Sustainable Cities and Society 132 (2025) 106793

Contents lists available at ScienceDirect

Sustainable Cities and Society

journal homepage: www.elsevier.com/locate/scs

Empowering SMEs with *SustainWater Bot* to advance urban water sustainability

Muhammad Arslan [a,*], Saba Munawar [b], Zainab Riaz [c]

[a] *School of Architecture and Environment, University of the West of England, Bristol, England*
[b] *Telecommunication Engineering, National University of Computer and Emerging Sciences, Islamabad, Pakistan*
[c] *College of Business, Abu Dhabi University, Abu Dhabi, United Arab Emirates*

ARTICLE INFO

Keywords:
Urban water decision-making
Sustainable transitions
Small and medium-sized enterprises (SMEs)
Large language models (LLMs)
Retrieval-augmented generation (RAG)

ABSTRACT

Climate change, population growth, and resource constraints are intensifying pressure on urban water systems (UWSs), prompting a shift toward integrated information management. Due to their agility and reach, small- and medium-sized enterprises (SMEs) are central to this transition. However, many SMEs lack access to robust information systems (ISs) that consolidate government initiatives, industry trends, and broader water-related data, impeding sustainable adoption. This study introduces *SustainWater Bot*, a chatbot driven by generative artificial Intelligence (GenAI), including large language models (LLMs) and retrieval-augmented generation (RAG). Designed to fill this information gap, *SustainWater Bot* addresses the shortcomings of conventional LLMs, such as information misalignment, over-complexity, and information deficiencies. RAG enables semantic consolidation of various sources, such as news, government reports, industry insights, academic research, and social media, into an integrated IS. The evaluation results showed that RAG with LLM-based methods outperformed traditional information retrieval (IR) techniques, with Llama3.2:3b achieving top scores in precision (95 %), completeness (95 %), and exact match (90 %). Traditional IR techniques such as term frequency-inverse document frequency (TF-IDF) and best matching 25 (BM25) performed lower but offered quicker responses. *SustainWater Bot* supports informed decision-making through a question-answering (QA) framework that delivers relevant insights on sustainable urban water initiatives (SUWIs). It is built on open-source technologies and offers SMEs a cost-effective, scalable, and sustainable solution to enhance eco-friendly water practices and operational efficiency.

1. Introduction

In 2019, the UK became the first major economy to incorporate a net-zero emissions target for 2050 into law (Water-UK, 2023). This transition involves reducing greenhouse gas emissions to nearly zero and offsetting any remaining emissions (Bataille, 2020). Key strategies for achieving this include water conservation, leakage reduction, and efficient UWSs (Water-UK, 2023; Werbeloff & Brown, 2016). UWSs are under increasing strain from climate change, population growth, resource limitations, and shifting community needs (Werbeloff & Brown, 2016). These pressures are leading to a shift toward more integrated and sustainable information management approaches. Effective urban water management (UWM) must address water security, flood risk, urban heat islands, and waterway degradation while improving liveability (Werbeloff & Brown, 2016). Achieving these objectives necessitates a profound shift toward environmentally sustainable practices. Emerging research on socio-technical and socio-ecological transitions provides insights into how systems can shift to more sustainable practices (Werbeloff & Brown, 2016). Sustainable urban water transitions (SUWTs) occur at multiple levels, such as individual, community, SME, and government levels (Dunn et al., 2017). Given their substantial influence in urban economic systems and their effect on environmental sustainability, the SMEs' role in these transformations is particularly crucial.

In 2023, SMEs, which constitute over 99 % of the UK's business population with approximately 5.5 million enterprises, significantly impact resource consumption and greenhouse gas emissions (Hutton, 2024). Their role in urban water transitions is crucial due to their involvement in key areas, such as water supply, wastewater management, and stormwater infrastructure (Dunn et al., 2017). By adopting sustainable practices, such as using water-efficient technologies, reusing water, and reducing runoff, SMEs can enhance the resilience of UWSs

* Corresponding author.
E-mail address: muhammad.arslan@uwe.ac.uk (M. Arslan).

https://doi.org/10.1016/j.scs.2025.106793
Received 4 January 2025; Received in revised form 3 August 2025; Accepted 4 September 2025
Available online 5 September 2025

(Borsacchi & Pinelli, 2020). Their agility in implementing innovative water management solutions allows SMEs to effectively tackle local water challenges and contribute to broader urban sustainability goals (Raya-Tapia et al., 2023; Tan & Yao, 2023). SMEs can reduce their environmental footprint and support community sustainability through water stewardship initiatives (Dunn et al., 2017; Borsacchi & Pinelli, 2020). Embracing sustainability-oriented business models, such as rainwater harvesting, greywater recycling, and green infrastructure, enables SMEs to drive transformative change in UWSs and set industry standards, influencing larger corporations and promoting urban sustainability.

Government policies, incentives, subsidies, and regulations support SMEs' transitions to sustainable urban water practices, which are crucial for encouraging these practices (Álvarez Jaramillo et al., 2019; Vidal-Lamolla et al., 2024; Garcia et al., 2024). In addition to governmental support, SMEs benefit from a wealth of information available in various media formats, including structured databases, semi-structured Extensible Markup Language (XML) files, and unstructured text documents (Adnan & Akbar, 2019). However, effectively organizing and accessing this diverse information is challenging (Klewitz & Hansen, 2014). The current literature (Kolditz et al., 2019; Liu et al., 2021; Muñoz-Pascual et al., 2021; Klewitz & Hansen, 2014; Naeem et al., 2024) identifies a gap in ISs that can manage and synthesize these varied data formats. A streamlined system (Rajaram & Tinguely, 2024) that allows SMEs to easily access and utilize the latest SUWIs and policies is urgently needed. Developing such an IS would enable SMEs to make informed decisions and utilize resources, thus supporting broader sustainable UWM efforts (Rajaram & Tinguely, 2024). This study seeks to address this challenge by focusing on key questions.

1) How can an IS semantically integrate data from various sources to benefit SMEs in the urban water sector using advanced open-source technologies?
2) How can this system improve the decision-making processes related to SMEs' SUWTs?
3) How can the system be developed to remain sustainable, enabling seamless upgrades and customization to adapt to the changing interests of SMEs over time?

Guided by these questions, the study is structured around the following research hypotheses:

A. A domain-specific IS built with RAG-enabled LLMs provides more accurate and contextually relevant responses than traditional IR-based methods (e.g., TF-IDF and BM25).
B. Curated, semantically structured information across key urban water topics enables users, particularly SMEs, to form more actionable ISs within the sustainability landscape.
C. A modular and adaptable IS architecture enables the long-term sustainability of IS and decision support capabilities, thereby reducing the dependence on constant manual information curation.

To address these research questions and hypotheses, this study proposes the development of the *SustainWater Bot*, a chatbot designed as an IS to semantically integrate data from multiple sources. Using advanced open-source GenAI technologies (Kurivella & Lauren, 2023), including LLMs and RAG, *SustainWater Bot* is specifically tailored for the water domain, offering several significant advantages:

1. *Comprehensive Integration: SustainWater Bot* semantically consolidates diverse data sources into an integrated system, including news outlets, government reports, industry publications, and social media data. This integration equips SMEs with a broad and up-to-date repository of information and trends on SUWIs.
2. *Enhanced Decision-Making*: The chatbot supports SMEs in making informed, data-driven decisions about sustainable water management by providing timely and relevant information. This capability enhances operational efficiency and helps SMEs to adopt effective and innovative practices more quickly.
3. *Adaptable, Cost-Effective, and Sustainable Solution: SustainWater Bot* is designed for easy upgrades and customization, ensuring it remains relevant as organizational needs evolve. Its adaptability and use of relevant datasets make it a powerful tool for continuous learning and improvement. Furthermore, *SustainWater Bot* offers a practical, economical, and sustainable solution by using existing LLMs for ongoing adaptation in UWM with an emphasis on minimizing development costs.

The paper is structured as follows: Section 2 outlines the background of the study and establishes the foundation. Section 3 details the dataset used to construct the IS. Moreover, the implementation of the IS, *SustainWater Bot,* is discussed, along with the system's evaluation. Section 4 offers a discussion on the system's benefits, limitations, and future works. Finally, Section 5 presents concluding remarks.

2. Background

The background has two sections: one on UWSs and sustainability, and the other on LLMs with RAG and their applications. It highlights the need for an integrated IS to provide SMEs with up-to-date water information.

2.1. Urban water systems (UWSs)

UWSs involve managing water in cities, including supply, treatment, and stormwater management (Daniell et al., 2015). Sustainable management faces challenges from environmental factors such as resource scarcity and climate change, socio-economic factors, and governance issues such as the need for global cooperation (Daniell et al., 2015). Understanding these drivers is essential for effective UWM. To address these issues, UWSs must achieve objectives such as ensuring a reliable water supply, protecting public health, managing flood risks, and preserving environmental and social facilities (Daniell et al., 2015). Various approaches are used: water supply systems focus on reliable distribution through centralized infrastructure and treatment; sewered systems handle wastewater collection and treatment; drained systems manage stormwater to prevent flooding (Risch et al., 2015); waterways systems restore natural waterways for ecosystem services and flood control (Schiessel Harvey, 2021); water cycle systems take a holistic view of sustainable use (Mitchell et al., 2001); water-sensitive urban systems integrate design to manage runoff and reduce pollution (Chadfield et al., 2024); and adaptive urban-rural systems address the urban-rural water interactions (Wang et al., 2024). Each approach is essential for achieving urban water sustainability.

To address the complexities of UWM, several innovative studies have been proposed and demonstrated different systems aimed at improving the understanding, planning, and efficiency of UWSs. One study, as described by Stewart et al. (2010), introduced a web-based knowledge management system designed to integrate smart metering, end-use water consumption data, wireless communication networks, and ISs. This system provides real-time insights into water consumption patterns, offering valuable information to both consumers and utilities. Another significant contribution is the dynamic urban water simulation model, developed by Willuweit and O'Sullivan (2013). It allows for the analysis of different urbanization scenarios and climatic changes, offering insights into their effects on the urban water cycle. The integrated urban water model (Sharvelle et al., 2017) forecasts urban water demand and evaluates the impacts of water conservation and reuse strategies. Another study (Nguyen et al., 2018) highlights the use of granular data from an intelligent water management system to help planners optimize efficiencies across different stages of the urban water cycle, including supply, distribution, and wastewater treatment.

Additionally, the creation of an urban building database (Monteiro et al., 2018) supports innovative city initiatives by supplying the detailed data required for urban building energy simulations, thus improving urban energy management systems. Moreover, a study by Van Ginkel et al. (2018) developed a dashboard with 56 indicators using the pressure-state-impact-response framework to examine urban water security, supporting strategic planning and governance. Furthermore, a study (Garajeh et al., 2024) proposed a geospatial-based method to optimize the location of drinking water relief posts in Zanjan, improving the efficiency of the urban water network. Another study (Salam, 2024) introduced an internet of things (IoTs)-based approach to environmental quality and monitoring. This approach generates insights into sustainable resource management by addressing the impacts of water contaminants. Lastly, Raman et al. (2024) outlined the development of a knowledge-based decision support system to enhance the decision-making capabilities of water planning experts through a framework for evaluating water supply management scenarios.

The existing studies (Monteiro et al., 2018; Garajeh et al., 2024; Van Ginkel et al., 2018; Willuweit & O'Sullivan, 2013) have made significant contributions to advancing UWM systems, particularly in promoting sustainability practices. Traditional IR methods, such as TF-IDF (i.e., measures how important a word is to a document in a dataset) (Aizawa, 2003; Roelleke & Wang, 2008) and BM25 (i.e., a ranking function that improves relevance scoring in search results) (Amati, 2009), have been employed in various systems (Li et al., 2025; Kim et al., 2022) to enable text-based querying on water-related datasets. However, these approaches primarily rely on lexical matching, which limits their ability to capture semantic relevance, especially in complex domains such as UWM (Marwah & Beel, 2020). Moreover, most of the current systems (Sharvelle et al., 2017; Nguyen et al., 2018) are constrained by their reliance on a single type or source of data, are not open-source, and lack transparency regarding their information sources.

Given the dynamic and multifaceted nature of UWM, it is essential to continuously update the systems that inform sustainable transitions with current information about water systems and their surrounding landscapes. SMEs, which play a crucial role in this sector, often encounter resource (i.e., time and finances) limitations that deter their ability to evaluate and adopt existing ISs tailored to their needs. Consequently, there is a growing demand for open-source, customizable, and adaptable solutions that can meet the unique requirements of SMEs. Currently, no comprehensive IS exists that enables SMEs to efficiently access up-to-date data on SUWIs and the broader urban water landscape (Roffia & Dabić, 2024; Kirchner-Krath et al., 2024; Maman et al., 2024). This gap highlights the relevance of LLMs augmented with RAG. The following section introduces the fundamentals of LLMs and RAG and discusses their potential to address the limitations of existing ISs while meeting the evolving needs of stakeholders in UWM.

2.2. RAG with LLMs through GenAI

GenAI refers to algorithms and models designed to create new content (i.e., text, images, etc.) by learning patterns and language structures from large datasets (Gupta et al., 2024). This technology enables machines to produce contextually relevant outputs that mimic human creativity and problem-solving abilities. LLMs, a key application of GenAI, are pre-trained on extensive datasets containing billions of parameters (Yang et al., 2024; Chang et al., 2024). By using deep learning techniques, LLMs excel at identifying complex patterns and relationships within language, enabling high performance across a wide range of natural language processing (NLP) tasks. However, they face challenges with domain-specific queries due to their reliance on broadly trained datasets, resulting in inaccuracies or irrelevant responses (Arefeen et al., 2024). The key limitations of LLMs include information misalignment, excessive complexity, and deficiencies in accessing domain-relevant data (Fig. 1). To address these limitations, RAG has been introduced (Lewis et al., 2020). RAG enhances LLMs by addressing challenges such as information misalignment, over-complexity, and information gaps. It improves accuracy by retrieving context-specific data, simplifies responses for clarity, and provides domain-specific information (Fig. 2). Eventually, RAG-based approaches bridge the gap between user queries and up-to-date knowledge, refining the LLM output.

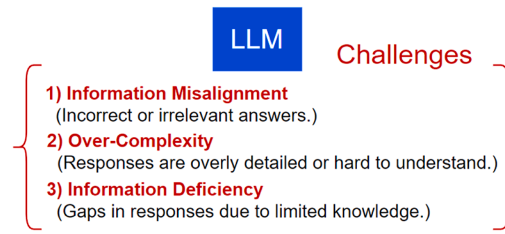

Fig. 1. Key challenges in using LLMs for information extraction.

RAG demonstrates its versatility in enhancing QA across various domains (Gao et al., 2023; Arslan et al., 2024a, 2024b, 2024c). For example, MIRAGE (Xiong et al., 2024) uses medical QA datasets to refine biomedical QA systems. In finance, the RAG (Jimeno Yepes et al., 2024) leverages financial reports to provide accurate responses to financial queries. In knowledge management, RAG-augmented knowledge graph reasoning (Sha et al., 2023) enriches knowledge graphs with commonsense QA, improving commonsense reasoning. LiVersa (Ge et al., 2023) applies RAG to enhance medical QA for liver health diagnostics, while Almanac (Zakka et al., 2024) integrates clinical guidelines and treatment recommendations to support clinical decision-making. RAG also benefits educational decision-making through the assessment of tutoring practices (Han et al., 2024), which analyzes middle-school dialogue transcripts. It handles out-of-domain scenarios with textbook QA in the life sciences and earth sciences (Alawwad et al., 2024) and automates IT project requests via automated form filling (Bucur, 2023).

In the health sector, the frontline health worker capacity building system (Al Ghadban et al., 2023) uses RAG to process pregnancy-related guidelines. Self-BioRAG (Jeong et al., 2024) integrates biomedical instruction sets for improved biomedical informatics. Additionally, RAG contributes to the technical and humanitarian fields, with RAG-Fusion (Rackauckas, 2024) enhancing technical QA using product datasheets and FloodBrain Colverd et al. (2023) supporting flood disaster reporting with timely information. Further applications of RAG with LLMs have been explored in studies by Fan et al. (2024) and Arslan et al. (2024c, 2024d).

The studies highlight the wide-ranging applications of RAG in improving IR and generative tasks across sectors such as biomedical, financial, technical, and humanitarian fields. These advancements are represented by the development of QA assistants, such as chatbots, which establish how RAG-based approaches enable interactions with complex systems. However, the combination of RAG with LLMs in the urban water sector remains limited. By combining RAG with LLMs, organizations can efficiently access and synthesize data from diverse sources, addressing the key limitations of LLMs, including information misalignment, excessive complexity, and difficulties in accessing domain-specific data. This approach can create powerful QA systems that uncover critical insights for informed decision-making in UWM. By integrating various open data sources, such as news outlets, government

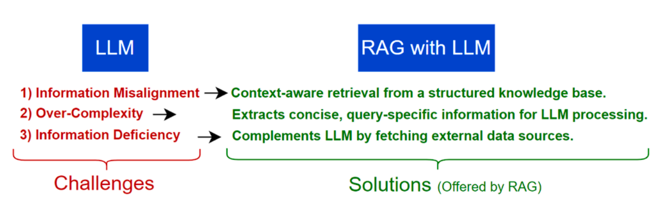

Fig. 2. Addressing LLM information challenges with RAG.

reports, and industry publications, RAG simplifies data aggregation and analysis. The following section will outline the proposed approach, using LLMs and RAG to develop a comprehensive IS that provides SMEs with an up-to-date repository of information on SUWIs.

3. Proposed system

This section starts by presenting the construction of a focused dataset to enhance an LLM integrated with a RAG-based IS, *SustainWater Bot*, a QA chatbot tailored to the urban water sector. The choice of QA chatbot is based on existing research highlighting its ability to improve user interaction and engagement with ISs (Arslan et al., 2024d). Chatbots streamline querying and IR, offering quick, accessible responses that are especially useful in complex fields, such as water management, where information must be communicated clearly and concisely (Nirala et al., 2022; Samuel et al., 2024; Jasmin et al., 2024). Following the dataset development process, this section presents the system's implementation using state-of-the-art tools and technologies. Finally, the system's performance is assessed by its ability to generate precise and relevant responses to user queries.

3.1. Dataset development

The dataset design began with a comprehensive literature review to support decision-making in SUWTs using the RAG approach with LLMs (Sharp, 2017). This process established the foundational scope and identified the core information needs within the urban water sector, especially for SMEs aiming to implement sustainable practices. Seven key information topics were identified as critical to this domain: funding opportunities, community initiatives, industry standards, technological innovations, emerging trends, regulatory guidelines, and case studies. These topics address the strategic and operational aspects of sustainability and are essential for evidence-based planning and action. Table 1 presents the significance of each topic and examples of relevant data sources. To construct the dataset, six primary information source types were utilised: government reports, industry publications, academic research, news articles, social media insights, and case studies. These were selected to ensure a diverse and balanced representation of perspectives across the urban water landscape. The collected datasets span multiple formats, including hypertext markup language (HTML) pages, portable document format (PDFs), and comma-separated values (CSV) files, as illustrated in Fig. 3. For the proof-of-concept implementation, the dataset was intentionally kept narrow to a curated set of authoritative data sources. Examples include UK Research and Innovation (UKRI, 2024) and Innovate UK (2024) for funding opportunities; Groundwork UK (2024) for community initiatives; BSI (2024) and Water UK (2024; 2025) for industry standards; The Water Report (Thewaterreport, 2024) for technological innovations; and BBC Science & Environment (Science & Environment, 2024) for emerging trends. Regulatory guidance is drawn from Ofwat (2024), Environment Agency (2024), and DEFRA (2024), while case studies are provided by UKWIR (2024).

Open data sources were selected based on transparent methodological criteria, including institutional authority, topical relevance, publication frequency, and UK-specific focus. Preference was given to regulatory bodies, government agencies, research institutions, and recognized industry publications known for reliability and consistent updates. Each topic area was supported by at least one independently verifiable source to ensure balanced representation and reduce bias. While the system architecture was designed to support future extensions and dynamic inputs, this prototype utilized a fixed, manually curated source set to support reproducibility and consistency in the evaluation. Data processing followed a structured pipeline encompassing source collection, cleaning, deduplication, and validation. Although automated scraping and crawling were not employed, all datasets were manually retrieved from official UK open data platforms to ensure precision and relevance. The collected documents underwent standardized cleaning, including the removal of formatting artifacts and conversion to plain text formats where needed. Figures were retained because the RAG-based architecture with LLMs can interpret both textual and visual information for downstream query tasks. Importantly, most of the collected documents included embedded metadata, such as publication dates and timestamps, specifically in HTML and PDF formats, thus eliminating the need for manual temporal mapping. A human-in-the-loop validation step was realized to verify the presence of this temporal metadata across all data sources. Because each source corresponded to a distinct topic and originated from a unique organization, the risk of data duplication was minimal. However, manual checks were conducted to ensure that the dataset remained free from redundancies and content overlap.

Table 1
Overview of topics with descriptions and examples.

No.	Topic	Description	Example of data
1	Funding Opportunities	Information on grants, subsidies, and other financial incentives available to support sustainable practices.	UKRI (2024), Innovate UK (2024), and Community Fund (2024).
2	Community Initiatives	SMEs can learn about grassroots movements and community-led water projects that they can support or collaborate with.	Groundwork UK (2024), Transitionnetwork (2024), and the wildlifetrust (2024).
3	Industry Standards	SMEs can adopt guidelines and benchmarks to improve water efficiency and sustainability.	The British Standards Institution BSI (2024), Water UK (2024; 2025), and the Construction Industry Research and Information Association (CIRIA, 2024).
4	Technological Innovations	Technologies and solutions in water management that SMEs can implement to enhance sustainability.	The Water Report (Thewaterreport, 2024), and Smartwatermagazine (2024).
5	Emerging Trends	Updates on water management, climate change, and urban sustainability provide insights into new challenges and opportunities.	Environment Agency (2024), the Department for Environment Food & Rural Affairs (DEFRA, 2024), The Guardian's Environment section (Environment, 2024), and Science and Environment (2024).
6	Regulatory Guidelines	Information on regulations and compliance requirements related to water usage, waste management, and sustainability practices.	DEFRA (2024), Ofwat (2024), Environment Agency (2024), Environment and climate change (2024), and Local Government Assoication (2024).
7	Case Studies	Successful initiatives in UWM, which can serve as models for SMEs.	UK Water Industry Research (Ukwir, 2024), Waterwise (2024), and local authorities.

The following section details the development of the *SustainWater Bot*, a QA IS built upon this collected dataset, designed to assist analysts in extracting relevant insights from the diverse and evolving body of knowledge surrounding sustainable UWM.

3.2. System implementation

This study developed an advanced version of the *SustainWater Bot* to equip SMEs with a powerful IS for exploring SUWIs, leveraging LLMs known for their capacity to perform complex reasoning across domains. Following a comprehensive evaluation of several state-of-the-art LLMs, including BLOOM (Le Scao et al., 2022), Falcon (Almazrouei et al., 2023), GPT-4 (Achiam et al., 2023), LLaMA (Touvron et al., 2023;

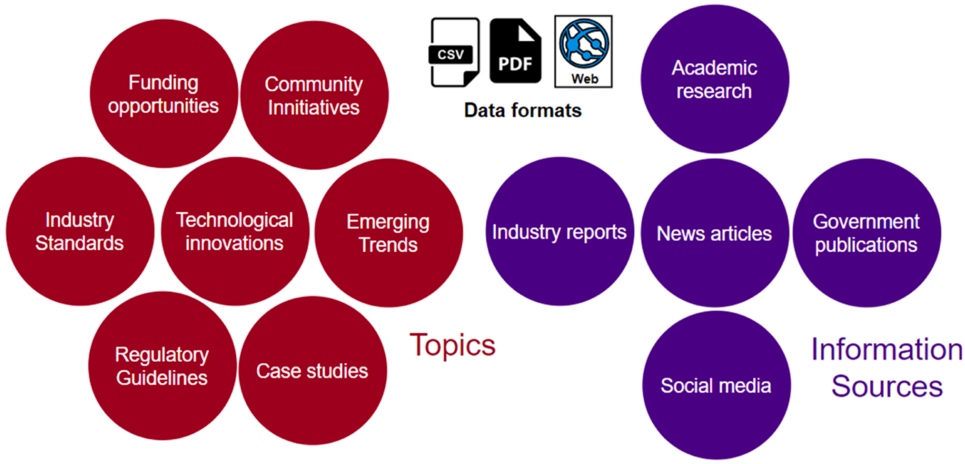

Fig. 3. Data collection on urban water topics from diverse sources.

Dubey et al., 2024), and Chinchilla (Hoffmann et al., 2022), three models were selected for integration: Llama3.2:3b (Touvron et al., 2023; Dubey et al., 2024), Gemma3:4b (Team et al., 2025), and Phi4-mini:3.8b (Abouelenin et al., 2025). These models were selected based on their strong performance in reasoning tasks, open-source availability, compact size (enabling efficient deployment), and extensive training on diverse datasets (de Curtò & de Zarzà, 2024). The use of models with relatively small but comparable parameter sizes allows for practical implementation on limited infrastructure while maintaining high-quality responses. Developers can select among these models based on specific resource constraints and performance needs; however, in this study, the three selected models were explicitly tailored for deployment in the urban water domain. The integration of RAG technology enables the *SustainWater Bot* to accurately interpret and respond to domain-specific queries by dynamically retrieving relevant content from curated datasets. The effectiveness of the system in addressing a wide range of UWM questions relies heavily on the quality, coverage, and semantic structure of the domain-specific datasets, which provide the contextual foundation for accurate and relevant responses.

The integration architecture outlined in Fig. 4 was implemented to adapt selected LLMs with RAG for urban water-specific QA tasks. This approach seamlessly combines LLMs with Hugging Face models (huggingface.co/models), which are renowned for their advanced NLP capabilities (Arslan et al., 2024a). The implementation begins with the installation of essential packages, including Transformers (Devlin et al., 2018), LangChain (2024), and LlamaIndex (2024), which are pivotal for model deployment, data handling, and embedding operations. LangChainEmbedding is specifically employed to transform text into vector representations (Thakur et al., 2020), forming the foundation for semantic understanding and IR. The workflow starts with information collection, where relevant datasets were gathered from diverse sources such as CSV files, PDFs, and HTML pages. These sources served as the foundation of an IS for addressing user queries and constructing a robust knowledge base. Once collected, the data undergoes preprocessing to ensure it is clean, structured, and ready for processing. During this phase, the data is divided into manageable "chunks", which break down large datasets into smaller, logical segments. This chunking process facilitates efficient storage, IR, and processing. Each chunk of data is then transformed into a numerical representation (i.e., embeddings) that captures the semantic meaning of the text. These embeddings are stored in a vector database, a specialized system optimized for managing and searching extensive collections of embeddings. The vector database serves as a foundation of RAG architecture, enabling rapid IR of the most relevant information required to address user queries accurately and efficiently. When a user submits a query through the system, it is converted into its embedding, which is then compared against the stored embeddings in the vector database using semantic search techniques. This process identifies the most relevant chunks of information to the

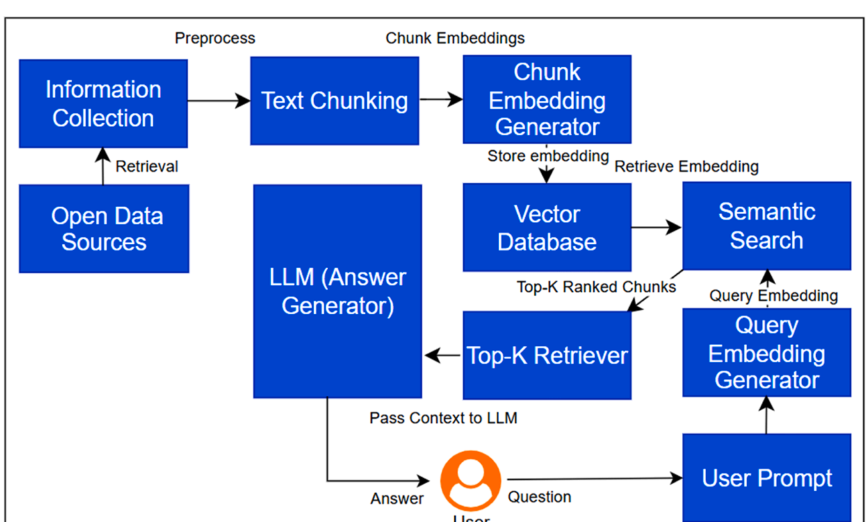

Fig. 4. Implementation of *SustainWater Bot* through RAG and LLM integration.

query. The retrieved chunks are ranked based on their relevance, ensuring that the most appropriate information is prioritized. These ranked results are passed to the LLM as contextual input, equipping it with highly relevant data for generating responses (Fig. 4).

Once the integration of RAG with LLM is completed for the development of the *SustainWater Bot*, it provides an advanced IS designed to support water analysts in conducting QA tasks related to SUWIs. The process (Fig. 5) begins with user interaction through the chatbot interface, where a water analyst initiates the process by posing a question. This query is routed to the system for processing and generating a relevant answer. *SustainWater Bot* improves the system's ability to deliver precise answers tailored to the specific needs of each query through RAG. The system is highly adaptable, allowing analysts to customize topics and data sources to suit their specific analytical needs. Its scope is defined by a diverse range of topics and information sources, which can be customized to meet particular analytical requirements. Analysts can enrich the system's database through targeted data curation efforts. For instance, one water analyst might focus on collecting data on industry standards, funding opportunities, community initiatives, and emerging trends, sourcing information from industry reports, academic studies, government publications, and social media. Simultaneously, another analyst may curate data on regulatory guidelines, technological innovations, and case studies using academic research, news articles, and governmental updates. Together, these collaborative efforts ensure that the centralized database remains accurate, up-to-date, and comprehensive. This customizability ensures the system's ongoing relevance for a wide array of applications in urban water sustainability.

3.3. System evaluation

A test dataset consisting of 105 manually created questions was developed to evaluate the system's performance. These questions were evenly distributed across the seven predefined urban water topics (Table 1), with 15 questions allocated to each topic to ensure thematic balance. The question design process was guided by the content within the constructed dataset linked to the LLM through RAG, ensuring that each question was contextually grounded in the available source material. A diversity of query types, including fact-based (i.e., seeking objective information), policy-related (i.e., rules, regulations, or governance), temporal (i.e., involving time), and entity-specific questions (i.e., focused on people, places, or organizations), was considered, thereby reflecting the range of information needs relevant to SMEs and analysts working within SUWTs. Although the dataset was manually created, an internal validation process ensured coverage across the full topical spectrum, avoiding the overrepresentation of any single source. Although this proof-of-concept evaluation did not focus on external validity, the methodology can be extended to incorporate domain experts and validated questions in future iterations.

To assess model performance on domain-specific QA tasks, a constructed question set was evaluated using four key metrics: precision, completeness, exact match, and average response time (RT). Precision (%), defined in *Eq. (1)*, reflects how accurately each method captured critical information units such as Named Entities (e.g., "£4.1 billion"), Core Concepts (e.g., "low-income households"), and Modifiers (e.g., "140 % increase", "£10/month"). Completeness (%), defined in *Eq. (2)*, measures the proportion of essential information included in a model's response (Le Bronnec et al., 2024). The exact match (%), defined in *Eq. (3)*, is the strictest metric, requiring complete coalition with gold-standard information units, not just partial matches. The computational efficiency (i.e., RT) of each method is quantified in seconds (s) as defined in *Eq. (4)*. Based on these evaluation criteria, a performance comparison was conducted in two stages using the same benchmark question set. In the first stage, the traditional IR techniques, TF-IDF and BM25, were assessed against the gold standard answers in terms of response quality and average RT. Table 2 presents a sample of 6 out of the 105 manually curated questions. These approaches demonstrated limited ability to handle semantically nuanced queries, highlighting the inherent constraints of lexical matching in complex domains such as SUWIs. In the second stage, three RAG-enabled LLMs (Llama3.2:3b, Gemma3:4b, and Phi4-mini:3.8b) were evaluated using the same benchmark, a representative sample of 6 questions selected from the full set of 105, as stated in Table 3. This stage enabled a more detailed analysis of the model outputs, particularly in terms of semantic relevance and adaptability to contextual queries. Finally, Table 4 presents a comparative summary of all evaluated methods, both traditional IR and LLM-based, across the four key metrics. This structured, two-stage evaluation reveals important trade-offs and emerging advantages of LLMs with RAG architectures, particularly in supporting efficient, accurate, and scalable QA for SMEs operating in the urban water sector.

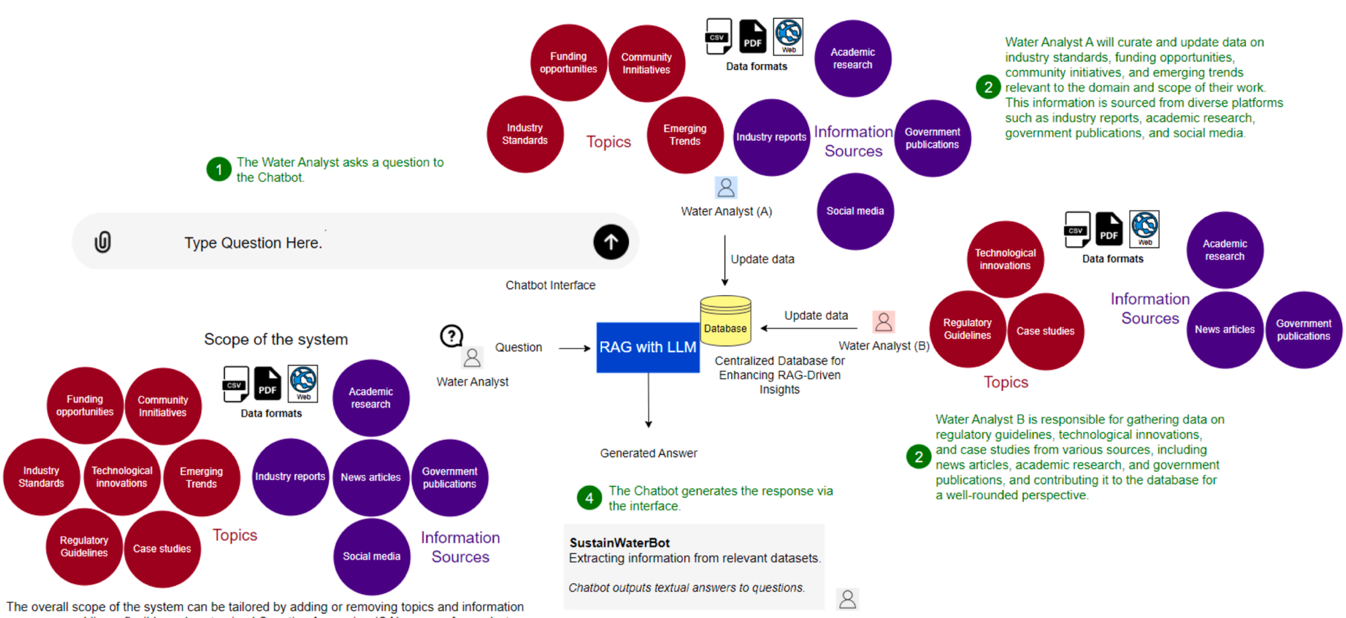

Fig. 5. *SustainWater Bot* using RAG with LLM-based QA system.

Table 2
Comparative performance and response time (RT) analysis of TF-IDF and BM25 approaches against gold standard QA results (Water UK, 2025).

No.	Question	Gold Standard Answer	TF-IDF		BM25	
			Response	RT (s)	Response	RT (s)
1	How much financial support will water companies provide for low-income households?	£4.1 billion will be allocated over five years, representing a 140 % real-terms increase compared to the previous five-year period.	Over 3 million households are set to benefit from reduced water bills as companies commit to a record level of investment.	0.003	Over 3 million households are set to benefit from reduced water bills as companies commit to a record level of investment.	0.002
2	What types of support are available to customers struggling with water bills?	Support includes social tariffs, WaterSure charges, payment breaks, debt forgiveness, and services via the Priority Services Register.	Customers facing financial difficulties can access various support options, such as payment breaks and debt relief	0.003	Customers facing financial difficulties can access various support options, such as payment breaks and debt relief.	0.002
3	What is the Priority Services Register?	It offers additional non-financial support for vulnerable groups, including people with disabilities and families with babies under 12 months.	A variety of assistance, beyond just financial aid, is available through the Priority Services Register for people with disabilities and parents of babies under one year old.	0.003	A variety of assistance, beyond just financial aid, is available through the Priority Services Register for people with disabilities and parents of babies under one year old.	0.002
4	What will be the average annual water bill after April 2025?	The average bill will be around £603 per year, or £1.65 per day.	This represents a £123 increase (26 %) compared to the previous year, with the average annual bill now at £603.	0.003	This represents a £123 increase (26 %) compared to the previous year, with the average annual bill now at £603.	0.002
5	Are 2025 water bills significantly higher than previous years?	Bills will only be about 5 % higher in real terms than in 2010.	Despite this increase, water bills remain just 5 % higher in real terms compared to 2010.	0.003	>3 million households are set to receive reduced bills and financial aid worth up to £4.1 billion over the next five years, representing a 140 % real-term increase compared to the last five years.	0.002
6	Compare the water bill increases in 2025 with the changes in real terms since 2010.	Bills rise by about £10/month in 2025 but are only 5 % higher in real terms than in 2010.	This unprecedented investment comes after more than a decade of real-term decreases in bills.	0.003	This unprecedented investment comes after more than a decade of real-term decreases in bills.	0.002

$$Precision\ (\%) = \frac{Number\ of\ Correctly\ Retrieved\ Information\ Units}{Total\ Number\ of\ Retrieved\ Information\ Units} \times 100 \quad (1)$$

$$Completeness\ (\%) = \frac{Matched\ Information\ Units}{Total\ Gold\ Standard\ Information\ Units} \times 100 \quad (2)$$

$$Exact\ Match = \begin{cases} 1, & if\ Matched\ Information\ Units = Total\ Gold\ Standard\ Information\ Units \\ 0, & otherwise \end{cases} \quad (3)$$

$$Average\ Response\ Time\ (RT)\ (s) = \frac{\sum Time\ per\ Response}{Total\ Number\ of\ Responses} \quad (4)$$

Llama3.2 leads with a precision of 95 %, followed closely by Gemma3 at 90 %. Phi4-mini achieves a lower precision at 75 %, whereas TF-IDF and BM25 lag at 70 % and 65 %, respectively. Again, Llama3.2 and Gemma3 perform strongly with 95 % and 90 % completeness, respectively. Phi4-mini follows at 80 %, whereas TF-IDF and BM25 fall short at 65 % and 60 %, often omitting supporting details. Llama3.2 achieves 90 % exact match, indicating highly consistent and thorough responses. Gemma3 records 80 % and Phi4-mini 60 %, showing a greater tendency toward partially correct answers. TF-IDF and BM25 trail with 40 %, reinforcing their limitations in capturing complete semantic content. The traditional IR methods are speedy (under 0.005 s), with BM25 slightly outperforming TF-IDF in speed. Among the LLMs, Llama3.2 is the fastest (2.29 s), followed by Gemma3 (2.91 s), while Phi4-mini is notably slower at 4.69 s, indicating its higher cost for more detailed outputs. Llama3.2:3b consistently ranks highest in all quality metrics, making it the most effective model in both precision and completeness with strong semantic coverage. Gemma3:4b performs similarly well but is slightly less consistent on the exact match. Phi4-mini:3.8b offers a good middle ground with solid completeness and acceptable precision but trades speed for depth. In contrast, TF-IDF and BM25, while fast, fall short in producing accurate and complete responses, making them less suitable for high-quality answer generation.

4. Discussion

First, it discusses how the proposed approach alleviates fragmented access to SUWI-related knowledge. Next, it evaluates model performance across various question types and examines the original hypotheses' outcomes. The discussion also emphasizes the influence of data quality on QA effectiveness and concludes with limitations and potential directions for future research.

4.1. Fragmented access to SUWI knowledge

SMEs lack an integrated IS to access and consolidate dispersed SUWI data, impeding progress on SUWTs. Existing systems often miss critical information, require significant development effort, and are costly (Cosgrove & Loucks, 2015). To address these limitations, this study offers a solution, i.e., the *SustainWater Bot*. This system integrates advanced LLMs with RAG technology to address common LLM limitations, such as information misalignment, excessive complexity, and information deficiencies. For instance, using the official document from (BGS, 2025), the first question in Fig. 6 demonstrates that the LLM generates an overly detailed yet imprecise response. Moreover, the LLM fails to produce an answer for the second question due to limitations in

Table 3
Comparative performance and response time (RT) analysis of LLMs against gold standard QA results (Water UK, 2025).

No.	Question	Gold Standard Answer	LLM: Llama3.2:3b		LLM: Gemma3:4b		LLM: Phi4-mini:3.8b	
			Response	RT (s)	Response	RT (s)	Response	RT (s)
1	How much financial support will water companies provide for low-income households?	£4.1 billion will be allocated over five years, representing a 140 % real-terms increase compared to the previous five-year period.	>3 million households are set to receive reduced bills and up to £4.1 billion in financial support over the next five years, representing a 140 % real-term increase compared to the last five years.	1.66	Water companies will provide £4.1 billion in financial support for low-income households over the next five years.	1.01	Water companies will provide £4.1 billion in financial support over five years as part of affordability packages including reduced bills specifically targeting >3 million, predominantly lower-income homes through 'social tariffs' or 'WaterSure' charges among other forms of assistance.	3.02
2	What types of support are available to customers struggling with water bills?	Support includes social tariffs, WaterSure charges, payment breaks, debt forgiveness, and services via the Priority Services Register.	Support measures for customers struggling with water bills include: Payment breaks, Debt forgiveness, 'Social tariffs' or 'WaterSure' charges that will reduce their water bills (targeted at low-income customers who use a lot of water for essential family or health reasons). Additionally, customers are encouraged to contact their water company directly to see what help is available.	1.96	The following types of support are available to customers struggling with water bills: Social tariffs, WaterSure, charges (specifically for low-income customers who use a lot of water for essential family or health reasons), Payment breaks, Debt forgiveness, A variety of measures through the Priority Services Register for customers with difficulties like sight, hearing, or mobility, or parents with babies under 12 months old.	3.01	Customers who struggle with their water bills can receive a range of financial and non-financial supports including 'social tariffs', 'WaterSure' charges that reduce the bill amount for eligible households. Additionally, there may be payment breaks (pausing payments), debt forgiveness measures available to help customers manage or eliminate outstanding debts related to unpaid utility services.	6.08
3	What is the Priority Services Register?	It offers additional non-financial support for vulnerable groups, including people with disabilities and families with babies under 12 months.	The Priority Services Register (PSR) appears to be a scheme that provides support services for vulnerable customers, including those with sight, hearing, or mobility difficulties, as well as parents with babies under 12 months old.	1.75	The Priority Services Register is designed to assist a broad spectrum of people, from those with sensory or mobility challenges to parents caring for babies under one year old.	1.12	The Priority Services Register (PSR) provides a variety of measures for people who need them, including those with sight or hearing difficulties as well as parents caring for babies under 12 months old. These services are not limited solely to financial assistance but also encompass other forms of support aimed at improving accessibility and quality of life for various individuals within the community.	1.78
4	What will be the average annual water bill after April 2025?	The average bill will be around £603 per year, or £1.65 per day.	The text does not explicitly state what the new average annual water bill will be. It mentions that customer bills will rise by about £10 a month from April 2025, and provides an example of the increased bill for households in England and Wales (£603) and businesses (£1.32 per day), but it does not provide information on the total amount or percentage increase.	1.51	The average annual water bill will be £603.	0.68	£603	1.18
5	Are 2025 water bills significantly higher than previous years?	Bills will only be about 5 % higher in real terms than in 2010.	Despite a £10 monthly increase starting April 2025, amounting to a £123 (26 %) rise over the previous year, water bills will be only 5 % higher in real terms compared to 2010, making the overall increase relatively modest.	1.67	Water bills will rise by around £10 monthly from April 2025, marking a £123 (26 %) increase over last year. The average annual bill will reach £603.	1.96	Yes, starting April 2025 customer bills will rise by about £10 per month. Despite this increase being equivalent to an average annual bill set at £603 (26 % on the previous year), in real terms they are only around 5 % higher compared to what customers paid back in 2010.	3.25
6	Compare the water bill increases in 2025 with the changes in real terms since 2010.	Bills rise by about £10/month in 2025 but are only 5 % higher in real terms than in 2010.	From April 2025, customers can expect their bills to rise by around £10 monthly, a 26 % increase (£123) over last year. Adjusted for inflation,	1.66	From April 2025, customer bills will rise by about £10 a month, which equates to an increase of £123 (26 %). In real terms, water bills are	2.14	In April 2025, customer bills will rise by about £10 per month. Despite this increase being noticeable on a monthly basis and an annual total of	4.24

(continued on next page)

Table 3 (*continued*)

No.	Question	Gold Standard Answer	LLM: Llama3.2:3b		LLM: Gemma3:4b		LLM: Phi4-mini:3.8b	
			Response	RT (s)	Response	RT (s)	Response	RT (s)
			water bills are still only about 5 % higher than they were in 2010.		now only around 5 % higher than they were in 2010.		around £123 (26 %), when adjusted for inflation to measure it against past costs from similar timespans or the baseline year provided in real terms since 2010—water bills are only approximately 5 % higher than they were back then, effectively making them comparable.	

Table 4
Performance comparison of information retrieval and LLMs on QA task.

No.	Method	Precision (%)	Completeness (%)	Exact Match (%)	Response Time (s)
1	TF-IDF	70	65	40	0.003
2	BM25	65	60	40	0.002
3	Llama3.2:3b	95	95	90	2.29
4	Gemma3:4b	90	90	80	2.91
5	Phi4-mini:3.8b	75	80	60	4.69

its training data. The *SustainWater Bot* semantically integrates diverse information sources by utilizing the RAG, offering a comprehensive and highly relevant information base to support decision-makers in the urban water sector.

4.2. Evaluating performance across question types rather than topics

The evaluation of the *SustainWater Bot* employed a diverse set of question types (see Table 3), including fact-based, list-based, comparative, summary, and impact-driven questions, to assess the performance of small-scale RAG-enabled LLMs (Llama3.2, Gemma3, Phi4-mini). The model's performance was influenced more significantly by the type of reasoning tasks it was asked to perform rather than the specific topic domain involved. This means that whether the question related to funding opportunities, community initiatives, industry standards, technological innovations, emerging trends, regulatory guidelines, or case studies signified less than whether the model needed to engage in reasoning, inference, or pattern recognition. Fact-based and list-type questions (e.g., Q1–Q2, Q4 in Table 3) were answered well by all models, particularly Llama3.2 and Gemma3. In contrast, impact-oriented and comparative questions (e.g., Q5–Q6) were more challenging and exposed meaningful differences. Phi4-mini excelled in nuanced reasoning and regulatory interpretation, Gemma3 offered the most balanced and structured summaries, and Llama3.2 was fastest but less robust on multi-layered logic.

4.3. Hypothesis validation and insights

Table 4 provides empirical support for Hypothesis A, confirming that RAG-enabled LLMs deliver more accurate and contextually relevant responses than traditional IR methods, such as TF-IDF and BM25. Precision and completeness scores were consistently higher across the evaluated models, suggesting that the semantic integration and generative capabilities of LLMs significantly enhance domain-specific IS. Hypothesis B is also supported: the constructed, semantically structured dataset, balanced across seven SUWT themes, enabled coherent responses, indicating that the system can help users, especially SMEs, form more actionable ISs within the sustainability landscape. The quality and clarity of its responses suggest strong potential for supporting users in understanding complex urban water issues. Although a direct evaluation of the IS's impact on user cognition was beyond the scope of this study. Finally, the system design aligns with Hypothesis C, demonstrating how a modular and adaptable architecture can support the long-term sustainability of an IS. The *SustainWater Bot* demonstrates cost-effectiveness and environmental sustainability (Beloskar et al., 2023; Weidinger et al., 2021). Rather than relying on resource-intensive end-to-end model training, the system integrates pre-trained, open-source LLMs via RAG, significantly reducing development time, computational overhead, and financial barriers, particularly for SMEs. This modular structure enables easy updates by simply replacing data sources, eliminating the need for full retraining or dependence on proprietary platforms. Although this proof-of-concept did not include detailed carbon accounting, existing studies have estimated that LLM inference typically consumes approximately 0.0003 kW-hour (kWh) per query (Maliakel et al., 2025; Jegham et al., 2025). In contrast, training LLMs from scratch can require thousands of kWh, leading to a significantly higher carbon footprint (Weidinger et al., 2021). By leveraging lightweight, domain-specific adaptations of pre-trained models, *SustainWater Bot* achieves substantial energy savings.

4.4. The critical role of data quality

Beyond the adaptability, cost-effectiveness, and modular architecture of the *SustainWater Bot*, the quality and relevance of its outputs are fundamentally dependent on the quality of the data it processes. It is essential to ensure that the input data meets the rigorous criteria outlined in Table 5 to sustain and enhance its effectiveness. These factors directly influence the system's ability to generate meaningful insights (Liu et al., 2024; Alan et al., 2024; Dam et al., 2024; Dubey et al., 2024). The results emphasize the need for systematic, high-quality dataset construction, although strict data selection was not the focus. In future work, robust methods should be developed to ensure reliable and context-aware chatbot performance in sustainable UWM.

4.5. Limitations and directions for future research

This study has several limitations that need to be acknowledged and addressed in future work.

1) *Manual Data Collection and Source Scope*: The dataset used for the RAG was collected manually from open data sources. Automated methods, such as web scraping or real-time application programming interfaces (APIs) integration, were not implemented. The system will incorporate programmatic data ingestion via APIs and automated parsers in future iterations, alongside periodic human validation to ensure data quality and enable real-time system updates. Additionally, while the *SustainWater Bot* focuses on a selected set of open-access UK-based sources, many valuable data streams remain unexploited. These include public opinion surveys (YouGov, 2024), market intelligence (IBISWorld UK, 2024; The Economist, 2024),

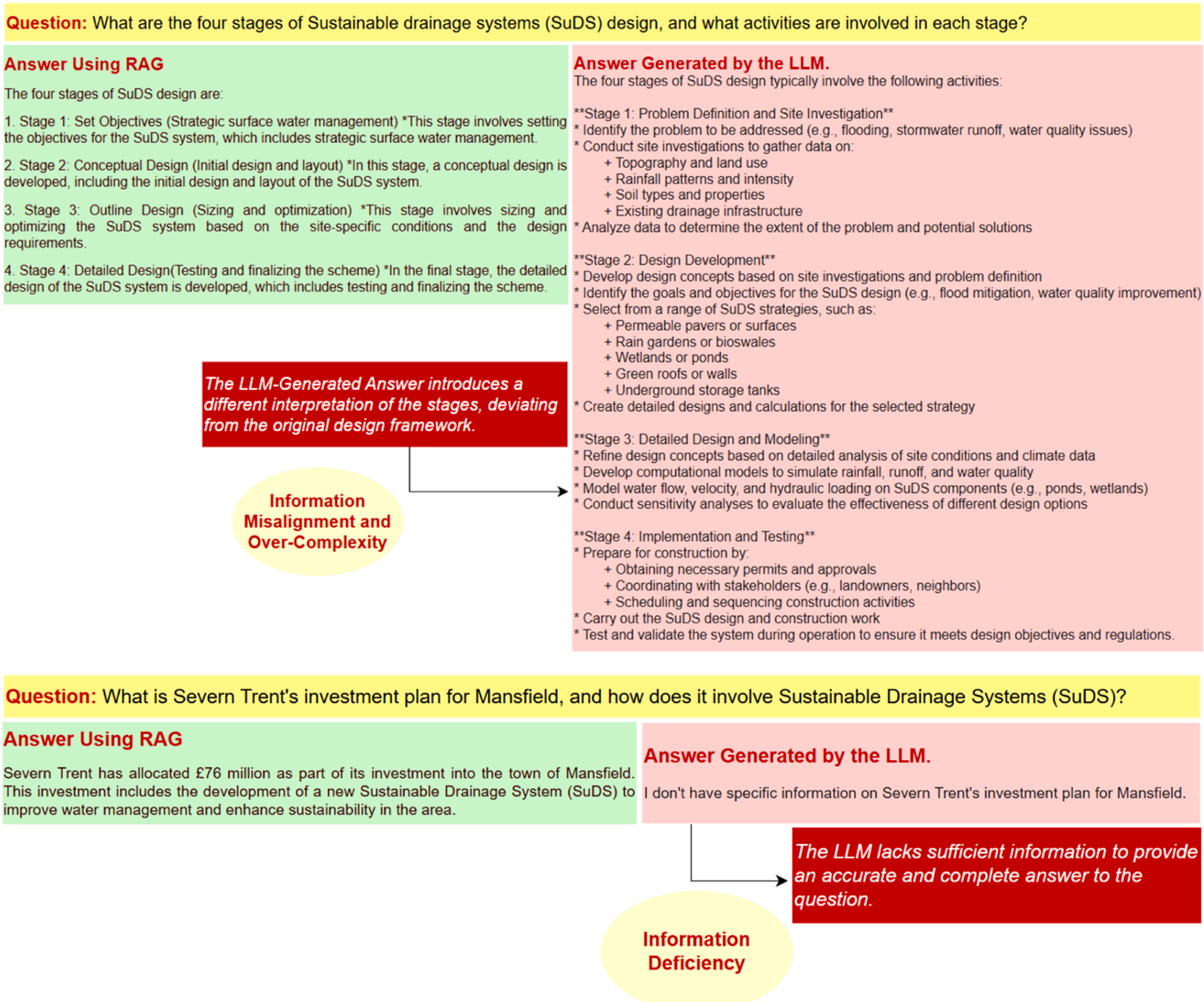

Fig. 6. *SustainWater Bot* addressing information misalignment, over-complexity, and information deficiencies in IR.

scientific data (UKCEH, 2024; BGS, 2024; The Royal Society, 2024), and long-term projections from research institutions (MetOffice, 2024; National Infrastructure Commission, 2024; UKCIP, 2024). Public sentiment and real-time operational data (Assets Publishing Service, 2023; Greenpeace, 2024) could also provide SMEs with rich insights. Integrating these sources would significantly expand the practical relevance of the system.

2) *Open* vs. *Proprietary Data Integration*: The current implementation relies exclusively on open-access datasets to ensure SMEs' broad accessibility and cost-effectiveness. However, the system architecture is designed to be modular and extensible, thereby supporting future integration of proprietary or subscription-based sources. This hybrid model, which combines open and licensed datasets, could significantly enhance the system's decision-making capabilities by incorporating high-resolution news data, detailed commercial analytics, or restricted policy documents.
3) *Geographic and Domain Specificity:* The *SustainWater Bot* has been specifically tailored to the UK context, drawing primarily on UK-based public datasets relevant to SUWTs. Although this targeted approach enhances the relevance for local decision-makers, it also introduces a potential limitation in terms of external validity. The current knowledge base of the system may reduce its applicability in other regional contexts. However, the system's modular architecture supports straightforward adaptation. By replacing the underlying data sources with region-specific or domain-relevant datasets (either open-access or proprietary), the chatbot can be extended to support water-related decision-making in other countries or entirely different domains, such as energy, waste, or agriculture. To improve relevance, SME experts who understand local and sectoral priorities should be involved in dataset selection. Since SME interests can vary, future work should explore methods to balance diverse perspectives and precisely integrate them into the data selection process.
4) *LLM Selection: SustainWater Bot* currently uses three selected LLMs for all RAG-based operations. These models were selected based on accessibility and baseline performance. However, future work should explore a broader range of LLMs, both open-source and commercial, to identify those best suited for specific query types, data characteristics, or operational limitations.
5) *Scalability, System Architecture, and Response Latency:* The current implementation of the *SustainWater Bot* supports only one user session at a time and operates on a basic input–output loop, which limits its scalability in multi-user settings. Although the response latency was reasonable, averaging between 2s and 4s across the tested prompts, the system is not yet optimized for concurrent access.

Table 5
Key factors influencing the quality of information in the Chatbot.

No.	Factor	Description
1	Relevance	■ *Domain-Specific Data*: Ensure that the selected data aligns with the specific goals of the system. ■ *Contextual Appropriateness*: Select data that directly answers or is related to the anticipated questions.
2	Quality and Credibility of Sources	■ *Peer-Reviewed Articles*: Prefer information from reputable academic journals. ■ *Official Data Sources*: Government publications, standards organizations, and established think tanks should be prioritized. ■ *Industry Reports and Case Studies*: Use well-regarded industry reports and case studies that are specific to the context of the system's focus.
3	Recency and Timeliness	■ *Up-to-Date Information*: Prioritize recent sources, especially for fast-evolving fields like technology or policy. ■ *Historical Context*: If historical data is important for comparison, make sure to distinguish between recent and older sources based on relevance.
4	Diversity and Coverage	■ *Variety of Perspectives*: Use a diverse set of sources to avoid bias, including academic, industry, and governmental perspectives. ■ *Geographical Representation*: Ensure the data includes regional differences, especially for fields like sustainability, which may have global or industry-specific variations.
5	Ethical and Legal Considerations	■ *Ethically Sourced Data*: Ensure the data is ethically collected and respects privacy laws. ■ *Licensing*: Verify that the data can be used legally, especially if feeding proprietary information into the system.
6	Data Verifiability	■ *Citations*: Prefer data that is well-cited, allowing users to trace the information back to its origin.
7	Granularity and Detail	■ *Sufficient Detail*: Ensure that the data is sufficiently granular to answer detailed queries.

Future iterations should adopt an asynchronous representational state transfer (RESTful) architecture, introducing stateless API endpoints, session management, and backend concurrency controls to support simultaneous users. Additionally, the system's reliance on the RAG technology may introduce additional latency as the dataset scales. Although only a selected set of sources is currently integrated, expanding the number of data sources or increasing query volumes could result in slower RTs. This limitation underlines the need for optimization of performance.

6) *System Evaluation*: The system was tested with a limited set of questions, which may not fully reflect the diversity of urban water queries. Future evaluations should include a broader range of real-world and user-generated questions for greater robustness. However, this proof-of-concept system did not focus on external validation. The evaluation framework can be expanded to include domain expert–validated queries, enabling more generalizable insights across various use cases and user groups. Such an approach will strengthen the credibility and applicability of the model in practical deployments.
7) *Performance and Sustainability Considerations*: Although the current implementation emphasizes cost-effectiveness, further research should measure the system's energy consumption per query and evaluate the environmental impact. Benchmarking against commercial models would help assess the trade-offs between performance, cost, and sustainability.

5. Conclusions

SMEs play a crucial role in the UK's journey towards net-zero emissions, especially by promoting efficient urban water systems. However, many SMEs face a substantial knowledge gap in sustainable UWM, leading to inefficiencies and higher costs. To make well-informed decisions, timely and comprehensive access to information is required, covering government policies, technological advances, market trends, and environmental considerations. Unfortunately, current systems often struggle to unify these diverse data sources, hindering SMEs' ability to implement sustainable practices effectively. This study presents the *SustainWater Bot*, an innovative LLM-based IS driven by RAG technology, designed to address challenges such as information misalignment, complexity, and data insufficiency. The system delivers relevant insights into the urban water sector while leveraging open-source technologies to reduce the carbon footprint of its development. Although promising, the *SustainWater Bot* requires further evaluation to refine its precision and efficiency. Future work should address challenges such as potential delays caused by incorporating additional data sources and exploring alternative LLM models to optimize performance. By addressing these areas, the system can better empower SMEs to achieve SUWTs.

Declaration of generative AI and AI-assisted technologies

As this study incorporates the use of Large Language Models (LLMs), the responses presented in Tables 2 and 3 were generated using Llama3.2, Gemma3, and Phi4-mini. These examples illustrate each model's capability to interpret and produce contextually relevant answers based on a curated dataset sourced from Water UK (2025), highlighting their applicability within the study's framework. To improve the clarity and coherence of the manuscript's language, the GPT-4-turbo model was used for proofreading and refinement purposes. After utilizing this tool, the authors thoroughly reviewed the language and accept full responsibility for the manuscript's content.

Funding details

The authors express their gratitude to the *University of the West of England* for funding this research.

Data availability

The dataset used in this study was sourced from open-access platforms, as detailed in Section 3 of this article.

CRediT authorship contribution statement

Muhammad Arslan: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis. **Saba Munawar:** Writing – review & editing, Data curation. **Zainab Riaz:** Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Muhammad Arslan reports financial support was provided by University of the West of England. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank Professor Jessica Lamond from the University of the West of England (UWE Bristol) for her generous time and valuable feedback on an early version of this manuscript and the proposed system. The authors also gratefully acknowledge UWE Bristol for funding this research.

Data availability

The data is available online. Only open datasets have been used.

References

Abouelenin, A., Ashfaq, A., Atkinson, A., Awadalla, H., Bach, N., Bao, J., & Zhou, X. (2025). Phi-4-mini technical report: Compact yet powerful multimodal language models via mixture-of-loras. arXiv preprint arXiv:2503.01743.

Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F.L., & McGrew, B. (2023). Gpt-4 technical report. arXiv preprint arXiv:2303.08774.

Adnan, K., & Akbar, R. (2019). Limitations of information extraction methods and techniques for heterogeneous unstructured big data. *International Journal of Engineering Business Management, 11*, Article 1847979019890771.

Aizawa, A. (2003). An information-theoretic perspective of tf–idf measures. *Information Processing & Management, 39*(1), 45–65.

Álvarez Jaramillo, J., Zartha Sossa, J. W., & Orozco Mendoza, G. L. (2019). Barriers to sustainability for small and medium enterprises in the framework of sustainable development—Literature review. *Business Strategy and the Environment, 28*(4), 512–524.

Al Ghadban, Y., Lu, H.Y., Adavi, U., Sharma, A., Gara, S., Das, N., & Hirst, J.E. (2023). Transforming healthcare education: Harnessing large language models for frontline health worker capacity building using retrieval-augmented generation. medRxiv, 2023-12.

Alan, A.Y., Karaarslan, E., & Aydin, O. (2024). A RAG-based question answering system proposal for understanding islam: Mufassirqas llm. arXiv preprint arXiv:2401.15378.

Alawwad, H.A., Alhothali, A., Naseem, U., Alkhathlan, A., & Jamal, A. (2024). Enhancing textbook question answering task with large language models and retrieval augmented generation. arXiv preprint arXiv:2402.05128.

Almazrouei, E., Alobeidli, H., Alshamsi, A., Cappelli, A., Cojocaru, R., Debbah, M., ... & Penedo, G. (2023). The falcon series of open language models. *arXiv preprint arXiv: 2311.16867*.

Amati, G. (2009). *BM25. In encyclopedia of database systems* (pp. 257–260). Boston, MA: Springer.

Arefeen, M. A., Debnath, B., & Chakradhar, S. (2024). Leancontext: Cost-efficient domain-specific question answering using llms. *Natural Language Processing Journal, 7*, Article 100065.

Arslan, M., Munawar, S., & Cruz, C. (2024a). Business insights using RAG–LLMs: A review and case study. *Journal of Decision Systems*, 1–30. https://doi.org/10.1080/12460125.2024.2410040

Arslan, M., Mahdjoubi, L., & Munawar, S. (2024b). Driving sustainable energy transitions with a multi-source RAG-LLM system. *Energy and Buildings, 324*, Article 114827. https://doi.org/10.1016/j.enbuild.2024.114827

Arslan, M., Ghanem, H., Munawar, S., & Cruz, C. (2024c). A survey on RAG with LLMs. *Procedia Computer Science, 246*, 3781–3790. https://doi.org/10.1016/j.procs.2024.09.178

Arslan, M., Munawar, S., & Riaz, Z. (2024d). Sustainable urban water decisions using generative artificial intelligence. In *2024 International Conference on Decision Aid Sciences and Applications (DASA)* (pp. 1–5). IEEE. https://doi.org/10.1109/DASA63652.2024.10836374.

Assets Publishing Service. (2023). List of councils in England by type. Retrieved from https://assets.publishing.service.gov.uk/media/6401d6a6d3bf7f25f813fc3a/List_of_councils_in_England_2023.pdf Accessed August 1, 2026.

Bataille, C. G. (2020). Physical and policy pathways to net-zero emissions industry. *Wiley Interdisciplinary Reviews: Climate Change, 11*(2), e633.

Beloskar, V. D., Haldar, A., & Rao, S. N. (2023). Socially responsible investments: A retrospective review and future research agenda. *Business Strategy and the Environment, 32*(7), 4841–4860.

BGS. (2024). Understanding our Earth. Retrieved from https://www.bgs.ac.uk/ Accessed July 19, 2024.

BGS (2025), Guide to Suds for Developers *British Geological Survey*, www.bgs.ac.uk/download/guide-to-suds-for-developers/. Accessed 5 Sept. 2025.

Borsacchi, L., & Pinelli, P. (2020). *Sustainable and innovative practices of small and medium-sized enterprises in the water and waste management sector* (pp. 255–290). Innovation Strategies in Environmental Science.

BSI. (2024). Sustainability. Retrieved from https://www.bsigroup.com/en-GB/ Accessed July 09, 2024.

Bucur, M. (2023). *Exploring Large Language Models and Retrieval Augmented Generation for Automated Form Filling (Bachelor's thesis*. University of Twente.

Chadfield, S. J., Wei, Y., & Lieske, S. N. (2024). Water sensitive communities: A systematic review with a complex adaptive systems perspective. *Journal of Environmental Planning and Management, 67*(5), 1077–1103.

Chang, Y., Wang, X., Wang, J., Wu, Y., Yang, L., Zhu, K., & Xie, X. (2024). A survey on evaluation of large language models. *ACM Transactions on Intelligent Systems and Technology, 15*(3), 1–45.

CIRIA. (2024). Ciria. Retrieved from *https://www.ciria.org/* Accessed July 09, 2024.

Colverd, G., Darm, P., Silverberg, L., & Kasmanoff, N. (2023). FloodBrain: Flood disaster reporting by web-based retrieval augmented generation with an llm. arXiv preprint arXiv:2311.02597.

Community Fund. (2024). Community Fund. Retrieved from https://www.tnlcommunityfund.org.uk/ Accessed July 29, 2024.

Cosgrove, W. J., & Loucks, D. P. (2015). Water management: Current and future challenges and research directions. *Water Resources Research, 51*(6), 4823–4839.

Dam, S.K., Hong, C.S., Qiao, Y., & Zhang, C. (2024). A complete survey on llm-based ai chatbots. arXiv preprint arXiv:2406.16937.

Daniell, K. A., Rinaudo, J. D., Chan, N. W. W., Nauges, C., & Grafton, Q. (2015). *Understanding and managing urban water in transition* (pp. 1–30). Springer Netherlands.

de Curtò, J., & de Zarzà, I. (2024). Comparative analysis of reasoning capabilities in foundation models. In *2024 2nd International Conference on Foundation and Large Language Models (FLLM)* (pp. 141–149). IEEE.

Department for Environment Food & Rural Affairs (DEFRA). (2024). Retrieved from https://www.gov.uk/government/organisations/department-for-environment-food-rural-affairs Accessed July 09, 2024.

Devlin, J., Chang, M.W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv: 1810.04805.

Dubey, A., Jauhri, A., Pandey, A., Kadian, A., Al-Dahle, A., Letman, A., & Ganapathy, R. (2024). The llama 3 herd of models. arXiv preprint arXiv:2407.21783.

Dunn, G., Brown, R. R., Bos, J. J., & Bakker, K. (2017). The role of science-policy interface in sustainable urban water transitions: Lessons from Rotterdam. *Environmental Science & Policy, 73*, 71–79.

Environment. (2024). Environment. Retrieved from https://www.theguardian.com/uk/environment Accessed July 09, 2024.

Environment Agency. (2024). Retrieved from *https://www.gov.uk/government/organisations/environment-agency* Accessed July 09, 2024.

Environment and climate change. (2024). Environment and climate change. Retrieved from https://www.gov.scot/environment-and-climate-change/ Accessed July 29, 2024.

Fan, W., Ding, Y., Ning, L., Wang, S., Li, H., Yin, D., & Li, Q. (2024). A survey on rag meeting llms: Towards retrieval-augmented large language models. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (pp. 6491–6501).

Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., & Wang, H. (2023). Retrieval-augmented generation for large language models: A survey. arXiv preprint arXiv: 2312.10997, 2(1).

Garajeh, M. K., Feizizadeh, B., Salmani, B., & Ghasemi, M. (2024). Analyzing urban drinking water system vulnerabilities and locating relief points for urban drinking water emergencies. *Water Resources Management, 38*(7), 2339–2358.

Garcia, C., López-Jiménez, P. A., Pérez-Sánchez, M., & Sanchis, R. (2024). *Methodology for assessing progress in sustainable development goals indicators in urban water systems. How far are we from the 2030 targets?* Sustainable Cities and Society, Article 105616.

Ge, J., Sun, S., Owens, J., Galvez, V., Gologorskaya, O., Lai, J.C., & Lai, K. (2023). Development of a liver disease-specific large language model chat interface using retrieval augmented generation. medRxiv.

Greenpeace. (2024). Take the 25 mile challenge. Retrieved from *https://www.greenpeace.org.uk/* Accessed July 19, 2024.

Groundwork. (2024). Groundwork takes practical action to create a fair and green future in which people, places, and nature thrive. Retrieved from *https://www.groundwork.org.uk/* Accessed July 19, 2024.

Gupta, P., Ding, B., Guan, C., & Ding, D. (2024). *Generative AI: A systematic review using topic modelling techniques*. Data and Information Management, Article 100066.

Han, Z.FeiFei, Lin, J., Gurung, A., Thomas, D.R., Chen, E., Borchers, C., Gupta, S., & Koedinger, K.R. (2024). Improving assessment of tutoring practices using retrieval-augmented generation. arXiv preprint arXiv:2402.14594.

Hoffmann, J., Borgeaud, S., Mensch, A., Buchatskaya, E., Cai, T., Rutherford, E., & Sifre, L. (2022). Training compute-optimal large language models. arXiv preprint arXiv: 2203.15556.

Hutton, G. (2024). Business statistics by House of Commons Library. Retrieved from *https://researchbriefings.files.parliament.uk/documents/SN06152/SN06152.pdf* Accessed July 09, 2024.

Ibisworld. (2024). IBISWorld for the UK. Retrieved from https://www.ibisworld.com/united-kingdom/ Accessed July 19, 2024.

Innovate UK. (2024). UK Research & Innovation. Retrieved from *https://www.ukri.org/councils/innovate-uk/* Accessed July 29, 2024.

Jasmin, A., Ramesh, P., & Tanveer, M. (2024). Development of Artificial intelligence-based chatbot for smart aquafarm practices. *Expert Systems, 41*(6), Article e13123.

Jegham, N., Abdelatti, M., Elmoubarki, L., & Hendawi, A. (2025). How hungry is AI? benchmarking energy, water, and carbon footprint of llm inference. arXiv preprint arXiv:2505.09598.

Jeong, M., Sohn, J., Sung, M., & Kang, J. (2024). Improving medical reasoning through retrieval and self-reflection with retrieval-augmented large language models. arXiv preprint arXiv:2401.15269.

Jimeno Yepes, A., You, Y., Milczek, J., Laverde, S., & Li, L. (2024). Financial report chunking for effective retrieval augmented generation. arXiv e-prints, arXiv-2402.

Kim, J., Kim, H. U., Adamowski, J., Hatami, S., & Jeong, H. (2022). Comparative study of term-weighting schemes for environmental big data using machine learning. *Environmental Modelling & Software, 157*, Article 105536.

Kirchner-Krath, J., Morschheuser, B., Sicevic, N., Xi, N., von Korflesch, H. F., & Hamari, J. (2024). Challenges in the adoption of sustainability information systems: A study on green IS in organizations. *International Journal of Information Management*, Article 102754.

Klewitz, J., & Hansen, E. G. (2014). Sustainability-oriented innovation of SMEs: A systematic review. *Journal of Cleaner Production, 65*, 57–75.

Kolditz, O., Rink, K., Nixdorf, E., Fischer, T., Bilke, L., Naumov, D., & Yue, T. (2019). Environmental information systems: Paving the path for digitally facilitated water management (Water 4.0). *Engineering, 5*(5), 828–832.

Kurivella, P. K., & Lauren, P. (2023). A comparative survey of open-source platforms for chat-based generative AI. In *Proceedings of the Future Technologies Conference* (pp. 164–185). Cham: Springer Nature Switzerland.

Langchain. (2024). *Applications that can reason*. Powered by LangChain. Retrieved from https://www.langchain.com/ Accessed July 05, 2024.

Le Bronnec, F., Vérine, A., Negrevergne, B., Chevaleyre, Y., & Allauzen, A. (2024). Exploring precision and recall to assess the quality and diversity of LLMs. In *62nd Annual Meeting of the Association for Computational Linguistics*.

Le Scao, T., Fan, A., Akiki, C., Pavlick, E., Ilić, S., Hesslow, D., & Al-Shaibani, M. S (2022). *Bloom: A 176b-parameter open-access multilingual language model*. arXiv. https://doi.org/10.48550/arXiv.2211.05100

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems, 33*, 9459–9474.

Li, Q., Jiang, A., Zhang, Z., Wang, H., & Wang, L. (2025). The application of TF-IDF in the research of flood and drought disaster defense system. *CE/Papers, 8*(2), 1910–1919.

Liu, Q., Yang, L., & Yang, M. (2021). Digitalisation for water sustainability: Barriers to implementing circular economy in smart water management. *Sustainability, 13*(21), Article 11868.

Liu, Y., Cao, J., Liu, C., Ding, K., & Jin, L. (2024). Datasets for large language models: A comprehensive survey. arXiv preprint arXiv:2402.18041.

llamaindex. (2024). Turn your enterprise data into production-ready LLM applications. Retrieved from https://www.llamaindex.ai/ Accessed July 05, 2024.

Local Government Assoication. (2024). We are the national voice of local government, working with councils to support, promote and improve. Retrieved from https://www.local.gov.uk/ Accessed July 19, 2024.

Maliakel, P.J., Ilager, S., & Brandic, I. (2025). Investigating energy efficiency and performance trade-offs in llm inference across tasks and dvfs settings. arXiv preprint arXiv:2501.08219.

Maman, A., Dias, J. G., & Bassi, F. (2024). Sustainability-oriented management in the SMEs. A multilevel analysis in the European Union. *Journal of Environmental Management, 365*, Article 121559.

Marwah, D., & Beel, J. (2020). Term-recency for TF-IDF, BM25 and USE term weighting. In *Proceedings of the 8th International Workshop on Mining Scientific Publications* (pp. 36–41).

MetOffice. (2024). UK climate projections (UKCP). Retrieved from https://www.metoffice.gov.uk/research/approach/collaboration/ukcp Accessed July 19, 2024.

Mitchell, V. G., Mein, R. G., & McMahon, T. A. (2001). Modelling the urban water cycle. *Environmental Modelling & Software, 16*(7), 615–629.

Monteiro, C. S., Costa, C., Pina, A., Santos, M. Y., & Ferrão, P. (2018). An urban building database (UBD) supporting a smart city information system. *Energy and Buildings, 158*, 244–260.

Muñoz-Pascual, L., Curado, C., & Galende, J. (2021). How does the use of information technologies affect the adoption of environmental practices in SMEs? A mixed-methods approach. *Review of Managerial Science, 15*(1), 75–102.

Naeem, K., Aloui, S., Zghibi, A., Mazzoni, A., Triki, C., & Elomri, A. (2024). A system dynamics approach to management of water resources in Qatar. *Sustainable Production and Consumption, 46*, 733–753.

National Infrastructure Commission. (2024). Impartial, expert advice on major long term infrastructure challenges. Retrieved from *https://nic.org.uk/* Accessed July 09, 2024.

Nguyen, K. A., Stewart, R. A., Zhang, H., Sahin, O., & Siriwardene, N. (2018). *Re-engineering traditional urban water management practices with smart metering and informatics. Environmental Modelling & Software, 101*, 256–267.

Nirala, K. K., Singh, N. K., & Purani, V. S. (2022). A survey on providing customer and public administration based services using AI: Chatbot. *Multimedia Tools and Applications, 81*(16), 22215–22246.

Ofwat. (2024). We regulate the water sector in England and Wales. Retrieved from *https://www.ofwat.gov.uk/* Accessed July 29, 2024.

Rackauckas, Z. (2024). RAG-Fusion: A new take on retrieval-augmented generation. arXiv preprint arXiv:2402.03367.

Rajaram, K., & Tinguely, P. N. (2024). Generative artificial intelligence in small and medium enterprises: Navigating its promises and challenges. *Business Horizons*.

Raman, H., Mohan, S., & Sethuraj, G. (2024). A knowledge based decision support system for water distribution management. *WIT transactions on information and communication technologies* (p. 6).

Raya-Tapia, A. Y., Cansino-Loeza, B., Sánchez-Zarco, X. G., Ramírez-Márquez, C., Martín, M., & Ponce-Ortega, J. M. (2023). A spatial and temporal assessment of resource security in the water, energy, food and waste nexus in Spain. *Sustainable Production and Consumption, 39*, 109–122.

Risch, E., Gutierrez, O., Roux, P., Boutin, C., & Corominas, L. (2015). Life cycle assessment of urban wastewater systems: Quantifying the relative contribution of sewer systems. *Water Research, 77*, 35–48.

Roelleke, T., & Wang, J. (2008). TF-IDF uncovered: A study of theories and probabilities. In *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 435–442).

Roffia, P., & Dabić, M. (2024). The role of management control and integrated information systems for the resilience of SMEs. *Review of Managerial Science, 18*(5), 1353–1375.

Salam, A. (2024). Internet of things in water management and treatment. *Internet of things for sustainable community development: Wireless communications, sensing, and systems* (pp. 273–298). Cham: Springer International Publishing.

Samuel, D. J., Sermet, M. Y., Mount, J., Vald, G., Cwiertny, D., & Demir, I. (2024). Application of large language models in developing conversational agents for water quality education, communication and operations. *Eartharxiv*. https://doi.org/10.31223/X5XT4K

Schiessel Harvey, N. (2021). The multiple values of urban waterways. *The palgrave handbook of climate resilient societies* (pp. 1–27). Cham: Springer International Publishing.

Science & Environment. (2024). Science & environment. Retrieved from *https://www.bbc.co.uk/news/science_and_environment* Accessed July 09, 2024.

Sha, Y., Feng, Y., He, M., Liu, S., & Ji, Y. (2023). Retrieval-augmented knowledge graph reasoning for commonsense question answering. *Mathematics, 11*(15), 3269. https://doi.org/10.3390/math11153269

Sharp, L. (2017). *Reconnecting people and water: Public engagement and sustainable urban water management*. Routledge.

Sharvelle, S., Dozier, A., Arabi, M., & Reichel, B. (2017). A geospatially-enabled web tool for urban water demand forecasting and assessment of alternative urban water management strategies. *Environmental Modelling & Software, 97*, 213–228.

Smartwatermagazine. (2024). Smart Water Magazine. Retrieved from *https://smartwatermagazine.com/* Accessed July 29, 2024.

Stewart, R. A., Willis, R., Giurco, D., Panuwatwanich, K., & Capati, G. (2010). Web-based knowledge management system: Linking smart metering to the future of urban water planning. *Australian Planner, 47*(2), 66–74.

Tan, S., & Yao, L. (2023). Managing and optimizing urban water supply system for sustainable development: Perspectives from water-energy-carbon nexus. *Sustainable Production and Consumption, 37*, 39–52.

Team, G., Kamath, A., Ferret, J., Pathak, S., Vieillard, N., Merhej, R., & Iqbal, S. (2025). Gemma 3 technical report. arXiv preprint arXiv:2503.19786.

Thakur, N., Reimers, N., Daxenberger, J., & Gurevych, I. (2020). Augmented SBERT: Data augmentation method for improving bi-encoders for pairwise sentence scoring tasks. arXiv preprint arXiv:2010.08240.

Thameswater. (2024). Water transfer from the River Severn to the River Thames. Retrieved from *https://www.thameswater.co.uk/about-us/regulation/strategic-water-resource-solutions/water-transfer-from-the-river-severn-to-the-river-thames* Accessed July 19, 2024.

The Economist. (2024). Our latest coverage of climate change. Retrieved from https://www.economist.com/climate-change Accessed July 19, 2024.

The Royal Society. (2024). Welcome to the royal society. Retrieved from *https://royalsociety.org/* Accessed July 09, 2024.

Thewaterreport. (2024). Expert analysis for a changing water industry. Retrieved from *https://www.thewaterreport.co.uk/* Accessed July 09, 2024.

Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., & Scialom, T. (2023). Llama 2: Open foundation and fine-tuned chat models. arXiv preprint arXiv:2307.09288.

UK Centre for Ecology & Hydrology (UKCEH). (2024). Environmental science for a world where people and nature prosper. Retrieved from https://www.ceh.ac.uk/ Accessed July 19, 2024.

UKCIP. (2024). UK Climate Impacts Programme. Retrieved from https://www.ukcip.org.uk/ Accessed July 19, 2024.

UKRI. (2024). UK Research & Innovation. Retrieved from https://www.ukri.org/ Accessed July 09, 2024.

Ukwir. (2024). The Collaborative Research platform for the UK and Ireland Water sector. Retrieved from https://ukwir.org/leading-the-water-industry-research-agenda Accessed July 19, 2024.

Van Ginkel, K. C., Hoekstra, A. Y., Buurman, J., & Hogeboom, R. J. (2018). Urban water security dashboard: Systems approach to characterizing the water security of cities. *Journal of Water Resources Planning and Management, 144*(12), Article 04018075.

Vidal-Lamolla, P., Molinos-Senante, M., Oliva-Felipe, L., Alvarez-Napagao, S., Cortés, U., Martínez-Gomariz, E., & Poch, M. (2024). Assessing urban water demand-side management policies before their implementation: An agent-based model approach. *Sustainable Cities and Society, 107*, Article 105435.

Wang, Z., Zhan, A., Tao, Y., Jian, Y., & Yao, Y. (2024). Sustainable governance of drinking water conservation areas based on adaptive thresholds. *Journal of Environmental Management, 351*, Article 119605.

Water UK. (2024a). Our world-leading plan to remove 4 million spills from rivers and seas. Retrieved from https://www.water.org.uk/Accessed. July 29, 2024.

Water UK. (2025). >3 million households to receive reduced bills as water companies announce record-breaking investment. Retrieved from https://www.water.org.

uk/news-views-publications/news/more-3-million-households-receive-reduced-bills-water-companies Accessed 29 July 2025.
Water-UK. (2023). Net zero 2030 routemap. Retrieved from https://www.water.org.uk/sites/default/files/2023-08/Water-UK-Net-Zero-2030-Routemap-Summary-updated.pdf Accessed July 09, 2024.
Waterwise. (2024). Waterwise. Retrieved from https://www.waterwise.org.uk/ Accessed July 29, 2024.
Weidinger, L., Mellor, J., Rauh, M., Griffin, C., Uesato, J., Huang, P.S., & Gabriel, I. (2021). Ethical and social risks of harm from language models. arXiv preprint arXiv: 2112.04359.
Werbeloff, L., & Brown, R. R. (2016). Using policy and regulatory frameworks to facilitate water transitions. *Water Resources Management, 30*, 3653–3669.
Willuweit, L., & O'Sullivan, J. J. (2013). A decision support tool for sustainable planning of urban water systems: Presenting the dynamic Urban Water Simulation Model. *Water Research, 47*(20), 7206–7220.
Xiong, G., Jin, Q., Lu, Z., & Zhang, A. (2024). Benchmarking retrieval-augmented generation for medicine. arXiv preprint arXiv:2402.13178.
Yang, J., Jin, H., Tang, R., Han, X., Feng, Q., Jiang, H., & Hu, X. (2024). Harnessing the power of llms in practice: A survey on chatgpt and beyond. *ACM Transactions on Knowledge Discovery from Data, 18*(6), 1–32.
Yougov. (2024). Yougov-UK. Retrieved from *https://yougov.co.uk/* Accessed July 29, 2024.
Zakka, C., Shad, R., Chaurasia, A., Dalal, A. R., Kim, J. L., Moor, M., & Hiesinger, W. (2024). Almanac—Retrieval-augmented language models for clinical medicine. *NEJM AI, 1*(2), Article AIoa2300068.