Improving access to groundwater data using GroundWaterML2

Bruce Simons
Information Modeller, CSIRO Land and Water, Clayton, Australia
E-mail: bruce.simons@csiro.au

Eloise Nation
Senior Hydrogeologist, Groundwater Unit, Water Resources Assessment, Bureau of Meteorology, Melbourne, Australia
E-mail: e.nation@bom.gov.au

Peter Dahlhaus
Senior Research Fellow, Centre for eResearch and Digital Innovation, Federation University Australia, Ballarat, Australia
E-mail: p.dahlhaus@federation.edu.au

Abstract

This paper describes the preliminary development and use in the Australian context of a standard designed to exchange groundwater related data. The storage and management of groundwater data is distributed across many agencies and organisations, in disparate databases and formats. Discovering, accessing, interpreting, reformatting and using this data can present considerable challenges for the end-user. Groundwater data interoperability consideration of the use of communication protocols to achieve technical interoperability, the use of common data models to achieve syntactic interoperability and the use of controlled vocabularies to achieve semantic interoperability.

GroundWaterML2 is a Geography Mark-up Language (GML) application developed by the Open Geospatial Consortium (OGC) Hydro Domain Working Group. It is intended as a standard for the transfer of groundwater feature data, including data about water wells, aquifers, and related entities. The OGC initiated an interoperability experiment to develop and test the model for commercial, technical, scientific, environmental and policy use cases. CSIRO, the Bureau of Meteorology and Federation University Australia contributed to the design of GroundWaterML2, and established separate OGC web services delivering data out of the National Groundwater Information System (NGIS) and Federation University Australia databases. These services delivered borehole location and construction details, downhole geology, hydrogeologic unit information, groundwater discharge properties, and groundwater fluid property observations.

Bringing these services to production would allow users and clients, such as the 'Visualising Victoria's Groundwater' and 'Australian Groundwater Explorer' portals, to access data from multiple providers in a standard format.

Keywords: groundwater, groundwaterml2, hydrogeology, interoperability, open geospatial consortium, standards, mark-up language.

1. INTRODUCTION

The expanding global demand for groundwater to supply human consumption, energy and food production has led to groundwater resource overexploitation (Gorelick and Zheng, 2015) with corresponding threats to environmental and ecological values (e.g. Nevill et al., 2010) and the sustainability of food production (e.g. Scanlon et al., 2012). These concerns, together with the predicted impacts of changing climates (Taylor et al., 2013) and the potential impacts of unconventional energy exploitation (AWA, 2012) have resulted in an urgent need for more globally accessible groundwater data. The need arises from the fact that groundwater is a vital resource, but hidden from view. Hence the nature of groundwater is generally misunderstood by most of the community and often the subject of myths (Price, 1996). Unlocking the potential of groundwater for human uses increasingly requires the social license to do so (e.g. Fernandez et al., 2014, Gill et al., 2014). Providing current data from authoritative and trusted sources together with the tools that allow groundwater systems to be visualised in 3D, and see the groundwater systems' responses to inputs and outputs would provide a powerful instrument to educate the broader community on the nature and function of their groundwater systems.

The current era of Big Data (e.g. Boyd and Crawford, 2012, Mayer-Schonberger and Cukier, 2013) provides unprecedented opportunities for sharing of groundwater data across national and international borders, and between government, industry and academia, to gain a deeper understanding and appreciation of our global environments, including hydrogeological environments. Data availability has vastly improved as governments in many countries adopt open data policies, acknowledging that data provision is essential for a wide range of environmental, economic and agricultural activities (Zuiderwijk and Janssen, 2014).

Yet despite this unprecedented availability of data, limitations remain on how to access and use the data to best develop water management policy and further the understanding of groundwater science (e.g. Loch et al., 2014). Part of the problem is that storage and management of groundwater data is distributed across many organisations, in disparate databases and formats. A comparison of two, reasonably similar groundwater databases from Federation University Australia and the Bureau's National Groundwater Information System (NGIS) is indicative of the kinds of problems faced by users, even for concepts as well-defined as boreholes (figure 1). Discovering, accessing, interpreting, reformatting and using this data presents considerable challenges for the end-user. In Australia for example, information and data on groundwater is distributed via various web-portals, web-based GIS tools, password protected sites, cloud storage, portable storage devices; hardcopy maps, theses, reports, newsletters, documents, videos and podcasts. This impressive resource of data, information and knowledge is largely ignored simply because most users do not have the knowledge, capability or desire to deal with the data access and use issues. The challenge of centralizing or federating diverse datasets is beyond the capacity or scope of any single organisation.

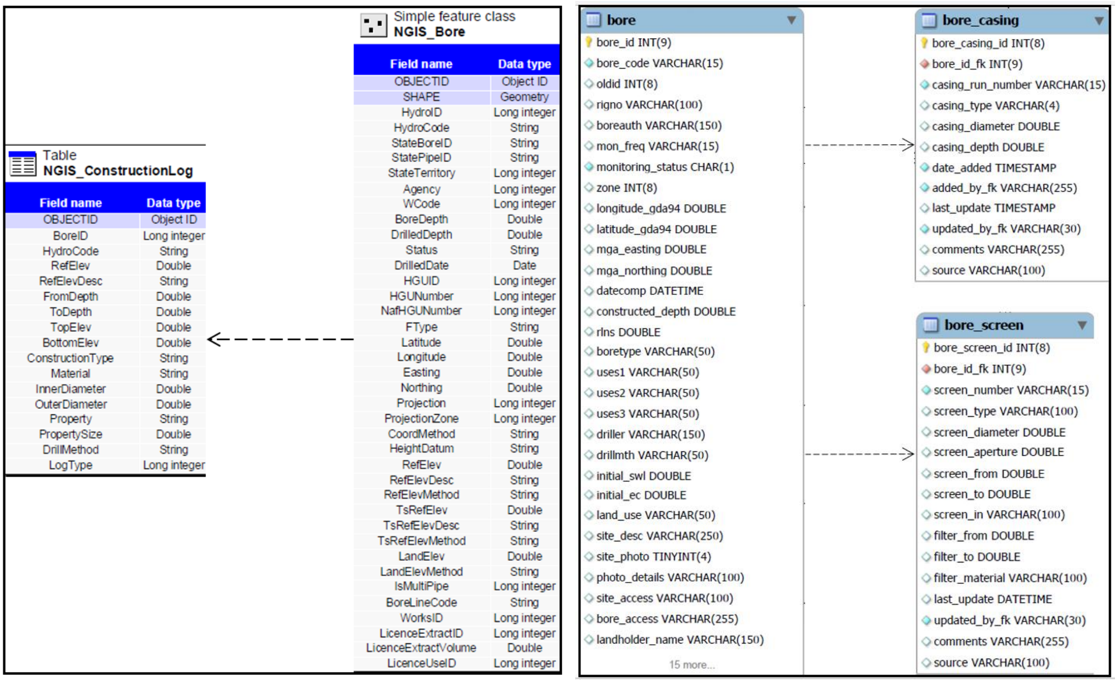

Figure 1 Database schema for borehole features from Australia's National Groundwater Information System (l) and Federation University Australia (r)

The development and deployment of web services built on a Spatial Data Infrastructure (SDI) to federate groundwater data from disparate database sources offers a solution. This interoperability provides the capacity to transfer and use information in a uniform and efficient manner across multiple organisations and information technology systems. It encompasses the use of communication protocols to achieve technical interoperability, the use of common data models to achieve schematic interoperability and the use of controlled vocabularies to achieve semantic interoperability (Brodaric & Gahegan 2006).

The global exemplar groundwater implementations of this approach are the Canadian ‘Groundwater Information Network’ (Boisvert and Brodaric, 2012); the European Commission's INSPIRE specifications (INSPIRE, 2013a, 2013b); the United States ‘National Ground-Water Monitoring Network’ (ACWI, 2013); ground water aspects of the French ‘Water Information System’ (BDLISA, 2013, 2014, ADES, 2015), the New Zealand SMART system (Klug and Kmoch, 2014); and the Australian ‘Visualising Victoria's Groundwater’ portal (Dahlhaus et al., 2012, Dahlhaus et al., 2015).

2. GROUNDWATER DATA INTEROPERABILITY

Since the early 1990s the development of open source software and open standards has been embraced by the academic, research and government sectors (Weber, 2005). The Open Geospatial Consortium (OGC) (www.opengeospatial.org) was launched in 1994 by the open source community to guide the development of standards, rules, decision making procedures and sanctioning mechanisms, and now comprises over 500 government, academic and private sector organisations (OGC, 2015). Raster data is usually delivered using OGC standard web mapping services (WMS). Groundwater vector, feature property and observational data, on the other hand, is commonly accessed via a web interface built to bespoke requirements over organization-specific databases, or simplified and delivered using simple web feature services (WFS). These interoperable groundwater data systems are usually based on SDI software conforming to the OGC standards and developed and distributed by the Open Source Geospatial Foundation (www.osgeo.org).

At its core, seamless information exchange of a domain as complex as groundwater relies on schematic interoperability; that is agreed formats, communication protocols and schemas for serving, querying and consuming data, along with agreed content (semantic interoperability). The Canadian Geological Survey developed a groundwater markup language—GroundWaterML1 (Boisvert and Brodaric, 2007, 2008, 2012) as a foundation for using WFS to deliver the full range of groundwater data commonly stored in relational databases. GroundWaterML1 is based on the geography markup language (GML), an XML grammar defined by the OGC to express geographic features.

2.1. The Groundwater Interoperability Experiment (GW2IE)

In recognition of the need for an internationally agreed groundwater data exchange standard, the OGC Hydro Domain Working Group implemented an international interoperability experiment (GW2IE) in 2012 to develop and test a standard for groundwater data exchange (OGC 2013). The imperative for the development of the Groundwater Mark-up Language (GroundWaterML2) was the need to harmonise and advance existing groundwater data modelling initiatives, such as the Canadian-USA developed GroundWaterML1 (Boisvert & Brodaric 2012), the European Union INSPIRE geology/hydrogeology package (INSPIRE 2013a) and environmental monitoring facilities (INSPIRE 2013b) models, and the IUGS-CGI GeoSciML geology model (Simons et al. 2006). GroundWaterML2 is intended as the authoritative international standard for the transfer of groundwater feature data, including data about water wells, aquifers, groundwater fluids and related entities.

The collaborators in the interoperability experiment were the Geological Surveys of Canada, USA, Germany, UK, France and Poland, as well as the European Commission, CSIRO and the Bureau of Meteorology (Australia), the International Groundwater Resources Assessment Centre (UNESCO), Salzburg University (Austria) and Federation University Australia.

The methodology adopted for GW2IE was:

- Specification of the use-cases that the exchange standard would meet;
- Comparison of existing groundwater information models;
- Development of a unified conceptual model, documented in UML;
- Development of a GML-compliant logical model, documented in UML;

- Generation of a GML-XML (XSD) schema;
- Documentation of constraints not contained in the XSD, such as content;
- Development of example GML-XML instance documents, based on actual data;
- Evaluation of the instance documents against the use-cases;
- Deployment of WFS, SOS, and WMS services utilizing GWML2; and
- Documentation of the GW2IE results as an OGC engineering report.

The GW2IE commenced in October 2012 and aimed to be completed by October 2015 with delivery of an engineering report to the OGC for public discussion. Follow-up work is intended to take the GW2IE results to a working standard for international groundwater data exchange.

For the interoperability experiment five use-cases were specified as typical data exchange scenarios:

(1) A commercial use-case focused on drilling water wells with knowledge of aquifers;
(2) A policy use case concerned with the management of groundwater resources;
(3) An environmental use-case that considers the role of groundwater in natural ecosystems;
(4) A scientific use-case concerned with modelling groundwater systems; and
(5) A technologic use-case concerned with interoperability between diverse information systems and associated data formats.

2.2. GroundWaterML2

The proposed GroundWaterML2 standard consists of three related objects: the *conceptual model*—a technology-neutral UML representation of the semantics of the groundwater domain; the *logical model*—a GML-specific schema that incorporates the OGC suite of standards; and the *XML schema*—an XML encoding of the logical model.

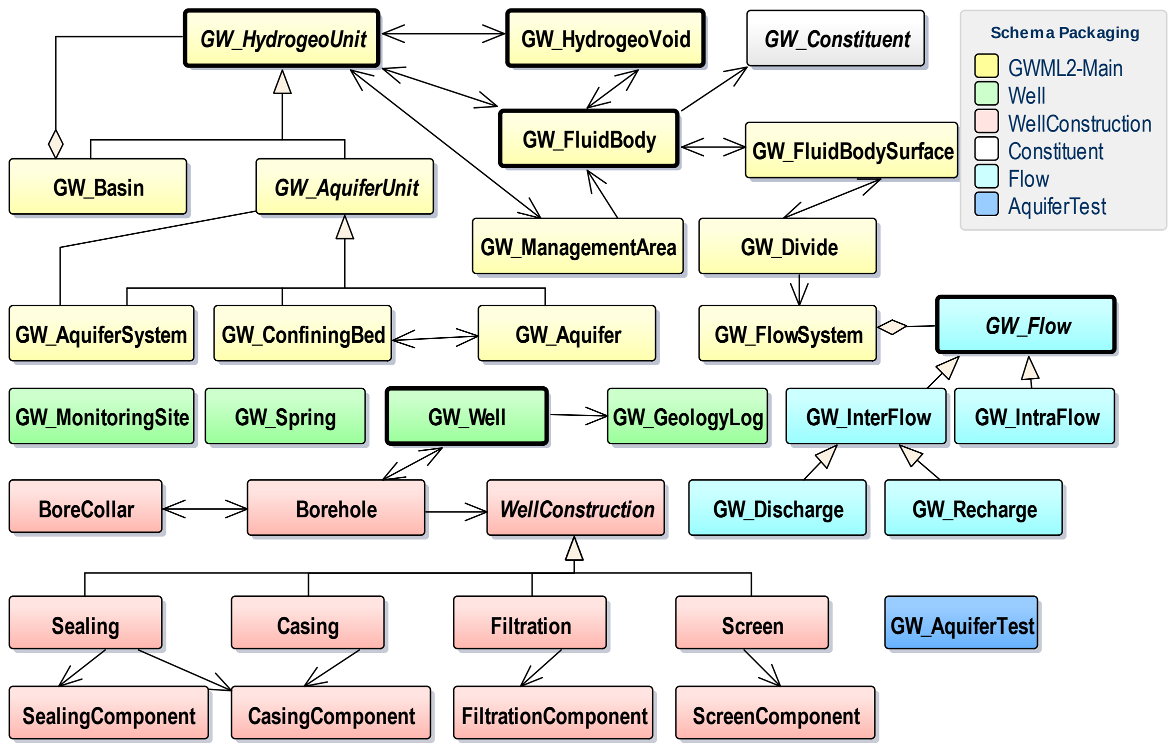

Figure 2 Simplified UML representation of the GroundWaterML2 conceptual model and associations. (Colouring represents the application schema packaging, italicised labels are abstract classes, and heavy outlines the main conceptual components).

The GroundWaterML2 conceptual model design identified the concepts required to meet the GW2IE use cases, as well as other entities in existing standards considered fundamental to the groundwater domain (Figure 2). The conceptual model consists of five main components, and related properties and entities: hydrogeological units, fluid bodies, voids, groundwater flow, and wells. These entities identify the fluid container (a unit or its materials), the fluid itself (fluid body), the spaces in the container potentially occupied by the fluid (voids), the flow of fluid within and between containers and their spaces (flow), and the natural and artificial artefacts used to withdraw, inject, or monitor fluid with respect to a container (wells, springs, monitoring sites) (OGC 2015a).

The logical model introduces OGC General Reference Model artefacts, such as classes, relations, properties, and organizing principles, to the conceptual model. It extends or integrates with other OGC GML standards, such as MD_Metadata, Observations and Measurements, Sensor Web Enablement and GeoSciML. The logical model design focus is on the **feature types**; that is, the packets of information that will be transferred between systems in the application schema.

The GroundWaterML2 logical model is organised into 6 modular packages (Figure 3):

(1) **GWML2-Main**: core elements such as aquifers, their pores, and fluid bodies;
(2) **Constituent**: the biologic, chemical, and material elements of a fluid body;
(3) **Flow**: groundwater flow within and between containers;
(4) **Well**: water wells, springs, and monitoring sites;
(5) **WellConstruction**: the components used to construct a well; and
(6) **AquiferTest**: pump testing results (independent of all other packages).

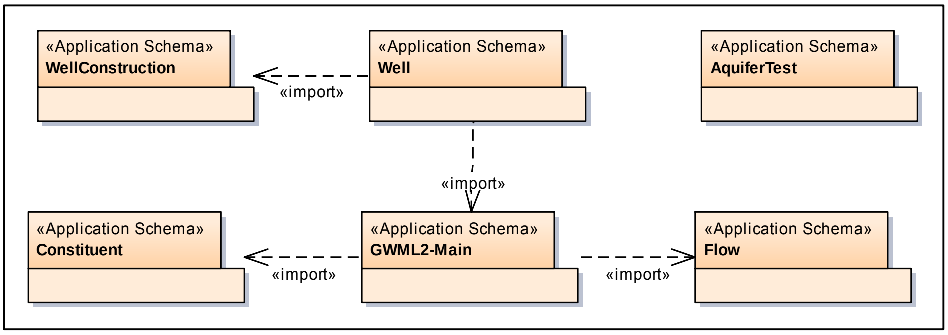

Figure 3 Dependencies between the GroundWaterML2 «Application Schema» packages.

The third component of the standard is the XML schema that specifies the encoding rules for correct machine interpretation of the data. In addition to groundwater specific concepts and properties, the specification builds on the following standards for encoding XML data:

- Observations & Measurements (OMXML);
- sweCommon (OGC 08-094r1);
- GML ISO 19136:2007 (OGC 07-036);
- ISO 19139 (Metadata);
- W3C XSD; and
- GeoSciML 3.2.

It accomplishes this by (a) extending these schema with groundwater specializations, (b) referring to a class in these schema to type a named property, or (c) using a class from the schemas as one of the participants in a relationship.

3. GROUNDWATER WEB SERVICES

In addition to contributing to the design and documentation of GW2IE, CSIRO, the Bureau of Meteorology and Federation University Australia established OGC-compliant GroundWaterML2 Web Feature Services using Australian groundwater data to test the model.

The process of establishing the test services entailed:

- Using UML and commented instructions in XML instance documents to map the National Groundwater Information System (NGIS) data hosted by the Bureau, and the Victorian groundwater data hosted by Federation University Australia, to GroundWaterML2;
- Designing ‘views’ of the databases that as closely as possible matched the target

GroundWaterML2 schema and implementing them as materialized views or replicated databases; and

- Configuring the open source software GeoServer (geoserver.org/) with application schema to deliver GroundWaterML2 compliant Web Feature Services.

The test services (http://external.opengis.org/twiki_public/HydrologyDWG/UseCaseWebServices) deliver ‘GW_Aquifer’, ‘GW_AquiferSystem’, ‘GW_Well’ and ‘Borehole’ GroundWaterML2 features. The Australian GW2IE Web Feature Services were established to demonstrate that data from different databases could use GroundWaterML2 to deliver different datasets in the same format (Figure 4). These services were deployed in a development environment and are not configured for performance or reliability.

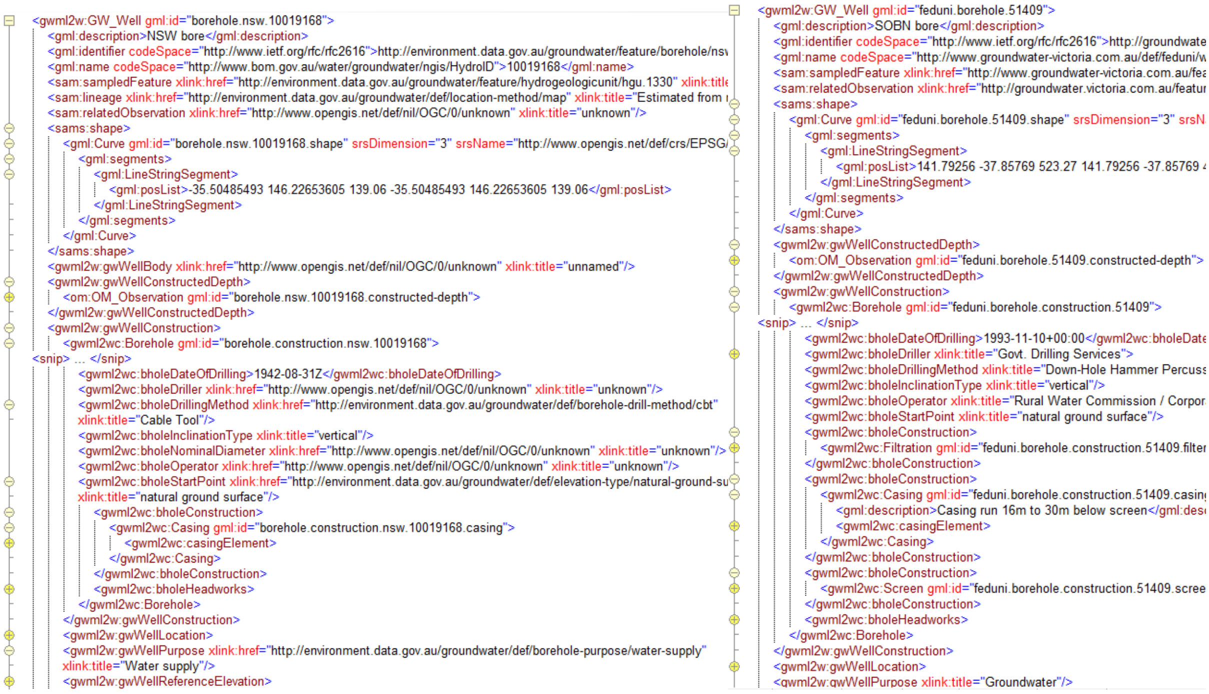

Figure 4 GWML2 XML well data responses from groundwater web feature services from NGIS (l) and Federation University Australia (r) databases. XML tags with ‘+’ have been compressed.

4. FUTURE WORK

Further GroundWaterML2 test services are planned to test the flow use case, which aims to capture data about groundwater flow within and between ‘containers’ (such as aquifers and groundwater dependent ecosystems). The performance of the existing services needs to be improved, and a detailed comparison of the different services each GWML2IE organization has established.

The GWML2IE Engineering Report needs to be finalized and presented for public discussion and feedback, and an OGC Standards Working Group established to take GroundWaterML2 through the formal standards process. Any changes made to the proposed standard as a result will need to be accommodated in the current services before the “Visualising Victoria’s Groundwater” or the “Australian Groundwater Explorer” web-clients can be modified to use the services.

5. CONCLUSIONS

The OGC Groundwater Interoperability Experiment generated a proposed data exchange standard, GroundWaterML2. The involvement of domain and informatics experts from a number of international groundwater and research institutes has meant that it includes the fundamental groundwater concepts required to meet the primary use cases associated with groundwater discovery, management and use. In Australia, the capacity for GroundWaterML2 to deliver required features and associated properties was tested using both national and regional data from two different data providers. Demonstration OGC Web Feature Services were established delivering data from these databases in a common format.

For the most part GroundWaterML2 allowed the delivery of the required data contained in the databases. Minor issues due to limitations with GroundWaterML2 or its associated dependencies were documented in the associated engineering report, and are expected to be addressed as part of an ongoing OGC activity.

6. ACKNOWLEDGMENTS

The GroundWaterML2 design work was conducted in collaboration with the participants of the OGC GW2IE: NRCan (Boyan Brodaric, Eric Boisvert, Francois Létourneau), BRGM (Sylvain Grellet, Santiago Gabillard, Laurence Chery), USGS (Jessica Lucido), GSG (Bernhard Wagner), and GNSS (Alex Kmoch). The support in setting up the Australian test Web Feature Services by Andrew Macleod at Federation University Australia and Jonathan Yu at CSIRO is acknowledged. Funding for CSIRO and Bureau of Meteorology work is provided by Water Information Research and Development Alliance (WIRADA), and funding for Federation University Australia work by the 'Centre for eResearch and Digital Innovation' and the 'Victorian Broadband Enabled Innovation Program'.

7. REFERENCES

ACWI (2013). *A National Framework for groundwater monitoring in the United States.* Report prepared by the Subcommittee on Ground Water of The Advisory Committee on Water Information, Reston, Virginia, USA. 182p. from: http://acwi.gov/sogw/ngwmn_framework_report_july2013.pdf.

ADES (2015). *Portail national d'access aux données sur les eaux souterraines*. http://www.ades.eaufrance.fr/Viewer.aspx (accessed 13 July 2015)

AWA (2012). *Coal seam gas mining and groundwater: a review of the controversies around and potential impacts of CSG in Australia.* Water: Journal of the Australian Water Association, 39(7), 46-48.

BDLISA (2013). *Dictionnaire des données – Référentiel hydrogéologique (Version 2).* http://www.sandre.eaufrance.fr/urn.php?urn=urn:sandre:dictionnaire:SAQ:FRA:::ressource:2.0:::pdf.

BDLISA (2014). *Base de Données des Limites des Systèmes Aquifères.* http://www.reseau.eaufrance.fr/geotraitements/viewer/bdlisa (accessed 13 July 2015).

Boisvert, E. and Brodaric, B. (2007). *GroundWater Markup Language (GWML): Extending GeoSciML for groundwater*. American Geophysical Union, Fall Meeting 2007, Dec 2007, Abstract #IN53C-03

Boisvert, E. and Brodaric, B. (2008). *GroundWater Markup Language Specification v. 1.0.* Natural Resources Canada. Retrieved 23 July 2010, from: http://ngwd-bdnes.cits.rncan.gc.ca/service/api_ngwds/en/gwml.html. Last updated: 30-12-2008.

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

Boyd, D. and Crawford, K. (2012). *Critical questions for big data.* Information, Communication & Society, 15(5), 662-679.

Brodaric, B. and Gahegan, M. (2006). *Representing geoscientific knowledge in cyberinfrastructure: Some challenges and implementations.* in Sinha, A.K., ed., *Geoinformatics: Data to Knowledge:* Geological Society of America Special Paper 397, p. 1-20, doi: 10.1130/2006.2397(01).

Dahlhaus, P.G., MacLeod, A. and Thompson, H. (2012). *An interoperable federation of hydrogeological research data for Victoria, Australia.* in *Confronting Global Change. 39th IAH Congress*, Niagara Falls, Canada, 16-21 September 2012. (Abstract)

Dahlhaus, P.G., Murphy, A., MacLeod, A., Thompson, H., McKenna, K. and Ollerenshaw, A. (2015). *Making the invisible visible: assessing the impact of federating groundwater data in Victoria, Australia.* Journal of Hydroinformatics, (in press).

Fernandez, S., Bouleau, G. and Treyer, S. (2014). *Bringing politics back into water planning scenarios in Europe.* Journal of Hydrology, 518(part A), 17-27.

Gill, B.C., Webb, J., Wilkinson, R. and Cherry, D. (2014). *Irrigator responses to groundwater resource management in northern Victoria, southeastern Australia.* Journal of Hydrology, 518(part A), 83-93.

Gorelick, S.M. and Zheng, C. (2015). *Global change and the groundwater management challenge.* Water Resources Research, (in press - available online first).

INSPIRE (2013a). *D2.8.II.4 INSPIRE Data Specification on Geology – Technical Guidelines*, D2.8.II.4_v3.0, http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_ER_v3.0.pdf (accessed June 29, 2015)

INSPIRE (2013b). *D2.8.III.7 INSPIRE Data Specification on Environmental monitoring Facilities – Draft Technical Guidelines*, D2.8.III.7_v3.0rc3, http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_EF_v3.0.pdf (accessed June 29, 2015).

Klug, H. and 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, 78-86.

Loch, A., Adamson, D. and Mallawaarachchi, T. (2014). *Role of hydrology and economics in water management policy under increasing uncertainty.* Journal of Hydrology, 518(part A), 5-16.

Mayer-Schonberger, V. and Cukier, K. (2013). *Big Data: A revolution that will transform how we live, work and think*. Houghton Mifflin Harcourt publishers, New York. 242p.

Nevill, J.C., Hancock, P.J., Murray, B.R., Ponder, W.F., Humphreys, W.F., Phillips, M.L. and Groom, P.K. (2010). *Groundwater-dependent Ecosystems and the Dangers of Groundwater Overdraft: A Review and An Australian Perspective.* Pacific Conservation Biology, 16(3), 187-208.

OGC (2013). *Hydrology Domain Working Group – GroundWater Interoperability Experiment 2.* OpenGIS Project Document 12-126: Activity Plan for an OGC Interoperability Experiment.

OGC (2015a). *OGC GroundWaterML 2.0.* OGC Implementation Standard (ed. Boyan Brodaric). Open Geospatial Consortium, Hydrology Domain Working Group. 177p. from: http://external.opengeospatial.org/twiki_public/pub/HydrologyDWG/GroundwaterInteroperabilityExperiment2/GroundWaterML2_v0.3.1_DRAFT.pdf.

OGC (2015b). "*OGC History (abbreviated).*" Open Geospatial Consortium. Retrieved 18/6/2015, from: http://www.opengeospatial.org/ogc/history.

Price, M. (1996). *Introducing Groundwater*. Nelson Thornes Ltd, Cheltenham, U.K. . 2nd ed. 289p.

Scanlon, B.R., Faunt, C.C., Longuevergne, L., Reedy, R.C., Alley, W.M., McGuire, V.L. and McMahon, P.B. (2012). *Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley.* Proceedings of the National Academy of Sciences of the United States of America, 109(24), 9320-9325.

Simons, B.A., Boisvert, E., Brodaric, B., Cox, S., Duffy, T. R., Johnson, B.R., Laxton, J.L. and Richard, S.M. (2006). *GeoSciML: Enabling the Exchange of Geological Map Data*. Australian Earth Sciences Convention, Melbourne, Australia.

Taylor, R.G., Scanlon, B., Doll, P., Rodell, M., van Beek, R., Wada, Y., Longuevergne, L., Leblanc, M., Famiglietti, J.S., Edmunds, M., Konikow, L., Green, T.R., Chen, J., Taniguchi, M., Bierkens, M.F.P., MacDonald, A., Fan, Y., Maxwell, R.M., Yechieli, Y., Gurdak, J.J., Allen, D.M., Shamsudduha, M., Hiscock, K., Yeh, P.J.F., Holman, I. and Treidel, H. (2013). *Ground water and climate change.* Nature Climate Change, 3(4), 322-329.

Uslander, T. (2005). *Trends of environmental information systems in the context of the European Water Framework Directive.* Environmental Modelling & Software, 20(12), 1532-1542.

Weber, S. (2005). *The Success of Open Source*. Harvard University Press, Cambridge, MA, USA. 320p.

Zuiderwijk, A. and Janssen, M. (2014). *Open data policies, their implementation and impact: A framework for comparison.* Government Information Quarterly, 31(1), 17-29.