

Article

AI-Augmented Water Quality Event Response: The Role of Generative Models for Decision Support

Stephen Mounce ^{1,2,*} , Richard Mounce ³  and Joby Boxall ² ¹ Mounce Hydrosmart Ltd., Beverley HU17 7GD, UK² School of Mechanical, Aerospace and Civil Engineering, University of Sheffield, Sheffield S10 2TN, UK; j.b.boxall@sheffield.ac.uk³ RBMTS Ltd., York YO26 5SX, UK; richard@rbmtrafficsolutions.com

* Correspondence: s.r.mounce@sheffield.ac.uk

Abstract

The global water sector faces unprecedented challenges from climate change, rapid urbanisation, and ageing infrastructure, necessitating a shift towards proactive, digital strategies. Historically characterised as “data rich but information poor,” the sector struggles with underutilised and siloed operational data. Traditional machine learning (ML) models have provided a foundation for smart water management, and subsequently deep learning (DL) approaches utilising algorithmic breakthroughs and big data have proved to be even more powerful under the right conditions. This paper explores and reviews the transformative potential of Generative Artificial Intelligence (GenAI) and Large Language Models (LLMs), enabling a paradigm shift towards data-centric thinking. GenAI, particularly when augmented with Retrieval-Augmented Generation (RAG) and agentic AI, can create new content, facilitate natural language interaction, synthesise insights from vast unstructured data (of all types including text, images and video) and automate complex, multi-step workflows. Focusing on the critical area of drinking water quality, we demonstrate how these intelligent tools can move beyond reactive systems. A case study is presented which utilises regulatory reports to mine knowledge, providing GenAI-powered chatbots for accessible insights and improved water quality event management. This approach empowers water professionals with dynamic, trustworthy decision support, enhancing the safety and resilience of drinking water supplies by recalling past actions, generating novel insights and simulating response scenarios.

Keywords: hydroinformatics; artificial intelligence; machine learning; deep learning; generative AI; data mining; decision support; drinking water quality; incident management



Academic Editor: Tianhong Li

Received: 19 September 2025

Revised: 10 November 2025

Accepted: 11 November 2025

Published: 14 November 2025

Citation: Mounce, S.; Mounce, R.; Boxall, J. AI-Augmented Water Quality Event Response: The Role of Generative Models for Decision Support. *Water* **2025**, *17*, 3260. <https://doi.org/10.3390/w17223260>

Copyright: © 2025 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

The global water sector is navigating a period of profound transformation, driven by the intersecting pressures of climate change, rapid urbanisation, population growth, increasing consumer expectations and ageing infrastructure, the latter of which is not far from being at a critical level in developed western countries such as the UK [1]. These challenges place unprecedented strain on the management of water resources, demanding a paradigm shift from traditional, often reactive, operational models toward more proactive, efficient and resilient strategies. Central to this transition is the concept of digitalisation, which promises to reshape how water utilities monitor, manage and optimise their complex networks. However, the water sector has historically been characterised as ‘data rich but

information poor (DRIP)', a condition where vast amounts of operational data are collected but remain siloed and underutilised, hindering the extraction of actionable knowledge. Water companies are actually interested in information and knowledge, not raw data. Knowledge comes from understanding the information about a subject, and then using it to make decisions, form judgements/opinions or make predictions.

For decades, the discipline of hydroinformatics has sought to bridge this gap by applying computational intelligence, information technology and data science to solve complex water management problems. Hydroinformatics is concerned with the development and hydrological application of mathematical modelling, information technology, data science and artificial intelligence (AI) tools. It provides computer-based decision-support systems that are solutions for areas, such as Smart Networks for leakage management. Early data-driven models, including Artificial Neural Networks (ANNs), Support Vector Machines (SVMs) and Random Forests (RFs), have been applied to various tasks such as prediction, classification and data mining, demonstrating the value of machine learning in the water domain [1]. These tools have provided the foundation for smarter water management, enabling applications in leakage detection, asset condition assessment and demand forecasting. Yet, these model-centric approaches often depend on high-fidelity and structured data, the scarcity of which can limit their effectiveness and scalability [2]. In addition, they remain hard-data centric, missing extensive softer—often linguistic—sources that embed high-value context and tacit expert knowledge which is now being lost to workforce ageing and retirement. With the rapid development of AI tools, it has been suggested that a paradigm shift toward more inclusive data-centric thinking and away from model-centric approaches (based on internal computational architectures and numerical data only) is needed in the broader water industry [3].

Deep learning (DL) sits at the convergence of large-scale datasets and advanced machine learning methods, enabling solutions to problems that are difficult for humans to explicitly specify or fully comprehend [4]. Its key advance is training ANNs (with architectures that can have hundreds of layers) to discover representations in stages: models first capture fine-grained patterns (e.g., strokes or character shapes) and then build up to abstract concepts and classes (e.g., words or objects). This hierarchical representation has reshaped AI over the past decade by supporting pipelines that move from raw data to operational decisions. Widespread adoption has been propelled by data abundance, improved model formulations and notable algorithmic progress. Such deep learning is at the heart of revolutionary technologies such as Generative AI.

The recent and rapid emergence of Generative Artificial Intelligence (GenAI), and particularly Large Language Models (LLMs), represents a fundamental technological inflection point with the potential to empower data-centric thinking and help overcome many traditional barriers [5]. Unlike earlier AI systems that primarily focused on quantitative analysis and prediction, generative models are capable of creating new, original content—including text, code, images and synthetic data—based on patterns learned from existing information [6]. This capability, powered by architectures like the Transformer [7], allows for human-like natural language interaction, sophisticated reasoning [8] and the synthesis of insights from vast quantities of unstructured data, such as technical manuals, operational logs and regulatory reports [5]. Whilst GenAI systems offer powerful capabilities, they are prone to so-called “hallucinations,” defined as instances where the model invents details not supported by data. Recognising this limitation is vital to mitigating risks in applied settings, as we shall discuss. The potential of GenAI to transform knowledge-driven fields is reshaping industries, and the water sector is now beginning to explore its applications, with early adoption and pilot projects beginning to demonstrate the value of GenAI. Figure 1 shows the results of a comprehensive search (using OpenAlex and Crossref across

journals, conferences, preprints, theses, etc.) in publications that mention “generative AI” or related terms (e.g., generative models, GANs) in the context of water resources and hydrology. The bar chart summarises the approximate counts per year (global coverage, all publication types). In 2021, there were only a handful of such works, but by 2025 the number had surged dramatically.

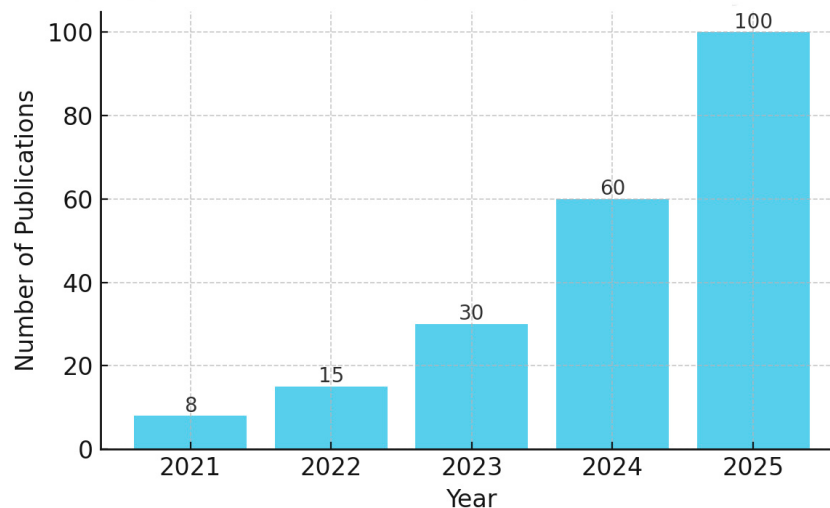


Figure 1. Rise in publications on generative AI in water resources/hydrology (2021–2025). Data gathered via OpenAlex/Crossref API searches (using year filters and keyword queries). 2025 is partial (up to Q3).

A chatbot is a program (powered by ML algorithms) within an app that interacts directly with users to help with simple tasks. ChatGPT allows the user to ask questions in natural language, to which the chatbot responds. Platforms such as OpenAI allow the creation of AI Assistants or specialised “GPTs” for particular domain tasks. Prompt engineering is utilised to craft instructions that guide the AI to generate desired outputs. A key enabling technology for these advanced applications is Retrieval-Augmented Generation (RAG), which grounds an LLM’s responses in external, verified knowledge sources. By retrieving relevant information from a utility’s internal databases or technical documents before generating a response, RAG ensures that the AI’s output is accurate, up-to-date, and context-specific, thereby mitigating the risk of factual errors or “hallucinations” [9]. A significant evolution in this space is the development of agentic AI, where AI systems, or “agents,” are designed to autonomously perceive their environment, make decisions and execute multi-step tasks to achieve specific goals [10]. This moves beyond simple question-answering systems to create a “virtual workforce” capable of orchestrating workflows, providing functionality such as data analysis, report generation and even operational control (with suitable safeguards). The concept of AI agents has been identified by Gartner as the top strategic technology trend for 2025 and holds immense promise for orchestrating the multifaceted operations of a smart water network. According to Futurum Group’s CIO Insights Q2-2025 global survey—covering 204 enterprise CIOs from Fortune 500 and Global 2000 companies across 17 industry verticals—89% consider agent-based AI a strategic priority for automation, decision-making and process orchestration [11]. Figure 2 demonstrates the progression of Generative AI application from improved searches, to chatbots and AI assistants, to more advanced decision support with specialised agents for tasks and workflows.

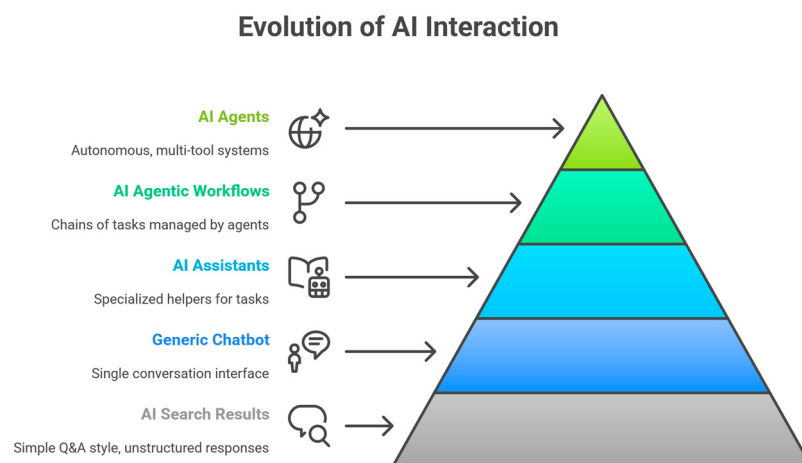


Figure 2. The evolution of AI: from static search to autonomous agents.

After exploring the digital transformation of the water sector and some of the latest work in GenAI, this paper focuses on a critical, high-stakes application area within the water sector: the management of drinking water quality. The provision of safe, clean water is a paramount public health responsibility, and incidents such as contamination events, discolouration or microbiological failures require rapid and effective decision-making to protect consumers and comply with strict regulatory standards. We then explore how LLMs, RAG and agentic AI can develop in the future to create a new generation of intelligent tools. These tools will not only recall past actions but also generate novel insights, simulate response scenarios, draft communications and orchestrate complex, multi-step workflows. Through a discussion of the state of the art and a case study focused on a GenAI-powered chatbot for water quality incident management, this paper explores and reviews the transformative potential of GenAI and LLMs to overcome barriers in the water sector. We demonstrate how these intelligent tools can move beyond reactive systems to enhance decision support in drinking water quality. The paper concludes by presenting a forward-looking roadmap for an AI-augmented future where water professionals are empowered with dynamic, intelligent and trustworthy decision support systems to ensure the safety and resilience of our drinking water supplies.

2. Background: Digital Transformation and AI in the Water Sector

Terms like the “Fourth Industrial Revolution” and the “Second Machine Age” have become shorthand for the sweeping digital transformation across manufacturing, public services and utilities such as water. As digitalisation reshapes daily operations, organisations must redesign workflows to ingest, integrate and act on new data streams. By rethinking processes around data, value can be created end-to-end rather than in silos. In the water sector, these technologies promise meaningful improvements in reliability, efficiency and customer outcomes. Recent analyses suggest that, by 2030, digital transformation technologies, especially AI and platform-based business models, could generate nearly US\$20 trillion in additional global economic value [12], with as much as 70% of the new value created over the next decade being digitally enabled [13].

High-fidelity data is the building block for this transformation. Innovations such as PipeBots are designed to provide this necessary data by enabling a shift from limited, fixed-location (Eulerian) sensing to pervasive, Lagrangian (mobile) sensing within buried pipe networks, generating potentially massive volumes of real-time information [14]. This increased data frequency and geo-distribution provide the granularity required to produce actionable information, thereby enabling the development of site-specific, continually up-

dated predictive models that support a move from reactive to truly proactive maintenance and system performance management. These virtual representations of water systems or “digital twins” of water assets (computational replicas of networks, plants and catchments) can provide operational awareness and near-real-time views of hydraulics and water quality. With continuous sensing and model updates, they can boost efficiency, strengthen resilience, reduce leakage and enable predictive maintenance rather than reactive fixes. More broadly, “digital water” lays the groundwork for utilities to embed data science into everyday practice and to turn dispersed data into coordinated, real-world decisions across the entire system.

In the UK, the infrastructure (pipe networks) for delivering clean drinking water is ageing and deteriorating and, for the water sector in particular, we live off a legacy of past over-engineering. Water utilities are increasingly moving from traditional, reactive operational models toward more proactive, efficient and data-driven strategies. This shift, often termed digital transformation, involves integrating advanced technologies like the Internet of Things (IoT), cloud computing and artificial intelligence (AI) to turn massive volumes of data into actionable knowledge. Water utilities collect vast amounts of data from sources like SCADA systems, GIS databases, asset records and customer contacts. However, this data often remains in isolated “silos,” stored in different formats across separate platforms, which hinders integrated analysis and decision-making. It has been estimated that some water utilities might use as little as 10% of the data they collect, highlighting a significant gap between data availability and its application. The journey toward a smarter water sector has been marked by the evolution of data-driven modelling techniques, not just collecting more data, but integrating it and applying advanced analytics to extract meaningful patterns and insights. To date, most digital efforts have centred on hard numerical data (SCADA time series, network telemetry, etc.), often leaving out text-based operational evidence. Emerging GenAI/LLM/MLLM tools can now parse and align unstructured text and other inputs with the hard numeric data, extracting entities and events with provenance so decisions are richer, faster and more auditable.

For decades, researchers and utilities have applied machine learning (ML) models to address specific challenges. These “model-centric” approaches focused on leveraging algorithms to find patterns in historical data. Areas include the following:

- Prediction and forecasting: Artificial Neural Networks (ANNs), Support Vector Machines (SVMs) and Random Forests have been widely used for tasks like water demand forecasting, water quality prediction and stream flow forecasting.
- Classification and anomaly detection: These models are also used for classification tasks, such as identifying pipes at high risk of failure or detecting anomalies in sensor data that could indicate a leak or contamination event.
- Asset management: In asset management, AI has been used for automated condition assessment. For example, computer vision combined with ML methods like Random Forest has been successfully applied to analyse CCTV footage of sewer pipes, automatically detecting faults like cracks and displaced joints with high accuracy [15].

The recent explosion in data volume, driven by the proliferation of IoT sensors and smart meters, has fuelled the adoption of deep learning [4], a subfield of ML that uses multi-layered neural networks [16,17]. Key architectures include the following:

- Convolutional Neural Networks (CNNs) are adept at capturing spatial patterns, making them ideal for analysing geospatial data like satellite imagery for land use classification or the remote sensing of water bodies.
- Recurrent Neural Networks (RNNs), and particularly their advanced variant Long Short-Term Memory (LSTM), are designed to handle sequential data, making them

state-of-the-art for time-series forecasting of hydrological variables like rainfall, stream flow and water demand.

- Hybrid Models: researchers are increasingly creating hybrid models that combine the strengths of different architectures. A CNN-LSTM model, for instance, can use CNN layers to extract spatial features from weather maps and LSTM layers to model the temporal evolution of a flood event.

While the aforementioned DL models primarily focus on analysis and prediction, the recent emergence of GenAI and Large Language Models (LLMs) represents a new paradigm. GenAI models, particularly LLMs like GPT-5 and Large Multimodal Models (MLLMs) that can process images and text, data, images and video are capable of creating seemingly new, original content (based on learned patterns) to help augment researchers' productivity and to help uncover patterns that were not previously recognised. This capability will transform how the water sector can address data scarcity, manage knowledge and interact with complex systems.

Applications range from enhancing operational efficiency [18] and predictive maintenance to improving customer engagement and knowledge management [5]. For instance, Multimodal Large Language Models (MLLMs) can unlock vast amounts of domain knowledge trapped in physical archives by accurately transcribing handwritten notes and complex forms from legacy documents [19]. This is particularly valuable for creating comprehensive historical records to train AI models for detecting rare events like pipeline leaks or water quality anomalies. GenAI is reshaping traditional engineering by enabling intuitive, natural language interactions with complex systems. LLMs can assist in generating solutions, analysing data or writing code and can be supplemented with technical manuals or integrated with analytical tools, providing context-specific information for knowledge management for areas such as asset management, predictive maintenance and demand management [20].

Some of the more recent academic research that demonstrates the possibilities of GenAI are the following:

- A powerful application of GenAI is in creating high-fidelity synthetic data to augment limited real-world datasets, which is crucial for training robust AI models. This is a crucial benefit over conventional statistical methods when high-fidelity data is scarce. A study directly compared deep generative models (CTGAN, TVAE and CopulaGAN) with a classical Gaussian-copula approach for synthetic water quality data and found the GenAI methods yielded better downstream model performance [21]. A generative AI approach was developed for spatiotemporal imputation and demand prediction in water distribution systems, exploiting sensor data to reconstruct missing data and improve forecasting accuracy [22]. Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs) have been used to generate realistic time-series data for water demand, synthetic acoustic signals for leak detection and patterns of water quality parameters to train anomaly detection systems. A recent study explored using LLMs to generate synthetic hydrological data (e.g., daily rainfall and reservoir levels) by providing prompts engineered with key statistical properties from a real dataset. Models trained on the augmented dataset showed improved forecasting accuracy [23]. Training generative tools like VAEs, GANs and diffusion models on digital twin data facilitates data augmentation, enabling the creation of high-fidelity synthetic water consumption samples which are already labelled to train supervised learning models. These models can further simulate different scenarios, including shifts in water demand, infrastructure failures or severe weather conditions [24].
- A multi-agent LLM framework has been presented to automate water-distribution optimisation [25]. An Orchestrating Agent coordinates knowledge, modelling (EPANET)

and coding agents for hydraulic model calibration and pump-operation optimisation. Tested on Net2 and Anytown, the coding agent performed most accurately across both benchmarks; natural-language reasoning lacked numerical precision, while tool integration improved reliability.

- Projects like WaterER and DRACO are creating standardised benchmarks to rigorously evaluate the performance of different LLMs (e.g., GPT-4, Gemini and Llama) on domain-specific tasks. The Drinking Water and Wastewater Cognitive assistant (DRACO) initiative [26] released a dataset for benchmarking open- and closed-source models and found that state-of-the-art open-weight models can match, and sometimes outperform, proprietary systems, which is promising for utilities seeking secure, on-premise deployments.
- Researchers can use LLMs to perform AI-driven literature mapping. For example, using an LLM and geocoding on 310,000 hydrology papers (1980–2023), a study mapped basin trends, collaborations, and hotspots [27]. Research has surged since the 1990s, shifting from groundwater/nutrients toward climate change and ecohydrology. Activity concentrates in North America and Europe, with rising China/South Asia attention but persistent gaps in extreme-rainfall regions.
- HydroSuite-AI is an LLM-enhanced web application that generates code snippets and answers questions related to open-source hydrological libraries, lowering the entry barriers for researchers and practitioners in the field without extensive programming expertise [9].
- A vast repository of knowledge in the water sector remains locked in non-machine-readable legacy documents like handwritten maintenance logs, scanned reports and old maps. Traditional Optical Character Recognition (OCR) engines like Tesseract struggle with so called “dark data,” especially with complex layouts and handwriting. A case study on Danish water utility documents demonstrated that a Multimodal Large Language Model (MLLM), Qwen-VL, dramatically outperformed Tesseract [18]. The MLLM reduced the word error rate by nearly nine-fold, accurately transcribing handwritten notes and correctly interpreting complex forms in a zero-shot setting.

3. Drinking Water Quality AI Analytics

3.1. Overview

The imperative to deliver safe, clean drinking water and manage wastewater effectively places water quality management at the forefront of the global water sector’s priorities. A US study recorded 7.15 million water-borne illnesses annually, with > 600 k emergency department visits, 118 k hospitalisations, 6630 deaths and an annual economic cost of \$3.33 billion [28]. Water quality is a multifaceted issue, influenced by an increasing array of pollutants, including micro-pollutants and trace contaminants (such as PFAS). Managing these threats is central to public health and environmental quality. Monitoring is inherently expensive and labour-intensive, spanning the entire water cycle from wastewater resource recovery facilities and combined sewer overflows (CSOs) to rivers, lakes and oceans. Many variables crucial for process optimisation, such as microbial activity, cannot be directly measured, necessitating inference from indirect measurements and domain knowledge. Moreover, water quality is often subjected to multiple stressors (e.g., climate change, land use and management practices) that can exert complex synergistic and antagonistic effects, making prediction and management challenging.

Water quality in the UK’s approximately 1500 water treatment works is well understood. The challenge is to monitor water quality in the ~300,000 km of pipe in the UK distribution network. Water quality data in the distribution network is obtained from manual sampling, portable water quality monitors and a relatively small but growing number

of fixed water quality monitors. Customers and regulators are demanding ever greater water quality; however, water utilities are unable to deliver because they have inadequate data and understanding or insights into what is happening in their networks. This critical domain is fraught with complexities on the operational management side, stemming from DRIP [1], the presence of legacy systems, disparate and siloed data sources (e.g., SCADA, GIS, IoT sensors and asset management databases) and varied data formats, often resulting in information silos, limiting real-time monitoring and hindering effective decision-making. Beyond the data layer, water systems function as high-surface-area physical, chemical and biological reactors, where interacting processes (mixing, adsorption/desorption, redox transformations, corrosion and scaling, biofilm growth and sloughing, pathogen die-off/regrowth and natural organic matter dynamics) evolve across multiple spatial and temporal scales. The data fragmentation and process uncertainty create an operational environment where situational awareness is partial, model confidence can be fragile and risk-informed decisions must be made under persistent uncertainty. Network water quality data such as temperature, turbidity, conductivity, pH, etc. is still collected inconsistently and with limited automation. As performance improves and whole-life costs drop, quality sensors could provide routine operational signals and become as ubiquitous as flow meters. For now, analytics remain underdeveloped because real-time water quality monitoring is only sparsely deployed across distribution networks [29,30]. Currently, data/information access is often limited to secure remote platforms only at the premises of the service provider or data generator or on non-standard and expensive portable platforms. Furthermore, critical contextual information, such as the physical and topological relationships between measurements, often remains in human-interpretable formats (e.g., SCADA synoptics) rather than machine-readable ones. Mobile sensing platforms, including robots and drones, could be game changing for providing intelligence for applications, including interpretation and mission planning [14].

3.2. Traditional Machine Learning and Deep Learning in Water Quality

In recent decades, data-driven models, particularly those based on machine learning and deep learning, have been deployed to address specific water quality challenges. These approaches aim to map inputs to outputs (for hard numerical data) without necessarily modelling the underlying physical processes in detail [31].

A range of application areas have been explored, many centring on prediction/forecasting, in the sphere of anomaly detection and for data mining. For water quality parameter prediction, early ANNs were applied to water quality modelling [32]. More recently, an LSTM-based encoder–decoder model has also been shown to predict water quality variables with satisfactory accuracy after data denoising [33]. Further advancements include hybrid encoder–decoder BiLSTM models with attention mechanisms, which outperform state-of-the-art algorithms by efficiently handling noise, capturing long-term correlations, and performing dimensionality reduction [34]. Other techniques such as ensemble decision tree ML algorithms have been applied, including for predicting low chlorine events [35].

In contamination event detection, ML and DL models are widely used for water quality anomaly detection, addressing point, contextual and collective anomalies [31,36–38]. Most data-driven methods suffer drawbacks associated with high computational cost and an imbalanced anomalous-to-normal data ratio, resulting in high false positive rates and poor handling of missing data [23].

Remote sensing data enables ML to serve as an efficient alternative to traditional manual laboratory analysis, offering real-time detection feedback which is crucial for rapid contamination detection [22]. Computer vision techniques are also being employed for tasks such as algal bloom monitoring. Efforts are directed towards using advanced image

description mechanisms to detect and analyse water pollution sources, facilitating prompt intervention and mitigation measures [39].

Self-organising maps (SOMs), a form of unsupervised Artificial Neural Network (ANN) [40], have been applied in data mining for data visualisation and knowledge discovery for water quality. These have been demonstrated for clustering of water quality, hydraulic modelling and asset data over multiple DMAs [41], in relating water quality and age in drinking water [42] and in exploring the rate of discolouration material accumulation in drinking water [43]. Other approaches to data-driven visualisation/mapping and cluster analysis include using k-means (typically with Euclidean distance), hierarchical clustering, distribution models (such as the expectation-maximisation algorithm), fuzzy clustering and density-based models, e.g., DBSCAN and Sammon's projection [44]. Work has also developed techniques to explore correlations (using semblance analysis) between water quality and hydraulic data streams in order to assess asset deterioration and performance [29].

3.3. Generative AI for Enhanced Water Quality Modelling

The advent of GenAI, LLMs and MLLMs offers a paradigm shift in how water quality challenges can be addressed, moving beyond mere analysis and prediction to content creation, advanced knowledge management and robust decision support. Crucially, approaches once confined to hard numerical data can now unlock and integrate “softer” evidence: operator logs, maintenance tickets, customer contacts, images/video and expert narratives via LLM-driven extraction and linking. This multimodal fusion strengthens signal detection and root-cause hypotheses by aligning qualitative cues with quantitative trends, reducing silos and improving interpretability. In practice, provenance tracking and uncertainty annotations helps keep the softer insights auditable while enhancing the value of the hard data. Some of the more recent academic research that applies GenAI in a potable water quality context are the following:

- Synthetic data generation to address scarcity. One of the significant limitations of traditional ML/DL models in water quality is the difficulty in obtaining large amounts of high-fidelity, domain-specific data, especially for rare events like contamination. Generative Adversarial Networks (GANs) are proving useful for creating synthetic data that mimics real measurements, thereby augmenting sparse datasets and improving model robustness [33].
- Contamination detection. A case study involved a novel GAN architecture that generated expected normal water quality patterns across multiple monitoring sites in China's Yantian network. By comparing real measurements against these GAN-generated baseline patterns, the system effectively detected contamination events, demonstrating the model's capability to capture complex spatiotemporal relationships with high detection performance and low false alarm rates [45].
- Multimodal monitoring and visualisation. MLLMs (e.g., GPT-4 Vision, Gemini, LLaVA and QWEN-VL) are being evaluated for hydrological applications such as water pollution management, demonstrating exceptional proficiency in interpreting visual data for water quality assessments. These models can integrate visual and textual data for real-time monitoring and analysis, enabling dynamic, real-time visualisations of complex water quality scenarios, including potential contaminant spread patterns [46].
- AI-assisted information retrieval. LLMs are being integrated into platforms to enable conversational queries about operational data. For instance, Klir's “Boots” chatbot, powered by ChatGPT, can summarise lab results, flag anomalies in water quality readings and assist with report generation to regulators by drawing on the utility's secure internal database [47].

- AI assistants for professionals. LLMs can serve as “water expert models” to assist professionals by answering questions, summarising compliance documents (enhanced decision support) and generating reports or code [48,49]. The WaterER benchmark suite has evaluated LLMs (e.g., GPT-4, Gemini and Llama3) for their ability to perform water engineering and research tasks, demonstrating their proficiency in generating precise research gaps for papers on “contaminants and related water quality monitoring and assessment” and creating appropriate titles for drinking water treatment research. It is expected that the field of hydrology can benefit from LLMs, by addressing various challenges within physical processes [50]. WaterGPT is a domain-adapted Large Language Model for hydrology [51]. The authors curated hydrology corpora, performed incremental pretraining, supervised fine-tuning and enabled multimodal inputs. WaterGPT supports knowledge-based Q&A, hydrological analysis and decision support across water resources tasks, showing improved performance versus general LLMs on domain benchmarks and case studies reported. Another perspective on customising general LLMs into domain-adapted “WaterGPTs” for water and wastewater management appears in [52]. The authors outline methods (including prompt engineering, knowledge and tool augmentation and fine-tuning), discuss dataset curation and ethics and propose benchmarking tasks and evaluation suites. A roadmap highlights reliability, safety and practical use cases across stakeholders.

In summary, AI, particularly with the advancements in generative models, is rapidly transforming water quality analytics (and the water sector in general). From synthetic data generation and advanced knowledge management to enhanced decision support and real-time monitoring, these technologies offer unprecedented opportunities. However, navigating the challenges related to data quality, explainability, ethical governance and workforce development will be crucial to fully realise AI’s potential in ensuring the safety and sustainability of water resources.

4. Case Study: Drinking Water Quality Event Management

4.1. Scope

Drinking water quality is consistently ranked as the top service priority for water companies. Water treatment is consistently produced to a high standard; however, that quality can be compromised during prolonged transit through extensive, ageing and deteriorating distribution networks. Treated water undergoes various changes whilst travelling through the ageing distribution system infrastructure and a link exists between the age of distributed water, persistent sporadic bacteriological failures and poor water quality. In drinking water supply systems, when a hazard occurs from water treatment works to consumers’ taps, it is considered an event. Events do occasionally occur and can be aesthetic, chemical and microbiological failures. Every water company in England and Wales must inform the Drinking Water Inspectorate (DWI) of all events that have affected, or are likely to affect, drinking water quality, sufficiency of supplies and where there may be a risk to health. If quality fails, actions will be flagged for consideration and certain sampling types will be required (3.52 million tests were conducted in England in 2022). For example, in 2023, a total of 533 events were reported to the DWI and by 2024 this had increased to 556 overall events, with 55% of these water quality events reported outside of the control of the water company [53]. The DWI evaluates the companies’ responses to the most significant events on an annual basis via documentary reporting. Existing solutions to the problem of tackling water quality events in water supply systems typically involve manual analysis of graphs and charts and following designated procedures and performing evaluation in conference calls where necessary. Priority will depend on the issue, so that when a non-critical event occurs a control centre operator would look at a set of criteria and

procedures (see Figure 3 for an example of the life cycle of an event caused by a burst main). These approaches are insular to each individual water company with varying types of procedures, working practices and response plans. This process is time-consuming, prone to human error and unlikely to optimally utilise past learning.

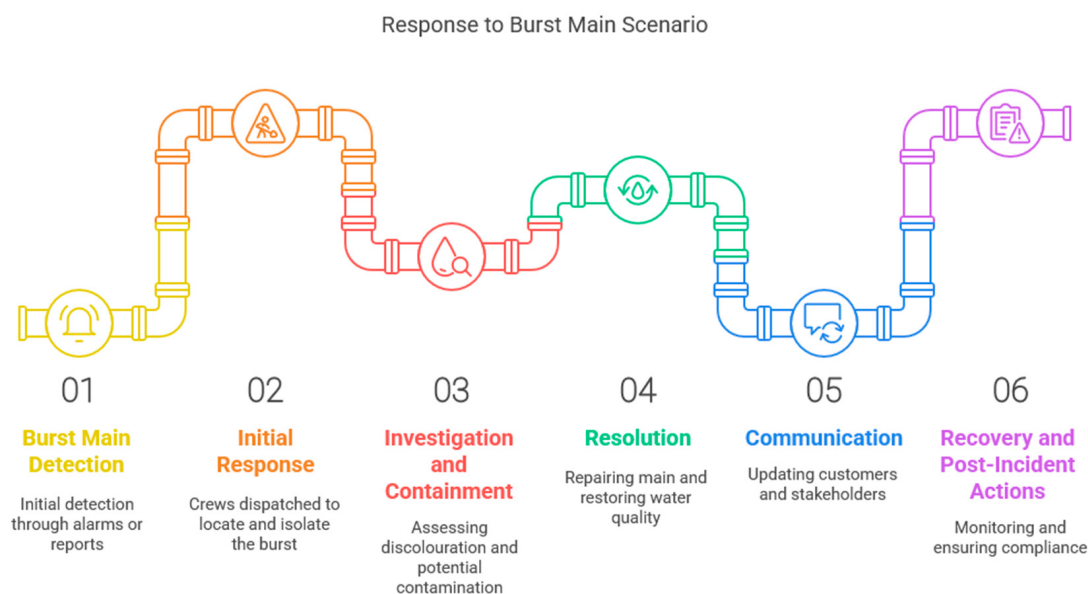


Figure 3. Phases dealing with a significant water quality event.

4.2. ACQUIRE Project

The ACQUIRE project [54] utilises and exploits DWI reports regarding water quality events submitted by water companies to mine the complex knowledge and derive understanding from the information buried within it. These reports form a valuable resource that can be explored and exploited for improved understanding, evaluation and learning for incident management solutions. By data mining many previous water quality event reports, valuable insights into the water quality events can be obtained, facilitating better responses by effectively providing a “triage” view of typical actions from sector-wide responses. This leads to better water quality management, reduced risks, improved regulatory compliance, enhanced customer satisfaction and water industry reputation.

Previous work [55] presented a prototype and findings on how water companies could deal more effectively with drinking water quality events by using information from previous incidents. It explored the use of case-based reasoning (CBR), as a useful tool for providing diagnosis and solutions which can be used for the interpretation of, and guidance on, water quality incidents in water distribution systems. CBR is a classic knowledge-based AI technique that relies on the reuse of past experience. Similar problems from the past have similar solutions and therefore deriving solutions to new problems can be effectively addressed by reusing (and adapting) past solutions. While promising, the early CBR prototype was limited by the need for expert scoring of every DWI finding; a limited historical case base; and a complex user interface which restricted its usage to expert or academic users. A subsequent study [56] expanded this foundation by creating a comprehensive database of over 2300 historical events from 14 years of DWI reports and used deep learning models to automate the sentiment scoring of regulatory outcomes. Further, LLMs were used to create AI assistant “chatbots” to provide accessible insights. Using GenAI can generate decision support insights, aiding efficient repair and response strategies and enhance emergency response in water systems by simulating novel scenarios beyond historical data, thereby improving readiness and adaptive capacity. But it is

important to underscore the need for water professionals to blend AI knowledge with engineering principles [20].

4.3. RAG for Generic Response Plan

LLMs can be both commercial/proprietary (examples: GPT 5.1 (OpenAI), Claude 4.5 (Anthropic), Gemini 2.5 (Google DeepMind) and Grok 4 (xAI)) and open source (examples: Llama 3.3 (Meta), DeepSeek V3.2-Exp and Mistral 7B v0.3). There are various benchmarks for rating the models, with various research groups deciding which benchmarks are suitable, so in practice the choice of model is application-dependent. A number of open source and rapidly developing GenAI delivery platforms and LLMs were used to build and deploy AI Assistant bots with the RAG concept of tuning on appropriate documents (see Figure 4). The RAG approach grounds the LLM's responses in factual, up-to-date information, with a curated and high-provenance document corpus and carefully constructed prompts, mitigating the risk of “hallucinations” [5,57]. RAG enhances the results achieved in both accuracy and reliability by retrieving information from trusted sources. For example, the prompt fed to the LLM can be enhanced with more up-to-date or technical information.

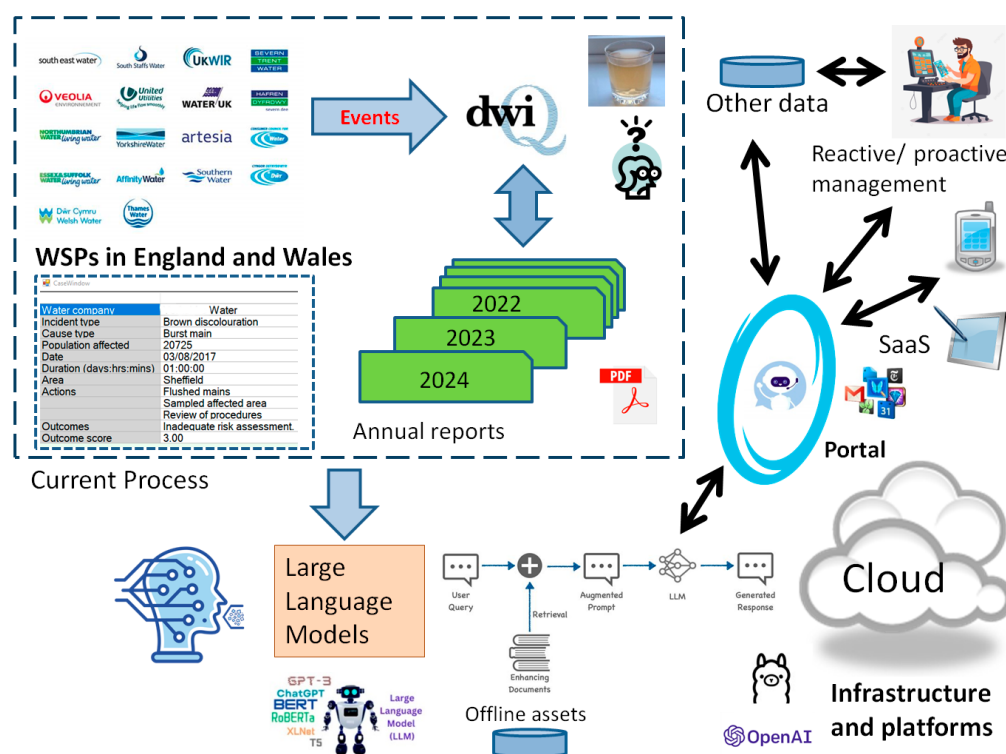


Figure 4. RAG workflow for application.

The system differs fundamentally from a standard, general purpose LLM such as off-the-shelf ChatGPT due to its specialised methodology, using RAG that strictly grounds generation in a curated corpus of DWI regulatory reports. Rather than relying on a general model's latent knowledge, every answer is carefully prompted to retrieve from high-provenance and versioned DWI sources, allowing audit trails. This architecture improves factual fidelity, trustworthiness and explainability, all essential for regulated decision support. Outputs are aligned to UK regulatory language and practices, improving local relevance and compliance in incident rooms and post-event reviews. Generic response plans are assembled by retrieving previously effective interventions and mapping them to current context, thereby accelerating safe and consistent decision-making.

Compared to manual analysis, these approaches offer several significant improvements. Firstly, they enable rapid identification of previously effective interventions for a given unfolding incident, as well as those to be avoided, providing valuable insights for control rooms and water quality scientists. They also aggregate and present actionable information for response plans, improving decision-making and response effectiveness. Furthermore, the tools can facilitate strategic data mining and provide inter-industry learning (including for the training of new staff). They could particularly benefit smaller water companies by providing whole-industry learning and helping to gradually refine response plans over time. Also, there are many water quality events that are not large enough to be reported (in 2024, there were a total of 70,507 consumer contacts in England regarding the acceptability of drinking water—up 12% since 2022), however, that companies must still address through good decision-making. Change needs to be cross-sector, by encouraging industry learning, promoting resilience and informing impact plans [19,58,59]. Event response is often challenged by the rarity of such events, which limits organisational learning, and the gradual loss of institutional knowledge as experienced personnel retire [19]. A notable challenge is the loss of “institutional memory” and ensuring that the best possible practice is being effectively shared across the sector [1]. Since the reports are public, originating from all companies, they are effectively “open data,” engendering data sharing and alleviating fears about data security and confidentiality by creating a transparent supply chain of datasets for analysts. Figure 5 shows an example of generating a response plan using RAG. A detailed sixteen-phase plan is produced starting with part 1—Immediate Incident Response.

Response Plan AI Assistant



Hi! Please briefly describe the water quality incident type and cause

Cryptosporidium incident caused by a possible faulty valve affecting a reservoir supplying approximately 40,000 residents



For a Cryptosporidium incident caused by a possible faulty valve affecting a reservoir supplying approximately 40,000 residents, a structured generic response plan based on best practices and lessons learned from historical incidents reported to the Drinking Water Inspectorate (DWI) would include the following actions:

1. Immediate Incident Response

- Isolate the affected reservoir by removing it from supply to prevent further distribution of potentially contaminated water.
- Bypass the reservoir if possible to maintain supply continuity.
- Increase chlorine residuals in the distribution system to provide additional disinfection.
- Issue a Boil Water Notice (BWN) or Do Not Drink Notice promptly to all affected consumers, with clear communication on the reason and expected duration.
- Provide alternative water supplies such as bottled water or water bowlers.

Describe the incident type and cause



Developed by ACQUIRE

Figure 5. Using the AI assistant to generate a response plan.

GenAI tools using such structured data with RAG can improve water quality management and help the water sector become more efficient and cost-effective, while protecting public health. LLMs can be fine-tuned as powerful query and reporting tools in water sector operations by capturing institutional knowledge.

4.4. Results and Findings

The implementation of the ACQUIRE AI Assistant for Response Plans within a real-world UK water company provided empirical grounding and concrete evidence demonstrating how a domain-adapted GenAI system differs functionally from a standard, generalised LLM system. Unlike a general chatbot system, ACQUIRE is enhanced using RAG to mine knowledge specifically from regulatory reports, steered by focused prompts and offering trustworthy, sector-aligned decision support for critical water quality events. Three months of feedback collected from a control room yielded specific insights into its unique benefits in a local operational context.

The tool's most significant benefit lay in supporting inexperienced staff with limited operational understanding, acting as a functional roadmap or starting point. During an incident caused by a storm, the tool proved particularly helpful for new personnel when a less experienced incident manager faced complications involving treatment failures and power issues. The ACQUIRE tool was used to immediately put high-level initial actions on the manager's radar, facilitating the incident management team's discussion on mitigation strategies, including alternative water, communications strategy and water quality testing. Conversely, experienced incident response managers utilised the tool in an augmentation capacity, running "in the background" to double-check potential actions and compare them against their own initial responses. This application highlighted ACQUIRE's ability to provide recommendations that might otherwise have been overlooked by highly experienced personnel, such as during the planning for a sewage treatment works (STW) shutdown.

Methodologically, ACQUIRE proved valuable in streamlining the maintenance of physical documents, specifically feeding its recommendations into the revision of incident response plans (which are kept as hard copies). It was used to update generic response plans for incidents like a "burst drinking water main" and "water treatment works incident response plans". This included systematically outlining both immediate considerations (e.g., locating the burst, isolating the section and assessing water quality impact) and secondary considerations (e.g., notifying regulators, managing road closures and assessing resilience). Empirical feedback further confirmed that, while the basic response plan was "good," it lacked value for experienced individuals whose actions were already planned or underway. The core value of ACQUIRE in the local context is providing a reliable, regulatory-aligned skeletal structure or action plan based on aggregated institutional knowledge, which can then be elaborated upon by other systems or human experts as needed, confirming its unique function as a domain-specific decision support foundation.

5. Discussion

5.1. Context

A key AI development is the distinction between generative AI and agentic AI. While generative AI functions as a sophisticated pattern-matching and completion system, agentic AI marks a fundamental evolution. Agentic AI systems operate with a degree of autonomy, capable of setting goals, formulating strategies and adapting their approach. This makes agentic AI more akin to a proactive colleague than a reactive assistant. For instance, while GenAI might answer a specific query on optimal chlorine dosing levels based on historical data, an agentic AI could proactively monitor water quality sensors for emerging risks, forecast potential KPI breaches, recommend pre-emptive set-point adjustments and schedule maintenance interventions. Human-in-the-loop (HITL) and human-on-the-loop (HOTL) approaches represent crucial safeguards as AI autonomy increases, especially in the regulated environment of water quality management. These paradigms ensure accountability, transparency and ethical alignment—by reserving human oversight (HITL) or supervisory control (HOTL), balancing efficiency with responsibility. For example, if

a traditional AI tool were to raise a burst alert based on abnormal flow signatures to the control room, a GenAI could then synthesise telemetry and customer contact data to narrow down the location and an agentic AI could then also recommend practical next steps, such as identifying repair materials required, team mobilisation options and traffic management considerations (all pending human approval).

AI is evolving beyond simple tools to become “agentic colleagues”, capable of processing low-level tasks, anticipating next steps in workflows and collaborating. Operationally, agentic AI fosters “smarter scale,” enabling smaller teams to achieve significant output by dramatically increasing efficiency, rather than just growing the headcount (empowering humans to shift from manual tasks to higher-value work, whilst AI agents handle repetitive duties).

Virtual employees of the future may possess corporate roles and network access with greater autonomy than current task-specific agents. This evolution necessitates that organisations fundamentally rethink cybersecurity strategies and access management to accommodate AI identities operating across multiple corporate systems. In industries like water management, agentic AI could optimise energy usage, manage networks and predict maintenance needs all with minimal (HITL), or potentially without, human intervention (HOTL).

The water sector is notoriously slow and resistant to adopt new technologies, partly due to existing ossified workflows. Coupled with an adverse data sharing culture, innovation adoption can be extremely slow and incremental. Further, despite the current buzz, AI technology can be perceived as a “threat to jobs” (as with all sectors). Water utilities are beginning to deploy LLM-powered chatbots to provide customers with instant, context-aware support and engagement and to equip internal teams with tools that can quickly summarise compliance documents or retrieve maintenance procedures. Major industry bodies, including the American Water Works Association (AWWA) and the Water Environment Federation (WEF), have initiated research to establish best practices for GenAI, signalling a sector-wide move toward adoption. Forward-looking water utilities, including those in the UK, the EU and North America, have started piloting generative AI and LLM solutions in real operations. A notable case is DC Water (Washington, D.C.’s water utility) which has actively encouraged its staff to use tools like ChatGPT for daily tasks such as writing reports, coding and analysing data, establishing an internal community of practice to explore GenAI’s potential. They have also applied them for improving customer engagement (e.g., chatbots for customer queries). This underscores that large utilities are beginning to deploy LLMs as knowledge management and decision support tools internally, albeit with cautious governance. Some utilities are also exploring custom LLMs, training smaller models on internal reports, maintenance logs and design manuals, so that staff can query these via chatbot and instantly retrieve answers with context, saving staff time on tedious tasks. In essence, the technical trend is towards hybrid systems: pairing LLMs with structured data sources, or combining physics-based models with generative ML, to leverage the strengths of both (under human oversight).

A white paper [60] published by the UK Water Partnership (UKWP) highlights how AI is poised to revolutionise the water industry and describes several real-world AI/GenAI use cases being developed. It highlights that, if AI is pointed at unstructured data without first improving the dataset standard, spurious results will likely follow. They recommend only pointing LLMs at curated data—by considering “fine-tuning” capabilities, i.e., by a human resource that is capable of training models appropriately.

In 2025, UK Water Industry Research (UKWIR) launched an AI-powered search tool using a customised large-language model (Llama) trained on their own library of reports to provide contextual search and generative AI summaries specific to water research [61]. The

UKWIR's tool is hosted onsite and trained exclusively on UKWIR's own library of reports and tools.

5.2. Roadmap

Figure 6 provides a high-level roadmap for considering when to deploy GenAI in real-world applications in the water sector.

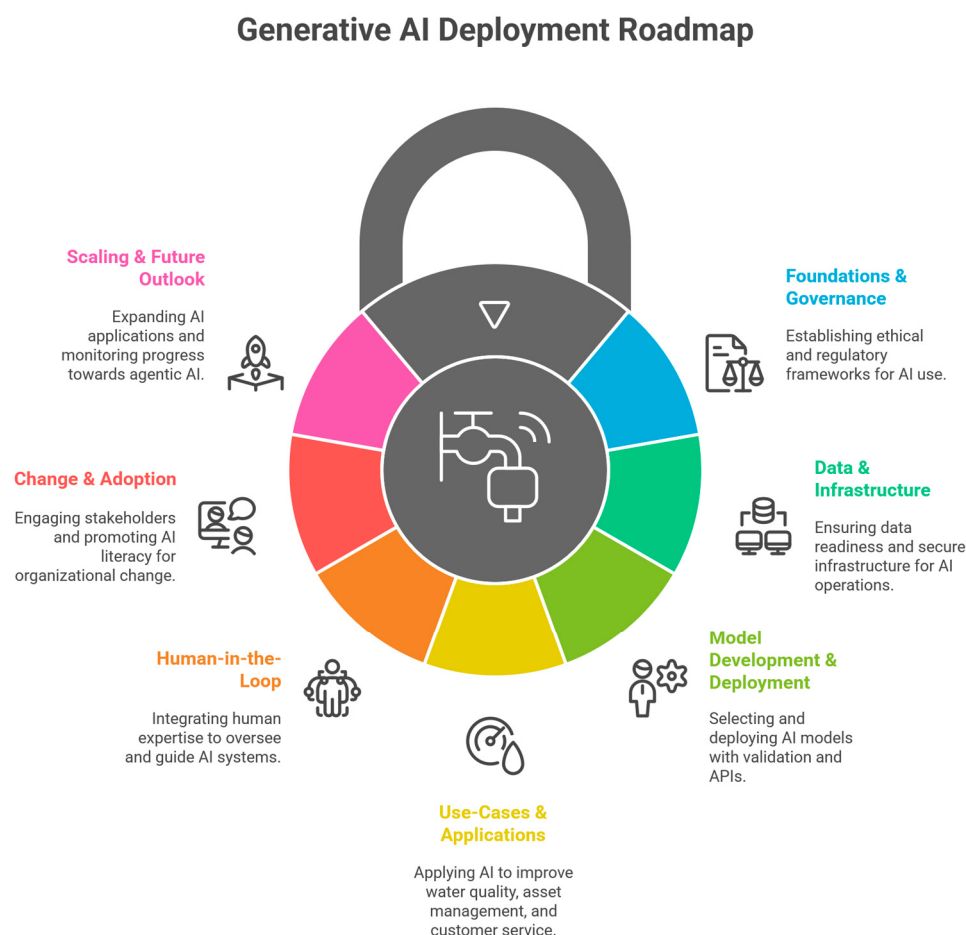


Figure 6. Roadmap for GenAI adoption in the water industry.

Key areas:

- **Foundations and governance.** Foundations and governance provide the strategic bedrock for GenAI deployment in the water sector. Establishing a clear business case is essential, ensuring that investment delivers measurable benefits in efficiency, resilience or customer service. Alignment with regulators safeguards compliance, while ethical frameworks address bias, fairness and transparency. Building trust in AI requires proactive communication and embedding responsible AI practices from the outset. Crucially, success rests on positioning AI as a partner to human expertise, enabling collaboration rather than substitution.
- **Data and infrastructure.** Data readiness entails ensuring quality, availability and interoperability across diverse sources such as SCADA, GIS and asset management systems [62]. Establishing targeted governance frameworks, focused on specific use cases, supports compliance without stifling innovation. Cloud, edge or hybrid architectures must be evaluated for scalability, latency and resilience, with cybersecurity protections built into every layer. With so much data available, it can be overwhelming to identify the most valuable insights for decisions. If data is machine-readable in

knowledge repositories then humans can be AI assisted to be more productive and effective. By addressing data fragmentation and prioritising metadata and standards, utilities can build robust pipelines that allow AI to process, contextualise and integrate operational data into actionable intelligence.

- **Model development and deployment.** Generative AI is emerging as a “game changer” by democratising access to complex AI tools through natural language interactions, enabling intuitive communication with complex systems and reducing the steep learning curve associated with traditional software and advanced analytics [19]. Model development involves the systematic selection, adaptation and validation of generative AI tools for water applications. Choosing between commercial and open-source Large Language Models requires balancing transparency, licencing and support. Fine-tuning models with sector-specific documentation enhances contextual accuracy. RAG offers an effective mechanism to reduce hallucinations and ground outputs in verified sources. Rigorous benchmarking and validation ensure robustness and reliability, while modular deployment strategies—whether on-premises, cloud-based or via APIs—facilitate integration into operational workflows and maintain flexibility as foundation models evolve.
- **Use-cases and applications.** The true value of GenAI lies in its application to priority use cases that deliver tangible benefits. When considering GenAI, keep an open mind—it is not all about chat functions. For example, AI can be incorporated into reporting/dashboard tools or to explore how automation and agentic AI could be used to enhance business processes and decision-making. In water quality monitoring, AI can synthesise laboratory data, regulatory limits and historical trends to flag emerging risks. Asset management and predictive maintenance are strengthened by analysing sensor data to anticipate failures and optimise investment. Customer service can be enhanced through AI-enabled chatbots that provide clear responses to queries. Furthermore, regulatory compliance activities benefit from automated report generation, structured audit trails and streamlined document retrieval, collectively driving efficiency, transparency and resilience.
- **Human-in-the-loop.** HITL/HOTL approaches ensure that AI deployment remains guided by expert judgement and domain-specific oversight. Rather than replacing operators, GenAI functions as a co-pilot, augmenting their capacity to interpret complex datasets and make informed decisions. Structured frameworks for human oversight safeguard against errors and reinforce accountability. Developing hybrid expertise, engineers proficient in both water systems and AI, will enable organisations to critically assess outputs. Fusion skills, particularly judgement integration, support the evaluation of AI suggestions for trustworthiness and relevance, ensuring that GenAI systems complement, rather than supplant, human analytical and ethical capabilities. This collaborative intelligence, where human ingenuity combines with AI systems, consistently outperforms either humans or machines working alone and helps address and mitigate job fears as human engineering experience is still essential.
- **Change and adoption.** Successful GenAI integration demands comprehensive change and adoption strategies. Users need to be involved in the process as the process: it should be about what people need from AI, not the technology itself. Integration and change management are key to ensuring any new AI solution is properly adopted and embedded in processes. Stakeholder engagement spanning regulators, operators and customers can foster legitimacy and confidence. Upskilling programmes in AI literacy will equip staff to use tools effectively, reducing resistance and building organisational trust. Building systems that excel at preserving institutional knowledge by processing historical records and creating accessible repositories of operational

expertise is particularly valuable during workforce transitions. Change management must also address cultural concerns, including fears of redundancy or technological disruption, by framing AI as an enabler of higher-order and more productive work. Incremental adoption through pilots builds confidence and delivers demonstrable value, while embedding AI into business processes requires strong leadership to champion innovation and sustain momentum across the organisation.

- **Scaling and future outlook.** Scaling GenAI from pilots to sector-wide adoption requires careful planning and sustained monitoring. Early projects should focus on quick wins, delivering measurable improvements that secure organisational confidence. Continuous retraining and adaptation are necessary to maintain accuracy, particularly as new foundation models and regulatory requirements emerge. Over time, adoption can progress from reactive generative tools towards agentic AI systems capable of autonomous goal-setting and adaptive strategy.

5.3. Challenges

Generative AI is poised to play an expanding role in the water industry, but its successful widespread adoption will require navigating several challenges and innovating in key areas. Challenges relate not only to technical feasibility but also to regulatory, ethical and environmental considerations. Addressing these issues will be essential for ensuring that AI fulfils its potential as a transformative force in water management.

A primary challenge lies in the development of domain-specific models. While the prospect of water-focused foundational models trained on regulations, design standards and operational datasets is compelling, progress is hindered by data scarcity and fragmentation. The localisation issue, particularly the under-representation of non-English documents and local operational contexts, directly contributes to AI model bias and regional inequity in AI performance. Models trained predominantly on data and standards from specific regions may struggle to generalise effectively when deployed in diverse global water systems. This problem is compounded by the inconsistency in quality standards across utilities and geographical areas. Such variation critically impacts data labelling, as the definition of acceptable water quality or a system anomaly may differ by region, preventing the creation of standardised, high-quality labelled datasets. Consequently, these inconsistencies also undermine model validation and benchmarking processes, making it difficult to rigorously compare AI performance or ensure that a model meeting compliance standards in one regulatory environment is trustworthy in another. Collaborative data-sharing initiatives, such as the EU's WATERVERSE platform [63], alongside standardisation efforts, will be crucial to enable the creation of robust and representative training corpora. Without such measures, model bias and limited generalisability remain persistent risks.

The application of generative AI for real-time decision support further illustrates the promise and complexity of this technological shift. Embedding models into treatment plant controls or distribution network monitoring could enable the predictive optimisation of dosing regimes, emergency response and adaptive network management. However, scalability and reliability are critical barriers: models must function across city-wide sensor networks without sacrificing speed or accuracy. Moreover, the conservatism of the water sector, rooted in its public health responsibilities, demands rigorous validation. HITP/HOTL oversight will remain indispensable until automated systems demonstrate consistent trustworthiness. Governance through this oversight can be retained while automation executes guard-railed, pre-approved actions, with post-action audit trails to build trust and expand autonomy over time. Moving beyond alarm-driven operations, GenAI enables proactive management: forecasting emerging risks, recommending pre-emptive set-point adjustments and scheduling interventions before KPIs drift toward a breach. In a proactive mode,

digital twins and prescriptive analytics could continually simulate “what-if” scenarios, ranking interventions by impact on compliance risk, resilience and costs.

Closely tied to these technical hurdles is the issue of explainability and trust. The black-box nature of many AI systems, particularly ANNs, is a long-standing issue. The opacity of deep learning systems poses difficulties in regulated environments, where transparency and accountability are essential. Explainable AI techniques will therefore play a pivotal role in translating complex model outputs into interpretable insights, enabling experts to validate predictions, trace contaminant pathways and design appropriate interventions. The absence of interpretability risks undermining regulatory approval, operator confidence and public trust.

The ethical, social and regulatory dimensions of AI adoption further complicate implementation. A secure regulatory environment is crucial, as legal grey areas and ethical quandaries regarding data breaches, copyright infringement or DEI rules can lead to cautious industries like the water sector to adopt a “wait-and-see” approach. Data privacy, accountability for AI-driven outcomes and the prevention of inequitable service allocation are pressing concerns. If training data embeds structural biases, AI systems may inadvertently prioritise certain communities exacerbating disparities. Regulators are likely to introduce new requirements, including auditability, certification and ethical guidelines for AI deployment in critical water services. Proactive governance frameworks will be necessary to align innovation with principles of equity and safety.

Another emerging issue is the resource efficiency of AI systems themselves. Training and running large generative models consume substantial energy and water, particularly for cooling data centres. Estimates suggest that individual queries to state-of-the-art models already entail measurable water use, raising questions of sustainability. Adopting “green AI” principles (such as energy-efficient architectures, solar-powered data centres and intelligent scheduling of computation) will be vital for reconciling AI’s benefits with environmental stewardship. There are even opportunities for innovation at the water–compute nexus, for example, through utilities hosting edge infrastructure cooled by treated wastewater.

Equally important are the human and workforce dimensions. Generative AI is unlikely to replace water professionals, but it will augment their roles by automating routine tasks and providing decision support. This shift will require significant upskilling and a cultural reframing of AI from threat to collaborator. Structured training programmes, integration of AI into engineering curricula and communities of practice will be essential in preparing the workforce for collaborative intelligence environments.

6. Conclusions and Future Work

The water sector is currently undergoing a period of profound digital transformation, driven by intersecting pressures from climate change, urbanisation and ageing infrastructure. Traditional data-driven methods, including machine learning (ML) and deep learning (DL), have provided necessary foundations for smart water management, particularly in prediction and forecasting. These conventional models, however, are fundamentally model-centric and often constrained by the scarcity of high-fidelity, structured data.

Although the digital uptake is increasing, the sector’s conservative approach means it still lags other industries in integrating smart technologies. Two key persistent sector-specific challenges are the persistence of siloed and underutilised operational data and the gradual loss of institutional knowledge as experienced personnel retire. This paper argues that the recent emergence of Generative AI (GenAI), Large Language Models (LLMs) and Retrieval-Augmented Generation (RAG) represents a paradigm shift toward data-centric thinking, moving beyond the limitations of model-centric approaches. By tapping into data assets, an accessible repository of aggregated institutional memory can be created

that transcends the boundaries of individual utilities, promoting cross-sector learning and resilience. GenAI unlocks vital “softer” evidence, including operator logs and customer contacts, through multimodal fusion of inputs like text, images and video. This integration strengthens root-cause hypotheses and provides richer, faster and more auditable insights for decision support.

The proposed roadmap provides a high-level plan for GenAI adoption across seven key areas, including governance, data infrastructure and rigorous model deployment. This strategy emphasises aligning investments with regulatory compliance and using RAG to ensure accuracy and reduce hallucinations. Crucially, the plan guides the sector’s evolution from reactive generative tools toward agentic AI systems capable of autonomous goal-setting and adaptive strategy over time. Scaling these systems requires navigating significant challenges related to data, governance and skills. In a sector characterised by public health responsibilities, adopting human-in-the-loop (HITL) and human-on-the-loop (HOTL) oversight remains indispensable to ensure accountability, transparency and consistent trustworthiness. This supervisory control paradigm can guide the safe expansion of AI autonomy, allowing automation to execute pre-approved actions while reserving human judgement for critical decision points.

The case study provides concrete evidence of the practical value derived from the LLM/RAG application. The findings confirm the successful mining of complex knowledge from DWI reports to generate accessible and trustworthy decision support for water quality events. Key empirical insights include the following:

1. Support for inexperienced staff: The ACQUIRE tool proved most beneficial to new and inexperienced staff with limited operational understanding, serving as an effective roadmap. During a storm incident, for example, the tool helped a less experienced incident manager immediately identify high-level initial actions necessary for managing treatment failures and power issues, spanning alternative water provision and communications strategies.
2. Augmentation for experts: For experienced incident response managers, the tool functioned in an augmentation capacity, running in the background to double-check potential actions and validate their initial responses, demonstrating its value even to highly experienced personnel for decision support.
3. Regulatory alignment and document streamlining: The GenAI-derived outputs proved reliable enough to be actively fed into the internal revision of hard-copy incident response plans for generic events like a “burst drinking water main” and “water treatment works incident response plans”. This capability highlights that the system effectively provides a reliable, regulatory-aligned skeletal structure or action plan based on aggregated institutional knowledge. Specific examples of outputs included systematically outlining both immediate considerations (e.g., locating the burst and isolating the section) and secondary considerations (e.g., notifying regulators and assessing resilience).

The practical application confirmed that the system’s primary strength in the local context is providing a reliable knowledge foundation. Furthermore, the use of GenAI/RAG, which is focused on leveraging aggregated public data (DWI reports), helps water companies address the high volume of smaller water quality events that limit organisational learning.

The water sector is not generally at the forefront of the technology adoption curve and thus not often seen as an attractive industry when compared to other cutting-edge industries for data science professionals. With AI becoming more prevalent across industries, there is a growing need to make it broadly available, accessible and applicable to engineers and scientists with varying specialisations. Engineers, not just data scientists, will drive the experimentation and adoption of AI in water industry applications. The

complexity of larger datasets, embedded applications and bigger development teams will drive solution providers towards interoperability, greater collaboration, reduced reliance on IT departments and higher productivity workflows. The key to realising these benefits will be addressing the challenges head-on: ensuring data quality, upskilling the workforce, maintaining ethical guardrails and sharing successes and failures openly across the water community. With thoughtful implementation, generative AI and LLMs could become invaluable tools in building a more resilient, efficient and smart water future.

The water sector is at a pivotal moment. Innovators are disrupting the old business models and there is little place for complacency in the water industry, with threats such as decentralised and distributed technology arising. Disruptive innovation can be beneficial. The evolution from traditional data analytics to the sophisticated capabilities of generative and agentic AI offers unprecedented opportunities to build more resilient, efficient and sustainable water systems. Realising this future will require a concerted effort to address the foundational challenges of data, skills and governance, fostering a collaborative ecosystem where technology enhances, rather than replaces, human expertise.

Looking forward, the water sector's adoption of AI is evolving beyond reactive generative tools toward agentic AI. Strategic integration of AI will move beyond the hype cycle, driving innovation in business models and operational practices. The rise in agentic AI promises more autonomous workflows, capable of proactively managing networks and responding dynamically to changing conditions.

Other future work will focus on developing domain-specific foundational models to overcome data scarcity and fragmentation as well as HITL/HOTL oversight and explainable AI techniques to address trust and accountability in regulated environments.

Synthetic data generation offers a solution to the scarcity of high-fidelity training datasets, particularly for rare events. As enterprise AI becomes increasingly democratised through open-source initiatives, accessibility and adaptability will improve, reducing barriers to adoption across organisations of varying sizes.

In parallel, adaptive organisational cultures will be required to sustain these innovations. Institutions must embrace continuous learning, experimentation and regulatory engagement to safely integrate AI. Transparency, accountability and ethical safeguards will remain foundational to ensuring public confidence in AI-enhanced water services.

Ultimately, realising the full potential of AI in water management hinges on cultivating a collaborative intelligence environment. This requires developing hybrid expertise through water professionals skilled in engineering principles and AI literacy, to critically assess and integrate AI outputs. By prioritising ethical governance, continuous upskilling and transparent deployment practices, GenAI and LLMs can move from novel applications to become trusted and indispensable collaborators/co-pilots in advancing the safety, efficiency and resilience of global drinking water supplies.

Author Contributions: S.M. primarily wrote the manuscript, conducted the review and developed Generative AI tools in the case study. R.M. contributed with design and developing tools in the case study and J.B. contributed on concepts and water quality expertise. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ofwat Discovery Challenge Innovation Fund grant ACQUIRE. Ofwat (the Water Services Regulation Authority) is the statutory body that regulates water and wastewater companies in England and Wales. No formal grant number was assigned to this award. No APC was charged.

Data Availability Statement: The original contributions presented in this study are included in the article. Further enquiries can be directed to the corresponding author.

Conflicts of Interest: Author Stephen R. Mounce was employed by the company Mounce Hydrosmart Ltd. Author Richard B. Mounce was employed by the company RBMTS Ltd. The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

References

1. Mounce, S.R. Data Science Trends and Opportunities for Smart Water Utilities. In *ICT for Smart Water Systems: Measurements and Data Science*; Han, D., Mounce, S., Scozzari, A., Soldovieri, F., Solomatine, D., Eds.; Handbook of Environmental Chemistry; Springer: Cham, Switzerland, 2020.
2. Bettin, A. Digital transformation for the water industry: How a data-driven business intelligence platform can improve operations. *Water Pract. Technol.* **2023**, *18*, 1599–1607. [\[CrossRef\]](#)
3. Zolghadr-Asli, B.; Ferdowsi, A.; Savić, D. A call for a fundamental shift from model-centric to data-centric approaches in hydroinformatics. *Camb. Prism. Water* **2024**, *2*, e7. [\[CrossRef\]](#)
4. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Taiwo, R.; Yussif, A.-M.; Zayed, T. Making waves: Generative artificial intelligence in water distribution networks: Opportunities and challenges. *Water Res. X* **2025**, *28*, 100316. [\[CrossRef\]](#)
6. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Networks. *arXiv* **2014**, arXiv:1406.2661. [\[CrossRef\]](#)
7. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 5998–6008.
8. Bubeck, S.; Chandrasekaran, V.; Eldan, R.; Gehrke, J.; Horvitz, E.; Kamar, E.; Lee, P.; Li, Y.; Lundberg, S.; Nori, H.; et al. Sparks of artificial general intelligence: Early experiments with GPT-4. *arXiv* **2023**, arXiv:2303.12712. [\[CrossRef\]](#)
9. Nishisako, S.; Saito, Y.; Nakayama, A.; Yasuda, H. Reducing hallucinations and trade-offs in responses from generative AI chatbots using retrieval-augmented generation (RAG). *JMIR Med. Inform.* **2025**, *13*, e27830.
10. World Economic Forum & Capgemini. *Navigating the AI Frontier: A Primer on the Evolution and Impact of AI Agents*; World Economic Forum: Geneva, Switzerland, 2024.
11. Futurum Group. *The Rise of Agentic AI: The Leading Solutions Transforming Enterprise Workflows in 2025*; Futurum Group: Austin, TX, USA, 2025.
12. International Data Corporation (IDC). The Global Impact of Artificial Intelligence on the Economy and Jobs. IDC. 2024. Available online: <https://my.idc.com/getdoc.jsp?containerId=prUS52600524> (accessed on 25 September 2025).
13. World Economic Forum. Digital Transition Framework. World Economic Forum. 2023. Available online: https://www3.weforum.org/docs/WEF_Digital_Transition_Framework_2023.pdf (accessed on 25 September 2025).
14. Mounce, S.R.; Shepherd, W.; Boxall, J.B.; Horoshenkov, K.; Boyle, J. Autonomous Robotics for Water and Sewer Networks. In *Hydrolink 2021/2*; International Association for Hydro-Environment Engineering and Research (IAHR): Madrid, Spain, 2021; pp. 55–62.
15. Li, Y.; Wang, H.; Dang, L.M. Vision-Based Defect Inspection and Condition Assessment for Sewer Pipes: A Comprehensive Survey. *Sensors* **2022**, *22*, 2722. [\[CrossRef\]](#)
16. Kumar, P.T.; Mishra, K. Deep learning in hydrology and water resources disciplines: Concepts, methods, applications, and research directions. *J. Hydrol.* **2024**, *628*, 130458.
17. Fu, G.; Jin, Y.; Sun, S.; Yuan, Z.; Butler, D. The role of deep learning in urban water management: A critical review. *Water Res.* **2022**, *223*, 118973. [\[CrossRef\]](#)
18. World Economic Forum & Accenture. *AI in Action: Beyond Experimentation to Transform Industry*; World Economic Forum & Accenture: Cologny, Switzerland, 2025.
19. Ramachandran, A.; Neergaard, T.F.B.; Maier, A.; Bayer, S. Digitization Framework of Water Utility Documents for Digital Twins. In Proceedings of the CCWI 2025—21st Computing & Control for the Water Industry Conference, Sheffield, UK, 1–3 September 2025.
20. Sela, L.; Housh, M.; Vairavamorthy, K. Making waves: The potential of generative AI in water utility operations. *Water Res.* **2025**, *272*, 122935. [\[CrossRef\]](#)
21. Chi, M.Y.; Koo, C.H.; Huang, Y.F.; Chan, W.D.; Pang, J.Y. Artificial Intelligence Generated Synthetic Datasets as the remedy for Data Scarcity in Water Quality Index Estimation. *Water Resour. Manag.* **2023**, *37*, 6183–6198. [\[CrossRef\]](#)
22. Velayudhan, N.K.; Devidas, A.R.; Savić, D.A. Generative AI for spatio-temporal multivariate imputation and demand prediction in water distribution systems. *Results Eng.* **2025**, *27*, 106178. [\[CrossRef\]](#)
23. Wala, J.; Rüppel, U. Leveraging large language models for synthetic data generation for AI applications for water infrastructure management. In *EG-ICE 2025: AI-Driven Collaboration for Sustainable and Resilient Built Environments Conference Proceedings*; Moreno-Rangel, A., Kumar, B., Eds.; University of Strathclyde Publishing: Glasgow, UK, 2025; ISBN 978-1-914241-82-6. [\[CrossRef\]](#)

24. Taloma, R.J.L.; Cuomo, F.; Communiello, D.; Pisani, P. Machine learning for smart water distribution systems: Exploring applications, challenges and future perspectives. *Artif. Intell. Rev.* **2025**, *58*, 120. [\[CrossRef\]](#)
25. Wang, J.; Fu, G.; Savić, D.A. Leveraging Large Language Models for Automating Water Distribution Network Optimization. *Water Research*. 2025. Available online: <https://www.sciencedirect.com/science/article/pii/S004313542501440X> (accessed on 25 September 2025).
26. Taormina, R.; Kyritsakas, G.; Sourlos, N.; van der Werf, J.A.; Soodan, A. DRACO: A Large Language Model Initiative for Urban Water Systems. In Proceedings of the CCWI 2025—21st Computing & Control for the Water Industry Conference, Sheffield, UK, 1–3 September 2025.
27. Miao, C.; Hu, J.; Moradkhani, H.; Destouni, G. Hydrological Research Evolution: A Large Language Model-Based Analysis of 310,000 Studies Published Globally Between 1980 and 2023. *Water Resour. Res.* **2024**, *60*, e2024WR038077. [\[CrossRef\]](#)
28. Collier, S.A.; Deng, L.; Adam, E.A.; Benedict, K.M.; Beshearse, E.M.; Blackstock, A.J.; Bruce, B.B.; Derado, G.; Edens, C.; Fullerton, K.E.; et al. Estimate of burden and direct healthcare cost of infectious waterborne disease in the United States. *Emerg. Infect. Dis.* **2021**, *27*, 140–149. [\[CrossRef\]](#)
29. Mounce, S.R.; Gaffney, J.W.; Boulton, S.; Boxall, J.B. Automated data driven approaches to evaluating and interpreting water quality time series data from water distribution systems. *ASCE J. Water Resour. Plan. Manag.* **2015**, *141*, 1–11. [\[CrossRef\]](#)
30. Gleeson, K.; Husband, S.; Boxall, J.B. Multi-parameter multi-sensor data fusion for drinking water distribution system water quality management. *AQUA—Water Infrastruct. Ecosyst. Soc.* **2025**, *74*, 451–465. [\[CrossRef\]](#)
31. Li, Z.; Ma, W.; Zhong, D.; Ma, J.; Zhang, Q.; Yuan, Y.; Liu, X.; Wang, X.; Zou, K. Applications of machine learning in drinking water quality management: A critical review on water distribution system. *J. Clean. Prod.* **2024**, *481*, 144171. [\[CrossRef\]](#)
32. Skipworth, P.J.; Saul, A.; Machell, J. Predicting water quality in distribution using artificial neural networks. *ICE J. Water Marit. Energy* **1999**, *136*, 1–8. [\[CrossRef\]](#)
33. Tripathy, K.P.; Mishra, A.K. Deep learning methods in hydrology and water resources: A critical review. *J. Hydrol.* **2024**, *628*, 130458. [\[CrossRef\]](#)
34. Bi, J.; Zhang, L.; Yuan, H.; Zhang, J. Multi-indicator water quality prediction with attention-assisted bidirectional LSTM and encoder-decoder. *Inf. Sci.* **2023**, *625*, 65–80. [\[CrossRef\]](#)
35. Kyritsakas, G.; Boxall, J.; Speight, V. A data-driven predictive model for disinfectant residual in drinking water storage tanks. *AWWA Water Sci.* **2024**, *6*, e1376. [\[CrossRef\]](#)
36. Eliades, D.G.; Vrachimis, S.G.; Moghaddam, A.; Tzortzis, I.; Polycarpou, M.M. Contamination event diagnosis in drinking water networks: A review. *Annu. Rev. Control* **2023**, *55*, 420–441. [\[CrossRef\]](#)
37. Gong, J.; Guo, X.; Yan, X.; Hu, C. Review of urban drinking water contamination source identification methods. *Energies* **2023**, *16*, 705. [\[CrossRef\]](#)
38. Dogo, E.M.; Nwulu, N.I.; Twala, B.; Aigbavboa, C. A survey of machine learning methods applied to anomaly detection on drinking-water quality data. *Urban Water J.* **2019**, *16*, 235–248. [\[CrossRef\]](#)
39. Kadiyala, L.; Mermer, O.; Samuel, D.J.; Sermet, Y.; Demir, I. The Implementation of Multimodal Large Language Models for Hydrological Applications: A Comparative Study of GPT-4 Vision, Gemini, LLaVa, and Multimodal-GPT. *Hydrology* **2024**, *11*, 148. [\[CrossRef\]](#)
40. Kalteh, A.M.; Hjorth, P.; Berndtsson, R. Review of the self-organizing map (SOM) approach in water resources: Analysis, modelling and application. *Environ. Model. Softw.* **2008**, *23*, 835–845. [\[CrossRef\]](#)
41. Speight, V.; Mounce, S.R.; Boxall, J.B. Identification of the causes of drinking water discolouration from machine learning analysis of historical datasets. *Environ. Sci. Water Res. Technol. (R. Soc. Chem.)* **2019**, *5*, 747–755. [\[CrossRef\]](#)
42. Blokker, E.J.M.; Furnass, W.R.; Machell, J.; Mounce, S.R.; Schaap, P.G.; Boxall, J.B. Relating water quality and age in drinking water distribution systems using self-organising maps. *Environments* **2016**, *3*, 10. [\[CrossRef\]](#)
43. Mounce, S.R.; Blokker, E.J.M.; Husband, S.P.; Furnass, W.R.; Schaap, P.G.; Boxall, J.B. Multivariate data mining for estimating the rate of discoloration material accumulation in drinking water distribution systems. *J. Hydroinf.* **2016**, *18*, 96–114. [\[CrossRef\]](#)
44. Sammon, J.W. A nonlinear mapping for data structure analysis. *IEEE Trans. Comput.* **1969**, *C-18*, 401–409. [\[CrossRef\]](#)
45. Li, Z.; Liu, H.; Zhang, C.; Fu, G. Generative adversarial networks for detecting contamination events in water distribution systems using multi-parameter, multi-site water quality monitoring. *Environ. Sci. Ecotechnol.* **2023**, *14*, 100231. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Samuel, D.J.; Sermet, Y.; Cwiertny, D.; Demir, I. Integrating vision-based AI and large language models for real-time water pollution surveillance. *Water Environ. Res.* **2024**, *96*, e11092. [\[CrossRef\]](#)
47. WaterWorld. 2023. Available online: <https://www.waterworld.com/smart-water-utility/press-release/14295040/klir-unveils-chatgpt-integration-for-its-water-utility-management-platform> (accessed on 25 September 2025).
48. Pursnani, V.; Erazo Ramirez, C.; Sermet, M.Y.; Demir, I. HydroSuite-AI: Facilitating hydrological research with LLM-driven code assistance. *Earth ArXiv* **2024**, 8121. [\[CrossRef\]](#)

49. Samuel, D.J.; Sermet, M.Y.; Mount, J.; Vald, G.; Cwiertny, D.; Demir, I. Application of large language models in developing conversational agents for water quality education, communication and operations. *Water Pract. Technol.* **2025**, *20*, 2094–2109. [CrossRef]
50. Pursnani, V.; Sermet, Y.; Kurt, M.; Demir, I. Performance of ChatGPT on the US fundamentals of engineering exam: Comprehensive assessment of proficiency and potential implications for professional environmental engineering practice. *Comput. Educ. Artif. Intell.* **2023**, *5*, 100183. [CrossRef]
51. Ren, Y.; Zhang, T.; Dong, X.; Li, W.; Wang, Z.; He, J.; Zhang, H.; Jiao, L. WaterGPT: Training a large language model to become a hydrology expert. *Water* **2024**, *16*, 3075. [CrossRef]
52. Xu, B.; Wu, G.; Li, Z.; Xu, G.; Zeng, H.; Tong, R.; Ng, H.Y. Towards domain-adapted large language models for water and wastewater management: Methods, datasets and benchmarking. *npj Clean Water* **2025**, *8*, 82. [CrossRef]
53. DWI. *Drinking Water 2024: The Chief Inspector's Report for Drinking Water in England*; Drinking Water Inspectorate: London, UK, 2025.
54. Ofwat. 2024. Available online: <https://waterinnovation.challenges.org/news/discovery-winners/> (accessed on 25 September 2025).
55. Mounce, S.R.; Mounce, R.B.; Boxall, J.B. Case-based reasoning to support decision making for managing drinking water quality events in distribution systems. *Urban Water J.* **2016**, *13*, 727–738. [CrossRef]
56. Mounce, S.R.; Mounce, R.B.; Boxall, J.B. Case-based reasoning and Generative AI decision support tools for managing drinking water quality events in distribution systems. In Proceedings of the 21st International Computing and Control for the Water Industry (CCWI2025), Sheffield, UK, 1–3 September 2025.
57. Leyden, T.; Hart-Davis, G. *Agentic AI For Dummies®*, *Integrail Special Edition*; Agentic Platforms and APIs: Global Market Outlook; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2025.
58. International Water Association (IWA). *Global Trends & Challenges in Water Science, Research and Management: A Compendium of Hot Topics and Features from IWA Specialist Groups*, 3rd ed.; International Water Association: London, UK, 2022.
59. Daniel, I.; Ajami, N.K.; Castelletti, A.; Savic, D.; Stewart, R.A.; Cominola, A. A survey of water utilities' digital transformation: Drivers, impacts, and enabling technologies. *npj Clean Water* **2023**, *6*, 51. [CrossRef]
60. The UK Water Partnership & Aiimi. *AI Within Reach (White Paper)*. The UK Water Partnership. 2025. Available online: <https://www.theukwaterpartnership.org/publications/ukwp-and-aiimi-ai-white-paper> (accessed on 25 September 2025).
61. Water Magazine. *AI Tool Transforms Access to UKWIR Research*. Water Magazine. 2025. Available online: <https://www.watermagazine.co.uk/2025/08/05/ai-tool-transforms-access-to-ukwir-research/> (accessed on 25 September 2025).
62. Chaudhary, M.; Pandey, S. Water Sector 5.0: Harnessing digital transformation opportunities and mitigating risks. *Water Resour. Manag.* **2025**, *39*, 2463–2477. [CrossRef]
63. 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. [CrossRef]

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.