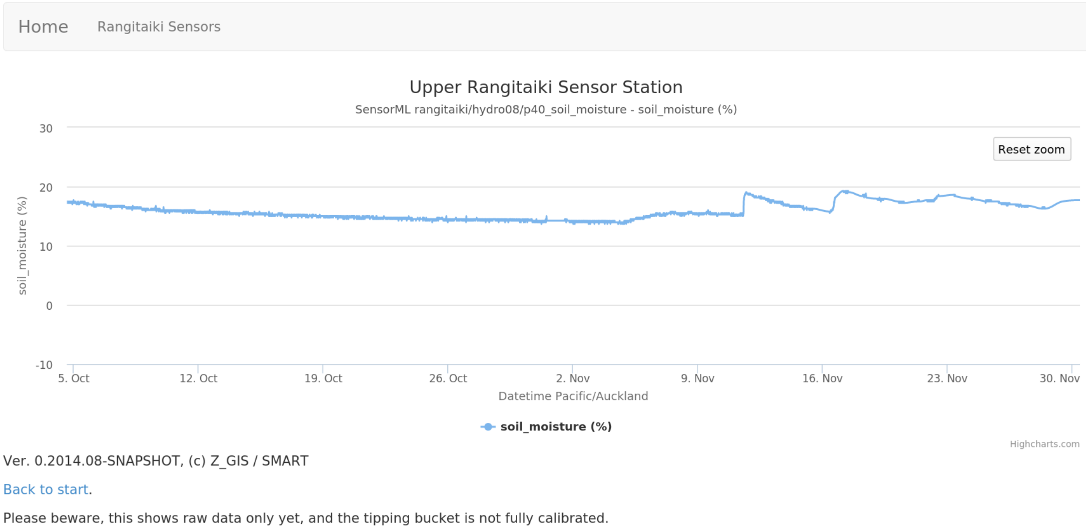


Figure 25: Screen shot of the groundwater portal website accessing and plotting the raw, ground truth data, from the Upper Rangitaiki sensor station data via the SOS server

The groundwater portal website as reference client accesses the raw data from the SOS server and plots data points within 5-10 minutes of field measurement as shown in Figure 25. Thus, this implementation provides proof of concept and evidence that it is possible to link the collected ground truth data, in near real time, directly to the groundwater SDI.

Adopting SWE for Observation Databases

The NGMP and CLIDB databases from the New Zealand regional and national data sources (Section 4.2) and SMART Aquifer characterisation Programme (Section 4.3) were used to link observation data from existing databases through SOS.

Both the NGMP and the CLIDB databases store point based sampling locations and time series of observations/measurements for certain environmental conditions. However, the means of measuring data, i.e. the procedure, varies. Table 14 demonstrates how the NGMP and CLIDB database elements can be mapped and transformed into OGC SOS2.0 and O&M 2.0 elements.

Both databases have their own domain specific database schemas as well as dedicated custom web interfaces for data access. To make these databases interoperable, two versions of the open source 52°North (52N, an open source geospatial software company) Sensor Observation Service have been modified to connect to each of the databases.

Table 14: The mapping of local domain data model semantics into OGC SOS 2.0 and O&M 2.0, source: Kmoch et al. (2015, tbl. 3,p. 11)

SOS 2.0/O&M 2.0	**NGMP**	**CLIDB**
SOS/`OM_Observed-Property`	`Parameter`, a numerical identifier of the associated entity, which has a name and a description, e.g. `1679` - 'Static Water Level' - 'Distance to standing water level'	`Measure`, an abbreviation (aka LADA Code) as identifier of the measured phenomenon as well as its temporal spacing, e.g. `223HOURLY` - 'hourly rainfall'
SOS/`OM_Procedure`	`Parameter` only, each `Parameter` has an `Equipment` associated with it of how it was measured, e.g. 'Heron Dipper', for simplicity the actual device is only described via SensorML_`Components`	`Station`, CLIDB only separates between `Measure` and `Station`, `Procedure` is mapped to `Station` according to 'OGC SOS Lightweight Station Profile', the different sensors (raingauge, anemometer etc), are described as SensorML_`Components`
SOS/`OM_FeatureOf-Interest`	`Feature`, the site feature entity where the measurements/samples have been taken, incl. name, identifier and spatial coordinates	`Station`, the CLIDB station entity contains name, identifier as well as spatial coordinates
SOS/`Offering`	`Parameter`, the offering collections are modelled as spatially distributed measurements of the same observed property	`Station`, for CLIDB the offering is the collection of all observations for each single station
Result Values	`OM_MeasurementType`, for EODP and SOS Hydrology Profile WaterML2.0 encoding is also supported	Intended default encoding is `SWE_DataArrayType`, for EODP and SOS Hydrology Profile WaterML2.0 encoding is also supported

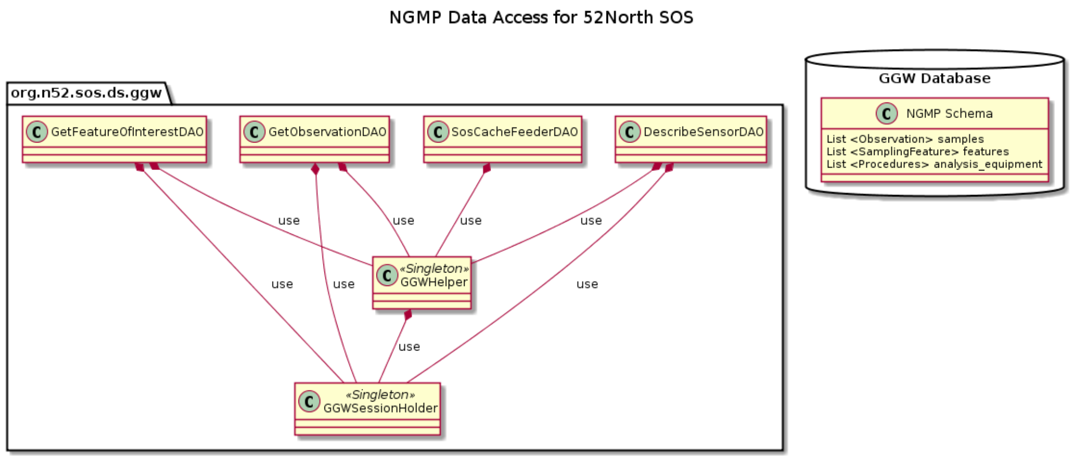


Figure 26: UML class diagram for Implementation of NGMP Data Access for 52N SOS, the GGW Database is queried via SQL, the SQL session is managed by the `GGWSessionHolder`

Figure 26 conceptualises how the 52N SOS server data model for its native database schema and its data access objects (DAOs) have been extended to fit the specific data models of NGMP and CLIDB and map their respective data models into the OGC SOS/O&M data model.

As a result of this software extension NGMP-SOS and the CLIDB-SOS can now deliver time series and observation data to the SDI. The reference web interface provided data viewing capabilities. Data was presented in traditional 2D web maps with sampling feature locations and related time series data graphs, as shown in Figure 27. The generated SensorML for the water level sampling protocol representation of the NGMP parameter '1679' (Static Water Level) is provided as an example in Appendix 21.

Inter-service Link from Sampling Locations to Observations

The clarity of the discovery process for environmental time series data might support or exacerbate the use of otherwise accessible data in a SDI. One approach has been documented in the EODP specification Ritchie, Schmidt, et al. (2014). For a simplified discovery process sampling features are publish through WFS. Related observation metadata, e.g. about the environmental time series data that was collected at these sampling locations is listed for each sampling feature. The actual result values in the observation entities are withheld, because the WFS is only intended for discovery via the spatial context of where observation data was collected with a 'step-by-step' discovery. The encoded related `OM_Observation` does not necessarily describe single ob-

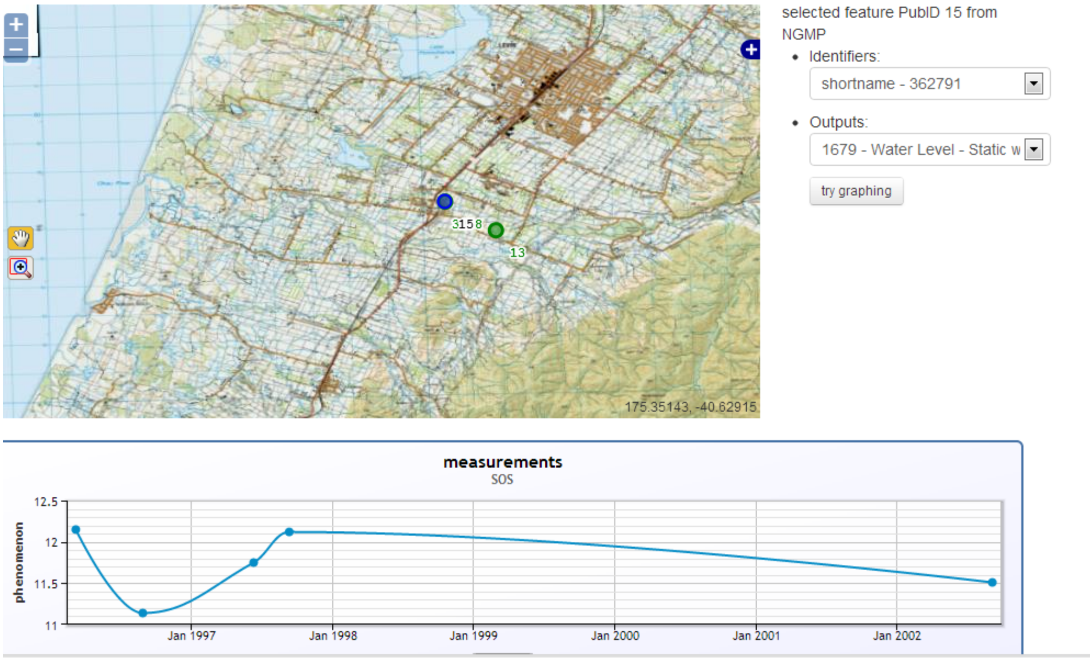


Figure 27: A view of groundwater observation data in the Horowhenua area in the Manawatu-Wanganui region as accessed from the portal via the NGMP-SOS service

servations or measurement, but rather a summary over a time period where appropriate.

The instance documents for WFS `SF_SpatialSamplingFeature` and described service linkage via `OM2:parameter:NamedValue` property from WFS to SOS is shown in Listing 1. The transformation was performed using the Geoserver 'app-schema' extension to publish the sampling metadata (i.e. sampling locations, observed properties, time period for measurements and procedure/equipment information) as `OM_Observations` from the NGMP database. A fully encoded document is provided as an exemplar in Appendix 19.

```
<om:parameter>
  <om:NamedValue>
    <om:name xlink:arcrole="observationdata" xlink:href="http://eodp.domain.nz/def/param-name/eodp/1.0/relatedSOSendpoint" xlink:role="relatedSOSendpoint"/>
    <om:value xlink:arcrole="protocol" xlink:href="http://portal.smart-project.info/sos-smart/sos/kvp" xlink:role="OGC:SOS-2.0"/>
  </om:NamedValue>
</om:parameter>
```

Listing 1: Service linkage via `parameter` property from WFS to SOS

Geoserver allowed querying the sampling features for related observations by time instants that intersect the summary time period of the described observation. However, a limitation of Geoserver (v2.6) at the time of testing was that it did not support querying via overlapping time periods.

MQTT Transport and Semantics for Sensor Web Integration

A specific challenge of environmental sensor networks is their requirement for low energy consumption and often the constraint of low uplink bandwidth. Also, radio or wireless transmission is a very energy consuming action. MQTT is a widely recommended protocol for application cases where network bandwidth and energy consumption for data transmission is constrained. These conditions are also true for current Internet of Things (IoT) applications (Jazayeri et al., 2015). Whilst classical OGC data interface standards depend on the OpenGIS®Web Services (OWS) Common specifications and are predominantly web-based with an HTTP transport model implied, OGC data encodings and semantics are not transferable one to one to an MQTT implementation.

Ghobaklou, Kmoch, Sallis (2013) and Klug, Kmoch, Reichel (2015) proposed a 'SOS/SPS-on- MQTT interlink to SDI' and modelled the original SOS and SPS semantics on MQTT. Figure 28 illustrates the bi-directional translation from high-level SWE SOS/SPS to low-level sensor network. It can be seen how standard OGC SOS/SPS HTTP interface outlets are kept for SDI integration, instead of introducing a new HTTP interface for observations and tasking. Nevertheless, MQTT sensor web linkage is converging towards a standardised means of SDI data integration with low level sensor networks.

In a similar approach, OGC SWE activities have produced the SensorThings API specification, which is an IoT sensor communication protocol, which roughly implements O&M semantics, and also provides an MQTT extension. Its design is separated in to the parts 1) Sensing, 'similar to SOS' and 2) Tasking, 'similar to SPS'. The data handling in the SensorThings API exposes a preference for data streams rather than single rich metadata observations. A tasking API has not yet been designed in the SensorThings API, but a current SPS-oriented tasking approach is presented here in this research.

The implemented infrastructure is explained in Figure 29. Both WSNs regularly push environmental data into the SOS via the shared MQTT queue (1, 2). A regular time based threshold process, implemented as a WPS process, retrieves the latest rainfall observations (3). The filtering process accumulates over a given time period, and triggers a reconfiguration event in the SPS when a threshold is met (4). The SPS issues a tasking message on the message queue, which the WSNs field computer consumes and then reconfigures (6) the phos-

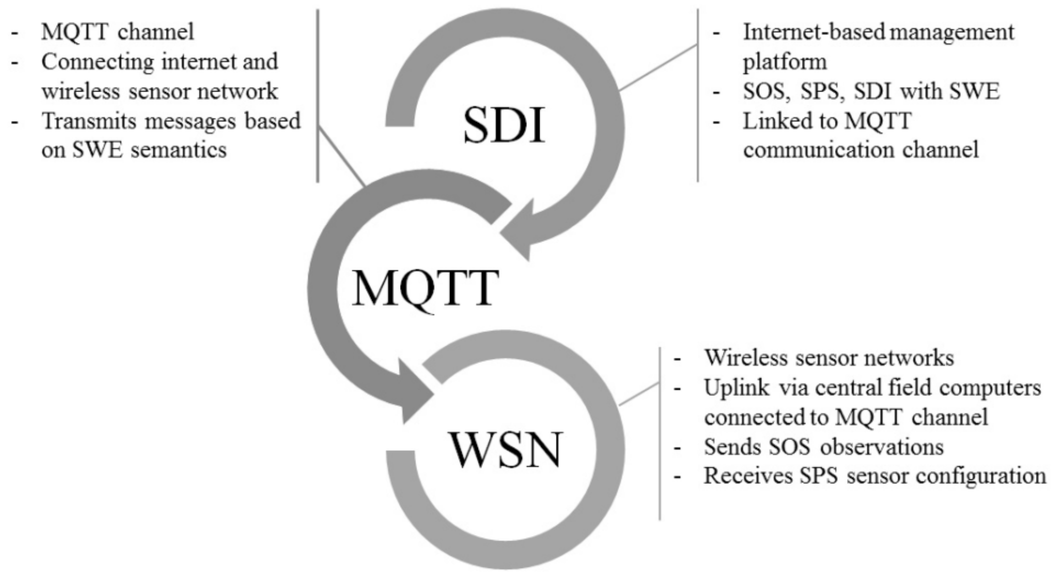


Figure 28: Bi-directional translation from high-level SWE SOS/SPS to low-level sensor network via MQTT linkage, source: own illustration from Klug et al. (2015)

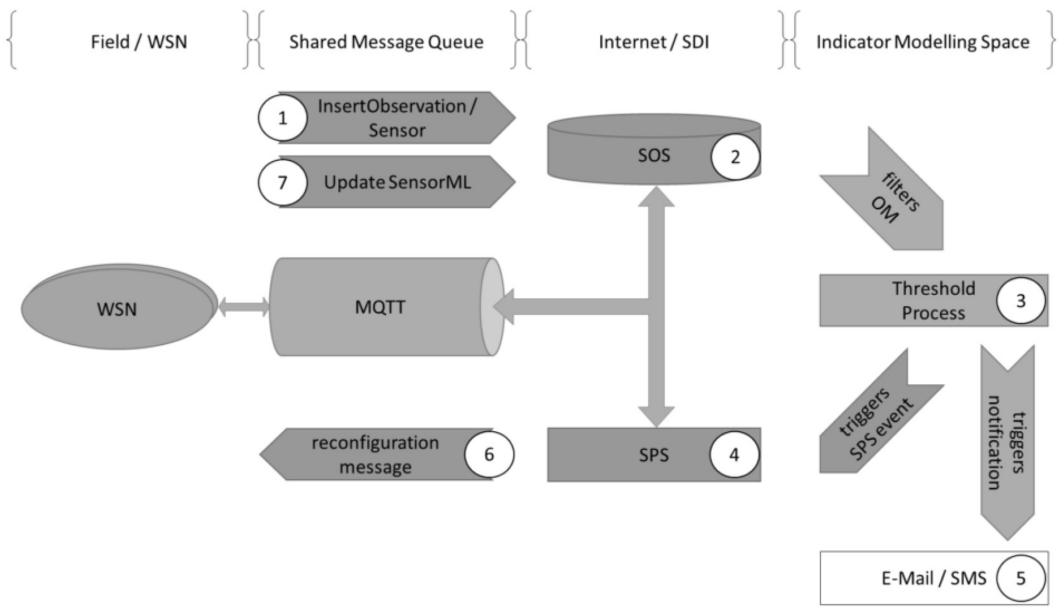


Figure 29: Reconfiguration message passing flow between high-level SWE SOS/SPS to low-level sensor network via MQTT linkage, the threshold process is realised as a WPS process that aggregates SOS Observations and triggers SPS reconfiguration for the sensor network if a threshold was exceeded, e.g. too high mm rainfall in last hour, source: own illustration from Klug et al. (2015)

phorus measuring device. The field computer updates the SensorML to reflect the configuration change (7). Simultaneously, the time based process (3) notifies decision makers via e-mail or SMS (5) about the rainfall event and the configuration change.

An exemplary SPS tasking message encoded in JSON which was transmitted as result of the threshold process to the field computer via MQTT is shown in Listing 2

```
// topic: sps/submitTask/sensorweb/admin/outbox/gateway0013A20040B5B303
{
  "messageUUID" : "f6fd0652-7a1a-4309-82bd-0f3f8d117a11",
  "task": "setWizMode",
  "parameter": {
    "mode": "LOW"
  }
}
```

Listing 2: Example SPS Tasking JSON on MQTT from Jenkins

When the data logger software process on the field computer receives the SPS tasking message via its subscription on the MQTT link, it issues an acknowledgement and initiates the sensor configuration on the low-level WSN. Subsequently, the sensor description of the configured sensor is updated in the SDI, in particular at its origin, which is the SensorML in the corresponding SOS server. There it becomes visible and accessible through the SDI. Listing 3 shows the updated SensorML elements for the measurement frequency for the `UpdateSensor` operation request.

```
<sml:capabilities name="measurementFrequency">
  <swe:SimpleDataRecord>
    <swe:field name="resolution">
      <swe:Text definition="urn:ogc:def:property:OGC:measuremode">
        <swe:value>LOW</swe:value>
      </swe:Text>
    </swe:field>
  </swe:SimpleDataRecord>
</sml:capabilities>
```

Listing 3: SensorML Capabilities of 'tasked' procedure from LOW to HIGH

6.2 METADATA MANAGEMENT

The overall implementation and application of the metadata design theory aspects highlight the integrative and linking aspects. Firstly, the ISO/ANZLIC metadata standard and encoding was adopted and metadata records created. An example XML ISO metadata record for rainfall recharge data set produced in the SMART programme is listed in Appendix 22. For the portal client capabilities where added that allowed creating standard conforming metadata records. The creation process is implemented as a guided web form, that addresses the following simple questions, What, Where, When, Who and How, which can be easily translated to the required basis elements to satisfy the ANZLIC/ISO metadata standards:

What?

- Title
- Keywords
- Abstract
- Topic (ANZLIC/ISO Category), e.g. `InlandWaters`, `Environment`, `GeoscientificInformation`
- Type of Resource, e.g. data set, service, sensor or model

Where?

- Geographical Scale
- Location Description
- Geographic or Projected Reference System
- Geographical Extent

When?

- Dates of Creation, Publication, or Revision of the Resource
- Lineage Information of the Resource
- Temporal Extent of the Resource

Who?

- Name of Contact Person for the Resource
- Phone number of the Contact Person
- Email Address of the Contact Person
- The Role of the Person in Relation to the Resource

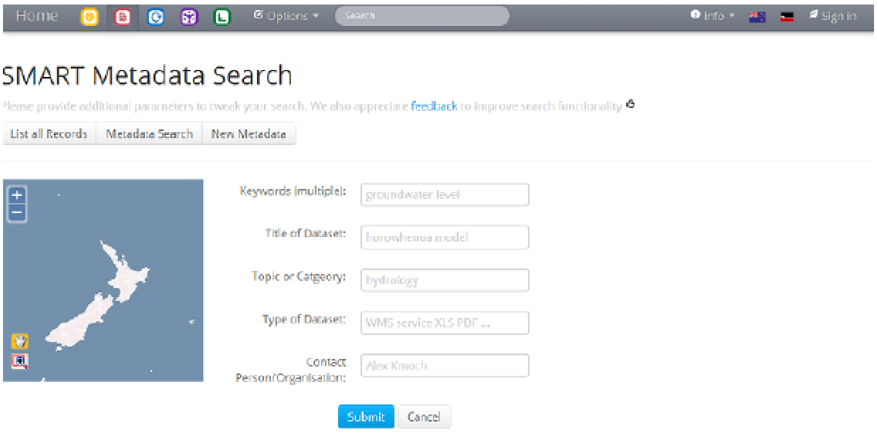


Figure 30: Portal Metadata Distributed Search

- Organisation (and/or Position) of the Contact Person
- A Weblink (Uniform Resource Locator (URL)) for the Organisation

How?

- License or other Constraints
- Type of Distribution Format
- Distribution Link

Through the portal client a user can now create and retrieve metadata records through a dedicated CSW catalogue server. As described in the Case Studies, many data producing, collecting or curating organisations in New Zealand publish metadata records for their data sets. Thus, the portal client was improved to allow querying many distributed CSW servers from the web front end. Figure 30 shows how a simplified query form was implemented. A slippy map on the left side shows the applied spatial bounding box, it can be zoomed and panned around to adjust the desired spatial context for the search. The generated search query was distributed to the registered CSW catalogue servers and the results collated in a list. For each returned record the portal client then summarises the metadata according to the simple questions 'What, Where, When, Who and How'.

An exemplary instance of an OWC document that links a WMS data source and corresponding CSW metadata record together as two offerings for one OWC entry is listed in Appendix 23. Such context documents could then be used by the portal client web mapping application. In the legend/layer panel of the web map where the currently loaded layers have entries, the corresponding catalogue entry could now be linked.

Section 5.2.3 describes an approach to make data discoverable through ISO/ANZLIC metadata records, which can be searched for in CSW-enabled catalogues. Table 8 (p. 76), illustrates that there are many unstructured data sets, for example reports and other written scientific documents, that do not exhibit any spatial reference. The Journal of Hydrology (New Zealand, Section 4.2) hosts a website where all articles from 1962 to today are accessible with author, title and abstract. The full text articles are available in all volumes inclusively until Volume 45 (until 2006). But they can not be searched via a spatial query.

Through an automated processing approach 607 actual single article html pages were extracted that uniquely identified and linked to a publication in the journal. However, data quality regarding authorship on the websites was not always consistent. Sometimes ',' and sometimes ';' were used as separators and author names abbreviations with '.' or without with an inconsistent usage of ' ' (space).

The journal features many editorials, news, book review or other types of written artefacts that were considered not to qualify as research articles. Based on titles and abstracts containing following or similar words or sentences have been removed from the process. These 'Stop words' have been selected and improved over the course of the analysis, e.g.:

- 'Book Review', 'Book Reviews', 'Book reviews', 'Editorial', 'Foreword',
- 'Journal of Hydrology (NZ) 5 Year Index Volumes 36 - 40',
- 'Metrication in scientific publications', 'Invited Editorial',
- 'Presidential Address', 'Forthcoming Events',
- 'Notices', 'Notes', 'Abstracts', 'Reports', 'Letters To The Editor', 'IHD Bulletin', 'Errata', 'Notice',
- 'Book Review', 'Publications received', 'Forthcoming meetings',
- 'New Publications and Forthcoming Meetings And Courses',
- 'IHD Bulletin New Zealand',
- 'No abstract', 'No Abstract', 'No abstract available.'

This resulted in 372 records from which 306 were freely available full text PDF articles. Those were downloaded and an algorithm applied that searched for place name occurrences separately in title, abstract and full text elements of the articles. The place names list used is retrieved from the Land Information New Zealand official gazetteer[10]. The gazetteer list contained 50510 geo-referenced names. LINZ provides a version with and without macrons in the place names. To simplify plain text matching the version without macrons was used. The gazetteer is implemented as simple features WFS with the following data scheme:

- ID: '15040'
- name: '15 Mile Creek'
- status: 'Official Approved'
- region: 'Nelson'
- projection: 'NZTM'
- northing: '5483525.2'
- easting: '1559021.0'
- geodetic datum: 'NZGD2000'
- latitude: '-40.79825'
- longitude: '172.514222'

A direct text-matching strategy was implemented over the list of used articles. For each element in the place names list, the search looked for a direct match in the articles' titles, abstracts and full text bodies. This implementation revealed reliability limitations because place names like 'Og' or 'Tor would be found in other compound place names like 'Bogs' or 'Tractor', or cause ambiguities like for 'Waikato' in 'Point Waikato', 'Waikato Region', or 'Waikato River'. Further development for a fully automated geo-referencing would require a thorough analysis of how large the error rate was. This has not been conducted in this study. Instead, a semi-automatic approach was chosen to be of greater immediate benefit, that would find the place names of a text and suggest that list of place names to the user at the stage of metadata creation. The generated metadata is a superset of the custodian's metadata which only had very few fields. The new metadata has many more fields that include the few former fields. There was no other reconciliation necessary with the metadata supplied by the data custodian. Thus the specific differentiations

10 LINZ, http://data.linz.govt.nz/#/search/category/gazetteer/, Gazetteer Names downloaded from LINZ 04.11.2014

of the described generated metadata records in comparison to the existing metadata on the website of the New Zealand Journal of Hydrology were:

1. the application of the extended, but standardised XML format of ISO/ANZLIC 19139 - Appendix 24 shows a complete metadata record for an article of the Journal of Hydrology (New Zealand);
2. the enablement of a spatial search through the additional, generated geographical bounding box based on the found location references;
3. and the interoperable provisioning via the CSW web service API

Furthermore, to support keywords based search as well as a web accessible and machine readable glossary by itself, the Freshwater Lexicon was translated into a simple RDF/SKOS vocabulary and published via a SPARQL endpoint. The concept identifiers were compliant to the URI scheme as proposed in the Linked Data literature, and for demonstration purposes made resolvable via HTTP through their URI identifier.Appendix 25 show the full encoding. In Figure 31 is shown how the portal client interacts with the SPARQL endpoint to list the concepts, i.e. the glossary terms.

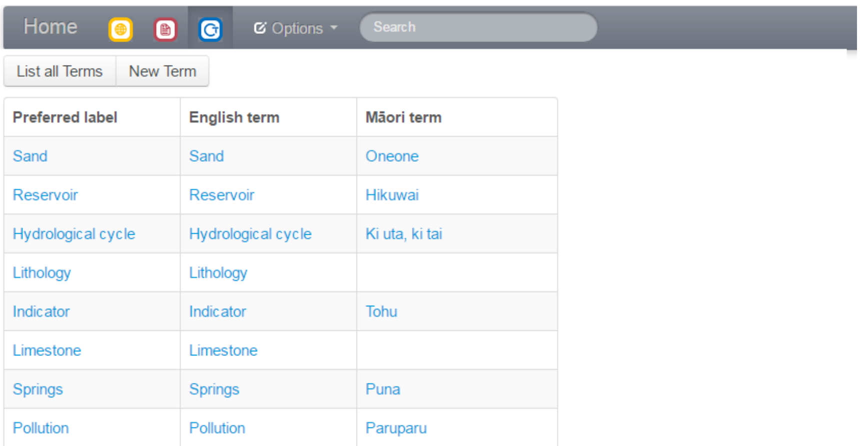


Figure 31: Portal NZHS Glossary

In the same fashion, the observed properties, i.e. the 'parameters', of the NGMP database where encoded and published, too. Listing 4 shows the concept encoding for the 'Water Level' property including the metadata from the database. Based on Linked Data concepts, the concept identifier is a URI, which can be resolved via HTTP. The portal client acts as the Linked Data proxy the SPARQL endpoint and issues a query to retrieve the graph properties that are related to the URI concept.

```
<skos:Concept rdf:about="http://vocab.smart-project.info/ngmp/phenomenon/1679">
  <skos:label>Water Level - Static water level m</skos:label>
  <dc:title>Water Level - Static water level m</dc:title>
  <skos:prefLabel xml:lang="en">Water Level - Static water level m</
    skos:prefLabel>
  <skos:definition xml:lang="en">[1679]: Water Level - Static water level m (in
    m)</skos:definition>
  <dc:description>[1679]: Water Level - Static water level m (in m)</
    dc:description>
  <dc:modified>2014-11-06T15:00:05.892+13:00</dc:modified>
  <skos:inCollection rdf:resource="http://vocab.smart-project.info/collection/
    ngmp/phenomena"/>
</skos:Concept>
```

Listing 4: NGMP Observed Property 1679 Standing Water Level as RDF/SKOS encoded

Figure 32 shows how the portal client renders the retrieved data as an HTML page. The full list of concepts and how they are embedded in a SKOS collection are shown in the Appendix 26. Analogously, the portal client would show the full collection as an HTML list as depicted for the glossary in Figure 31.

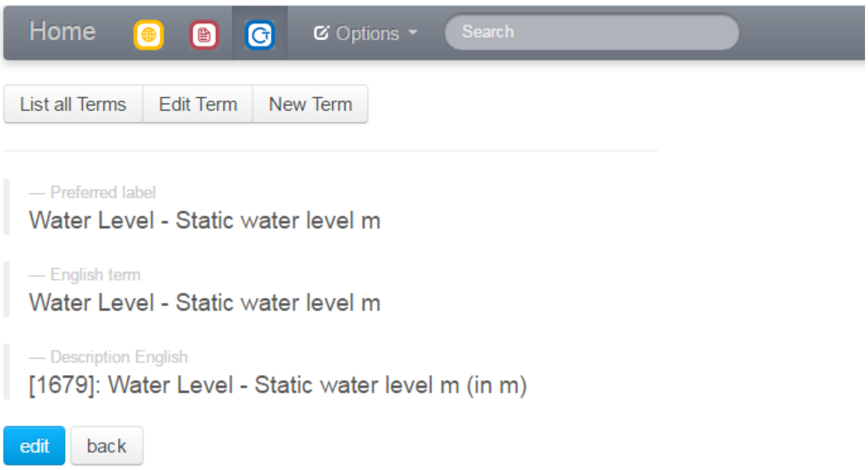


Figure 32: Portal NGMP Observed Property URI Vocab resolved http://vocab.smart-project.info /ngmp /phenomenon/1679

This has the advantage that a client, which does not know where a specific vocabulary concept entry would be maintained for an identifier, could just try to load the URI as a HTTP GET URL. Based on HTTP content negotiation a result may be returned that the client understands. Thus, using resolvable URIs as identifiers eases the immediate interlinkage based on HTTP infrastructure.

6.3 enabling the model web through geoprocessing

The geological and water budget assessments for the Horowhenua area in southern Manawatu-Wanganui region of the lower North Island of New Zealand serve as case studies to build an experimental, internet-enabled interactive groundwater model based on MODFLOW. The distributed data and processing services can be linked to prepare an on-demand analysis and exploration of hydrogeological data to support the characterisation of the Ohau and Waikawa groundwater aquifers (subcatchments in the Horowhenua district case study)

To create the necessary input grids for MODFLOW GeoTIFF raster layers from a New Zealand elevation model as well as from geological stratigraphic layers were used to defined the model dimensions. Discrete observation data for groundwater levels was rasterised and translated into basic model parameters like hydraulic head and conductivity. The sources for these data sets were OGC web services, and could also be mapped e.g. in GIS tools or web maps. Subsequently, mediation routines processed MODFLOW outputs and promoted the data to the higher OGC groundwater SDI processing layer which allowed presentation and storing of the results in the web data sources again. Figure 33 shows the input mapped from the web service data sources in QGIS, which highlights the advantage of open interfaces. Furthermore, GWML2 is an OGC GML application schema that describes groundwater related feature data types such as aquifers, wells and their inherent properties. To visualise the spatio-temporal distribution, level and flow of groundwater, observations and feature (rainfall, well and river flow measurements, aquifer properties, geology) data and are exposed to an OGC WPS and thus can be integrated into web based modelling systems.

A very important parameter for a groundwater flow model, amongst others, is the hydraulic conductivity of the rock or earth material. Defining aquifer properties over the catchment area for the MODFLOW data inputs could be done with generalising values from feature data from GWML2 aquifer unit descriptions and related attributes. But specifically defining a conductivity and other aquifer property values is mostly a manual step which needs to be done by an experienced groundwater scientist. Thus, the actual flow model was not scientifically viable and only served as a technical implementation example.

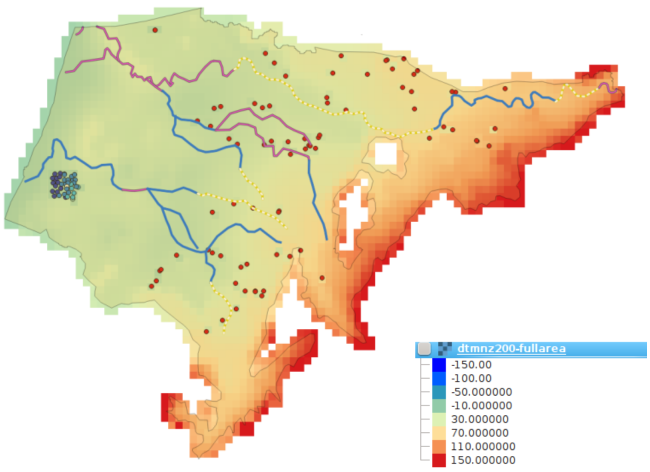


Figure 33: Assumptions and very simple model with a one layered model grid, groundwater levels as `OM_Observations` from the SDI as boundary conditions and catchment as model boundary (Ohau and Waikawa catchments, South Horowhenua).

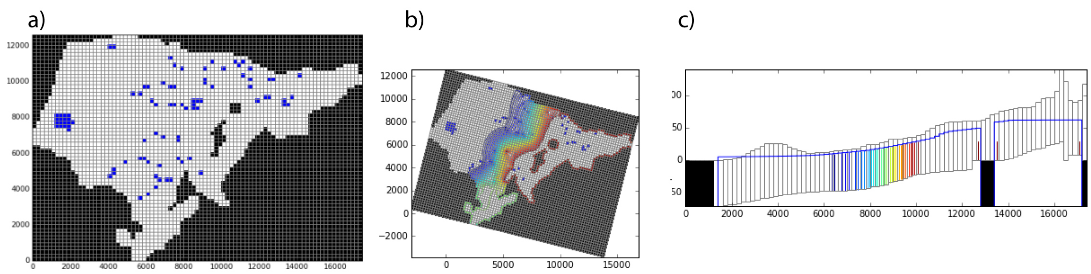


Figure 34: Plots of MODFLOW model prepared with SDI data for Ohau and Waikawa catchments (South Horowhenua), a) plot of boundary conditions (groundwater level measurements), b) simulated yield array for active zone, c) central cross-section from west to east

WPS Configuration and MODFLOW Integration

The WPS process input fields are satisfied with the linkage to OGC data sources. After MODFLOW has been successfully run exemplary grids are plotted and returned as result as shown in Figure 34, the MODFLOW grid with hydraulic conditions and a cross-section.

The WPS-enabled MODFLOW software module has the capability to run pre-configured and calibrated groundwater models with latest available observation data from the groundwater SDI. Furthermore, the interoperable groundwater model was linked with other WPS processes, and thus created the foundation of a so called ’model web’.

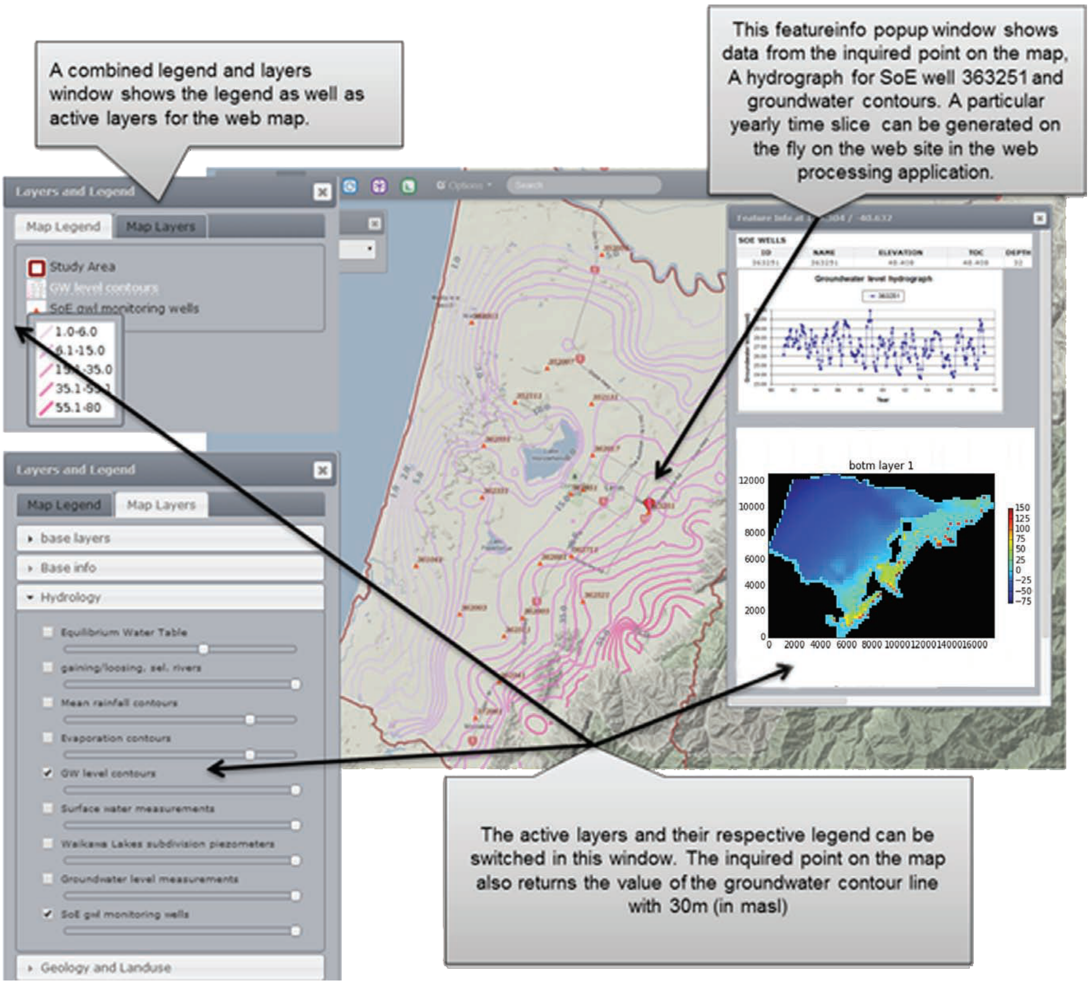


Figure 35: Web map viewing application displays 2D representations of geospatial groundwater data layers for the Horowhenua area via a variety of OGC compliant data sources, incl. resulting data sets from geoprocessing

A web map viewing application (Figure 35) displays 2D representations of geospatial groundwater data layers via OGC compliant WMS, WFS and WCS layers from remote servers and graphs for hydrological measurement data from SOS.

6.4 development of 3d scene generation as wps process

This section describes the implementation of the on-demand 3D visualisation of hydrogeological data in the web. A framework process manages the interaction between data and processing services. In Kmoch and Klug (2014) successfully trialled and tested WMC to demonstrate the integration of spatio-temporal observation data via SOS into a web mapping context. The initial message to the framework process is the request for a 3D view of a map context, which is defined via a OWC document.

Such an OWC context document was then forwarded to a specifically developed WPS process. This WPS process derived basic layout information from the context document, like the spatial bounding box. Then the process iterated over the entries of the context document and based on the type of the entry it retrieved the referenced data set and built the required X3D scene elements. In OWC context documents also WPS process requests can be used as entries.

A 52N WPS has been implemented to run the X3D scene generation from a OWC (and formerly trialled WMC) input document (30). The Java files in Appendix listings 27, 28 and 29 show exemplary the classes and implemented interfaces, needed to 1) runs a WPS process within the 52N WPS software package, 2) parse the WMC/OWC document and retrieve layers, geographic bounding box and geographic reference system to 3) build the complete scene layout. Appendix Figure 43 also shows a UML class diagram of the program.

For OWC documents in fact only a different input document parser needed to be implemented. Furthermore, X3D is enriched with a variety of metadata elements, so that the scene contains valuable information about its data origins. Excerpts from the generated scene are shown in Appendix 31. Table 13 (p. 116.) illustrates which OGC service and data elements are mapped and transformed to X3D elements. Figure 36 shows a sequence diagram of the implementation. The algorithm checks the type of each source and then generates the necessary X3D elements/nodes based on Table 13.

The OWC document is used to define the data sources, which were subsequently inquired and the returned data transformed into the respective X3D

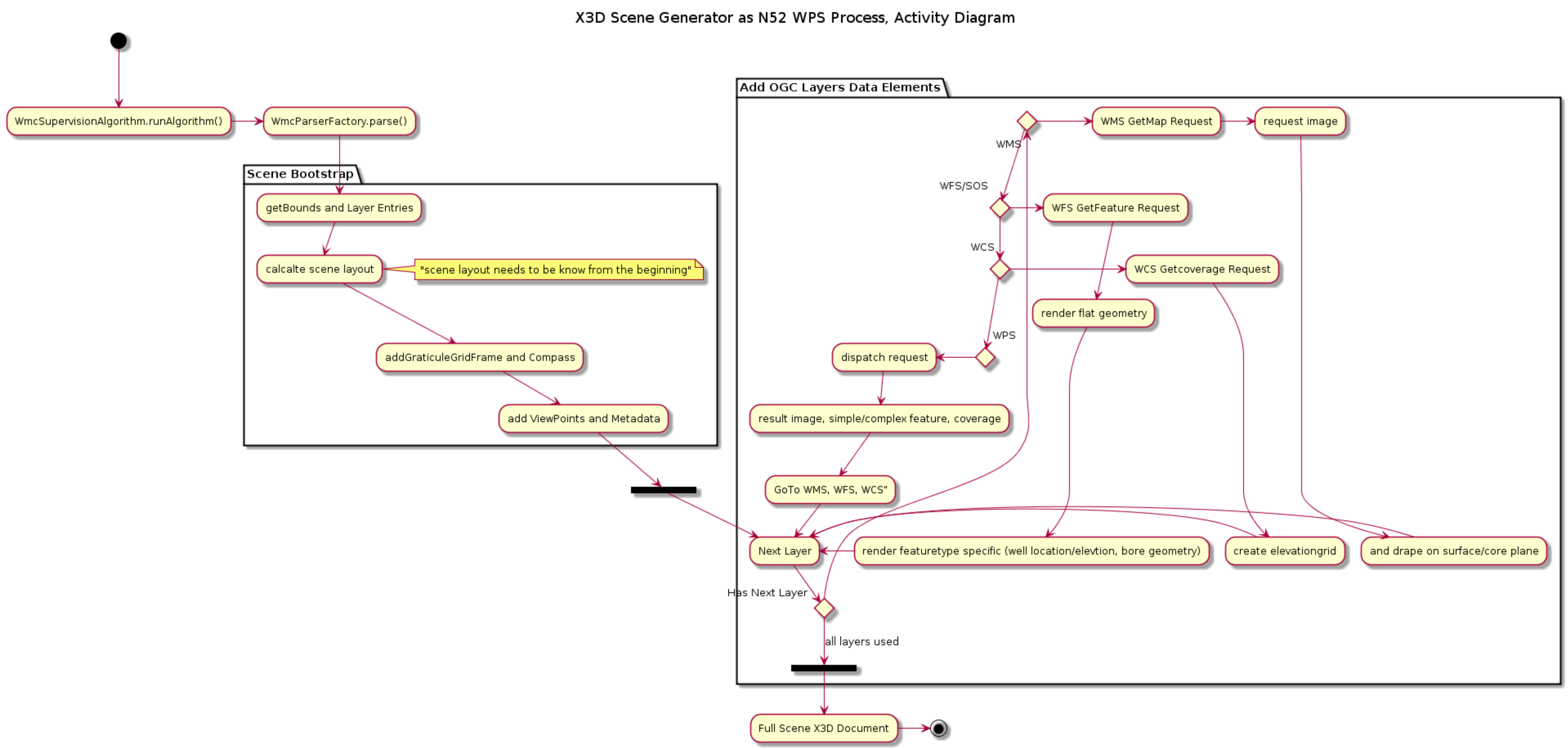


Figure 36: UML Activity Diagramm for Implementation of the X3D scene generation as 52N WPS algorithm/process

elements. The portal controller initiates this project on request for a view by a user. The controller takes the current map context and sends it to the WPS process (Appendix 32). The generated scene is then loaded from the portal controller into a view in the portal (Appendix 33) and rendered via the 'x3dom' toolkit with WebGL support into an interactive scene in the browser. The following images show different instantiations of X3D scenes generated from different contexts rendered on the reference client website:

- Figure 37 - a generated X3D scene from the OWC context via x3dom in the browser, basic DEM surface, and WMS images draped on top, a graticule grid shows dimensions and directions, based on the bounding box and coordinate system from the context document
- Figure 38 - a generated X3D scene with basic DEM surface, and WMS images draped on top, a graticule grid and wells (well-representing features from WFS), generated as columns, they referenced via their top elevation and length/depth attribute
- Figure 39 - a generated X3D scene with basic DEM surface (green), and surfaces of geological layers (yellow, orange and red); WCS coverages created from the EarthVision exports
- and Figure 40 - the generated X3D scene, basic DEM surface (green), and surface of an equilibrium water table (WCS GeoTIFF data set from SMART Satellite Remote Sensing

Three-dimensional web visualisation capabilities for environmental data in New Zealand are presently limited and established 3D models are based on cost-intensive proprietary desktop software solutions. With this 3D visualisation method founded on distributed data sources and a scene creation service insights into the subsurface and hence available (ground) water resources should be enabled. This integrated and visualised information is not only useful for supporting policy making, but also provides a common ground for competitive groundwater management and use. The integration of ground truth time series from sensor stations helps characterising the input and output relations of the catchment in near real-time.

3D information, particular high resolution surfaces and raster tend to be problematic in lightweight environments like web browsers. Thus, low-resolution image requests were built in the image draping step of the scene generation. The scene creation process samples the overall scene size down,

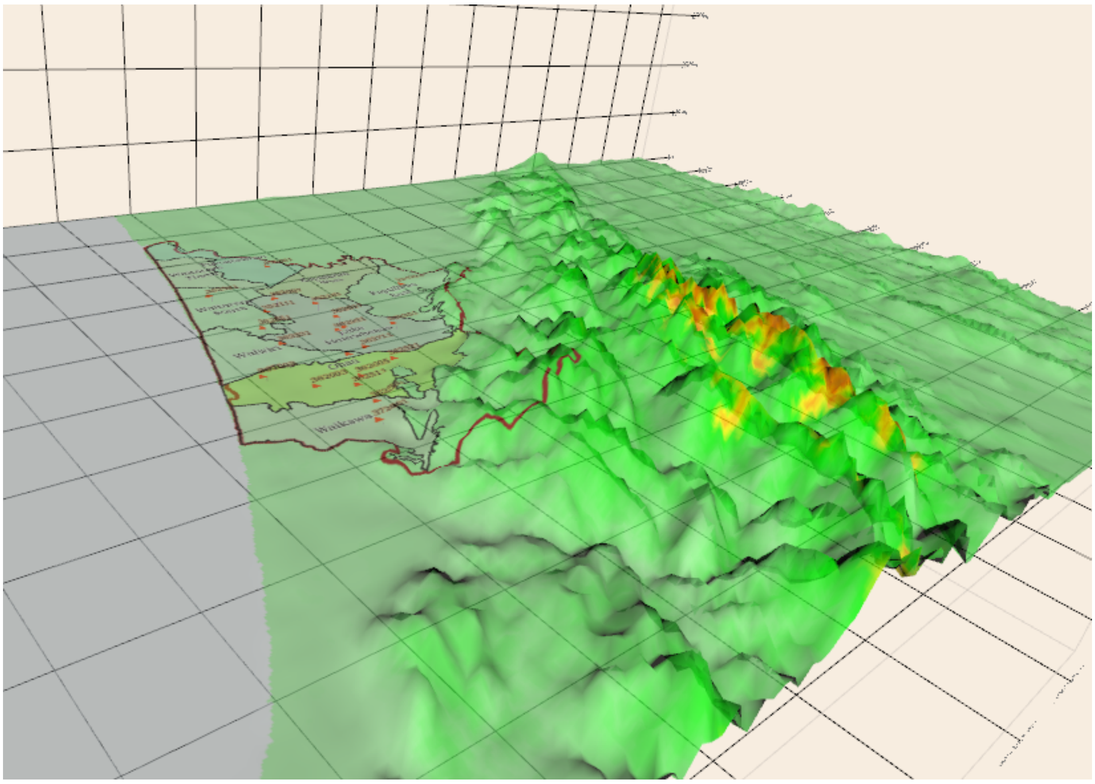


Figure 37: Groundwater Portal 3D Viewer Module rendering the generated X3D scene from the OWC context via x3dom in the browser, basic DEM surface, and WMS images draped on top, a graticule grid shows dimensions and directions, based on the bounding box and coordinate system from the context document

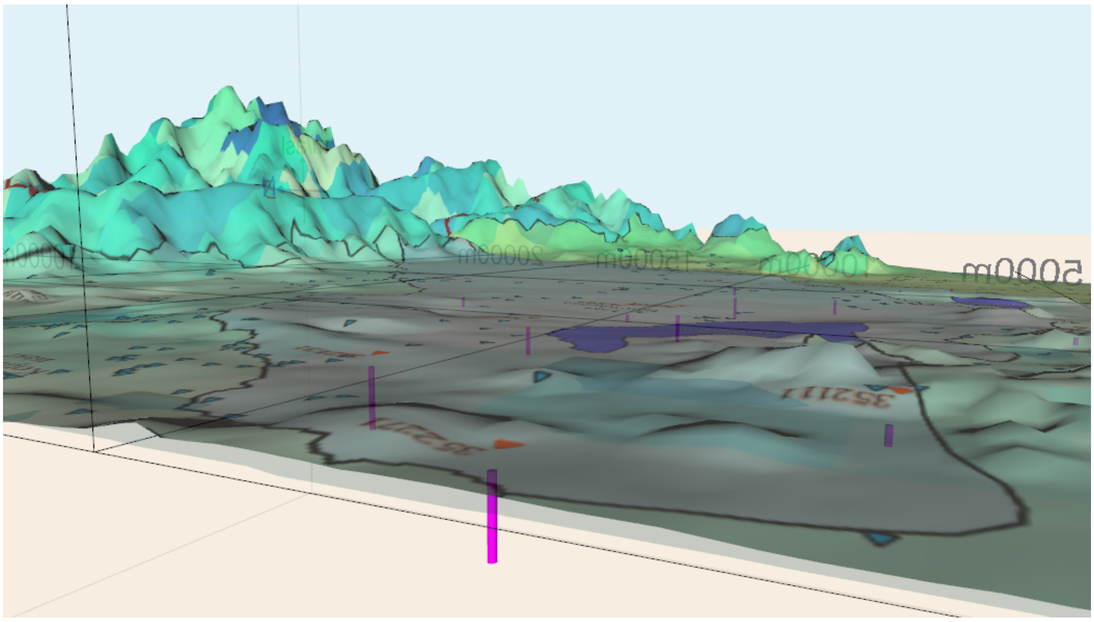


Figure 38: Groundwater Portal 3D Viewer Module, zoomed into the generated X3D scene with basic DEM surface, and WMS images draped on top, a graticule grid and wells (well-representing features from WFS) in purple, generated as columns, referenced via their top elevation and length/depth attribute

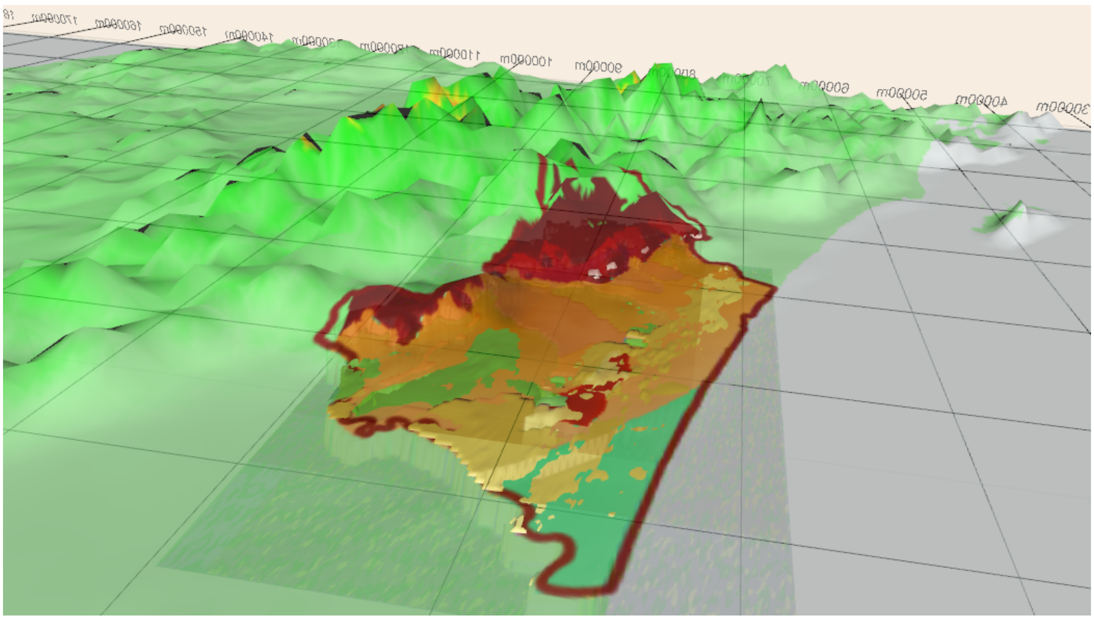


Figure 39: Groundwater Portal 3D Viewer Module rendering, basic DEM surface (green), and surfaces of geological layers (yellow, orange and red; WCS coverages created from the EarthVision exports

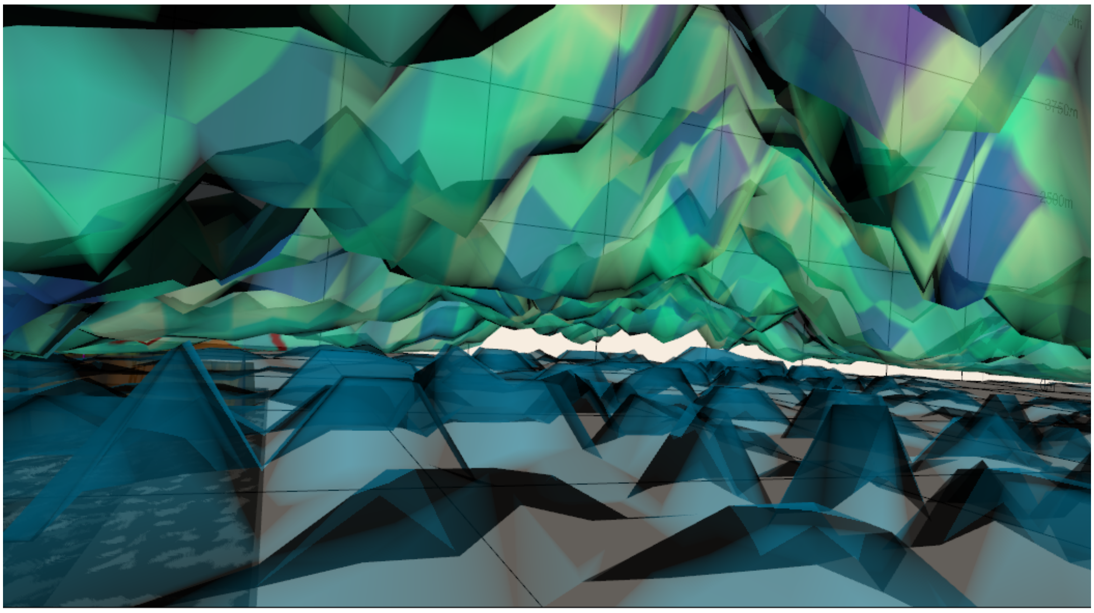


Figure 40: Groundwater Portal 3D Viewer Module, zoomed into the generated X3D scene, basic DEM surface (green), and surface of an equilibrium water table (WCS GeoTIFF data set from SMART Satellite Remote Sensing

so that grids and graticules and therefore image textures do not extend over 1000 pixels in side length.

Although X3D supports Level of Detail (LOD) in general, the dynamic creation in the presented project did not allow for pre-rendering and 'depositing' tiles on public HTTP accessible web space. Thus, to achieve scenes with different element resolution level, a new scene generation sequence is currently required.

6.5 FRAMEWORK EVALUATION

The implementation applies the research method to the designated use cases and evaluates the developed framework in an experimental set-up. By qualitative comparative analysis information loss through multiple technical and semantic transformations with typical processing chains in the use cases is examined.

The following three lists recapitulate the case studies with their main groups of data set types and services. The abbreviations are be used in the coverage matrix (Figure 41, p. 150) in order to gain an understanding of how many cases the implementation covered successfully, partially successfully or unsuccessfully (includes no implementation, too) from the case studies. Subsequently, the implementation is tested against the developed evaluation criteria from Table 2 (p. 60), and summarised in the evaluation matrix shown in Figure 42 (p. 152).

Horowhenua local case study data set types:

- CS:HORO GEOST - Table 3 (p. 73): Data sets describing geological units/structures, elevation and surfaces
- CS:HORO GEOBS - Table 4 (p. 74): Data sets containing well descriptions, borelogs, lithology observations
- CS:HORO HYMG - Table 5 (p. 74): Data sets about catchments and management zones
- CS:HORO HYOBS - Table 6 (p. 75): Data sets from hydrological and hydrogeological monitoring sites and related measurements
- CS:HORO MOD - Table 7 (p. 76): Data sets from hydrological and hydrogeological numerical model results/outputs
- CS:HORO RPRT - Table 8 (p. 76): Data sets solely for map visualisation and unstructured data set like reports and images

New Zealand Environmental Information Services case study, from Chapter 4.2 p. 79:

- CS:NZ LINZ - LINZ Data Service
- CS:NZ LRIS - Landcare LRIS data portal
- CS:NZ NEIB - NIWA Environmental Information Browser (EIB)

- CS:NZ GEOVT - LINZ public geodata catalogue
- CS:NZ DOC - DoC Geoportal
- CS:NZ CLIDB - NIWA CLIDB SOS
- CS:NZ NFLOW - NIWA Hydro and Flows archive Kisters SOS
- CS:NZ EWK - Waikato Regional Council Hydro Kisters SOS
- CS:NZ HHTOP - Horizons Regional Council Hilltop SOS Testbed
- CS:NZ GNSGEOL - GNS Geoserver with Geological Map
- CS:NZ GNSEBOF - GNS Earth Beneath Our Feet (EBOF)
- CS:NZ LAWA - LAWA Website
- CS:NZ JOHNZ - Journal of Hydrology (New Zealand)

SAC programme case study data sets and services, from Chapter 4.3 p. 83:

- CS:SAC EWT - Equilibrium Water Table data set
- CS:SAC AQ - NZ Aquifers (White 2011) data set
- CS:SAC GWZON - NZ Groundwater Zones Information data set (Lovett & Cameron, 2014)
- CS:SAC HALON - Halon Sampling data set
- CS:SAC NGMP - NGMP SOS Service
- CS:SAC URSEN - Upper Rangitaiki Sensor station and SOS service

The coverage matrix (Figure 41, p. 150) shows that generally most case studies - 21 out of 25, with 4 case studies only partially or not implemented at all - have been used in the implementation/evaluation cycles and that all the implementations been tested across the board. The highlights and conclusions from the implementation coverage are explained as follows:

1. The implementations for 'Groundwater SDI' and 'Metadata Management' have been most comprehensively tested. The reason is, that these methods and techniques are also already best understood and well described in the literature. Also the technological adoption in the hydrology domain in New Zealand is the most progressed here.

Case Studies	Groundwater SDI	Sensor Web Integration	Metadata Management	Geoprocessing Model Web	3D Geovisualisation Process	Sums
CS:HORO GEOST	x	-	x	x	x	4/0/1
CS:HORO GEOBS	x	partially	x	-	partially	2/2/1
CS:HORO HYMG	x	-	x	x	x	4/0/1
CS:HORO HYOBS	x	x	x	x	-	4/0/1
CS:HORO MOD	x	-	x	-	x	3/0/2
CS:HORO RPRT	x	-	x	-	x	3/0/2
CS:NZ LINZ	x	-	x	-	x	3/0/2
CS:NZ LRIS	x	-	x	-	x	3/0/2
CS:NZ NEIB	x	-	x	-	-	2/0/3
CS:NZ GEOVT	x	-	x	-	-	2/0/3
CS:NZ DOC	x	-	x	-	x	3/0/2
CS:NZ CLIDB	x	x	partially	x	-	3/1/1
CS:NZ NFLOW	partially	-	-	-	-	0/1/4
CS:NZ EWK	partially	-	-	-	-	0/1/4
CS:NZ HHTOP	partially	x	partially	-	-	1/2/2
CS:NZ GNSGEOL	x	-	x	partially	x	3/1/1
CS:NZ GNSEBOF	x	-	x	x	x	4/0/1
CS:NZ LAWA	partially	partially	-	-	-	0/2/3
CS:NZ JOHNZ	x	-	x	-	-	2/0/3
CS:SAC EWT	x	-	x	x	x	4/0/1
CS:SAC AQ	x	-	x	x	x	4/0/1
CS:SAC GWZON	x	-	x	-	-	2/0/3
CS:SAC HALON	x	-	x	-	x	3/0/2
CS:SAC NGMP	x	x	x	x	x	5/0/0
CS:SAC URSEN	x	x	x	x	-	4/0/1
Sums	21/4/0	5/2/18	20/2/3	9/1/15	13/1/11	

Legend:	Successfully implemented	partially successful	not successful, not implemented	high overall score	low overall score

Figure 41: Coverage matrix for the implementation over the case studies data sets and services

2. The case studies that have been tested across all or most implementations comprise of data sets from Horowhenua case study, the NIWA CLIDB, data sets and services from GNS and the SAC programme. In particular the NGMP database got major attention during the course of this study.

3. The 'Sensor Web Integration' did naturally not achieve high scores because of the high specificity in focus on hydro-climate and hydrogeological time-series data. However, case studies that were covered by the implementation scored above-average (with the exception of the Hilltop Testbed, c.f. below)

4. In reverse, the NIWA Hydro Flows archive Kisters SOS (CS:NZ NFLOW) and the Waikato Regional Council Hydro Kisters SOS (CS:NZ EWK) have not been tested in specific during the course of this study because of technical or administrative roadblocks from the providers' side in combination with lack of time. Thus, the implementations did not cover these services.

5. Horizons Regional Council Hilltop SOS Testbed (CS:NZ HHTOP) did not score high implementation coverage due to the experimental nature of the provider's service implementation. While OGC SOS and WFS support was proclaimed, the Hilltop server did not behave as expected compared to other SOS services in the case studies.

6. The coverage of successfully tested implementations for 'Geoprocessing Model Web' and '3D Geovisualisation Process' was average. This should still be interpreted as a good result as the innovation stretch towards new methods and techniques was the greatest here.

7. Eventually, LAWA was more a of a complementary client to the above listed hydrological observation data services. The website's managing team announced that LAWA was successfully loading hydrological data via OGC SOS in WaterML2 format across the regions.

The evaluation matrix (Figure 42, p. 152) shows that all evaluation criteria (c.f. Table 2, p. 60) have have scored positive in at least 1 or 2 out of the 5 implementations and that 4 out of 5 implementations have passed 16-18 out 20 criteria with a positive outcome. The generally very positive result of evaluating the criteria against all the implementations supports the unifying character of the theory and design for an interoperable distributed groundwater infrastructure, which the implemented artefacts strive to adhere to. The

Criterion ID	Groundwater SDI	Sensor Web Integration	Metadata Management	Geoprocessing Model Web	3D Geovisualisation Process	Sums
HYNZ1	x	x	x	x	x	5/0/0
HYNZ2	x	x	x	x	x	5/0/0
SDI1	x	x	x	x	partially	4/1/0
SDI2	x	partially	x	x	x	4/1/0
SDI3	x	x	x	x	partially	4/1/0
SDI4	-	x	-	x	partially	2/1/2
SDI5	x	x	x	x	x	5/0/0
SDI6	x	x	x	x	x	5/0/0
SWE1	x	x	x	x	x	5/0/0
SWE2	partially	x	-	-	-	1/1/3
MOD1	partially	x	x	partially	x	3/2/0
MOD2	x	x	x	x	x	5/0/0
MOD3	x	x	x	x	x	5/0/0
MOD4	partially	x	x	partially	-	2/2/1
MOD5	x	x	x	x	x	5/0/0
VIZ1	x	x	x	x	x	5/0/0
VIZ2	x	x	partially	x	x	4/1/0
VIZ3	x	partially	partially	partially	partially	1/4/0
VIZ4	x	x	x	x	partially	4/1/0
VIZ5	x	x	x	x	x	5/0/0
Sums	16/3/1	18/2/0	16/2/2	16/3/1	13/5/2	

Legend: criterion met | partially met | criterion not met | high overall score | low overall score

Figure 42: Evaluation results for the implementation, criteria applied from Table 2, p. 60, developed in the Methodology Chapter 3.4.

highlights and conclusions from the evaluation matrix are explained as follows:

1. Again, the slightly less successfully evaluated implementations for 'Geoprocessing Model Web' and especially '3D Geovisualisation Process' can be interpreted as result of the innovation stretch towards new methods and techniques that was the greatest here.

2. The 'Sensor Web Integration' did achieve high scores because of the high focus during the course of this study on the essential hydro-climate and hydrogeological time-series data. The importance of having up-to-data environmental monitoring data available was another driving factor. The Upper Rangitaiki sensor station in the SAC programme and the NGMP database got major attention during the course of this study.

3. Criterion SDI4 ('the design paradigm surpasses classic information systems design approaches') was not evaluated positively across all implementation, because 'Groundwater SDI' and 'Metadata Management' are already well understood and widely adopted methods. Here existing knowledge was leveraged but not developed further.

4. Criterion SWE2 ('data protocols are appropriate for low bandwidth unreliable network connections') did not score high because of the high specificity in focus on data transfer in remote and rural areas. This was really only tested in the 'Sensor Web Integration' implementation for the remote Upper Rangitaiki sensor station. Beyond these special use cases the standard OGC web-services and formats were used. Under a mobile device use case this criterion might be revisited for the other implementations. However, the typical use case in the course of this study was usage via a web browser from a networked computer.

5. Criterion MOD4 ('modelling workflow can be automated and operationalised from data capture via modelling and prepared information visualisation for decision making') was evaluated to be only partially successful. The main reason is that for groundwater modelling much testing and experimentation with the model by a groundwater scientist is requirement and thus a fully automated operationalisation was only attempted but a reasonable groundwater flow model was not successfully completed.

6. Criterion VIZ3 ('provides declarative translation from geoscience data to visual model') was also evaluated to be only partially successful. This

relates back to the implementation coverage for the '3D Geovisualisation Process' which demonstrated the concept of a declarative mapping of OGC service and data elements to X3D elements (c.f. Table 13 (p. 116). However, while gridded raster data formats like GeoTIFF can be used to represent a variety of distributions of environmental spatial distributions, GML application schema-based data transfer standards like GWML and GeoSciML each feature type distinctly. The translation has only been demonstrated for several main types like geological units, aquifers, wells and springs. This can be used as a guideline for the many other feature types. Furthermore, no temporal visualisations have been implemented and therefore the evaluation can only allow for a partial result.

6.6 SUMMARY

Based on the latest available international and national standards and the identified existing gaps in New Zealand, this service-oriented modular based concept has been realised in the SMART groundwater portal as shown in Figure 22, p. 121. Multiple interfaces provide the means to deliver, use and visualise hydrogeological data. The main processing, transport and mediation happens within the 'Data Interface and Services Layer', which is part of the back end application running on the web server.

Existing data providers can be registered and connected to the infrastructure via standardised OGC web services. The web interface, also referred to as the 'front end' or user interface, is the part of the service that is visible to the user. This interface presents the data from the 'Data Interface and Services Layer' and includes web applications providing data viewing capabilities, like traditional 2D web maps with sampling feature locations and related time series data graphs and attribute tables, as shown in Figure 27, p. 129.

Furthermore, the interface includes utilities to query, search, and discover registered data sets. Data is also directly accessible via the explicit OGC web service interfaces for use with GIS and other hydrological software applications that support OGC web services. Likewise, data sets and related metadata can be queried through the portal's CSW interface, which will distribute CSW GetRecords search queries across the major geoportals in New Zealand, including LINZ Data Service, geodata.govt.nz, NIWA Environmental Information Browser, Landcare LRIS Portal and DoC Geoportal and collate the search results with reference back to their original catalogue.

Part IV

DISCUSSION

7 DISCUSSION

This chapter discusses the developed grounded design theory for a 'Hydrogeology Infrastructure' and how the reference implementation addressed challenges and short comings from existing methods that were identified in the literature review. Subsequently, the chapter reflects on the novel GIScience research method developed as part of this research, and the degree to which the application of this method is suitable for designing interoperable data and analytics infrastructures for Earth Sciences domains.

The main goal of this research was to enable an holistic and integrated framework design for groundwater and groundwater-related data sets and models through advancing the understanding of interoperability for hydrogeology. This study aimed to identify and fill gaps, solve issues and remove ambiguities that exist in current (inter-)national standards, models and techniques. The focus was on shortcomings that might inhibit the use of such standards, models and techniques in a comprehensive interoperable hydrogeological data framework. The Grounded Design method was developed to examine which standards could be implemented and integrated in order to create a new data enriched perspective of New Zealand's groundwater resources.

7.1 IMPROVING GROUNDWATER DATA INTEROPERABILITY

Hydrogeological research in New Zealand would benefit from better data integration. The monetary benefit is hard to quantify, but from the literature review it was concluded that there is a need for improvement, and also what the areas for improvement were. Since the onset of this research the understanding for framing existing standards, file formats and internet technologies into a (re-)useable agreed-on set has increased in New Zealand. But tools and data resources for hydrogeological assessments are not yet well integrated, thus, hampering effective and efficient use of data, and consequently such assessments consume a large share of research time and budget.

In this study the topic was approached from a holistic, integrative GIScience perspective. Methods, standards, formats and encodings addressing interoperable geographic and environmental data exchange were sourced, analysed and critiqued, as part of a systematic literature review. Case studies, relevant applied technologies and software systems were also investigated. A grounded theory was developed from the case studies and compared to New Zealand groundwater data management practices and analysis tools. Subsequently, the theory was advanced into a grounded design for a hydrogeology infrastructure in reconciliation with the literature and based on the reviewed methods, standards, formats and technologies.

The case studies were essential as they gave immediate context to the developed theory. It was also helpful that this study was embedded in the SAC programme which provided a strong foundation in actual hydrogeological research practices in New Zealand. Participation in the international OGC GW2IE provided immediate access to the latest developments in groundwater data modelling and interoperable data sharing.

Through both, the thorough literature review and the participation in the GW2IE, the researcher gained an international view of the field in general. In the eye of the researcher, the practical understanding and application by governmental bodies of this type of interoperability in New Zealand, in particular in the field of hydrogeology is lacking behind, for example, Canada, Australia, the US, or Europe with its INSPIRE-related efforts. However, there is a strong, small community active in New Zealand, which instead of imposing new standards, fosters framing and adoption of existing standards into national profiles, e.g. EODP, to lower the entry level of usage.

The literature review of hydrogeology in New Zealand revealed two major drivers for this research, 1) the need for effective, transparent and sustainable water resources management methods and supporting tools; and 2) that effective and efficient water data management could greatly improve the former, but it was 'hard' and a recurring, unsolved challenge.

Spatial Data Infrastructures have been well described in the literature, and New Zealand employs various methods. Yet, for water data management there was no national consent on 'How' to manage such data. This study provides a comprehensive, yet integrated concept for groundwater data management with directly linked processing, modelling, and visualisation capabilities. Through the course of this research it also became apparent, that the data collection through environmental wireless sensor networks could be linked into the framework via the same conceptual SDI semantics (i.e. via SWE). Through abductive reasoning from the case studies, different mechanisms for the transmission of such environmental observation data into the

SDI were considered. This innovative approach evolved SDI theory under the perpetuation of the SDI and SWE semantics without breaking with its foundations.

The developed method of 'Grounded Design' for GIScience is a new research paradigm which explicitly prescribes an exploratory and abstracted data analysis. Based on abductive reasoning it allows better understandable evolution of the design theory instead of a revolutionary one. This new perception augments data-heavy interdisciplinary research with interoperable data integration methods from GIScience.

7.2 DATA SOURCES

Although technology and policy have provided ready access to more data, there are limitations in how to transform this data into an improved understanding of hydrogeological science and water management.

Systems rely on SDI, an infrastructure that uses communication protocols and common data models to achieve technical interoperability, and the use of controlled vocabularies to achieve semantic interoperability.

The web-enabled open standards-based framework developed in this research, for New Zealand, provides improved discovery of and access to distributed groundwater-related geospatial features and observation data through the internet.

The OGC employs the notion of interoperability experiments and testbeds to test new standards. These interoperability experiments and testbeds typically focus on a small number of standards at a time and aim to mature those standards. This promotes the maturing of the involved standards and their implementation in software artefacts thereof.

For example, in the current development of GWML2 through the GW2IE, the focus is a domain data model and its encoding. In the development phase and subsequently in the interoperability tests WFS, Filter Encoding (FE) and GML are tested to see if they work together in real world implementations in order to effectively deliver GWML2 encoded data. This research could also be understood as a 'meta interoperability experiment' that addresses the scientific domain as use cases (groundwater in New Zealand) from an holistic and opportunistic perspective. This study looks beyond current methods, as it became apparent through the case studies, that not all cases from groundwater research, modelling and visualisation could be satisfied by GWML2. GWML2 was intended to be integrated with a suite of data management, modelling and visualisations tools, beyond storage and exchange of harmonised data sets.

For storing and subsequently making complex feature data accessible via the internet, a database is more appropriate than shapefiles and property files, which have been used in the prototyped infrastructure. However, the mapping from relational database tables into a complex features representation encoded in GWML2 is a database design and feature mapping exercise. There are established methods for generating database schemas either from the physical schema or the logical UML schema.

Treating high level data and metadata access from New Zealand's significant observation databases like CLIDB and GGW/NGMP in the same way as sensor networks greatly streamlines data collection, search and acquisition. This required an analytical act to be able to expose the data model through an O&M observation data model as well as the metadata of the collecting procedure (be it field procedure, lab analysis or sensors). Subsequently, the implementation iterations could evaluate if the data which was produced from the SOS servers as O&M/WaterML also represents the underlying data of the legacy database.

Timely automated updates of the harmonised NGMP/GGW database provided the most recent picture of the environment at any time without labour intensive workloads. On-going quality control and assurance of the original data sets can be made accessible to clients on-demand and add more value. Furthermore, the integration of monitoring observation time series from sensor networks will enable characterising input and output relations of the catchment in near real-time.

However, the New Zealand setting challenged the traditional OGC SDI transport patterns via HTTP and XML. In rural and remote areas, like the Upper Rangitaiki sensor station case study location, the GPRS/3G cell phone coverage and reception was patchy, and internet connectivity and available bandwidth were unreliable. Thus, an innovative approach needed to be found for efficient implementation of SWE semantics in the inherently unreliable and low bandwidth uplink connections of wireless sensor networks.

While problems related to low bandwidth networks is currently a concern, with increasing 3G coverage, bandwidth in urban areas is a slowly lessening concern. However, in remote rural locations connectivity, available bandwidth, and energy-consumption the heavy XML-based semantics of OGC HTTP web services have to be considered. For this reason an MQTT transport mechanism with a light-weight JSON encoding for the payload was chosen for New

Zealand. This choice was made based on the literature, as there was no equivalent in New Zealand.

The developed design theory recommended clearly SWE semantics, but most of OGC interface standards rely on an HTTP transport model. However, this is currently being challenged through the SensorThings API, which also extends the mapping of the conceptual models of SWE onto the MQTT transport domain. Here, observations (aka Part 1: Sensing) are implemented on MQTT in a new conceptual design, which deviates from the classical SWE SOS and SPS model. Furthermore, the observation part (Sensing) is passed as an OGC implementation standard, and the planned parts 2 and 3 (Tasking, and Notifications) are also envisioned as remodelled concepts of existing OGC standards or draft standards (e.g SPS for Tasking; and Web Notification Service or Sensor Event Service for Notifications).

SOAP is a widely employed protocol for web services, and many OGC standards define a SOAP web service interface. SOAP requires XML as transport encoding. There are also other current developments regarding the reliance on HTTP in particular with respect to the Linked Data convention and the REST paradigm. Both use HTTP verbs as a means of expressing actions with an URI addressed resource. REST is not a protocol in itself, but rather a convention. REST-enabled web services are used in many current modern web and mobile applications where JSON is becoming a preferred data transport format. Here, a convergence towards a more lightweight web can be observed with modern web applications, which shift from XML to JSON as their data transport encoding between server (web application) and the client (smart phone or web browser).

Through the distributed nature of geodata services the primary client is the user who connects to the infrastructure with mobile and web clients to a potentially unconstrained number of servers where data and processing resources are hosted. Both, SOAP and MQTT are designated machine to machine protocols. Both resemble the classic client server model of network computing. The inter-server communication and the end user client (and the sensor web) linkage into the server network are the main categories of participants. Sensor networks also connect to the server infrastructure and need to have lightweight bilateral capabilities (tasking, and sensing, which is primarily a push mode), whereas mobile and web clients do not necessarily need to be tasked. Client or web browser notifications are used for interactive purposes at the user interface, which work predominately in pull mode.

A conclusion from that situation is that users with lightweight mobile or web clients mainly need tailored access to the resources of the infrastructure (i.e. data, processing, and visualisation), as they will not download, process and visualise huge data sets locally. End users such as scientists are an exception as they tend to use desktop computers for data processing, analysis and visualization because of the need for sufficient processing power. Using lightweight mobile and web clients means that large datasets cannot be directly downloaded, processed, or visualised locally and that a tailored access solution is required for such capabilities for end users. Thus, the convergence of SDI, sensor web, and message passing should be observed as an important step towards a new SDI paradigm.

7.3 METADATA

Being able to search major New Zealand catalogues for research articles and data sets from one place is practical. It is possible and useful to link the research articles, for example of the Journal of Hydrology New Zealand, into the standardised metadata catalogue search. One outcome of this study was that the ISO 19115/ 19119/ 19139 metadata standards provide reliable and sufficient guidelines to produce appropriate metadata records. While the ANZLIC metadata standard is a profile of the ISO metadata standard itself it does not seem necessary to develop an additional specialised metadata standard to create metadata records from the data sets of the case studies for search and discovery in CSW-enabled catalogue.

End users might request data repositories at any time and request the same data set many times. For example, the same data set might be accessed several times when running data modelling routines. The consumer does not want to be billed multiple times for the same data set just because it has been accessed more than once. Furthermore, those recurrent data requests cause extra network and system load on the data provider side.

However, in terms of explicitly licensing environmental data products for multiple accesses there seems to exists an inertness in New Zealand in particular concerning hydrological, climatological and geological data sets from CRIs, regional and central government agencies. For NZGOAL licensing, and the different funding sources of research and data collection which results in various data products, data providers should be aware that with emerging web service architecture in particular the existing soft licence agreements might need to be revised. NZGOAL states that, under most conditions, agencies should publish data with the most open (i.e., the most permissive) form of Creative Commons licence. But there are various degrees of permissiveness with the Creative Commons licence.

The declarative approach referencing data sets through OWC context documents in the way described in the metadata section of the design theory allows for the handling of a well-defined collection of resources from the Hydrogeology Infrastructure including complete metadata records. Consequently, this context-based passing around of data resembles a 'by reference' type definition of data inputs for WPS processes. The actual data does not need to be copied before an actual processing function is applied. WPS models themselves can be addressed via references, including a metadata record that provides more information about the WPS process itself. Thus, more comprehensive scenarios become available for how data quality, provenance, and licensing information for resulting subsequent data products could be derived, because the respective metadata records for the data sets are also included in such a collection of references. This approach could be compared to the way open source projects define their product license based on the dependencies of the upstream software libraries that they link with. Software modules that are licensed under popular open source licenses like the strict GNU Public Licenses (GPL), the more liberal Apache Software License and very permissive licenses like the MIT or BSD software licenses, can mostly be combined freely in a new software product. Software modules that are integrated into other software must have the modules' licenses explicitly shipped with the software. The terms and conditions of a software product which links such libraries needs to comply to all license conditions of the amalgamated upstream software, which results typically in selecting the most appropriate software license. Thus, it becomes possible to infer the resultant license based on the 'input' licenses. This practice is indispensable in the software ecosystem and it certainly is necessary in the scientific domain where more and more data products are created automatically. This has not been attempted in this study, but by explicitly carrying around the references to the originating authoritative metadata records within OWC documents such a paradigm could be operationalised for use in particular in the 'model web'.

Searching for specific place names or regions might be more powerful with the OpenStreetMap (OSM) gazetteer, as it contains additional volunteered geographical information provided by the public community. The continuously updated OSM Nominatim geocoding service and the OpenStreetMap, which is made available for reuse under the Open Database Licence (ODBL) sharealike licence could be used to complement the LINZ gazetteer. However, more place names will not necessary solve the challenge of accurately identifying occurrences of place names in documents. Furthermore, more place names will significantly increase processing time. This aspect would need to be investigated further to ensure the solution is employable on a large scale.

Beyond the seemingly straightforward task of literal comparison of place names from lists like the LINZ gazetteer several challenges arose which were identified but not addressed further in this study. These challenges were in part related to ambiguities which arise from the textual context e.g. 'Waikato' could not be differentiated from the word comparisons between for example 'Waikato river', the 'Waikato region', or the 'Waikato river catchment'. Additional difficulties stem from contents and quality of the LINZ gazetteer register of place names, which holds official as well as unofficial records. This is further aggravated by the fact, that the place names used in publications can vary even further by referencing of geological formations. The LINZ gazetteer only holds point geometries for locations. Thus, the approach to find a spatial bounding box for a metadata record is neither accurate nor precise, but a reasonably pragmatic approach. Eventually, the automated geo-referencing method that was developed in this study was not be evaluated on a quantitative basis. But the overall approach to provide user support for the task of on-demand geo-coding of written documents to make them spatially discoverable in CSW catalogue services is undoubtedly of advantage.

Technical aspects of semantic web methodologies provide means of improving semantic interoperability and data harmonisation only so far. Geosciences communities have been continuously focussing on semantic methodologies using ontologies and their machine-readable encoding.

The current situation would lead to the conclusion that these ontologies are not complete. The method of grounded design that was employed in this study suggests that local use case orientated applications of ontologies should be based on agreements of communities working together. This approach would allow for the adoption of existing ontologies, and for the extension of these ontologies or the creation of local ontologies that link to larger more general ontologies, without competing with the general ontologies.

The three examples encountered in this research were the NZHS freshwater glossary, the NGMP observed properties parameters, and the NIWA CLIDB LADA codes. Cross-checking for keywords and integration of thesauri/vocabularies to improve search semantics, as well as explicitly describe units of measurement definitions were beyond the scope of this research. However, this aspect of harmonisation was examined both conceptually and pragmatically, via implementation, to determine whether or not such integration was possible. Further development of those ontologies would require domain experts in order to encapsulate their knowledge for the purpose of automatic translations between data services. This translation step would then deal 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.

Semantic descriptions of the observed environmental phenomena are increasingly viewed as a pathway to data interoperability between regional councils, CRIs and other private and governmental organisations. A community effort 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.

In this study, for the gazetteer service a WFS/GML Simple Features implementation was used, whereas for the thesaurus SPARQL with RDF+XML/SKOS was employed. There is considerable debate in the general research domain about which method to use for which case. Particularly at the boundary between feature implementation and thesaurus the delineations become blurry. There are alternative approaches. For example, the thesaurus could be expressed in OWL which would provide more detailed modelling capabilities. The transport encoding could be one of the more popular triple notions or even JSON (JSON-LD). Here again, it can be observed that the field is in flux and convergence might be ahead. The discussion as to whether GeoSPARQL or WFS Gazeetter, XML or JSON-based encodings, or other triple notations should be used and when raises questions about the practicality, usability, effectiveness and efficiency of their respective implementations.

From a low-level software implementation perspective the selection might matter but on a conceptual infrastructure level it matters less. OGC service interface specifications (OWS Commons) and SPARQL can be considered to be independent of either SOAP or REST because they operated on top of well understood basic HTTP. Thus, the ubiquitousness of HTTP in that domain should be honoured.

Pragmatically from a data provider's stance, security is an issue: the transmission and access to sensitive datasets needs to be secure and have integrity. HTTP can be transparently secured via a variety of mechanisms, e.g. Transport Layer Security (TLS) encryption and HTTP basic authentication. Other approaches like API keys have become popular, too. More complex schemes, like OAuth, work with REST, whereas security extension were designed for SOAP protocols, e.g. the Security Assertion Markup Language (SAML). After all, authentication sources (i.e. user identity) are again independent from the process and are purely community or organisational considerations.

These considerations have been removed from this study where the research was striving for a parsimonious theory in correspondence with the idea of Occam's razor. And while it must be acknowledged that there is stronger usage of advanced ontologies and more detailed data modelling pursued outside of New Zealand, the discussion and decisions around OWL vs. SKOS in this research were guided by the desire for simplicity and applicability. In particular in the context of the case studies improving complexity would not add value to the design theory.

7.4 THE GROUNDWATER SDI

A new key aspect of these interoperable systems is that the source data resides with the custodians and is accessed on demand through the internet. This ensures the currency and validity of the different services and that user activities are connected and interdependent. Compliant systems are capable of accessing, displaying, and harvesting the distributed spatial (meta) data to create a modelling environment made of orchestrated distributed data and geoprocessing services.

From 2007, when LINZ published NZGO and NZGOAL, to today, data providers in New Zealand are still primarily concerned with the challenges surrounding the publication of their data.

The starting point to address the objectives of this study and answer the research questions were: 1) the ever increasing collection of specifically hydrogeological and environmental data, in a wide variety of formats; 2) the evolving data formats, data and database models and formalised standards; and 3) their documented application to data collection, storage, harmonisation, discovery, access and transfer in the scientific literature.

The search, retrieval, and pre-processing of data sets consumes valuable time and personnel resources on both the data provider and data user side. Inconsistent data formats and lack of consistency and reliability of metadata are of significant concern and contribute significantly to data user costs. These concerns have already resulted in the development of new data management techniques, tools, and data models to effectively manage the vast amount of hydrological data being collected. However, putting these tools, models, and techniques to use in a broader interdisciplinary context was difficult.

Through seamless continuous linkage and guidance from data collection to data discovery and sharing, this work brings together the advances from different fields of SDI research and applications to refine technical aspects of SDI in the domain of Hydrogeology. The organisational, policy and economic (as in business models) aspects of SDI however were not examined in-depth. From the case studies and from the literature review the framing was taken, and the design theory focussed on technological integration aspects. The reason for this focus was that implementation and integration challenges were a major obstacle to any further developments in the field and thus this obstacle needed to be overcome first.

Data (re-)usability, scientific modelling and advanced visualisation are all complex and elaborate tasks. Thus, combining these tasks into a seamless framework is an even more complex and elaborate task. This complexity means that the development of solutions for this task are rarely accomplished, or only accomplished with high development costs. Consequently, groundwater resources management problems that could be solved better with a combined solution (data (re-)usability, scientific modelling and advanced visualisation) suffer from a sparsity of viable cost-effective implementations. In contrast to a sparsity problem, an abundance problem is a type of problem for which an abundance of cost-effective methods and tools are available.

This research presents an integrated framework design that combines data (re-)usability with seamless scientific modelling and advanced visualisation using open standards. The result is a framework that has the potential to transform groundwater resources management from a sparsity problem into an abundance problem.

Subsequently, organisational, policy informing, economic and educational data and science requirements for groundwater resources management can be re-evaluated. A new ease of groundwater science accessibility and reproducibility could catalyse societal acknowledgement in science and demand increase in public availability of interactive and online groundwater knowledge.

With this framework the horizon beyond data publishing with OGC standards was examined. Gaps along the way from data collection (sensor networks), data storage, discovery, access and sharing (classic SDI), to use in scientific tools and visualisation of raw as well as processed data informed and contribute to this new Hydrogeology Infrastructure. More developments on single data sets available at the regional councils' and research institutes' data repositories can be made accessible online in the agreed on target schemas as described.

In the literature review it was stated that the situation of data interoperability for groundwater data management in New Zealand according to Lowry et al. (2003, p. 579) 'appear to be predominantly information and implementation issues.' Implementation constraints included inconsistencies, across organisations, in data formats, standardised APIs, parameter naming, and parameter descriptions. These issues were drivers for this research as at the on-set of this research the issues still existed, despite the fact that technologies, methods and techniques, albeit disparate ones which require careful consideration and integration, now exist.

In the initial research context of this study it was identified that unstructured hydrogeological research outputs, and in particular textual documents like PDF files, were barely accessible through classical SDI techniques. This has been put into perspective in the course of this study, as geographical and semantic referencing methods for 'unstructured' and otherwise inaccessible documents was demonstrated successfully. Thus a spatial or topical search may now include this legacy scientific information as viable results in single point of access CSW queries.

Once OGC-enabled data services are readily available, they can be reused in a multi-project context. This enabled multifunctional and multi-user public use cases through services that are accessible through standardised data exchange. Preferably users would connect interoperable applications to the OGC-enabled web services.

In this study two approaches to connecting interoperable applications were examined. Firstly, standardisation of access through agreed on (web-) service interfaces and data formats (i.e. standardised API). This allows systems and software modules to be reused to access a variety of distributed data sources and load the data sets. And secondly, semantic inter-linkage of concepts has emerged as a sensible approach for the following reasons: While agreed on community vocabularies, definition of concepts (e.g. concrete definitions of measurement concepts) are more likely to be accepted by immediately collaborating data users like scientists and council officers, a general, possibly even global problem for standardisation is that different interpretations and meanings may exist between user groups.

From a rational management efficiency point of view it seems useful to have 'one definition to rule them all', but in reality reaching agreement at higher levels may be almost impossible. Thus, as it needs to be used, needs to be applicable, starting with definitions from smaller, but practising groups is a great bottom up approach which may lead to an output which could be widely adopted. Thus, the semantic technologies move in the background and enable the community to work.

Coming up with a joint vocabulary for hydrogeologists is also challenging due to the multidisciplinary nature of the domain. Different vocabularies are used in the relevant disciplines. For example, geology and meteorology: both have their own disciplinary language but are also integral to hydrogeology with its own language and vocabulary.

These communities obviously have overlap. Particularly for New Zealand using defined geological units to conclude over hydrogeological units is unfortunately not working (Lovett & Cameron, 2014). Thus, they still need to link to lithological and stratigraphical definitions of the geologists, but will have to add own characteristics that form different unit boundaries of importance.

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. This has been done for the New Zealand Geological Map, but for hydrogeological units and lithological classifications no attempts have been made yet in New Zealand.

Hydrogeologists need distinctive well-defined groundwater chemistry terms, but also need to employ relatable terms from geology. For meteorological/hydrological interaction, there needs to be a clear and shared understanding of the definitions for rainfall and river flow etc.

For inter-cultural international settings such as the European Union reliable standards and multi-lingual data sharing capabilities are of utmost importance. In order to ensure transregional, interdisciplinary (hydrology, climatology, land use science, soil science or earth sciences in general), and transcultural usability for a wide community of different user groups (e.g. public sector bodies, private companies and citizens), the related services need to consider aspects of multilingualism (e.g. Te Reo Māori and English) and indigenous knowledge approaches. The INSPIRE directive is very ambitious when compared to the New Zealand situation. But to avoid extensive work load in establishing the thesaurus, the General Multilingual Environmental Thesaurus (GEMET) could complement New Zealand community efforts. Here the ontological relationships, expressed as concepts, encoded in standardised formats, can be linked as direct or narrow matches with other domain ontologies and broader general ontologies like NASA SWEET[1]. After all, community vocabularies and ontologies are local artefacts for example, but the integrative theory is applicable globally.

Thus, technical interoperability through interfaces and encodings are easier to prescribe standards, than concept definitions. Here heterogeneous definitions across domains seem more acceptable when they they are linked through concept matching to various degrees.

7.5 GEOPROCESSING

Geoprocessing is a classic GIScience discipline. Many GIS implementations with geoprocessing toolboxes are available with algorithms for the breadth of geographic problems. However, GIS interoperability for Earth Sciences (with focus on Hydrogeology) is typically limited to surface hydrology, soils and basic geology as specialisations of physical geography. This might stem from a different understanding of the natural processes at work, which are strongly conceptually, structurally and numerically modelled aspects in geological and groundwater. In GIScience spatial data management and geoprocessing is a well developed field of research. Thus, groundwater modelling with a GIS geoprocessing method was attempted. Did it work? Yes, technically this gap between traditional groundwater modelling with manual data management and this integrated and online data linkage approach was closed. But the developed understanding revealed new challenges, in particular challenges related to user interfaces development and usability. These challenges extend beyond the scope of this research, because the design of effective user interfaces was not addressed through the research questions, nor the focus of this research. The on-set of linking integrated environmental modelling with real-time environmental sensor data has been proposed as a necessary roadmap for the future (Laniak et al., 2013). To better understand the interlinked nature of an integrated modelling environment there were two approaches discussed in this study:

1. tight coupling, which creates very complex and specific software (OpenMI, MPI), where it is harder to scale out over a network and flexibly linking of computing and data resources.

1 NASA SWEET Ontology, website: https://sweet.jpl.nasa.gov/

2. loose coupling, based on standardised data exchange over web services, where computing and data resources are linked via web services.

Due to the nature of data infrastructure categories stemming from the case studies, the modelling integration research in this study focussed on loose coupling, aiming for logical and technological consistency. Nevertheless, both approaches need to be discussed to understand the context of the implications.

ArcGIS model builder is an advanced tool for composing modular geoprocesses into a larger modelling workflow. But the data files need to be accessible through the computer's file system. Advanced commercial tools increasingly are able to load a variety of data sources into their modelling work space to preprocess data sets from remote data into the required formats. The ArcGIS model runs only on the local computer. Advances in commercial implementations demonstrated the use of proprietary protocol extensions to spread computing over local network resources. Thus, automated interchanges and digital representations of model results were specialised solutions and neither generalisable nor extendible beyond their proprietary purposes.

An alternative implementation strategy was used in this research in order to encapsulate the model logic into one or several WPS processes. MODFLOW, USGS Soil Water Balance (SWB) model[2] and Soil Water Assessment Tool (SWAT) model[3] are widely used groundwater flow and hydrological modelling software packages that are implemented as independent command line tools. Wrappers are suggested in order to include these models in a web-enabled execution environment accessible through the standardised WPS interface. For groundwater modelling with MODFLOW this has not really been explored beyond traditional GIS approaches before.

Recent trends in 'Big Data' and 'Cloud Computing' have also been addressed in the literature, and OGC web services for SDI and WPS for processing and modelling in the web to tackle the large amounts and variety of data and to unify currently still very disconnected environmental modelling routines. A 'WPS mediation' could be used to dispatch models as services on different computing back-ends. This idea was taken further conceptually in this research in order to increase the practicality and usefulness of 'WPS mediation'
.

2 SWB website: http://pubs.usgs.gov/tm/tm6-a31/
3 SWAT website: http://swat.tamu.edu/

The integration of legacy models through wrapping their legacy code into OGC interoperable interfaces now becomes a software engineering exercise in terms of exposing all of the many MODFLOW software parameters effectively through the WPS interface. For a comprehensive design theory, additional important aspects need to be considered. The data quality of the used data, as well as uncertainty of intermediate modelling results need to be propagated through the full process chain.

Here three places became apparent where data quality and uncertainty could be expressed in relation to a data set, be it raw data, processed or a model output from many inputs. This can be presented in the form of a metadata data record holding data quality, lineage and provenance information. This metadata record is stored in the catalogue and acts as the root for data traversal and evaluation.

Another typical variant is the data set itself if it provides the necessary encoding, e.g. GML-based, O&M in particular, but also multidimensional encodings like NetCDF provide the capabilities and elements that can be used to transport the information along the actual data.

A third variant is uncertainy representations of data sets in a separate data set. Thus, it represents rather a new data set. Typical cases are raster data sets.

Klug and Kmoch (2015) highlighted that repeatability and transferability of the modelling exercises are hampered, because current modelling tools and software modules are often designed to work standalone and not within a 'network of models and data'. It was also noted that over-specialised networked or file-based data exchange between models and users via an easily accessible SDI are usually too loosely defined. In this research the WPS-based encapsulation with well-defined standards for the purpose of integrating with and using data from the SDI was determined to be the best solution given the hydrogeological context.

Ultimately, this study emphasised a declarative approach by referencing data sets and models from the SDI with the help of OWC documents. Such documents explicitly hold together data sources, models, result data sets as well as the related metadata records via their references. This approach enables a completely new perspective for a distributed work environment. It increases reliability, credibility and reproducibility of hydrogeological research.

This study showed that classic geoprocessing techniques from GIS modelling could be applied to hydrogeology, thus, broadening the field of applied GIS methods. However, a drawback of the tested implementation of the MOD-

FLOW WPS was the limited usability. The test was only successful for a small number of input parameters. It might turn out, that is not feasible in terms of usability to expose all of the many input parameters that MODFLOW supports in the WPS process description. Instead, it might make more sense to create different modules, particularly to be able to provide specific functionality in terms of modelling outcomes, e.g. be it flow direction and velocity, recharge and water budget, steady-state or transient model. However, it was not focus of this study to determine how to integrate MODFLOW best into an SDI, but rather to evaluate if it could be done under the context of SDI data provisioning.

Another hydrologically critical point was that when designing architectures for environmental software applications linked with live sensor services it becomes even more necessary to add proper measures for validation and improvement of model (auto-)calibration. But this is hard to achieve. For groundwater models like MODFLOW this is particularly true, and typically many realisations of a model will be generated and parameters adjusted based on scientists experience. Here post-processing tools like PEST, for model-Independent parameter estimation and uncertainty analysis[4], are widely used for optimisation. Again, eventual judgements will be made by scientist based on their understanding of the hydrogeological system.

While it has not been examined nor discussed in greater detail in this study, a recursively cascaded WPS execution which may take other WPS outputs as inputs would be of great advantage. This can be understood as a workflow but also as in functional declaration of dependent WPS processes. A deeper examination and evaluation of recursively cascaded WPS execution would be one potential avenue for further research using quantitative DSR research methods.

The meaning of the inputs and outputs descriptions of WPS processes in general and their semantic definitions needs further improvement. The current WPS 1.0.0 standard process descriptions only allows for the definition of interoperable data exchange on technical and syntactical level. To allow for better semantic interoperability for geoprocesses and subsequent visualisation, this thesis' GIScience research approach has revealed several suggestions from annotations to additional semantic mediation services.

However, all these approaches assume that the ontologies used are complete. It may be concluded that ontology inter-linkage could support the matchmaking and mediation from source data to required model input data

4 PEST website: http://www.pesthomepage.org/

based on syntactical (file formats, encodings) and semantic level (rainfall, river flow, groundwater level). Consequently, the use of CSW queries that employ multiple thesauri and controlled vocabularies would support filling the WPS process input requirements from an SDI.

To take this further, SensorML and Sensor Web Enablement (OGC SWE) components provide much more complex data in- and output definitions and can be easily interlinked with Linked Data vocabularies. With the release of the WPS 2.0 specification in 2015 the process model itself has been largely decoupled from the WPS protocol allowing the use of SensorML descriptions of processes. So far there is no implementation of a SensorML-enabled WPS server. This could not be further examined in this study due to lack of time and available implementations. This study alternatively employed OWS context documents with WPS processes, in particular to reference utilised WPS processes in a compact transport encoding.

7.6 GEOVISUALISATION

In this study the 3D visualisation of hydrogeological data was presented as a dynamic transformation of data from the SDI into an X3D scene description. X3D is an open visualisation standard that can be rendered in current web browsers based on HTML5 and WebGL. The approach can be labelled as a declarative way of visualising data as the visual model is declared and described using higher level geometric 1D (points), 2D (lines and polygons) and 3D (surfaces and bodies) primitives that can be decorated with textures.

This resembles a functional orchestration of SDI data sources and processes to generate intermediate elements from raw SDI data into an interactive 3D scene.

CityGML rendering and viewing services using CityGML as exchange format can be compared to the analogous application for the groundwater domain. A CityGML feature is a GML complex feature similar to features from GWML and GeoSciML. While the underlying method of CityGML rendering applications are not publicly accessible and the actual visual model also remains inaccessible in 3D, the described X3D method contributes to open and public standardisation efforts. It also provides guidelines for experimental (DSR) instantiations towards better understanding of the transformation from simple and complex GML features into 3D shape directives.

To date many visualisation formats have already been successfully implemented and combined in 2D web mapping applications. WMS and WMTS can be understood as portrayal visualisations of ground-truth discrete or continuous data sets, which may or may not be accessible via WFS or WCS. WMS rendering from complex features is currently only realised via 'flattened' complex features and therefore as simple feature representation, due to a lack of ability to select the relevant geometries and data elements and combine them into a coherent visualisation. Similarly, the current X3D scene rendering makes assumptions about the input formats and their meaning.

Several technical issues became apparent in the evaluation. X3D is a powerful language to model 3D scene but the actual rendering capability was dependent on the JavaScript library x3dom which did not implement all X3D capabilities. Thus, the full specification of X3D was not accessible through the web browser during the course of this study.

The abstraction of geovisualisation as another application of geoprocessing yielded a practical and pragmatical manifestation of the functional projection from data over models and software modules into the visual domain. As geoprocessing concepts are well understood in the GIScience domain, the understanding of the generation of geovisualisation (e.g. the creation of an X3D scene) as geoprocessing algorithms provides a valuable encapsulation of complex visualisation from well defined GIS-type input data.

The 3D scene generation process, in the way it was implemented, has to wait for all elements to be available from the SDI before it can assemble the complete scene graph. This is partially necessary because the scene bootstrapping applies several transformations on groups of shapes (e.g. the vertical exaggeration) before it will be send to the client. An X3D element-wise streaming or ad-hoc 'level-of-detail' pyramids approach could be a possibility to improve usability and flexibility in the web browser. Furthermore, too many elements and big point clouds can become too 'heavy' (consist of too much data) for the browser to handle. A light-weight representation is necessary. This has been addressed by manual down-sampling, for example for WCS requests, and also requires the initial scene bootstrap.

Due to the lack of support by x3dom at the time of the implementation, the geo-referencing and respect for geographical scale were not addressed sufficiently. For simple 3D visualisation purposes of regional scale geological model this does not cause issues, on large scale and for analytical purposes this method is clearly not accurate enough.

Scalable Vector Graphics (SVG) is a declarative visual XML language for vector graphics. SVG is a standard format which can be rendered by current browsers natively. In the way that W3C and the Web3D group advanced VRML to X3D there is a possibility that X3D could eventually become a native 3D browser format, too.

WMS is a specialised interface for rendering geographic data sets based on attribute selection and styling through, e.g. SLD/SE or Cascading Style Sheets (CSS, which are heavily used with HTML). The current understanding of 3D portrayal variants from OGC focussed on the interface semantics and did not take the underlying data semantics into account. The new declarative 3D rendering method could therefore serve as an explicit basis for portrayal services. While this has not been attempted selective styling could be employed, in the same way as for WMS, to the 3D visualisations.

As a result of the 3D visualizations, proof of concept implementation, some semantic issues became evident. For example, if several grid coverages were available in a given context document it was impossible to automatically determine which was the designated surface DTM. In another case, interpolated groundwater level or concentrations of constituents could be visualised in a pictorial WMS fashion (for example, as heat map or contour lines). Alternatively, such 'value grids' could be visualised like an elevation (aka 2.5 surface) in a 3D scene. While both forms of visualisation have their time and relevance, the automatic selection of the desired form of visualisation might become possible through the incorporation of ontologies to infer a dynamic mapping of OGC geodata elements into visual elements.

The overall topic could be elaborated on more extensively. For example, if we treat this 'mapping' as a functional of the set of OGC data input to the set of (X3D) visual elements we might be able to include more visualisation elements. Geoscientific and hydrological data sets, properties distribution and time series might then not only represented as maps in their true-to-life forms but also as interactive adaptive 3D bar charts for time series in 3D view.

7.7 THE HYDROGEOLOGY INFRASTRUCTURE

The literature review revealed that New Zealand aquifers are not well enough understood to make comprehensive informed policy decisions. To allow for more effective policies and better informed decision making, water resources management data needs to be used more effectively and efficiently.

The overall vision at the onset of this research was to improve the time consuming process of retrieving field observations and to deliver the data in a way that it could be easily used and analysed. Typically the environmental data life cycle begins with real time ground truth field observations which are then collated in same way as a data set. The data is quality assessed and preprocessed for use in hydrogeological systems assessment and characterisation, which usually involves numerical modelling and predictions. The results are presented in reports which are then considered in water resources management and used for compliance and monitoring purposes according to the RMA.

This study developed a hydrogeology infrastructure design that attempted to address the full data life cycle in the context of hydrogeology in New Zealand. Formerly disconnected and distributed data sets could be discovered and combined with a harmonised representation and distributed within one data portal. For example to permit direct access to groundwater data, a user can directly access the data sources via WFS and SOS without the need of labour intense pre-processing and harmonisation of data sets.

Easy access to hydrogeological data and information would be a key enabler for efficient assessment and management of groundwater resources for New Zealand. While this study focussed on integration of data processing, analysis and visualisation in a comprehensive framework easy access was a prerequisite that was examined and developed further.

The presented developments are of benefit for consistent and transparent decision support in catchment-wide approaches and for analysis at the national level. Contexts (represented as OWC documented) can be visualised in 2D and 3D data viewer to facilitate a spatial understanding of a catchment. These contexts can also be sent to more sophisticated hydrological modelling processes that need a lot of different input parameters, e.g. the exemplary MODFLOW module and the X3D visualisation process.

Traditionally, these sophisticated modelling tools run in isolation from the data preparation process which makes the modelling process very time consuming. Also these tools outputs were not directly reusable in the OGC web spatial context so the WPS process would have to reprocess these outputs back into a format that could be handled in the Spatial Data Infrastructure (SDI). Such modelling tools are often used to calculate predictions with the result types resembling discrete observations (as in water sheds at a point, or water balance) or gridded data such as spatial distribution of environmental properties.

A possibility could be to rewrite legacy software modules like MODFLOW for modern demands like cloud parallel execution capabilities and alignment with web-enabled data management. Such improvement could even open the door to 'big data' type analyses of groundwater settings based on available data in the SDI and would support policy making as well as provide a common ground for competitive groundwater management and use as the used data sets of such scientific studies could be referenced for verification purposes.

In this study the technical performance of the implementations were not validated. This was due to the overarching aspect of examining the web services and data exchange interplay. A variety of implementations were tested to fulfil and validate the design theory. Thus, the body of data analysed in this study is from scientific literature, formal standards developed by expert groups, and the data sets and technologies from the case studies.

Now it can be confidently stated that more quantitative measures should be used to sensibly evaluate technical performance of the data and processing services as well as the usability of the data portal reference client.

In terms of generalizability of the developed theory it could be said that similar conditions and understanding of SDI have been found and implemented in particular in Australia, North America and Europe. The New Zealand case study exposes several supporting characteristics like NZGO, NZGOAL and well developed Information and Communications Technologies (ICT) infrastructure. However, open data legislation, internet accessibility and available bandwidth were not compared with other countries and might differ.

Conditions for stricter internet and data access regulations as they exist in China and Russia have not been examined and might imply stronger differences in data exposure and network security.

Also greatly varying infrastructure maturity, ICT access, and usage patterns as they exist in India or Africa might require different considerations for end user accessibility and uptake of such a Hydrogeology Infrastructure. Also many countries do not yet have basic groundwater bore databases (e.g. many of the developing countries) thus the application of the developed design and infrastructure would be inapplicable to them.

Thus, as a conservative conclusion, the developed design theory might fit very well with the western societies in particular: New Zealand, Australia, United States, Canada and Europe. For other communities the overall theory might still be of great value, but a grounded theory review of their local specifics would be necessary to better address requirements of these communities.

7.8 grounded design with case studies

In retrospect, the developed research philosophy was a successful guiding principle for the context of this study.

Grounded Theory by itself is not intended to provide an immediate solution to a problem. Its capability is to yield better understandings of certain research contexts. Therefore, it needs sufficient variety as well as depth of source data. Although the constant comparative method of GT uncovers the emerging recurring patterns of the source data into account, this is not necessarily represented in the statistical relationships and distributions in the resulting grounded theory. The constant comparative method is a tool to abstract away the variety of the data and to capture its essence (given there is valuable essence for theory) - it is not a statistical exercise.

The case studies represent New Zealand in the original derived grounded theory coding. Design Science in return is aiming to generate new or improve theory thus creating new knowledge through moving forward. Community participation provided essential input for the grounded theory development and the published research contributions as described in the introduction were essential parts of the design science feedback loop to validate and improve the overall design theory along the way.

The departure points for the methodology were existing theory and artefacts. The flaws and issues of the current situation were identified from the formed grounded theory and subsequently, through reconciliation with a literature review. Thus, the need for improvement of some kind was pointed out and comprised the initial context for the Design Science iterations. From there the research added a problem-solving dimension. It did so by implementing new artefacts and comparing their characteristics with the existing ones.

Through this circumscription new knowledge is generated as the implementation served as instantiation of a new idea or partial method which was then incorporated into the greater design theory for the particular problem domain. Also, in DSR, progress is achieved when existing technologies are replaced by more effective ones. Based on the problem space maturity in the DSR knowledge contribution framework from Gregor and Hevner (2013), most parts of the design theory for the Hydrogeology Infrastructure could be attributed as cases of 'Exaptation' (non-trivial extension of known solutions for new problems). However, the encapsulation of legacy groundwater modelling software in geoprocesses to enable the model and the reshaping of the understanding of geovisualisation as an extension of the model web, the X3D WPS process, could be understood as an 'Invention' (new solutions for new problems).

Grounded Theory allows for, even prescribes, a deferring and delaying of theorising. This forces the researcher to work through the data and avoid premature judgements and investigation into causality. The actual theory is put together during the write-up phase from the categories. Thus, submission to GT forced the researcher to be open for patterns in data without immediately theorising reasons.

It is a human trait to try to find meaning in everything, thus every bit of information wants to be put into context and be theorised. This can distract and cloud the researcher's ability to let the data flow openly and impart his ability to abstract concepts effectively. By avoiding theorising in the early stages of the research progress, new pathways opened up from the concepts that emerged, and not from the theory that might be constantly formed in the researcher's mind.

Through constant awareness the researcher strived to avoid finding reasons why a concept emerged or why it related to another concept, because this would limit the abstracting and comprehensive capability of GT. Consequently, the researcher could stay 'uninvolved' with the data and theory until the method of coding and constant comparison from an abstracting exercise yielded satiated categories.

In the synergistic triangulation with grounded theory, GT provides a real context, given that the data for GT sufficient. The case studies provide a well-defined field of input for GT. Eventually, the researcher theorises and reasons over the results and thus can yield a more comprehensive and less biased theory.

From this foundation the design process with its inferential and abductive clues make the improvement steps more clear and even reproducible to an extent. And it was not until the step towards the DSR process that the research tries to solve a gap or improve a situation.

The next step in the presented triangulation of methods was to establish the design theory in DSR artefacts. From there the DSR process took over and was more concerned with the implementation, practicability and efficiency (to some extent). This is the main reason why MQTT was not captured in the design theory. The design theory concluded that SWE semantics should be applied, but through DSR and a real world instantiation it became apparent that some standards could not be applied in some situations (here, the Upper Rangitaiki case study).

In this study the integration, interoperability and eventually the convergence of design patterns were the main output. This was hardly quantifiable. Through the strong GT aspects this study has developed traits of a meta study which is rooted in and subsequently, evaluated with concrete case studies.

8 CONCLUSION AND OUTLOOK

An explicit research method was developed to provide a scientific framework by which the research context was analysed, solutions derived and a reference implementation evaluated and discussed. The analytical process strongly depended on the case studies because the initial theory was grounded in the case studies. the design research process reconciled the grounded theory with the state-of-the-art literature and through real world implementations the design theory was improved and evaluated using case studies. The result is a grounded design theory which stems from the current context in the real world (case studies) but also yielded innovative and improved theory aspects, and thus created new knowledge.

The liberty for this unconventional approach to SDI design is founded within GIScience itself as an holistic connecting discipline and provides new guidance for data research in hydrogeology and the Earth Sciences in general. The interdisciplinary domain aspect is congruent with the interoperability data system integration aspects.

Like all eventual social constructs technology is either used or not and standards are either implemented or not. And while there are reasons for adoption or lack of adoption of technology and standards, they often do not seem to appeal to logic, performance, or other means of effectiveness. A great example is how DVD became the globally accepted standard of a media disc format after competing with the LaserDisc format. While the latter was the technologically superior format the former one was adopted by a wider industry and thus, caused the downfall of the latter. While this historical anecdote's outcome was driven purely by commercial desires of their inventors, the OGC in contrast comprises a global research community. Thus, acceptance and convergence might be broader and more general. However, each OGC domain working group consists of only a few specialists who may or may not represent the greater community. Insofar, research into the development of standards and the dynamics between the groups, as well as inside the groups provides valuable understanding of the greater embedding of such standards.

Overall, this study revealed the partial tendency of converging and consolidating data standards. All aspects of data and metadata capturing, (re-)

organisation and data harmonisation, as well as semantic and technical interoperability were considered in the matured design theory. It is now possible to produce a comprehensive geospatial information framework on groundwater into a robust platform product, to amass New Zealand's groundwater information from heterogeneous (distributed) sources into one seamless view across New Zealand and for better decision making support and business utilisation in the future.

8.1 answering the research questions

Question 1: What are the required interfaces and data formats to comprehensively capture describe and share hydrogeological and groundwater related data sets in New Zealand to enable data interoperability for hydrogeology ('A Groundwater SDI')?

The Grounded Theory approach provided the main pathway towards reaching the research objectives 1 ('Review methods, standards, formats') and 2 ('Identify, select and discuss an appropriate set thereof') and towards answering RQ 1 (Chapter 4 and Chapter 5). Existing SDI methods were examined in relation to New Zealand and to the domain of Hydrogeology and were evaluated on the Case Studies (Chapter 4).

Table 9 (p. 101) and Table 10 (p. 108) summarise 'the required interfaces and data formats to comprehensively capture describe and share hydrogeological and groundwater related data sets in New Zealand to enable data interoperability for hydrogeology'.

More precisely, Table 10 (p. 108) summarises the data type and domain categories and concepts which relate to overall metadata provenance. The resulting interface standards and data formats (as selected from the case studies and in reconciliation with the literature) fulfil the case studies requirements for the groundwater SDI part of the hydrogeology infrastructure and thus contribute to research objectives 1 and 2 and help to the answer this RQ.

The required interfaces and data formats to comprehensively capture, describe and share hydrogeological and groundwater related data sets in New Zealand were tested and evaluated on a reference implementation that arose from the application of the proposed design theory. Data interoperability for hydrogeology ('Groundwater SDI') was demonstrated and consequently, RQ1 can be considered as having been answered comprehensively.

Question 2: Can scientific groundwater flow and water balance simulation models and codes be linked into a networked environment to enable interoperable models ('A Model Web')?

RQ2 was answered through sharpening and refinement of the perspective of how standardised data formats and their accessibility via standardised interfaces enabled interoperability. This refined understanding was applied to existing scientific groundwater flow and water balance simulation software modules.

The developed 'Hydrogeology Infrastructure' provides a new framework for designing interoperable data and analytics infrastructures based on the selected methods, standards, formats and technologies determined in answering RQ1. Through the exploratory and iterative implementations, improvements and design evaluations detailed in the 'Geoprocessing' sections of Chapter 5 and Chapter 6 the answer to RQ2 was verified. Yes, it is possible to link groundwater flow and water balance simulation models so that they are interoperable within a networked web environment.

Limits to the current tested implementation of the MODFLOW WPS was the constrained usability. No advanced user interface was available to better guide the user and assist in including user decisions into the modelling process. The test was only successful for a small number of input parameters. It might not be feasible in terms of usability to expose all of the many input parameters that MODFLOW supports in the WPS process description. Also linking live sensor services into the modelling process brought the necessity out to add proper measures for model validation and to improve automated model calibration.

Question 3: Can such models be provisioned with available interoperable data sets ('Context-based Data Provisioning')?

Similar to RQ2, RQ3 was answered through applying the understanding of existing scientific groundwater flow and water balance simulation software modules under a geoprocessing paradigm with the verification of RQ2.

Furthermore, through recognition and linkage of the context-referencing capabilities of OWC described in the Metadata sections of Chapter 5 and Chapter 6 and provisioning of such a focal data context to a complex geoprocessing algorithm, this RQ has been answered affirmatively, based on the selected methods, standards, formats and technologies from RQ1.

Through the exploratory and iterative implementations, improvements and design evaluations in 'Geoprocessing' sections in Chapter 5 and Chapter 6 the answer to RQ3 was successfully verified.

Question 4: Can the visualisation of interoperable data and models be realised within the same design principles ('Visualisation as an interoperable Model')?

By reshaping the current understanding of the creation of a visualisation as a specialisation of the model web, visualisation can be realised as interoperable model that is provisioned with data from the groundwater SDI. Thus, the answer to this RQ built upon the answers to RQ1, RQ2 and RQ3. Through the implementation and evaluations of the X3D WPS process in Chapter 5 and Chapter 6 RQ4 was answered and 3D visualizations developed that are founded on the interoperable data and data models in an encompassing design framework.

Question 5: Does the new (theory of the) framework design overcome the challenges and close the gaps identified from an analysis of literature?

DSR guides the work on advancing the theory for designing interoperable data and analytics infrastructures. For overarching frameworks, none or only few have applied available techniques or methods, and standards consistently as a distributed web-based framework theory. In the synergistic triangulation with GT, the case studies provide a well-defined field of input from a real-world context for GT. The use of case studies served as an explicit body of data for GT and DSR. Relevant evaluation criteria were established based on the state-of-the-art best practices and identified gaps from the literature. The 'Hydrogeology Infrastructure' application and implementation was evaluated in Chapter 6 (Section 6.5, p. 148), and critically discussed in Chapter 7 as the new (theory of the) framework design overcame the outlined gaps from the literature. Finally, through the triangulation of the described research methods, the envisioned 'Hydrogeology Infrastructure' was completed and the answer to RQ5 was successfully verified. This design process with its inferential and abductive reasoning processes made the developed theory and 'Hydrogeology Infrastructure' design clear and reproducible.

Ultimately, this study emphasised a declarative approach by referencing data sets and models from the SDI with the help of OWC documents. These

OWC documents hold the references to data and sophisticated hydrological modelling processes (e.g. the exemplary MODFLOW module and the X3D visualisation process). The described encapsulation and referencing capabilities enabled the model web and the reshaping of the understanding of geovisualisation as an extension of the model web. This study and the novel method developed contribute to public and open standardisation efforts. It enables a completely new perspective for a distributed work environment. The use of Grounded Design Theory and artefacts developed through applying this novel theory increases reliability, credibility and reproducibility of hydrogeological information and research.

8.2 research contributions and implications

Based on Grounded Theory coding and emerged patterns, a well threaded and consistent theory was developed. Subsequently, with its iterative implementation and evaluation the groundwater framework design and its artefacts thereof evolved.

The final result is a design theory of a coherent, yet loosely coupled, distributed groundwater data analytics network, which is based on SDI principles, with perspective on integration of data, processing, visualisation and simulation capabilities.

The developed hydrogeology infrastructure extends the notion of a traditional SDI as its interoperability principles include geoprocessing, modelling and visualisation as fundamental considerations. This approach enabled the presentation of hydrogeological data and models, by overcoming the challenges inherent in data silos and in preparing hydrogeological models (due to diversity of data formats and legacy scientific codes) while still acknowledging organisational boundaries.

An important aspect of the metadata handling suggested in this study allows for improved IP management and more transparent considerations for the ownership of data and metadata, through integrated metadata propagation through the context-based data referencing. Before, data was held back because of unclear conditions. But now, with provenance and data quality at hand for the user to address, assess and communicate, data use scenarios become more transparent, reliable, and reproducible.

Furthermore, the application of the overall theory will support policy making as well as provide a common ground for competitive groundwater management and use. The importance of integrated transdisciplinary knowledge capitalisation with respect to groundwater-related data has also been acknowledged.

The design theory as well as the implementation and evaluation pathway of this study are built on the shoulder of giants: OGC, ISO and W3C standards, as well as the incredible variety and sophistication of available free and open source software implementations.

During the community interactions I learned that for New Zealand OGC and EODP support is now planned for many of the tools to web-enable the hydro(geo)logical data management systems currently in use in New Zealand.

Through on-going cooperation with the community stakeholders, interest increased throughout the environmental data community in New Zealand and sparked the foundation of a joint cross-agency working group to formalise available standards for data services into a refined agreed upon behaviour and improve the EODP specification with community profiles.

8.3 OUTLOOK AND FUTURE WORK

In the future it should be possible in New Zealand to link and display surface hydrological models like Overseer, Topnet (Topmodel), hydrogeological models, and uncertainty or so called data worth models, for water management zones and catchments in a collated web map view with available data and other research outputs. Then it would be possible to evaluate what data has most or least impact and what data collection would be most important to improve uncertainty, or which hydro(geo)logical models requires the same addressable data set catalogue access visualise the immediate model linkage.

The approach of using OGC OWC context document encoding comprehensively as an abstract data package could be developed further for more flexibility (especially in describing utilised WPS processes). The developed WPS dispatch process for the X3D generation could be generalised to handle OWC collections to provide the necessary input linking for WPS processes in general.

With the release of the OGC WPS 2.0 specification for example, the OGC SensorML encoding could now be employed as a supported description format for WPS process parameters. WPS input fields could be describe as SWE data record. SWE components provide more complex data in- and output description and could be readily interlinked with Linked Data Vocabularies. This would allow a rich set of semantics.

Furthermore, if WPS processes, their inputs and consequently their outputs could be expressed in such clear type descriptions, a functional composition of WPS processes becomes possible. But not in the object oriented way of a model builder or a business process workflow, but in a scientific modelling way. Then reproducibility is a matter of an OWC document and the accessibility of the addressed resources and processes due to declarative model and data capture in a joint context.

The second logical projection is based on the theory produced during this thesis and understanding of the rendering process from diverse source data into the X3D scene. The explicit handling of the generation of the 3D scene as a WPS process sharpens the profile of a geovisualisation pipeline. It also seems a logical conclusion to follow the paradigm of the web infrastructure, and thus the 3D web capabilities that WebGL provide.

Furthermore, the declarative method of transforming (not exclusively) GML to X3D could be examined in more detail. The rendering from WFS features into WMS images is still a challenge, too, because there is no reasonable strategy so far to select properties from complex features for a rendering pipeline in WMS. The way how WMS rendering works in conjunction with SLD and SE forms a logical new foundation for 3D rendering. Translated into 3D rendering of geometric elements (simple features, a geometry, select attribute to apply style) this should be applied in particular to advance the declarative directives for 3D shapes (in reconciliation with X3D for example).

Eventually, there is also more research necessary to incorporate ontologies and preferably enable the inference for dynamic mapping of OGC geodata elements into X3D models, again with consideration of with SLD and SE for further convergence.

Finally, in this study free and open source products were considered as well as proprietary and commercial software products. More importantly, publicly accessible national and international standards were employed to set-up a New Zealand groundwater data and knowledge framework. The resulting theory and its implementation illustrate a convergence of the distilled design elements. The development of such a grounded design theory for a Hydrogeology Infrastructure and its evaluation have not been exercised to that extent in New Zealand. The general feasibly was demonstrated in the course of this research and is promising a paradigm shift in the understanding and development of analytic geodata infrastructures.

BIBLIOGRAPHY

Akyildiz, I. F., Su, W., Sankarasubramaniam, Y., & Cayirci, E. (2002). Wireless sensor networks: a survey. *Computer Networks*, *38*(4), 393–422. doi: 10.1016/S1389-1286(01)00302-4

Albrecht, J. (1999). Geospatial information standards. A comparative study of approaches in the standardisation of geospatial information. *Computers & Geosciences*, *25*(1), 9–24. doi: 10.1016/S0098-3004(98)00102-2

Ames, D. P., Horsburgh, J. S., Cao, Y., Kadlec, J., Whiteaker, T., & Valentine, D. (2012). HydroDesktop: Web services-based software for hydrologic data discovery, download, visualization, and analysis. *Journal of Environmental Modelling & Software*, *37*(0), 146–156. doi: 10.1016/j.envsoft.2012.03.013

Andrews, F. T., Croke, B. F. W., & Jakeman, A. J. (2011). An open software environment for hydrological model assessment and development. *Journal of Environmental Modelling & Software*, *26*(10), 1171–1185. doi: 10.1016/j.envsoft.2011.04.006

Antoine Isaac, V. U. A., & Ed Summers, L. O. C. (2008). *SKOS Simple Knowledge Organization System Primer* (W3C Workin ed.). Retrieved from http://www.w3.org/TR/2008/WD-skos-primer-20080221/

Arheimer, B., Wallman, P., Donnelly, C., Nystrom, K., & Pers, C. (2011). E-HypeWeb: Service for Water and Climate Information - and Future Hydrological Collaboration across Europe? In J. Hrebicek, G. Schimak, & R. Denzer (Eds.), *Environmental software systems: Frameworks of eenvironment* (Vol. 359, pp. 657–666).

Atkinson, R., Dornblut, I., & Smith, D. (2012). An international standard conceptual model for sharing references to hydrologic features. *Journal of Hydrology*(0). doi: 10.1016/j.jhydrol.2011.12.002

Bailey, J. E., & Chen, A. (2011). The role of Virtual Globes in geoscience. *Computers & Geosciences*, *37*(1), 1–2. doi: 10.1016/j.cageo.2010.06.001

Baker, T., Bechhofer, S., Isaac, A., Miles, A., Schreiber, G., & Summers, E. (2013). Key choices in the design of Simple Knowledge Organization System (SKOS). *Journal of Web Semantics*, *20*, 35–49. doi: 10.1016/j.websem.2013.05.001

Ballagh, L. M., Raup, B. H., Duerr, R. E., Khalsa, S. J. S., Helm, C., Fowler, D., & Gupte, A. (2011). Representing scientific data sets in KML: Methods and challenges. *Computers & Geosciences*, *37*(1), 57–64. doi: 10.1016/j.cageo.2010.05.004

Bandaragoda, C., Tarboton, D., & Maidment, D. (2006, jan). Hydrology's efforts toward the cyberfrontier. *Eos, Transactions American Geophysical Union*, *87*(1), 2–6. doi: 10.1029/2006eo010005

Baronti, P., Pillai, P., Chook, V. W. C., Chessa, S., Gotta, A., & Hu, Y. F. (2007). Wireless sensor networks: A survey on the state of the art and the 802.15.4 and ZigBee standards. *Computer Communications*, *30*(7), 1655–1695. doi: 10.1016/j.comcom.2006.12.020

Bastin, L., Cornford, D., Jones, R., Heuvelink, G. B. M., Pebesma, E., Stasch, C., . . . Williams, M. (2013). Managing uncertainty in integrated environmental modelling: The UncertWeb framework. *Journal of Environmental Modelling & Software*, *39*(0), 116–134. doi: 10.1016/j.envsoft.2012.02.008

Batcheller, J. K., & Reitsma, F. (2010). Implementing feature level semantics for spatial data discovery: Supporting the reuse of legacy data using open source components. *Computers, Environment and Urban Systems*, *34*(4), 333–344. doi: 10.1016/j.compenvurbsys.2010.02.008

Beard, D. (2006). Using VRML to share large volumes of complex 3D geoscientific information via the Web. In *Web3d 2006 proceedings of the eleventh international conference on 3d web technology* (pp. 163–167). Columbia, Maryland: ACM. doi: 10.1145/1122591.1122614

Beck, R., Weber, S., & Gregory, R. W. (2013). Theory-generating design science research. *Information Systems Frontiers*, *15*(4), 637–651. doi: 10.1007/s10796-012-9342-4

Behr, J., Eschler, P., Jung, Y., & Zöllner, M. (2009). X3DOM: a DOM-based HTML5/X3D integration model. In *Web3d 2009 proceedings of the 14th international conference on 3d web technology* (pp. 127–135). Darmstadt, Germany: ACM. doi: 10.1145/1559764.1559784

Behr, J., Jung, Y., Franke, T., & Sturm, T. (2012, aug). In *Proceedings of the 17th international conference on 3d web technology - web3d '12* (p. 17). New York, New York, USA: ACM Press. doi: 10.1145/2338714.2338717

Beran, B., & Piasecki, M. (2009). Engineering new paths to water data. *Computers & Geosciences*, *35*(4), 753–760. doi: 10.1016/j.cageo.2008.02.017

Bermudez, L., Cook, T., Forrest, D., Bogden, P., Galvarino, C., Bridger, E., . . . Graybeal, J. (2009). Web feature service (WFS) and sensor observation service (SOS) comparison to publish time series data. In *Collaborative technologies and systems, 2009. cts '09. international symposium on* (pp. 36–

43). doi: 10.1109/cts.2009.5067460

Bhaskar, R. (1998). Philosophy and Scientific Realism. In *Critical realism: Essential readings* (pp. 16–47). doi: 10.1163/156920601794750701

Birks, D. F., Fernandez, W., Levina, N., & Nasirin, S. (2013). Grounded theory method in information systems research: its nature, diversity and opportunities. *European Journal of Information Systems*, *22*(1), 1–8. doi: 10.1057/ejis.2012.48

Blaney, H. F., & Criddle, W. D. (1950). Determining Water Requirements in Irrigated Areas from Climatological and Irrigation Data. *US Dep. Agr., Soil Sci. Techn. Paper No. 96.*.

Blaschke, T., Strobl, J., & Donert, K. (2011). Geographic Information Science: Building a Doctoral Programme Integrating Interdisciplinary Concepts and Methods. *Procedia - Social and Behavioral Sciences*, *21*, 139–146. doi: 10.1016/j.sbspro.2011.07.013

Boisvert, E., & Brodaric, B. (2012). GroundWater Markup Language (GWML) - enabling groundwater data interoperability in spatial data infrastructures. *Journal of Hydroinformatics*, *14*(1), 93–107.

Booch, G. (2010). Enterprise architecture and technical architecture. *IEEE Software*, *27*(2). doi: 10.1109/MS.2010.42

Botts, M., Percivall, G., Reed, C., & Davidson, J. (2008). OGC® Sensor Web Enablement: Overview and High Level Architecture. In S. Nittel, A. Labrinidis, & A. Stefanidis (Eds.), *Geosensor networks* (Vol. 4540, pp. 175–190). Springer Berlin Heidelberg. doi: 10.1007/978-3-540-79996-2_10

Brodaric, B., Boisvert, E., Lucido, J., Simons, B., Dahlhaus, P., Wagner, B., ... Kmoch, A. (2016). *OGC GroundWaterML 2 - GW2IE FINAL REPORT* (Tech. Rep.). Open Geospatial Consortium. Retrieved from https://portal.opengeospatial.org/files/?artifact_id=64688

Brodaric, B., Booth, N., Boisvert, E., & Lucido, J. (2015, oct). Groundwater data network interoperability. *Journal of Hydroinformatics*, jh2015242. doi: 10.2166/hydro.2015.242

Brodaric, B., Sharpe, D., Boisvert, E., Logan, C., Russell, H., Julien, H., ... Létourneau, F. (2011). *Groundwater Information Network: Recent Developments and Future Directions.* Quebec, Canada, 28.-31.08.2011.

Bröring, A., Echterhoff, J., Jirka, S., Simonis, I., Everding, T., Stasch, C., ... Lemmens, R. (2011). New Generation Sensor Web Enablement. *Sensors*, *11*(3), 2652–2699.

Buccella, A., Cechich, A., Gendarmi, D., Lanubile, F., Semeraro, G., & Colagrossi, A. (2011). Building a global normalized ontology for integrating geographic data sources. *Computers & Geosciences*, *37*(7), 893–916. doi: 10.1016/j.cageo.2011.02.022

C. Reed, C. T. O. (2009). *The OGC and REST* (Tech. Rep.). Retrieved from portal.opengeospatial.org/files/?artifact{_}id=32222?

Car, N. J., Cox, S. J. D., Stenson, M. P., Atkinson, R. A., & Fitch, P. (2015). A Services' Frameworks And Support Services For Environmental Information Communities. In *Proceedings of the 11th conference on hydroinformatics, hic 2014.* Retrieved from http://hdl.handle.net/123456789/1545

Carleton, C. J., Dahlgren, R. A., & Tate, K. W. (2005). A relational database for the monitoring and analysis of watershed hydrologic functions: I. Database design and pertinent queries. *Computers & Geosciences*, *31*(4), 393–402. doi: 10.1016/j.cageo.2004.10.007

Castronova, A. M., Goodall, J. L., & Elag, M. M. (2013). Models as web services using the Open Geospatial Consortium (OGC) Web Processing Service (WPS) standard. *Journal of Environmental Modelling & Software*, *41*(0), 72–83. doi: 10.1016/j.envsoft.2012.11.010

Castronova, A. M., Goodall, J. L., & Ercan, M. B. (2013). Integrated modeling within a Hydrologic Information System: An OpenMI based approach. *Journal of Environmental Modelling & Software*, *39*(0), 263–273. doi: 10.1016/j.envsoft.2012.02.011

Chen, D., Shams, S., Carmona-Moreno, C., & Leone, A. (2010). Assessment of open source GIS software for water resources management in developing countries. *Journal of Hydro-environment Research*, *4*(3), 253–264. doi: 10.1016/j.jher.2010.04.017

Chen, J., Díaz, M., Rubio, B. B., Troya, J. M. J. M., Diaz, M., Rubio, B. B., & Troya, J. M. J. M. (2013, jun). PS-QUASAR: A publish/subscribe QoS aware middleware for Wireless Sensor and Actor Networks. *Journal of Systems and Software*, *86*(6), 1650–1662. doi: 10.1016/j.jss.2013.02.028

Chen, N., Hu, C., Chen, Y., Wang, C., & Gong, J. (2012). Using SensorML to construct a geoprocessing e-Science workflow model under a sensor web environment. *Computers & Geosciences*, *47*(0), 119–129. doi: 10.1016/j.cageo.2011.11.027

Chen, N., Wang, X., & Yang, X. (2013). A direct registry service method for sensors and algorithms based on the process model. *Computers & Geosciences*, *56*(0), 45–55. doi: 10.1016/j.cageo.2013.03.008

Choor alil, V. S., & Gopi nathan, E. (2015). A Semantic Web query Optimization Using Resource Description Framework. *Procedia Computer Science*, *70*, 723–732. doi: 10.1016/j.procs.2015.10.110

Cools, J., Broekx, S., Vandenberghe, V., Sels, H., Meynaerts, E., Vercaemst, P., ... Huygens, M. (2011). Coupling a hydrological water quality model and an economic optimization model to set up a cost-effective emission reduction scenario for nitrogen. *Journal of Environmental Modelling & Software*, *26*(1), 44–51. doi: 10.1016/j.envsoft.2010.04.017

Couclelis, H. (2009). Ontology, Epistemology, Teleology: Triangulating Geographic Information Science. In *Research trends in geographic information science* (pp. 3–15). doi: 10.1007/978-3-540-88244-2

Cox, M. E., James, A., Hawke, A., & Raiber, M. (2013). Groundwater Visualisation System (GVS): A software framework for integrated display and interrogation of conceptual hydrogeological models, data and time-series animation. *Journal of Hydrology*, *491*(0), 56–72. doi: 10.1016/j.jhydrol.2013.03.023

Cox, S. J. D., & Botts, M. (2007). *OpenGIS SensorML Encoding Standard* (Tech. Rep.). Open Geospatial Consortium. Retrieved from http://www.opengeospatial.org/standards/sensorml

Cox, S. J. D., Simons, B., & Yu, J. (2014). A harmonised vocabulary for water quality. In *Proceedings, 11th international conference on hydroinformatics - hic 2014.* doi: 10.13140/RG.2.1.2490.4404

Creswell, J. W. (2003). *Research Design: Qualitative, Quantitative. and Mixed Methods Approaches* (2nd Editio ed.). Nebraska: SAGE Publications. Retrieved from http://isites.harvard.edu/fs/docs/icb.topic1334586.files/2003{_}Creswell{_}AFrameworkforDesign.pdf

Creswell, J. W., & Miller, D. (2000). Determining validity in qualitative inquiry. *Theory into practice*, *39*(July 2015), 124–130. doi: 10.1207/s15430421tip3903

Crompvoets, J., Vancauwenberghe, G., Bouckaert, G., & Vandenbroucke, D. (2011). *Practices to Develop Spatial Data Infrastructures: Exploring the Contribution to E-Government* (S. Assar, I. Boughzala, & I. Boydens, Eds.). New York: Springer. doi: 10.1007/978-1-4419-7533-1_13

Crout, N. M. J., Tarsitano, D., & Wood, A. T. (2009). Is my model too complex? Evaluating model formulation using model reduction. *Journal of Environmental Modelling & Software*, *24*(1), 1–7. doi: 10.1016/j.envsoft.2008.06.004

Cruz, S. A. B., Monteiro, A. M. V., & Santos, R. (2012). Automated geospatial Web Services composition based on geodata quality requirements. *Computers & Geosciences*, *47*(0), 60–74. doi: 10.1016/j.cageo.2011.11.020

Dahlhaus, P., MacLeod, A., & Thompson, H. (2011). *UB Spatial - An interoperable groundwater database* (Vol. Programme). Rydges Tradewinds, Cairns. Retrieved from http://www.vvg.org.au/cb{_}pages/files/Cairns{_}poster.pdf

Dahlhaus, P., Murphy, A., MacLeod, A., Thompson, H., McKenna, K., & Ollerenshaw, A. (2015, jul). Making the invisible visible: the impact of federating groundwater data in Victoria, Australia. *Journal of Hydroinformatics*, jh2015169. doi: 10.2166/hydro.2015.169

Daughney, C. J., & Reeves, R. R. (2006). Analysis of temporal trends in New Zealand's groundwater quality based on data from the National Groundwater Monitoring Programme. *Journal of Hydrology (NZ)*, *45*(1), 41–62.

de Andrade, F. G., Baptista, C. d. S., & Leite Jr, F. L. (2011). Using Federated Catalogs to Improve Semantic Integration among Spatial Data Infrastructures. *Transactions in GIS*, *15*(5), 707–722. doi: 10.1111/j.1467-9671.2011.01286.x

de Groot, R. (2006). Function-analysis and valuation as a tool to assess land use conflicts in planning for sustainable, multi-functional landscapes. *Journal of Landscape and Urban Planning*, *75*, 175–186. doi: 10.1016/j.landurbplan.2005.02.016

de Jesus, J., Walker, P., Grant, M., & Groom, S. (2012, oct). WPS orchestration using the Taverna workbench: The eScience approach. *Computers & Geosciences*, *47*, 75–86. doi: 10.1016/j.cageo.2011.11.011

Devaraju, A., Kuhn, W., & Renschler, C. S. (2015). A formal model to infer geographic events from sensor observations. *International Journal of Geographical Information Science*, *29*(1), 1–27. doi: 10.1080/13658816.2014.933480

Díaz, L., Bröring, A., McInerney, D., Libertá, G., & Foerster, T. (2013). Publishing sensor observations into Geospatial Information Infrastructures: AÂ use case in fire danger assessment. *Journal of Environmental Modelling & Software*, *48*(0), 65–80. doi: 10.1016/j.envsoft.2013.06.002

Dobson, P. J. (2002). Critical realism and information systems research : why bother with philosophy ? *Information Research*, *7*(2), 1–16. Retrieved from http://www.informationr.net/ir/7-2/paper124.html

Dongarra, J. (1994). Preface, Introduction To MPI. *International Journal of High Performance Computing Applications*, *3-4*(8), 167-168. doi: 10.1177/109434209400800301

Dubé, L., & Paré, G. (2003). Rigor in Information Systems Positivist Case Research: Current Practices, Trends, and Recommendations. *MIS Quarterly*, *27*(4), 597–635. doi: 10.2307/30036550

Dublin Core: DCMI Metadata Terms. (2010). Retrieved 2016-07-27, from http://dublincore.org/documents/2010/10/11/dcmi-terms/

Easterbrook, S., Singer, J., Storey, M.-A., & Damian, D. (2008). Selecting Empirical Methods for Software Engineering Research. In *Guide to advanced empirical software engineering* (pp. 285–311). London: Springer London. doi: 10.1007/978-1-84800-044-5_11

Elliott, S., Ritchie, A., & Snow, V. (2012). *Frameworks for interoperable freshwater models. Initial evaluation.* (Vol. Client Rep; Tech. Rep.). Retrieved from https://teamwork.niwa.co.nz/download/attachments/10158088/FIFM+Screening+report+final.pdf

Erickson, J., & Siau, K. (2008). Web Services, Service-Oriented Computing, and Service-Oriented Architecture: Separating Hype from Reality. *Journal of Database Management*, *19*(3), 42–54. doi: 10.4018/jdm.2008070103

EU Directive. (2007). *Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE).* Retrieved from http://tinyurl.com/65pzem5

Evangelidis, K., Ntouros, K., Makridis, S., & Papatheodorou, C. (2014, feb). Geospatial services in the Cloud. *Computers & Geosciences*, *63*, 116–122. doi: 10.1016/j.cageo.2013.10.007

Evans, A., Romeo, M., Bahrehmand, A., Agenjo, J., & Blat, J. (2014, jun). 3D graphics on the web: A survey. *Computers & Graphics*, *41*(0), 43–61. doi: 10.1016/j.cag.2014.02.002

Farnaghi, M., & Mansourian, A. (2013, sep). Disaster planning using automated composition of semantic OGC web services: A case study in sheltering. *Computers, Environment and Urban Systems*, *41*, 204–218. doi: 10.1016/j.compenvurbsys.2013.06.003

Feast, L., & Melles, G. (2010). Epistemological Positions in Design Research: A Brief Review of the Literature. In *Proceedings of the international conference on design education (connected)* (pp. 1–5). Retrieved from http://swinburne.academia.edu/attachments/1357366/download

Fielding, R. T. (2000). *Architectural Styles and the Design of Network-based Software Architectures* (PhD Dissertation, University of California, Irvine). Retrieved from https://www.ics.uci.edu/{~}fielding/pubs/dissertation/fielding{_}dissertation.pdf

Gahegan, M., Luo, J., Weaver, S. D., Pike, W., & Banchuen, T. (2009). Connecting GEON: Making sense of the myriad resources, researchers and concepts that comprise a geoscience cyberinfrastructure. *Computers & Geosciences*, *35*(4), 836–854. doi: 10.1016/j.cageo.2008.09.006

Gao, J., Gao, Y., Zhao, G., & Hörmann, G. (2010). Minimum ecological water depth of a typical stream in Taihu Lake Basin, China. *Quaternary International*, *226*(1-2), 136–142. doi: 10.1016/j.quaint.2010.03.004

GeoSciML. (2011). *GeoSciML.* Retrieved from http://www.geosciml.org/

Ghobakhlou, A., Kmoch, A., & Sallis, P. (2013). Integration of Wireless Sensor Network and Web Services. In J. Piantadosi, Anderssen, R. S., & J. Boland (Eds.), *Modsim2013, 20th international congress on modelling and simulation. modelling and simulation society of australia and new zealand, december 2013, isbn: 978-0-9872143-3-1.* doi: 10.13140/2.1.3508.0001

GIN. (2011). *Groundwater Information Network (GIN).* http://gw-info.net/: Natural Resources Canada.

Giuliani, G., Nativi, S., Lehmann, A., & Ray, N. (2012). WPS mediation: An approach to process geospatial data on different computing backends. *Computers & Geosciences*, *47*(0), 20–33. doi: 10.1016/j.cageo.2011.10.009

Giuliani, G., Ray, N., & Lehmann, A. (2011). Grid-enabled Spatial Data Infrastructure for environmental sciences: Challenges and opportunities. *Future Generation Computer Systems*, *27*(3), 292–303. doi: 10.1016/j.future.2010.09.011

Glaser, B. G. (2008). Conceptualization: On theory and theorizing using grounded theory. *International Journal of Qualitative Methods*, *1*, 23–38. doi: 10.1177/160940690200100203

Glaser, B. G., & Strauss, A. L. (1967). The discovery of grounded theory. *International Journal of Qualitative Methods*, *5*, 1–10. doi: 10.2307/588533

Glimm, B., Horridge, M., Parsia, B., & Patel-Schneider, P. F. (n.d.). *A Syntax for Rules in OWL 2.* Retrieved from http://webont.org/owled/2009/papers/owled2009{_}submission{_}16.pdf

Goles, T., & Hirschheim, R. (2000). The paradigm is dead, the paradigm is dead…long live the paradigm: the legacy of Burrell and Morgan. *Omega*, *28*(3), 249–268. doi: 10.1016/S0305-0483(99)00042-0

Goodall, J. L., Horsburgh, J. S., Whiteaker, T. L., Maidment, D. R., & Zaslavsky, I. (2008). A first approach to web services for the National Water Information System. *Journal of Environmental Modelling & Software*, *23*(4), 404–411. doi: 10.1016/j.envsoft.2007.01.005

Goodall, J. L., Robinson, B. F., & Castronova, A. M. (2011). Modeling water resource systems using a service-oriented computing paradigm. *Environmental Modelling & Software*, *26*(5), 573–582. doi: 10.1016/j.envsoft.2010.11.013

Goodchild, M. F. (1992, jan). Geographical information science. *International journal of geographical information systems*, *6*(1), 31–45. doi: 10.1080/02693799208901893

Goodchild, M. F. (2006, nov). GIScience Ten Years After Ground Truth. *Transactions in GIS*, *10*(5), 687–692. doi: 10.1111/j.1467-9671.2006.01022.x

Goodchild, M. F. (2009, nov). Geographic information systems and science: today and tomorrow. *Procedia Earth and Planetary Science*, *15*(1), 1037–1043. doi: 10.1080/19475680903250715

Grabaum, O., & Meyer, B. (1998). Multicriteria optimization of landscapes using GIS-based functional assessments. *Journal of Landscape and Urban Planning*, *43*, 21–34.

Granell, C., Díaz, L., & Gould, M. (2010). Service-oriented applications for environmental models: Reusable geospatial services. *Environmental Modelling & Software*, *25*(2), 182–198. doi: 10.1016/j.envsoft.2009.08.005

Granell, C., Díaz, L., Schade, S., Ostländer, N., & Huerta, J. (2013). Enhancing integrated environmental modelling by designing resource-oriented interfaces. *Journal of Environmental Modelling & Software*, *39*(0), 229–246. doi: 10.1016/j.envsoft.2012.04.013

Granell, C., Havlik, D., Schade, S., Sabeur, Z., Delaney, C., Pielorz, J., ... Mon, J. L. (2016, apr). Future Internet technologies for environmental applications. *Environmental Modelling & Software*, *78*, 1–15. doi: 10.1016/j.envsoft.2015.12.015

Gray, D. E. (2014). Theoretical perspectives and research methodologies. In *Doing research in the real world* (pp. 16–38). Retrieved from http://www.uk.sagepub.com/books/Book239646{#}tabview=toc

Green, C. H., & van Griensven, A. (2008). Autocalibration in hydrologic modeling: Using SWAT2005 in small-scale watersheds. *Journal of Environmental Modelling & Software*, *23*(4), 422–434. doi: 10.1016/j.envsoft.2007.06.002

Gregor, S. (2002). Design Theory in Information Systems. *Australasian Journal of Information Systems*(December), 14–22. doi: 10.3127/ajis.v10i1.439

Gregor, S. (2006). The nature of theory in information systems. *MIS Quarterly*, *30*(3), 611–642. doi: 10.1080/0268396022000017725

Gregor, S., & Hevner, A. R. (2013). Positioning and Presenting Design Science Research for Maximum Impact. *MIS Quarterly*, *37*(2), 337–355. doi: 10.2753/MIS0742-1222240302

Gregory, R. W. (2010). Design Science Research and the Grounded Theory Method : Characteristics, Differences , and Complementary Uses. *Proceedings of the 18th European Conference on Information Systems (ECIS)*, 111–127. doi: 10.1007/978-3-7908-2781-1_6

Gröger, G., Kolbe, T., Nagel, C., & Häfele, K.-H. (2012). *OGC City Geography Markup Language (CityGML) En-coding Standard* (Tech. Rep.). Open Geospatial Consortium. Retrieved from https://portal.opengeospatial.org/files/?artifact{_}id=47842

Gröger, G., & Plümer, L. (2012). *CityGML - Interoperable semantic 3D city models* (Vol. 71). doi: 10.1016/j.isprsjprs.2012.04.004

Harbaugh, A. (2005). *MODFLOW-2005, the U.S. Geological Survey modular ground-water model - the Ground-Water Flow Process.* U.S. Geological Survey.

Harding, J., Mosley, P., Pearson, C., & Sorrell, B. (2004). *Freshwaters of New Zealand* (N. Z. H. S. I. Inc. & N. Z. L. Society, Eds.). Wellington.

Hart, J. K., & Martinez, K. (2006). Environmental Sensor Networks: A revolution in the earth system science? *Earth-Science Reviews*, *78*(3-4), 177–191. doi: 10.1016/j.earscirev.2006.05.001

Heavner, M. J., Fatland, D. R., Hood, E., & Connor, C. (2011). SEAMONSTER: A demonstration sensor web operating in virtual globes. *Computers & Geosciences*, *37*(1), 93–99. doi: 10.1016/j.cageo.2010.05.011

Henson, C. A., Pschorr, J. K., Sheth, A. P., & Thirunarayan, K. (2009). SemSOS: Semantic sensor Observation Service. In *Collaborative technologies and systems, 2009. cts '09. international symposium on* (pp. 44–53). doi: 10.1109/cts.2009.5067461

Hevner, A., & Chatterjee, S. (2010). *Design Research in Information Systems: Theory and Practice* (Vol. 22). doi: 10.1007/978-1-4419-5653-8

Hildebrandt, D., & Döllner, J. (2010). Service-oriented, standards-based 3D geovisualization: Potential and challenges. *Computers, Environment and Urban Systems*, *34*(6), 484–495. doi: 10.1016/j.compenvurbsys.2010.05.003

Hinkelmann, K., Gerber, A., Karagiannis, D., Thoenssen, B., Van Der Merwe, A., & Woitsch, R. (2016). A new paradigm for the continuous alignment of business and IT: Combining enterprise architecture modelling and enterprise ontology. *Computers in Industry*, *79*, 77–86. doi: 10.1016/j.compind.2015.07.009

Hirschheim, R. a. (1985). Information Systems Epistemology: An Historical Perspective. *Research Methods in Information Systems*, 13–38.

Holliday, A. (2010). Analysing qualitative data. In B. Phakiti & A. Paltridge (Eds.), *Continuum companion to research methods in applied linguistics* (pp. 98–110). London: Continuum International Publishing Group.

Horak, J., Orlik, a., & Stromsky, J. (2007). Web services for distributed and interoperable hydro-information systems. *Hydrology and Earth System Sciences Discussions*, *4*(3), 1879–1891. doi: 10.5194/hessd-4-1879-2007

Horrocks, I., Patel-Schneider, P. F., & van Harmelen, F. (2003). From SHIQ and RDF to OWL: the making of a Web Ontology Language. *Web Semantics: Science, Services and Agents on the World Wide Web*, *1*(1), 7–26. doi: 10.1016/j.websem.2003.07.001

Horsburgh, J. S., Tarboton, D. G., Hooper, R. P., & Zaslavsky, I. (2014). Managing a community shared vocabulary for hydrologic observations. *Journal of Environmental Modelling & Software*, *52*(0), 62–73. doi: 10.1016/j.envsoft.2013.10.012

Horsburgh, J. S., Tarboton, D. G., Piasecki, M., Maidment, D. R., Zaslavsky, I., Valentine, D., & Whitenack, T. (2009). An integrated system for publishing environmental observations data. *Journal of Environmental Modelling & Software*, *24*(8), 879–888. doi: 10.1016/j.envsoft.2009.01.002

Hughes, J., Bakker, M., White, J., Langevin, C., Post, V., Fienen, M., & Starn, J. (2015). FloPy Version 3 - a Python package for MODFLOW-based models. In *Modflow and more 2015: Modeling a complex world.* Golden, Colorado, May 31-June 3, 2015.

Hunkeler, U., Lombriser, C., Truong, H. L., & Weiss, B. (2013). A case for centrally controlled wireless sensor networks. *Computer Networks*, *57*(6), 1425–1442. doi: 10.1016/j.comnet.2012.12.019

Informatics, A. (2015). *AQUARIUS Time-Series* (Vol. 2015) (No. 15.05.2015). Retrieved from http://aquaticinformatics.com/products/aquarius-time-series/

IQUEST. (2011). *The environmental monitoring blog. http://www.iquest.co.nz/environmental-monitoring-blog/index.php/2011/10/05/kisters-using-new-ogc-standards* (No.

15.05.2013). Retrieved from http://www.iquest.co.nz/environmental-monitoring-blog/index.php/2011/10/05/kisters-using-new-ogc-standards

ISO 19103. (2005). *Geographic information - Conceptual schema language.*

ISO 19115. (2003). *Geographic information - Metadata.*

ISO 19118. (2011). *Geographic information - Encoding.*

ISO 19136. (2007). *Geographic information - Geography Markup Language (GML).*

ISO 19139. (2007). *Geographic Information - Metadata - XML schema implementation (Encoding of Metadata).*

ISO 19156. (2011). *Geographic information - Observations and Measurements.*

ISO/IEC. (1997). *Virtual Reality Modeling Language - ISO/IEC 14772-1:1997 .* http://www.web3d.org/x3d/specifications/vrml/ISO-IEC-14772-VRML97/: Author.

ISO/IEC. (2008). *Extensible 3D (X3D) - ISO/IEC 19775-1.2:2008.* http://www.web3d.org/x3d/specifications/: Author.

Iwanaga, T., El Sawah, S., & Jakeman, A. (n.d.). Design and implementation of a web-based groundwater data management system. *Mathematics and Computers in Simulation*(0). doi: 10.1016/j.matcom.2012.11.009

Janowicz, K., Scheider, S., Pehle, T., & Hart, G. (2012). Geospatial semantics and linked spatiotemporal data - Past, present, and future. *Semantic Web*, *3*, 321–332. doi: 10.3233/SW-2012-0077

Jasiewicz, J., & Metz, M. (2011). A new GRASS GIS toolkit for Hortonian analysis of drainage networks. *Computers & Geosciences*, *37*(8), 1162–1173. doi: 10.1016/j.cageo.2011.03.003

Jazayeri, M., Liang, S., & Huang, C.-Y. (2015, sep). Implementation and Evaluation of Four Interoperable Open Standards for the Internet of Things. *Sensors*, *15*(9), 24343–24373. doi: 10.3390/s150924343

Jirka, S., Bröring, A., Kjeld, P., Maidens, J., & Wytzisk, A. (2012). A Lightweight Approach for the Sensor Observation Service to Share Environmental Data across Europe. *Transactions in GIS*, *16*(3), 293–312. doi: 10.1111/j.1467-9671.2012.01324.x

Jones, D., Jones, N., Greer, J., & Nelson, J. (2015). A cloud-based MODFLOW service for aquifer management decision support. *Computers & Geosciences*, *78*(1), 81–87. doi: 10.1016/j.cageo.2015.02.014

Jorgensen, L. F., Refsgaard, J. C., & Hojberg, A. L. (2007). The inadequacy of monitoring without modelling support. *Journal of Environmental Monitoring*, *9*(9), 931–942. doi: 10.1039/b706617c

Kao, S., Ranatunga, K., Squire, G., Pratt, A., & Dee, D. (2011). Visualisation of hydrological observations in the water data transfer format. *Journal of Environmental Modelling & Software*, *26*(12), 1767–1769. doi: 10.1016/j.envsoft.2011.07.005

Kiehle, C. (2006). Business logic for geoprocessing of distributed geodata. *Computers & Geosciences*, *32*(10), 1746–1757. doi: 10.1016/j.cageo.2006.04.002

Kisters. (2015a). *Hydstra Modules. http://www.kisters.net/hydstra-modules.html.* Retrieved from http://www.kisters.com.au/english/html/au/homepage.html

Kisters. (2015b). *KISTERS Web Interoperability Solution (KiWIS). http://kiwis.kisters.de/KiWIS/* (No. 15.05.2015).

Kisters. (2015c). *Wiski. http://www.kisters.net/wiski.html* (No. 15.05.2015). Retrieved from http://www.kisters.net/wiski.html

Klug, H. (2012). An integrated holistic transdisciplinary landscape planning concept after the Leitbild approach. *Journal Ecological Indicators*, *23*(0), 616–626. doi: 10.1016/j.ecolind.2012.05.019

Klug, H., & Bretz, B. (2012). *Discover INSPIRE compliant harmonised soil data and services. Final product summary of the GS Soil Project on "Assessment and strategic development of INSPIRE compliant Geodata-Services for European Soil Data". Co-funded by the community programme eConten.* Salzburg.

Klug, H., Daughney, C., Verhagen, F., Westerhoff, R., & Dudley Ward, N. (2011). Freshwater resources management: Starting SMART characterization of New Zealand's aquifers. *EarthZine Magazine*, Online Article. Retrieved from http://www.earthzine.org/2011/12/13/freshwater-resources-management-starting-smart-characterization-of-new-zealands-aquifers/

Klug, H., & Kmoch, A. (2014, aug). A SMART groundwater portal: An OGC web services orchestration framework for hydrology to improve data access and visualisation in New Zealand. *Computers & Geosciences*, *69*(0), 78–86. doi: 10.1016/j.cageo.2014.04.016

Klug, H., & Kmoch, A. (2015, jan). Operationalizing environmental indicators for real time multi-purpose decision making and action support. *Ecological Modelling*, *295*(0), 66–74. doi: 10.1016/j.ecolmodel.2014.04.009

Klug, H., Kmoch, A., & Reichel, S. (2015, nov). Adjusting the Frequency of Automated Phosphorus Measurements to Environmental Conditions. *GI_Forum 2015 - Journal for Geographic Information Science - Geospatial*

Minds for Society, *1*, 590–599. doi: 10.1553/giscience2015s590

Kmoch, A., & Klug, H. (2014). Visualization of 3D Hydrogeological Data in the Web. *GI_Forum 2014 - Geospatial Innovation for Society*(2014), 16–24. doi: 10.1553/giscience2014s16.Visualization

Kmoch, A., Klug, H., Ritchie, A. B. H., Schmidt, J., & White, P. A. (2015, oct). A Spatial Data Infrastructure Approach for the Characterization of New Zealand's Groundwater Systems. *Transactions in GIS*, *Early View*, xx–xx. doi: 10.1111/tgis.12171

Knapen, R., Janssen, S., Roosenschoon, O., Verweij, P., de Winter, W., Uiterwijk, M., & Wien, J.-E. (2013). Evaluating OpenMI as a model integration platform across disciplines. *Journal of Environmental Modelling & Software*, *39*(0), 274–282. doi: 10.1016/j.envsoft.2012.06.011

Kragt, M. E., Newham, L. T. H., Bennett, J., & Jakeman, A. J. (2011). An integrated approach to linking economic valuation and catchment modelling. *Journal of Environmental Modelling & Software*, *26*(1), 92–102. doi: 10.1016/j.envsoft.2010.04.002

Krysanova, V., Hattermann, F., & Wechsung, F. (2007). Implications of complexity and uncertainty for integrated modelling and impact assessment in river basins. *Journal of Environmental Modelling & Software*, *22*(5), 701–709. doi: 10.1016/j.envsoft.2005.12.029

Lacasta, J., Nogueras-Iso, J., López-Pellicer, F. J., Muro-Medrano, P. R., & Zarazaga-Soria, F. J. (2007). ThManager: An Open Source Tool for Creating and Visualizing SKOS. *Information Technology and Libraries*, *26*(3), 39–51. Retrieved from http://www.ala.org/lita/ital/sites/ala.org.lita.ital/files/content/26/3/lcst.pdf

Land Information New Zealand. (2007). *Understanding our geographic information landscape. A New Zealand Geospatial Strategy. A coordinated approach to location information.* Retrieved from http://www.linz.govt.nz/docs/geospatial-office/nz-geospatial-strategy-2007.pdf

Laniak, G. F., Olchin, G., Goodall, J., Voinov, A., Hill, M., Glynn, P., ... Hughes, A. (2013). Integrated environmental modeling: A vision and roadmap for the future. *Journal of Environmental Modelling & Software*, *39*(0), 3–23. doi: 10.1016/j.envsoft.2012.09.006

Latre, M. Á., Lopez-Pellicer, F. J., Nogueras-Iso, J., Béjar, R., Zarazaga-Soria, F. J., & Muro-Medrano, P. R. (2013, oct). Spatial Data Infrastructures for environmental e-government services: The case of water abstractions authorisations. *Environmental Modelling & Software*, *48*, 81–92. doi: 10.1016/j.envsoft.2013.06.005

Leighton, B., Cox, S. J. D., Car, N., Stenson, M. P., Vleeshouwer, J., & Hodge, J. (2015). A Best of Both Worlds Approach to Complex, Efficient, Time Series Data Delivery. *Environmental Software Systems. Infrastructures, Services and Applications*, *448*(JANUARY), 371–379. doi: 10.1007/978-3-319-15994-2_37

Li, X., Di, L., Han, W., Zhao, P., & Dadi, U. (2010). Sharing geoscience algorithms in a Web service-oriented environment (GRASS GIS example). *Computers & Geosciences*, *36*(8), 1060–1068. doi: 10.1016/j.cageo.2010.03.004

Liang, S. H. L., Croitoru, A., & Tao, C. V. (2005). A distributed geospatial infrastructure for Sensor Web. *Computers & Geosciences*, *31*(2), 221–231. doi: 10.1016/j.cageo.2004.06.014

Lindström, G., Pers, C. P., Rosberg, R., Strömqvist, J., & Arheimer, B. (2010). Development and test of the HYPE (Hydrological Predictions for the Environment) model - A water quality model for different spatial scales. *Hydrology Research*, *41.3*(4), pp295–319.

Lovett, A. P., & Cameron, S. G. (2014). *Development of a National Groundwater Atlas for New Zealand, GNS Science Report 2014/30* (Tech. Rep. No. 30). Taupo: GNS Science.

Lowry, T. S., Bright, J. C., Close, M. E., Robb, C. A., White, P. A., & Cameron, S. G. (2003). Management gaps analysis: a case study of groundwater resource management in New Zealand. *International Journal of Water Resources Development*, *19*(4), 579–592. doi: 10.1080/0790062032000161382

Lutz, M., J.Sprado, E.Klien, C.Schubert, Christ, i., Sprado, J., ... Schubert, C. (2009). Overcoming semantic heterogeneity in spatial data infrastructures. *Computers &Geosciences*, *35*(4), 739–752. doi: 10.1016/j.cageo.2007.09.017

Ma, X., Carranza, E. J. M., Wu, C., & van der Meer, F. D. (2012, mar). Ontology-aided annotation, visualization, and generalization of geological time-scale information from online geological map services. *Computers & Geosciences*, *40*, 107–119. doi: 10.1016/j.cageo.2011.07.018

Ma, X., Carranza, E. J. M., Wu, C., van der Meer, F. D., & Liu, G. (2011). A SKOS-based multilingual thesaurus of geological time scale for interoperability of online geological maps. *Computers & Geosciences*, *37*(10), 1602–1615. doi: 10.1016/j.cageo.2011.02.011

Maguire, D. J., & Longley, P. A. (2005). The emergence of geoportals and their role in spatial data infrastructures. *Computers, Environment and Urban Systems*, *29*(1), 3–14. doi: 10.1016/j.compenvurbsys.2004.05.012

Makler-Pick, V., Gal, G., Gorfine, M., Hipsey, M. R., & Carmel, Y. (2011). Sensitivity analysis for complex ecological models - A new approach. *Journal of Environmental Modelling & Software*, *26*(2), 124–134. doi: 10.1016/j.envsoft.2010.06.010

Manso, M.-Á., & Wachowicz, M. (2009, may). GIS Design: A Review of Current Issues in Interoperability. *Geography Compass*, *3*(3), 1105–1124. Retrieved from http://doi.wiley.com/10.1111/j.1749-8198.2009.00241.x doi: 10.1111/j.1749-8198.2009.00241.x

Manso, M.-Á., Wachowicz, M., & Bernabé, M.-Á. (2009, feb). Towards an Integrated Model of Interoperability for Spatial Data Infrastructures. *Transactions in GIS*, *13*(1), 43–67. Retrieved from http://doi.wiley.com/10.1111/j.1467-9671.2009.01143.x doi: 10.1111/j.1467-9671.2009.01143.x

March, S. T., & Smith, G. F. (1995, dec). Design and natural science research on information technology. *Decision Support Systems*, *15*(4), 251–266. doi: 10.1016/0167-9236(94)00041-2

Mark, D. M. (2002). Geographic Information Science : Defining the Field. *Foundations of Geographic Information Science*, 3–18. doi: 10.1201/9780203009543.ch1

Matthies, M., Giupponi, C., & Ostendorf, B. (2007). Environmental decision support systems: Current issues, methods and tools. *Journal of Environmental Modelling & Software*, *22*(2), 123–127. doi: 10.1016/j.envsoft.2005.09.005

Meek, S., Jackson, M., & Leibovici, D. G. (2016, feb). A BPMN solution for chaining OGC services to quality assure location-based crowdsourced data. *Computers & Geosciences*, *87*, 76–83. doi: 10.1016/j.cageo.2015.12.003

Minacapilli, M., Iovino, M., & D'Urso, G. (2008). A distributed agro-hydrological model for irrigation water demand assessment. *Journal of Agricultural Water Management*, *95*(2), 123–132. doi: 10.1016/j.agwat.2007.09.008

Mingers, J. (2004, apr). Real-izing information systems: critical realism as an underpinning philosophy for information systems. *Information and Organization*, *14*(2), 87–103. doi: 10.1016/j.infoandorg.2003.06.001

Ministry for the Environment. (2006). *Snapshot of Water Allocation in New Zealand. Prepared by Aqualinc Research Limited.*

Ministry for the Environment New Zealand. (2013). *Resource Management Act 1991* (Vol. 1991 No 69). Wellington, New Zealand: Ministry for the

Environment New Zealand. Retrieved from http://www.legislation.govt.nz/act/public/1991/0069/latest/096be8ed80b45308.pdf

Morozov, I., Chubak, G., & Blyth, S. (2009). Interactive 3D/2D visualization for geophysical data processing and interpretation. *Computers & Geosciences*, *35*(7), 1397–1408. doi: 10.1016/j.cageo.2008.10.005

Mouton, C., Behr, J., Jung, Y., Parfouru, S., Jeulin, C., Dutertre, C., ... Stein, C. (2014, aug). Enhancing the plant layout design process using X3DOM and a scalable web3D service architecture. In *Proceedings of the nineteenth international acm conference on 3d web technologies - web3d '14* (pp. 125–132). New York, New York, USA: ACM Press. doi: 10.1145/2628588.2628592

NIWA. (2013). *TiDeDa - time dependent data. http://www.niwa.co.nz/software/tideda-time-dependent-data* (Vol. 2013) (No. 15.05.2013). Retrieved from http://www.niwa.co.nz/software/tideda-time-dependent-data

OGC. (2007). *OpenGIS® Geography Markup Language (GML) Encoding Standard (ISO 19136), 2007-08-27, OGC 07-036, Version: 3.2.1, Clemens Portele.* http://www.opengeospatial.org/standards/is: The Open Geospatial Consortium (OGC).

OGC. (2015). *OGC Standards.* http://www.opengeospatial.org/standards: Open Geospatial Consortium. Retrieved from http://www.opengeospatial.org/standards

OGC, & Baumann, P. (2006). *OGC WCS 1.1.0 Interface Standard, Web Coverage Service (WCS) Implementation Specification.* http://www.opengeospatial.org/standards/is: The Open Geospatial Consortium (OGC).

OGC, & Baumann, P. (2010). *OGC WCS 2.0 Interface Standard - Core, v2.0.* http://www.opengeospatial.org/standards/is: The Open Geospatial Consortium (OGC).

OGC, & Beaujardiere, J. d. L. (2006). *OpenGIS Web Map Service (ISO 19128) WMS v1.3.0 .* http://www.opengeospatial.org/standards/is: The Open Geospatial Consortium (OGC).

OGC, Brackin, R., & Gonçalves, P. (2014a). *OGC OWS Context Atom Encoding Standard 1.0.*

OGC, Brackin, R., & Gonçalves, P. (2014b). *OGC OWS Context Conceptual Model 1.0* (Vol. OGC IS; Tech. Rep.). Retrieved from http://www.opengeospatial.org/

OGC, Bröring, A., Stasch, C., & Echterhoff, J. (2011). *OGC Sensor Planning Service Implementation Standard.* http://www.opengeospatial.org/standards/sps; http://www.opengis.net/doc/IS/SPS/2.0: The Open Geospatial Consortium (OGC).

OGC, Bröring, A., Stasch, C., & Echterhoff, J. (2012). *OGC Sensor Observation Service Interface Standard, v2.0.* http://www.opengeospatial.org/standards/is: The Open Geospatial Consortium (OGC).

OGC, & Cox, S. J. D. (2007). *Observations and Measurements - Part 2 - Sampling Features.* http://www.opengeospatial.org/standards/is: The Open Geospatial Consortium (OGC). Retrieved from http://portal.opengeospatial.org/files/?artifact{_}id=22467

OGC, & Cox, S. J. D. (2011). *Observations and Measurements - XML Implementation (O&M v2.0 conceptual model also published as ISO/DIS 19156), O&M v2.0.* http://www.opengeospatial.org/standards/is: The Open Geospatial Consortium (OGC). Retrieved from http://www.opengeospatial.org/standards/is

OGC, Liang, S., Huang, A. C.-y., Khalafbeigi, T., & Kim, K. (2015). *OGC SensorThings API Candidate Standard Part 1: Sensing* (Tech. Rep.). Open Geospatial Consortium.

OGC, Mueller, M., & Pross, B. (2015). *OGC WPS 2.0 Interface Standard* (Tech. Rep.). Open Geospatial Consortium.

OGC, Nebert, Whiteside, Vretanos, & Editors. (2007). *OpenGIS Catalogue Service Implementation Specification (ISO 19115), v2.0.2.* http://www.opengeospatial.org/standards/is: The Open Geospatial Consortium (OGC).

OGC, Peter, P., & Vretanos, A. (2010). *OpenGIS Web Feature Service 2.0 (ISO 19142) WFS v2.0.* http://www.opengeospatial.org/standards/is: The Open Geospatial Consortium (OGC).

OGC, & Robin, A. (2011). *OGC SWE Common Data Model Encoding Standard, v2.0.* http://www.opengeospatial.org/standards/is: The Open Geospatial Consortium (OGC).

OGC, Schilling, A., & Kolbe, T. H. (2010). *Draft for Candidate Web 3D Service Interface Standard, v0.4.0.* The Open Geospatial Consortium (OGC). Retrieved from http://www.opengeospatial.org/standards/is

OGC, & Schut, P. (2007). *OpenGIS® Web Processing Service WPS v1.0.0, 2007-06-08, OGC 05-007r7, Peter Schut.*

http://www.opengeospatial.org/standards/is: The Open Geospatial Consortium (OGC).

OGC, & Taylor, P. (2012). *WaterML 2.0: Part 1- Timeseries.* http://www.opengeospatial.org/standards/is: The Open Geospatial Consortium (OGC).

OGC, & Vretanos, P. (2005a). *OGC Web Map Context, WMC 1.1.0 .* http://www.opengeospatial.org/standards/is: The Open Geospatial Consortium (OGC).

OGC, & Vretanos, P. (2005b). *OpenGIS Web Feature Service, WFS v1.1.0 .* http://www.opengeospatial.org/standards/is: The Open Geospatial Consortium (OGC).

Owen, C. (1998). Design research: building the knowledge base. *Design Studies*, *19*(1), 9–20. doi: 10.1016/S0142-694X(97)00030-6

Pahl-Wostl, C. (2007). The implications of complexity for integrated resources management. *Journal of Environmental Modelling & Software*, *22*(5), 561–569. doi: 10.1016/j.envsoft.2005.12.024

Pebesma, E., Cornford, D., Dubois, G., Heuvelink, G. B. M., Hristopulos, D., Pilz, J., ... Skøien, J. O. (2011). INTAMAP: The design and implementation of an interoperable automated interpolation web service. *Computers & Geosciences*, *37*(3), 343–352. doi: 10.1016/j.cageo.2010.03.019

Peckham, S. D., & Goodall, J. L. (2013). Driving plug-and-play models with data from web services: A demonstration of interoperability between CSDMS and CUAHSI-HIS. *Computers & Geosciences*, *53*(0), 154–161. doi: 10.1016/j.cageo.2012.04.019

Perry, M., & Herring, J. (2012). OGC GeoSPARQL-A geographic query language for RDF data. *OGC Candidate Implementation Standard*, 57. Retrieved from http://www.opengis.net/doc/IS/geosparql/1.0

Pickles, J. (1997, jun). Tool or Science?GIS, Technoscience, and the Theoretical Turn. *Annals of the Association of American Geographers*, *87*(2), 363–372. doi: 10.1111/0004-5608.00058

Porra, J., Hirschheim, R., & Parks, M. S. (2014). The Historical Research Method and Information Systems Research. *Journal of the Association for Information Systems*, *15*(9), 536–576.

Portele, C., Hobona, G., Percivall, G., Heylen, C., van der Horn, A. P., Borrebaek, M., ... van der Horn, A. P. (2009). *GIGAS Technical Note - D2.2b Data Harmonisation and Semantic Interoperability* (Tech. Rep.). Retrieved from http://www.thegigasforum.eu/project/material/deliverables.html

Potschin, M., & Klug, H. (2010). Planning landscape visions and its implementation. *Journal Futures*, *42*(7), 653–655. doi: 10.1016/j.futures.2010.04.002

Pulido, J. R. G., Ruiz, M. A. G., Herrera, R., Cabello, E., Legrand, S., & Elliman, D. (2006). Ontology languages for the semantic web: A never completely updated review. *Knowledge-Based Systems*, *19*(7), 489–497. doi: 10.1016/j.knosys.2006.04.013

Raj Shrestha, R., & Rode, M. (2008). Multi-objective calibration and fuzzy preference selection of a distributed hydrological model. *Journal of Environmental Modelling & Software*, *23*(12), 1384–1395. doi: 10.1016/j.envsoft.2008.04.001

Ramos, L. (2015, oct). Semantic Web for Manufacturing, Trends and Open Issues. *Computers & Industrial Engineering*, *90*, 444–460. doi: 10.1016/j.cie.2015.10.013

Ranatunga, K., Walker, G., & Sheahan, P. A. (2011). Implementation of groundwater observations in the Water Data Transfer Format. *Journal of Environmental Modelling & Software*, *26*(4), 549–550. doi: 10.1016/j.envsoft.2010.10.005

Ratto, M., Castelletti, A., & Pagano, A. (2012). Emulation techniques for the reduction and sensitivity analysis of complex environmental models. *Journal of Environmental Modelling & Software*, *34*(0), 1–4. doi: 10.1016/j.envsoft.2011.11.003

Refsgaard, J. C., van der Sluijs, J. P., Højberg, A. L., & Vanrolleghem, P. A. (2007). Uncertainty in the environmental modelling process - A framework and guidance. *Journal of Environmental Modelling & Software*, *22*(11), 1543–1556. doi: 10.1016/j.envsoft.2007.02.004

Resch, B., Wohlfahrt, R., & Wosniok, C. (2014). Web-based 4D visualization of marine geo-data using WebGL. *Cartography and Geographic Information Science*, *41*(3), 235–247. doi: 10.1080/15230406.2014.901901

Ritchie, A., Hodges, S., Kmoch, A., Schmidt, J., Watson, B., & White, P. (2014). *New Zealand Environmental Observation Data Profile* (Tech. Rep.). New Zealand Environmental Observation Data Working Group. Retrieved from https://github.com/EODP-NZ/eodp-dev/blob/master/doc/spec-eodp-core-1.0.pdf

Ritchie, A., Schmidt, J., Hodges, S., Watson, B., Kmoch, A., & White, P. (2014). Development of a New Zealand Hydrology Time Series Data Exchange Standard. In *2014 water symposium - joint conference of the new zealand hydrological society, new zealand freshwater sciences society and the ipenz rivers* *group*. Blenheim. Retrieved from http://www.2014watersymposium.co.nz/asset/library/Ritchie.pdf

Rodgers, M. (2015). *The DataTamer suite. http://www.hilltop.co.nz/* (Vol. 2015) (No. 15.05.2015). Retrieved from http://www.hilltop.co.nz/

Rosen, M. R., & White, P. A. (2001). *Groundwaters of New Zealand* (N. Z. H. S. Inc., Ed.). Wellington.

Rutledge, D., Elliott, S., Snow, V., Ritchie, A., & Turek, G. (2012). *Development of a Prototype Framework to facilitate interoperable Freshwater Modelling in New Zealand*. Nelson.

Saunders, M., Lewis, P., & Thornhill, A. (2012). *Research Methods for Business Students*.

Schimak, G., Rizzoli, A. E., & Watson, K. (2010). Sensors and the environment - Modelling & ICT challenges. *Environmental Modelling & Software*, *25*(9), 975–976. doi: 10.1016/j.envsoft.2010.03.022

Schmidt, B., & May, M. (2012). *TRITURUS: A Java-based Framework for 3D Geo-Visualization Applications. White Paper* (Tech. Rep.).

Scott, D. (2007). Resolving the quantitative- qualititive dilemma: A critical realist approach. *Journal of Research and Method in Education*, *30*(1), 3–17. doi: 10.1080/17437270701207694

Sen, M., & Duffy, T. (2005). GeoSciML: Development of a generic GeoScience Markup Language. *Comput. Geosci.*, *31*(9), 1095–1103. doi: 10.1016/j.cageo.2004.12.003

Sheth, A. (1999). Interoperating Geographic Information Systems. *Interoperating Geographic Information Systems*, 5–30. Retrieved from http://www.mendeley.com/catalog/changing-focus-interoperability-information-systems-system-syntax-structure-semantics/ doi: 10.1007/978-1-4615-5189-8

Sikolia, D., Biros, D., Mason, M., Weiser, M., Sikolia, D., & Weiser, M. (2013). Trustworthiness of grounded theory methodology research in information systems. In *Proceedings of the eighth midwest association for information systems conference* (Vol. 24, pp. 1–5). Illinois: Association for information systems. Retrieved from http://aisel.aisnet.org/mwais2013/16

SMART. (2016). *SMART Aquifer Characterisation (SAC) project homepage*. http://www.gns.cri.nz/Home/Our-Science/Environment-and-Materials/Groundwater/Research-Programmes/SMART-Aquifer-Characterisation/About: SMART consortium.

Smith, T. M., & Lakshmanan, V. (2011). Real-time, rapidly updating severe weather products for virtual globes. *Computers & Geosciences*, *37*(1), 3–12.

doi: 10.1016/j.cageo.2010.03.023

Stein, C., Limper, M., & Kuijper, A. (2014, aug). Spatial data structures for accelerated 3D visibility computation to enable large model visualization on the web. In *Proceedings of the nineteenth international acm conference on 3d web technologies - web3d '14* (pp. 53–61). New York, New York, USA: ACM Press. doi: 10.1145/2628588.2628600

Stock, K., Stojanovic, T., Reitsma, F., Ou, Y., Bishr, M., Ortmann, J., & Robertson, A. (2012). To ontologise or not to ontologise: An information model for a geospatial knowledge infrastructure. *Computers & Geosciences*, *45*(0), 98–108. doi: 10.1016/j.cageo.2011.10.021

Strauss, A. L., & Corbin, J. M. (1994). *Grounded theory methodology.* doi: 10.1007/BF00988593

Strauss, A. L., & Corbin, J. M. (1998). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* (Vol. 2nd).

Takeda, H., Veerkamp, P., Tomiyama, T., & Yoshikawa, H. (1990). Modeling Design Processes. *AI Magazine*, *11*, 37–48. doi: 10.1609/aimag.v11i4.855

The OpenMI Association Technical Committee, Bröring, A., Stasch, C., & Echterhoff, J. (2010). *OpenMI Standard 2 Specification for the OpenMI (Version 2.0)* (Tech. Rep.). The OpenMI Association.

Thomas, B., Steidl, J., Dietrich, O., & Lischeid, G. (2011). Measures to sustain seasonal minimum runoff in small catchments in the mid-latitudes: A review. *Journal Hydrology*, *408*(3-4), 296–307. doi: 10.1016/j.jhydrol.2011.07.045

Thornberg, R. (2012). Informed Grounded Theory. *Scandinavian Journal of Educational Research*, *56*(3), 243–259. doi: 10.1080/00313831.2011.581686

Tian, Y., & Huang, M. (2012). Enhance discovery and retrieval of geospatial data using SOA and Semantic Web technologies. *Expert Systems with Applications*, *39*(16), 12522–12535. doi: 10.1016/j.eswa.2012.04.061

Tobler, W. R. (1970). A computer movie simulating urban growth in the detroit region. *Economic Geography*, *46*, 234-240. Retrieved from http://www.jstor.org/stable/143141

Tress, B., Tress, G., van Apeldoorn, R., & Fry, G. (2003). *Interdisciplinarity and Transdisciplinary Landscape Studies: Potentials and Limitations*. Wageningen.

Urquhart, C., Lehmann, H., & Myers, M. D. (2010). Putting the 'theory' back into grounded theory: Guidelines for grounded theory studies in information systems. *Information Systems Journal*, *20*(4), 357–381. doi: 10.1111/j.1365-2575.2009.00328.x

Usländer, T., Jacques, P., Simonis, I., & Watson, K. (2010). Designing environmental software applications based upon an open sensor service architecture. *Environmental Modelling & Software*, *25*(9), 977–987. doi: 10.1016/j.envsoft.2010.03.013

Vaishnavi, V., & Kuechler, B. (2004). Design Science Research in Information Systems Overview of Design Science Research. *Ais*, 45. doi: 10.1007/978-1-4419-5653-8

van den Brink, L., Portele, C., & Vretanos, P. A. (2012). *Geography Markup Language (GML) simple features profile (Corrigendum)* (Tech. Rep.). Open Geospatial Consortium. Retrieved from http://portal.opengeospatial.org/files/?artifact{_}id=42729

Vitolo, C., Elkhatib, Y., Reusser, D., Macleod, C. J., & Buytaert, W. (2015, jan). Web technologies for environmental Big Data. *Environmental Modelling & Software*, *63*, 185–198. doi: 10.1016/j.envsoft.2014.10.007

Voinov, A., & Bousquet, F. (2010). Modelling with stakeholders. *Journal of Environmental Modelling & Software*, *25*(11), 1268–1281. doi: 10.1016/j.envsoft.2010.03.007

Wallis, I., Moore, C., Post, V., Wolf, L., Martens, E., & Prommer, H. (2014). Using predictive uncertainty analysis to optimise tracer test design and data acquisition. *Journal of Hydrology*, *515*, 191–204. doi: 10.1016/j.jhydrol.2014.04.061

Walsh, I. (2014). *Using Grounded Theory to Avoid Research Misconduct in Management Science.* Retrieved 2016-01-05, from http://groundedtheoryreview.com/2014/06/22/using-grounded-theory-to-avoid-research-misconduct-in-management-science/

WebGL. (2012). *WebGL - OpenGL ES 2.0 for the Web.* Retrieved from http://www.khronos.org/webgl/

West, G. A. W. (2015). The Semantic Web and its Relevance to Advanced Globe Processing. In *9th symposium of the international society for digital earth isde.* Halifax.

White, P. A. (2001). Groundwater resources in New Zealand. In M. R. Rosen & P. A. White (Eds.), *Groundwaters of new zealand* (pp. 47–75). Wellington: New Zealand Hydrological Society.

White, P. A. (2006). Some Future Directions in Hydrology. *Journal of Hydrology (NZ)*, *45*((2)), 63–68.

White, P. A. (2007). *New Zealand groundwater volumes - Update for June 2005. Unpublished letter report to Statistics New Zealand.* (Tech. Rep.). GNS Science.

White, P. A., Raiber, M., Della Pasqua, F., Zarour, H., & Meilhac, C. (2010). *Three-dimensional geological model of the Horowhenua area. GNS Science Consultancy Report 2009/310* (Tech. Rep.). Taupo, New Zealand: GNS Science.

White, P. A., Raiber, M., Zarour, H., Meilhac, C., & Green, S. (2010). *Horowhenua water resources: water budget and groundwater-surface water interaction, GNS Science Consultancy Report 2010/22* (Tech. Rep.). GNS Science.

White, P. A., & Reeves, R. R. (2002). *The volume of groundwater in New Zealand 1994 to 2001. GNS Science Client Report 2002/79* (Tech. Rep.). Taupo, New Zealand: GNS Science.

White, P. A., Sharp, B. M. H., & Kerr, G. N. (2001). Economic valuation of the Waimea Plains groundwater system. *Journal of Hydrology (NZ)*, *40*(1), 59–76.

William, G., Milke, M. W., & Raffensperger, J. F. (2009). Survey of New Zealand hydrologists on information needs. *Journal of Hydrology (NZ)*, *48*(1), 1–12.

Wojda, P., & Brouyère, S. (2013). An object-oriented hydrogeological data model for groundwater projects. *Journal of Environmental Modelling & Software*, *43*(0), 109–123. doi: 10.1016/j.envsoft.2013.01.015

Wriedt, G., Van der Velde, M., Aloe, A., & Bouraoui, F. (2009). Estimating irrigation water requirements in Europe. *Journal Hydrology*, *373*(3-4), 527–544. doi: 10.1016/j.jhydrol.2009.05.018

Yang, C., Raskin, R., Goodchild, M., & Gahegan, M. (2010). Geospatial Cyberinfrastructure: Past, present and future. *Computers, Environment and Urban Systems*, *34*(4), 264–277. doi: 10.1016/j.compenvurbsys.2010.04.001

Yick, J., Mukherjee, B., & Ghosal, D. (2008). Wireless sensor network survey. *Computer Networks*, *52*(12), 2292–2330. doi: 10.1016/j.comnet.2008.04.002

Yin, R. K. (1994). Case Study Research: Design and Methods. *Applied Social Research Methods Series*, *5*, 219. doi: 10.1097/FCH.0b013e31822dda9e

Yu, G. E., Zhao, P., Di, L., Chen, A., Deng, M., & Bai, Y. (2012, oct). BPELPowerâ[[[ERROR FOR PACKAGE inputenc]]][[[ERROR FOR PACKAGE inputenc]]]A BPEL execution engine for geospatial web services. *Computers & Geosciences*, *47*, 87–101. doi: 10.1016/j.cageo.2011.11.029

Yue, P., Wei, Y., Di, L., He, L., Gong, J., & Zhang, L. (2011, jul). Sharing geospatial provenance in a service-oriented environment. *Computers, Environment and Urban Systems*, *35*(4), 333–343. doi: 10.1016/j.compenvurbsys.2011.02.006

Zachariadis, M., Scott, S., & Barrett, M. (2010). Exploring critical realism as the theoretical foundation of mixed-method research: Evidence from the economics of IS innovations. *Working Paper Series - Cambridge Business School*.

Zachman, J. A. (1987). A framework for information systems architecture. *IBM Systems Journal*, *26*(3), 276–292. doi: 10.1147/sj.263.0276

Zachman, J. A. (2015). *A Historical Look at Enterprise Architecture with John Zachman.* Retrieved 2017-05-02, from https://blog.opengroup.org/2015/01/23/a-historical-look-at-enterprise-architecture-with-john-zachman/

Zhang, Y., Meratnia, N., & Havinga, P. J. M. (2010). Ensuring high sensor data quality through use of online outlier detection techniques. *International Journal of Sensor Networks*, *7*(3), 141–151. doi: 10.1504/ijsnet.2010.033116

Zhao, P., Di, L., Yu, G., Yue, P., Wei, Y., & Yang, W. (2009). Semantic Web-based geospatial knowledge transformation. *Computers & Geosciences*, *35*(4), 798–808. doi: 10.1016/j.cageo.2008.03.013

Zhao, P., Foerster, T., & Yue, P. (2012). The Geoprocessing Web. *Computers & Geosciences*, *47*(0), 3–12. doi: 10.1016/j.cageo.2012.04.021

ACRONYMS

API Application Programming Interface

ANZLIC Australia New Zealand Land Information Council

ArcGRID A non-proprietary ASCII raster format, also known as an ARC/INFO ASCII GRID

BOM Australian Bureau of Meteorology

BPEL Business Process Execution Language

BPMN Business Process Markup Notation

CC Creative Commons

CityGML City Geographic Markup Language

CLIDB Climate and Flow Database

CRI Crown Research Institute

CSW Catalogue Service for the Web

CUAHSI Consortium of Universities for the Advancement of Hydrologic Science, Inc.

DEM Digital Elevation Model

DNS Domain Name System

DoC Department of Conservation

DOM Document Object Model

DTM Digital Terrain Model

DSR Design Science Research

EA Enterprise Architecture

EODP Environmental Observation Data Profile

ESB Enterprise Service Bus

FE Filter Encoding

GeoJSON Geographic JSON

GeoSciML Geoscientific Markup Language

GEOSPARQL Geographic SPARQL

GGW Geothermal and Groundwater Database

GIScience Geographical Information Science

GIS Geographic Information System

GML Geographic Markup Language

GNS Institute of Geological and Nuclear Sciences

GT Grounded Theory

GW2IE Groundwater Interoperability Experiment 2

GWML also GroundWaterML, Groundwater Markup Language

HPC High Performance Computing

HTML Hyper Text Markup Language

HTTP Hyper Text Transfer Protocol

INSPIRE Infrastructure for Spatial Information in Europe

ICT Information and Communications Technologies

ISO International Organization for Standardization

JSON JavaScript Object Notation

JSON-LD JSON Linked Data

LADA Land Data Catalogue

LAWA Land, Air, Water, Aotearoa

LINZ Land Information New Zealand

LRIS Land Resource Information Systems

MfE Ministry for the Environment

MPI Message Passing Interface

MQTT Message Queue Telemetry Transport

NEMS National Environmental Monitoring Standards

NetCDF Network Common Data Form

NGMP National Groundwater Monitoring Programme

NIWA National Institue for Water and Atmospheric Research

NZGO New Zealand Geospatial Office

NZGOAL New Zealand Government Open Access and Licensing

NZMG New Zealand Map Grid

NZTM2000 New Zealand Transverse Mercator 2000

OASIS Organization for the Advancement of Structured Information Standards

OGC Open Geospatial Consortium

O&M Observations & Measurements

OMS Object Modelling System

OpenMI Open Modelling Interface

OWC OpenGIS®Web Services Context

OWL Web Ontology Language

OWS OpenGIS®Web Services

PDF Portable Document Format

RDF Resource Description Framework

RDFS RDF Schema

REST Representational State Transfer

RMA Resource Management Act

SAC SMART Aquifer Characterisation programme

SDI Spatial Data Infrastructure

SensorML Sensor Model Language

SE Symbol Encoding

SF Simple Features

SKOS Simple Knowledge Organization System

SLD Styled Layer Descriptor

SMART simple, measurable, achievable, repeatable, and time-based

SOAP Simple Object Access Protocol

SOA Service Oriented Architecture

SOE State of the Environment

SOS Sensor Observation Service

SPARQL SPARQL Protocol And RDF Query Language

SPS Sensor Planning Service

SQL Structured Query Language

SWE Sensor Web Enablement

UML Unified Modeling Language

URI Uniform Resource Identifier

URL Uniform Resource Locator

USGS United States Geological Survey

VRML Virtual Reality Modeling Language

W3C world Wide Web Consortium

WaterML Water Markup Language

WCPS Web Coverage Processing Service

WCS Web Coverage Service

WDTF Water Data Transfer Format

WFS Web Feature Service

WGS84 World Geodetic System 1984

WMC Web Map Context

WMO World Meteorological Organisation

WMS Web Map Service

WMTS Web Map Tile Service

WPS Web Processing Service

WSDL Web Service Description Language

WSN Wireless Sensor Network

X3D Extensible 3D - A 3D Scene Graph Markup Language

XML Extensible Markup Language

XSD XML Schema Definition

Part V

APPENDIX

A SWE FOR OBSERVATION DATABASES LISTINGS

A.1 SOS INTERFACES FOR CLIDB DATABASE

```java
package org.n52.sos.ds.clidb;
   public CLIDBHelper() throws OwsExceptionReport {
      ConnectorProperties properties = null;
      // should those maps be actually synchroniszed?
      kStationVMeasures = new HashMap<String, List<String>>();
       kMeasureVStations = new HashMap<String, List<String>>();
       allCrossCheckedProcedures = new HashMap<String, Procedure>();
       allCrossCheckedFeatures = new HashMap<String, FeatureOfInterest>();
       generatedOfferings = new HashMap<String, Offering>();
       generatedObservableProperties = new HashMap<String, ObservableProperty>();
       generatedObservationConstallations = new ArrayList<ObservationConstellation>();
       statementsCounter = 0;

      try {
         properties = ConnectorProperties.getInstance();
         this.connector = new CLIDBConnectorImpl(properties.getProperty("server"), new Integer(properties.getProperty("port")),
                 properties.getProperty("sid"), properties.getProperty("username"), properties.getProperty("password"));
         setCLIDBConnector(connector);

   public static CLIDBHelper getInstance() throws OwsExceptionReport {
      if(instance == null) {
            instance = new CLIDBHelper();
       }
      return instance;
   }
```

Listing 5: Excerpts from CLIDB-SOS `CLIDBHelper` class

```java
package org.n52.sos.ds.clidb;
/**
 * Implementation of the interface IDescribeSensorDAO
 *
 */
public class DescribeSensorDAO extends AbstractDescribeSensorDAO {
    private final CLIDBSessionHolder sessionHolder = new CLIDBSessionHolder();
    private final HibernateProcedureConverter procedureConverter = new HibernateProcedureConverter();
                final Procedure procedure = CLIDBHelper.getInstance().getProcedureForIdentifier(childProcID, session);
                SosProcedureDescription childProcedure = procedureConverter.createSosProcedureDescription(procedure, childProcID,
      outputFormat);
```

Listing 6: Excerpts from CLIDB-SOS `DescribeSensorDAO` class

```java
package org.n52.sos.ds.clidb;
public class GetFeatureOfInterestDAO extends AbstractGetFeatureOfInterestDAO {

    private CLIDBSessionHolder sessionHolder = new CLIDBSessionHolder();
    private static final Logger LOGGER = LoggerFactory.getLogger(GetFeatureOfInterestDAO.class);
                    SosFeatureCollection featureCollection =
                            new SosFeatureCollection(getConfigurator().getFeatureQueryHandler().getFeatures(
                                    new ArrayList<String>(foiIDs), request.getSpatialFilters(), session,
                                    request.getVersion(), -1));
```

Listing 7: Excerpts from CLIDB-SOS `GetFeatureOfInterestDAO` class

```java
package org.n52.sos.ds.clidb;

         CLIDBDataSeries cds = session.getData(stationId, meas, start, end);

         // vocaburi forward mapping!
         String vocabUriMeas = CLIDBHelper.mapConceptUriForward(meas.code);

         List<TemporalCLIDBNode> result = cds.data;
         CLIDBDataSeriesHeader header = cds.header;
         String unit = header.measure.unit.typeName;
         for (TemporalCLIDBNode node : result) {
```

Listing 8: Excerpts from CLIDB-SOS `GetObservationDAO` class

```java
package org.n52.sos.ds.clidb;
        try {
         session = getSession();
            LOGGER.debug("Cache Update, trying to get CLIDBConnector instance: " + session.toString());
         CLIDBHelper.getInstance().generateCLIDBCache();
```

```
            InitialCacheUpdate update = new InitialCacheUpdate(getCacheThreadCount());

            update.setCache(cache);
            update.setErrors(errors);
            update.setSession(session);
            update.execute();
        } catch (Exception he) {
            LOGGER.error("Error while updating ContentCache!", he);
        } finally {
            returnSession(session);
        }
```

Listing 9: Excerpts from CLIDB-SOS `SosCacheFeederDAO` class

A.2 SOS INTERFACES FOR NGMP DATABASE

```
package org.n52.sos.ds.ggw;
public class GGWHelper {

   /**
     * logger
     */
    private static final Logger LOGGER = LoggerFactory.getLogger(GGWHelper.class);

    /** instance attribut, due to the singleton pattern */
    private static GGWHelper instance = null;

    /** cheap-o-cache maps, due to the singleton pattern */
    private Map<String, Set<String>> kSiteVParam;
    private Map<String, Set<String>> kParamVSite;
    private Map<String, FeatureOfInterest> allCrossCheckedFeatures;
    private Map<String, Procedure> allCrossCheckedProcedures;
    private Map<String, Offering> generatedOfferings;
    private Map<String, ObservableProperty> generatedObservableProperties;
    private Set<ObservationConstellation> generatedObservationConstallations;
    private Set<String> allObservationIds;

    public GGWHelper() throws OwsExceptionReport {

      kSiteVParam = new HashMap<String, Set<String>>();
       kParamVSite = new HashMap<String, Set<String>>();
       allCrossCheckedProcedures = new HashMap<String, Procedure>();
       allCrossCheckedFeatures = new HashMap<String, FeatureOfInterest>();
       generatedOfferings = new HashMap<String, Offering>();
       generatedObservableProperties = new HashMap<String, ObservableProperty>();
       generatedObservationConstallations = new HashSet<ObservationConstellation>();
       allObservationIds = new HashSet<String>();
    }


   public static GGWHelper getInstance() throws OwsExceptionReport {
      if(instance == null) {
            instance = new GGWHelper();
       }
      return instance;
   }

   public void generateGGWCache(Connection session) throws OwsExceptionReport {

      LOGGER.debug("-- generateGGWCache() -- ");
      long millis = System.currentTimeMillis();

      HashMap<String, Set<String>> t_kSiteVParam = new HashMap<String, Set<String>>();
      HashMap<String, Set<String>> t_kParamVSite = new HashMap<String, Set<String>>();
      HashMap<String, Procedure> t_allCrossCheckedProcedures = new HashMap<String, Procedure>();
      HashMap<String, FeatureOfInterest> t_allCrossCheckedFeatures = new HashMap<String, FeatureOfInterest>();
      HashMap<String, Offering> t_generatedOfferings = new HashMap<String, Offering>();
      HashMap<String, ObservableProperty> t_generatedObservableProperties = new HashMap<String, ObservableProperty>();
      HashSet<ObservationConstellation> t_generatedObservationConstallations = new HashSet<ObservationConstellation>();
      HashSet<String> t_allObservationIds = new HashSet<String>();

      // query all features
      // sql all features with event and project and not confidential
      String query = "SELECT DISTINCT FT.FEATURE, FT.FEATURE_PUBLIC_ID, FT.FEAT_DISP_NAME, "
            + " FT.FEAT_TOTAL_LENGTH, FT.FEAT_DESCR, FT.FEAT_LOCATION_DESCR, FT.FEAT_CALC_NZGD1949_LATITUDE, "
            + " FT.FEAT_CALC_NZGD1949_LONGITUDE, FT.FEAT_CALC_NZGD1949_LATLONG_SRC, FT.FEAT_TYPE_LEAF,  PF.PROJECT"
            + " FROM GGW.FEATURE FT"
            + " RIGHT OUTER JOIN GGW.PROJ_FEAT PF"
            + " ON (FT.FEATURE = PF.FEATURE)"
            + " WHERE (PF.PROJ_FEAT_IS_VALID = 1) and (PF.PROJ_FEAT_IS_CNFDTL = 0) "
            + " AND "
            + addWhereEProjectOrFilter(allProjects,
                  "PF.");

      String queryObsProp = "SELECT R.SAMPLE_ANALYSIS_RESULT_PUB_ID, R.PARAM_TYPE, R.PARAM_TYPE_DISP_NAME, "
            + " E.FEATURE_PUBLIC_ID, E.EVENT_START_TSTZ_AS_DATE, E.EVENT_START_NZST "
            + " FROM GGW.V_EXP_SAMPLE_ANALYSIS_RESULT R "
            + " RIGHT OUTER JOIN GGW.V_EXP_SAMPLE_ANALYSIS A "
            + " ON (R.SAMPLE_ANALYSIS = A.SAMPLE_ANALYSIS)"
            + " RIGHT OUTER JOIN GGW.V_EXP_SAMPLE S"
            + " ON (A.SAMPLE = S.SAMPLE) "
            + " RIGHT OUTER JOIN GGW.V_EXP_EVENT E "
```

```
            + " ON (S.EVENT = E.EVENT)  "
            + "WHERE "
            + addWhereRParamOrFilter(allParams,
                "R.")
            + " AND "
            + addWhereEProjectOrFilter(allProjects,
                "E.");
```

Listing 10: Excerpts from NGMP-SOS `GGWHelper` class

```
package org.n52.sos.ds.ggw;

/**
 * Implementation of the interface IDescribeSensorDAO
 *
 */
public class DescribeSensorDAO extends AbstractDescribeSensorDAO {
    private final GGWSessionHolder sessionHolder = new GGWSessionHolder();
                final Procedure procedure = GGWHelper.getInstance().getProcedureForIdentifier(childProcID, session);
                SosProcedureDescription childProcedure = procedureConverter.createSosProcedureDescription(procedure, childProcID,
      outputFormat);
```

Listing 11: Excerpts from NGMP-SOS `DescribeSensorDAO` class

```
package org.n52.sos.ds.ggw;

public class GetFeatureOfInterestDAO extends AbstractGetFeatureOfInterestDAO {

    private GGWSessionHolder sessionHolder = new GGWSessionHolder();
                    SosFeatureCollection featureCollection =
                            new SosFeatureCollection(getConfigurator().getFeatureQueryHandler().getFeatures(
                                    new ArrayList<String>(foiIDs), request.getSpatialFilters(), session,
                                    request.getVersion(), -1));
```

Listing 12: Excerpts from NGMP-SOS `GetFeatureOfInterestDAO` class

```
package org.n52.sos.ds.ggw;
      // by querying all samples
      StringBuilder queryFullGetObservationData = new StringBuilder();
      queryFullGetObservationData.append("SELECT R.SAMPLE_ANALYSIS_RESULT_PUB_ID, R.PARAM_TYPE, R.RESULT_AS_NMBR, R.
       PARAM_TYPE_DISP_NAME, "
            + " R.PARAM_TYPE_UNIT_ABRV, E.FEATURE, E.FEATURE_PUBLIC_ID, E.EVENT_START_NZST_AS_DATE, E.EVENT_START_TSTZ_AS_DATE, E.
       EVENT_START_NZST, E.PROJECT "
            + " FROM GGW.V_EXP_SAMPLE_ANALYSIS_RESULT R "
            + " RIGHT OUTER JOIN GGW.V_EXP_SAMPLE_ANALYSIS A "
            + " ON (R.SAMPLE_ANALYSIS = A.SAMPLE_ANALYSIS)"
            + " RIGHT OUTER JOIN GGW.V_EXP_SAMPLE S"
            + " ON (A.SAMPLE = S.SAMPLE) "
            + " RIGHT OUTER JOIN GGW.V_EXP_EVENT E "
            + " ON (S.EVENT = E.EVENT)  "
            + "WHERE " + GGWHelper.addWhereEProjectOrFilter(allProjects, "E.") );
         // execute query
         prepStmtGetObsData = session.prepareStatement(queryFullGetObservationData.toString());
```

Listing 13: Excerpts from NGMP-SOS `GetObservationDAO` class

```
package org.n52.sos.ds.ggw;
        try {

         GGWHelper.getInstance().generateGGWCache(conn);

            InitialCacheUpdate update = new InitialCacheUpdate(getCacheThreadCount());
            // session = getSession();
            update.setCache(cache);
            update.setErrors(errors);
            update.setSession(conn);
            update.execute();
        } catch (Exception he) {
            LOGGER.error("Error while updating ContentCache!", he);
        } finally {
         LOGGER.debug("SosCacheFeeder updateCache returnSession()");
            returnSession(conn);
        }
```

Listing 14: Excerpts from NGMP-SOS `SosCacheFeederDAO` class

B

SDI INSTANCE DOCUMENTS

B.1 GROUNDWATERML2 INSTANCE DOCUMENTS


```xml
<?xml version="1.0" encoding="UTF-8"?>
<gwml2:GW_ManagementArea xmlns:gwml2="http://www.opengis.net/gwml-main/2.2"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="http://www.opengis.net/gwml-main/2.2 http://ngwd-bdnes.cits.nrcan.gc.ca/service/gwml/schemas/2.2/gwml2-well.xsd"
    xmlns:gmd="http://www.isotc211.org/2005/gmd" xmlns:gwml2f="http://www.opengis.net/gwml-flow/2.2"
    xmlns:gco="http://www.isotc211.org/2005/gco"
    xmlns:swe="http://www.opengis.net/swe/2.0" xmlns:sam="http://www.opengis.net/sampling/2.0"
    xmlns:gwml2w="http://www.opengis.net/gwml-well/2.2" xmlns:om="http://www.opengis.net/om/2.0"
    xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:gml="http://www.opengis.net/gml/3.2"
    xmlns:sams="http://www.opengis.net/samplingSpatial/2.0" gml:id="horo.cm.4">
    <gml:description>
        The Ohau catchment in the South of Horowhenua district, area in hectares 54.1335337952
    </gml:description>
    <gml:identifier codeSpace="http://www.gns.cri.nz">catchments/Ohau</gml:identifier>
    <gml:name codeSpace="http://www.gns.cri.nz">Ohau</gml:name>
    <gml:boundedBy>
        <gml:Envelope srsDimension="2" srsName="urn:ogc:def:crs:EPSG::4326">
            <gml:lowerCorner>-40.70908827419566 175.1512623098973</gml:lowerCorner>
            <gml:upperCorner>-40.64650261982838 175.34035218665645</gml:upperCorner>
        </gml:Envelope>
    </gml:boundedBy>
    <gwml2:gwAreaShape>
        <gml:MultiSurface gml:id="cm.shape.4" srsDimension="2" srsName="urn:ogc:def:crs:EPSG::4326">
            <gml:surfaceMember>
                <gml:Polygon gml:id="cm.shape.4.1" srsDimension="2">
                    <gml:exterior>
                        <gml:LinearRing>
                            <gml:posList> -40.65216490093651 175.25293586637216 -40.65228163932043
                                175.25293979876676 -40.65216699213812 175.25282884136962
                                -40.65216490093651 175.25293586637216 </gml:posList>
                        </gml:LinearRing>
                    </gml:exterior>
                </gml:Polygon>
            </gml:surfaceMember>
            <gml:surfaceMember>
                <gml:Polygon gml:id="cm.shape.4.2" srsDimension="2">
                    <gml:exterior>
                        <gml:LinearRing>
                            <gml:posList> -40.64860741008362 175.1992372188969 -40.64887750673649
                                175.1992461020351 -40.64887074073361 175.19960069226167
                                -40.64914081649706 175.19960957621328 -40.649134049408815
                                175.19996416625122 -40.64940412626912 175.19997305173933
                                175.19922833648945 -40.64860741008362 175.1992372188969
                            </gml:posList>
                        </gml:LinearRing>
                    </gml:exterior>
                    <gml:interior>
                        <gml:LinearRing>
                            <gml:posList> -40.67801752302447 175.25845963640833 -40.67965313431198
                                175.2561532664341 -40.680624158457846 175.2560043748437
                                -40.681522525822324 175.2560346959848 -40.682925932149814
                                175.25499202871387 -40.68377116507043 175.25420300798461
                                -40.6845348702513 175.25404708838624 -40.68483069251186
                                175.25660055416145 -40.68656281083376 175.25895483109502
                                -40.68695137647072 175.26130508911976 -40.68609185128348
                                175.26282030053326 -40.68708857058016 175.26405874509175
                                -40.687484241249884 175.26586524818276 -40.68766007215999
                                175.26620346447106 -40.688444741487515 175.26707328728165
                                -40.688346463228505 175.2674900821607 -40.68751283804191
                                175.26842335065285 -40.6869034446236 175.26871923731775
                                -40.68635103182839 175.26886616628096 -40.68524630990728
                                175.26846772787516 -40.6845989729919 175.2681735775108
                                -40.68433580549244 175.26827115241738 -40.684053842674
                                175.26863730896608 -40.68349584639927 175.26906805344194
                                -40.682721146968696 175.26952397583523 -40.682374548078215
                                175.26951220801422 -40.682051247588674 175.26923201906592
                                -40.681748811525736 175.26834903494404 -40.68150586677524
                                175.26787632948324 -40.68119133226316 175.267608282823
                                -40.68032424245744 175.26726528913608 -40.679724033007844
                                175.26709405988018 -40.678724168555114 175.26703102348154
                                -40.6780079575942 175.26599818994575 -40.678525563821076
                                175.26428995430905 -40.67862984671428 175.26340650242213
                                -40.67802259250908 175.26172968139338 -40.67801752302447
                                175.25845963640833 </gml:posList>
                        </gml:LinearRing>
                    </gml:interior>
                </gml:Polygon>
            </gml:surfaceMember>
        </gml:MultiSurface>
    </gwml2:gwAreaShape>
    <gwml2:gwAreaType>Groundwater Management Zone</gwml2:gwAreaType>
```

```xml
<!-- a water budget from a report for the horowhenua district, the numbers might not be representative
the groundwater system is not a well known aquifer system -->
<gwml2:gwAreaWaterBudget>
    <gwml2f:GW_WaterBudget>
        <gwml2f:gwBudgetAmount>
            <om:OM_Observation gml:id="om_budget_1432124_1">
                <gml:description>estimated water budget in Horowhenua area</gml:description>
                <gml:name codeSpace="http://gns.cri.nz/client_reports">2010-22</gml:name>
                <om:phenomenonTime>
                    <gml:TimePeriod gml:id="om_budget_1432124_ti_1">
                        <gml:beginPosition>2009-01-01T00:00:00</gml:beginPosition>
                        <gml:endPosition>2010-01-01T00:00:00</gml:endPosition>
                    </gml:TimePeriod>
                </om:phenomenonTime>
                <om:resultTime>
                    <gml:TimeInstant gml:id="om_budget_1432124_ti_2">
                        <gml:timePosition>2010-06-30T12:00:00</gml:timePosition>
                    </gml:TimeInstant>
                </om:resultTime>
                <om:procedure xlink:href="http://gns.cri.nz/consultancy"
                    xlink:title="Commercial Research"/>
                <om:observedProperty xlink:href="http://some.vocab/water-budget"
                    xlink:title="Water Budget"/>
                <om:featureOfInterest xlink:href="#om_budget_1432124_1"/>
                <om:result>
                    <om:result xmlns:ns="http://www.opengis.net/swe/2.0"
                        xsi:type="ns:QuantityType">
                        <ns:uom xlink:href="http://www.opengis.net/def/uom/UCUM/0/m3"
                            xlink:title="cubic meters"/>
                        <ns:value>0.0</ns:value>
                    </om:result>
                </om:result>
            </om:OM_Observation>
        </gwml2f:gwBudgetAmount>
        <gwml2f:gwBudgetDischarge xsi:nil="true" nilReason="withheld"/>
        <gwml2f:gwBudgetRecharge>
            <gwml2f:GW_Recharge gml:id="qin_all_1">
                <gml:description>
                    in-flow into study area, combined from prec, gw, surface
                </gml:description>
                <gwml2f:gwFlowPersistence xlink:href="http://somevocab.org/gwml/2.0/notes"
                    xlink:title="notspecified"/>
                <gwml2f:gwFlowProcess xlink:href="http://somevocab.org/gwml/2.0/notes"
                    xlink:title="infiltration"/>
                <gwml2f:gwFlowTime>
                    <gml:TimePeriod gml:id="flowt_qin_all_1">
                        <gml:beginPosition>2009-01-01T00:00:00</gml:beginPosition>
                        <gml:endPosition>2010-01-01T00:00:00</gml:endPosition>
                    </gml:TimePeriod>
                </gwml2f:gwFlowTime>
                <gwml2f:gwFlowVelocity xsi:nil="true"/>
                <gwml2f:gwFlowVolumeRate>
                    <om:OM_Observation gml:id="om_flowt_qin_all_1">
                        <gml:description>estimated water inflow over all</gml:description>
                        <gml:name codeSpace="http://gns.cri.nz/">combined-inflow</gml:name>
                        <om:phenomenonTime>
                            <gml:TimePeriod gml:id="om_flowt_qin_all_1_ti_1">
                                <gml:beginPosition>2009-01-01T00:00:00</gml:beginPosition>
                                <gml:endPosition>2010-01-01T00:00:00</gml:endPosition>
                            </gml:TimePeriod>
                        </om:phenomenonTime>
                        <om:resultTime>
                            <gml:TimeInstant gml:id="om_flowt_qin_all_1_ti_2">
                                <gml:timePosition>2010-06-30T12:00:00</gml:timePosition>
                            </gml:TimeInstant>
                        </om:resultTime>
                        <om:procedure xlink:href="http://gns.cri.nz/consultancy"
                            xlink:title="Commercial Research"/>
                        <om:observedProperty xlink:href="http://some.vocab/water-budget"
                            xlink:title="Water Budget"/>
                        <om:featureOfInterest xlink:href="http://the.gwhdrogeounit.x"
                            xlink:title="Hydrogeological Unit"/>
                        <om:result>
                            <om:result xmlns:ns="http://www.opengis.net/swe/2.0"
                                xsi:type="ns:QuantityType">
                                <ns:uom
                                    xlink:href="http://www.opengis.net/def/uom/UCUM/0/million_cubic_metres_per_year"
                                    xlink:title="Million Cubik Meters per Year"/>
                                <ns:value>326.7</ns:value>
                            </om:result>
                        </om:result>
                    </om:OM_Observation>
                </gwml2f:gwFlowVolumeRate>
                <gwml2f:gwFlowDestinationBody xlink:href="http://the.gwwaterbody.x"
                    xlink:title="Water body in aquifersystem"/>
                <gwml2f:gwFlowDestinationContainer xlink:href="http://the.gwhdrogeounit.x"
                    xlink:title="Hydrogeological Unit"/>
                <gwml2f:gwFlowLocation xlink:href="#cm.shape.4"/>
                <gwml2f:gwFlowSourceBody xlink:href="http://the.environment"
                    xlink:title="natural environment"/>
                <gwml2f:gwFlowSourceContainer xlink:href="http://the.environment"
                    xlink:title="natural environment"/>
            </gwml2f:GW_Recharge>
        </gwml2f:gwBudgetRecharge>
        <gwml2f:gwBudgetValidTime>
            <gml:TimePeriod gml:id="wbom_tp_1">
                <gml:beginPosition>2009-01-01T00:00:00</gml:beginPosition>
                <gml:endPosition>2010-01-01T00:00:00</gml:endPosition>
            </gml:TimePeriod>
```

```xml
            </gwml2f:gwBudgetValidTime>
        </gwml2f:GW_WaterBudget>
    </gwml2:gwAreaWaterBudget>
    <gwml2:relatedManagementArea xlink:href="http://link.to/containing/district/horowhenua"
        xlink:title="Horowhenua District"/>
    <gwml2:gwManagedUnit xlink:href="http://the.gwhdrogeounit.x" xlink:title="Hydrogeological Unit"/>
</gwml2:GW_ManagementArea>
```

Listing 15: Instance encoding example for catchments/management zones as `gwml2:GW_ManagementArea`

```xml
<?xml version="1.0" encoding="UTF-8"?>
<gwml2w:GW_MonitoringSite xmlns:gwml2w="http://www.opengis.net/gwml-well/2.2"
    gml:id="soe.mon.352271" xmlns:swe="http://www.opengis.net/swe/2.0"
    xmlns:sam="http://www.opengis.net/sampling/2.0"
    xmlns:sams="http://www.opengis.net/samplingSpatial/2.0"
    xmlns:om="http://www.opengis.net/om/2.0" xmlns:xlink="http://www.w3.org/1999/xlink"
    xmlns:gml="http://www.opengis.net/gml/3.2" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="http://www.opengis.net/gwml-well/2.2 http://ngwd-bdnes.cits.nrcan.gc.ca/service/gwml/schemas/2.2/gwml2-well.xsd">
    <gml:description>state of the environment monitoring well with a depth of 93m</gml:description>
    <gml:identifier codeSpace="http://www.gns.cri.nz">soe.352271</gml:identifier>
    <gml:name codeSpace="http://www.gns.cri.nz">352271</gml:name>
    <gml:location>
        <gml:Point gml:id="loc.mon.352271" srsDimension="2">
            <gml:pos>175.25156953952836 -40.575382977771476</gml:pos>
        </gml:Point>
    </gml:location>
    <sam:sampledFeature xsi:nil="true" nilReason="unknown"/>
    <!-- the vertical line of the monitoring well -->
    <sams:shape>
        <gml:Curve gml:id="shape.mon.352271" srsDimension="3"
            srsName="http://www.opengis.net/def/crs/EPSG/0/4327">
            <gml:segments>
                <gml:LineStringSegment>
                    <gml:posList>175.25156953952836 -40.575382977771476 19.0 175.25156953952836
                        -40.575382977771476 112.2</gml:posList>
                </gml:LineStringSegment>
            </gml:segments>
        </gml:Curve>
    </sams:shape>
    <gwml2w:gwSiteReferenceElevation>
        <gwml2w:Elevation>
            <gwml2w:elevation srsDimension="1" srsName="http://www.opengis.net/def/crs/EPSG/0/4440"
                uomLabels="m NZVD2009">14.295</gwml2w:elevation>
            <gwml2w:elevationType xlink:href="http://somevocab.org/gwml/2.0/notes"
                xlink:title="Relative Level Natural Surface"/>
            <gwml2w:elevationMeasurementMethod xlink:href="http://somevocab.org/gwml/2.0/notes"
                xlink:title="MeasuredInSitu"/>
        </gwml2w:Elevation>
    </gwml2w:gwSiteReferenceElevation>
    <gwml2w:gwSiteType xlink:href="http://somevocab.org/gwml/2.0/notes" xlink:title="Water well"/>
</gwml2w:GW_MonitoringSite>
```

Listing 16: Instance encoding example for NGMP monitoring sites as `gwml2:GW_MonitoringSite` which is also a specialisation of `sams:SF_SpatialSamplingFeature`

```xml
<?xml version="1.0" encoding="UTF-8"?>
<gwml2w:GW_Spring xmlns:gwml2w="http://www.opengis.net/gwml-well/2.2"
    xmlns:sam="http://www.opengis.net/sampling/2.0"
    xmlns:sams="http://www.opengis.net/samplingSpatial/2.0"
    xmlns:gco="http://www.isotc211.org/2005/gco" xmlns:xlink="http://www.w3.org/1999/xlink"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="http://www.opengis.net/gwml-well/2.2
    http://ngwd-bdnes.cits.nrcan.gc.ca/service/gwml/schemas/2.2/gwml2-well.xsd"
    xmlns:gml="http://www.opengis.net/gml/3.2" gml:id="ngmp.pub.7">
    <gwml2w:gwSpringName>Te Waikoropupu Springs ('Pupu Springs')</gwml2w:gwSpringName>
    <gwml2w:gwSpringLocation>
        <gml:Point srsName="http://www.opengis.net/gml/srs/epsg.xml#4326" gml:id="loc_ngmp_7">
            <gml:coordinates cs="," ts=" " decimal=".">172.768,-40.849</gml:coordinates>
        </gml:Point>
    </gwml2w:gwSpringLocation>
    <gwml2w:gwSpringReferenceElevation>
        <gwml2w:Elevation>
            <gwml2w:elevation srsName="http://www.opengis.net/gml/srs/epsg.xml#4440"
                srsDimension="1" uomLabels="m above sea level">3.7</gwml2w:elevation>
            <gwml2w:elevationType xlink:href="http://some.vocab.org/MeasuredInSitu"
                xlink:title="in situ"/>
            <gwml2w:elevationMeasurementMethod xlink:href="http://some.vocab/some-gps-gear"
                xlink:title="D-GPS method"/>
        </gwml2w:Elevation>
    </gwml2w:gwSpringReferenceElevation>
    <gwml2w:gwSpringType xlink:href="http://somevocab.org/gwml/2.0/notes" xlink:title="mineral"/>
    <gwml2w:gwSpringCauseType xlink:href="http://somevocab.org/gwml/2.0/notes"
        xlink:title="artesian"/>
    <gwml2w:gwSpringGeology xlink:href="http://maps.gns.cri.nz/geology/wfs?"
        xlink:title="QMAP NZ Geology"/>
    <gwml2w:gwSpringUnit xsi:nil="true" nilReason="unknown"/>
    <gwml2w:gwSpringConstruction xsi:nil="true"/>
    <gwml2w:gwSpringPersistence xlink:href="http://somevocab.org/gwml/2.0/notes"
        xlink:title="permanent"/>
```

```xml
    <gwml2w:gwSpringLicence xsi:nil="true"/>
</gwml2w:GW_Spring>
```

Listing 17: Instance encoding example for Horowwhenua springs data sets as `gwml2:GW_Spring`

```xml
<?xml version="1.0" encoding="UTF-8"?>
<gwml2w:GW_Well xmlns:gwml2w="http://www.opengis.net/gwml-well/2.2"
    xmlns:gmd="http://www.isotc211.org/2005/gmd"
    xmlns:swe="http://www.opengis.net/swe/2.0" xmlns:sam="http://www.opengis.net/sampling/2.0"
    xmlns:gwml2wc="http://www.opengis.net/gwml-wellconstruction/2.2"
    xmlns:om="http://www.opengis.net/om/2.0" xmlns:xlink="http://www.w3.org/1999/xlink"
    xmlns:gml="http://www.opengis.net/gml/3.2"
    xmlns:sams="http://www.opengis.net/samplingSpatial/2.0"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="http://www.opengis.net/gwml-well/2.2 http://ngwd-bdnes.cits.nrcan.gc.ca/service/gwml/schemas/2.2/gwml2-well.xsd"
    gml:id="soe.well.352271">

    <gml:identifier codeSpace="http://www.gns.cri.nz">
        http://vocab.smart-project.info/ngmp/soe/well/352271
    </gml:identifier>
    <gml:name codeSpace="http://www.gns.cri.nz">352271</gml:name>
    <gml:boundedBy>
        <gml:Envelope srsDimension="2" srsName="urn:ogc:def:crs:EPSG::4326">
            <gml:lowerCorner>-40.575382977771476 175.25156953952836</gml:lowerCorner>
            <gml:upperCorner>-40.575382977771476 175.25156953952836</gml:upperCorner>
        </gml:Envelope>
    </gml:boundedBy>
    <sam:sampledFeature xsi:nil="true" nilReason="missing"/>
    <!-- the vertical line of the monitoring well -->
    <sams:shape>
        <gml:Curve gml:id="shape.mon.352271" srsDimension="3"
            srsName="http://www.opengis.net/def/crs/EPSG/0/4327">
            <gml:segments>
                <gml:LineStringSegment>
                    <gml:posList>175.25156953952836 -40.575382977771476 19.0 175.25156953952836
                        -40.575382977771476 112.2</gml:posList>
                </gml:LineStringSegment>
            </gml:segments>
        </gml:Curve>
    </sams:shape>
    <gwml2w:gwWellLocation>
        <gml:Point gml:id="loc.mon.352271" srsDimension="2" srsName="urn:ogc:def:crs:EPSG::4326">
            <gml:pos>-40.575382977771476 175.25156953952836</gml:pos>
        </gml:Point>
    </gwml2w:gwWellLocation>
    <gwml2w:gwWellPurpose xlink:href="http://somevocab.org/gwml/2.0/notes" xlink:title="Monitoring"/>
    <gwml2w:gwWellReferenceElevation>
        <gwml2w:Elevation>
            <gwml2w:elevation srsDimension="1" srsName="http://www.opengis.net/def/crs/EPSG/0/4440"
                uomLabels="m NZVD2009">14.295</gwml2w:elevation>
            <gwml2w:elevationType xlink:href="http://somevocab.org/gwml/2.0/notes"
                xlink:title="Relative Level Natural Surface"/>
            <gwml2w:elevationMeasurementMethod xlink:href="http://somevocab.org/gwml/2.0/notes"
                xlink:title="MeasuredInSitu"/>
        </gwml2w:Elevation>
    </gwml2w:gwWellReferenceElevation>
    <gwml2w:gwWellStatus xlink:href="http://somevocab.org/gwml/2.0/notes" xlink:title="standby"/>
    <gwml2w:gwWellTotalLength>
        <swe:Quantity>
            <swe:uom xlink:href="http://qudt.org/vocab/unit#Meter" xlink:title="metre"/>
            <swe:value>93.0</swe:value>
        </swe:Quantity>
    </gwml2w:gwWellTotalLength>
    <gwml2w:gwWellWaterUse xlink:href="http://somevocab.org/gwml/2.0/notes" xlink:title="monitoring"/>
    <gwml2w:gwWellGeology>
        <gwml2w:GW_GeologyLog gml:id="soe.well.lithology.352271">
            <gml:identifier codeSpace="http://www.gns.cri.nz">
                http://vocab.smart-project.info/ngmp/soe/log/352271
            </gml:identifier>
            <om:phenomenonTime xlink:href="http://www.opengis.net/def/nil/OGC/0/unknown"/>
            <om:resultTime xlink:href="http://www.opengis.net/def/nil/OGC/0/unknown"/>
            <om:procedure xlink:href="http://www.opengis.net/def/nil/OGC/0/unknown"/>
            <om:observedProperty
                xlink:href="http://www.opengis.net/def/gwml/2.2/observedProperty/earthMaterial"
                xlink:title="lithology"/>
            <om:featureOfInterest xlink:href="http://vocab.smart-project.info/ngmp/soe/well/352271"/>
            <om:result>
                <gwml2w:GW_GeologyLogCoverage gml:id="soe.well.lithology.coverage352271">
                    <gml:identifier codeSpace="http://www.gns.cri.nz">
                        http://vocab.smart-project.info/ngmp/soe/log-cov/352271
                    </gml:identifier>
                    <gwml2w:element>
                        <gwml2w:LogValue>
                            <gwml2w:fromDepth>
                                <swe:Quantity>
                                    <swe:uom xlink:href="http://qudt.org/vocab/unit#Meter"
                                        xlink:title="metre"/>
                                    <swe:value>0.0</swe:value>
                                </swe:Quantity>
                            </gwml2w:fromDepth>
                            <gwml2w:toDepth>
                                <swe:Quantity>
                                    <swe:uom xlink:href="http://qudt.org/vocab/unit#Meter"
                                        xlink:title="metre"/>
```

```
                <swe:value>0.3</swe:value>
            </swe:Quantity>
        </gwml2w:toDepth>
        <gwml2w:value>
            <swe:DataRecord
                definition="http://www.opengis.net/def/gwml/2.2/datarecord/earthMaterial"
                id="log.id.1672">
                <swe:field name="lithology">
                    <swe:Category
                        definition="http://www.opengis.net/def/gwml/2.2/observedProperty/earthMaterial">
                        <swe:codeSpace
                            xlink:href="http://resource.geosciml.org/classifierscheme/cgi/201211/simplelithology"
                            xlink:title="Simple lithology"/>
                        <swe:value>Top Soil</swe:value>
                    </swe:Category>
                </swe:field>
                <swe:field name="description">
                    <swe:Text
                        definition="http://www.gwml.org/def/gwml/2.2/datarecord/FreeText">
                        <swe:value>topsoil and sand</swe:value>
                    </swe:Text>
                </swe:field>
            </swe:DataRecord>
        </gwml2w:value>
    </gwml2w:LogValue>
</gwml2w:element>
<gwml2w:element>
    <gwml2w:LogValue>
        <gwml2w:fromDepth>
            <swe:Quantity>
                <swe:uom xlink:href="http://qudt.org/vocab/unit#Meter"
                    xlink:title="metre"/>
                <swe:value>0.3</swe:value>
            </swe:Quantity>
        </gwml2w:fromDepth>
        <gwml2w:toDepth>
            <swe:Quantity>
                <swe:uom xlink:href="http://qudt.org/vocab/unit#Meter"
                    xlink:title="metre"/>
                <swe:value>2.4</swe:value>
            </swe:Quantity>
        </gwml2w:toDepth>
        <gwml2w:value>
            <swe:DataRecord
                definition="http://www.opengis.net/def/gwml/2.2/datarecord/earthMaterial"
                id="log.id.1673">
                <swe:field name="lithology">
                    <swe:Category
                        definition="http://www.opengis.net/def/gwml/2.2/observedProperty/earthMaterial">
                        <swe:codeSpace
                            xlink:href="http://resource.geosciml.org/classifierscheme/cgi/201211/simplelithology"
                            xlink:title="Simple lithology"/>
                        <swe:value>Peat</swe:value>
                    </swe:Category>
                </swe:field>
                <swe:field name="description">
                    <swe:Text
                        definition="http://www.gwml.org/def/gwml/2.2/datarecord/FreeText">
                        <swe:value>organic</swe:value>
                    </swe:Text>
                </swe:field>
            </swe:DataRecord>
        </gwml2w:value>
    </gwml2w:LogValue>
</gwml2w:element>
<gwml2w:element>
    <gwml2w:LogValue>
        <gwml2w:fromDepth>
            <swe:Quantity>
                <swe:uom xlink:href="http://qudt.org/vocab/unit#Meter"
                    xlink:title="metre"/>
                <swe:value>2.4</swe:value>
            </swe:Quantity>
        </gwml2w:fromDepth>
        <gwml2w:toDepth>
            <swe:Quantity>
                <swe:uom xlink:href="http://qudt.org/vocab/unit#Meter"
                    xlink:title="metre"/>
                <swe:value>14.6</swe:value>
            </swe:Quantity>
        </gwml2w:toDepth>
        <gwml2w:value>
            <swe:DataRecord
                definition="http://www.opengis.net/def/gwml/2.2/datarecord/earthMaterial"
                id="log.id.1674">
                <swe:field name="lithology">
                    <swe:Category
                        definition="http://www.opengis.net/def/gwml/2.2/observedProperty/earthMaterial">
                        <swe:codeSpace
                            xlink:href="http://resource.geosciml.org/classifierscheme/cgi/201211/simplelithology"
                            xlink:title="Simple lithology"/>
                        <swe:value>Sand</swe:value>
                    </swe:Category>
                </swe:field>
                <swe:field name="description">
                    <swe:Text
                        definition="http://www.gwml.org/def/gwml/2.2/datarecord/FreeText">
                        <swe:value>sand</swe:value>
                    </swe:Text>
```

```
                                </swe:field>
                            </swe:DataRecord>
                        </gwml2w:value>
                    </gwml2w:LogValue>
                </gwml2w:element>
                <gwml2w:element>
                    <gwml2w:LogValue>
                        <gwml2w:fromDepth>
                            <swe:Quantity>
                                <swe:uom xlink:href="http://qudt.org/vocab/unit#Meter"
                                    xlink:title="metre"/>
                                <swe:value>14.6</swe:value>
                            </swe:Quantity>
                        </gwml2w:fromDepth>
                        <gwml2w:toDepth>
                            <swe:Quantity>
                                <swe:uom xlink:href="http://qudt.org/vocab/unit#Meter"
                                    xlink:title="metre"/>
                                <swe:value>17.0</swe:value>
                            </swe:Quantity>
                        </gwml2w:toDepth>
                        <gwml2w:value>
                            <swe:DataRecord
                                definition="http://www.opengis.net/def/gwml/2.2/datarecord/earthMaterial"
                                id="log.id.1675">
                                <swe:field name="lithology">
                                    <swe:Category
                                        definition="http://www.opengis.net/def/gwml/2.2/observedProperty/earthMaterial">
                                        <swe:codeSpace
                                            xlink:href="http://resource.geosciml.org/classifierscheme/cgi/201211/simplelithology"
                                            xlink:title="Simple lithology"/>
                                        <swe:value>Gravel</swe:value>
                                    </swe:Category>
                                </swe:field>
                                <swe:field name="description">
                                    <swe:Text
                                        definition="http://www.gwml.org/def/gwml/2.2/datarecord/FreeText">
                                        <swe:value>fine stones and blue clay</swe:value>
                                    </swe:Text>
                                </swe:field>
                            </swe:DataRecord>
                        </gwml2w:value>
                    </gwml2w:LogValue>
                </gwml2w:element>
                <gwml2w:element>
                    <gwml2w:LogValue>
                        <gwml2w:fromDepth>
                            <swe:Quantity>
                                <swe:uom xlink:href="http://qudt.org/vocab/unit#Meter"
                                    xlink:title="metre"/>
                                <swe:value>17.0</swe:value>
                            </swe:Quantity>
                        </gwml2w:fromDepth>
                        <gwml2w:toDepth>
                            <swe:Quantity>
                                <swe:uom xlink:href="http://qudt.org/vocab/unit#Meter"
                                    xlink:title="metre"/>
                                <swe:value>93.2</swe:value>
                            </swe:Quantity>
                        </gwml2w:toDepth>
                        <gwml2w:value>
                            <swe:DataRecord
                                definition="http://www.opengis.net/def/gwml/2.2/datarecord/earthMaterial"
                                id="log.id.1676">
                                <swe:field name="lithology">
                                    <swe:Category
                                        definition="http://www.opengis.net/def/gwml/2.2/observedProperty/earthMaterial">
                                        <swe:codeSpace
                                            xlink:href="http://resource.geosciml.org/classifierscheme/cgi/201211/simplelithology"
                                            xlink:title="Simple lithology"/>
                                        <swe:value>Rock</swe:value>
                                    </swe:Category>
                                </swe:field>
                                <swe:field name="description">
                                    <swe:Text
                                        definition="http://www.gwml.org/def/gwml/2.2/datarecord/FreeText">
                                        <swe:value>greywacke greywacke</swe:value>
                                    </swe:Text>
                                </swe:field>
                            </swe:DataRecord>
                        </gwml2w:value>
                    </gwml2w:LogValue>
                </gwml2w:element>
            </gwml2w:GW_GeologyLogCoverage>
        </om:result>
        <gwml2w:startDepth>
            <swe:Quantity>
                <swe:uom xlink:href="http://qudt.org/vocab/unit#Meter" xlink:title="metre"/>
                <swe:value>0.0</swe:value>
            </swe:Quantity>
        </gwml2w:startDepth>
        <gwml2w:endDepth>
            <swe:Quantity>
                <swe:uom xlink:href="http://qudt.org/vocab/unit#Meter" xlink:title="metre"/>
                <swe:value>93.2</swe:value>
            </swe:Quantity>
        </gwml2w:endDepth>
    </gwml2w:GW_GeologyLog>
</gwml2w:gwWellGeology>
```

```
</gwml2w:GW_Well>
```

Listing 18: Instance encoding example for wells with bore lithology as `gwml2:GW_Well`

B.2 SENSOR WEB INSTANCE DOCUMENTS


```
<?xml version="1.0" encoding="UTF-8"?>
<sams:SF_SpatialSamplingFeature gml:id="ngmp.380"  xmlns:xs="http://www.w3.org/2001/XMLSchema"   xmlns:gmd="http://www.isotc211.org
      /2005/gmd"   xmlns:gco="http://www.isotc211.org/2005/gco"   xmlns:swe="http://www.opengis.net/swe/2.0"   xmlns:sam="http://www.
      opengis.net/sampling/2.0"  xmlns:om="http://www.opengis.net/om/2.0"   xmlns:xlink="http://www.w3.org/1999/xlink"   xmlns:gml="
      http://www.opengis.net/gml/3.2"   xmlns:sams="http://www.opengis.net/samplingSpatial/2.0"   xmlns:xsi="http://www.w3.org/2001/
      XMLSchema-instance"  timeStamp="2016-06-01T12:12:04.491Z"   xsi:schemaLocation="http://www.isotc211.org/2005/gco http://www.
      isotc211.org/2005/gco/gco.xsd http://www.opengis.net/gml/3.2    http://schemas.opengis.net/gml/3.2.1/gml.xsd    http://www.
      opengis.net/samplingSpatial/2.0   http://schemas.opengis.net/samplingSpatial/2.0/spatialSamplingFeature.xsd http://www.isotc211
      .org/2005/gmd   http://www.isotc211.org/2005/gmd/gmd.xsd http://www.opengis.net/om/2.0 http://schemas.opengis.net/om/2.0/
      observation.xsd">
  <gml:identifier codeSpace="http://www.gns.cri.nz">http://vocab.smart-project.info/ngmp/feature/380</gml:identifier>
  <gml:name codeSpace="http://www.gns.cri.nz">WCRC-AGNEW</gml:name>
  <sam:type xlink:href="http://www.opengis.net/def/samplingFeatureType/OGC-OM/2.0/SF_SamplingPoint"/>
  <sam:sampledFeature xsi:nil="true" nilReason="missing"/>
  <sam:relatedObservation>
    <om:OM_Observation gml:id="ngmp.om.id.107">
      <gml:identifier codeSpace="http://www.gns.cri.nz">http://vocab.smart-project.info/ngmp/da/om.id.107</gml:identifier>
      <om:type xlink:href="http://www.opengis.net/def/observationType/OGC-OM/2.0/OM_Measurement"/>
      <om:phenomenonTime>
        <gml:TimePeriod>
          <gml:beginPosition>1998-12-07T15:30:00.000+13:00</gml:beginPosition>
          <gml:endPosition>2014-06-24T15:00:00.000+12:00</gml:endPosition>
        </gml:TimePeriod>
      </om:phenomenonTime>
      <om:resultTime>
        <gml:TimeInstant>
          <gml:timePosition>2014-09-18T00:00:00.000+12:00</gml:timePosition>
        </gml:TimeInstant>
      </om:resultTime>
      <om:procedure xlink:href="http://vocab.smart-project.info/ngmp/procedure/1005"/>
      <om:parameter>
        <om:NamedValue>
          <om:name xlink:arcrole="observationdata" xlink:href="http://eodp.domain.nz/def/param-name/eodp/1.0/relatedSOSendpoint"
      xlink:role="relatedSOSendpoint"/>
          <om:value xlink:arcrole="protocol" xlink:href="http://portal.smart-project.info/sos-smart/sos/kvp" xlink:role="OGC:SOS-2.0"
      />
        </om:NamedValue>
      </om:parameter>
      <om:observedProperty xlink:href="http://vocab.smart-project.info/ngmp/phenomenon/1005"/>
      <om:featureOfInterest xlink:href="http://vocab.smart-project.info/ngmp/feature/380"/>
      <om:resultQuality xlink:href="http://www.opengis.net/def/nil/OGC/0/withheld"/>
      <om:result xlink:href="http://www.opengis.net/def/nil/OGC/0/withheld"/>
    </om:OM_Observation>
  </sam:relatedObservation>
  <sam:relatedObservation>
    <om:OM_Observation gml:id="ngmp.om.id.253">
      <gml:identifier codeSpace="http://www.gns.cri.nz">http://vocab.smart-project.info/ngmp/da/om.id.253</gml:identifier>
      <om:type xlink:href="http://www.opengis.net/def/observationType/OGC-OM/2.0/OM_Measurement"/>
      <om:phenomenonTime>
        <gml:TimePeriod>
          <gml:beginPosition>1998-12-07T15:30:00.000+13:00</gml:beginPosition>
          <gml:endPosition>2014-06-24T15:00:00.000+12:00</gml:endPosition>
        </gml:TimePeriod>
      </om:phenomenonTime>
      <om:resultTime>
        <gml:TimeInstant>
          <gml:timePosition>2014-09-18T00:00:00.000+12:00</gml:timePosition>
        </gml:TimeInstant>
      </om:resultTime>
      <om:procedure xlink:href="http://vocab.smart-project.info/ngmp/procedure/1106"/>
      <om:parameter>
        <om:NamedValue>
          <om:name xlink:arcrole="observationdata" xlink:href="http://eodp.domain.nz/def/param-name/eodp/1.0/relatedSOSendpoint"
      xlink:role="relatedSOSendpoint"/>
          <om:value xlink:arcrole="protocol" xlink:href="http://portal.smart-project.info/sos-smart/sos/kvp" xlink:role="OGC:SOS-2.0"
      />
        </om:NamedValue>
      </om:parameter>
      <om:observedProperty xlink:href="http://vocab.smart-project.info/ngmp/phenomenon/1106"/>
      <om:featureOfInterest xlink:href="http://vocab.smart-project.info/ngmp/feature/380"/>
      <om:resultQuality xlink:href="http://www.opengis.net/def/nil/OGC/0/withheld"/>
      <om:result xlink:href="http://www.opengis.net/def/nil/OGC/0/withheld"/>
    </om:OM_Observation>
  </sam:relatedObservation>
  <sam:relatedObservation>
    <om:OM_Observation gml:id="ngmp.om.id.392">
      <gml:identifier codeSpace="http://www.gns.cri.nz">http://vocab.smart-project.info/ngmp/da/om.id.392</gml:identifier>
      <om:type xlink:href="http://www.opengis.net/def/observationType/OGC-OM/2.0/OM_Measurement"/>
      <om:phenomenonTime>
        <gml:TimePeriod>
          <gml:beginPosition>1998-12-07T15:30:00.000+13:00</gml:beginPosition>
          <gml:endPosition>2014-06-24T15:00:00.000+12:00</gml:endPosition>
        </gml:TimePeriod>
      </om:phenomenonTime>
```

```
      <om:resultTime>
        <gml:TimeInstant>
          <gml:timePosition>2014-09-18T00:00:00.000+12:00</gml:timePosition>
        </gml:TimeInstant>
      </om:resultTime>
      <om:procedure xlink:href="http://vocab.smart-project.info/ngmp/procedure/1111"/>
      <om:parameter>
        <om:NamedValue>
          <om:name xlink:arcrole="observationdata" xlink:href="http://eodp.domain.nz/def/param-name/eodp/1.0/relatedSOSendpoint" xlink:role="relatedSOSendpoint"/>
          <om:value xlink:arcrole="protocol" xlink:href="http://portal.smart-project.info/sos-smart/sos/kvp" xlink:role="OGC:SOS-2.0" />
        </om:NamedValue>
      </om:parameter>
      <om:observedProperty xlink:href="http://vocab.smart-project.info/ngmp/phenomenon/1111"/>
      <om:featureOfInterest xlink:href="http://vocab.smart-project.info/ngmp/feature/380"/>
      <om:resultQuality xlink:href="http://www.opengis.net/def/nil/OGC/0/withheld"/>
      <om:result xlink:href="http://www.opengis.net/def/nil/OGC/0/withheld"/>
    </om:OM_Observation>
  </sam:relatedObservation>
  <sam:relatedObservation>
    <om:OM_Observation gml:id="ngmp.om.id.540">
      <gml:identifier codeSpace="http://www.gns.cri.nz">http://vocab.smart-project.info/ngmp/da/om.id.540</gml:identifier>
      <om:type xlink:href="http://www.opengis.net/def/observationType/OGC-OM/2.0/OM_Measurement"/>
      <om:phenomenonTime>
        <gml:TimePeriod>
          <gml:beginPosition>1998-09-15T10:45:00.000+12:00</gml:beginPosition>
          <gml:endPosition>2014-06-24T15:00:00.000+12:00</gml:endPosition>
        </gml:TimePeriod>
      </om:phenomenonTime>
      <om:resultTime>
        <gml:TimeInstant>
          <gml:timePosition>2014-09-18T00:00:00.000+12:00</gml:timePosition>
        </gml:TimeInstant>
      </om:resultTime>
      <om:procedure xlink:href="http://vocab.smart-project.info/ngmp/procedure/1416"/>
      <om:parameter>
        <om:NamedValue>
          <om:name xlink:arcrole="observationdata" xlink:href="http://eodp.domain.nz/def/param-name/eodp/1.0/relatedSOSendpoint" xlink:role="relatedSOSendpoint"/>
          <om:value xlink:arcrole="protocol" xlink:href="http://portal.smart-project.info/sos-smart/sos/kvp" xlink:role="OGC:SOS-2.0" />
        </om:NamedValue>
      </om:parameter>
      <om:observedProperty xlink:href="http://vocab.smart-project.info/ngmp/phenomenon/1416"/>
      <om:featureOfInterest xlink:href="http://vocab.smart-project.info/ngmp/feature/380"/>
      <om:resultQuality xlink:href="http://www.opengis.net/def/nil/OGC/0/withheld"/>
      <om:result xlink:href="http://www.opengis.net/def/nil/OGC/0/withheld"/>
    </om:OM_Observation>
  </sam:relatedObservation>
  <sam:relatedObservation>
    <om:OM_Observation gml:id="ngmp.om.id.698">
      <gml:identifier codeSpace="http://www.gns.cri.nz">http://vocab.smart-project.info/ngmp/da/om.id.698</gml:identifier>
      <om:type xlink:href="http://www.opengis.net/def/observationType/OGC-OM/2.0/OM_Measurement"/>
      <om:phenomenonTime>
        <gml:TimePeriod>
          <gml:beginPosition>1998-09-15T10:45:00.000+12:00</gml:beginPosition>
          <gml:endPosition>2014-06-24T15:00:00.000+12:00</gml:endPosition>
        </gml:TimePeriod>
      </om:phenomenonTime>
      <om:resultTime>
        <gml:TimeInstant>
          <gml:timePosition>2014-09-18T00:00:00.000+12:00</gml:timePosition>
        </gml:TimeInstant>
      </om:resultTime>
      <om:procedure xlink:href="http://vocab.smart-project.info/ngmp/procedure/1428"/>
      <om:parameter>
        <om:NamedValue>
          <om:name xlink:arcrole="observationdata" xlink:href="http://eodp.domain.nz/def/param-name/eodp/1.0/relatedSOSendpoint" xlink:role="relatedSOSendpoint"/>
          <om:value xlink:arcrole="protocol" xlink:href="http://portal.smart-project.info/sos-smart/sos/kvp" xlink:role="OGC:SOS-2.0" />
        </om:NamedValue>
      </om:parameter>
      <om:observedProperty xlink:href="http://vocab.smart-project.info/ngmp/phenomenon/1428"/>
      <om:featureOfInterest xlink:href="http://vocab.smart-project.info/ngmp/feature/380"/>
      <om:resultQuality xlink:href="http://www.opengis.net/def/nil/OGC/0/withheld"/>
      <om:result xlink:href="http://www.opengis.net/def/nil/OGC/0/withheld"/>
    </om:OM_Observation>
  </sam:relatedObservation>
  <sam:relatedObservation>
    <om:OM_Observation gml:id="ngmp.om.id.931">
      <gml:identifier codeSpace="http://www.gns.cri.nz">http://vocab.smart-project.info/ngmp/da/om.id.931</gml:identifier>
      <om:type xlink:href="http://www.opengis.net/def/observationType/OGC-OM/2.0/OM_Measurement"/>
      <om:phenomenonTime>
        <gml:TimePeriod>
          <gml:beginPosition>1998-09-15T10:45:00.000+12:00</gml:beginPosition>
          <gml:endPosition>2014-06-24T15:00:00.000+12:00</gml:endPosition>
        </gml:TimePeriod>
      </om:phenomenonTime>
      <om:resultTime>
        <gml:TimeInstant>
          <gml:timePosition>2014-09-18T00:00:00.000+12:00</gml:timePosition>
        </gml:TimeInstant>
      </om:resultTime>
      <om:procedure xlink:href="http://vocab.smart-project.info/ngmp/procedure/1662"/>
      <om:parameter>
        <om:NamedValue>
```

```
        <om:name xlink:arcrole="observationdata" xlink:href="http://eodp.domain.nz/def/param-name/eodp/1.0/relatedSOSendpoint"
     xlink:role="relatedSOSendpoint"/>
        <om:value xlink:arcrole="protocol" xlink:href="http://portal.smart-project.info/sos-smart/sos/kvp" xlink:role="OGC:SOS-2.0"
     />
      </om:NamedValue>
    </om:parameter>
    <om:observedProperty xlink:href="http://vocab.smart-project.info/ngmp/phenomenon/1662"/>
    <om:featureOfInterest xlink:href="http://vocab.smart-project.info/ngmp/feature/380"/>
    <om:resultQuality xlink:href="http://www.opengis.net/def/nil/OGC/0/withheld"/>
    <om:result xlink:href="http://www.opengis.net/def/nil/OGC/0/withheld"/>
  </om:OM_Observation>
 </sam:relatedObservation>
 <sam:relatedObservation>
  <om:OM_Observation gml:id="ngmp.om.id.1065">
    <gml:identifier codeSpace="http://www.gns.cri.nz">http://vocab.smart-project.info/ngmp/da/om.id.1065</gml:identifier>
    <om:type xlink:href="http://www.opengis.net/def/observationType/OGC-OM/2.0/OM_Measurement"/>
    <om:phenomenonTime>
      <gml:TimePeriod>
        <gml:beginPosition>1998-09-15T10:45:00.000+12:00</gml:beginPosition>
        <gml:endPosition>2014-06-24T15:00:00.000+12:00</gml:endPosition>
      </gml:TimePeriod>
    </om:phenomenonTime>
    <om:resultTime>
      <gml:TimeInstant>
        <gml:timePosition>2014-09-18T00:00:00.000+12:00</gml:timePosition>
      </gml:TimeInstant>
    </om:resultTime>
    <om:procedure xlink:href="http://vocab.smart-project.info/ngmp/procedure/1679"/>
    <om:parameter>
      <om:NamedValue>
        <om:name xlink:arcrole="observationdata" xlink:href="http://eodp.domain.nz/def/param-name/eodp/1.0/relatedSOSendpoint"
     xlink:role="relatedSOSendpoint"/>
        <om:value xlink:arcrole="protocol" xlink:href="http://portal.smart-project.info/sos-smart/sos/kvp" xlink:role="OGC:SOS-2.0"
     />
      </om:NamedValue>
    </om:parameter>
    <om:observedProperty xlink:href="http://vocab.smart-project.info/ngmp/phenomenon/1679"/>
    <om:featureOfInterest xlink:href="http://vocab.smart-project.info/ngmp/feature/380"/>
    <om:resultQuality xlink:href="http://www.opengis.net/def/nil/OGC/0/withheld"/>
    <om:result xlink:href="http://www.opengis.net/def/nil/OGC/0/withheld"/>
  </om:OM_Observation>
 </sam:relatedObservation>
 <sams:shape>
  <gml:Point gml:id="sams.shape.380" srsDimension="2">
    <gml:pos>171.04178451280865 -42.83874742863528</gml:pos>
  </gml:Point>
 </sams:shape>
</sams:SF_SpatialSamplingFeature>
</wfs:FeatureCollection>
```

Listing 19: Instance encoding example for a basic EODP sampling location of an NGMP site

```
<?xml version="1.0" encoding="UTF-8"?>
<swes:DescribeSensorResponse xmlns:swes="http://www.opengis.net/swes/2.0" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
     xmlns:gml="http://www.opengis.net/gml/3.2" xsi:schemaLocation="http://www.opengis.net/swes/2.0 http://schemas.opengis.net/swes
     /2.0/swesDescribeSensor.xsd http://www.opengis.net/gml/3.2 http://schemas.opengis.net/gml/3.2.1/gml.xsd http://www.opengis.net/
     gml http://schemas.opengis.net/gml/3.1.1/base/gml.xsd http://www.opengis.net/sensorML/1.0.1 http://schemas.opengis.net/sensorML
     /1.0.1/sensorML.xsd http://www.opengis.net/swe/1.0.1 http://schemas.opengis.net/sweCommon/1.0.1/swe.xsd">
 <swes:procedureDescriptionFormat>http://www.opengis.net/sensorML/1.0.1</swes:procedureDescriptionFormat>
 <swes:description>
  <swes:SensorDescription>
    <swes:validTime>
      <gml:TimePeriod gml:id="tp_10D16AA308FE3B6299F009AB3627287C68CFEAF3">
        <gml:beginPosition>2015-07-22T01:47:33.360Z</gml:beginPosition>
        <gml:endPosition indeterminatePosition="unknown"/>
      </gml:TimePeriod>
    </swes:validTime>
    <swes:data>
      <sml:SensorML xmlns:sml="http://www.opengis.net/sensorML/1.0.1" xmlns:swe="http://www.opengis.net/swe/1.0.1" xmlns:gml="http:
     //www.opengis.net/gml" version="1.0.1">
        <sml:member>
          <sml:System>
            <sml:keywords>
              <sml:KeywordList>
                <sml:keyword>soil moisture</sml:keyword>
                <sml:keyword>climate</sml:keyword>
                <sml:keyword>weather</sml:keyword>
              </sml:KeywordList>
            </sml:keywords>
            <sml:identification>
              <sml:IdentifierList>
                <sml:identifier name="uniqueID">
                  <sml:Term definition="urn:ogc:def:identifier:OGC:uniqueID">
                    <sml:value>http://vocab.smart-project.info/sensorweb/procedure/rangitaiki/hydro08/p40_soil_moisture</sml:value>
                  </sml:Term>
                </sml:identifier>
                <sml:identifier name="longName">
                  <sml:Term definition="urn:ogc:def:identifier:OGC:1.0:longName">
                    <sml:value>/sensorweb/procedure/rangitaiki/hydro08/p40_soil_moisture</sml:value>
                  </sml:Term>
                </sml:identifier>
                <sml:identifier name="shortName">
                  <sml:Term definition="urn:ogc:def:identifier:OGC:1.0:shortName">
                    <sml:value>p40_soil_moisture</sml:value>
                  </sml:Term>
```

```
                </sml:identifier>
                <sml:identifier name="modelNumber">
                  <sml:Term definition="urn:ogc:def:identifier:OGC:modelNumber">
                    <sml:value>10HS</sml:value>
                  </sml:Term>
                </sml:identifier>
                <sml:identifier name="manufacturer">
                  <sml:Term definition="urn:ogc:def:identifier:OGC:manufacturer">
                    <sml:value>Decagon</sml:value>
                  </sml:Term>
                </sml:identifier>
              </sml:IdentifierList>
            </sml:identification>
            <sml:classification>
              <sml:ClassifierList>
                <sml:classifier name="intendedApplication">
                  <sml:Term definition="urn:ogc:def:classifier:OGC:application">
                    <sml:value>weather</sml:value>
                  </sml:Term>
                </sml:classifier>
                <sml:classifier name="sensorType">
                  <sml:Term definition="urn:sensor:classifier:sensorType">
                    <sml:codeSpace xlink:href="http://sweet.nasa.jp/ontologoy/#atmosphere"/>
                    <sml:value>soil_moisture</sml:value>
                  </sml:Term>
                </sml:classifier>
              </sml:ClassifierList>
            </sml:classification>
            <sml:validTime>
              <gml:TimePeriod>
                <gml:beginPosition>2015-07-22T01:47:33.360Z</gml:beginPosition>
                <gml:endPosition indeterminatePosition="unknown"/>
              </gml:TimePeriod>
            </sml:validTime>
            <sml:capabilities name="featureOfInterest">
              <swe:SimpleDataRecord>
                <swe:field name="FeatureOfInterestID">
                  <swe:Text definition="FeatureOfInterest identifier">
                    <swe:value>http://vocab.smart-project.info/sensorweb/feature/hydro08</swe:value>
                  </swe:Text>
                </swe:field>
              </swe:SimpleDataRecord>
            </sml:capabilities>
            <sml:capabilities name="offerings">
              <swe:SimpleDataRecord>
                <swe:field name="Offering for sensor http://vocab.smart-project.info/sensorweb/procedure/rangitaiki/hydro08/p40_soil_moisture">
                  <swe:Text definition="http://www.opengis.net/def/offering/identifier">
                    <swe:value>http://vocab.smart-project.info/sensorweb/offering/rangitaiki/hydro08/p40_soil_moisture</swe:value>
                  </swe:Text>
                </swe:field>
              </swe:SimpleDataRecord>
            </sml:capabilities>
            <sml:capabilities name="measurementFrequency">
              <swe:SimpleDataRecord>
                <swe:field name="resolution">
                  <swe:Text definition="urn:ogc:def:property:OGC:measuremode">
                    <swe:value>LOW</swe:value>
                  </swe:Text>
                </swe:field>
              </swe:SimpleDataRecord>
            </sml:capabilities>
            <sml:contact>
              <sml:ContactList>
                <sml:member>
                  <sml:ResponsibleParty>
                    <sml:individualName>Hermann Klug</sml:individualName>
                    <sml:organizationName>Z_GIS</sml:organizationName>
                    <sml:positionName>Researcher</sml:positionName>
                    <sml:contactInfo>
                      <sml:phone>
                        <sml:voice>+43 662 8044-7561</sml:voice>
                      </sml:phone>
                      <sml:address>
                        <sml:deliveryPoint>Schillerstr. 30</sml:deliveryPoint>
                        <sml:city>Salzburg</sml:city>
                        <sml:postalCode>5020</sml:postalCode>
                        <sml:country>Austria</sml:country>
                      </sml:address>
                      <sml:onlineResource xlink:href="http://hermannklug.com"/>
                    </sml:contactInfo>
                  </sml:ResponsibleParty>
                </sml:member>
              </sml:ContactList>
            </sml:contact>
            <sml:position name="sensorPosition">
              <swe:Position fixed="false" referenceFrame="urn:ogc:def:crs:EPSG::4326">
                <swe:location>
                  <swe:Vector>
                    <swe:coordinate name="easting">
                      <swe:Quantity axisID="x">
                        <swe:uom code="degree"/>
                        <swe:value>-38.8194</swe:value>
                      </swe:Quantity>
                    </swe:coordinate>
                    <swe:coordinate name="northing">
                      <swe:Quantity axisID="y">
                        <swe:uom code="degree"/>
                        <swe:value>176.4875</swe:value>
```

```xml
                        </swe:Quantity>
                      </swe:coordinate>
                      <swe:coordinate name="altitude">
                        <swe:Quantity axisID="z">
                          <swe:uom code="m"/>
                          <swe:value>750.0</swe:value>
                        </swe:Quantity>
                      </swe:coordinate>
                    </swe:Vector>
                  </swe:location>
                </swe:Position>
              </sml:position>
              <sml:inputs>
                <sml:InputList>
                  <sml:input name="soil_moisture">
                    <swe:ObservableProperty definition="http://vocab.smart-project.info/sensorweb/phenomenon/soil_moisture"/>
                  </sml:input>
                </sml:InputList>
              </sml:inputs>
              <sml:outputs>
                <sml:OutputList>
                  <sml:output name="soil_moisture">
                    <swe:Quantity definition="http://vocab.smart-project.info/sensorweb/phenomenon/soil_moisture">
                      <swe:uom code="%"/>
                    </swe:Quantity>
                  </sml:output>
                </sml:OutputList>
              </sml:outputs>
            </sml:System>
          </sml:member>
        </sml:SensorML>
      </swes:data>
    </swes:SensorDescription>
  </swes:description>
</swes:DescribeSensorResponse>
```

Listing 20: Instance encoding example for a `SOS/OM_Procedure` sensor description of a soil moisture sensor at the Upper Rangitaiki case study site

```xml
<?xml version="1.0" encoding="UTF-8"?>
<swes:InsertSensor
    xmlns:swes="http://www.opengis.net/swes/2.0"
    xmlns:sos="http://www.opengis.net/sos/2.0"
    xmlns:swe="http://www.opengis.net/swe/1.0.1"
    xmlns:sml="http://www.opengis.net/sensorML/1.0.1"
    xmlns:gml="http://www.opengis.net/gml"
    xmlns:xlink="http://www.w3.org/1999/xlink"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" service="SOS" version="2.0.0" xsi:schemaLocation="http://www.opengis.net/sos/2.0 http://schemas.opengis.net/sos/2.0/sosInsertSensor.xsd   http://www.opengis.net/swes/2.0 http://schemas.opengis.net/swes/2.0/swes.xsd">
    <swes:procedureDescriptionFormat>http://www.opengis.net/sensorML/1.0.1</swes:procedureDescriptionFormat>
    <swes:procedureDescription>
<sml:SensorML  version="1.0.1">
          <sml:member>
            <sml:ProcessModel>
              <gml:description>aggregated sub-procedures (methods, equipments, sources) of obtaining measurements</gml:description>
              <sml:identification>
                <sml:IdentifierList>
                  <sml:identifier name="shortname">
                    <sml:Term definition="urn:ogc:def:identifier:OGC:1.0:shortname">
                      <sml:value>1679</sml:value>
                    </sml:Term>
                  </sml:identifier>
                  <sml:identifier name="uniqueID">
                    <sml:Term definition="urn:ogc:def:identifier:OGC:uniqueID">
                      <sml:value>http://vocab.smart-project.info/ngmp/procedure/1679</sml:value>
                    </sml:Term>
                  </sml:identifier>
                  <sml:identifier name="offering">
                    <sml:Term definition="urn:ogc:def:identifier:OGC:offeringID">
                      <sml:value>http://vocab.smart-project.info/ngmp/offering/1679</sml:value>
                    </sml:Term>
                  </sml:identifier>
                  <sml:identifier name="longname">
                    <sml:Term definition="urn:ogc:def:identifier:OGC:1.0:longname">
                      <sml:value>Water Level - Static water level m (Distance of standing water level from datum plane under natural equilibrium conditions)</sml:value>
                    </sml:Term>
                  </sml:identifier>
                </sml:IdentifierList>
              </sml:identification>
              <sml:classification>
                <sml:ClassifierList>
                  <sml:classifier name="characteristic">
                    <sml:Term definition="gns.cri.nz:ggw:procedure:308:1:101">
                      <sml:value>comments. pressure (Pa)/ 9806.65;equip. Wika tronic line manameter;equip. desc. Wika Tronic Line Manometer;method. Unknown Analysis Method;method. descr. Unknown method of analysis, done at an unknown location;src. Calculated;src. descr. Source for parameter was from other data;</sml:value>
                    </sml:Term>
                  </sml:classifier>
                  <sml:classifier name="characteristic">
                    <sml:Term definition="gns.cri.nz:ggw:procedure:378:1:2">
                      <sml:value>equip. Pressure Gauge;equip. desc. Pressure Gauge (kPa);method. Unknown Analysis Method;method. descr. Unknown method of analysis, done at an unknown location;src. Measured;</sml:value>
                    </sml:Term>
                  </sml:classifier>
```

```
            <sml:classifier name="characteristic">
              <sml:Term definition="gns.cri.nz:ggw:procedure:377:1:2">
                <sml:value>abs. err. .01;equip. Solinst GW probe;equip. desc. Solinst water level meter which relies on electric circuit being completed when slim electrode makes contact with water;method. Unknown Analysis Method;method. descr. Unknown method of analysis, done at an unknown location;src. Measured;</sml:value>
              </sml:Term>
            </sml:classifier>
            <sml:classifier name="characteristic">
              <sml:Term definition="gns.cri.nz:ggw:procedure:363:1:2">
                <sml:value>perc. err. .01;min. detect. 7;max. detect. 7030;equip. Druck DPI800 Pressure Indicator;equip. desc. Pressure indicator;method. Unknown Analysis Method;method. descr. Unknown method of analysis, done at an unknown location;src. Measured;</sml:value>
              </sml:Term>
            </sml:classifier>
            <sml:classifier name="characteristic">
              <sml:Term definition="gns.cri.nz:ggw:procedure:334:1:2">
                <sml:value>equip. Manual tape with weight on the end;equip. desc. Simple tape measure with weight attached to end;method. Unknown Analysis Method;method. descr. Unknown method of analysis, done at an unknown location;src. Measured;</sml:value>
              </sml:Term>
            </sml:classifier>
            <sml:classifier name="characteristic">
              <sml:Term definition="gns.cri.nz:ggw:procedure:328:1:2">
                <sml:value>abs. err. .02;equip. Druck water level meter;equip. desc. Druck waterlevel meter;method. Unknown Analysis Method;method. descr. Unknown method of analysis, done at an unknown location;src. Measured;</sml:value>
              </sml:Term>
            </sml:classifier>
            <sml:classifier name="characteristic">
              <sml:Term definition="gns.cri.nz:ggw:procedure:312:1:2">
                <sml:value>equip. Colmark pressure;equip. desc. Colmark pressure device;method. Unknown Analysis Method;method. descr. Unknown method of analysis, done at an unknown location;src. Measured;</sml:value>
              </sml:Term>
            </sml:classifier>
            <sml:classifier name="characteristic">
              <sml:Term definition="gns.cri.nz:ggw:procedure:332:1:2">
                <sml:value>abs. err. .01;equip. Seba water level meter;equip. desc. Seba water level meter;method. Unknown Analysis Method;method. descr. Unknown method of analysis, done at an unknown location;src. Measured;</sml:value>
              </sml:Term>
            </sml:classifier>
            <sml:classifier name="characteristic">
              <sml:Term definition="gns.cri.nz:ggw:procedure:333:1:2">
                <sml:value>abs. err. .05;equip. Solinst water level meter;equip. desc. Solinst water level meter which relies on electric circuit being completed when slim electrode makes contact with water;method. Unknown Analysis Method;method. descr. Unknown method of analysis, done at an unknown location;src. Measured;</sml:value>
              </sml:Term>
            </sml:classifier>
            <sml:classifier name="characteristic">
              <sml:Term definition="gns.cri.nz:ggw:procedure:330:1:2">
                <sml:value>abs. err. .001;equip. Heron water level meter;equip. desc. Heron water level meter which relies on electric circuit being completed when slim electrode makes contact with water;method. Unknown Analysis Method;method. descr. Unknown method of analysis, done at an unknown location;src. Measured;</sml:value>
              </sml:Term>
            </sml:classifier>
            <sml:classifier name="characteristic">
              <sml:Term definition="gns.cri.nz:ggw:procedure:1:1:2">
                <sml:value>equip. Unknown Equipment;equip. desc. Unknown Equipment;method. Unknown Analysis Method;method. descr. Unknown method of analysis, done at an unknown location;src. Measured;</sml:value>
              </sml:Term>
            </sml:classifier>
          </sml:ClassifierList>
        </sml:classification>
  <sml:capabilities name="offerings">
                <!-- Special capabilities used to specify offerings. -->
                <!-- Parsed and removed during InsertSensor/UpdateSensorDescription,
        added during DescribeSensor. -->
                <!-- Offering is generated if not specified. -->
                <swe:SimpleDataRecord>
                    <!-- Field name is used for the offering's name -->
                    <swe:field name="Offering for 1679">
                        <swe:Text definition="urn:ogc:def:identifier:OGC:offeringID">
                            <swe:value>http://vocab.smart-project.info/ngmp/offering/1679</swe:value>
                        </swe:Text>
                    </swe:field>
                </swe:SimpleDataRecord>
            </sml:capabilities>
        <sml:contact>
          <sml:ContactList>
            <sml:member>
              <sml:ResponsibleParty>
                <sml:individualName>A. Kmoch</sml:individualName>
                <sml:organizationName>Institute of Geological &amp; Nuclear Sciences Ltd (GNS)</sml:organizationName>
                <sml:positionName>Student</sml:positionName>
                <sml:contactInfo>
                  <sml:phone>
                    <sml:voice>+64 7 374 8211</sml:voice>
                  </sml:phone>
                  <sml:address>
                    <sml:deliveryPoint>114 Karetoto Road</sml:deliveryPoint>
                    <sml:city>Taupo</sml:city>
                    <sml:postalCode>3384</sml:postalCode>
                    <sml:country>New Zealand</sml:country>
                  </sml:address>
                  <sml:onlineResource xlink:href="http://www.gns.cri.nz/" />
                </sml:contactInfo>
              </sml:ResponsibleParty>
            </sml:member>
          </sml:ContactList>
        </sml:contact>
        <sml:inputs>
```

```
                <sml:InputList>
                  <sml:input name="Water Level - Static water level m">
                    <swe:Quantity definition="http://vocab.smart-project.info/ngmp/phenomenon/1679">
                      <gml:description>Distance of standing water level from datum plane under natural equilibrium conditions in Metres, Measurement unit of Metres</gml:description>
                      <swe:uom code="m" />
                    </swe:Quantity>
                  </sml:input>
                </sml:InputList>
              </sml:inputs>
              <sml:outputs>
                <sml:OutputList>
                  <sml:output name="Water Level - Static water level m">
                    <swe:Quantity definition="http://vocab.smart-project.info/ngmp/phenomenon/1679">
                      <gml:description>Distance of standing water level from datum plane under natural equilibrium conditions in Metres, Measurement unit of Metres</gml:description>
                      <swe:uom code="m" />
                    </swe:Quantity>
                  </sml:output>
                </sml:OutputList>
              </sml:outputs>
              <sml:method>
                <sml:ProcessMethod>
                  <sml:rules>
                    <sml:RulesDefinition>
                      <gml:description>The procedure 'http://vocab.smart-project.info/ngmp/procedure/1679' generates the following output(s): 'http://vocab.smart-project.info/ngmp/phenomenon/1679'. The input(s) is/are unknown (this description is generated).</gml:description>
                    </sml:RulesDefinition>
                  </sml:rules>
                </sml:ProcessMethod>
              </sml:method>
            </sml:ProcessModel>
          </sml:member>
        </sml:SensorML>
    </swes:procedureDescription>
    <!-- multiple values possible -->
    <swes:observableProperty>http://vocab.smart-project.info/ngmp/phenomenon/1679</swes:observableProperty>
    <swes:metadata>
        <sos:SosInsertionMetadata>
            <sos:observationType>http://www.opengis.net/def/observationType/OGC-OM/2.0/OM_Measurement</sos:observationType>
            <!-- multiple values possible -->
            <sos:featureOfInterestType>http://www.opengis.net/def/samplingFeatureType/OGC-OM/2.0/SF_SamplingPoint</sos:featureOfInterestType>
        </sos:SosInsertionMetadata>
    </swes:metadata>
</swes:InsertSensor>
```

Listing 21: Instance encoding example for a `SOS/OM_Procedure` process description for groundwater sampling protocol for the NGMP database as generated through the adopted 52N SOS software

C FRAMEWORK METADATA LISTINGS


```xml
<?xml version="1.0" encoding="UTF-8"?>
<!-- pycsw 1.10.2 --><csw:GetRecordByIdResponse xmlns:csw="http://www.opengis.net/cat/csw/2.0.2"
                            xmlns:dc="http://purl.org/dc/elements/1.1/"
                            xmlns:os="http://a9.com/-/spec/opensearch/1.1/"
                            xmlns:atom="http://www.w3.org/2005/Atom"
                            xmlns:inspire_common="http://inspire.ec.europa.eu/schemas/common/1.0"
                            xmlns:gco="http://www.isotc211.org/2005/gco"
                            xmlns:gmd="http://www.isotc211.org/2005/gmd"
                            xmlns:xs="http://www.w3.org/2001/XMLSchema"
                            xmlns:srv="http://www.isotc211.org/2005/srv"
                            xmlns:apiso="http://www.opengis.net/cat/csw/apiso/1.0"
                            xmlns:rim="urn:oasis:names:tc:ebxml-regrep:xsd:rim:3.0"
                            xmlns:soapenv="http://www.w3.org/2003/05/soap-envelope"
                            xmlns:wrs="http://www.opengis.net/cat/wrs/1.0"
                            xmlns:sitemap="http://www.sitemaps.org/schemas/sitemap/0.9"
                            xmlns:ogc="http://www.opengis.net/ogc"
                            xmlns:dct="http://purl.org/dc/terms/"
                            xmlns:ows="http://www.opengis.net/ows"
                            xmlns:xlink="http://www.w3.org/1999/xlink"
                            xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
                            xmlns:dif="http://gcmd.gsfc.nasa.gov/Aboutus/xml/dif/"
                            xmlns:gml="http://www.opengis.net/gml"
                            xmlns:fgdc="http://www.opengis.net/cat/csw/csdgm"
                            xmlns:inspire_ds="http://inspire.ec.europa.eu/schemas/inspire_ds/1.0"
                            xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
                            xsi:schemaLocation="http://www.opengis.net/cat/csw/2.0.2 http://schemas.opengis.net/csw/2.0.2/CSW-discovery.xsd">
   <gmd:MD_Metadata xsi:schemaLocation="http://www.isotc211.org/2005/gmd http://schemas.opengis.net/csw/2.0.2/profiles/apiso/1.0.0/apiso.xsd">
      <gmd:fileIdentifier>
         <gco:CharacterString>urn:uuid:2c5f1309-d721-4299-88bf-e462c577b99a-horowhenua_ws:ewt_nzprj_new</gco:CharacterString>
      </gmd:fileIdentifier>
      <gmd:language>
         <gco:CharacterString/>
      </gmd:language>
      <gmd:contact/>
      <gmd:dateStamp>
         <gco:Date/>
      </gmd:dateStamp>
      <gmd:metadataStandardName/>
      <gmd:metadataStandardVersion/>
      <gmd:identificationInfo>
         <gmd:MD_DataIdentification>
            <gmd:citation>
               <gmd:CI_Citation>
                  <gmd:title>
                     <gco:CharacterString>ewt_nzprj_new</gco:CharacterString>
                  </gmd:title>
                  <gmd:date>
                     <gmd:CI_Date>
                        <gmd:date>
                           <gco:Date>2016-05-17</gco:Date>
                        </gmd:date>
                        <gmd:dateType>
                           <CI_DateTypeCode xmlns="http://www.isotc211.org/2005/gmd"
                                            codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/Codelist/ML_gmxCodelists.xml#CI_DateTypeCode"
                                            codeListValue="revision"/>
                        </gmd:dateType>
                     </gmd:CI_Date>
                  </gmd:date>
               </gmd:CI_Citation>
            </gmd:citation>
            <gmd:abstract>
               <gco:CharacterString>Generated from GeoTIFF</gco:CharacterString>
            </gmd:abstract>
            <gmd:pointOfContact>
               <gmd:CI_ResponsibleParty>
                  <gmd:individualName>
                     <gco:CharacterString>Rogier Westerhoff</gco:CharacterString>
                  </gmd:individualName>
                  <gmd:organisationName>
                     <gco:CharacterString>GNS Science</gco:CharacterString>
                  </gmd:organisationName>
                  <gmd:positionName>
                     <gco:CharacterString>Hydrogeology</gco:CharacterString>
                  </gmd:positionName>
                  <gmd:contactInfo>
                     <gmd:CI_Contact>
                        <gmd:address>
                           <gmd:CI_Address>
                              <gmd:electronicMailAddress>
                                 <gco:CharacterString>r.westerhoff@gns.cri.nz</gco:CharacterString>
                              </gmd:electronicMailAddress>
                           </gmd:CI_Address>
                        </gmd:address>
```

```
               </gmd:CI_Contact>
            </gmd:contactInfo>
            <gmd:role>
               <gmd:CI_RoleCode codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/Codelist/ML_gmxCodelists.xml#CI_RoleCode"
                                codeListValue="custodian">custodian</gmd:CI_RoleCode>
            </gmd:role>
         </gmd:CI_ResponsibleParty>
      </gmd:pointOfContact>
      <gmd:descriptiveKeywords>
         <gmd:MD_Keywords>
            <gmd:keyword>
               <gco:CharacterString>WCS</gco:CharacterString>
            </gmd:keyword>
            <gmd:keyword>
               <gco:CharacterString>GeoTIFF</gco:CharacterString>
            </gmd:keyword>
            <gmd:keyword>
               <gco:CharacterString>ewt_nzprj_new</gco:CharacterString>
            </gmd:keyword>
         </gmd:MD_Keywords>
      </gmd:descriptiveKeywords>
      <gmd:resourceConstraints>
         <gmd:MD_SecurityConstraints>
            <gmd:classification>
               <MD_ClassificationCode xmlns="http://www.isotc211.org/2005/gmd"
                                      codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/Codelist/ML_gmxCodelists.xml#MD_ClassificationCode"
                                      codeListValue="unclassified"/>
            </gmd:classification>
         </gmd:MD_SecurityConstraints>
      </gmd:resourceConstraints>
      <gmd:resourceConstraints>
         <gmd:MD_LegalConstraints>
            <gmd:accessConstraints>
               <MD_RestrictionCode xmlns="http://www.isotc211.org/2005/gmd"
                                   codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/Codelist/ML_gmxCodelists.xml#MD_RestrictionCode"
                                   codeListValue="license"/>
            </gmd:accessConstraints>
            <gmd:useConstraints>
               <MD_RestrictionCode xmlns="http://www.isotc211.org/2005/gmd"
                                   codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/Codelist/ML_gmxCodelists.xml#MD_RestrictionCode"
                                   codeListValue="license"/>
            </gmd:useConstraints>
            <gmd:otherConstraints>
               <gco:CharacterString>Creative Commons and Disclaimer</gco:CharacterString>
            </gmd:otherConstraints>
         </gmd:MD_LegalConstraints>
      </gmd:resourceConstraints>
      <gmd:spatialRepresentationType>
         <MD_SpatialRepresentationTypeCode xmlns="http://www.isotc211.org/2005/gmd"
                                           codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/Codelist/ML_gmxCodelists.xml#MD_SpatialRepresentationTypeCode"
                                           codeListValue="grid"/>
      </gmd:spatialRepresentationType>
      <gmd:language>
         <LanguageCode xmlns="http://www.isotc211.org/2005/gmd"
                       codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/Codelist/ML_gmxCodelists.xml#LanguageCode"
                       codeListValue="eng"/>
      </gmd:language>
      <gmd:topicCategory>
         <gmd:MD_TopicCategoryCode>climatologyMeteorologyAtmosphere</gmd:MD_TopicCategoryCode>
      </gmd:topicCategory>
      <gmd:extent>
         <gmd:EX_Extent>
            <gmd:geographicElement>
               <gmd:EX_GeographicBoundingBox>
                  <gmd:extentTypeCode>
                     <gco:Boolean>true</gco:Boolean>
                  </gmd:extentTypeCode>
                  <gmd:westBoundLongitude>
                     <gco:Decimal>160.00</gco:Decimal>
                  </gmd:westBoundLongitude>
                  <gmd:eastBoundLongitude>
                     <gco:Decimal>179.14</gco:Decimal>
                  </gmd:eastBoundLongitude>
                  <gmd:southBoundLatitude>
                     <gco:Decimal>-52.51</gco:Decimal>
                  </gmd:southBoundLatitude>
                  <gmd:northBoundLatitude>
                     <gco:Decimal>-30.95</gco:Decimal>
                  </gmd:northBoundLatitude>
               </gmd:EX_GeographicBoundingBox>
            </gmd:geographicElement>
         </gmd:EX_Extent>
      </gmd:extent>
   </gmd:MD_DataIdentification>
</gmd:identificationInfo>
<gmd:distributionInfo>
   <gmd:MD_Distribution>
      <gmd:distributionFormat>
         <gmd:MD_Format>
            <gmd:name>
               <gco:CharacterString>GeoTIFF</gco:CharacterString>
            </gmd:name>
            <gmd:version>
```

```xml
                    <gco:CharacterString>1.0</gco:CharacterString>
                 </gmd:version>
              </gmd:MD_Format>
           </gmd:distributionFormat>
           <gmd:transferOptions>
              <gmd:MD_DigitalTransferOptions>
                 <gmd:onLine>
                    <gmd:CI_OnlineResource>
                       <gmd:linkage>
                          <gmd:URL>http://portal.smart-project.info/geoserver/horowhenua_ws/wcs</gmd:URL>
                       </gmd:linkage>
                       <gmd:name>
                          <gco:CharacterString>horowhenua_ws:ewt_nzprj_new</gco:CharacterString>
                       </gmd:name>
                       <gmd:description>
                          <gco:CharacterString>OGC-Web Coverage Service</gco:CharacterString>
                       </gmd:description>
                       <gmd:function>
                          <CI_OnLineFunctionCode xmlns="http://www.isotc211.org/2005/gmd"
                                                 codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/Codelist/ML_gmxCodelists.xml#CI_OnLineFunctionCode"
                                                 codeListValue="download"/>
                       </gmd:function>
                    </gmd:CI_OnlineResource>
                 </gmd:onLine>
              </gmd:MD_DigitalTransferOptions>
           </gmd:transferOptions>
        </gmd:MD_Distribution>
     </gmd:distributionInfo>
  </gmd:MD_Metadata>
</csw:GetRecordByIdResponse>
```

Listing 22: Instance example of SMART Rainfall Recharge dataset ANZLIC/ISO metadata record

```xml
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<feed xmlns="http://www.w3.org/2005/Atom"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:georss="http://www.georss.org/georss"
  xmlns:gml="http://www.opengis.net/gml"
  xmlns:owc="http://www.opengis.net/owc/1.0"
  xmlns:ol="http://openlayers.org/context"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xml:lang="en">
  <link rel="profile"
    href="http://www.opengis.net/spec/owc-atom/1.0/req/core"
    title="This file is compliant with version 1.0 of OGC Context"/>
  <id>http://portal.smart-project.info/context/smart-nz</id>
  <title>New Zealand</title>
  <subtitle type="text/plain">
    New Zealand Overview, Other Databases
  </subtitle>
  <author>
    <name>Alex Kmoch</name>
    <email>a.kmoch@gns.cri.nz</email>
  </author>
  <updated>2016-02-20T17:26:23Z</updated>
  <owc:display>
    <owc:pixelWidth>800</owc:pixelWidth>
    <owc:pixelHeight>600</owc:pixelHeight>
  </owc:display>
  <rights>
    Copyright (c) 2011-2016.  Some rights reserved.  This feed
    licensed under a Creative Commons Attribution 3.0 License.
  </rights>
  <georss:where>
    <gml:Envelope srsName="EPSG:4326" srsDimension="2">
      <gml:lowerCorner>168 -45</gml:lowerCorner>
      <gml:upperCorner>182 -33</gml:upperCorner>
    </gml:Envelope>
  </georss:where>
  <!-- <dc:date>2009-01-23T09:08:56.000Z/2009-01-23T09:14:08.000Z</dc:date> -->
  <link rel="self" type="application/atom+xml" href="http://portal.smart-project.info/context/smart-nz.atom.xml"/>
  <link rel="icon" type="image/png" href="http://portal.smart-project.info/fs/images/nz_m.png"/>
  <category scheme="view-groups" term="nz_overview" label="Overview"/>
  <category scheme="view-groups" term="nz_other" label="Other Databases"/>
  <entry>
    <id>http://portal.smart-project.info/context/smart-nz_other-layer-765</id>
    <title>Linz NZ Terrain Relief (Topo50)</title>
    <updated>2016-02-21T11:58:23Z</updated>
    <dc:publisher>GNS</dc:publisher>
    <dc:creator>GNS</dc:creator>
    <dc:rights>Fees:none / Contraints:none</dc:rights>
    <content type="html">abstract about data</content>
    <link rel="related" href="http://portal.smart-project.info/" type="text/html"/>
    <category scheme="view-groups" term="nz_other" label="Other Databases"/>
    <georss:where>
      <gml:Envelope srsName="EPSG:4326" srsDimension="2">
        <gml:lowerCorner>168 -45</gml:lowerCorner>
        <gml:upperCorner>182 -33</gml:upperCorner>
      </gml:Envelope>
    </georss:where>
    <owc:offering code="http://www.opengis.net/spec/owc-atom/1.0/req/wms">
      <owc:operation code="GetCapabilities" method="GET" type="application/xml" href="https://data.linz.govt.nz/services;key=a8fb9bcd52684b7abe14dd4664ce9df9/wms?VERSION=1.3.0&REQUEST=GetCapabilities"/>
```

```
    <owc:operation code="GetMap" method="GET" type="image/png" href="https://data.linz.govt.nz/services;key=a8fb9bcd52684b7abe14dd4664ce9df9/wms?VERSION=1.3&REQUEST=GetMap&SRS=EPSG:4326&BBOX=168,-45,182,-33&WIDTH=800&HEIGHT=600&LAYERS=layer-765&FORMAT=image/png&TRANSPARENT=TRUE&EXCEPTIONS=application/vnd.ogc.se_xml"/>
  </owc:offering>
  <owc:offering code="http://www.opengis.net/spec/owc-atom/1.0/req/csw">
    <owc:operation code="GetCapabilities" method="GET" type="application/xml" href="http://data.linz.govt.nz/feeds/csw?SERVICE=CSW&VERSION=2.0.2&REQUEST=GetCapabilities"/>
    <owc:operation code="GetRecordById" method="POST" type="application/xml" href="http://data.linz.govt.nz/feeds/csw">
      <owc:request type="application/xml">
        <csw:GetRecordById xmlns:csw="http://www.opengis.net/cat/csw/2.0.2"
          xmlns:gmd="http://www.isotc211.org/2005/gmd/" xmlns:gml="http://www.opengis.net/gml"
          xmlns:ogc="http://www.opengis.net/ogc" xmlns:gco="http://www.isotc211.org/2005/gco"
          xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
          outputFormat="application/xml" outputSchema="http://www.isotc211.org/2005/gmd"
          service="CSW" version="2.0.2">
          <csw:Id>d181a201-b138-9336-7a73-d0f88123a6bd</csw:Id>
          <csw:ElementSetName>full</csw:ElementSetName>
        </csw:GetRecordById>
      </owc:request>
    </owc:operation>
  </owc:offering>
  <!--
  <owc:minScaleDenominator>2500</owc:minScaleDenominator>
  <owc:maxScaleDenominator>25000</owc:maxScaleDenominator>
  -->
  <!-- WMC OL2 extension -->
  <ol:transparent>true</ol:transparent>
  <ol:isBaseLayer>false</ol:isBaseLayer>
  <ol:opacity>0.9</ol:opacity>
  <ol:displayInLayerSwitcher>false</ol:displayInLayerSwitcher>
</entry>

<entry>
  <id>http://portal.smart-project.info/context/smart-nz_other-layer-767</id>
  <title>Linz NZ Mainland Topo50</title>
  <updated>2016-02-21T11:58:23Z</updated>
  <dc:publisher>GNS</dc:publisher>
  <dc:creator>GNS</dc:creator>
  <dc:rights>Fees:none / Contraints:none</dc:rights>
  <content type="html">abstract about data</content>
  <link rel="related" href="http://portal.smart-project.info/" type="text/html"/>
  <category scheme="view-groups" term="nz_other" label="Other Databases"/>
  <georss:where>
    <gml:Envelope srsName="EPSG:4326" srsDimension="2">
      <gml:lowerCorner>168 -45</gml:lowerCorner>
      <gml:upperCorner>182 -33</gml:upperCorner>
    </gml:Envelope>
  </georss:where>
  <owc:offering code="http://www.opengis.net/spec/owc-atom/1.0/req/wms">
    <owc:operation code="GetCapabilities" method="GET" type="application/xml" href="https://data.linz.govt.nz/services;key=a8fb9bcd52684b7abe14dd4664ce9df9/wms?VERSION=1.3.0&REQUEST=GetCapabilities"/>
    <owc:operation code="GetMap" method="GET" type="image/png" href="https://data.linz.govt.nz/services;key=a8fb9bcd52684b7abe14dd4664ce9df9/wms?VERSION=1.3&REQUEST=GetMap&SRS=EPSG:4326&BBOX=168,-45,182,-33&WIDTH=800&HEIGHT=600&LAYERS=layer-767&FORMAT=image/png&TRANSPARENT=TRUE&EXCEPTIONS=application/vnd.ogc.se_xml"/>
  </owc:offering>
  <owc:offering code="http://www.opengis.net/spec/owc-atom/1.0/req/csw">
    <owc:operation code="GetCapabilities" method="GET" type="application/xml" href="http://data.linz.govt.nz/feeds/csw?SERVICE=CSW&VERSION=2.0.2&REQUEST=GetCapabilities"/>
    <owc:operation code="GetRecordById" method="POST" type="application/xml" href="http://data.linz.govt.nz/feeds/csw">
      <owc:request type="application/xml">
        <csw:GetRecordById xmlns:csw="http://www.opengis.net/cat/csw/2.0.2"
          xmlns:gmd="http://www.isotc211.org/2005/gmd/" xmlns:gml="http://www.opengis.net/gml"
          xmlns:ogc="http://www.opengis.net/ogc" xmlns:gco="http://www.isotc211.org/2005/gco"
          xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
          outputFormat="application/xml" outputSchema="http://www.isotc211.org/2005/gmd"
          service="CSW" version="2.0.2">
          <csw:Id>d181a201-b138-9336-7a73-d0f88123a6bd</csw:Id>
          <csw:ElementSetName>full</csw:ElementSetName>
        </csw:GetRecordById>
      </owc:request>
    </owc:operation>
  </owc:offering>
  <!--
  <owc:minScaleDenominator>2500</owc:minScaleDenominator>
  <owc:maxScaleDenominator>25000</owc:maxScaleDenominator>
  -->
  <!-- WMC OL2 extension -->
  <ol:transparent>true</ol:transparent>
  <ol:isBaseLayer>false</ol:isBaseLayer>
  <ol:opacity>0.9</ol:opacity>
  <ol:displayInLayerSwitcher>false</ol:displayInLayerSwitcher>
</entry>

<entry>
  <id>http://portal.smart-project.info/context/smart-nz_other-NZL_GNS_1M_Lithostratigraphy</id>
  <title>NZ GNS QMAP</title>
  <updated>2016-02-21T11:58:23Z</updated>
  <dc:publisher>GNS</dc:publisher>
  <dc:creator>GNS</dc:creator>
  <dc:rights>Fees:none / Contraints:none</dc:rights>
  <content type="html">abstract about data</content>
  <link rel="related" href="http://portal.smart-project.info/" type="text/html"/>
  <category scheme="view-groups" term="nz_other" label="Other Databases"/>
  <georss:where>
    <gml:Envelope srsName="EPSG:4326" srsDimension="2">
      <gml:lowerCorner>168 -45</gml:lowerCorner>
      <gml:upperCorner>182 -33</gml:upperCorner>
    </gml:Envelope>
  </georss:where>
```

```
    <owc:offering code="http://www.opengis.net/spec/owc-atom/1.0/req/wms">
      <owc:operation code="GetCapabilities" method="GET" type="application/xml" href="http://maps.gns.cri.nz/geology/wms?VERSION=1.3.0&REQUEST=GetCapabilities"/>
      <owc:operation code="GetMap" method="GET" type="image/png" href="http://maps.gns.cri.nz/geology/wms?VERSION=1.3&REQUEST=GetMap&SRS=EPSG:4326&BBOX=168,-45,182,-33&WIDTH=800&HEIGHT=600&LAYERS=gns:NZL_GNS_1M_Lithostratigraphy&FORMAT=image/png&TRANSPARENT=TRUE&EXCEPTIONS=application/vnd.ogc.se_xml"/>
    </owc:offering>
    <owc:offering code="http://www.opengis.net/spec/owc-atom/1.0/req/csw">
      <owc:operation code="GetCapabilities" method="GET" type="application/xml" href="http://data.gns.cri.nz/metadata/srv/eng/csw?SERVICE=CSW&VERSION=2.0.2&REQUEST=GetCapabilities"/>
      <owc:operation code="GetRecordById" method="POST" type="application/xml" href="http://data.gns.cri.nz/metadata/srv/eng/csw">
        <owc:request type="application/xml">
          <csw:GetRecordById xmlns:csw="http://www.opengis.net/cat/csw/2.0.2"
            xmlns:gmd="http://www.isotc211.org/2005/gmd/" xmlns:gml="http://www.opengis.net/gml"
            xmlns:ogc="http://www.opengis.net/ogc" xmlns:gco="http://www.isotc211.org/2005/gco"
            xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
            outputFormat="application/xml" outputSchema="http://www.isotc211.org/2005/gmd"
            service="CSW" version="2.0.2">
            <csw:Id>f892d778-c62e-481a-9d01-a4cfeaa1e1ab</csw:Id>
            <csw:ElementSetName>full</csw:ElementSetName>
          </csw:GetRecordById>
        </owc:request>
      </owc:operation>
    </owc:offering>
    <!--
    <owc:minScaleDenominator>2500</owc:minScaleDenominator>
    <owc:maxScaleDenominator>25000</owc:maxScaleDenominator>
    -->
    <!-- WMC OL2 extension -->
    <ol:transparent>true</ol:transparent>
    <ol:isBaseLayer>false</ol:isBaseLayer>
    <ol:opacity>0.8</ol:opacity>
    <ol:displayInLayerSwitcher>false</ol:displayInLayerSwitcher>
  </entry>

  <entry>
    <id>http://portal.smart-project.info/context/smart-nz_other-nz-dtm-100x100</id>
    <title>NZ DTM 100x100</title>
    <updated>2016-02-21T11:58:23Z</updated>
    <dc:publisher>GNS</dc:publisher>
    <dc:creator>GNS</dc:creator>
    <dc:rights>Fees:none / Contraints:none</dc:rights>
    <content type="html">abstract about data</content>
    <link rel="related" href="http://portal.smart-project.info/" type="text/html"/>
    <category scheme="view-groups" term="nz_other" label="Other Databases"/>
    <georss:where>
      <gml:Envelope srsName="EPSG:4326" srsDimension="2">
        <gml:lowerCorner>168 -45</gml:lowerCorner>
        <gml:upperCorner>182 -33</gml:upperCorner>
      </gml:Envelope>
    </georss:where>
    <owc:offering code="http://www.opengis.net/spec/owc-atom/1.0/req/wms">
      <owc:operation code="GetCapabilities" method="GET" type="application/xml" href="http://portal.smart-project.info/geoserver/wms?VERSION=1.3.0&REQUEST=GetCapabilities"/>
      <owc:operation code="GetMap" method="GET" type="image/png" href="http://portal.smart-project.info/geoserver/wms?VERSION=1.3&REQUEST=GetMap&SRS=EPSG:4326&BBOX=168,-45,182,-33&WIDTH=800&HEIGHT=600&LAYERS=horowhenua_ws:nz-dtm-100x100&FORMAT=image/png&TRANSPARENT=TRUE&EXCEPTIONS=application/vnd.ogc.se_xml"/>
    </owc:offering>
    <owc:offering code="http://www.opengis.net/spec/owc-atom/1.0/req/csw">
      <owc:operation code="GetCapabilities" method="GET" type="application/xml" href="http://portal.smart-project.info/pycsw/csw?SERVICE=CSW&VERSION=2.0.2&REQUEST=GetCapabilities"/>
      <owc:operation code="GetRecordById" method="POST" type="application/xml" href="http://portal.smart-project.info/pycsw/csw">
        <owc:request type="application/xml">
          <csw:GetRecordById xmlns:csw="http://www.opengis.net/cat/csw/2.0.2"
            xmlns:gmd="http://www.isotc211.org/2005/gmd/" xmlns:gml="http://www.opengis.net/gml"
            xmlns:ogc="http://www.opengis.net/ogc" xmlns:gco="http://www.isotc211.org/2005/gco"
            xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
            outputFormat="application/xml" outputSchema="http://www.isotc211.org/2005/gmd"
            service="CSW" version="2.0.2">
            <csw:Id>urn:uuid:1f542dbe-a35d-46d7-9dff-64004226d23f-nz-dtm-100x100</csw:Id>
            <csw:ElementSetName>full</csw:ElementSetName>
          </csw:GetRecordById>
        </owc:request>
      </owc:operation>
    </owc:offering>
    <!--
    <owc:minScaleDenominator>2500</owc:minScaleDenominator>
    <owc:maxScaleDenominator>25000</owc:maxScaleDenominator>
    -->
    <!-- WMC OL2 extension -->
    <ol:transparent>true</ol:transparent>
    <ol:isBaseLayer>false</ol:isBaseLayer>
    <ol:opacity>0.8</ol:opacity>
    <ol:displayInLayerSwitcher>false</ol:displayInLayerSwitcher>
  </entry>
  </entry>
</feed>
```

Listing 23: OWC Context example New Zealand overview

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- pycsw 1.10.2 --><csw:GetRecordByIdResponse xmlns:csw="http://www.opengis.net/cat/csw/2.0.2"
                           xmlns:dc="http://purl.org/dc/elements/1.1/"
                           xmlns:os="http://a9.com/-/spec/opensearch/1.1/"
                           xmlns:atom="http://www.w3.org/2005/Atom"
                           xmlns:inspire_common="http://inspire.ec.europa.eu/schemas/common/1.0"
                           xmlns:gco="http://www.isotc211.org/2005/gco"
```

```xml
                        xmlns:gmd="http://www.isotc211.org/2005/gmd"
                        xmlns:xs="http://www.w3.org/2001/XMLSchema"
                        xmlns:srv="http://www.isotc211.org/2005/srv"
                        xmlns:apiso="http://www.opengis.net/cat/csw/apiso/1.0"
                        xmlns:rim="urn:oasis:names:tc:ebxml-regrep:xsd:rim:3.0"
                        xmlns:soapenv="http://www.w3.org/2003/05/soap-envelope"
                        xmlns:wrs="http://www.opengis.net/cat/wrs/1.0"
                        xmlns:sitemap="http://www.sitemaps.org/schemas/sitemap/0.9"
                        xmlns:ogc="http://www.opengis.net/ogc"
                        xmlns:dct="http://purl.org/dc/terms/"
                        xmlns:ows="http://www.opengis.net/ows"
                        xmlns:xlink="http://www.w3.org/1999/xlink"
                        xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
                        xmlns:dif="http://gcmd.gsfc.nasa.gov/Aboutus/xml/dif/"
                        xmlns:gml="http://www.opengis.net/gml"
                        xmlns:fgdc="http://www.opengis.net/cat/csw/csdgm"
                        xmlns:inspire_ds="http://inspire.ec.europa.eu/schemas/inspire_ds/1.0"
                        xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
                        xsi:schemaLocation="http://www.opengis.net/cat/csw/2.0.2 http://schemas.opengis.net/csw/2.0.2/CSW-discovery.xsd">
<gmd:MD_Metadata>
   <gmd:fileIdentifier>
      <gco:CharacterString>8254ec62-3e1d-445e-b840-2b5e83de5e2f</gco:CharacterString>
   </gmd:fileIdentifier>
   <gmd:language>
      <gco:CharacterString>eng</gco:CharacterString>
   </gmd:language>
   <gmd:characterSet>
      <gmd:MD_CharacterSetCode codeList="http://www.isotc211.org/2005/resources/codeList.xml#MD_CharacterSetCode"
                               codeListValue="utf8"/>
   </gmd:characterSet>
   <gmd:contact>
      <gmd:CI_ResponsibleParty>
         <gmd:individualName>
            <gco:CharacterString>Keller, H.M.</gco:CharacterString>
         </gmd:individualName>
         <gmd:organisationName>
            <gco:CharacterString>NZ Hydrological Society - Publications - Journal</gco:CharacterString>
         </gmd:organisationName>
         <gmd:contactInfo>
            <gmd:CI_Contact>
               <gmd:phone>
                  <gmd:CI_Telephone>
                     <gmd:voice>
                        <gco:CharacterString>+64 6 357 1605</gco:CharacterString>
                     </gmd:voice>
                  </gmd:CI_Telephone>
               </gmd:phone>
               <gmd:address>
                  <gmd:CI_Address>
                     <gmd:electronicMailAddress>
                        <gco:CharacterString>admin@hydrologynz.org.nz</gco:CharacterString>
                     </gmd:electronicMailAddress>
                  </gmd:CI_Address>
               </gmd:address>
               <gmd:onlineResource>
                  <gmd:CI_OnlineResource>
                     <gmd:linkage>
                        <gmd:URL>http://www.hydrologynz.org.nz/index.php/nzhs-publications/nzhs-journal</gmd:URL>
                     </gmd:linkage>
                  </gmd:CI_OnlineResource>
               </gmd:onlineResource>
            </gmd:CI_Contact>
         </gmd:contactInfo>
         <gmd:role>
            <gmd:CI_RoleCode codeList="http://www.isotc211.org/2005/resources/codeList.xml#CI_RoleCode"
                             codeListValue="author"/>
         </gmd:role>
      </gmd:CI_ResponsibleParty>
   </gmd:contact>
   <gmd:dateStamp>
      <gco:DateTime>2014-11-04T23:29:25.247+13:00</gco:DateTime>
   </gmd:dateStamp>
   <gmd:metadataStandardName/>
   <gmd:metadataStandardVersion/>
   <gmd:identificationInfo>
      <gmd:MD_DataIdentification>
         <gmd:citation>
            <gmd:CI_Citation>
               <gmd:title>
                  <gco:CharacterString>Sources of streamflow in a small high country catchment in Canterbury, New Zealand</gco:CharacterString>
               </gmd:title>
               <gmd:date>
                  <gmd:CI_Date>
                     <gmd:date>
                        <gco:Date>2016-05-17</gco:Date>
                     </gmd:date>
                     <gmd:dateType>
                        <CI_DateTypeCode xmlns="http://www.isotc211.org/2005/gmd"
                                         codeList="http://www.isotc211.org/2005/resources/codeList.xml#CI_DateTypeCode"
                                         codeListValue="revision"/>
                     </gmd:dateType>
                  </gmd:CI_Date>
               </gmd:date>
            </gmd:CI_Citation>
         </gmd:citation>
         <gmd:abstract>
```

```xml
          <gco:CharacterString>Camp Stream Basin, which has been investigated for its sources of streamflow, is located in the headwaters of Broken River, Craigieburn Range, Canterbury, New Zealand. It is a 234-acre basin with about half each if its area above&#xD;
                    and below bushline. Discharge, temperature and specific conductivity readings at nine locations within the basin were taken during baseflow conditions and used to trace the sources of streamflow. The measurements used the dye-dilution method and the&#xD;
                    catchment area was sub-divided into three main zones: bush, transition (partly bush and alpine), and alpine. The measurements indicated that all three zones contributed about equal parts to streamflow. On a unit area basis, however, the transition&#xD;
                    and alpine zones yielded almost 14 times as much as the bush zone. Streamflow seemed to originate in little springs and creeks above the bushline, but in numerous seepage horizons near the channel in the bush. Stream temperatures seemed to be&#xD;
                    affected only slightly by air temperatures. Conductivities showed a very strong relationship with discharge at the same location. New gauging methods on this basis are therefore suggested. This conductivity/discharge relationship varied considerably&#xD;
                    between the locations of measurement with a clear trend of increasing values of b (the regression coefficient ) with distance from the mouth of the basis. The hypothesis is made that this conductivity/discharge relationship is characteristic of how&#xD;
                    physical and chemical weathering of vegetation cover, soil and parent material affects the water discharged from the catch area above the point of measurement. This is supported by conductivity measurements from seepages and creeks affluent to the&#xD;
                    stream; they showed generally much higher values in the bush zone than in the alpine. Further studies to check this hypothesis are suggested.</gco:CharacterString>
      </gmd:abstract>
      <gmd:pointOfContact>
         <gmd:CI_ResponsibleParty>
            <gmd:individualName>
               <gco:CharacterString>Keller, H.M.</gco:CharacterString>
            </gmd:individualName>
            <gmd:organisationName>
               <gco:CharacterString>NZ Hydrological Society - Publications - Journal</gco:CharacterString>
            </gmd:organisationName>
            <gmd:contactInfo>
               <gmd:CI_Contact>
                  <gmd:phone>
                     <gmd:CI_Telephone>
                        <gmd:voice>
                           <gco:CharacterString>+64 6 357 1605</gco:CharacterString>
                        </gmd:voice>
                     </gmd:CI_Telephone>
                  </gmd:phone>
                  <gmd:address>
                     <gmd:CI_Address>
                        <gmd:electronicMailAddress>
                           <gco:CharacterString>admin@hydrologynz.org.nz</gco:CharacterString>
                        </gmd:electronicMailAddress>
                     </gmd:CI_Address>
                  </gmd:address>
                  <gmd:onlineResource>
                     <gmd:CI_OnlineResource>
                        <gmd:linkage>
                           <gmd:URL>http://www.hydrologynz.org.nz/index.php/nzhs-publications/nzhs-journal</gmd:URL>
                        </gmd:linkage>
                     </gmd:CI_OnlineResource>
                  </gmd:onlineResource>
               </gmd:CI_Contact>
            </gmd:contactInfo>
            <gmd:role>
               <gmd:CI_RoleCode codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/Codelist/ML_gmxCodelists.xml#CI_RoleCode"
                                codeListValue="author">author</gmd:CI_RoleCode>
            </gmd:role>
         </gmd:CI_ResponsibleParty>
      </gmd:pointOfContact>
      <gmd:descriptiveKeywords>
         <gmd:MD_Keywords>
            <gmd:keyword>
               <gco:CharacterString>hydrology</gco:CharacterString>
            </gmd:keyword>
            <gmd:keyword>
               <gco:CharacterString>New Zealand</gco:CharacterString>
            </gmd:keyword>
            <gmd:keyword>
               <gco:CharacterString>Journal of Hydrology</gco:CharacterString>
            </gmd:keyword>
            <gmd:type>
               <MD_KeywordTypeCode xmlns="http://www.isotc211.org/2005/gmd"
                                   codeList="http://www.isotc211.org/2005/resources/codeList.xml#MD_KeywordTypeCode"
                                   codeListValue="theme"/>
            </gmd:type>
         </gmd:MD_Keywords>
      </gmd:descriptiveKeywords>
      <gmd:resourceConstraints>
         <gmd:MD_Constraints>
            <gmd:useLimitation>
               <gco:CharacterString>The Journal of Hydrology (New Zealand) (ISSN 0022-1708.) - We have loaded all abstracts to Volume 51 and papers for free viewing up to Volume 46 as at 12 October 2011.</gco:CharacterString>
            </gmd:useLimitation>
         </gmd:MD_Constraints>
      </gmd:resourceConstraints>
      <gmd:resourceConstraints>
         <gmd:MD_LegalConstraints>
            <gmd:useLimitation>
               <gco:CharacterString>The Journal of Hydrology (New Zealand) (ISSN 0022-1708.) - We have loaded all abstracts to Volume 51 and papers for free viewing up to Volume 46 as at 12 October 2011.</gco:CharacterString>
            </gmd:useLimitation>
            <gmd:accessConstraints>
               <MD_RestrictionCode xmlns="http://www.isotc211.org/2005/gmd"
```

```
                                   codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/Codelist/ML_gmxCodelists.xml#MD_RestrictionCode"
                                   codeListValue="otherRestrictions"/>
            </gmd:accessConstraints>
            <gmd:useConstraints>
               <MD_RestrictionCode xmlns="http://www.isotc211.org/2005/gmd"
                                   codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/Codelist/ML_gmxCodelists.xml#MD_RestrictionCode"
                                   codeListValue="copyright"/>
            </gmd:useConstraints>
         </gmd:MD_LegalConstraints>
      </gmd:resourceConstraints>
      <gmd:spatialResolution>
         <gmd:MD_Resolution>
            <gmd:equivalentScale>
               <gmd:MD_RepresentativeFraction>
                  <gmd:denominator>
                     <gco:Integer>50000</gco:Integer>
                  </gmd:denominator>
               </gmd:MD_RepresentativeFraction>
            </gmd:equivalentScale>
         </gmd:MD_Resolution>
      </gmd:spatialResolution>
      <gmd:spatialResolution>
         <gmd:MD_Resolution>
            <gmd:distance>
               <gco:Distance uom="m"/>
            </gmd:distance>
         </gmd:MD_Resolution>
      </gmd:spatialResolution>
      <gmd:language>
         <LanguageCode xmlns="http://www.isotc211.org/2005/gmd"
                       codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/Codelist/ML_gmxCodelists.xml#LanguageCode"
                       codeListValue="eng">eng</LanguageCode>
      </gmd:language>
      <gmd:characterSet>
         <MD_CharacterSetCode xmlns="http://www.isotc211.org/2005/gmd"
                              codeList="http://www.isotc211.org/2005/resources/codeList.xml#MD_CharacterSetCode"
                              codeListValue="utf8"/>
      </gmd:characterSet>
      <gmd:topicCategory>
         <gmd:MD_TopicCategoryCode>inlandWaters</gmd:MD_TopicCategoryCode>
      </gmd:topicCategory>
      <gmd:extent>
         <gmd:EX_Extent>
            <gmd:geographicElement>
               <gmd:EX_GeographicBoundingBox>
                  <gmd:extentTypeCode>
                     <gco:Boolean>true</gco:Boolean>
                  </gmd:extentTypeCode>
                  <gmd:westBoundLongitude>
                     <gco:Decimal>169.129339</gco:Decimal>
                  </gmd:westBoundLongitude>
                  <gmd:eastBoundLongitude>
                     <gco:Decimal>174.097074</gco:Decimal>
                  </gmd:eastBoundLongitude>
                  <gmd:southBoundLatitude>
                     <gco:Decimal>-52.553725</gco:Decimal>
                  </gmd:southBoundLatitude>
                  <gmd:northBoundLatitude>
                     <gco:Decimal>-39.25153</gco:Decimal>
                  </gmd:northBoundLatitude>
               </gmd:EX_GeographicBoundingBox>
            </gmd:geographicElement>
         </gmd:EX_Extent>
      </gmd:extent>
   </gmd:MD_DataIdentification>
</gmd:identificationInfo>
<gmd:distributionInfo>
   <gmd:MD_Distribution>
      <gmd:distributionFormat>
         <gmd:MD_Format>
            <gmd:name>
               <gco:CharacterString>Journal Article in PDF</gco:CharacterString>
            </gmd:name>
            <gmd:version>
               <gco:CharacterString>1.3 scanned / 1.6 electronic articles</gco:CharacterString>
            </gmd:version>
         </gmd:MD_Format>
      </gmd:distributionFormat>
      <gmd:transferOptions>
         <gmd:MD_DigitalTransferOptions>
            <gmd:onLine>
               <gmd:CI_OnlineResource>
                  <gmd:linkage>
                     <gmd:URL>http://www.hydrologynz.co.nz/journal.php?article_id=536</gmd:URL>
                  </gmd:linkage>
                  <gmd:function>
                     <CI_OnLineFunctionCode xmlns="http://www.isotc211.org/2005/gmd"
                                            codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/Codelist/ML_gmxCodelists.xml#CI_OnLineFunctionCode"
                                            codeListValue="download"/>
                  </gmd:function>
               </gmd:CI_OnlineResource>
            </gmd:onLine>
         </gmd:MD_DigitalTransferOptions>
      </gmd:transferOptions>
   </gmd:MD_Distribution>
```

```
      </gmd:distributionInfo>
      <gmd:dataQualityInfo>
         <gmd:DQ_DataQuality>
            <gmd:scope>
               <gmd:DQ_Scope>
                  <gmd:level>
                     <gmd:MD_ScopeCode codeList="http://www.isotc211.org/2005/resources/codeList.xml#MD_ScopeCode"
                                       codeListValue="dataset"/>
                  </gmd:level>
               </gmd:DQ_Scope>
            </gmd:scope>
         </gmd:DQ_DataQuality>
      </gmd:dataQualityInfo>
      <gmd:metadataConstraints>
         <gmd:MD_Constraints>
            <gmd:useLimitation>
               <gco:CharacterString>The Journal of Hydrology (New Zealand) (ISSN 0022-1708.) - We have loaded all abstracts to Volume 51 and papers for free viewing up to Volume 46 as at 12 October 2011.</gco:CharacterString>
            </gmd:useLimitation>
         </gmd:MD_Constraints>
      </gmd:metadataConstraints>
   </gmd:MD_Metadata>
</csw:GetRecordByIdResponse>
```

Listing 24: Instance example of JoHNZ ANZLIC/ISO metadata record

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"    xmlns:dc="http://purl.org/dc/elements/1.1/"
    xmlns:dcterms="http://purl.org/dc/terms/"     xmlns:foaf="http://xmlns.com/foaf/0.1/"     xmlns:gml="http://www.opengis.net/gml"
    xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"    xmlns:skos="http://www.w3.org/2004/02/skos/core#"
    xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <!-- Collection -->
  <skos:Collection rdf:about="http://vocab.smart-project.info/collection/glossary/terms">
    <rdfs:label>NZ HS Freshwater Database Glossary</rdfs:label>
    <dc:title>NZ HS Freshwater Database Glossary</dc:title>
    <dc:description>SMART project and GNS collected geoscience terms</dc:description>
    <dc:creator>
      <foaf:Organization>
        <foaf:name>GNS</foaf:name>
      </foaf:Organization>
    </dc:creator>
    <dc:rights>CC-SA-BY-NC 3.0 NZ</dc:rights>
    <dcterms:issued>2014-11-17T20:55:27.460+13:00</dcterms:issued>
    <dcterms:modified>2014-11-10T20:55:27.460+13:00</dcterms:modified>
<skos:member>http://vocab.smart-project.info/glossary/term/2144</skos:member>
<skos:member>http://vocab.smart-project.info/glossary/term/2145</skos:member>
<skos:member>http://vocab.smart-project.info/glossary/term/2146</skos:member>
<skos:member>http://vocab.smart-project.info/glossary/term/2147</skos:member>
<skos:member>http://vocab.smart-project.info/glossary/term/2148</skos:member>
```

Listing 25: RDF/SKOS encoded NZHS Freshwater Glossary

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"    xmlns:dc="http://purl.org/dc/elements/1.1/"
    xmlns:dcterms="http://purl.org/dc/terms/"     xmlns:foaf="http://xmlns.com/foaf/0.1/"     xmlns:gml="http://www.opengis.net/gml"
    xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"    xmlns:skos="http://www.w3.org/2004/02/skos/core#"
    xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <!-- Collection -->
  <skos:Collection rdf:about="http://vocab.smart-project.info/collection/ngmp/phenomena">
    <rdfs:label>National Groundwater Monitoring Programme, observed properties</rdfs:label>
    <dc:title>National Groundwater Monitoring Programme, observed properties</dc:title>
    <dc:description>SMART project and GNS collected geoscience terms</dc:description>
    <dc:creator>
      <foaf:Organization>
        <foaf:name>GNS</foaf:name>
      </foaf:Organization>
    </dc:creator>
    <dc:rights>CC-SA-BY-NC 3.0 NZ</dc:rights>
    <dcterms:issued>2014-11-17T20:55:00.215+13:00</dcterms:issued>
    <dcterms:modified>2014-11-10T20:55:00.215+13:00</dcterms:modified>
<skos:member>http://vocab.smart-project.info/ngmp/phenomenon/1001</skos:member>
<skos:member>http://vocab.smart-project.info/ngmp/phenomenon/1002</skos:member>
<skos:member>http://vocab.smart-project.info/ngmp/phenomenon/1004</skos:member>
<skos:member>http://vocab.smart-project.info/ngmp/phenomenon/1005</skos:member>
<skos:member>http://vocab.smart-project.info/ngmp/phenomenon/1645</skos:member>
<skos:member>http://vocab.smart-project.info/ngmp/phenomenon/1652</skos:member>
<skos:member>http://vocab.smart-project.info/ngmp/phenomenon/1662</skos:member>
<skos:member>http://vocab.smart-project.info/ngmp/phenomenon/1676</skos:member>
<skos:member>http://vocab.smart-project.info/ngmp/phenomenon/1679</skos:member>
<skos:member>http://vocab.smart-project.info/ngmp/phenomenon/1681</skos:member>
<skos:member>http://vocab.smart-project.info/ngmp/phenomenon/1682</skos:member>
<skos:member>http://vocab.smart-project.info/ngmp/phenomenon/1686</skos:member>
<skos:Concept rdf:about="http://vocab.smart-project.info/ngmp/phenomenon/1662">
    <skos:label>Temperature Deg. C</skos:label>
    <dc:title>Temperature Deg. C</dc:title>
    <skos:prefLabel xml:lang="en">Temperature Deg. C</skos:prefLabel>
    <skos:definition xml:lang="en">[1662]: Temperature Deg. C (in B0C)</skos:definition>
    <dc:description>[1662]: Temperature Deg. C (in B0C)</dc:description>
    <dc:modified>2014-11-06T15:00:05.881+13:00</dc:modified>
    <skos:inCollection rdf:resource="http://vocab.smart-project.info/collection/ngmp/phenomena"/>
</skos:Concept>
<skos:Concept rdf:about="http://vocab.smart-project.info/ngmp/phenomenon/1676">
    <skos:label>Water Level - Draw down m</skos:label>
    <dc:title>Water Level - Draw down m</dc:title>
```

```xml
    <skos:prefLabel xml:lang="en">Water Level - Draw down m</skos:prefLabel>
    <skos:definition xml:lang="en">[1676]: Water Level - Draw down m (in m)</skos:definition>
    <dc:description>[1676]: Water Level - Draw down m (in m)</dc:description>
    <dc:modified>2014-11-06T15:00:05.900+13:00</dc:modified>
    <skos:inCollection rdf:resource="http://vocab.smart-project.info/collection/ngmp/phenomena"/>
</skos:Concept>
<skos:Concept rdf:about="http://vocab.smart-project.info/ngmp/phenomenon/1679">
    <skos:label>Water Level - Static water level m</skos:label>
    <dc:title>Water Level - Static water level m</dc:title>
    <skos:prefLabel xml:lang="en">Water Level - Static water level m</skos:prefLabel>
    <skos:definition xml:lang="en">[1679]: Water Level - Static water level m (in m)</skos:definition>
    <dc:description>[1679]: Water Level - Static water level m (in m)</dc:description>
    <dc:modified>2014-11-06T15:00:05.892+13:00</dc:modified>
    <skos:inCollection rdf:resource="http://vocab.smart-project.info/collection/ngmp/phenomena"/>
</skos:Concept>
<skos:Concept rdf:about="http://vocab.smart-project.info/ngmp/phenomenon/1681">
    <skos:label>Carbon mg/L All forms as CaCO3 - total</skos:label>
    <dc:title>Carbon mg/L All forms as CaCO3 - total</dc:title>
    <skos:prefLabel xml:lang="en">Carbon mg/L All forms as CaCO3 - total</skos:prefLabel>
    <skos:definition xml:lang="en">[1681]: Carbon mg/L All forms as CaCO3 - total (in mg/L)</skos:definition>
    <dc:description>[1681]: Carbon mg/L All forms as CaCO3 - total (in mg/L)</dc:description>
    <dc:modified>2014-11-06T15:00:06.048+13:00</dc:modified>
    <skos:inCollection rdf:resource="http://vocab.smart-project.info/collection/ngmp/phenomena"/>
</skos:Concept>
<skos:Concept rdf:about="http://vocab.smart-project.info/ngmp/phenomenon/1682">
    <skos:label>Carbon mg/L All forms as HCO3 - total</skos:label>
    <dc:title>Carbon mg/L All forms as HCO3 - total</dc:title>
    <skos:prefLabel xml:lang="en">Carbon mg/L All forms as HCO3 - total</skos:prefLabel>
    <skos:definition xml:lang="en">[1682]: Carbon mg/L All forms as HCO3 - total (in mg/L)</skos:definition>
    <dc:description>[1682]: Carbon mg/L All forms as HCO3 - total (in mg/L)</dc:description>
    <dc:modified>2014-11-06T15:00:06.122+13:00</dc:modified>
    <skos:inCollection rdf:resource="http://vocab.smart-project.info/collection/ngmp/phenomena"/>
</skos:Concept>
<skos:Concept rdf:about="http://vocab.smart-project.info/ngmp/phenomenon/1686">
    <skos:label>Bicarbonate mg/L as HCO3 - filterable</skos:label>
    <dc:title>Bicarbonate mg/L as HCO3 - filterable</dc:title>
    <skos:prefLabel xml:lang="en">Bicarbonate mg/L as HCO3 - filterable</skos:prefLabel>
    <skos:definition xml:lang="en">[1686]: Bicarbonate mg/L as HCO3 - filterable (in mg/L)</skos:definition>
    <dc:description>[1686]: Bicarbonate mg/L as HCO3 - filterable (in mg/L)</dc:description>
    <dc:modified>2014-11-06T15:00:06.042+13:00</dc:modified>
    <skos:inCollection rdf:resource="http://vocab.smart-project.info/collection/ngmp/phenomena"/>
</skos:Concept>
<skos:Concept rdf:about="http://vocab.smart-project.info/ngmp/phenomenon/1687">
    <skos:label>Nitrate mg/L as N - total</skos:label>
    <dc:title>Nitrate mg/L as N - total</dc:title>
    <skos:prefLabel xml:lang="en">Nitrate mg/L as N - total</skos:prefLabel>
    <skos:definition xml:lang="en">[1687]: Nitrate mg/L as N - total (in mg/L)</skos:definition>
    <dc:description>[1687]: Nitrate mg/L as N - total (in mg/L)</dc:description>
    <dc:modified>2014-11-06T15:00:06.045+13:00</dc:modified>
    <skos:inCollection rdf:resource="http://vocab.smart-project.info/collection/ngmp/phenomena"/>
</skos:Concept>
<skos:Concept rdf:about="http://vocab.smart-project.info/ngmp/phenomenon/2002">
    <skos:label>Salinity ppt</skos:label>
    <dc:title>Salinity ppt</dc:title>
    <skos:prefLabel xml:lang="en">Salinity ppt</skos:prefLabel>
    <skos:definition xml:lang="en">[2002]: Salinity ppt (in ppt)</skos:definition>
    <dc:description>[2002]: Salinity ppt (in ppt)</dc:description>
    <dc:modified>2014-11-06T15:00:05.923+13:00</dc:modified>
    <skos:inCollection rdf:resource="http://vocab.smart-project.info/collection/ngmp/phenomena"/>
</skos:Concept>
</rdf:RDF>
```

Listing 26: RDF/SKOS encoded NGMP Observed Properties from NGMP Parameters

D GEOVISUALISATION PROCESS

```java
package org.n52.wps.server.algorithm.smart;
/**
 * This algorithm loads a WMC doc and transforms it to an X3D scene.
 *
 * @author Alex Kmoch
 *
 */
@Algorithm(version = "1.0.0", abstrakt = "This algorithm loads a WMC doc and transforms it to an X3D scene.")
public class WmcSupervisionAlgorithm extends AbstractAnnotatedAlgorithm {

   private static Logger LOGGER = Logger.getLogger(WmcSupervisionAlgorithm.class);

   public WmcSupervisionAlgorithm() {
      super();
   }

   private String result;
   private String wmcdoc;

   private final String proxyUrl = "http://portal.smart-project.info/smart/proxy2?url=";

   @LiteralDataOutput(identifier = "result", binding = LiteralStringBinding.class)
   public String getResult() {
      return result;
   }

   @LiteralDataInput(identifier = "wmcDoc", abstrakt = "The WMC Document", minOccurs = 1)
   public void setWmcDoc(String wmcdoc) {
      this.wmcdoc = wmcdoc;
   }

   @Execute
   public void runAlgorithm() throws IOException, XmlException {
            BoundsModel bounds = WmcParserFactory.getBounds(wmcType);
            List<X3DGridLayerModel> testList = WmcParserFactory.parseXmlBean(wmcType, bounds);

            // here create X3D scene
            SimpleX3DStaxWriter x3dWriter = new SimpleX3DStaxWriter();
            x3dWriter.setLayerlist(testList);
            x3dWriter.setBounds(bounds);

            // TODO alternatively
            // new X3DUtil approach



            try {
               File temp = File.createTempFile("x3d-scene-", ".tmp");
               String tmpFileName = temp.getAbsolutePath();
               x3dWriter.setOutputfile(tmpFileName);
               x3dWriter.setProxyUrl(proxyUrl);
               x3dWriter.writeX3dLayer();
               InputStream in = new FileInputStream(tmpFileName);
               result = IOUtils.toString(in, "UTF-8");

            } catch (XMLStreamException e) {
               e.printStackTrace();
               result = e.getLocalizedMessage();

            } catch (FileNotFoundException e) {
               e.printStackTrace();
               result = e.getLocalizedMessage();
            }
```

Listing 27: Excerpts from `WmcSupervisionAlgorithm` class

```java
package org.n52.wps.server.algorithm.smart.Components;
public class X3DUtil {

   public static X3D createProperRootNode(X3DDocument theRootNodeDoc) {

      X3D rootNode = theRootNodeDoc.addNewX3D();

      rootNode.setProfile(X3DConstants.PROFILE_FULL);
      rootNode.setVersion(X3DConstants.X3D_VERSION);

      XmlCursor cursor2 = rootNode.newCursor();

      cursor2.toLastAttribute();
      cursor2.insertAttributeWithValue("noNamespaceSchemaLocation", X3DConstants.X3D_SCHEMALOCATION);
      cursor2.insertAttributeWithValue("width", "1200px");
      cursor2.insertAttributeWithValue("height", "600px");
      cursor2.insertAttributeWithValue("id", "rootNode");
      cursor2.insertAttributeWithValue("showStat", "true");
```

```java
    cursor2.insertAttributeWithValue("showLog", "true");
    cursor2.insertAttributeWithValue("showProgress", "bar");
    cursor2.dispose();

    return rootNode;
  }
  public static Head createHeadNode(List<X3DGridLayerModel> layerList, BoundsModel bounds, Map<String, String> metaList) {
  public static void basicSceneBootstrap(Scene scene, BoundsModel bounds) {

    // NOPE sky and goundcolour
    // scene.addNewBackground().setSkyColor("0.9 1 1");
    // scene.getBackgroundArray(0).setGroundColor("0.5 0.37 0");
    Background background = scene.addNewBackground();
    background.setDEF("basic_background");
    background.setSkyColor(getColor("#E0F2F7"));
    background.setGroundColor(getColor("#F8ECE0"));
    MetadataString bgmeta = MetadataString.Factory.newInstance();
    bgmeta.setName("background");
    background.setMetadataString(bgmeta);

    // NAV speed on 2
    NavigationInfo navinfo = scene.addNewNavigationInfo();
    navinfo.setSpeed(2.0f);
    navinfo.setType(X3DConstants.NAV_TYPES);
    MetadataString speedmeta = MetadataString.Factory.newInstance();
    speedmeta.setName("speedmeta");
    navinfo.setMetadataString(speedmeta);
```

Listing 28: Excerpts from `X3DUtil` class

```java
package org.n52.wps.server.algorithm.smart.models;
public class WmcParserFactory {

  private static Logger LOGGER = Logger.getLogger(WmcParserFactory.class);

  public static BoundsModel getBounds(ViewContextType wmcType) {

    BoundsModel bounds = new BoundsModel();

    // [0] minx="176.1" [1] miny="-38.11" [2] maxx="176.2" [3] maxy="-38.1"
    // SRS="EPSG:4326"
    // window size width/height px

    if (wmcType.getGeneral().getWindow() != null) {
      bounds.window[0] = wmcType.getGeneral().getWindow().getWidth().intValue();
      bounds.window[1] = wmcType.getGeneral().getWindow().getHeight().intValue();
    }

      if (wmcType.getGeneral().getBoundingBox() != null) {
      bounds.ogcSrs = wmcType.getGeneral().getBoundingBox().getSRS();
      bounds.bbox[0] = wmcType.getGeneral().getBoundingBox().getMinx().doubleValue();
      bounds.bbox[1] = wmcType.getGeneral().getBoundingBox().getMiny().doubleValue();
      bounds.bbox[2] = wmcType.getGeneral().getBoundingBox().getMaxx().doubleValue();
      bounds.bbox[3] = wmcType.getGeneral().getBoundingBox().getMaxy().doubleValue();
    }

      return bounds;
  }

  public static List<X3DGridLayerModel> parseXmlBean(ViewContextType wmcType, BoundsModel bounds) {

      List<X3DGridLayerModel> x3dLayerList = new ArrayList<X3DGridLayerModel>();

      String wmcTitle = "scene";

      // BoundsModel bounds
      // [0] minx="176.1" [1] miny="-38.11" [2] maxx="176.2" [3] maxy="-38.1"
    // SRS="EPSG:4326"
    // window size width/height px

      if (wmcType.getGeneral().getTitle() != null) {
      wmcTitle = wmcType.getGeneral().getTitle();
    }

    if (wmcType.getLayerList().getLayerArray() != null) {
      LayerType[] layerList = wmcType.getLayerList().getLayerArray();

      for (LayerType layerType : layerList) {
        X3DGridLayerModel layer = parseLayer(layerType, bounds);
        if (!layer.layerName.equalsIgnoreCase("NULL")) {
          x3dLayerList.add(layer);
        }
      }
    }

    return x3dLayerList;
  }
```

Listing 29: Excerpts from `WmcParserFactory` class

```xml
<ViewContext xmlns="http://www.opengis.net/context" xmlns:ol="http://openlayers.org/context"
  xmlns:sld="http://www.opengis.net/sld" version="1.1.0"
  id="OpenLayers_Context_66"
  xsi:schemaLocation="http://www.opengis.net/context http://schemas.opengis.net/context/1.1.0/context.xsd"
```

```
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <General>
    <Window width="800" height="600" />
    <BoundingBox minx="175" miny="-40.8"
      maxx="175.5" maxy="-40.4" SRS="EPSG:4326" />
    <Title>Horowhenua Context</Title>
  </General>
  <LayerList>
    <Layer queryable="1" hidden="0">
      <Server service="OGC:WMS" version="1.1.1">
        <OnlineResource xlink:type="simple"
          xmlns:xlink="http://www.w3.org/1999/xlink" xlink:href="http://portal.smart-project.info/geoserver/horowhenua_ws/wms" />
      </Server>
      <Name>horowhenua_ws:study_area</Name>
      <Title>Horowhenua Area</Title>
      <FormatList>
        <Format current="1">image/png</Format>
      </FormatList>
      <StyleList>
        <Style current="1">
          <Name />
          <Title>Default</Title>
        </Style>
      </StyleList>
      <Extension>
            <ol:transparent>true</ol:transparent>
            <ol:isBaseLayer>false</ol:isBaseLayer>
            <ol:opacity>0.6</ol:opacity>
            <ol:displayInLayerSwitcher>true</ol:displayInLayerSwitcher>
        </Extension>
    </Layer>
    <Layer queryable="1" hidden="0">
      <Server service="OGC:WMS" version="1.1.1">
        <OnlineResource xlink:type="simple"
          xmlns:xlink="http://www.w3.org/1999/xlink" xlink:href="http://portal.smart-project.info/geoserver/horowhenua_ws/wms" />
      </Server>
      <Name>horowhenua_ws:soe_gwl_monitoring_wells</Name>
      <Title>SoE gwl monitoring wells</Title>
      <FormatList>
        <Format current="1">image/png</Format>
      </FormatList>
      <StyleList>
        <Style current="1">
          <Name />
          <Title>Default</Title>
        </Style>
      </StyleList>
      <Extension>
            <ol:transparent>true</ol:transparent>
            <ol:isBaseLayer>false</ol:isBaseLayer>
            <ol:opacity>0.8</ol:opacity>
            <ol:displayInLayerSwitcher>true</ol:displayInLayerSwitcher>
        </Extension>
    </Layer>
  </LayerList>
</ViewContext>
```

Listing 30: Excerpts from a Horowhenua layer context document

```
<X3D profile="Full" noNamespaceSchemaLocation="http://www.web3d.org/specifications/x3d-3.3.xsd"
  width="1200px" height="600px" id="rootNode" showProgress="bar" showStat="false" showLog="false" version="">
  <head>
    <meta name="BBOX" content="174.949249 -40.782002 175.608429 -40.535882"/>
    <meta name="BBOX size in m x/y" content="55642.400642.2 27330.987156.2"/>
    <meta name="Layers" content="7"/>
    <meta name="WMC Window Size X/Y" content="1920 945"/>
    <meta name="WPS Process" content="org.n52.wps.server.algorithm.smart.WmcSupervisionAlgorithm"/>
    <meta name="Author" content="Alex K"/>
    <meta name="Project" content="SMART Aquifer Characterisation Programme"/>
    <meta name="WMC Title" content="Default WMC Title"/>
    <meta name="OGC:SRS" content="EPSG:4326"/>
    <meta name="LayerName_1" content="layer-767"/>
    <meta name="Layerref_1" content="https://data.linz.govt.nz/services;key=a8fbsbcd52684b7ase14dd4664ce9df9/wms"/>
    <meta name="LayerName_2" content="horowhenua_ws:study_area"/>
    <meta name="Layerref_2" content="http://portal.smart-project.info/geoserver/wms"/>
    <meta name="LayerName_3" content="horowhenua_ws:catchments"/>
    <meta name="Layerref_3" content="http://portal.smart-project.info/geoserver/wms"/>
    <meta name="LayerName_4" content="horowhenua_ws:lakes"/>
    <meta name="Layerref_4" content="http://portal.smart-project.info/geoserver/wms"/>
    <meta name="LayerName_5" content="horowhenua_ws:soe_gwl_monitoring_wells"/>
    <meta name="Layerref_5" content="http://portal.smart-project.info/geoserver/wms"/>
    <meta name="LayerName_6" content="horowhenua_ws:soe_gwl_monitoring_wells"/>
    <meta name="Layerref_6" content="http://portal.smart-project.info/geoserver/wfs"/>
    <meta name="LayerName_7" content="horowhenua_ws:soe_gwl_monitoring_wells"/>
    <meta name="Layerref_7" content="http://portal.smart-project.info/52nSOSv3.5.0/sos"/>
  </head>
  <Scene>
    <Background DEF="basic_background" skyColor="0.88 0.95 0.97" groundColor="0.97 0.93 0.88">
      <MetadataString name="background"/>
    </Background>
    <NavigationInfo speed="2.0" type="ANY">
      <MetadataString name="speedmeta"/>
    </NavigationInfo>
    <Viewpoint description="vpSouth" DEF="vpSouth" orientation="0 1 0 0" position="27821.200321 12000.000000 54661.974311 "
       centerOfRotation="27821.200321 0.000000 13665.493578 ">
      <MetadataString name="vpSouth"/>
    </Viewpoint>
```

```
    <Viewpoint description="vpNorth" DEF="vpNorth" orientation="0 1 0 3.14" position="27821.200321 12000.000000 -27330.987156 "
      centerOfRotation="27821.200321 0.000000 13665.493578 ">
      <MetadataString name="vpNorth"/>
    </Viewpoint>
    <Viewpoint description="vpWest" DEF="vpWest" orientation="0 1 0 4.56" position="-27330.987156 12000.000000 13665.493578 "
      centerOfRotation="27821.200321 0.000000 13665.493578 ">
      <MetadataString name="vpWest"/>
    </Viewpoint>
    <Viewpoint description="vpEast" DEF="vpEast" orientation="0 1 0 1.56" position="82973.387797 12000.000000 13665.493578 "
      centerOfRotation="27821.200321 0.000000 13665.493578 ">
      <MetadataString name="vpEast"/>
    </Viewpoint>
    <Viewpoint description="vpAbove" DEF="vpAbove" orientation="-1 0 0 1.57" position="27821.200321 78690.237631 13665.493578 "
      centerOfRotation="27821.200321 0.000000 13665.493578 ">
      <MetadataString name="vpAbove"/>
    </Viewpoint>
    <Transform DEF="compasscross_transform" translation="60279.267362 4636.866720 31967.853876 " scale="4636.866720 4636.866720
      4636.866720 ">
      <Shape DEF="shape_OGC:WMS_horowhenua_ws:study_area">
        <IndexedFaceSet ccw="false" coordIndex="0 1 2 3 -1" texCoordIndex="0 1 2 3 -1" solid="false" colorPerVertex="true"
      normalPerVertex="false" creaseAngle="0.3" DEF="faceset_OGC:WMS_horowhenua_ws:study_area">
          <Coordinate point="0.000000 0.000000 27330.987156 55642.400642 0.000000 27330.987156 55642.400642 0.000000 0.000000
      0.000000 0.000000 0.000000"/>
          <TextureCoordinate point="0 0 0 1 1 1 0 1">
            <MetadataString name="faceset-texture-coordinates"/>
          </TextureCoordinate>
        </IndexedFaceSet>
        <Appearance>
          <Material transparency="0.3" specularColor="0.88 0.95 0.97">
            <MetadataString name="material_OGC:WMS_horowhenua_ws:study_area"/>
          </Material>
          <ImageTexture DEF="texture_OGC:WMS_horowhenua_ws:study_area" url="http://portal.smart-project.info/geoserver/wms?SERVICE=
      WMS&REQUEST=GetMap&VERSION=1.1.1&LAYERS=horowhenua_ws:study_area&FORMAT=image/png&SRS=EPSG:4326&BBOX
      =174.94924926044294,-40.78200189557148,175.60842894781922,-40.53588171238633&TRANSPARENT=true&STYLES=&WIDTH=1920&HEIGHT=945">
            <MetadataString name="horowhenua_ws:study_area"/>
          </ImageTexture>
        </Appearance>
      </Shape>
    <Transform DEF="wfstrans" rotation="1 0 0 3.141593" translation="0 0 27330.987156">
      <Transform translation="25521.951851 1.000000 22942.665909 " scale="20 4.0 20">
        <Shape DEF="">
          <Appearance>
            <Material diffuseColor="1.00 0.00 1.00" transparency="0.0" shininess="0.0">
              <MetadataInteger value="length=0.0m"/>
            </Material>
          </Appearance>
          <Cylinder radius="1.0" height="0.0">
            <MetadataString name=""/>
          </Cylinder>
        </Shape>
      </Transform>
```

Listing 31: Excerpts from X3D scene encoding for the Horowhenua scene

```
// smartportal, package views;

public class WmcTest extends Controller {
   public static Promise<Result> wpsRenderTo3d() {
      ...
      xml = XmlObject.Factory.parse(wmcdoc, xmlOptions);

      if (!(xml instanceof ViewContextDocument) {
         equestBuilder
         .append("<?xml version=\"1.0\" encoding=\"UTF-8\" standalone=\"yes\"?>"
         + "<wps:Execute service=\"WPS\" version=\"1.0.0\""
         + "  xmlns:wps=\"http://www.opengis.net/wps/1.0.0\" xmlns:ows=\"http://www.opengis.net/ows/1.1\""
         + "  xmlns:xlink=\"http://www.w3.org/1999/xlink\" xmlns:xsi=\"http://www.w3.org/2001/XMLSchema-instance\""
         + "  xsi:schemaLocation=\"http://www.opengis.net/wps/1.0.0 "
         + "  http://schemas.opengis.net/wps/1.0.0/wpsExecute_request.xsd\">"
         + "  <ows:Identifier>org.n52.wps.server.algorithm.smart.WmcSupervisionAlgorithm</ows:Identifier>"
         + "  <wps:DataInputs>"
         + "      <wps:Input>"
         + "         <ows:Identifier>wmcDoc</ows:Identifier>"
         + "         <wps:Data>"
         + "            <wps:LiteralData><![CDATA[");

         requestBuilder.append(wmcdoc);

         requestBuilder.append("]]></wps:LiteralData>" + "        </wps:Data>"
         + "      </wps:Input>" + " </wps:DataInputs>"
         + "   <wps:ResponseForm>"
         + "      <wps:RawDataOutput mimeType=\"text/xml\">"
         + "         <ows:Identifier>result</ows:Identifier>"
         + "      </wps:RawDataOutput>" + "  </wps:ResponseForm>"
         + "</wps:Execute>");

         final String wpsRequest = requestBuilder.toString();

         Promise<Result> resultPromise = WS.url(wpsUrl)
         .setContentType("application/xml").post(wpsRequest)
         .map(new Function<WSResponse, Result>() {
         public Result apply(WSResponse response) {

         // InputStream input = response.getBodyAsStream();
         String result = response.getBody();
         return ok(views.html.x3dlegacy.render(result,
         "testing wps rendered x3d insert.", lang,
```

```
        userName));

        }
        });
        return resultPromise;
..
```

Listing 32: Excerpts from `WmcTest`

```
// smartportal, package views;

@(theScene: String, message: String, lang: String = "en", loggedin: String = "guest")

<!-- the x3dom stuff  -->
<link rel="stylesheet" media="screen" href="@routes.Assets.at("stylesheets/x3dom.css")">
<script src="@routes.Assets.at("javascripts/x3dom.js")" type="text/javascript"></script>

<div id="x3dcontainer2" class="x3dcontainer2">
  @Html(theScene)
</div>
..
```

Listing 33: Excerpts from portal views Scala template with the `<div x3dcontainer2>` element that will be extended with the WPS generated X3D document and subsequenly rendered by the x3dom software

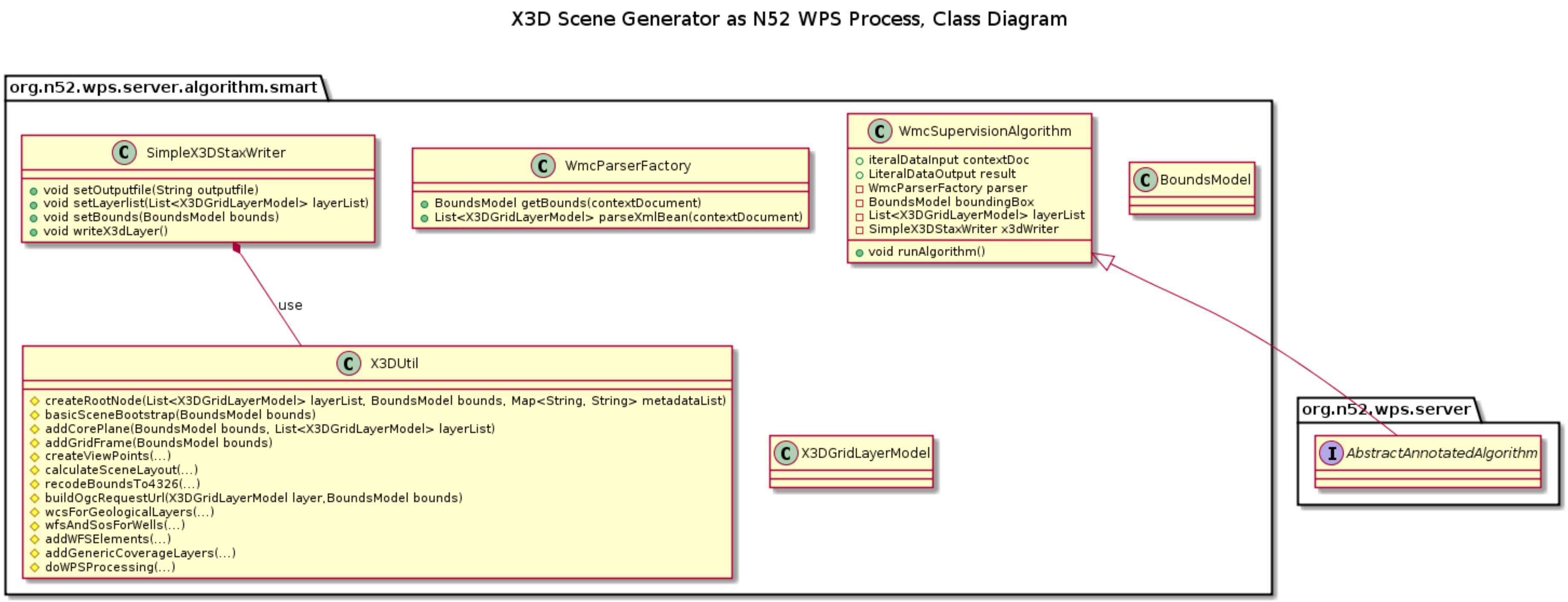


Figure 43: UML Class Diagramm for Implementation of the X3D scene generation as 52N WPS algorithm

E

EARTHVISION XYZ EXPORT TO GEOTIFF TRANSFORM

```java
package info.smart.sensorweb.geopublish;


import static org.junit.Assert.*;

import org.geotools.coverage.Category;
import org.geotools.coverage.CoverageFactoryFinder;
import org.geotools.coverage.GridSampleDimension;
import org.geotools.coverage.grid.GridCoverage2D;
import org.geotools.coverage.grid.GridCoverageBuilder;
import org.geotools.coverage.grid.GridCoverageBuilder.Variable;
import org.geotools.coverage.grid.GridCoverageFactory;
import org.geotools.coverage.grid.io.AbstractGridFormat;
import org.geotools.coverage.grid.io.imageio.GeoToolsWriteParams;
import org.geotools.factory.GeoTools;
import org.geotools.factory.Hints;
import org.geotools.gce.geotiff.GeoTiffFormat;
import org.geotools.gce.geotiff.GeoTiffWriteParams;
import org.geotools.gce.geotiff.GeoTiffWriter;
import org.geotools.geometry.Envelope2D;
import org.geotools.geometry.jts.ReferencedEnvelope;
import org.geotools.referencing.CRS;
import org.geotools.resources.i18n.Vocabulary;
import org.geotools.resources.i18n.VocabularyKeys;
import org.geotools.util.NumberRange;
import org.junit.Test;
import org.opengis.coverage.ColorInterpretation;
import org.opengis.coverage.SampleDimensionType;
import org.opengis.geometry.Envelope;
import org.opengis.parameter.GeneralParameterValue;
import org.opengis.parameter.ParameterValue;
import org.opengis.parameter.ParameterValueGroup;
import org.opengis.referencing.FactoryException;
import org.opengis.referencing.NoSuchAuthorityCodeException;
import org.opengis.referencing.crs.CoordinateReferenceSystem;

import info.smart.sensorweb.geopublish.models.LittlePointy;
import it.geosolutions.geoserver.rest.GeoServerRESTPublisher;

import java.awt.Color;
import java.awt.Graphics2D;
import java.awt.image.BufferedImage;
import java.awt.image.DataBuffer;
import java.awt.image.Raster;
import java.awt.image.WritableRaster;
import java.io.BufferedReader;
import java.io.File;
import java.io.FileInputStream;
import java.io.IOException;
import java.io.InputStream;
import java.io.InputStreamReader;
import java.util.ArrayList;
import java.util.Arrays;
import java.util.HashMap;
import java.util.HashSet;
import java.util.List;
import java.util.Map;
import java.util.Set;

import javax.imageio.ImageIO;
import javax.measure.quantity.Length;
import javax.measure.unit.SI;
import javax.measure.unit.Unit;
import javax.media.jai.PropertySourceImpl;
import javax.media.jai.RasterFactory;

public class WriteTiffFromHoroData {

   @Test
   public void readDatAndReadGeoTiff() throws NumberFormatException, IOException, NoSuchAuthorityCodeException, FactoryException {

      String dat1 = "Greywacke_top_100.dat";
      String dat2 = "Holocene_top_100.dat";
      String dat3 = "Q6_top_100.dat";
      String dat4 = "Q2Q3Q4_top_100.dat";
      String dat5 = "Q5_top_100.dat";

      List<String> datfiles = new ArrayList<String>();
      datfiles.addAll(Arrays.asList(new String[] { dat1, dat2, dat3, dat4, dat5 }));

      for (String dat : datfiles) {

         String line = null;
         File asciigrid = new File(buildFilePath(dat));

         InputStream is = new FileInputStream(asciigrid);
```

```
        InputStreamReader isr = new InputStreamReader(is);
        BufferedReader br = new BufferedReader(isr);

        Set<String[]> fieldset = new HashSet<String[]>();

        List<LittlePointy> lps = new ArrayList<LittlePointy>();

        String s_srs = "EPSG:27200";
        String t_srs = "EPSG:4326";

        double NODATA = -9999;

        double min_x = NODATA;
        double max_x = NODATA;
        double min_y = NODATA;
        double max_y = NODATA;

        double min_elevation = NODATA;
        double max_elevation = NODATA;

        double x_spacing = NODATA;
        double y_spacing = NODATA;

        int x_dim = 0;
        int y_dim = 0;

        double t_x = NODATA;
        double t_y = NODATA;
        double t_x_space = NODATA;
        double t_y_space = NODATA;
        double elevation = NODATA;

        double uncert1 = NODATA;
        double uncert2 = NODATA;

        while ((line = br.readLine()) != null) {
           String[] fields = line.split(" ");
           fieldset.add(fields);
           assertTrue(fields.length == 5);

           LittlePointy lp = new LittlePointy();
           lp.setSrs(s_srs);

           // 2695200 6048000 150 53 1
           for (int i = 0; i < fields.length; i++) {
              String t_string = fields[i];

              if (i==0) {
                 // 20xxxxx is longitude, laenge, x
                 t_x = Double.parseDouble(t_string);
                 lp.setX(t_x);

                 if (t_x > max_x && !(max_x == NODATA)) {
                    max_x = t_x;
                 } else if (max_x == NODATA) {
                    max_x = t_x;
                 }
                 if (t_x < min_x && !(min_x == NODATA)) {
                    min_x = t_x;
                 } else if (min_x == NODATA) {
                    min_x = t_x;
                 }
              }
              if (i==1) {
                 // 65xxxx is latitude, breite, y
                 t_y = Double.parseDouble(t_string);
                 lp.setY(t_y);

                 if (t_y > max_y && !(max_y == NODATA)) {
                    max_y = t_y;
                 }else if (max_y == NODATA) {
                    max_y = t_y;
                 }
                 if (t_y < min_y && !(min_y == NODATA)) {
                    min_y = t_y;
                 } else if (min_y == NODATA) {
                    min_y = t_y;
                 }
              }
              if (i==2) {
                 // elevation relative to sealevel
                 elevation = Double.parseDouble(t_string);
                 lp.setElev(elevation);

                 if (elevation > max_elevation && !(max_elevation == NODATA)) {
                    max_elevation = elevation;
                 }else if (max_elevation == NODATA) {
                    max_elevation = elevation;
                 }
                 if (elevation < min_elevation && !(min_elevation == NODATA)) {
                    min_elevation = elevation;
                 } else if (min_elevation == NODATA) {
                    min_elevation = elevation;
                 }
              }
              if (i==3) {
                 // uncert 1
                 uncert1 = Double.parseDouble(t_string);
                 lp.setUncert1(uncert1);
```

```
               }
               if (i==4) {
                  // uncert 2
                  uncert2 = Double.parseDouble(t_string);
                  lp.setUncert2(uncert2);
               }

            }
            lps.add(lp);
         }

         log(dat + " number of fields: " + fieldset.size());

         // check for smallest x_spacing and y_spacing
         for (LittlePointy lp_outer : lps) {

            for (LittlePointy lp_inner : lps) {

               if (!(lp_outer.getX() == lp_inner.getX())) {

                  t_x_space = Math.abs(lp_outer.getX() - lp_inner.getX());

                  if (t_x_space < x_spacing  && !(x_spacing == NODATA)) {
                     x_spacing = t_x_space;
                  } else if (x_spacing == NODATA) {
                     x_spacing = t_x_space;
                  }
               }

               if (!(lp_outer.getY() == lp_inner.getY())) {

                  t_y_space = Math.abs(lp_outer.getY() - lp_inner.getY());

                  if (t_y_space < y_spacing  && !(y_spacing == NODATA)) {
                     y_spacing = t_y_space;
                  } else if (y_spacing == NODATA) {
                     y_spacing = t_y_space;
                  }
               }
            }
         }

         x_dim = (int) (( max_x - min_x ) / x_spacing);
         y_dim = (int) (( max_y - min_y ) / y_spacing);

         log(String.format("%s min_x %f ,min_y %f , max_x %f , max_y %f, x_spacing %f , y_spacing %f , x_dim %d , y_dim %d, min_elevation %f , max_elevation %f ",
               dat, min_x,min_y, max_x, max_y, x_spacing, y_spacing, x_dim, y_dim, min_elevation, max_elevation));

         // how do we fill the pixel values
         double[] pixelarray = new double[x_dim * y_dim];

         int counter = 0;
         for (int y = 0; y < y_dim; y++) {

            for (int x = 0; x < x_dim; x++) {

               double calcval = getCalcValue(lps, x, y, x_spacing, y_spacing, min_x, min_y, NODATA);
               // int pixelpos = (y*x_dim) + x;
               // now reverse filling, still linewise, from last "line" starting
               int pixelpos = ( (y_dim - 1 - y) *x_dim ) + x;
               pixelarray[pixelpos] = calcval;

               if (calcval != NODATA ) {
                  counter++;
               }
            }
         }
         log(String.format("%s hits overall %d, of known fields %d ", dat, counter, lps.size()));

         // set output file
         String tiffile = dat.replace(".dat", ".tif");
         File output = new File(buildFilePath(tiffile));
         GeoTiffWriter writer = new GeoTiffWriter(output);

           GridCoverage2D coverage = createCovFromRaster(pixelarray, s_srs, min_x, max_x, min_y, max_y, x_dim, y_dim, min_elevation, max_elevation, NODATA, dat);
         assertNotNull(coverage);
         writer.write(coverage, null);
         coverage.dispose(true);
           assertTrue(output.canRead());
         // TODO beware test
           // assertTrue(publishGeoTIFF(tiffile));

//         break;
      }
   }

   /**
    *
    * @param pixelfloat
    * @param s_srs
    * @param min_x
    * @param max_x
    * @param min_y
    * @param max_y
    * @param x_dim
    * @param y_dim
    * @param min_elevation
```

```java
     * @param max_elevation
     * @param NODATA
     * @param datname
     * @return
     * @throws NoSuchAuthorityCodeException
     * @throws FactoryException
     */
    public GridCoverage2D createCovFromRaster(double[] pixelarray, String s_srs, double min_x, double max_x, double min_y, double
        max_y,
           int x_dim, int y_dim, double min_elevation, double max_elevation,
           double NODATA, String datname) throws NoSuchAuthorityCodeException, FactoryException {

       WritableRaster raster = RasterFactory.createBandedRaster(DataBuffer.TYPE_FLOAT, x_dim, y_dim, 1, null);
       for (int y=0; y<y_dim; y++) {
            for (int x=0; x<x_dim; x++) {
                // We exploit the clamping capabilities of the sample model.
             int pixelpos = (y*x_dim) + x;
                raster.setSample(x, y, 0, (float) pixelarray[pixelpos]);
            }
        }
       // GeoTools.getDefaultHints()
         GridCoverageFactory factory = CoverageFactoryFinder.getGridCoverageFactory(GeoTools.getDefaultHints());

       CoordinateReferenceSystem crs = CRS.decode(s_srs, true);
       ReferencedEnvelope env = new ReferencedEnvelope(min_x, max_x, min_y, max_y, crs);

       writeRefPNG(x_dim, y_dim, raster, datname.replace(".dat", "_rasterfloat.dat"));

       Category nan = new Category(Vocabulary.formatInternational(VocabularyKeys.NODATA), new Color(0, 0, 0, 0), 0);
       Category values = new Category("Elevation", new Color[] { new Color(255, 0, 0, 0) }, NumberRange.create(1,255), NumberRange.
        create((int) min_elevation, (int ) max_elevation));

       final Unit<Length> uom = SI.METER;
       final GridSampleDimension band = new GridSampleDimension(datname.replace(".dat", "_rasterfloat"), new Category[] { nan, values
        }, uom).geophysics(true);
       final Map<String, Double> properties = new HashMap<String, Double>();
       properties.put("GC_NODATA", new Double(NODATA));

       // GridCoverage2D coverage = factory.create(datname.replace(".dat", "_rasterfloat"), raster, env );
       // GridCoverage2D coverage = factory.create(datname.replace(".dat", "_rasterfloat"), raster, env, new GridSampleDimension[] {
        band }, null, properties);
       GridCoverage2D coverage = factory.create(datname.replace(".dat", "_rasterfloat"), raster, env, new GridSampleDimension[] { band
         });

       return coverage;
    }

    // test only
    public void writeRefPNG(int x_dim, int y_dim, WritableRaster raster, String datname) {

       BufferedImage image = new BufferedImage(x_dim, y_dim, BufferedImage.TYPE_BYTE_GRAY);
       image.setData(raster);
       try {
          ImageIO.write(image, "png", new File(datname.replace("dat", "png")));
       } catch (IOException e) {
          // TODO Auto-generated catch block
          e.printStackTrace();
       }
    }

    /**
     *
     * @param lps
     * @param x
     * @param y
     * @param x_spacing
     * @param y_spacing
     * @param min_x
     * @param min_y
     * @param NODATA
     * @return
     */
    private double getCalcValue(List<LittlePointy> lps, int x, int y, double x_spacing, double y_spacing, double min_x, double min_y,
        double NODATA) {
       double x_coord = (x * x_spacing) + min_x;
       double y_coord = (y * y_spacing) + min_y;

       for (LittlePointy lp : lps) {
          if (x_coord == lp.getX() && y_coord == lp.getY()) {
             // log(String.format("hit x_coord %f y_coord %f lp.getElev() %f", x_coord , y_coord , lp.getElev()) );
             return lp.getElev();
          }
       }
       return NODATA;
    }

    /**
     *
     * @param fileName
     * @return
     */
    public String buildFilePath(String fileName) {

       String workingdir = null;
       String new_workingdir = null;

       String os = System.getProperty("os.name").toLowerCase();
       String pathdelim = "/";
       String currentDir = System.getProperty("user.dir");
```

```
        if (os.indexOf("win") >= 0) {

            pathdelim = "\\";
            workingdir = new File(currentDir).getPath();
            new_workingdir = workingdir + pathdelim + "misc";

        } else if (os.indexOf("nix") >= 0 || os.indexOf("nux") >= 0) {

            pathdelim = "/";
            workingdir = new File(currentDir).getPath();
            new_workingdir = workingdir + pathdelim + "misc";
        }

        return new_workingdir + pathdelim + fileName;
    }

    private void log(String msg) {
        System.out.println(msg);
    }
}
```

Listing 34: Excerpts from EarthVision to GeoTIFF transform program