

WSL: How data sharing can identify leakages in vital water infrastructure

Liz Varga, Professor of Complex Systems at Loughborough University Director of the Centre in Artificial Intelligence for the Built Environment (AI4BE) (formerly at University College London)



Introduction

The Water Systems Leakage (WSL) project was led by DAFNI Champion and Strategy Board Member Professor Liz Varga as part of the 2024-2025 data sharing project 'Data Infrastructure for National Infrastructure' (DINI). This exciting programme of work was instigated by the Department for Science, Innovation and Technology (DSIT), aimed at better and safer use of data in research and funded by UK Research and Innovation (UKRI)'s [Digital Research Infrastructure \(DRI\)](#) programme. The 'DRI' is a £129m initiative aimed at developing a research data system that's interconnected, human, FAIR and sustainable.

With almost 20% of water resources in the UK wasted in seepage and escapes, this project used data to enhance leakage detection in water distribution systems, identify barriers to data sharing, and propose solutions that would facilitate cooperation between stakeholders. The research focused on the importance of consistent data collection and standardisation, the creation of Trusted Research Environments and collaborative platforms for data confidentiality, to enable access to infrastructure data for faster detection of leaks.

The DAFNI-DINI project allowed Professor Varga and her team at University College London (UCL) to further explore the efficacy of water systems, building on an existing project with DSIT, which involved creating algorithms for predicting water systems leakage.



WSL: How data sharing can identify leakages in vital water infrastructure

The challenges

Water is a critical infrastructure, providing clean, safe water into homes. However, there are huge barriers in accessing and sharing water data, with concerns regarding data protection and the potential to expose security vulnerabilities.

Previously, water data had only been shared via individual trusted working partnerships. Whilst the team were able to obtain data from water companies, many of those providers were reluctant. **“We had to build relationships and restart earlier conversations in order to convince water companies that exploratory work on new methods for systems leakage management was worthy of investigation,”** explained Professor Varga.

Data security was a major challenge, with companies varying in their concerns regarding data quality and how up-to-date records were.

Professor Varga says, **“The key research gaps became quickly apparent. The absence of a comprehensive method to overcome data-sharing barriers. This was coupled with an absence of tailored data standards for leakage management, and an urgent need for a holistic framework to assess benefits and prioritise barriers.”**

The approach

Fellow DINI project, Icebreaker One was also finding it a challenge to speak to the relevant people and access the right data. DAFNI commissioned IB1 to explore the requirements and impact of supporting an improved sharing of national infrastructure data with publicly funded researchers, focusing on energy, water and transportation.

IB1 devised a framework outlining the many different types of barriers to data sharing, such as commercial sensitivity and legality. The WSL team used this framework and followed the same mixed methodology approach, comprised of a literature review, workshops and interviews.

The literature review involved collecting reports from public databases and water supply companies, then compiling this data to establish data standards and ontology. Journal papers and case studies on water systems leakage were examined and found to include evidence that supported the barriers uncovered by Icebreaker One.

These barriers were further validated via interviews with utilities, academics, and technicians, facilitating discussions on potential solutions, such as AI-based tools for leakage detection.



WSL: How data sharing can identify leakages in vital water infrastructure

Key findings

With the identification of 22 barriers and 19 solutions to data sharing in water systems, the team decided to create a comprehensive spreadsheet containing links to all the different types of data that would be required by companies to reduce leakage. Their findings covered not only leakage data, but also data on repair work, customer complaints and soil type.

The analysis revealed that one fifth of all the companies' future demand will come from reducing their leakage. This would result in significant cost savings for both provider and customer, thanks to a significant amount of water that companies do not have to clean and pump.



Water companies are currently preoccupied with new projects, from building treatment plants to desalination. There is huge concern over keeping demand down with the team's data revealing that the larger the house, the more likely you are to use more water. The use of treated water for hot tubs and pools during the summer drastically increases demand, when rainwater from water butts could be used instead. Even toilet flushing does not require treated water.

WSL: How data sharing can identify leakages in vital water infrastructure

The team spoke to organisations, such as the Open Data Institute and [STREAM](#), based at Northumbria University. Stream is a collaboration between UK water companies, supported by industry and civil society partners with a vision to unlock the potential of water data to benefit customers, society, and the environment. Their missions for open data have so far been unsuccessful due to the barriers identified. However, STREAM has been successful in creating a collaborative network with most of the UK's water companies.

“The public ultimately overpays for leakage in water infrastructure”, explained Professor Varga. **“Using proven methods from Artificial Intelligence to identify and prioritise leakage management leads to direct savings as well as positive environmental impact.”**

Benefits of data sharing

Data sharing would lead to improved water leakage detection, with real-time sensors enabling early and proactive discovery, rather than a reliance on customer reports. It would foster innovation and allow researchers to validate models using real-world scenarios.

Transparent access to leakage data would inform evidence-based decision-making for water companies and contribute to sustainable development goals, such as cost-effective maintenance scheduling. Standardisation of datasets would provide for consistency the replication of solutions across regions and drives forward the potential for mutual aid and exchange between companies.

Barriers to data sharing

The 22 barriers identified covered both cultural and technical barriers such as discoverability and reliability, respectively. GDPR and commercial sensitivities limit access to business data, therefore reducing possibilities for collaboration and transparency. Data access for both academic researchers and the public is often extremely limited, affecting the validation of methods. In addition, sensor data is often incomplete, due to transmission and equipment failure issues, reducing its usefulness for advanced analysis and modelling.

WSL: How data sharing can identify leakages in vital water infrastructure

The main recommendations from the 19 solutions

1. A legal framework must be developed for data sharing agreements that protect commercial interests and allow researchers access.
2. Standardisation initiatives need to be adopted for units, terminology and sampling protocols.
3. Anonymisation techniques and datasets must be used to balance privacy concerns with data availability.
4. Stakeholders need to be trained on the benefits and methods of data sharing, eliminating cultural resistance.

How could this work benefit society as a whole?

The main and immediate benefit would be cheaper water bills as households are currently paying for leaks inherent in existing infrastructure. With leakages stemmed, different uses for water savings could be realised, such as creating green hydrogen for cleaner energy systems.

Next steps

On the DAFNI platform, researchers can access the spreadsheet of data or 'wiki' as termed by the team, as well as the methodology for use cases. A journal paper has been published, which contains proposals for recommendations - 'Addressing data sharing challenges for leakage management in water distribution networks: a multi-criteria decision-based (MCDM) assessment of barriers and solutions', Journal of Environmental Management, 391, September 2025. DOI: <https://doi.org/10.1016/j.jenvman.2025.126481>

The team are now exploring potential collaborations and funding opportunities.

"I think the main impact of the WSL project has been in raising the profile of the genuine challenges in obtaining data about water systems challenges," says Professor Varga. **"The DAFNI platform has great potential to help with sharing data, using standards, ensuring data is securely managed, and that privacy is maintained. DAFNI enables sharing into one trusted research platform for use by multiple universities. Society would reap benefits from such controlled data sharing."**



WSL: How data sharing can identify leakages in vital water infrastructure

Who's involved?

Ruoqing Yin, PhD student at UCL led the project, with Professor Varga, and doctoral students Haonan Xu and Jiaqian Wei, from UCL.

When did the project start and finish?

The project started in August 2024 and completed in January 2025.

