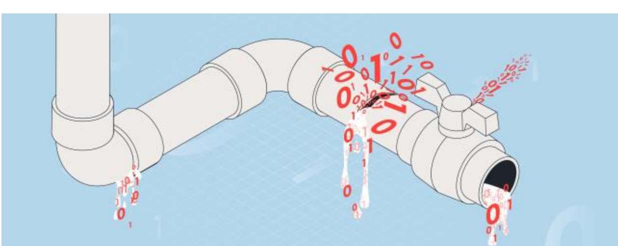
FOCUSING DATA ON SUSTAINABLE WATER DELIVERY

Accomplishing more with less to meet essential human needs and preserve water resources worldwide

Well-run water-system infrastructure is critical to supplying water to communities. However, sustainable provision of water is increasingly jeopardized by aging pipelines as well as excessive consumption, the effects of climate change and a growing and shifting global population.

Circa 70 percent of the Earth is covered in water, but only about 3 percent is freshwater, and less than 1 percent of that amount is available for human use.[1] Uneven distribution and polluted sources where supply is severely limited further strain the ability to meet the needs of humanity.

According to the United Nations, water scarcity can mean scarcity in availability due to physical shortage, or scarcity in access due to the failure of institutions to ensure a regular supply, or scarcity due to a lack of adequate infrastructure.[2] Water scarcity is a pressing and complex issue that water utility operators can help address through data management. Effective data management optimizes operational decision making in the treatment of water, and wastewater, and the distribution of water for diverse purposes, including drinking, bathing, and use by industries and in agriculture.



In the following Q&A, WSP experts discuss how data, digitization and intelligent software jointly provide promising opportunities for utility operators to make supply go further.

WSP experts: Matthew Gallaugher, Director of Technology, Digital, Australia; David Rawlinson, Lead Data Scientist, Australia; Anna Dahlman Petri, Senior Water Advisor, Sweden; and Morten Engedal Sørensen, Business Development Manager, Informatics, Denmark.

What are the key points to understand about effective data management?

David Rawlinson: Data management encompasses many different aspects of data, including how it is collected, stored, shared and used. Different technologies suit different situations, and for data management to be effective, the right tools must be used for each job.

Increasingly, cloud infrastructure enables multiple databases to be part of a single "picture," drawing from many sources. Using accurate and timely data as the basis for a wide range of decisions is a key quality of a digitally mature organization. This is harder than it sounds, because the various data sources were until recently not easily available to decision-makers. Digitization and database projects aim to fix this issue by bringing all the data together to create a "single source of truth"—a coherent picture for all users.

[1] NGWA, The Groundwater Association

[2] United Nations, Water Scarcity

Creating a single, coherent view of data is one of the first prerequisites for projects involving artificial intelligence [AI] and machine learning [ML], as these technologies often draw on data throughout an organization. Frequently, this view creates an immediate and valuable return on investment even before any smart algorithms have been written, simply because of the increased agility for sharing and integrating timely data from various sources.

Typically, a water network operator will use a variety of technologies to complete their digitization strategy. A collection of databases and other sources is known as a data "warehouse"; the term conveys the heterogenous nature of the contents. Although relational databases manage data integrity internally, data integrity is usually an application-layer concern in a data warehouse. Often, databases in a warehouse are replicas of standalone legacy databases that have been brought into the warehouse to provide a single point of access.

A data "lake" differs from a warehouse in that it tends to capture large quantities of less structured raw data, such as timeseries measurements or documents. These might be stored in newer "No-SQL"[3] databases that offer higher performance but fewer guarantees, such as referential integrity and consistency.

Creating and maintaining a clear understanding of the relations between data in different databases is a crucial ongoing task of data management.

Anna Dahlman Petri: Whatever the technologies used, an effective database must be in equal measure updated, accurate and accessible. These three elements can turn a database from a static list into a tool that can be relied on, even in pressured, time-critical situations. Accurate and detailed data is also, of course, the starting point for a digital twin.[4] By supplying modelling software with live data regarding pressure, flow and demand at key points throughout the water network, the operator can simulate in real-time the impact of changes in conditions throughout the water system. This aids in making the best decisions in daily operations and allows utilities to identify and even predict and prevent unwanted events, such as a supply shortage, a pipe break or contamination event.

Morten Engedal Sørensen: To effectively manage and use data for optimizing water systems, the fundamental design of the data model infrastructure is important. The data model used is usually developed in computer aided design systems (CAD). Securing a strong data model is the starting point for effective asset management; a strong data model sets the fundamental rules and logic related to the system water flow.

A digital twin then enables utilities to efficiently manage data concerning the water flow. Data from SCADA, IOT and other surrounding applications can be stored in a structured form as a homogeneous data set and applied to the digital twin. This process ultimately brings about transparency by collecting and displaying the same data to all stakeholders involved in operating the system; it also enhances the usage of artificial intelligence and machine learning.

For a utility company to be able to address current and future challenges, access to high-

[3] SQL: Structured Query Language

[4] A digital twin is a virtual replica of an asset that incorporates associated real-time data during operation of that asset. It provides an immersive and integrated visualization of previously siloed information and enables use of modern digital analysis techniques, such as condition-based monitoring and predictive analysis, to plan for the continued functioning of infrastructure. See "Digital Twins Contribute to Infrastructure Resilience," wsp.com

quality data is crucial. A digital twin equips utilities with the accurate and timely information needed to make decisions and take actions that address urgent issues and prevent future ones from arising.

> On the way to the customer's tap, many water utilities are losing an estimated 20% to 35% of the fresh, treated drinking water they produce.
>
> Source: Nickel Institute [5]

Can you further explore how data management and high-quality data help water utilities address system challenges such as waste?

David Rawlinson: When demand increases but supply is constrained, attention naturally turns to being more careful about waste, such as transmission losses. This means that we need to measure more carefully and pay better attention to anomalies and trends to find and prevent losses. Accurate and detailed data can reduce waste and improve efficiency.

Anna Dahlman Petri: In order to reduce waste, it is necessary to see where waste is occurring or where it is likely to happen in the system. This need points to one of the greatest challenges facing all water utility owners—almost all the water distribution system is underground, making it much more difficult to spot and fix any issues, often requiring costly digging operations and road closures.

In view of this reality, the best operational method is not to work on the physical network first but to create a model, such as a digital twin, of the water distribution network. A high-quality database vastly reduces the cost of creating a model while increasing the usefulness of the model in forecasting anything from probable weak pipes to network performance during predicted peak loads.

Once it is finished, a model is a great tool for developing reliable scenarios for decision making. These scenarios help us to understand possible impacts to the surroundings and what customers might experience in the case of a breakdown in aging infrastructure. Scenario development enables us to be prepared and take measures at the right time and place to mitigate potential damages. The consequences of a network failure might be anything from physical damages due to erosion from underground leakage, to the societal costs of leaving people without access to clean water.

Data management is therefore itself an effective tool to avoid time-consuming investigations and the need to rehabilitate damages. And no tool is better than its input data.

Matthew Gallaugher: The preventative role of data management is significant; more effective management of aging infrastructure can provide high value and reduce risks for asset owners. In fact, the data demands of aging infrastructure may be greater than those of modern infrastructure.

First, aging infrastructure suffers a comparative lack of sensing and measurement devices designed into the system—sensors have become much cheaper and more readily networked, meaning that we can automatically have a continual stream of relevant measurements. In older systems, retrofitted sensors, or ad-hoc measurement procedures, typically make data more indirectly or sparsely acquired. This situation creates a more heterogeneous and complex dataset.

---

[5] Nickel Institute, Lose the Leaks, January 30, 2020

Second, with infrastructure that has aged, there is greater uncertainty about the condition of the infrastructure due to material degradation and damage caused by external environmental factors over time. We generally know less about material quality and processes of older constructions, and quality control was less consistent. This means we need to generate new, additional data to measure these things in-situ or try to infer endpoints directly without having a precise model of the infrastructure.

Both these factors tend to require more complex and bespoke inference to gain a clear and accurate picture of the infrastructure. In a modern system, data streams neatly into a well-designed database with little technical debt. However, in older systems, a patchwork of different technologies provide data to various legacy systems; as discussed earlier, harmonizing data can be a significant challenge. Often, the first task is to provide digital infrastructure—such as databases—that capture all key data sources and present a coherent, unified view. We also need to keep the data up-to-date as new measurements arrive, performing curation and quality control. Frequently, the data remains fragmented, or is only periodically re-integrated with other sources. That is why digital maturity indices emphasize continual demonstration of data provenance, versioning, backup, and disaster recovery. None of these things are possible without good data management.

Morten Engedal Sørensen: Aging infrastructure certainly brings greater complexity. The expected conditions of the infrastructure must be validated from insights based on relevant data collected from different sources.

By correlating data that conveys the age of the pipes with data about soil composition, groundwater level, vegetation, and measured flow, it is possible for data scientists to provide a valid understanding of the condition of the infrastructure. In addition, retrieving data from the surrounding SCADA and IoT systems and applying the logic and rules from the data model paves the way for the application of artificial intelligence and machine learning; this process leads to valuable predictions about the current state of the assets underground and provides vital information for supporting maintenance decisions or replacement activities.

What role does data have in helping utility operators control costs and manage demand?

Matthew Gallaugher: Not all demands are equally important, and demand overall is elastic. Generally speaking, demand emerges from a population of users who have different needs and requirements. Data is critical to understanding these users, when and where their demand varies and how responsive the population is to various price levels at different times. Input costs such as electrical power also vary over time, meaning that optimum management is a juggling act of forecasting demand to adjust pricing. For many reasons, an element of uncertainty is unavoidable, which means that a variety of outcomes must be considered and “priced in” at all times.

Anna Dahlman Petri: A key challenge that water utilities face is they must be able to deliver on peak days when the demand can be twice the average. If that peak could be predicted and managed, the risk of failure would be mitigated, and the size and complexity of the water system could be optimized.

Accurate, real-time data could be used to create price models benefitting all parties—of course, with regulations in place and relevant data available; for example, a forecast can lead to a reduced price during windy or sunny days when sustainable electricity is more abundant, allowing industry to optimize their usage, and thereby helping the overall network.

A good understanding of the demand base and the actual supply can also allow for more accurate tariff rates on a more granular basis. This is especially useful when considering areas with the most-water-intensive industries, which tend to be more responsive to water price adjustments.

Morten Engedal Sørensen: Within the network, the key indicators concerning flows and pressures in all parts of the network are known to the utility operators by collecting data from the many infrastructural components in many different locations. By using the data, it is possible to make corrections to the parameters in the network and thereby take different actions to adjust the water supply to flow to particular areas or sites.

To be cost-effective, it is necessary to predict how much water will be consumed. Data can be used for that prediction; and aligned pressure potentially translates into lower electricity cost for utilities. By using data, and artificial intelligence, operators can optimize pump use and use of other components in the system to deliver water in the most efficient way.

All of the above is only possible when the water supply is sufficient. In case of breakage or leaks, recent events show how local disasters can cause huge losses of water and lead to extreme scarcity of water in regions that normally have a steady supply of water. In such cases, a digital twin can compute an area isolation plan to ensure that the part of the main with the leak or breakage could be isolated to prevent loss of water, while areas that are not affected can remain supplied.

How can artificial intelligence and machine learning contribute to informed decision making?

David Rawlinson: Artificial intelligence, in particular machine learning, has matured enormously in the past 10 years, especially when applied to sensing and perception. This technology contributes to decision making by allowing detailed and comprehensive audits of physical infrastructure, enabling 3D reconstruction, automatic defect detection and other capabilities.

But measurement itself is rarely the end goal. Instead, the objective is to construct a more detailed and accurate picture of current operating conditions to allow optimization of other processes as such maintenance, network operations and risk management. Optimization might involve finding ideal parameters for a policy, calibrating a simulation, or predicting the outcomes of various network configurations or scenarios.

Machine learning can help elsewhere too. It can improve situational awareness by filtering and consolidating alarms and providing tools—such as chatbots—to communicate with users and manage their accounts or recommend services. Hopefully, the picture that's emerging is that artificial intelligence and machine learning is not a separate entity, but a type of technology that can be integrated into many business areas.

Our experience is that the most effective role of artificial intelligence and machine learning is "decision support" for skilled operators and managers, rather than attempting total automation. Real-world systems—particularly featuring complex and geographically distributed infrastructure—tend to experience far too many strange and unanticipated events that cannot be incorporated into more narrowly defined software functions, even with advanced AI. Human subject matter experts are still the best generalists. In an ideal world, artificial intelligence and machine learning removes boring, repetitive and well-defined tasks while providing contextualized, insightful data views that enable experts to perform a "detective" role, exploring and understanding unusual events.

United Nations Sustainable Development Goal 6: Clean Water and Sanitation—Ensure availability and sustainable management of water and sanitation for all

In what ways can data management facilitate wise use of the world's water resources?

Anna Dahlman Petri: The water cycle is the continuous movement of water within the Earth and atmosphere. It is a complex system. Water never really disappears but occurs in different qualities and quantities in different places on earth. How we use it and what for—you could say we borrow it within the cycle—has a great impact on sustainability. Good data management creates the ability to understand future needs and helps all of us "borrow wisely" today, to conserve water and preserve our natural resources.

Morten Engedal Sørensen: When it comes to wise use of water, data can play an important new role.
The demand for water keeps growing along with the growth in world population. Increasing demand combined with pollution and climate change impacts put pressure on nature's process of replenishment within the water cycle. There is plenty of water in the world, but these and other factors are draining the usable freshwater that is accessible to humans.[6]

Besides legislative initiatives to encourage behavioral changes among consumers to avoid overuse, wise use also calls for actions in the utility sector to minimize loss of water within the network. It is imperative to make available water go further and to create and maintain an equilibrium with nature.

From a future-ready perspective, data from a large variety of data sources contribute to disclosing the important processes of the water cycle[7] and thereby the various actions needed to make sure that consumption is in balance with this vital but scarce natural resource.

[6] GreenFacts, Facts on Health and the Environment, Water Resources 4.3

[7] "Quantifying renewable groundwater stress with GRACE," June 16, 2015, AGU – Advancing Earth and Space Science, Water Resources Research

Contacts

Matthew Gallaugher
Director of Technology, Digital
Australia

Matthew.Gallaugher@wsp.com

David Rawlinson
Lead Data Scientist
Australia

David.Rawlinson@wsp.com

Anna Dahlman Petri
Senior Water Advisor
Sweden

Anna.Dahlman.Petri@wsp.com

Morten Engedal Sørensen
Business Development Manager,
Informatics
Denmark

Morten.Engedal@wsp.com

About WSP

WSP is one of the world's leading professional services consulting firms. We are dedicated to our local communities and propelled by international brainpower. We are technical experts and strategic advisors including engineers, technicians, scientists, architects, planners, surveyors and environmental specialists, as well as other design, program and construction management professionals. We design lasting solutions in the Transportation & Infrastructure, Property & Buildings, Environment, Power & Energy, Resources and Industry sectors, as well as offering strategic advisory services. Our talented people around the globe engineer projects that will help societies grow for lifetimes to come. wsp.com