engineering proceedings

MDPI

Proceeding Paper

WATERVERSE: Strategies in Stakeholder Engagement for the Digitalization of a Water Data Management Ecosystem †

Mollie Torello [1,*], Siddharth Seshan [1], Suze van der Meulen [2] and Lydia Vamvakeridou-Lyroudia [1,3]

[1] KWR Water Research Institute, 3433 Nieuwegein, The Netherlands; siddharth.seshan@kwrwater.nl (S.S.); lydia.vamvakeridou-lyroudia@kwrwater.nl (L.V.-L.)
[2] N.V. Pwn Waterleidingbedrijf Noord-Holland (PWN), 1991 Velserbroek, The Netherlands; suze.van.der.meulen@pwn.nl
[3] Centre for Water Systems, University of Exeter, Exeter EX4 4QF, UK
* Correspondence: mollie.torello@kwrwater.nl
† Presented at the 3rd International Joint Conference on Water Distribution Systems Analysis & Computing and Control for the Water Industry (WDSA/CCWI 2024), Ferrara, Italy, 1–4 July 2024.

Abstract: WATERVERSE utilize Multi-Stakeholder Forums (MSFs) to engage diverse stakeholders from the water sector. Forums foster dialogue to establish a current state and future vision of data ecosystems. Upcoming forums will explores risks, barriers, opportunities, and governance needs. In the Netherlands pilot, PWN developed a chloride prediction model, using WDME for streamlined data processing and dashboard insights. This initial implementation demonstrates the WDME's potential, addressing key stakeholder requirements and supporting data-driven water management decisions with future developments planned through 2025.

Keywords: FAIR; data management; resiliency; stakeholder engagement; sustainability

check for updates

Citation: Torello, M.; Seshan, S.; van der Meulen, S.; Vamvakeridou-Lyroudia, L. WATERVERSE: Strategies in Stakeholder Engagement for the Digitalization of a Water Data Management Ecosystem. *Eng. Proc.* **2024**, *69*, 209. https://doi.org/10.3390/engproc2024069209

Academic Editors: Stefano Alvisi, Marco Franchini, Valentina Marsili and Filippo Mazzoni

Published: 2 December 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

1.1. Introduction to WATERVERSE

WATERVERSE (2022–2025), funded by the EU, strives to create a Water Data Management Ecosystem (WDME) to make data management practices FAIR (Findable, Accessible, Interoperable, Reusable), affordable, secure, and easy to use. Adherence to the FAIR principles will enhance the (re)usability of data and promote interoperability in computational processes [1]. WATERVERSE is proposing results to lower the entry barriers to data spaces, enhance resilience, and develop market opportunities [2].

Seventeen partners, including six case studies (the Netherlands, Germany, Cyprus, United Kingdom, Spain, and Finland), across the EU represent various aspects of the water domain.

1.2. Introduction to PWN Case Study

Puur water and Natuur (PWN) is a drinking water utility located in the Province of North Holland, the Netherlands (NL), which provides drinking water to over 800,000 customers. The company recognizes the emergence of digital water practices and is actively engaging in a digital transformation. A key component of this initiative is for the establishment and development of a data lake and the formulation of a digitalization strategy and implementation roadmap.

The current challenges in data management are mainly based on how to match data from multiple sources and bring them together in an accessible manner for an organization and its stakeholders. Presently, raw data files are shared through email and the processes of validation, cleaning, and pre-processing are performed manually by individual users. PWN faces typical problems related to time series data, including column names, missing values, data–time formats, duplicates, and encoding issues.

Through WATERVERSE, PWN aims to establish advanced dashboarding to provide insights into the chloride concentrations levels in the IJsselmeer. This will assist in making more informed decisions on when to stop the intake of raw water at certain periods of time throughout the year. The end product will analyze and monitor data intakes and model predictions to provide alarming systems within the visualized dashboards.

2. Methodology

2.1. Stakeholder Engagement

In WATERVERSE, stakeholders are involved through Multi-Stakeholder Forums (MSFs) organized in a minimum of three meetings for each case study. The MSFs aim to create an engaging environment for stakeholders to assist in the development of the Water Data Management Ecosystem (WDME). Members of the project team created a preliminary stakeholder map which included the project target groups, the operators and local authorities dealing with the water cycle, enterprises working in Artificial Intelligence (AI) and data science, research centers and universities, innovation hubs/networks/clusters, emergency response services, citizen initiatives, policy makers, and governments.

Each MSF invites stakeholders to discuss specific topics relevant to digital water spaces. The first MSF, which took place in 2023, asked stakeholders to set the scene of the current situation and develop a future vision. The second MSF, set for 2024, will explore the opportunities and barriers with an emphasis on intervention measures, risk perceptions, and end-user acceptance. The third MSF, set for 2025, will focus on implementation support: behavioral changes, policy development, and governance framework.

The WATERVERE MSF framework and moderation techniques build off and extend past EU Horizon projects of NextGen, Ultimate, BINGO, and StopIT [3].

2.2. Technology Assessment

To facilitate the collection of user requirements, the requirements were broken down into terms of data pipelines, describing the steps taken from data collection to storage and use. Questionnaires were sent to each case study on these management requirements. Additionally, a data quality framework was established to set objectives, metrics, procedures, and validation criteria for the WDME.

After collecting requirements from the appropriate users and stakeholders, the requirements across the six case studies were analyzed and integrated into the WDME design [4,5].

3. Results

3.1. Stakeholder Engagement

Fourteen participants across six target groups attended the first MSF at PWN. Key issues identified were the uncertainty and distrust around data management. Little collaboration occurs between stakeholders and there is a noted lack of knowledge and skills within the water data sector. Participants saw benefits in faster, cheaper, and higher-quality data solutions, especially with the pre-existing open data sources. Stakeholders' functional requirements included security by design, an easy-to-use user interface, data processing, traceability, explainability, and expandability. Data requirements included quality labels, data standards (smart data models, anonymization, exchange format, validation), and meta data.

3.2. Technology Assessment

The three use cases for the NL pilot included (1) collection and harmonization of heterogeneous data, (2) sharing of the collected data after their preparation, and (3) federation of external data to make them available to PWN teams. Therefore, two systems were required for data gathering (i.e., sensors/systems, databases) and communications (API, SCADA, etc.). Preconditions to setting up the basic flow included open-source data, which are accessible through API, internal SCADA data, a sandbox environment for running and deploying scripts, and IoT sensors.

The basic PWN flow for the chloride prediction model steps is found in Table 1.

Table 1. Technology assessment results for basic flow to achieve the chloride prediction model.

Step Number	Description [1]
1	Collect relevant raw data from Rijkswaterstaat (RWS) via API
2	Collect raw data from Royal Netherlands Meteorological Institute (KNMI) via API
3	Collect raw data from PWN's IT infrastructure
4	Collection of other raw data
5	Process raw data and quality control
6	Merge all processed data into a single dataset
7	Save resulting dataset into database
8	Run chloride prediction model for scenario length
9	Save model results into a database
10	Data and predictions available to end-users via dashboard updates

[1] Results published for all case studies' descriptions as part of WATERVERSE Deliverable D2.1 [3].

3.3. First Pilot Iteration

In the first pilot iteration, data pipelines were developed for the different use cases that led to the collection of raw data from heterogeneous data sources (such as from KNMI and RWS), which were harmonized to a FIWARE-compliant standardized format, following the NGSI-LD specifications (Figure 1a). Additionally, the harmonized data that are stored in a data portal made available within the WDME were accessed and inputted to PWN's chloride prediction model, providing predictions which were also harmonized in NGSI-LD format to allow seamless usage of the data (Figure 1b). Such an initial implementation showcased the technical solutions that the WDME provides in fulfilling the requirements of key stakeholders and its potential usage as a Water Data Space.

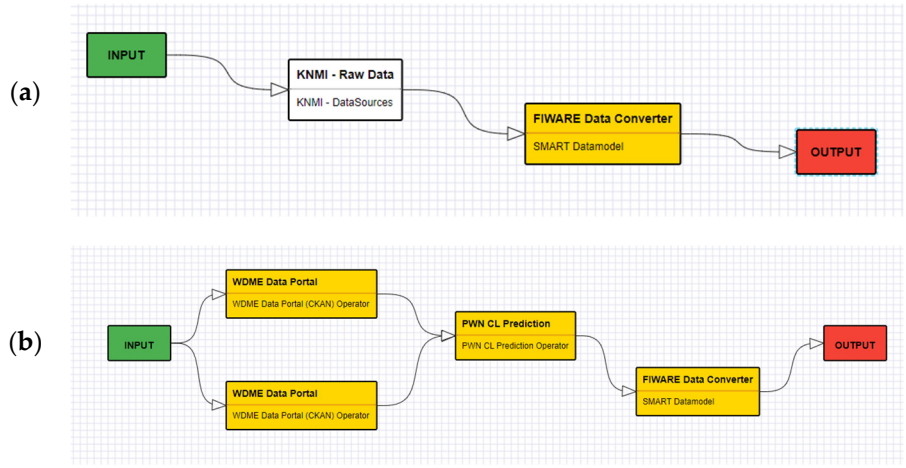

Figure 1. (**a**) Example of WDME data pipeline editor requirements for data input. (**b**) Example of WDME data pipeline editor for generated data via a prediction model.

4. Discussion

Stakeholders that participated in the WATERVERSE NL Pilot shared a strong desire to remain involved. Stakeholders viewed this as a learning experience and desired more microlearning content. Therefore, the core project team has committed to facilitate discussion about key issues (i.e., who owns the IJsselmeer forecasting model, why certain data are not shared publicly, etc.). After completion of the first iteration, the requirements of stakeholders were largely fulfilled from a technological perspective. The exception is real-time data from the PWN IT infrastructure. Through the WDME, more data sources which are relevant to the natural system were integrated. Additionally, the WDME provided a

sharing mechanism of various new insightful data points, such as the prediction mode, to stakeholders within PWN and all relevant external stakeholders.

5. Conclusions

The establishment of the WDME has been meticulously developed through the use of stakeholder input across the six pilot sites. Through the use of stakeholder engagement strategies in WATERVERSE, the first iteration of the developed chloride prediction model was integrated within the WDME to promote the sharing of the results and key insights with various stakeholder groups. Stakeholders' requirements and needs guided the development, including that of the data resources available. As more data sources are integrated and there are further developments within the WATERVERSE project, the provided data management solutions of the WDME coupled with applications such as the prediction model will lead to improved decision making related to the intake of raw water from IJsselmeer.

6. Further Research

WATERVERSE research will continue through 2025. During this time, MSF 2 and 3 will be completed focusing on the risks and barriers and implementation support, respectively. The second pilot iteration, focusing on internally operating the WDME, will occur. Additionally, we will conduct an analysis of overall stakeholder engagement across all case studies focusing on behavioral changes and future governance approaches.

Author Contributions: Methodology, M.T.; software, S.S.; formal analysis, M.T. and S.S.; investigation, M.T., S.S., and S.v.d.M.; resources, S.S.; data curation, S.S. and S.v.d.M.; writing—original draft preparation, M.T. and S.S.; writing—review and editing, L.V.-L.; supervision, L.V.-L.; funding acquisition, L.V.-L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was co-funded by the European Union, grant number 101070262.

Institutional Review Board Statement: This study was conducted in accordance with the ethics procedure of Horizon Europe and ensures full compliance with EU ethical principles and legislations.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: These datasets were derived from the following resources available in the public domain: https://dataplatform.knmi.nl/ (accessed on 28 February 2024) and https://data.overheid.nl/ (accessed on 28 February 2024). The raw data supporting the conclusions of this article for chloride predictions will be made available by the authors on request.

Acknowledgments: The authors thank the entire WATERVERSE consortium for their contributions to the overall project. They also thank the CETAqua team working with WATERVERSE for their contributions in the development of the technology assessment methodology.

Conflicts of Interest: The authors declare no conflicts of interest or commercial conflicts of interest.

References

1. FAIR Principles. Available online: https://www.go-fair.org/fair-principles/ (accessed on 27 January 2023).
2. Water Data Management Ecosystem for Water Data Spaces. Available online: https://waterverse.eu/ (accessed on 18 July 2023).
3. Torello, M.; Frijns, J.; van der Meulen, S.; Wulfert, M.; Haro, J.; Liljanto, N.; Charalambous, S.; Baena-Miret, S.D. 2.2 Stakeholder Engagement Activities and Feedback. WATERVERSE Open Access Publication. Available online: https://waterverse.eu/storage/2024/07/WATERVERSE_D2.2_Stakeholder-engagement-activities-and-feedback.pdf (accessed on 1 April 2024).
4. Baena-Miret, S.; Giménez Esteban, R.; Sarria Montón, M.; Vargiu, E.; Basile, M.; López Aguilar, F.; Seshan, S.D. *2.1 WATERVERSE WDME Design*; WATERVERSE Open Access Publication: Cornwall, UK, 2023.
5. Water Data Management Ecosystem for Water Data Spaces. Available online: https://cordis.europa.eu/project/id/101070262 (accessed on 24 October 2022).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.