

Data and Services Discovery projects - Transformative Data Collections

Title

Developing the Water and Energy Supply and Consumption (WESC) Data Standard

Introduction

This report summarises the scoping study activities for the project “Developing the Water and Energy Supply and Consumption (WESC) Data Standard”, which were held between 14 June and 29 November 2019. The WESCML standard was originally developed by CSIRO and AURIN in 2013-2015, then deployed in the [AURIN-CSIRO WESC data hub](https://www.wescml.org/) proof-of-concept. This consisted of several layers including an information model, vocabularies, tools and data, based on user requirements (Yu et al., 2014, Simons et al. 2016, Simons & Yu 2015) (<https://www.wescml.org/>).

Approach

The project comprised three main activities: data collection, analysis and reporting (Table 1).

Activity	Description	Date performed	Outputs
Data collection			
Human research ethics (HRE)	Submit application to The University of Melbourne, approved 9 August 2019. This was a precondition for all data collection activities.	Jul – Aug 2019	• Plain language statements, consent forms
Desktop review	Review of WESC data context in Australia including curation, legal and regulatory frameworks, policy drivers, other standards and reports.	Aug – Oct 2019	• Summary
Surveys and interviews	Conduct interviews and online surveys with data users and providers across Australia.	Aug – Oct 2019	• Responses • Transcripts
Forum	Convene a forum to bring data users, providers and others together at The University of Melbourne.	6 Sep 2019	• Transcripts • Notes
Analysis			
Data/contextual analysis	Summarise data collected, including data provider and user value perceptions and barriers to data.	Sep - Oct 2019	• Summary
Gap analysis	Identify gaps and propose solutions.	Oct – Nov 2019	• Summary
Reporting			
Interim report	Deliver interim report based on summaries and preliminary analyses.	8 Oct 2019	• Interim report
Presentation	Present project summary at the ARDC Data and Services Summit, Brisbane.	21 Oct 2019	• Presentation
Presentation	Present project to obtain feedback: • eResearch Australasia 2019, Brisbane • Review with key participants	23 Oct 2019 29 Nov 2019	• Presentations
Final report	Deliver final report to ARDC.	2 Dec 2019	• Final report

Table 1: List of scoping study activities, descriptions, outputs, dates and progress

FAIR

Part of the desktop review activity for data collection focused on reviewing the existing AURIN-CSIRO WESC data hub, to determine how findable, accessible, interoperable and reusable (FAIR) it is for data users. As this is a scoping study only, the review focused on AURIN's infrastructure and tools that serve (meta) data in both machine and human readable formats (this excludes a priori technical work at CSIRO to process data into the WESCML).

Following the formal assessment, the data hub was found to be moderately FAIR with several recommendations for improvement (Appendix 1). Proposed future work includes: use of globally unique, citable and persistent identifiers for data and vocabularies; maintaining the persistence of archived (meta) data; adoption of provenance standards and linking. Note, while only one dataset was assessed, the findings and recommendations apply to all others in the hub.

Collaboration and coverage

The scoping study engaged a wide variety of representatives from research, industry and government* (Table 2), which were from both Australia and overseas. Requests for contributions and collaborations were publicised through existing AURIN and CSIRO communication networks^ and participants were not incentivised.

Category	Participation counts		
	Survey/ Interview	Forum	Total
Research	11	14	25
Industry	6	6	12
Government	2	3	5
Total	19	23	42

Table 2: Collaboration and coverage of activities as of 29 November 2019

*Human research ethics requires that participants and their affiliations remain anonymous.

^Engagement campaign 9 Jul to 6 Sep 2019 had 30 social media announcements with 18,000+ impressions and 250+ likes/clicks

Learnings

The scoping study commenced by investigating the value of and barriers to sharing WESC data, which were identified as priorities during initial interviews and surveys with data users and providers. While the WESCML standard was found to be technically sufficient to support the provision of data overall, as illustrated by its adoption in the National Energy Analytics Research (NEAR) program; data providers are not actively adopting the standard for several reasons.

Barriers

Existing barriers for data providers to adopt the standard relate to the time and resources required and the low perceived return on investment. Current data priorities are geared towards supporting stakeholders (e.g. shareholders and customers) and meeting legislative obligations to share data with government authorities. Data teams reported that they are already having to cope with increasing data volumes, competing data priorities, privacy and ethics matters,

responding to ongoing staff turnover, skills shortages and other resource constraints. Any additional work to adopt the standard will add pressure to resourcing, be seen as detracting from business as usual (BAU) and have a negative value. These observations concord with related research from CSIRO on the role of social architecture to facilitate information infrastructures (Box and Lemon, 2015).

Seeking to support their data-driven work, data users rely heavily on data provided by open data portals, however they reported that current (meta) data supplied by data providers is of little value, lacking appropriate geographic scale and granularity, being of poor quality and frequently insufficiently licenced. In search of higher value data, users need to lodge ad hoc requests with providers, however find it very difficult to connect and obtain anything beyond what is publicly available. For these reasons, users feel that data providers have a 'take it or leave it' attitude, which is blocking their contributions to research and the quadruple bottom line (social, cultural, economic, environmental)¹. Interestingly, while the NEAR program seeks to facilitate data availability for energy supply and consumption, there appears no national equivalent in the water sector; particularly for consumption data. This presents an opportunity for future work.

Developing value propositions

To work towards the standard's adoption and the further development of the WESC data hub, data providers require a strong value proposition that principally demonstrates its economic and operational benefits. If value propositions are suitably formed, e.g. Smart (2017), they may be presented to get 'buy in' from executive management in data provider organisations and other relevant industry bodies. When considered within a broader context such as the interests of government, propositions may be extended to consider other social, cultural and environmental values.

It is important to note that value may be multi-faceted and include insights gained from research enabled by improved data standardisation, access, coordination and use. These feedback loops are typically in the order of 3-10 years due to the time required for project startup, analysis and publication, which is comparably much longer than the time needed by data providers to demonstrate value (e.g. annual reporting). Therefore, opportunity exists to explore potential short-term gains alongside longer-term benefits to develop more attractive propositions.

Future data challenges

While current data challenges are based on existing delivery of WESC data, data providers and governments are facing a range of new ones from disruptions such as the emergence of new technologies influencing both supply and consumption. First, data providers are deploying new sensors and automated systems that create and utilise improved data outputs for managing water and energy, but this huge amount of data is adding to existing data pressures. Second, new residential-based supply such as solar PV and waste/recycled water, together with real-time data, are facilitating innovative projects such as P2P (peer to peer) trading that will likely change actual water and energy consumption behaviour and empower local communities,

¹ANZ Guidelines for Fresh & Marine Water Quality, <https://www.waterquality.gov.au/anz-guidelines/resources/key-concepts/quadruple-bottom-line>

particularly through innovative water-energy nexus applications. Additionally, data is being collected by new solutions driven by startups and technology companies, which often use bespoke supply and consumption solutions. Yet as a data standard is not currently being developed or adopted for this market, data is not being shared beyond individual systems.

In addition to these disruptions, much broader challenges from climate change, population growth and associated land use change are already taking effect on planning and governance. The immediacy, scale and complexity of these may constitute drivers for the overhaul of existing systems and associated linear economic models, which might feed on a range of data-driven activities (Stahel, 2016). It is here that data demand is increasing and requires a higher level of connectivity between data users such as researchers and policy makers, and WESC data providers.

Recommendations

These current and future concerns lead to questions relating to the future feasibility of the WESCML standard and the AURIN-CSIRO data hub. As solutions to increasing the standard's adoption sit predominantly within the social architecture space, key recommendations and priority tasks in governance, agreements and participation areas need to be considered (Box and Lemon, 2015), complementing the technical and information engineering work done to date.

We propose future activities via a follow-on project that would further develop the social architectures which will in turn enable greater adoption, use and future technical development of the WESCML standard and WESC data hub. The following next steps are proposed:

Recommendation 1. Demonstrate the value of WESCML

Strong case studies are needed to illustrate the standard's value to WESC data providers, industry and government. This requires:

- Documenting current practices and research barriers in Australia
- Identifying international initiatives for potential collaboration

Recommendation 2. Raise the awareness of WESCML and WESC data

Promote awareness of the standard and its data across Australia and internationally via WESC related data working groups and relevant community events. Water examples include:

- Industry: Water Services Association of Australia (WSAA), Victorian Intelligent Water Network (IWN), VicWater, Australia Water Association (AWA) and events such as OzWater conferences
- Research: Research Data Alliance, International Conference Hydroinformatics
- Government: Department of Energy and Environment (DEE, Federal) and Department of Agriculture (DA, Federal), Department of Environment, Water, Land and Planning (DEWLP, Victoria), Municipal Association of Victoria Technology (MAV Tech, Local)

Present at other working groups and coordinating initiatives to explore opportunities to align WESCML with other relevant standards, with a view to co-develop a harmonised international standard at bodies such as the Open Geospatial Consortium (OGC). Water examples include:

- Water Data Transfer Format (WDTF²) – Australia (currently supply data only)
- Water Data Exchange (WADE³) – United States
- Site-Specific Water-Use Data System (SWUDS⁴) – United States
- Water Information System for Europe (WISE⁵) – Europe

Recommendation 3. Develop partnerships and coordinate initiatives from industry and government in order to articulate value propositions and define projects

Develop pathways from industry and government to help data provider BAU to handle existing and new types of data, work with privacy and ethics issues and estimate value. Initiatives to fund these pathways may include co-designing and forming project proposal/funding schemes with strong partnership requirements with private, public and research sectors that encourage collaboration and coordination in WESC data.

Understanding the opportunity for a national water analytics platform, example scoping for funding may include:

- Explore and implement extensions to the BoM's Data Connect (used by utilities to upload water supply data via the WDTF format for National Performance Reporting (NPR) purposes) and the Water Data Online⁶ platform, extending them to accommodate urban water supply and consumption data via the WESCML standard, as well as extension points for other relevant data
- Explore and fund partnership activities with WSAA and partners via the Digital Strategy and IoT forum, including a range of immediate short time benefits (low-hanging fruit) alongside longer-term objectives.

Engagement with other socio-technical initiatives such as the Location Integration Capability (Loc-I, CSIRO), and other international standards, where location and linked datasets are key, transformative components.

Partner with other existing data initiatives such as the National Energy Analytics Research (NEAR) project.

Sustainability

The ongoing sustainability of WESCML and the data hub relies heavily on both the social and technical architectures for reasons of relevance and value. Two roles are necessary to enact the

² WDTF - Bureau of Meteorology, <http://www.bom.gov.au/water/standards/wdtf/>

³ WADE - Western States Water Council, <http://wade.westernstateswater.org/>

⁴ SWUDS - United States Geological Survey, <https://water.usgs.gov/watuse/data/>

⁵ WISE - European Commission and European Environment Agency, <https://water.europa.eu/>

⁶ Water Data Online - Bureau of Meteorology, <http://www.bom.gov.au/waterdata/>

recommendations above: data relationships and technical development, which would need to interrelate. AURIN is an NCRIS capability tasked to support social sciences and urban planning research that is committed to pursuing data relationships on both data user and provider fronts. Given that there are over 100 water and energy utilities across Australia alone, establishing flexible data relationships must be a priority together with developing strong case studies to promote the value of adopting WESCML, while developing partnerships and initiatives in parallel.

Once adopted, two potential pathways can be formed from the three recommendations above to work towards the future sustainability of the WESC data hub. In the first, water and energy data would continue to be manually uploaded by utilities into the WESC data hub by way of CSIRO's data infrastructure; consisting of virtual machines with geospatial technology stacks. In the second pathway, the manual process of uploading water consumption data would be integrated with BoM's Data Connect. Here the WDTF standard would need to be investigated to include water consumption classes and attributes from WESCML, then the value of linking BoM's Data Connect to the AURIN-CSIRO WESC data hub needs to be clearly illustrated. If successful, this would provide value back to utilities via access to a national water consumption dataset (that does not currently exist) and work towards establishing a national water supply and consumption analytics platform for Australia. Similar connections relating to regulatory reporting may then be explored within energy data spaces with models, processes and lessons shared.

In both pathways, knowledge gained would provide critical input into the ongoing technical development of WESCML and the hub to ensure its relevance and operation into the future. Here sustainable research software best practices would need to be followed⁷, together with maintaining the environment where software is developed and used (currently a Github [repository](#)). Coordination of strategic direction and oversight would require authority structures, which might be established via peak bodies or other organisations (Box and Lemon, 2015).

Impact

Numerous outcomes and impacts are anticipated from this project. Summarised above, this project has engaged with a range of research, industry and government groups in Australia and overseas, which had enthusiastic conversations on the topics of value, barriers and standards that are expected to continue into 2020 in different forms. In this direction, the project team will increase its coverage by meeting with peak water and energy bodies in the coming months to discuss the standard with utilities across Australia. Work will continue to investigate international standards in water supply and consumption, which may value add to the WESCML standard. Pathways will be developed to ensure the project's future sustainability and impact. Outputs from this project will be drawn into a conference paper expected to be presented in 2020.

Due to the importance of WESC data, the expected impacts of this project are likely to be broad and cross-cutting, particularly as disruptive technologies roll out over the next 5-10 years. Impacts are anticipated beyond data users and providers to government, communities and the

⁷ ARDC, Working with research software, <https://ardc.edu.au/resources/working-with-research-software/> Accessed 7 October 2019.

wider public through improved data-driven planning, modelling of consumer behaviour and WESC services provision. Examples at an economic level may include the collection's value increasing to a national digital asset, adoption of circular economic models and helping data providers focus more on BAU and less on ad hoc data requests. Societal benefits could be diverse; from data-driven local citizen science projects, to P2P trading schemes and improved large-scale precinct planning. These would flow into environment sectors, with example benefits including improved water supply resilience, waterway health and wastewater augmentation.

References

- Simons, B., Yu, J., and Leighton, B. (2016): WESCML: A Data Standard for Exchanging Water and Energy Supply and Consumption Data, 12th International Hydroinformatics Conference, Incheon, South Korea, August 2016., IWA. <https://doi.org/10.1016/j.proeng.2016.07.451>
- Simons, B. and Yu, J. (2015): WESCML Information Model. v1. CSIRO. Data Collection. <http://doi.org/10.4225/08/574D1DEEA50DD>
- Yu, J., Simons, B., Caradoc-Davies, B., Leighton, B., Goh, R. and Inman, M. (2014), AURIN Lens 6/2 Protocols Report. CSIRO, Australia, <http://hdl.handle.net/102.100.100/92973?index=1>.
- Box, P. and Lemon, D. (2015) The role of social architecture in information infrastructure: A report for the national environmental information infrastructure (NEII). CSIRO, Australia, <https://doi.org/10.5072/83/5849a28b08365>.
- Stahel, W. R. (2016) The circular economy. *Nature*, vol. 531, no. 7595, pp. 435–438.
- Smart, A. (2017) Economic value of spatial information in NSW: Estimated for 2017 and 2022. ACIL ALLEN CRCSI, Australia.: <http://www.crcsi.com.au/news/economic-value-of-spatial-information-in-nsw/>.

Report prepared by Michael Rigby, AURIN / The University of Melbourne and Jonathan Yu, CSIRO. Date: 2 December