THE Water Research FOUNDATION

American Water Works Association

wef

WHITE PAPER

Project 5321

GenAI as a Catalyst for Water Sector Transformation: Summary of WRF Project 5321 Findings

How GenAI/LLMs Can Help Solve Challenges Facing Water Utilities Globally

September 2025

Executive Summary

Purpose

The Water Research Foundation (WRF) project 5321, *The Role of Generative AI (GenAI) for the Global Water Sector*, brings together researchers from the American Water Works Association (AWWA), Water Environment Foundation (WEF), and Karmous Edwards Consulting to build a foundational understanding of the role that GenAI powered by public large language models (LLMs) could play for municipal water and wastewater utilities.

The project connects major sector-wide initiatives. It supports AWWA's Water 2050 vision of advancing innovation for a sustainable and resilient water future, while also aligning with WEF's Circular Water Economy initiative, which emphasizes reducing, recovering, and regenerating resources across the water cycle. By integrating both perspectives, WRF project 5321 contributes to the water sector's ongoing evolution and its ability to meet future challenges.

This research project aims to increase awareness, share knowledge, and facilitate experimentation with GenAI/LLMs to support municipal water and wastewater utilities in adopting and gaining value from GenAI/LLMs. This project explores practical applications of GenAI/LLMs to improve efficiency and operational outcomes.

The four objectives of this project are to help utilities develop a foundational understanding of GenAI/LLMs techniques by:

- Building a common vocabulary and definition of terms. The project explores key concepts such as retrieval augmented generation (RAG), Agentic AI, the role of unstructured data, data cleansing and analytics, optical character recognition (OCR), voice applications, and computer vision
- Identifying practical applications of GenAI/LLMs using available public and open source LLMs via well-documented use cases.

- Developing useful guidelines for global water utilities on their GenAI journey.
- Developing a research roadmap for the future role of GenAI/LLMs in the water sector.

AI Concepts and Definitions

At its core, WRF project 5321 will share information about GenAI/LLMs to raise awareness among water professionals showcasing how existing GenAI tools powered by public facing LLMs such as ChatGPT-5, Claude, Gemini, etc., can be safely and effectively leveraged. Through regular knowledge exchanges, webinars, and workshops, the research team gained insights into how GenAI/LLMs can support daily operations, optimize workflows, and enhance decision-making.

GenAI refers to a subset of artificial intelligence (AI) models that can generate new content, rather than just classifying data from existing data sets or answering questions based on existing information. These models can create new content in the form of text or other forms of media based on patterns and examples they have been trained on.

AI is a broad field in computer science that focuses on developing smart machines capable of performing tasks that typically require human intelligence. These tasks include problem-solving, pattern recognition, decision-making, and language understanding. AI encompasses various subfields, including machine learning, deep learning, natural language processing, and transformer-architecture-based LLMs, each contributing to advancements in automation and intelligent computing.

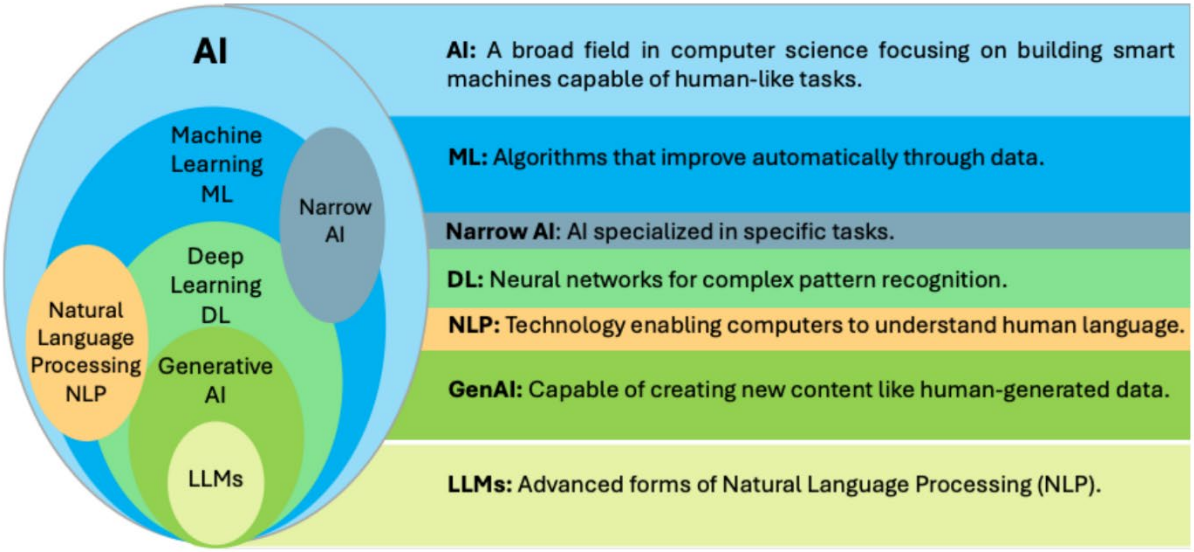

Figure 1. AI Concepts Defined.

Machine learning (ML) is a subset of AI that enables systems to learn and improve from data via algorithms that analyze large datasets, detect patterns, and make predictions. A subset of ML is deep learning (DL), which uses neural networks designed to mimic the human brain's functioning by detecting layers of patterns. DL models are particularly effective for tasks like image recognition, speech processing, and complex pattern recognition.

Natural language processing (NLP) is another crucial AI field that enables computers to understand, interpret, and generate human language. It is the foundation for chatbots, translation services, and text analysis tools. Within NLP, LLMs represent a significant advancement, as it allows AI models to generate new human-like content, such as text, images, videos, and even code, based on input data. LLMs are a powerful subset of GenAI that have been trained on massive amounts of data such as content from the internet and textbooks. LLMs are a subset of GenAI, meaning all LLMs are a form of GenAI, but not all GenAI models are LLMs. GenAI refers to any AI model that creates new, original content, including text, images, music, and code. For example, Midjourney and DALL-E are GenAI models that create images, but they are not LLMs because they do not create new content of text form. However, multimodal LLMs are a more advanced type of LLM that handles multiple data types and therefore referred to as multimodal.

This project primarily focuses on the practical applications of GenAI by experimenting with existing public LLMs. The research explores how these GenAI/LLMs tools can support water utilities in areas such as data cleansing, summarizing unstructured documents, and detecting anomalies through AI-powered computer vision, and more. By leveraging the capabilities of LLMs, utilities can improve operational efficiencies, enhance decision-making, and facilitate knowledge sharing across their organizations.

Retrieval Augmented Generation (RAG)

Another key concept that leverages GenAI/LLMs is Retrieval-Augmented Generation (RAG), a method that improves the accuracy of LLMs by allowing them to retrieve and incorporate relevant external data (users' data) during their response generation process. Traditional language models are trained on vast datasets but lack specific, real-time, or organization-specific context. RAG enables models to query structured and unstructured data sources, retrieve relevant information, and incorporate it into their responses, improving precision and contextual relevance. This also reduces the chances of hallucination, a term that describes when an AI system presents false, misleading, or distorted information as factual, since the LLM is mainly going to the local sources of information the user provided for the query.

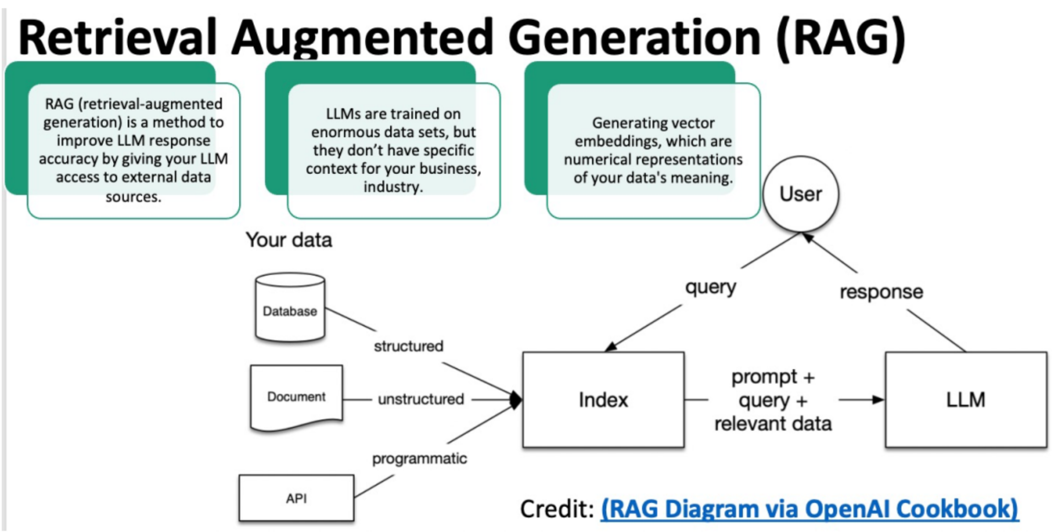

Figure 2. Retrieval Augmented Generation.
Source: Theja 2023.

A key component of RAG is the use of vector embeddings, which are numerical representations of textual, visual, or other data types. Embeddings allow language models to process and understand relationships between pieces of information efficiently. When unstructured data such as documents or images are indexed, they are converted into vector embeddings, which capture semantic meaning in a way that makes them easily searchable. When a user submits a query, the system retrieves the most relevant embeddings based on similarity scores, ensuring that the language model has contextually appropriate information before generating a response.

Recently, there has been much talk of Agentic AI, systems that go beyond responding to single prompts. Agentic AI software operates as autonomous agents capable of planning, executing, and adapting across multi-step tasks. These AI agents can make decisions, invoke other tools, access databases, and iterate based on feedback or changing conditions. The research team believes Agentic AI could be used to monitor SCADA anomalies, retrieve maintenance histories, draft reports, or even schedule follow-up tasks, without needing to be prompted for each step. For example, an agent could detect a suspected pump failure, cross-reference asset logs, summarize relevant repair instructions, and notify staff via an internal chatbot. This task-oriented autonomy can reduce time for operators, increase workflow efficiency, and help utilities stay proactive in managing infrastructure. As agent frameworks mature, they hold immense promise for streamlining complex utility operations in a safe, auditable, and scalable way.

A short glossary of terms:

- **Artificial Intelligence (AI):** Broad field of computer science focused on building smart machines capable of human-like tasks (problem-solving, decision-making, language, pattern recognition). AI is the broadest field.
- **Narrow AI:** AI systems specialized in specific tasks.
- **Machine Learning (ML):** Subset of AI that learns and improves automatically from data using algorithms.
- **Deep Learning (DL):** Subset of ML that uses neural networks to detect complex patterns (e.g., image or speech recognition).
- **Natural Language Processing (NLP):** Enables computers to understand, interpret, and generate human language (foundation for chatbots, translation, text analysis).
- **GenAI:** AI that creates new content (text, images, code, video, music) based on patterns in training data.
- **Large Language Models (LLMs):** Advanced GenAI models trained on massive datasets that can generate and explain human-like text; some are multimodal (can handle text, images, audio, etc.).
- **Retrieval Augmented Generation (RAG):** RAG means the AI can “look things up” in trusted documents before answering, so its responses are based on real information the user provides, not just what it learned during training.
- **Agentic AI:** Agentic AI is artificial intelligence that can plan and carry out multi-step tasks on its own, like a digital assistant that not only answers questions but also takes action, while still allowing humans to stay in control.
- **Traditional AI:** Refers to systems that analyze large sets of structured data to classify, predict, or optimize outcomes. These models are designed for specific, well-defined tasks like detecting leaks, forecasting demand, or recognizing images.

GenAI/LLMs vs. Traditional AI

Traditional AI and ML are primarily designed to classify, predict, or optimize outcomes based on large sets of structured data, such as SCADA data, weather, and sensor data. These systems usually require specialized expertise in data science, software development, and deep knowledge of the domain being analyzed. In contrast, GenAI, powered by LLMs, can create and explain new content, including text, code, images, and even voice interactions while using natural human language as the interface. This is a breakthrough because staff no longer need to write code or master advanced analytics to work with complex data, whether structured (tables, sensor readings) or unstructured (PDFs, handwritten notes, images, voice memos).

There is still an essential role for data scientists, system experts, and software developers, especially to design, validate, and maintain digital AI systems. However, GenAI/LLMs democratize access by allowing non-technical staff to extract insights from data that were previously out of reach. This project has shown that with GenAI/LLMs, a utility operator can describe what they need in plain language, and the system can generate insights, or even produce the code to achieve the result. This shift makes advanced AI accessible to non-technical staff, lowers the barrier to entry, and is why GenAI/LLMs represent a true game-changer for the water sector.

The role of unstructured data becomes critical because most utility information is unstructured. Gartner explains that 80% of enterprise data is unstructured. GenAI/LLMs can convert unstructured data into searchable, structured data and produce quick summaries and charts. With RAG, answers are grounded in an organization's standard operating procedures (SOPs), permits, manuals, and SCADA exports, improving relevance, while significantly reducing hallucinations. GenAI/LLMs also writes and reviews code, such as Python, to help with data analytics, including cleaning and joining datasets and producing visualizations. It supports voice as an interface for hands-free queries and field updates, and emerging Agentic AI can run multi-step workflows with human-in-the-loop oversight. Together, these capabilities form the foundation for the pilots, governance, and adoption roadmaps that follow.

> "Together, we aim to leverage the cost-effective yet sophisticated capabilities of generative AI to enhance utility operations, bridge the digital divide among utilities of all sizes, and establish a research roadmap that will propel global digital transformation in the water sector."
>
> Gigi Karmous-Edwards, Technical Lead and Co-Principal Investigator

Technology Enablers

GenAI/LLMs are usable today with two model paths for utilities: public LLM services using enterprise versions for security purposes (GPT-5, Claude, Gemini, etc.) and open-source LLM models (GPT-OSS, Llama, etc.) that can run in an organization's cloud or on-premises. These tools are all multimodal (text, images, PDFs, some have voice) and increasingly Agentic AI (able to plan and execute multi-step tasks with human-in-the-loop). Utilities can choose models based on security controls (audit logs, no data sharing, etc.), cost, latency, and deployment options. Utilities can build new efficient applications in three ways:

- No-code (custom GPTs and chat workflows).
- Low-code (notebooks, workflow builders, simple API wrappers).
- Full-code.

The research team describes Vibe coding as a way for non-software developers to describe application concepts in plain language and have the model develop the code and analytics to start the implementation. Utilities can utilize RAG so their GenAI/LLMs application, GPT, or prompt can summarize for users their own SOPs, permits, manuals, and SCADA exports. Another promising application is the use of voice to capture for hands-free field input, or any application input.

Agentic AI is now being used across various applications. They are a focus area for the main LLM companies like OpenAI and Gemini. OpenAI has put a focus on Agentic AI with OpenAI's CEO, Sam Altman, predicting that AI agents could begin to join the workforce in 2025 and significantly impact companies' output (Broomfield 2025). There has been a real effort in the tools to build agents at OpenAI with key releases in 2025:

- The Operator Agent, powered by the Computer-Using Agent (CUA) model, was released in January 2025 as a research preview and developer API. It is designed to interact with graphical user interfaces (GUIs) to automate tasks like coding and booking travel.
- The Agents SDK was released in March 2025, providing developers with tools to build their own agentic AI applications. It includes features for web search, file search, and computer use.
- ChatGPT Agent was introduced in mid-July 2025. This AI tool can perform multi-step tasks such as online shopping, creating presentations, and generating spreadsheets. An "all-in-one" agent was also launched in mid-July 2025, which Superhuman AI described as allowing for the quick, no-code deployment of AI agents.

Beyond awareness, the project facilitated experimentation by enabling utilities to test and validate small- to large-scale GenAI/LLMs capabilities. These proof-of-concept (PoC) experiments focus on practical applications to address water utility challenges. The project will document, analyze, and share findings from experiments as best practices with both participants and the global water sector, producing a utility guidebook, case studies, and a research roadmap to support GenAI/LLMs integration in utilities.

By democratizing access to GenAI/LLMs knowledge and tools, the initiative aims to ensure that all utilities, even under-resourced utilities, can benefit from technological advancements. Through collaboration with global utilities and industry experts, this study will help bridge the digital divide and promote practical, secure, and beneficial GenAI/LLMs adoption across the sector.

> "The voices that we need now include technology experts who can identify technology concepts that have the capability to be game changers in the water community – including Generative AI. This proposed project serves to do just that, to facilitate the development of best practices and guidance in this innovative and important new era."
>
> David LaFrance, CEO, AWWA

GenAI/LLMs Solutions to Challenges Facing the Water Sector

The global water sector is currently facing critical challenges with aging infrastructure, climate variability, resource constraints, workforce shortages, and customer affordability. GenAI/LLMs emerge with timely and accessible solutions, including:

- Simplifying workflows through employee digital assistants.
- Supporting asset maintenance and customer service.
- Optimizing operations.
- Preserving institutional knowledge.
- Delivering high-level capabilities at low cost through an easy-to-use interface.
- Streamlining optimizations across all operations.

This work aims to move beyond conceptual discussions and instead ground the research in utility-led, field-tested pilot/experimentation projects, and the everyday efforts of utility staff working across operations, customer service, and compliance roles.

Project Participants and Objectives

Project leadership, water utility participants, key objectives, and deliverables are summarized in the figure below. The leadership team carefully chose a diverse mix of utilities. Selection of the Global GenAI Utility Group (G3) comprised of utilities of different sizes, water and wastewater services, and global locations. This ensured that GenAI is not just for large utilities; our findings show GenAI works for smaller utilities too.

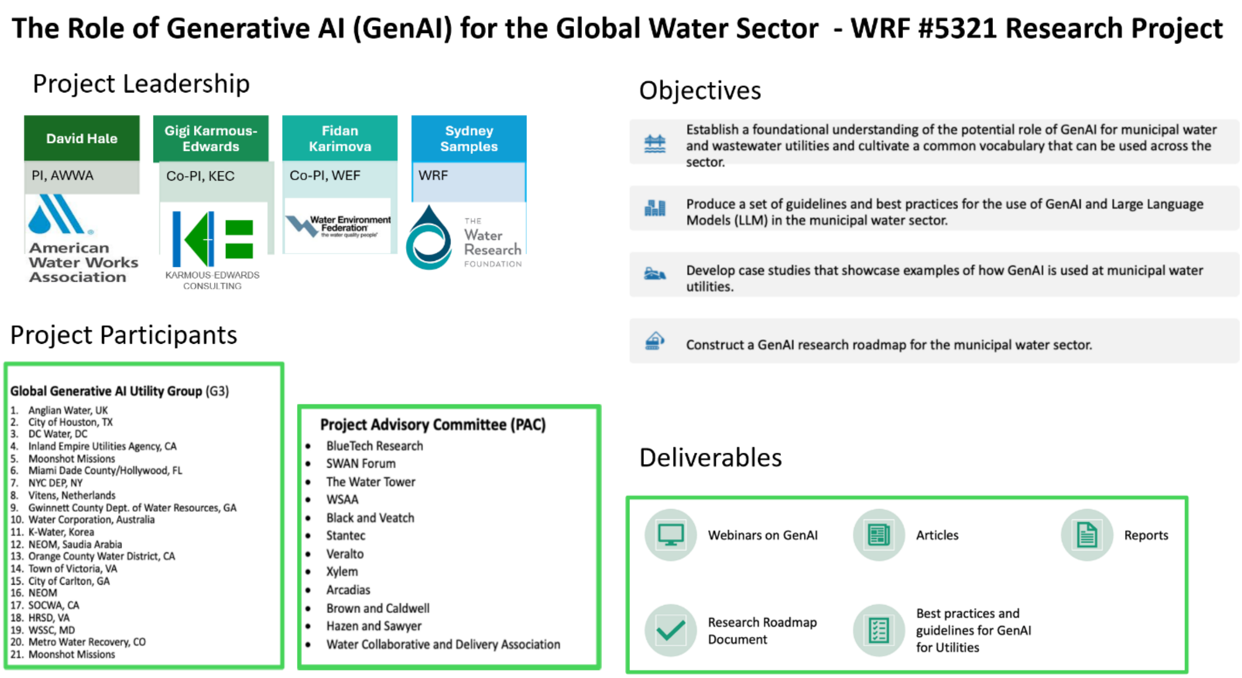

Figure 3. WRF Project 5321 Overview.

Main Findings

The research shows that GenAI/LLMs offer a practical, scalable, and affordable set of tools that can help address many of the sector's most persistent challenges. Rather than requiring large IT investments, utilities can begin applying GenAI/LLMs capabilities such as optical character recognition (OCR), computer vision, data analytics, and RAG with all types of structured and unstructured data sets, as well as PDFs, voice memos, images, engineering drawings, and handwritten documents. Importantly, these tools are accessible through secure enterprise versions, typically costing only $25–$60 per user per month, making them both affordable and aligned with data protection needs.

The research team discovered that a human-language interface to the enterprise version of GenAI/LLMs makes it easy for nearly all staff roles to start experimenting with enterprise versions of LLMs in a secure manner. Human-language interfaces have the potential to lower barriers for staff across all roles. When utilities hesitate to adopt GenAI/LLMs, it is usually for two main reasons: fears of job loss and concerns about data security and compliance. This project demonstrates GenAI/LLMs work best as a digital assistant that augments staff rather than replaces them. With enterprise deployments, added required security enabled, and model-training/data-sharing turned off, utilities can meet compliance requirements for digital applications.

Key Challenges to Scalable Adoption of GenAI/LLMs

Through numerous conversations and workshop dialogues, the participants discovered that utilities need to make safe, practical data governance the top priority by expanding policies and controls to cover both structured and unstructured information. Looking to ensure that a high quality of data is achieved and maintained because this determines the quality and reliability of the outputs generated. GenAI/LLMs tools can analyze both structured data and unstructured data making more information available for data-driven insights. The utilities in the G3 played a major role in creating practical recommendations to guide this effort for the report.

Data Governance

The G3 team focused on developing recommendations for effective data governance as water utilities adopt GenAI/LLMs, not just for compliance, but for enabling reliable, and impactful outcomes. The report outlines a strategic roadmap for utilities to govern their structured and unstructured data, covering everything from foundational data charters and inventories to tiered access, metadata standards, and decision-data capture. It emphasizes the importance of integrating cybersecurity, scaling governance to utility size, and continuously adapting to change. With more than 80% of enterprise data unstructured, GenAI/LLMs offer an unprecedented opportunity, but only if paired with intentional, right-sized governance. Here is a summary of topics:

- Start with a data charter.
- Inventory and classify data.
- Govern decision-data.
- Use tiered access controls.
- Scale governance to fit.

In summary, data governance has moved from important to mission critical. Managing unstructured content with the same rigor as SCADA tags or billing data will be essential for safe GenAI adoption.

Workforce and Change Management

Alongside data governance, the G3 focused on three additional pillars for successful GenAI/LLMs adoption, mainly related to workforce and people-related aspects:

1. Establishing a utility-wide community of practice.
2. Training staff on the safe and effective use of GenAI/LLMs.
3. Securing executive buy-in for enterprise-level deployments.

In parallel, the group documented several other common challenges, including concerns about job loss and issues of security, privacy, and compliance. A common concern among utility staff is that adopting GenAI/LLMs will eventually replace their jobs. However, the team's experiments show the opposite. The research team has observed these tools as a co-pilot, not a replacement. These tools help staff work more efficiently by automating repetitive tasks, surfacing insights from unstructured data, and preserving institutional knowledge, while leaving critical judgment, decision-making, and community engagement to people. In practice, GenAI/LLMs reduce the burden of paperwork and data searching, giving employees more time to focus on higher-value work. Rather than eliminating roles, it strengthens them, making the workforce more capable, confident, and future-ready.

There is also an awareness gap because many confuse GenAI/LLMs with traditional AI (mainly considered to be data and analytics). Many employees are unclear on the difference between GenAI/LLMs, traditional analytics, and narrow AI tools already in use. In contrast, GenAI/LLMs go beyond prediction, since it can create new content such as text, images, code, and even insights from unstructured data, all through natural language interfaces. This makes GenAI/LLMs far more accessible, allowing non-technical staff to use advanced AI capabilities without needing data science or programming skills.

Utilities often operate in silos, with different departments struggling to collaborate. This lack of shared terminology slows internal discussions and early planning, making it harder to align efforts. The result is fragmented pilots, duplication of work, and mistrust between teams. A further challenge is the divide between technical and non-technical staff, which creates uncertainty about roles as plain-language tools may bypass traditional IT processes. This research provides a common framework and language to help utilities bridge these gaps and work together more effectively. One key strategic risk is that utilities delaying GenAI/LLM adoption may find it harder to attract and retain top talent, further widening the digital divide.

Research participants found the following approaches to be good practices:

- Position GenAI/LLMs as an assistant**,** not a replacement, with visible executive sponsorship.
- Provide layered training for all roles and create a community of practice.
- Put clear policies in place for acceptable use, governance, and data security.
- Run small, measurable pilots, share results, and celebrate wins to build confidence, attract and retain talent, and preserve institutional knowledge.

In summary, the G3 team found that the success of GenAI/LLMs in the water sector depends not only on data or tools, but on people. Adoption will stall without a workforce that understands, trusts, and can effectively apply AI capabilities in daily workflows. The report explores how utilities can foster a shared vocabulary, create targeted training paths, and embed GenAI/LLMs into culture and operations through intentional design. Utilities that invest in role-relevant training, pilot coordination, and early success stories are best positioned to translate isolated and siloed AI trials into scalable, system-wide transformation.

Broader Challenges to Adoption

G3 participants examined the broader sector challenges associated with utilities adopting GenAI/LLMs. Some of the challenges identified will require further research and sector-wide coordination. Alongside the above challenges of data governance and workforce preparedness, are fragmented efforts, thin budgets, and unclear policies. The report explores practical ways the sector can start to address these issues. For water and wastewater utilities that do not embrace GenAI they face a significant risk of talent loss, as top professionals, especially young ones, gravitate towards organizations that prioritize innovative.

Participants also explored the growing water and energy demand related to GenAI/LLMs and the emerging resource crunch today from the unprecedented acceleration in building data centers across the country and the global constraints for capital, human, and supply chain resources in this globally competitive environment. There is also a great deal of effort from these GenAI/LLMs companies to move towards zero-water. Microsoft, for example, recently announced the successful deployment of a new closed-loop, chip-level cooling system that virtually eliminates evaporative water use (Ambros 2025).

These innovations dovetail with WEF's Circular Water Economy initiative. Although these innovations are currently in the works, today, most data centers being built require millions of gallons of water for cooling. The global AI race is driving the acceleration for the building of U.S. data centers. Also, the current U.S. Administration's Executive Order 14179 further adds to the momentum and adoption needs. Released in July 2025, this order includes an AI action plan with the following three pillars:

- **Accelerating AI Innovation:** Removing regulatory hurdles and encouraging broad AI adoption.
- **Building American AI Infrastructure:** Streamlining the permitting process for AI data centers, chips, and energy infrastructure, and prioritizing domestic sourcing.
- **Leading in International AI Diplomacy and Security:** Using U.S. influence to promote American AI standards globally, strengthen export controls against adversaries, and protect U.S. AI innovations.

For water utilities, this growth has direct, near-term impacts: large, continuous water demands for cooling; new peaks in electricity use that tighten the water–energy nexus; and accelerated timelines for permitting, capacity planning, and infrastructure upgrades. Many of these data centers are being built in rural areas due to lower land costs and lower taxes, yet the utilities in these areas have limited staff, thin budgets, and constrained supplies, making it hard to absorb the sudden industrial loads.

Bridging the Digital Divide with GenAI

Small and rural water utilities often lack funding, IT staff, and modern systems, which slow their digital progress. Staff are already stretched thin, and vital knowledge often leaves with retiring employees. Traditional grant processes and outdated technology create further barriers. Many small or low-resourced utilities are sometimes dependent on paper logs, siloed spreadsheets, or phone calls to keep operations running.

> "In an era when utilities face rapidly expanding challenges—from extreme weather to emerging contaminants and the retirement of experienced personnel—water leaders need a lifeline to help face these challenges in an affordable way. Generative AI can offer this lifeline to capture and share knowledge which can then be customized to deliver predictive and proactive recommendations to any employee at any size utility in any circumstance in any place. This project will provide essential insights on how this tool could evolve to be the most important addition to the management of water in a generation, perhaps any generation."
>
> George Hawkins, CEO, Moonshot Missions

GenAI/LLMs show a glimpse of being a game-changer with the potential to bridge the digital divide between small and larger utilities. WRF project 5321 shows that GenAI/LLMs could be a practical path forward for small and rural utilities that is affordable.

By using enterprise versions of public GenAI/LLMs at a modest monthly cost, utilities can digitize handwritten notes, photos, and voice memos, turning them into searchable, actionable data without major infrastructure upgrades. Natural-language interfaces mean staff do not necessarily need coding skills, or to be a data analyst, to use the tools. They can use human language as a potential interface to ask important questions or give voice commands to generate reports, spot trends, or access SOPs.

Early pilots show that GenAI/LLMs can preserve institutional knowledge, cut down on double entry, and enable real-time insights, helping even the smallest utilities modernize workflows and compete on a more level playing field.

Proof-of-Concepts

WRF project 5321 findings raise awareness, share knowledge, and enable hands-on experimentation with GenAI/LLMs to improve efficiency and knowledge sharing across utilities. Through ongoing knowledge exchanges, webinars, and workshops, G3 participants learned how to use current tools (e.g., enterprise versions of LLMs) safely and effectively to support daily operations, streamline workflows, and strengthen decision-making.

The experiments drove value and momentum. Utilities ran proof-of-concepts across topics including rate structure studies, sensor maintenance and placement, coding of operational scripts, computer vision, and capturing institutional knowledge.

These proof-of-concepts range from simple, high-value RAG applications to more complex efforts, code-based operations optimization, OpenAI's GPTs (generative pre-trained transformers) for rate-structure calculations, and exploratory Agentic AI for digital twins.

Sharing each pilot's method and results sparked new experiments at other utilities. Every completed pilot delivered a measurable benefit and great learning opportunities to better understand risks and existing challenges (which are recorded in the report). Each experiment produced time or cost savings and/or higher accuracy, proving that simply running the pilot creates immediate benefit.

The team noted that GenAI/LLMs accelerate digital transformation at any maturity level by making unstructured information usable and by enabling plain-language into workflows while using existing technology such as mobile phones. This is especially significant for small and under-resourced utilities.

For small utilities, they found benefits of GenAI/LLMs such as digitizing paper processes, interpreting existing utility documents, and automating reporting without large investments in IT projects or new infrastructure. GenAI/LLMs helps bridge the digital divide between small and large utilities at accessible cost.

Findings are captured as best practices, utility guidelines, case studies, and the beginning of a GenAI/LLMs research roadmap for the water sector. The final report will widen access to GenAI/LLMs for utilities around the globe and help bridge the digital divide.

Sample Use Cases

Example 1: Computer Vision for Safety and Asset Monitoring in Wastewater Facilities

Partner Utility: Hampton Roads Sanitation District (HRSD), VA

This case study documents Hampton Roads Sanitation District's effort to use computer vision to detect leaks outside of pipes, corrosion, human presence, and smoke/fire across wastewater facilities. The goal is to supplement (not replace) operator shift rounds with continuous automated monitoring and timely alerts. The solution prototype began May 2025, with significant code authored using Anthropic's AI tool called "Claude" within the Cursor IDE Technologies. Early work focuses on site-ready detection datasets, use of classification and anomaly detection models, and integration paths with existing operations.

Example 2: Water Synchronicity (Rate Structure Assistant) GPT

Partner Utility: City of Carlton, and TWT, GA

Water Synchronicity ChatGPT is a GenAI/LLM-assisted decision-support tool in the form of a chatbot that helps small water utilities explore alternative rate structures aligned with goals for affordability, conservation, and revenue stability. Built on affordability research authored/co-authored by Dr. Ben Rachunok and tested with the City of Carlton case study, the tool combines community-specific spreadsheets (system data and low-to-moderate-income data) with a LLM trained on a curated corpus of affordability research, regulatory guidance, and rate-design best practices. It operates in a closed (no-internet) mode for data protection and can incorporate workforce salary needs into scenario planning.

Users state their priorities and weights (e.g., protect low- and moderate-income (LMI) households, encourage conservation, maintain revenue stability) and receive tailored rate structure suggestions with transparent trade-offs, accelerating analysis, surfacing options, and informing stakeholder discussions. Outputs are starting points, not final rate designs. Although safeguards reduce it, the model can over-recommend flat rates, which are often unsuitable; results should be reviewed and refined through utility-specific analysis and engagement. As an exploratory assistant, the tool speeds scenario testing, improves clarity around trade-offs, and supports more informed, data-driven rate studies.

Example 3: Houston Water's AI Agents for Dynamic CIP Planning

Partner Utility: City of Houston, TX

Houston Water Infrastructure Planning Group is piloting an AI Agent for Dynamic Capital Improvement Planning (CIP) to modernize how infrastructure investments are prioritized. The agent integrates hydraulic modeling, asset condition, and growth data to continuously update project priorities based on evolving needs and constraints. By enabling scenario testing and risk-informed strategies, it shifts utilities from static five-year CIP cycles to adaptive, data-driven planning, delivering more resilient, and cost-effective infrastructure investment decisions.

Example 4: WaterGPT LLM for Houston Water Infrastructure Planning

Partner Utility: City of Houston, TX

Houston Water Infrastructure Planning Group is piloting WaterGPT to enable more informed planning and decision making. WaterGPT is a domain-specific LLM (an on premise downloaded opensource LLM) built for the water sector at City of Houston, designed to unify regulatory frameworks, planning documents, and operational datasets into a secure, conversational platform. It provides utilities with rapid access to complex regulatory and engineering knowledge, while supporting data-driven analyses for capital planning, risk management, and operational efficiency.

The system is architected for offline deployment to safeguard data sovereignty and can scale from regulatory compliance to advanced applications such as main break prediction, demand forecasting, and digital twin integration. By embedding strong governance protocols and piloting AI with non-sensitive datasets, WaterGPT builds trust with the users, while advancing the sector toward predictive and resilient water management.

Example 5: Modern-Day Customer Service (AI-Enabled)

Partner Utility: DC Water, Washington, D.C.

DC Water's Customer Care Department is committed to enhancing customer satisfaction and service quality by implementing GenAI-powered self-service capabilities. Applying the 80:20 Pareto Principle to customer service, the team first identified which 20% of the set of questions or support issues from customers drives 80% of our support volume. In other words, most of the effort is spent on a narrow set of commonly asked questions. Most of the questions that are constantly being asked have for the most part the same answers, which was incorporated into the GenAI Self-service.

In response, DC Water is in the process of automating portions of its customer service through GenAI/LLMs. The aim is to enable customer support 24x7x365 in multiple languages. This will be accomplished through a combination of: (i) conversational GenAI, (ii) text chatbots, and (iii) an interactive voice response (IVR) system. This includes creating and modifying service orders, accessing emergency updates, reporting emergency issues, retrieving account details, and making payments, among many other functions.

Example 6: Voice SOP at South Orange County Wastewater Authority

Partner Utility: South Orange County Wastewater Authority (SOCWA), CA

This case study explores Voice SOP, a voice-activated AI assistant designed to enable lab technicians in water utilities to query and access standard operating procedures (SOPs) hands-free, without removing gloves. The solution leverages GenAI/LLMs for natural language processing and voice commands to retrieve and verbalize SOP details, improving safety, efficiency, and compliance in lab environments. The main outcomes include:

- Reduced manual interruptions and time savings on SOP consultations.
- Enhanced accuracy in procedure adherence.
- Better overall workflow in glove-restricted settings.

Example 7: What Would Jerry Do (WWJD) – Chlorine GPT

Partner Utility: City of Carlton and TWT, GA

This GPT, known as “WWJD – Chlorine,” serves as a specialized expert assistant for rural and small water system operators, particularly those managing chlorine residuals in well water systems. Modeled after a seasoned rural water engineer, Jerry, it delivers clear, experienced-based advice tailored to the unique challenges of small utilities, like limited staffing, aging infrastructure, variable water quality, and strict regulatory oversight under the Safe Drinking Water Act. It provides guidance on chlorine dosing, residual monitoring, seasonal impacts, disinfection byproduct (DBP) control, equipment troubleshooting, and regulatory compliance.

The GPT is grounded in authoritative sources, including U.S. Environmental Protection Agency small system research, case studies on chlorine residuals and DBPs, and field-ready SOPs from the Rural Water Association, as well as a two-hour detailed conversation with Jerry on specific and frequently asked questions he gets daily. Its tone blends technical rigor with calm, practical wisdom, helping rural operators make sound, affordable decisions for safe and compliant drinking water. Operators that normally would have called Jerry for an answer to their question can now first try to chat with this chatbot via their mobile phone or computer to get the answers they were looking for.

Vision

WRF 5321 demonstrates GenAI/LLMs can provide practical and timely solutions for municipal water and wastewater utilities; the next step is scale. The research team approach is inclusive and human centered. GenAI/LLMs serve as a virtual assistant, augmenting utility professionals rather than replacing them. As a timely, accessible digital assistant, G3 members have shown that GenAI/LLMs simplify workflows, support asset maintenance, optimize operations, enhance customer service, preserve institutional knowledge, and democratize information, by delivering high capability at low cost through easy-to-use interfaces (e.g., voice, mobile phone, text).

Because most utilities share far more similarities than differences, such as SOPs, assets, vendors, regulatory workflows, the research team believes the sector must move away from one-off pilots to create robust tools that can be adapted for multiple scenarios. The team has documented pilots for replication; however, in the last few months the focus has been on scaling, not just duplicating. With AWWA, WEF, and WRF convening the sector to provide best practices, the team will help inform the sector on some of the core building blocks of GenAI/LLMs-enabled tools including unstructured-data ingestion, OCR/RAG, evaluation and safety guardrails, and role-based interfaces. This will enable utilities to take advantage of tools that bridge the digital divide, preserve institutional knowledge, and boost workforce productivity.

This research has made important progress in showing the practical, secure, and affordable role of GenAI/LLMs in the water sector. At the same time, it has surfaced key questions that must be addressed for the sector to continue growing and evolving in this space. These questions will form the foundation of a research roadmap to guide future efforts, ensuring that utilities, researchers, and industry partners can move forward together in a coordinated way. The upcoming second in-person workshop will be a key moment to capture these priorities, refine the roadmap, and set the direction for the next phase of research and innovation. The final report will include the research roadmap.

For further questions or suggestions, please contact:

Gigi Karmous-Edwards, Karmous Edwards Consulting (KEC), gigi@gigikarmous.com.

Authors by Organization

Organization	Name
Anglian Water	Annesley Crisp
Anglian Water	Andrew Smith
Anglian Water	Darren Coleman
American Water Works Association	David Hale
City of Carlton	Amanda Willis
City of Hollywood (FL)	Hardeep Anand
City of Houston	Satish Tripathi
DC Water	Robert Bornhofen
DC Water	Nadine Zerrini
Gwinnett County Department of Water Resources	Melissa Meeker
Gwinnett County Department of Water Resources	Kevin Middlebrooks
Hampton Roads Sanitation District	Jeff Sparks
Hampton Roads Sanitation District	Clay Wise
Inland Empire Utilities Agency	Michael Hurley
Karmous Edwards Consulting	Gigi Karmous-Edwards
Metro Water Recovery	Josh Goldman
North Carolina State University	Ben Rachunok
Moonshot Missions	Andy Koester
New York City Department of Environmental Protection	Michael Shum
Orange County Water Districts	Gina Ayala
Orange County Water Districts	Berch, Dani
Orange County Water Districts	Haney, Lisa
South Orange County Wastewater Authority	Amber Boone
Vitens	Martin Hutten
Water Environment Federation	Fidan Karimova
The Water Research Foundation	Sydney Samples
WSSC Water	Michael Abobor

References

Ambros. 2025. "Microsoft's Zero-Water Cooling: Lessons for Data Centers." May 5. https://www.serverion.com/uncategorized/microsofts-zero-water-cooling-lessons-for-data-centers/.

Broomfield, Benjamin. 2025. "AI agents | OpenAI CEO: We could be working alongside virtual employees by the end of 2025." *HR Grapevine*. January 7. https://www.hrgrapevine.com/us/content/article/2025-01-07-openai-ceo-we-could-be-working-alongside-virtual-employees-by-the-end-of-2025.

Theja, Ravi. 2023. "Evaluate RAG with LlamaIndex." *OpenAI Cookbook*. https://cookbook.openai.com/examples/evaluation/evaluate_rag_with_llamaindex.