

# Appendix:

## Water company contributions

**Matt Edwards**  
 Anglian Water Head of Enterprise Data, Chief Data Officer.

**“As with any major utility company, Anglian Water constantly wrestles with managing the high volumes and variety of data, content, and documents about its operational assets, sites, and associated health and safety practices.**

There are thousands of employees and partners across a wide and dispersed geography, from so many different disciplines, across operations, maintenance, engineering, asset management, and supply chain management etc., all constantly requiring access to the latest data and information, or potentially needing to update or add to existing data. Just like the assets themselves, much of the data and content has a long lifespan and could have been created many years ago, often pre-dating the connected and mobile working we have today. It’s easy to imagine how data hoarding, silos, duplication, and, worst of all, complete loss of data and documents can occur when people are searching for reliable instant access to information in remote field-based locations, regional company offices, or even a partner or contractor’s office.

With all these challenges in mind, Anglian Water is driving the creation of a virtual library for all of our key operational, maintenance, and engineering reference data and content. We’re doing this not only with the future opportunities of generative AI in mind, but also by using AI capabilities to create a source of trusted, understandable, and findable data and information that generative AI can query

against. We are deploying an AI search and discovery engine to find and profile critical historic and current data, documents, and content across a growing landscape of digital folders, drives, and applications. Once we connect to a new data source, the AI platform learns how to understand the subject of the data and content, and proactively auto-classifies or ‘tags’ the data by the assets, sites, or network area to which the data relates. We’re now using it to consistently classify documents into priority reference types, such as operational manuals, maintenance instructions, or safety and compliance information, that are all the basis of user search terms. The AI also supports identifying the latest, most relevant, and most accessed data and documents, all to understand and promote to users the most trusted master information that everyone should be accessing. Whereas in the past it may have taken hours or longer to find the latest master documents for specific sites, and days or even weeks to collate or refresh the contents into hard copies for holding on remote sites, by moving to digital formats where the master information can be accessed from anywhere in seconds, significant risks are being removed and efficiencies realised. By separating and removing the noise and information risks from old and out-of-date information, Anglian Water is not only realising value for

employees through finding the right information at the right time, but, crucially, is also preparing for generative AI opportunities as well.

Now that we are building a data source we can trust, the next opportunity is beginning to train generative AI to rapidly, consistently, and confidently react to user-specific questions and deliver instant answers. Using natural language Q&A, there are countless opportunities for further increasing the speed and reliability of new information and insight creation – for example, “Tell me which sites have had construction restrictions imposed as a result of asbestos?”, or “Tell me the last date when specific critical equipment was calibrated at a specific site”. The vision is to be able to create our own internal web of information that can support Q&A around all aspects of asset management, across all stages of the asset and operational lifecycle. There are countless opportunities for not only direct savings in time and cost, but also increases in the quality of decision-making at speed and the mitigation of decision-making risks. We see ensuring the base information being queried is as high quality as possible as a key ingredient to successfully exploiting generative AI, and to make that happen across such vast and varied data, we really need AI to deliver as much of the preparation as possible.”



# Anglian Water Case Study: Virtual 'Blue Box'

Water company contributions

**Matt  
Edwards**  
Anglian  
Water Head  
of Enterprise  
Data, Chief  
Data Officer.

Anglian Water's landmark Virtual Blue Box project employs AI-powered search and discovery to create a valuable single source of reliable asset information from reams of unstructured data.

The large blue plastic box is a ubiquitous sight for contractors, managers, engineers and others attending an Anglian Water site. It is the physical container for all of the key information relating to the assets and health, safety, operational and other considerations for on that site. In other words, everything that any stakeholder needs to know needs to be contained in the box. Between sites across the company, each box should have the same scope and quality of information organised in a consistent manner.

As with any company in the utility industries, Anglian Water found that the complexity and size of sites and the data hoarding, duplication and loss behaviours (amongst others) combine to create a perfect storm of data disappearance, degradation and differences. Hence, these blue plastic boxes end up being highly varied in terms of the value that they contain – some entirely meet the scope of the information required; others almost entirely fail to enable stakeholders to start or complete their tasks.

Think of the effort that individuals and teams then have to do – having assumed, perhaps, that they would be able to work with pertinent information right out of that blue box. At best, too much data is challenging to sift through to identify where the critical elements of information are. Too little, missing or outdated data raises concerns for data quality where there is insufficient data to enable effective, safe, efficient decision-making. At worst, there may be incomplete and inconsistent data where critical assets or risks are not detailed – and where the humans are unaware of what they don't know. This situation may only be resolved by an expensive error when, for example, sub-surface assets are damaged by excavation works or an accident occurs.

And yet, all of the data to appropriately populate each blue box does exist within the organisation – in tacit and formal sources. However, finding it may be extremely difficult. Maintaining it is a manageable task but challenging. Ensuring consistency across or within sites, also non trivial. The blue box is a clear example of where automation is the route to success.

The Virtual Blue Box project has culminated in the creation and maintenance of a non-physical repository for all of the key data on assets on sites and that massively important

summary of the health, safety and other considerations. The project began with large scale discovery across the siloes of the organisation. Rather than assuming where data lies by looking where Anglian Water would hope it would be, AI is used to undertake a generalised Asset search pre-filtration. In other words – to ensure that there are no hidden corners or human biases preventing discovery of this essential data.

Whereas in the past it may take weeks or months to collate or refresh the contents of a physical Blue Box, the Virtual Blue Box project has compressed this process into a matter of days. Rather than having a team of individuals focused on sleuthing activities to find data no matter where it may have been (mis) filed – now the AI provides individuals with the information that they need extracted from this data in near real time. The engineer, for example, can be significantly more confident that they know where there are critical assets lying just beneath the ground and where there are no impediments.

While the application of generative AI solutions was not necessarily a primary consideration for Anglian Water's Virtual Blue Box project, what has been achieved already by using AI-powered search and discovery has created a valuable single source of asset information from reams of unstructured data.



# Development of AI-powered decision-making at Yorkshire Water

Water company contributions

**Christopher Smith**  
Yorkshire Water Data Science Manager

Headlines about the Water sector moving into AMP8 rightly focus on the unprecedented investment in assets and infrastructure. While this substantial capital investment will ultimately revolutionise the industry, it will take some time before its positive effects are seen in relation to service, regulatory and environmental outcomes within the AMP. As such, effective operational decision-making will remain essential to improving these outcomes in the near-term. Recent advancements in artificial intelligence, particularly those relating to agentic AI, offer a transformative opportunity to enhance these traditionally manual and subjective processes with powerful real-time intelligence.

To further explore these opportunities, Yorkshire Water's Data Science team has been collaborating with colleagues from our operational control centre to develop an AI-powered decision-making solution. This project has so far delivered a pilot focussed on our wastewater network assets and the prevention of environmental impacts, such as watercourse pollution and storm and emergency overflow activations. The initial output has been the delivery of a wet-weather planning solution, which utilises a mixture of machine learning, statistical and GenAI models to predict which assets are likely to be at risk during these events based on forecast localised rainfall. The output of these predictive models is combined with dynamic situational data, such as asset availability and network capacity to generate a prioritised list of recommended actions, such as mobilising resource and the execution of reactive and preventative maintenance activities, which are then validated and triaged by colleagues in the control centre.

The next step is training the AI engine to discover the most relevant documents relating to any Anglian Water assets and sites, before then using generative AI to rapidly, consistently and confidently prepare user-specific insights. Example benefits will be shown by the automated integration of diverse sources of relevant data. For health and safety, for example, asbestos survey data would be very useful to have tagged to sites and assets with no human intervention – and discovered through natural language Q&A (e.g., "Tell me which sites have had construction restrictions imposed as a result of asbestos?"). For the efficient design and delivery of capital projects or post-incident recovery (e.g., replacing a damaged spill point), being able to access feasibility studies which have been considered and / or implemented on other similar sites would be valuable. Both reduce risks, costs and manual work and rework.

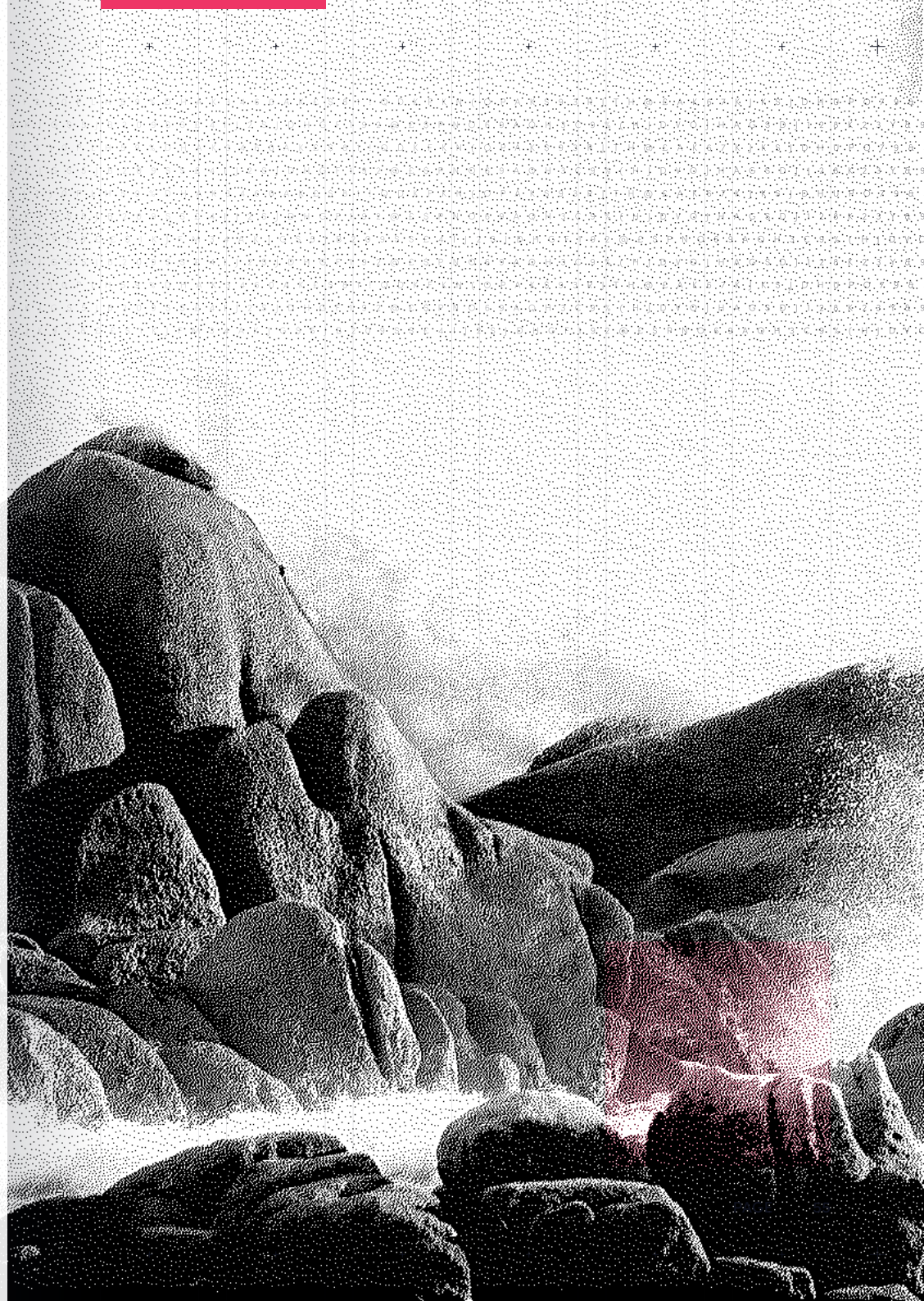


Continuous data and process discovery has been central to the success of this pilot. Any solution aimed at enhancing decision-making must integrate both the intricacies of existing processes and the reality of the data landscape in which they exist. These subjects are complex and situationally dependent, making it challenging to thoroughly explore them within the limited time typically allocated for 'data discovery'. We have benefited from strong collaboration among internal colleagues and a genuine shared sense of ownership of the solution. This has resulted in high engagement throughout the development process and an eagerness to test, validate and adopt any outputs.

Demonstrating proof of value has been a crucial aspect of developing this solution and maintaining stakeholder engagement. Backtesting the solution has been an integral step in this process, involving the application of models to historical high-rainfall and storm events and comparing the predictive outcomes to actual observations. This comparison allowed us to evaluate AI-generated recommendations against interventions enacted through the existing process. The results highlighted the potential value, particularly in mitigating pollution risk, and provided valuable insights into how the solution could be further optimised.

The extensive and continually evolving opportunities presented by AI pose challenges for organisations, especially those accustomed to delivering projects with well-defined scopes and outcomes. Organisations like ours often face difficulties in translating these opportunities into clearly articulated AI/ML problems, as our business use cases are so bespoke and complex. This pilot, along with the iterative development approach we have adopted, has provided us with a framework to effectively leverage AI. Importantly, it has also demonstrated the practical applications of AI to our stakeholders, allowing them to move beyond the hype and understand the genuine benefits. This has allowed us to create a comprehensive roadmap for expanding this solution to other operational processes during AMP8.

This initiative has highlighted how AI can play a leading role in enabling Yorkshire Water to realise its strategic vision of creating a thriving Yorkshire, right for customers, right for the environment. We are excited to see how the industry embraces the potential of AI in AMP8 and beyond.





01

# Foreword by UKWP Director Chris Newsome.

**Our industry is continuously adapting to the increasingly acute pressures of climate change, efficiency, and cost drivers. Customers' expectations of our services and their experiences have never been higher; we need to meet these challenges with reliability, resilience, and excellent value for money for consumers.**

Regulation (e.g., Environmental Information Regulations) and industry initiatives (e.g., Stream Open Data, Spring Innovation Challenge etc.) have gone some way to improving transparency in the industry. While there is more to do, it's only right that

performance shortfalls are highlighted, so that we know where to focus innovation, funding, and our finite knowledge resources, aligning to – and learning from – the UK Government's Industrial Strategy.

Infrastructure investment in the water sector is now the highest it has ever been, and with costs continually rising, technology can accelerate capital delivery, as well as downstream operations. Water bills are increasing in 2025, so in the context of customer affordability, how do we demonstrate value for money with those increased costs?

Automation, connectivity, and digital technologies are transforming the way our industry works: we deliver better services with fewer people; asset management is enhanced through data insights; and our people, customers, and infrastructure are more connected to one another than ever before.

Artificial Intelligence is changing so many aspects of our lives: ever-present on our devices and communications channels, even tracking – and influencing – our behaviours and habits. The question is: What's holding us back from unleashing the full power of AI within the water sector? What should we be

worried about and how can we address those concerns? How can we use this technology to improve transparency and trust with our customers and key stakeholders?

Never has there been more focus on the water industry's responsibility to protect the environment. This extends to the use of AI, where we must consider the 'circular economy' associated with water and energy resource demands through the use of emerging AI technologies.

This discussion paper aims to explore how generative AI can help us respond to our greatest challenges now and create a legacy of knowledge and expertise that's accessible, valuable, and helpful for future generations.

A X 4 E 1 2 ) 0 P L H 5 9 % D H A K 3 4 6 0 A K :  
M 3 4 A K J G S 5 J B 2 0 9 S S E 4 I J 2 S J D H  
A D W K J K 9 2 8 E E S 1 K J F M 3 4 A K J G S 5  
M 3 4 A K J G S 5 J B 2 0 9 S S E 4 I J 2 S J D H  
A X 4 E 1 2 ) 0 P L H 5 9 % D H A K 3 4 6 0 A K :  
M 3 4 A K J G S 5 J B 2 0 9 S S E 4 I J 2 S J D H



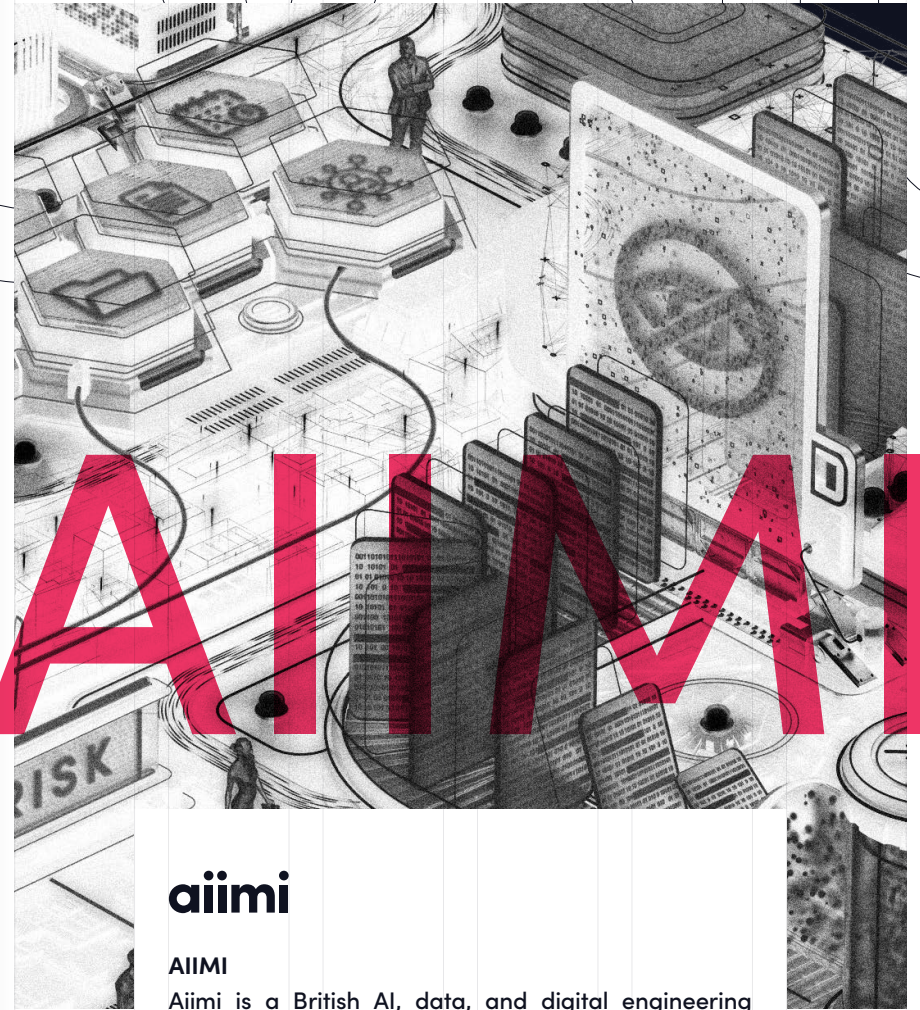
02  
Introductions.

# UKWP



## UKWP

The UK Water Partnership (UKWP) was established in 2015 to unite the UK water sector and the organisations who work within it. Its chief aims are to drive thought leadership on key emerging topics with a view to support new water research, encourage collaboration and innovation, and promote economic growth in the water sector.



# AIIMI

## aiimi

### AIIMI

Aiimi is a British AI, data, and digital engineering company, with long-standing water customers including Anglian Water, Northumbrian Water Group, Welsh Water, Severn Trent, and Yorkshire Water. Through its software and services, Aiimi delivered over 150 projects in AMP7, and has been a member of The UK Water Partnership since 2019. Aiimi is also the technical and data lead on Stream, the sector's Ofwat-funded open data project.



# Contributors:

**Aiimi** —————

Mike Hull  
Michael Jones  
Josh Swords  
Paul Maker  
Steve Salvin  
Sneha Eashwer Singhraj  
David Ford

**The UK Water Partnership** —————

Chris Newsome  
Hans Jensen  
Tom Flood

**Anglian Water** —————

Matt Edwards

**Yorkshire Water** —————

Christopher Smith

# Contact Us

**Aiimi Ltd**

[enquiries@aiimi.com](mailto:enquiries@aiimi.com)  
[www.aiimi.com](http://www.aiimi.com)



**The UK Water Partnership**

[info@theukwaterpartnership.org](mailto:info@theukwaterpartnership.org)  
[www.theukwaterpartnership.org](http://www.theukwaterpartnership.org)



**aiimi**

×



THE UK WATER PARTNERSHIP