



Enabling global exchange of groundwater data: GroundWaterML2 (GWML2)

Boyan Brodaric¹ · Eric Boisvert² · Laurence Chery³ · Peter Dahlhaus⁴ · Sylvain Grellet³ · Alexander Knoch⁵ · François Létourneau² · Jessica Lucido⁶ · Bruce Simons⁷ · Bernhard Wagner⁸

Received: 12 February 2017 / Accepted: 16 February 2018 / Published online: 21 March 2018
© The Author(s) 2018

Abstract

GWML2 is an international standard for the online exchange of groundwater data that addresses the problem of data heterogeneity. This problem makes groundwater data hard to find and use because the data are diversely structured and fragmented into numerous data silos. Overcoming data heterogeneity requires a common data format; however, until the development of GWML2, an appropriate international standard has been lacking. GWML2 represents key hydrogeological entities such as aquifers and water wells, as well as related measurements and groundwater flows. It is developed and tested by an international consortium of groundwater data providers from North America, Europe, and Australasia, and facilitates many forms of data exchange, information representation, and the development of online web portals and tools.

Keywords General hydrogeology · Geographic information systems · GroundWaterML2 · Groundwater data standards

Introduction

Opportunities have never been greater for sharing hydrogeological data across national and international borders, and across governmental, industry and academic sectors. These opportunities are driven primarily by a pressing need to better understand our natural environment at all scales, especially continental and global (e.g. Boyd and Crawford 2012; Mayer-Schonberger and Cukier 2013). Major challenges include meeting the expanding demand for water resources associated with human consumption, food production and energy resources

(Gorelick and Zheng 2015; Scanlon et al. 2012), the potential impact of unconventional energy exploitation (e.g. AWA 2012; Vidic et al. 2013; Vengosh et al. 2014), the predicted impacts of changing climate (e.g. Taylor et al. 2013), as well as other ecological and environmental impacts (e.g. Nevill et al. 2010).

The growth of hydrogeological data is another significant driver, magnified by its much greater availability due to technological advances and governmental open data policies (Zuiderwijk and Janssen 2014). However, collecting more data and increasing its availability is only part of the solution to the major challenges, as there are limitations in how to transform this data into improved understanding of hydrogeological science and water management (e.g. Loch et al. 2014). These limitations are largely due to the distributed custodianship and heterogeneity of groundwater data, making it difficult for the end-user to discover, access and harmonize the data. Although there may be an impressive volume of global groundwater data available, few users have the capacity or desire to overcome these difficulties.

An additional part of the solution is making disparate data interoperable—that is, making the data usable in a seamless manner regardless of its original heterogeneity. Data interoperability typically requires the transformation of data into a common representation readily understood by both the data supplier and consumer (Brodaric et al. 2016). Often this representation takes the form of a data standard developed and

✉ Boyan Brodaric
boyan.brodaric@canada.ca

¹ Geological Survey of Canada, Canada (GSC), Ottawa, Canada
² Geological Survey of Canada (GSC), Quebec City, Canada
³ Bureau de Recherches Géologiques et Minières (BRGM), Orléans, France
⁴ Federation University Australia (FedUni), Mount Helen, Australia
⁵ University of Salzburg (Z_GIS), Salzburg, Austria
⁶ U.S. Geological Survey (USGS), Reston, VA, USA
⁷ CSIRO Land and Water, Canberra, Australia
⁸ Bavarian Environment Agency (LfU), Hof, Germany

maintained by key stakeholders, one that is typically used within two main technological environments: (1) within data networks that supply heterogeneous data to users, or (2) within workflow systems that chain together diverse software for modularized processing of data. While online groundwater workflow systems are somewhat nascent (e.g. Klug and Kmoch 2014), web-based groundwater data networks have become more prevalent. Such networks consist of distributed data sources that are managed autonomously and that possess heterogeneous data structure and content. Government initiated examples of such groundwater data networks include the Canadian ‘Groundwater Information Network’ (Boisvert and Brodaric 2012; Brodaric et al. 2016); the United States ‘National Ground-Water Monitoring Network’ (Blodgett et al. 2016; Brodaric et al. 2016); groundwater aspects of the French ‘Water Information System’ (BDLISA 2013, 2014), the New Zealand SMART system (Klug and Kmoch 2014; Kmoch et al. 2015); and the Australian ‘National Groundwater Information System’ (BOM 2015; Carrara et al. 2015). Outside of government, the Australian ‘Visualizing Victoria’s Groundwater’ portal (Dahlhaus et al. 2012, 2016) is an interoperable data network developed by university researchers. It is noteworthy that the growth of such data networks is not limited to groundwater science, but echoes a larger trend in the sciences generally and in hydrology particularly (Brodaric and Piasecki 2016).

A consequence of the rise of water data networks is a growing imperative for their connection, which recognizes the need to build networks of networks to supply data for large-area science problems (e.g. Nativi et al. 2014). As this need becomes more evident, particularly across political boundaries, the absence of an international groundwater data representation becomes more acute. In this paper, this gap is addressed through the establishment of GroundWaterML2 (GWML2), a new global standard for groundwater data. GWML2 has been developed by a group of international groundwater data providers collaborating within the Groundwater Standards Working Group of the Open Geospatial Consortium (OGC). It comprises the groundwater component of the WaterML2 suite of hydrologic data standards and best practices (WaterML2 2017), and consists of data structures and encoding guidelines for groundwater data. GWML2 has been successfully implemented and tested by a variety of major groundwater data providers from North America, Europe, and Australasia, and it is expected to complement emerging data standards for both surface water and atmospheric water to facilitate data exchange for the complete water cycle.

This paper describes the motivation behind the development of GWML2, the development methods, the representation itself, its testing and evaluation, its relation to other standards and relevant work, and a summary including possible future directions. Overall, it is shown that GWML2 is an

effective data transfer mechanism and a useful interoperability aid for groundwater data.

Usage scenarios for GWML2

Five usage scenarios motivate GWML2. These represent common and important data delivery situations, categorized as commercial, policy-oriented, environmental, scientific, or technological. They are selected to guide the design of GWML2 and to provide evaluation criteria, inasmuch as the data required by these scenarios should be representable in GWML2 and transmittable from major groundwater data sources to enable the scenarios to be carried out.

The commercial scenario involves finding water wells and springs within an area to enable estimation of the cost to complete a new water supply well—for example, a consultant or water well driller could use a web tool to explore the local geology and inspect wells located near the target area. By investigating the surficial sediment and rock materials, water level, yield and total depth at each well, the consultant could infer the distance and materials to the water table, as well as the expected yield, and the driller could estimate the cost and timeline for drilling. The public are also able to assess online water well records and make independent estimates. This not only aids evaluation of drilling potential, but consequently might also influence property purchases. Key entities in this scenario include water wells, related measurements such as water levels, as well as various hydrogeological units such as aquifers or confining beds.

The policy scenario involves reporting on the state of groundwater in specific administrative districts, motivated in large part by European requirements. Involved is the collection and evaluation of geological and hydrogeological characteristics, as well as the quantitative and qualitative monitoring of chemical and physical indicators. This aids the overall assessment of a management area, especially one that crosses political borders, and typically requires synchronization of information collected by multiple state water authorities. Key entities include management areas, related hydrogeological units, and monitored information.

The environmental scenario enables environmental managers, water managers, and legislators to assess threats to groundwater dependent ecosystems. The role of groundwater in sustaining environmental values is of growing importance globally, particularly in arid countries such as Australia. Key items include depth to water table, monitored information on groundwater chemistry and biology, and flow between groundwater and surface water.

The scientific scenario focuses on the delivery of data for use in groundwater flow modeling and soil-water balance modeling, as one among a myriad of possible scientific activities. Key entities for such modeling include the

hydrogeological and geophysical properties of aquifers and related measurements, as well as characteristics of water wells, water bodies, and water use.

The technologic scenario determines compatibility with other hydrogeological data representations such as database schemas and exchange formats, via conversion to and from GWML2. This is particularly important to enable data interoperability (1) within a groundwater data network, by converting between local databases and GWML2, or (2) between different data networks, by converting between GWML2 and local data formats such as the European-wide INSPIRE (INSPIRE 2013a,b) or North American GWML1 (Boisvert and Brodaric 2012). Involved are all entities from the previous scenarios.

Methods for developing GWML2

GWML2 developers include hydrogeologists and information scientists from a cross-section of major governmental and academic groundwater data providers. The development of GWML2 follows five steps, as shown in Fig. 1, beginning with a needs assessment and concluding with a tested design published as an international standard. Step 1 involves the establishment of prototypical usage situations, as described in section ‘Usage scenarios for GWML2’, which form the requirements and testing criteria for GWML2. In step 2, key terms are identified from the usage scenarios and authoritative definitions are attached to ensure a common conceptual foundation as well as to frame GWML’s scope (GWML2 Vocabulary, 2017). Step 3 involves the development of a technical design (Brodaric 2017) expressed as three related information models that culminate in a specification for encoding data in XML (Extensible Markup Language; W3C 2008). In step 4 example encodings are developed for each component of GWML2, and these are made permanently available online alongside the design specification (GWML2 XML Examples

2017; GWML2 XML Schema 2017). Step 5 tests the encodings in implementations across North America, Europe and Australasia to ensure satisfaction of the usage scenarios, thereby providing proof of feasibility (see section ‘Implementation and evaluation of GWML2’).

Results: the GWML2 design

As sketched in Fig. 2, GWML2 can represent a wide variety of entities associated with hydrogeological units, subsurface bodies of water, and man-made artifacts such as wells. A novel aspect of its design is the refinement of a water container pattern (from Boisvert and Brodaric 2012), which conceptually distinguishes (1) a container such as a rock body or sand unit, (2) the spaces (voids) hosted by the container, and (3) the fluids occupying those spaces. These distinctions enable properties to be attributed to each of the containers, fluids, or voids individually, as well as to their relations. For example, an aquifer vulnerability estimate is a property of the unit or its part, the volume of a fluid is a property of the fluid body, and the volume occupied by voids such as fractures or pores, is a property of the voids. In contrast, porosity is a property of the relation between a container and its hosted voids, if understood as the proportion of void volume to total unit volume.

Types of hydrogeological units in GWML2 are aquifers, aquifer systems, confining beds, and groundwater basins. Such units are understood to be subsurface volumes of earth material that serve as containers for fluids. Their physical boundaries are demarcated by hydrogeological characteristics or, in the case of basins, by groundwater flow divides. As fluid containers, units variously host voids and are associated with a wide variety of properties such as porosity, transmissivity, storativity, hydraulic conductivity, yield and so on. GWML2 allows hydrogeological unit properties, and many others, to be further categorized by the data provider—for example, the porosity property can be additionally categorized as effective porosity, total porosity, or any other type of porosity, and the yield property can be similarly further categorized as safe yield, sustainable yield, aquifer yield, or any other relevant notion of yield. Importantly, these categorizations do not replace the standard GWML2 property, rather they serve as an optional and supplementary description for it. In addition to such augmented categorization, GWML2 also allows multiple occurrences of property values. This enables, for instance, not only both total and effective porosity to be specified for a hydrogeological unit, possibly as summary values for the unit as a whole, but it also permits the specification of multiple property values scattered throughout the unit to capture their spatial distribution, which is especially required by the scientific modeling usage scenario. This flexible mode of representing property values is due to re-use of the ISO/OGC Observations and Measurements standard (O&M; Cox

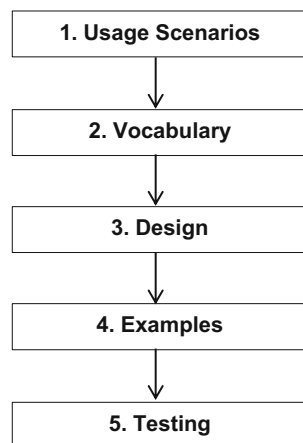
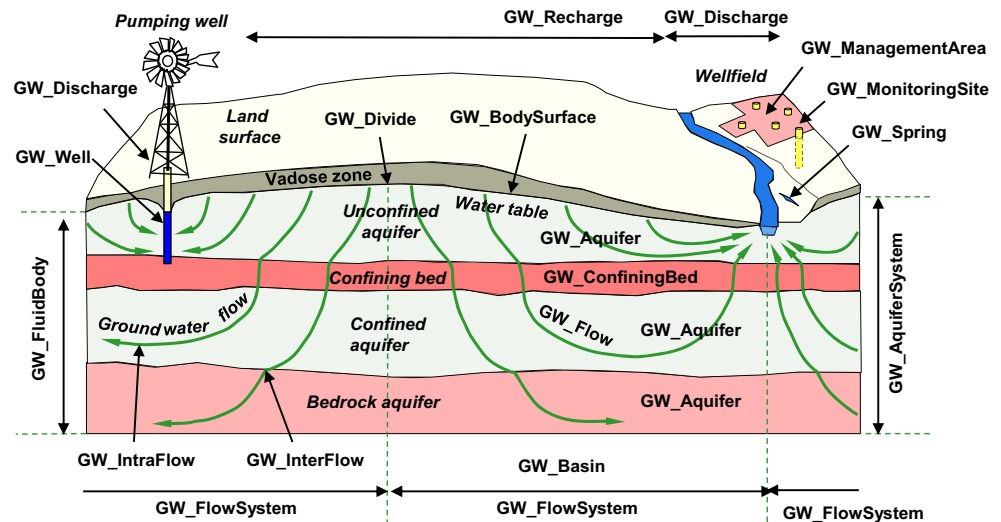


Fig. 1 GWML2 development method

Fig. 2 Main entities represented by GWML2, labeled with GW



2011, 2013), for representing measured values, and re-use of the OGC GeoSciML standard (GeoSciML, 2017), for representing a full geological description of a hydrogeological unit, including its materials, age, stratigraphic relations, and associated properties.

Figure 3 illustrates the use of the GeoSciML and O&M standards in a partial GWML2 encoding for a Canadian aquifer system, namely the Appalachian External Zone. Encoded are multiple material compositions from GeoSciML (i.e. `gsml:composition`), prefixed with “`gsml`”, and with the “Shale” instance fully open. Also shown are some key characteristics of the hydrogeological unit, prefixed by “`gwml2`” such as recharge (i.e. `gwml2:gwUnitRecharge`), discharge (i.e. `gwml2:gwUnitDischarge`), and the fully open hydraulic conductivity description (i.e. `gwHydraulicConductivity`). The value for this latter property is represented as a complex Observation, prefixed with “`om`”, and includes supplemental categorization of the property as a statistical median via the assignment of the “Hydraulic Conductivity – Median” qualifier (i.e. for `om:observedProperty`). Note the omission of several sections in Fig. 3 for reasons of space and readability, as indicated by “[...snip...]”.

Hydrogeological units in GWML2 can also be associated with water budgets, discharge and recharge estimates and their geographical zones, as well as management areas. Fluid bodies in GWML2 refer to the fluid contained in the voids of hydrogeological units, including their biologic, chemical and material constituents. Fluid bodies can also be subdivided into nested parts such as plumes, and can host various surfaces such as a water table, piezometric or potentiometric surface. These surfaces can in turn host divides—that is, lines projected onto the surface to denote the divergence of flow direction within the fluid body. Types of flow represented by GWML2 include those within containers, as well as between

containers. Intra-container flow is exemplified by the movement of fluid within a unit, and inter-container flow is exemplified by recharge and discharge.

Man-made artifacts in GWML2 consist of water wells and monitoring sites, as well as related measurements and calculations recorded as single values or multi-valued time series. Water wells can be described geologically, via reference to their host hydrogeological units, as well as by well logs that describe the lithology of the different units or various measurements along the length of the borehole such as those from borehole geophysics logging. Water wells can, of course, also be described by various hydrogeological characteristics such as water level depth and yield, and by construction characteristics associated with casings, screens, and seals. GWML2 monitoring sites are places where sensing devices periodically measure significant hydrogeological properties such as water level, flow rate, water temperature, and chemical composition, or where samples can be taken. Such sites are closely related to other entities in GWML2, for example water wells or springs, which can host the devices or the entity being monitored or sampled.

GWML2 also provides a framework for representing a wide variety of aquifer tests, from which hydrogeological properties can be determined. This might involve pumping or injecting fluid at known rates, and then observing changes in the water table over time, or injecting a tracer at some location and following its progression at observation points. GWML2 representation of such tests then interrelates data about initial conditions, test parameters and procedures, as well as sampling sites and pertinent entities such as wells and the targeted hydrogeological unit, in addition to the measured and calculated values for various GWML2 properties. If perchance an exotic property is missing in GWML2, it can be added as a supplementary categorization to an observation description. In this way, a wide variety of aquifer tests can


```

- <gwml2:GW_AquiferSystem
  [... snip ...]
  <gml:name codeSpace="http://gw-info.net">Appalachian External Zone</gml:name>
  [... snip ...]
- <gsml:composition>
  - <gsml:CompositionPart>
    <gsml:role nilReason="unknown"/>
    <gsml:material xlink:title="Shale"
      xlink:href="http://resource.geosciml.org/classifier/cgi/lithology/shale"/>
    </gsml:CompositionPart>
  </gsml:composition>
+ <gsml:composition>
+ <gsml:composition>
  [... snip ...]
  <gwml2:gwUnitMedia xlink:title="granular" xlink:href="http://gw-info.net/media/granular"/>
  <gwml2:gwUnitWaterBudget xsi:nil="true"/>
+ <gwml2:gwUnitRecharge>
+ <gwml2:gwUnitRecharge>
  [... snip ...]
+ <gwml2:gwUnitDischarge>
+ <gwml2:gwUnitDischarge>
  [... snip ...]
  - <gwml2:gwHydraulicConductivity>
    - <om:OM_Observation gml:id="gu.fprop.1.1">
      <gml:description>Hydraulic conductivity for the hydrogeological unit. The median value for
        hydraulic conductivity is  $10^{-6}$  0.2 m/s, but it ranges from  $10^{-7}$  0.8 to  $10^{-4}$  0.8 m/s.
        Hydraulic regional conductivity The decreasing trend of transmissivity with depth in the
        rock is observed in all contexts. The average hydraulic conductivity is  $10^{-3}$  0.9 m/s near
        the top of bedrock (z = 1 m) which gradually loses an order of magnitude at a depth of 10
        m ( $10^{-4}$  0.9 m/s), 25 m ( $10^{-5}$  0.9 m/s), 60 m ( $10^{-6}$  0.9 m/s) and 200 m ( $10^{-7}$  0.9
        m/s).</gml:description>
      <om:phenomenonTime xlink:href="http://www.opengis.net/def/uom/nil" nilReason="missing"/>
      <om:resultTime xlink:href="http://www.opengis.net/def/uom/nil" nilReason="missing"/>
      <om:procedure xlink:href="http://www.opengis.net/def/uom/nil" nilReason="missing"/>
      <om:observedProperty xlink:title="Hydraulic Conductivity - Median" xlink:href="http://ngwd-
        bdnes.cits.nrcan.gc.ca/Reference/uri-
        cgi/classifier/ca.gin/NRCanGroundWaterTopics/1404"/>
      <om:featureOfInterest xlink:href="#gin.1"/>
      <om:result uom="http://sweet.jpl.nasa.gov/2.3/reprSciUnits.owl#meterPerSecond"
        xsi:type="gml:MeasureType">0.2E-6</om:result>
    </om:OM_Observation>
  </gwml2:gwHydraulicConductivity>
</gwml2:GW_AquiferSystem>

```

Fig. 3 Partial GWML2 encoding for an aquifer system and its median hydraulic conductivity. “[...snip...]” denotes sections deleted for reasons of readability and space

be represented (see Brodaric 2017, pp. 106–115). A full description of the entities represented by GWML2, as well as their technical design, can be found in the standards specification (Brodaric 2017).

Implementation and evaluation of GWML2

GWML2 is successfully implemented by nine major providers of groundwater data, with the implementations collectively delivering a wide variety of data from diverse systems. Each component of GWML2 is served online from at least one data source, and many components are served by several sources (Brodaric 2016), enabling GWML2 to meet the following criteria:

- *Deployability*: implementation with existing open standards and technologies was relatively uncomplicated across a broad range of data systems, and the design was shown to be compatible with diverse software environments.

- *Completeness*: GWML2 was shown to cover the wide variety of key data types available in the data sources with minimal information loss. Moreover, successful encoding of each GWML2 component demonstrated conformance with the intended design.
- *Usability*: the information delivered by GWML2 satisfied the primary needs of the usage scenarios, as demonstrated by the successful delivery of all key entities from disparate databases. While efficiency was not measured quantitatively, in all cases the implementations returned results in times acceptable to human users, as indicated in ad-hoc reports.

Overall usability, including satisfaction of several usage scenarios, is further demonstrated by the consumption of GWML2 data by various web applications. Figure 4 illustrates satisfaction of the commercial scenario as GWML2 water well information is visualized from the Canadian (Fig. 4a) and US (Fig. 4b) groundwater data networks (Brodaric et al. 2016) via the Canadian web portal. This visualization also demonstrates satisfaction of an important aspect of the technological

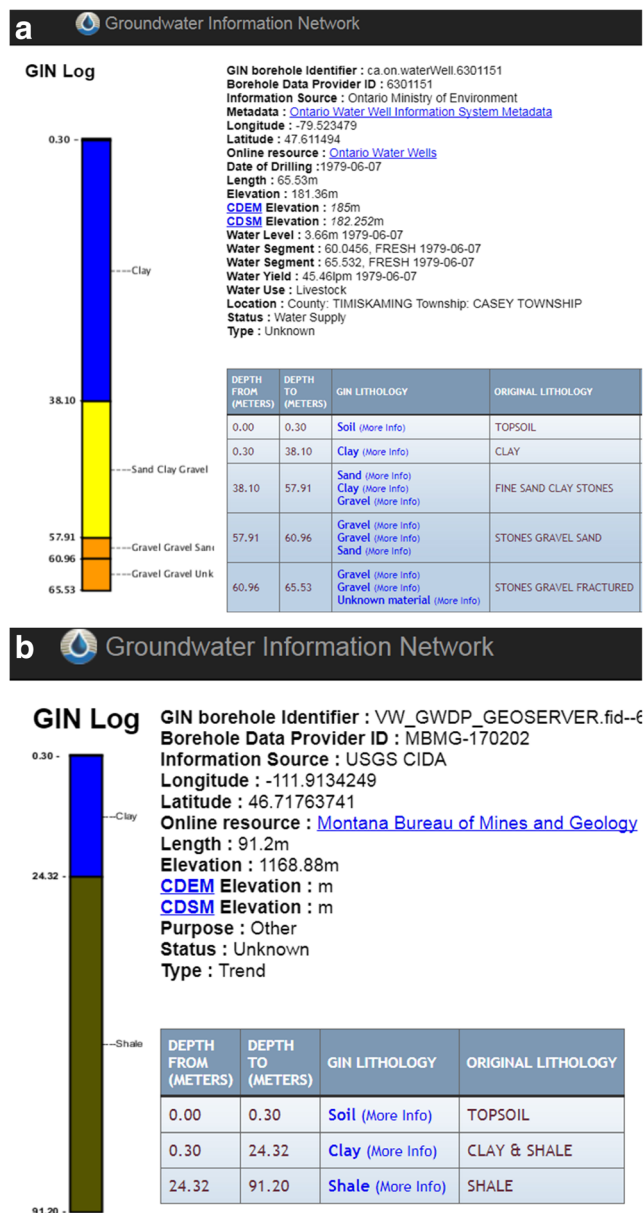


Fig. 4 **a** GWML2 water wells from the Canadian Groundwater Information Network (GIN), displayed in the GIN web portal (Brodaric et al. 2016). **b** GWML2 water wells from the US National Groundwater Monitoring Network (NGWMN), displayed in the GIN web portal (Brodaric et al. 2016)

scenario, namely the conversion of GWML1 to GWML2. Satisfaction of the environmental scenario is demonstrated in Fig. 5, which shows GWML2 groundwater chemistry in support of groundwater dependent ecosystems displayed in the Australian VVG (Visualizing Victoria's Groundwater) web portal, and originating from a data network of federal, state and academic groundwater data providers (Dahlhaus et al. 2016). Figure 6 illustrates development of a groundwater flow model in MODFLOW using GWML2 input data from a management area in New Zealand. When considered in tandem with the 3D visualization of the management area and

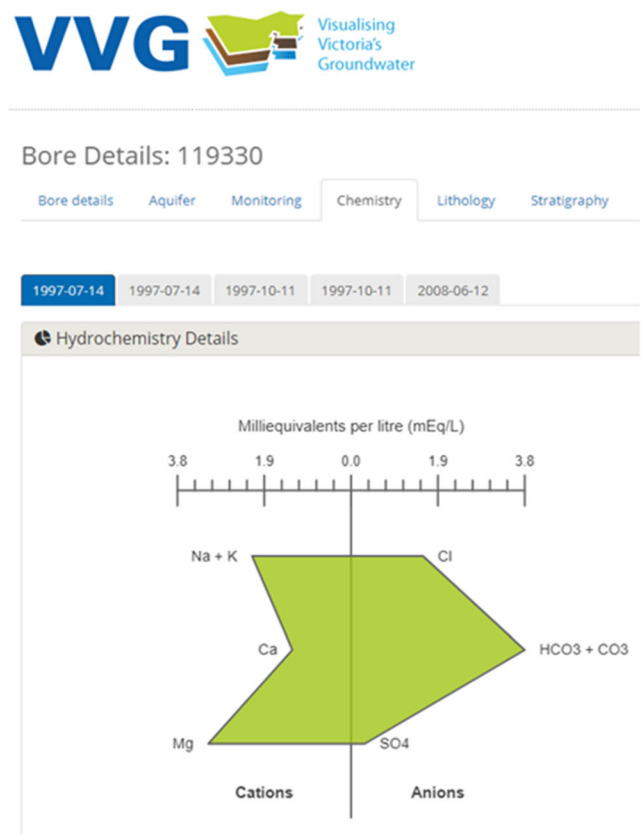


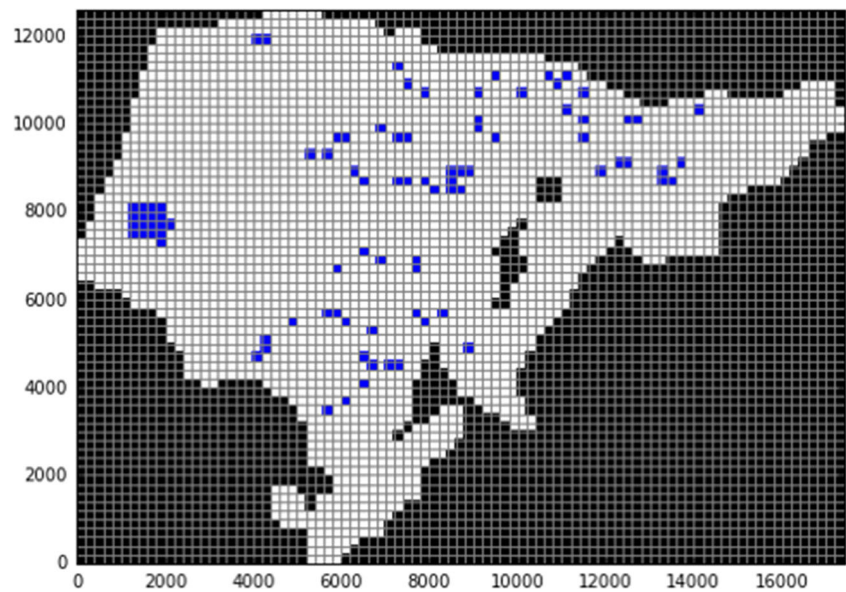
Fig. 5 GWML2 groundwater chemistry shown in the Australian VVG web portal (Dahlhaus et al. 2016)

model in Fig. 7 (Kmoch 2017), it is apparent that the scientific and policy scenarios are satisfied, inasmuch as a flow model was successfully constructed from GWML2 data, and from this several hydrogeological characteristics were attributed to the management area. These uses of GWML2 also highlight a significant additional benefit, namely that it facilitates the development of sharable hydrogeological software that can be built over a single normative data format.

Related work on groundwater data representation

Although data representations in hydrology are numerous, those focusing on groundwater are less common and can be categorized as (1) database schemas, (2) ontologies, and (3) data exchange formats. Groundwater database schemas provide a structure for organizing and managing groundwater information in data repositories, either for restricted scope (e.g. Chesnaux et al. 2011; Gogu et al. 2001; Steward and Bernard, 2006; Wojda et al. 2010; Oulidi et al. 2009), or for broad use (Kisters 2017; Nešetřil and Šembera 2014; Strassberg et al. 2007). Ontologies are formally arranged vocabularies used by data networks to help find and use distributed data, with a focus on hydrology in general (Atkinson and

Fig. 6 USGS MODFLOW single layer volumetric grid generated from GWML2 input data for the Horowhenua district management area, New Zealand. White cells denote the hydrogeologically active area, and blue cells denote sites for input observations from water wells (from Kmoch 2017). Cells are volumes with variable elevation and depth, possessing many hydrogeological properties from GWML2 such as hydraulic conductivity. The total area shown is approx. 12 km × 16 km, with distance in metres indicated along each axis



Dornblut 2012; Buttigieg et al. 2013; Raskin and Pan 2005), or groundwater in particular (Brodaric and Probst 2009; Tripathi and Babaie 2007; Yang et al. 2010). Groundwater data exchange formats are used to transfer data between a data supplier and consumer, and include the INSPIRE hydro/geology specification (INSPIRE 2013a), GWML1 (Boisvert and Brodaric 2012) and Hg2O (Wojda and Brouyère 2013). GWML2 clearly falls into this third category of data transfer format, but it is conceptually differentiated from representations in all categories by its adaptation of the water container pattern (from Boisvert and Brodaric 2012; see section ‘Results: the GWML2 design’), which distinguishes

containers (e.g. aquifers), water bodies, voids (e.g. pore spaces), and their properties and relations, enabling distinct description of each, and thus a more precise expression of the data.

Summary and future directions

GWML2 is as an international standard for exchanging a wide variety of groundwater data, to meet increasing demands for such data in many scientific and societal endeavors. Testing has shown that GWML2 can be used successfully to deliver

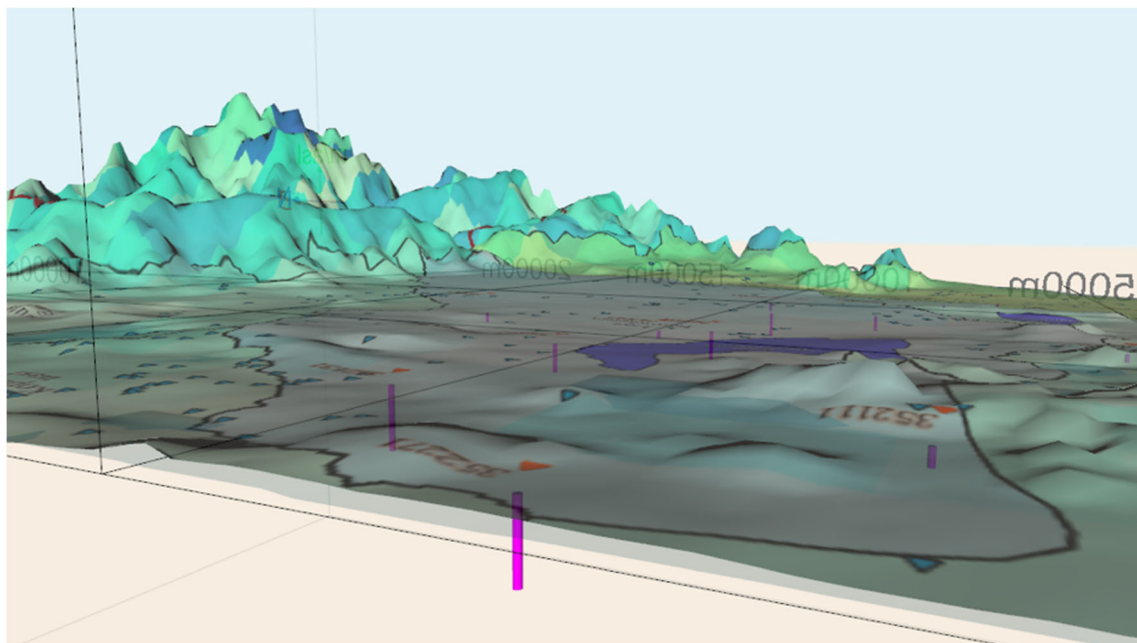


Fig. 7 3D visualization of GWML2 data for New Zealand management area, including related water wells from Fig. 6 (from Kmoch 2017)

groundwater information from significant data providers to meet key usage scenarios. In this process, GWML2 has already impacted the design of groundwater databases maintained by these providers, and is being incorporated into their operational data delivery mechanisms. It is further informing the nascent development of ontology-based knowledge structures for the hydrological domain (Hahmann et al. 2016).

Next steps for GWML2 involve its possible submission to the World Meteorological Organization for endorsement as a global hydrological standard, in concert with its WaterML2 siblings, to complement its current standing as a global geospatial standard. Possible future technical directions include the development of additional GWML2 encodings in languages other than XML, as well as the enhancement of standard content. While GWML2 provides a rigorous standard structure for groundwater data, in many cases, the items that populate that structure, in the form of controlled vocabularies, are not constrained—for example, GWML2 captures porosity or yield values associated with an aquifer and provides a means for specifying the particular type of porosity or yield. It does not, however, enumerate a controlled vocabulary of acceptable types of porosity or yield, leaving it to the data provider to determine these. As a result, GWML2 documents from different data sources will be structurally identically, but much of the content is likely to be heterogeneous. In essence, GWML2 assures structural interoperability, but only partial semantic interoperability of content. Advancement of standard content vocabularies for GWML2 is a future challenge.

Other potential work involves a “lite” version, to lower the entry barrier to GWML2, as well as advancement of the scientific modeling usage scenario to further test the efficacy of GWML2 in supplying data to, and between, groundwater modeling tools. Augmentation of the technologic usage scenario is also a possibility, to determine alignment with important standards such as with the ArcHydro Groundwater standard database schema (Strassberg et al. 2007), and with emerging surface water schemas such as HY_Features (Atkinson and Dornblut 2012). It is anticipated that these and other future directions for GWML2 will continue to be coordinated by the OGC Groundwater Standards Working Group in collaboration with related activities in industry, academia, and government, to promote GWML2 adoption and further the overall aim of enabling groundwater data exchange and interoperability.

Acknowledgements Many thanks to the editors as well as to the formal and informal reviewers for their very helpful guidance, which led to significant improvements in the paper. Included among them are J. Devlin, D. Blodgett, T. Ferre, C. Langevin, and C. Rivard. The Federation University of Australia acknowledges the support of several contributors in the development and testing of GWML2, most notably Andrew MacLeod.

Funding Information 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’. Funding for BRGM work is provided by its joint IT Research Center ‘INSIDE’ dedicated to innovation in Water Information Systems with the French National Agency for Biodiversity (AFB). Funding for the SMART Aquifer Characterization (SAC) programme at GNS Science, New Zealand, and the Interfaculty Department for Geoinformatics (Z_GIS), University of Salzburg, is provided by the Ministry of Business, Innovation, and Employment (MBIE), New Zealand. Funding for Natural Resources Canada (NRCAN) participants is provided by the NRCAN Groundwater Geoscience Program. Funding for U.S. Geological Survey participants is provided by the USGS Groundwater and Streamflow Information Program.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Atkinson R, Dornblut I (2012) OGC HY_Features: a common hydrologic feature model. OGC Discussion Paper 11-039r2, 49 pp. <http://www.opengis.net/doc/DP/hy-features>. Accessed 01 September 2017
- AWA (2012) Coal seam gas mining and groundwater: a review of the controversies around and potential impacts of CSG in Australia. *Water* 39(7):46–48
- BDLISA (2013) Dictionnaire des données: référentiel hydrogéologique (version 2) [Data dictionary: hydrogeological reference (version 2)]. <http://www.sandre.eaufrance.fr/urn.php?urn=urn:sandre:dictionnaire:SAQ:FRA::ressource:2.0::pdf>. Accessed 13 July 2015
- BDLISA (2014) Base de Données des Limites des Systèmes Aquifères [Aquifer System Limits database]. <http://www.reseau.eaufrance.fr/geotraitement/viewer/bdlisa>. Accessed 13 July 2015
- Blodgett D, Lucido J, Kreft J (2016) Progress on water data integration and distribution: a summary of select U.S. Geological Survey data systems. *J Hydroinf* 18(2):226–237
- Boisvert E, Brodaric B (2012) GroundWater markup language (GWML): enabling groundwater data interoperability in spatial data infrastructures. *J Hydroinf* 14(1):93–107
- BOM (2015) National Groundwater Information System. Bureau of Meteorology, Melbourne, Australia. <http://www.bom.gov.au/water/groundwater/ngis/>. Accessed 22 March 2015
- Boyd D, Crawford K (2012) Critical questions for big data, information. *Commun Soc* 15(5):662–679
- Brodaric B (ed) (2016) OGC GroundWaterML2: GW2IE final report. Open Geospatial Consortium engineering report 15-082, v2.1, 171 pp. https://portal.opengeospatial.org/files/?artifact_id=64688. Accessed 01 September 2017
- Brodaric B (ed) (2017) OGC WaterML2: part 4—GroundWaterML2 (GWML2). Open Geospatial Consortium Standard 16-032r2, v2.2.0, 160 pp. <http://docs.opengeospatial.org/is/16-032r2/16-032r2.html>. Accessed 01 September 2017
- Brodaric B, Booth N, Boisvert E, Lucido J (2016) Groundwater data network interoperability. *J Hydroinf* 18(2):198–225
- Brodaric B, Piasecki M (2016) Water data networks: foundations, technologies and systems, implementations and uses. *J Hydroinf* 18(2):149–151
- Brodaric B, Probst F (2009) Enabling cross-disciplinary e-science by integrating geoscience ontologies with DOLCE. *IEEE Intell Syst* 24(1):66–77
- Buttigieg PL, Morrison N, Smith B, Mungall CJ, Lewis SE, ENVO Consortium (2013) The environment ontology: contextualising

- biological and biomedical entities. *J Biomed Semantics* 4(1):43. <https://doi.org/10.1186/2041-1480-4-43>
- Carrara E, Nation E, Sharples J (2015) Groundwater information at the Bureau: providing a national picture. In: HWSR2015, 36th Hydrology and Water Resources Symposium, Hobart, Tasmania, 7–10 Dec 2015, Engineers Australia, Barton, Australia
- Chesnaux R, Lambert M, Walter J, Fillastre U, Hay M, Rouleau A, Daigneault R, Moisan A, Germaneau D (2011) Building a geodatabase for mapping hydrogeological features and 3D modeling of groundwater systems: application to the Saguenay-Lac-St-Jean region, Canada. *Comput Geosci* 37(11):1870–1882
- Cox S (2011) Observations and measurements: XML implementation v2.0. Open Geospatial Consortium Standard 10-025r, 77 pp
- Cox S (2013) Geographic information: observations and measurements v2.0. Open Geospatial Consortium Standard 10-004r3, 54 pp
- Dahlhaus PG, MacLeod A, Thompson H (2012) Federating hydrogeological data to visualise Victoria's groundwater. Fourth Circular, 34th International Geological Congress, Brisbane, Australia, 5–10 August 2012. Congress Abstracts (Abstract no. 1382), 592 pp
- Dahlhaus PG, Murphy A, MacLeod A, Thompson H, McKenna K, Ollerenshaw A (2016) Making the invisible visible: assessing the impact of federating groundwater data in Victoria, Australia. *J Hydroinf* 18(2):238–255
- GeoSciML Modelling Team (2017) Geoscience Markup Language 4.1 (GeoSciML). Open Geospatial Consortium Standard 16-008, v4.1, 247 pp
- Gogu RC, Carabin G, Hallet V, Peters V, Dassargues A (2001) GIS-based hydrogeological databases and groundwater modelling. *Hydrogeol J* 9(6):555–569
- Gorelick SM, Zheng C (2015) Global change and the groundwater management challenge. *Water Resour Res* 51:3031–3051. <https://doi.org/10.1002/2014wr016825>
- GWML2 XML Examples (2017) GWML2 instances. Open Geospatial Consortium Schema Repository, GWML2. <http://schemas.opengis.net/gwml/2.2/instances/>. Accessed 01 September 2017
- GWML2 XML Schema (2017) GWML2 schema. Open Geospatial Consortium Schema Repository, GWML2. <http://schemas.opengis.net/gwml/2.2/>. Accessed 01 September 2017
- GWML2 Vocabulary (2017) GWML2 final entities List. Hydro Domain Working Group wiki. http://external.opengis.org/twiki_public/HydrologyDWG/Gwml2FinalFeaturesList. Accessed 01 September 2017
- Hahmann T, Stephen S, Brodaric B (2016) Semantically refining the groundwater markup language (GWML2) with the help of a reference ontology. In: Miller J, O'Sullivan D, Wiegand N (eds) Proceedings of the 9th International Conference on GIScience, 27–30 Sept. 2016, Montreal, Canada
- 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_GE_v3.0.pdf. Accessed 08 March 2018
- INSPIRE (2013b) D2.8.III.11 INSPIRE data specification area management/restriction/regulation zones and reporting units, D2.8.III.11_v3.0. http://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_AM_v3.0rc2.pdf. Accessed 07 March 2018
- Kisters (2017) Kisters Hydstra. <http://kisters.com.au/hydstra-features.html>. Accessed 01 September 2017
- 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. *Comput Geosci* 69:78–86
- Kmoch A (2017) A context-based groundwater data infrastructure. PhD Thesis, Australia University of Technology, 268 pp, <https://aut.researchgateway.ac.nz/handle/10292/10740>. Accessed 01 September 2017
- Kmoch A, Klug H, Ritchie ABH, Schmidt J, White PA (2015) A spatial data infrastructure approach for the characterization of New Zealand's groundwater systems. *Trans GIS* 20(4):626–641. doi: <https://doi.org/10.1111/tgis.12171>
- Loch A, Adamson D, Mallawaarachchi T (2014) Role of hydrology and economics in water management policy under increasing uncertainty. *J Hydrol* 518(part A):5–16
- Mayer-Schonberger V, Cukier K (2013) Big data: a revolution that will transform how we live, work and think. Harcourt, New York, 242 pp
- Nativi S, Mazzetti P, Craglia M, Pirrone N (2014) The GEOSS solution for enabling data interoperability and integrative research. *Environ Sci Pollut Res* 21(6):4177–4192
- Nešetřil K, Šembera J (2014) Groundwater data management system. In: Proceedings of the 28th EnviroInfo 2014 Conference, Oldenburg, Germany, September 10–12, 2014
- Nevill JC, Hancock PJ, Murray BR, Ponder WF, Humphreys WF, Phillips ML, Groom PK (2010) Groundwater-dependent ecosystems and the dangers of groundwater overdraft: a review and an Australian perspective. *Pac Conserv Biol* 16(3):187–208
- Oulidi HJ, Ralf L, Benaabidate L, Wachter J (2009) HydrIS: an open source GIS decision support system for groundwater management (Morocco). *Geo-spatial Inf Sci* 12(3):212–216
- Raskin RG, Pan MJ (2005) Knowledge representation in the semantic web for earth and environmental terminology (SWEET). *Comput Geosci* 31(9):1119–1125
- Scanlon BR, Faunt CC, Longuevergne L, Reedy RC, Alley WM, McGuire VL, McMahon PB (2012) Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley. *Proc Natl Acad Sci U S A* 109(24):9320–9325
- Steward DR, Bernard EA (2006) The synergistic powers of AEM and GIS geodatabase models in water resources studies. *Ground Water* 44:56–61
- Strassberg G, Maidment DR, Jones NL (2007) A geographic data model for representing ground water systems. *Groundwater* 45(4):515–518
- Taylor RG, Scanlon B, Doll P, Rodell M, van Beek R, Wada Y, Longuevergne L, Leblanc M, Famiglietti JS, Edmunds M, Konikow L, Green TR, Chen J, Taniguchi M, Bierkens MFP, MacDonald A, Fan Y, Maxwell RM, Yechieli Y, Gurdak JJ, Allen DM, Shamsudduha M, Hiscock K, Yeh PJF, Holman I, Treidel H (2013) Ground water and climate change. *Nat Clim Chang* 3(4):322–329
- Tripathi A, Babaie HA (2007) Developing a modular hydrogeology ontology by extending the SWEET upper-level ontologies. *Comput Geosci* 34(9):1022–1033
- Vengosh A, Jackson RB, Warner N, Darrah TH, Kondash A (2014) A critical review of the risks to water resources from unconventional shale gas development and hydraulic fracturing in the United States. *Environ Sci Technol* 48(15):8334–8348
- Vidic RD, Brantley SL, Vandenbossche JM, Yoxtheimer D, Abad JD (2013) Impact of shale gas development on regional water quality. *Science* 340(6134)
- WaterML2 (2017) OGC WaterML2 hydrological information standards. <http://www.opengispatial.org/standards/waterml>. Accessed 01 September 2017
- Wojda P, Brouyère S (2013) An object-oriented hydrogeological data model for groundwater projects. *Environ Model Softw* 43:109–123
- Wojda P, Brouyère S, Derouane J, Dassargues A (2010) HydroCube: an entity-relationship hydrogeological data model. *Hydrogeol J* 18(8):1953–1962
- World Wide Web Consortium (W3C) (2008) Extensible Markup Language (XML) 1.0, 5th Edn. <https://www.w3.org/TR/REC-xml/>. Accessed 01 September 2017
- Yang X, Steward DR, de Lange WJ, Lauwo S, Chubb RM, Bernard EA (2010) Data model for system conceptualization in groundwater studies. *Int J Geogr Inf Sci* 24(5):677–694
- Zuiderwijk A, Janssen M (2014) Open data policies, their implementation and impact: a framework for comparison. *Gov Inf Q* 31(1):17–29