Driving sustainable energy transitions with a multi-source RAG-LLM system

Muhammad Arslan [a,*], Lamine Mahdjoubi [a], Saba Munawar [b]

[a] Architecture and Environment, University of the West of England, Bristol, United Kingdom
[b] Electrical (Telecommunication) Engineering, National University of Computer and Emerging Sciences (NUCES), Islamabad, Pakistan

ARTICLE INFO

ABSTRACT

By 2035, the UK aims to upgrade all homes to achieve a net-zero economy by 2050, thereby reducing energy consumption, household costs, and improving living conditions. Small and Medium-sized Enterprises (SMEs) play a crucial role in this transition. However, many SME contractors lack essential information on Sustainable Energy Initiatives (SEIs) and the relevant Energy landscape necessary for driving Sustainable Energy Transitions (SETs). This knowledge gap poses risks to SME interventions, potentially leading to increased costs and inefficiencies. Accessing timely information on SEIs including government policies, funding, technologies, and environmental impacts from various media sources is essential for guiding effective SETs and understanding the relevant Energy landscape, thereby facilitating informed decision-making. Currently, SMEs lack an integrated system that consolidates data from diverse media sources into a centralized Information System (IS), limiting their ability to effectively navigate SEIs. To address this gap, this research introduces an Energy Chatbot, a sustainable IS that utilizes Large Language Models (LLMs) integrated with multi-source Retrieval Augmented Generation (RAG). This system encompasses diverse media sources, including news articles, government reports, industry publications, academic research, and social media. The Energy Chatbot is designed to enhance decision-making for SMEs by providing comprehensive Energy sector insights through a Question Answering (QA) system. Key findings emphasize that this approach reduces costs by utilizing open-source models. Moreover, the Energy Chatbot provides SMEs with access to up-to-date information, enabling them to identify long-term sustainability strategies and maintain a competitive edge in the evolving Energy landscape.

1. Introduction

The UK has the least energy-efficient housing stock in Europe [1]. To meet its climate targets, the UK aims to upgrade all homes to Energy Performance Certificate (EPC) band C standard by 2035 [1]. Upgrading homes helps lower energy consumption, decrease household energy costs, and improve living conditions, all of which are crucial for achieving a net-zero economy by 2050. But only 29 % of homes today meet this standard, and the UK's current policy approach is nowhere near ambitious enough to the remaining 71 % [1]. More than 99.7 % of the UK's 340,000 construction companies are SMEs [2]. Achieving the government's net zero target by 2050 centres significantly on how effectively SMEs leverage current SEIs, which encompass policies, programs, and initiatives promoting sustainable energy practices. SMEs, often the primary point of contact for homeowners and with extensive experience in small to medium-scale projects, play a crucial role in this transition [3]. However, there is a notable shortage of SME contractors equipped with the necessary information about current SEIs essential for

driving SETs [4]. SETs involve transitioning from fossil fuels to renewable sources like solar and wind to mitigate climate change and achieve long-term environmental and economic sustainability by reducing carbon emissions [5]. The lack of comprehensive and up-to-date knowledge about SEIs and the relevant Energy landscape can lead to significant negative effects in SET efforts by SMEs, including increased costs and wasted time [6].

Accessing the latest information about existing SEIs is vital for driving SETs [7,8]. This information spans various domains, including the latest government policies, funding, and support schemes. Understanding current market trends, available technologies, environmental impacts, economic analyses, and social acceptance is crucial [9]. Additionally, knowledge of global and regional initiatives, success stories, and the risks involved in adopting new methods is essential [10]. These components collectively drive SETs by providing detailed and integrated knowledge about the evolving energy landscape. Diverse media sources such as news websites, government reports, industry publications, academic research, and social media all contribute to understanding SEIs

---

[11]. News websites provide the latest updates on SEIs, keeping stakeholders informed about recent developments and trends [12]. Government reports offer insights into policies, regulations, and funding opportunities essential for planning and implementation [13]. Industry publications and academic research deliver in-depth analyses, case studies, and technological advancements, offering practical and theoretical knowledge [14]. Social media platforms facilitate the exchange of ideas and experiences, enabling real-time discussions and the dissemination of best practices [15].

Currently, SMEs lack access to an integrated IS that consolidates data from diverse media sources into a centralized information hub [16–18]. This gap is critical because a centralized IS is essential for helping SMEs drive SETs by providing all available and updated information about SEIs and the related Energy landscape. The absence of such a system often results in inefficiencies and increased costs, as existing systems fail to aggregate information effectively. To address this issue, the research poses the following key questions:

- How can an integrated IS consolidate data from multiple sources to benefit SMEs in the Energy sector using state-of-the-art open-source technologies?
- How can this system enhance SMEs' decision-making processes concerning SETs?
- How can the system be designed to ensure sustainability, allowing it to remain both upgradable and customizable to adapt to the evolving needs and interests of organizations over time?

To address these research questions, this study proposes developing an Energy Chatbot, an integrated IS designed to effectively consolidate data from multiple sources. Utilizing advanced open-source technologies such as LLMs and multi-source RAG, this Chatbot is tailored specifically tailored to the domain of "Energy" and offers several key benefits: The Energy Chatbot integrates a wide range of media sources, including news outlets, government reports, industry publications, and social media, into a cohesive system. This integration provides SMEs with comprehensive, up-to-date information on SEIs and energy trends, significantly enhancing their decision-making capabilities. By delivering timely and precise information, the Chatbot supports SMEs in making data-driven decisions regarding SETs. It enhances operational efficiency by streamlining information gathering and analysis, ensuring that SMEs can quickly adopt the most effective and innovative practices. The system is designed to be a sustainable solution by allowing for easy upgrades and customization. It utilizes datasets relevant to the needs of organizations, ensuring the Chatbot remains adaptable and pertinent as requirements evolve. Furthermore, its design minimizes development costs, making it a cost-effective option for continuous learning and adaptation. In short, the Energy Chatbot provides SMEs with a centralized, efficient tool for accessing industry trends and technological advancements while supporting long-term sustainability goals and compliance with regulations.

The paper is structured as follows: Section 2 provides the background of the study. Section 3 introduces our novel IS i.e., Energy Chatbot, based on multi-source RAG with LLM. Section 4 covers the system's evaluation. Section 5 discusses the system's advantages and limitations. Finally, Section 6 offers concluding remarks.

2. Background

The relevant literature for this research is divided into two sections. The first part reviews existing studies that utilize various media sources to capture and understand the Energy landscape, presenting the current systems employed for Energy-related IS. The second part focuses on the introduction of LLMs and RAG, exploring how these technologies have been successfully applied to enhance IS and generative tasks across different domains. By examining these studies, we aim to identify gaps and opportunities for integrating RAG with LLMs to create a robust and sustainable IS for Energy, ultimately contributing to more informed decision-making processes in the energy sector.

2.1. Existing studies driving SETs

SETs entail moving from fossil fuels to renewables to mitigate carbon emissions and improve sustainability [19]. Extensive literature examines how diverse media sources contribute to acquiring essential Energy-related knowledge and developing IS that empower organizations to effectively advance SETs. For instance, Energy Storage (ES) technology is crucial for SETs but faces socio-political challenges [20]. Chen & Rowlands [20] analyzed 156 ES-related articles from People's Daily and China Daily using the Socio-Political Evaluation of Energy Deployment (SPEED) framework, revealing a promising pathway for ES development in China in uncertainties and local institutional inertia, highlighting the supportive role of the Chinese government. Piselli et al. [21] investigated the role of online information sources in promoting energy communities [22], using the Semantic Brand Score (SBS) indicator to analyze online news data, revealing distinct trends and information gaps.

Ganowski & Rowlands [23] examined national news coverage of ES in Canada and the UK from 2008 to 2017, highlighting how media discourse influences public acceptance of ES technologies. Dehler-Holland et al. [24] analyzed media framing of the German Renewable Energy Act from 2000 to 2017, revealing a shift from optimism about renewable energies to concerns about costs. Marzouki et al. [25] analyzed public discourse on Sustainable Development Goal (SDG) 11 (i.e., Sustainable Cities and Communities) using Twitter, finding a decline in engagement during the COVID-19 pandemic and growing importance of Artificial intelligence (AI) and Internet of things (IoT) technologies. Rommetveit et al. [26] explored the evolution of energy practices in Norway, highlighting the shift towards greater automation and the integration of digital technologies and renewables in electricity grids. Krzywda et al. [27] examined Poland's energy transition during its heavy reliance on coal, analyzing media coverage to illustrate divergent perspectives and strategies employed by different social actors. Ibar-Alonso et al. [28] used social listening on Twitter to analyze public sentiment during the Ukrainian–Russian conflict, finding that it impacted societal attitudes towards green energy.

Moreover, Lyytimäki [29] examined long-term biogas coverage in Finnish newspapers, identifying four key storylines in public debate. Walker et al. [30] explored the portrayal of renewable energy in Indigenous communities within Canadian media, finding that coverage is dominated by indigenous opposition to large-scale hydro projects. Labonte & Rowlands [31] examined how Twitter data can shed light on the socio-political aspects of sustainability transitions, finding that a small group of highly engaged users drives the conversation. Furthermore, Lyytimäki et al. [32] investigated the role of media in energy transitions by analyzing newspaper coverage of biogas in Finland, revealing that biogas coverage was minimal in the early 2000 s but saw a rapid increase and stabilization over time. Antal & Karhunmaa [33] analyzed media coverage from the UK, Finland, and Hungary between 2011 and 2015, finding that discussions were primarily techno-economic and supply-oriented, reflecting local visions of an ideal society.

In addition to the ISs built using diverse Energy-related datasets, several studies have focused on developing conversational agents, or Chatbots, which function as ISs for energy management. Gnewuch et al. [34] created design principles for energy feedback agents, which were tested through a focus group session. Suresan et al. [35] developed a Chatbot for real-time interaction to help consumers manage peak energy usage. Rocha et al. [36] introduced a self-reading Chatbot integrated with image processing to record energy consumption, particularly useful during the COVID-19 pandemic. Fontecha et al. [37] presented Green-MoCA, a system that monitors energy use in smart homes with a conversational assistant providing real-time feedback. Milano et al. [38]

developed EcoBot, which uses persuasion strategies to encourage energy savings, showing higher user satisfaction despite mixed results in effectiveness. Lastly, Onile et al. [39] demonstrated the effectiveness of a hybrid digital twin asset modeling approach in enhancing user engagement and recommendation quality for energy conservation.

As evidenced by the studies mentioned, online media sources including news articles, social media, and other platforms are crucial for enabling organizations to lead SETs. These sources offer a comprehensive view of the complex Energy landscape, whether through ISs or chatbots. SMEs often lack the Research and Development (R&D) resources to identify and test existing open-source systems, making it difficult to find solutions that fit their specific needs [40,41]. The evolving energy landscape requires continuous updates to systems and datasets, necessitating both time and financial resources [42]. For most resource-limited SMEs, this approach becomes impractical. Additionally, many of the studies mentioned above often rely on a single type of dataset, focusing only on one aspect of the Energy landscape. This approach limits the breadth and depth of information available, hindering comprehensive decision-making in energy transitions. This narrow focus can introduce bias and overlook valuable insights that could emerge from integrating multiple sources of information, such as news websites, government reports, industry publications, academic research, and social media. Adding more information sources increases the complexity of information linking and semantic enrichment, which involves connecting and enhancing data with meaningful context [43,44]. Semantic enrichment plays a crucial role in connecting diverse data streams to achieve a comprehensive understanding of Energy-related information [45].

For instance, consider the example of policy updates and market trends [46]. Relying solely on government reports may overlook real-time market reactions captured in social media or the latest technological advancements discussed in industry publications. By applying semantic enrichment techniques [47], such as linking policy updates with real-time market data and detailed industry analysis, organizations can bridge these gaps. This integrated approach enables users to extract relevant insights and present a holistic landscape of information essential for informed decision-making in energy transitions. To achieve effective integration of information, organizations must adopt a holistic approach leveraging diverse Energy-related information sources. Applying state-of-the-art semantic enrichment methods [47–49] is crucial as they facilitate connecting information from multiple sources based on similar contexts, thereby developing a comprehensive understanding of the energy landscape. Instead of relying on complex semantic enrichment methods [44], which come with various benefits and computational costs, using LLMs based on NLP offers a sustainable alternative [50].

2.2. LLMS and RAG

LLMs are advanced AI models based on NLP [51]. They exhibit impressive language comprehension and generation capabilities, but their performance varies depending on task complexity, training data quality, and the subtleties of human communication [51,52]. This enables tasks such as text generation, classification, sentiment analysis, information retrieval, and content summarization [52]. However, these transformative capabilities come at a significant computational cost, particularly during training and inference phases [53]. Substantial computing power, typically in the form of General Processing Unit (GPU) chips, is required. Training powerful LLMs necessitates thousands of GPUs, each with an environmental footprint from manufacturing to disposal. To align with sustainability goals, organizations must address the environmental impact of LLMs in their technological advancements. A key strategy is maximizing the reusability of LLMs to reduce their environmental footprint [54]. Instead of each organization training its own LLM, they can opt for open-source LLMs customized to their specific needs. This approach promotes responsible investment practices,
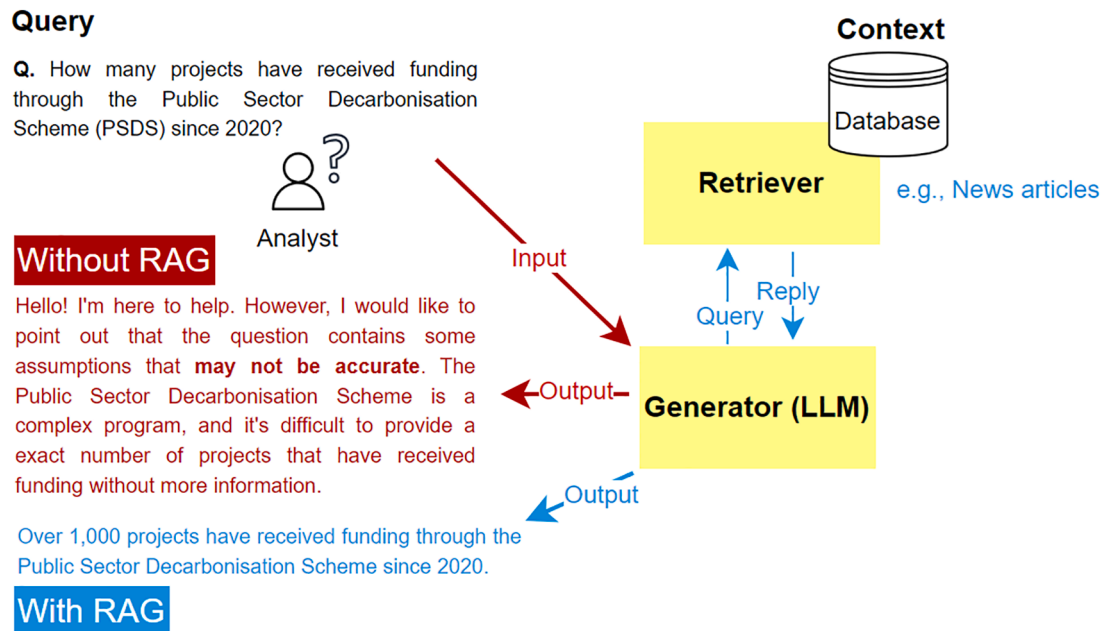
integrating environmental, social, and governance (ESG) criteria into investment decisions to foster sustainable financial returns and positive societal impact. The system tasked with investigating and analyzing information from diverse sources should prioritize sustainability through responsible investment, aiming to minimize its carbon footprint. To achieve this, we utilize advanced LLMs that seamlessly integrate multiple information sources using RAG, facilitating a comprehensive understanding without relying on complex semantic enrichment methods or information integration complexities.

RAG enhances LLM performance in generative tasks [55] by introducing an initial step where LLMs retrieve relevant information from external sources before generating text (see Fig. 1). This integration of external data retrieval improves the accuracy and relevance of the generated output, reducing errors and enhancing overall quality. This approach increases the practical applicability of LLMs in real-world scenarios, ensuring that the generated content is evidence-based and dependable [55]. Recent studies have explored various applications of RAG with LLMs across different domains. For instance, Alawwad et al. [56] discuss using RAG to handle out-of-domain scenarios in textbook Question Answering (QA) covering life science domains. Bucur et al. [57] apply RAG for automated form filling within enterprise search contexts, focusing on web page files for IT project request forms. Moreover, Zhang et al. [58] employ RAG for financial sentiment analysis using Twitter financial news, specializing in sentiment classification. Moreover, Al Ghadban et al. [59] highlight RAG's role in frontline health worker capacity building through pregnancy-related guidelines, enhancing health education QA. Jeong et al. [60] introduce Self-BioRAG, a framework for biomedical text instruction sets aimed at generating biomedical explanations and retrieving domain-specific documents. Furthermore, Xia et al. [61] propose a hybrid RAG approach to enhance user writing speed and accuracy. Rackauckas [62] uses RAG-Fusion to obtain technical product information from product datasheets. Shi et al. [63] develop RACE for retrieval-augmented commit message generation for code intelligence across multiple programming languages. Colverd et al. [64] use FloodBrain for flood disaster reporting via RAG to assist in humanitarian assistance.

The studies mentioned above illustrate RAG's versatility with LLMs in constructing ISs, enabling generative tasks such as QA across various domains, including biomedical, financial, technical, and humanitarian fields. These ISs, functioning as Chatbots, facilitate user interaction with the system, enabling seamless engagement [65]. Despite these advancements, the potential application of RAG combined with LLMs within the Energy sector remains largely unexplored. By integrating RAG with LLMs, SMEs can effectively access and synthesize data from a wide array of sources including news websites, government reports, industry publications, academic research, and social media. This integrated approach facilitates the development of sustainable QA systems capable of extracting meaningful patterns, trends, and insights essential for informed decision-making in SETs. Importantly, it streamlines the process by bypassing the complexities typically associated with integrating disparate sources of information. In the following section, we outline our integrated approach using a multi-source RAG with LLMs across various Energy-related information sources. This approach aims to develop a QA Chatbot that facilitates analysts in understanding the current Energy landscape.

3. Proposed system

This section begins with an overview of constructing the dataset designed to enhance the RAG with an LLM-based IS, specifically tailored for the Energy domain, functioning as an Energy Chatbot for QA. The choice of a QA Chatbot is supported by existing literature that highlights its benefits in enhancing user interaction and engagement with IS. Chatbots streamline the process of querying and retrieving information, providing quick responses and improving accessibility for users [66,67]. They are particularly effective in domains like Energy, where complex

Query

Q. How many projects have received funding through the Public Sector Decarbonisation Scheme (PSDS) since 2020?

Analyst

Without RAG

Hello! I'm here to help. However, I would like to point out that the question contains some assumptions that **may not be accurate**. The Public Sector Decarbonisation Scheme is a complex program, and it's difficult to provide a exact number of projects that have received funding without more information.

Over 1,000 projects have received funding through the Public Sector Decarbonisation Scheme since 2020.

With RAG

Context

Database

Retriever

e.g., News articles

Input

Query — Reply

Output

Generator (LLM)

Output

Fig. 1. Advantages of RAG technology in enhancing LLMs for domain-specific query response using external datasets: An example with latest news articles.

information needs to be synthesized and communicated in a user-friendly manner. Following the dataset construction, we provide implementation details on how the system was executed using existing tools. Finally, we evaluate the system's accuracy in generating responses to user queries.

3.1. Dataset

To develop a dataset essential for building an IS using RAG with LLMs to facilitate decision-making in SETs, we conducted a comprehensive literature review as outlined in the previous section. This review highlighted critical topics related to SEIs and the broader Energy landscape, crucial for advancing SETs. These topics include Policy Updates, Market Trends, Technological Innovations, Environmental Impacts, Economic Analysis, Social Acceptance and Community Engagement, Global and Regional Initiatives, Case Studies and Success Stories, Risk Assessment and Resilience, as well as Public Opinion and Stakeholder Perspectives. To ensure updated and relevant information for these topics, six key information sources such as news websites, government reports, industry publications, academic research, and social media data can be utilized (see Table 1). Each Energy-related topic mentioned above plays a significant role in delivering relevant information extracted from diverse sources to experts involved in SETs. Below, we detail the importance of each Energy-related topic in providing essential insights that support informed decision-making in this crucial domain (see Fig. 2).

1. *Policy Updates*: It focus on government policies and regulations related to renewable energy adoption, carbon pricing, energy efficiency standards, and incentives for sustainable energy projects. This category relies on government reports and news websites to provide insights into evolving energy policies. Examples include the UK Government Energy Policies [68], which detail official documents and updates from the UK Department for Business, Energy & Industrial Strategy (BEIS), and the reports from the Office of Gas and Electricity Markets' (Ofgem) on electricity and gas market regulations in Great Britain [69].

2. *Market Trends*: It provide insights into current and future developments in renewable energy investment and technologies. This category draws upon industry publications, academic

research, and news websites to report on the deployment of renewable technologies such as solar, wind, hydro, and geothermal. Examples include the Renewable Energy Guarantees of Origin (REGO) Scheme, tracked through government reports [69], and BEIS Energy Trends, which offers statistical data on energy production, consumption, and market growth rates in the UK [68].

3. *Technological Innovations*: It cover advancements in renewable energy technologies, including improvements in solar panel efficiency, energy storage solutions, grid integration technologies, and innovations in wind turbine design. This category utilizes industry publications, academic research, and news websites. Examples include data from UK Research and Innovation (UKRI) on various renewable energy projects and technological advancements [70] and Innovate UK's case studies and information on innovative projects within the UK's renewable energy sector [71].

4. *Environmental Impacts*: It analyze the environmental effects of different energy sources, including carbon emissions, air and water pollution, land use changes, and biodiversity impacts. This category relies on academic research, government reports, and industry publications. Examples include data from the UK Environmental Agency on air quality and emissions sourced from government reports [72], and insights into carbon emissions and environmental impacts provided by the Carbon Trust through industry publications [73].

5. *Economic Analysis*: It focuses on the economic aspects of renewable energy compared to fossil fuels, discussing trends in energy prices, cost reductions in renewable technologies, and the economic benefits of transitioning to sustainable energy sources. This category draws upon industry publications, academic research, and government reports. Examples include economic data from the Office for National Statistics (ONS) related to the energy sector [74] and the comprehensive economic data on the UK energy market provided by the BEIS Digest of UK Energy Statistics (DUKES) [75].

6. *Social Acceptance and Community Engagement*: It explore public perception and community interactions related to renewable energy projects, covering public sentiment, community engagement initiatives, and the social acceptance of renewable energy.

Table 1
Categorization of topics and potential sources for SETs.

No.	Topic	Description	Example of datasets	Potential information sources				
				News websites	Government reports	Industry publications	Academic research	Social media
1	**Policy Updates**	Provides information on government policies and regulations related to: Renewable energy adoption Carbon pricing Energy efficiency standards Incentives for sustainable energy projects	• **UK Government Energy Policies**: Official documents and updates from the UK Department for Business, Energy & Industrial Strategy (BEIS) on energy policies, regulations, and initiatives [68]. **Ofgem (Office of Gas and Electricity Markets)**: Reports and policy updates from the regulator for the electricity and gas markets in Great Britain [69].	✓	✓	×	×	×
2	**Market Trends**	Includes reports on: Trends in renewable energy investment Deployment of renewable technologies (e.g., solar, wind, hydro, geothermal) Market growth rates Forecasts for future developments	• **Renewable Energy Guarantees of Origin (REGO) Scheme**: Data from Ofgem on renewable energy certificates and market trends [69]. **Business, Energy & Industrial Strategy (BEIS) Energy Trends**: Regular publications by BEIS providing statistical data on energy production, consumption, and market developments in the UK [68].	×	✓	✓	✓	✓
3	**Technological Innovations**	Covers advancements in renewable energy technologies, such as: Improvements in solar panel efficiency Energy storage solutions Grid integration technologies Innovations in wind turbine design	• **UK Research and Innovation (UKRI)**: Data and reports on research and development projects in renewable energy technologies funded by UKRI [70]. **Innovate UK**: Case studies and data on innovative projects in the UK's renewable energy sector [71].	✓	×	✓	✓	✓
4	**Environmental Impacts**	Analyzes the environmental impacts of different energy sources, including: Carbon emissions Air and water pollution Land use changes Biodiversity impacts	• **UK Environmental Agency**: Data on air quality, emissions, and other environmental impacts of energy production and consumption [72]. **Carbon Trust**: Reports and data on carbon emissions and environmental impacts of various energy sources [73].	×	✓	✓	✓	×
5	**Economic Analysis**	Discusses the economic feasibility and cost competitiveness of renewable energy compared to fossil fuels, including: Trends in energy prices Cost reductions in renewable technologies Economic benefits of transitioning to sustainable energy sources	• **Office for National Statistics (ONS)**: Economic data related to the energy sector, including price indices, cost analyses, and economic impact assessments [74]. **BEIS Digest of UK Energy Statistics (DUKES)**: Comprehensive data on the economic aspects of the UK energy market [75].	×	✓	✓	✓	×
6	**Social Acceptance and Community Engagement**	Covers public perception, community engagement initiatives, and social acceptance of renewable energy projects, including: Challenges in local communities Successes in local communities	• **YouGov**: Surveys and data on public opinion regarding renewable energy and climate change in the UK [76]. **UK Social Media Platforms (Twitter, Facebook)**: Social media data on public sentiment and community engagement in renewable energy projects.	✓	×	×	✓	✓
7	**Global and Regional Initiatives**	Provides updates on: International agreements (e.g., Paris Agreement) Regional energy policies Collaborations between countries to promote sustainable energy transitions and address climate change	• **Climate Change Committee (CCC)**: Data and reports on the UK's progress towards meeting its climate goals and international agreements like the Paris Agreement [78]. **European Energy Portal**: Data on UK's regional energy policies and collaborations within Europe [79].	✓	✓	✓	×	×

Table 1 (*continued*)

No.	Topic	Description	Example of datasets	Potential information sources				
				News websites	Government reports	Industry publications	Academic research	Social media
8	**Case Studies and Success Stories**	Presents examples of successful renewable energy projects, including: Case studies of communities achieving significant renewable energy targets Lessons learned from implementation	• **Energy Systems Catapult**: Case studies of innovative energy projects and success stories in the UK [80]. **RenewableUK**: Case studies and reports on successful renewable energy projects and initiatives in the UK [81].	✓	×	✓	✓	×
9	**Risk Assessment and Resilience**	Offers information on the risks associated with energy transitions, including: Energy security concerns Technological risks Geopolitical implications Strategies for enhancing resilience in the energy sector	• **National Grid ESO (Electricity System Operator)**: Data and reports on risk assessments and resilience planning for the UK's electricity grid [82]. **UK Climate Projections (UKCP)**: Data and analysis on climate risks and resilience strategies relevant to the energy sector [83].	×	✓	✓	✓	×
10	**Public Opinion and Stakeholder Perspectives**	Includes polls, surveys, and interviews reflecting: Public opinion on energy policies Perspectives from industry leaders Perspectives from environmental organizations Perspectives from community groups advocating for or against specific energy initiatives	• **Ipsos MORI**: Polls and surveys on public opinion regarding energy policies and renewable energy projects in the UK [84]. **Stakeholder Reports (e.g., Greenpeace UK, Friends of the Earth)**: Reports and surveys reflecting the perspectives of various stakeholders in the renewable energy sector [85,86].	✓	×	✓	✓	✓

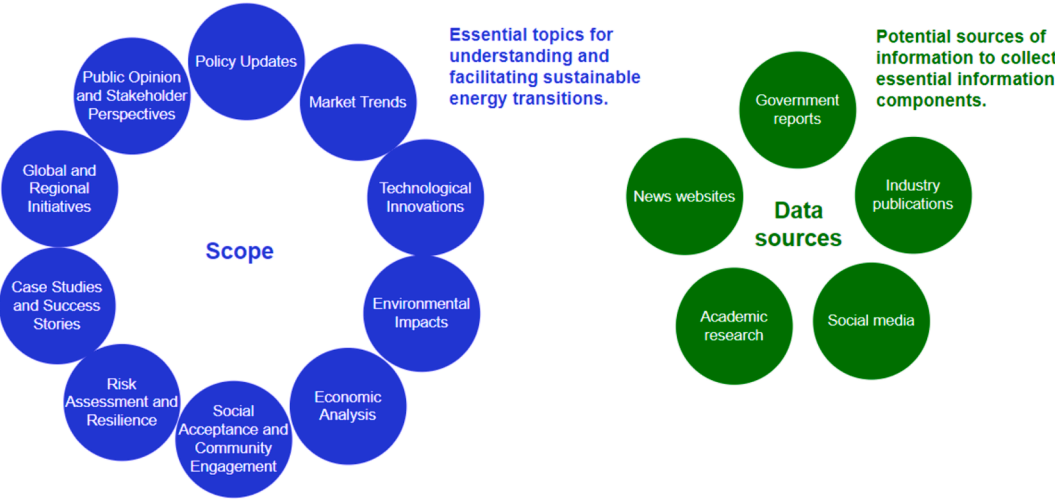✓Contributes significantly
×Does not contribute significantly

Fig. 2. Topics for Energy updates: Customizable scope and selected data sources.

This category utilizes news websites, social media platforms, and academic research. Examples include YouGov's surveys on public opinion regarding renewable energy and climate change in the UK [76], and insights from UK social media platforms like Twitter and Facebook on public sentiment and community engagement in renewable energy projects.

7. *Global and Regional Initiatives*: It provide updates on international agreements such as the Paris Agreement, regional energy policies, and collaborations between countries to promote SETs and address climate change [77]. This category draws upon government reports, news websites, and industry publications. Examples include insights from the Climate Change Committee (CCC) on the UK's progress towards climate goals [78] and the European Energy Portal's data on regional energy policies and collaborations within Europe [79].

8. *Case Studies and Success Stories*: It highlight examples of successful renewable energy projects, case studies of communities achieving significant renewable energy targets, and lessons learned from implementation. This category utilizes academic research, industry publications, and news websites. Examples include case

studies from the Energy Systems Catapult [80] on innovative energy projects and success stories documented by RenewableUK in the UK's renewable energy sector [81].

9. *Risk Assessment and Resilience*: It provide information on the risks associated with energy transitions, including energy security concerns, technological risks, geopolitical implications, and strategies for enhancing resilience in the energy sector. This category relies on government reports, academic research, and industry publications. Examples include risk assessments from the National Grid ESO on the UK's electricity grid [82] and academic research and analysis on climate risks relevant to the energy sector provided by UK Climate Projections (UKCP) [83].

10. *Public Opinion and Stakeholder Perspectives*: It encompass polls, surveys, and interviews reflecting public opinion on energy policies, as well as perspectives from industry leaders, environmental organizations, and community groups advocating for or against specific energy initiatives. This category utilizes social media data, news websites, academic research, and industry publications. Examples include Ipsos MORI's academic research on public opinion regarding energy policies [84] and stakeholder reports from organizations like Greenpeace UK [85] and Friends of the Earth [86]. providing industry perspectives on renewable energy projects and policies.

As detailed earlier, various information sources, including news websites, government reports, industry publications, academic research, and social media data contribute crucial information related to SEIs that drive SETs. While there may be additional sources providing insights into different aspects of the Energy landscape beneficial for SME

decision-making, we have chosen to focus on a single source for each component in our prototype IS to streamline its scope. For instance, we used Ofgem data [69] for Policy Updates, BEIS Energy Trends data [68] for Market Trends, UK Research and Innovation (UKRI) data [70] for Technological Innovations, Carbon Trust data [73] for Environmental Impacts, Office for National Statistics (ONS) data [74] for Economic Analysis, YouGov data [76] for Social Acceptance and Community Engagement, European Energy Portal data [79] for Global and Regional Initiatives, RenewableUK data [81] for Case Studies and Success Stories, National Grid ESO data [82] for Risk Assessment and Resilience, and Ipsos MORI data [84] for Public Opinion and Stakeholder Perspectives. In the next section, we will explain how the data collected from various sources will be used to construct a multi-source RAG system integrated with a LLM-based IS, functioning as an Energy Chatbot for QA.

3.2. System implementation

To empower SMEs with an interactive tool for exploring information on SEIs and the relevant Energy landscape, we developed an Energy Chatbot. Leveraging the capabilities of LLMs, renowned for their proficiency in complex reasoning tasks across diverse domains, was fundamental in creating this IS. We studied various existing LLM releases including BLOOM [87], Falcon [88], Generative Pre-trained Transformer (GPT)-4 [89], Llama [90], and Chinchilla [91]. Ultimately, we selected Llama (specifically Llama2, Llama3, and Llama3.1 developed by Meta), which are updated versions trained on extensive datasets and renowned for their superior performance compared to other open-source models. [90]. While developers have the flexibility to choose any of these models based on their preferences and accessibility, we
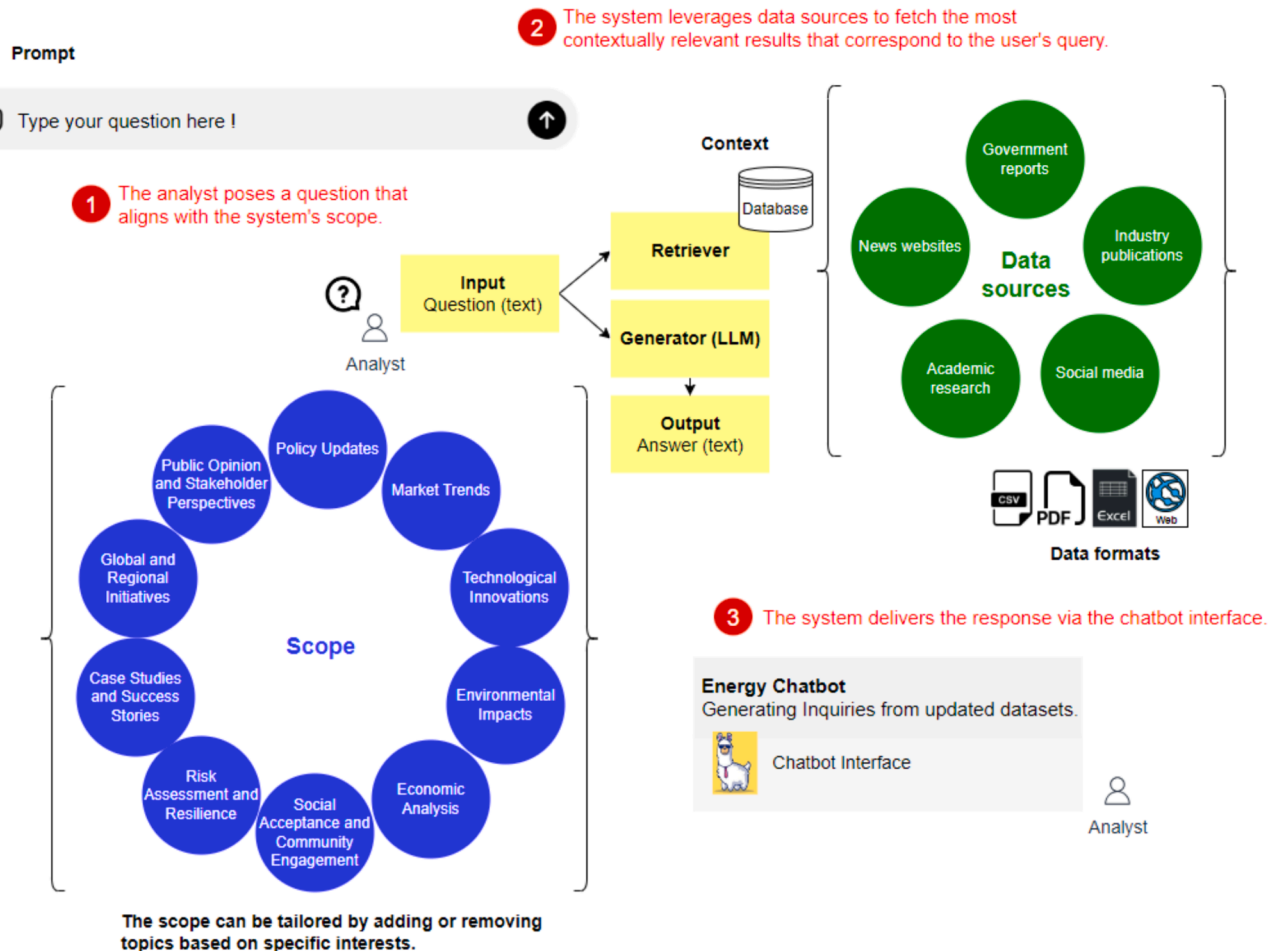

Fig. 3. Integrating datasets from diverse Sources to build an Energy Chatbot with LLM and multi-source RAG.

focused on customizing Llama models for Energy domain. This involved integrating RAG technology into our solution (see Fig. 3), enabling the Chatbot to effectively comprehend and respond to Energy-related queries using information extracted from diverse sources in various formats (such as CSVs, PDFs, spreadsheets, and website data). The effectiveness of the Chatbot in handling a wide range of Energy-related queries centres on the quality and comprehensiveness of the Energy dataset as detailed in the preceding section. To enable our chosen Llama models for Energy-related QA, we have implemented the Algorithm 1.

Algorithm 1: Implementing the RAG with LLM for Energy-related QA

Objective: Developing the Energy Chatbot for Querying Energy-Related Data.
1) Input: Energy-related dataset, user question
2) Process:
Step:1 Install Necessary Libraries
– Langchain, Langchain-Community, Sentence-Transformers, FAISS for GPU, PyPDF.
Step:2 Load and Prepare the Document
2(a) Load the Document:
– Use PyPDFLoader to load a document from the specified file path.
– Call loader.load() to extract the content of a document into a variable named documents.
2(b) Split the Document into Chunks:
– Initialize a CharacterTextSplitter with a chunk size of 1000 characters and an overlap of 30 characters.
– Use the splitter to divide the documents into manageable chunks.
Step 3: Load and Apply Embeddings
3(a) Load the Embedding Model:
– Select the HuggingFace embedding model sentence-transformers/all-mpnet-base-v2 with HuggingFaceEmbeddings.
– Set the model to run on the GPU by specifying model_kwargs = {"device": "cuda"}.
3(b) Generate Embeddings for Document Chunks:
– Apply the embedding model to the document chunks to generate vector representations.
Step 4: Store Embeddings with FAISS
4(a) Create a Vector Store with FAISS:
– Use FAISS to store the generated embeddings by calling FAISS.from_documents (docs, embeddings).
4(b) Persist the Vector Store:
– Save the FAISS index locally with vectorstore.save_local("faiss_index_").
4(c) Load the Vector Store from Disk:
– Reload the stored FAISS index using FAISS.load_local("faiss_index_", embeddings, allow_dangerous_deserialization = True).
4(d) Create a Retriever:
– Use persisted_vectorstore.as_retriever() to create a retriever for efficient information retrieval from the vector store.
Step 5: Set Up the LLM Model
5(a) Install Langchain-Ollama package.
5(b) Set Up Ollama Environment:
– Use colab-xterm to execute terminal commands within the notebook, ensuring ollama is installed and the relevant LLM model is pulled.
– Start the Ollama service and pull the required LLM by running the necessary commands within an xterm session.
5(c) Initialize the LLM:
– Instantiate the Ollama model using Ollama(model="*LLM-Name*").
Step 6: Query and Response Handling
6(a) Orchestrate with RetrievalQA Chain:
– Use the RetrievalQA.from_chain_type() method to create an orchestration chain that combines the Llama model and the retriever.
6(b) Accept User Query:
– Prompt the user for input with input("Type your query: \n").
6(c) Process the Query:
– If the query is not "exit", start a timer, execute the query through the qa.run(query) method, and measure the execution time.
– Print the model's response along with the execution time.
6(d) Exit Condition:
– If the user inputs "exit", display a message ("Exiting the program.") and end the execution.
7) Output: The text produced in response to the posed question.

The objective of the Algorithm 1 is to effectively query Energy-related information using an LLM (Llama) and models from Hugging Face [91]. The first step is to install the necessary libraries such as Langchain, Sentence-Transformers, and FAISS for managing document embeddings and retrieval efficiently [92–95]. The document containing Energy-related data is loaded using PyPDFLoader by Langchain, and then split into manageable chunks for processing. HuggingFace's embedding model is employed to generate vector representations of these chunks, which are stored using Facebook AI Similarity Search (FAISS), a powerful tool for efficient similarity search. This vector store then persisted locally and can be reloaded for future queries. A retriever is set up to facilitate quick information retrieval from the stored vectors. Next, the Ollama (i.e., Ollama.com) environment is configured to run Llama models. The chatbot uses a RetrievalQA chain to combine the Llama model with the retriever, ensuring precise and contextually relevant responses. Users can input their Energy-related queries, and the system will process these, providing answers while measuring execution time to ensure efficiency. The Chatbot is designed to offer accurate, detailed responses, and users can exit the system at any time by typing "exit".

3.3. System evaluation

To evaluate the Chatbot's answers, we have selected Precision, Recall, Accuracy, and F1 score as our key metrics [96–98]. These metrics are commonly used for assessing LLM-based systems [97]. The evaluation focuses on the quality of information present in the Chatbot's answers, which depends on the relevance and completeness of the information retrieved. To systematically assess the information, we define Information Components (ICs), which can include organizations, individuals, monetary amounts, topics, actions, locations, and dates (see Fig. 4). These ICs form the basis for calculating our performance evaluation metrics. Precision measures the effectiveness of the retrieval process by calculating the proportion of relevant ICs among all components retrieved. It underscores the system's capability to exclude irrelevant data effectively (see Equation (1). Recall, or completeness, assesses the proportion of relevant ICs that were successfully retrieved from the total available, reflecting the system's effectiveness in capturing all pertinent information (see Equation (2). Accuracy provides a holistic view of the system's performance by determining the proportion of correctly identified ICs both relevant and irrelevant, out of all ICs evaluated. It captures the overall effectiveness in retrieval and classification (see Equation (3). The F1 score (see Equation (4) combines Precision and Recall into a single metric, offering a balanced measure of performance where both metrics are crucial. By integrating these metrics, the F1 score delivers a comprehensive framework for evaluating LLM-based systems' performance.

$$Precision = \frac{Number of Relevant ICs Retrieved}{Total Number of ICs Retrieved} \tag{1}$$

$$Recall = \frac{Number of Relevant ICs Retrieved}{Total Number of Relevant ICs Available} \tag{2}$$

$$Accuracy = \frac{Number of Correctly Identified ICs}{Total Number of ICs Evaluated} \tag{3}$$

$$F1 score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{4}$$

We employ Cosine Similarity [99] to measure the similarity between ICs, which is instrumental in calculating various evaluation metrics. Cosine similarity measures the cosine of the angle between two vectors in a multi-dimensional space, providing a normalized metric that captures the orientation rather than the magnitude of the vectors [99]. By applying this metric, we can ensure a detailed and comprehensive evaluation of the Chatbot's performance, as it effectively compares the alignment and relevance of ICs across different responses. This approach helps in assessing how well the Chatbot's outputs match expected or reference information, thereby enhancing the accuracy and effectiveness of the performance evaluation process.

To demonstrate the application of ICs in calculating the evaluation metrics, we present five distinct cases based on the question: *"How much funding will public institutions receive to help install low carbon heating and energy-saving measures?"* The Gold Standard Answer is: *"Schools,*
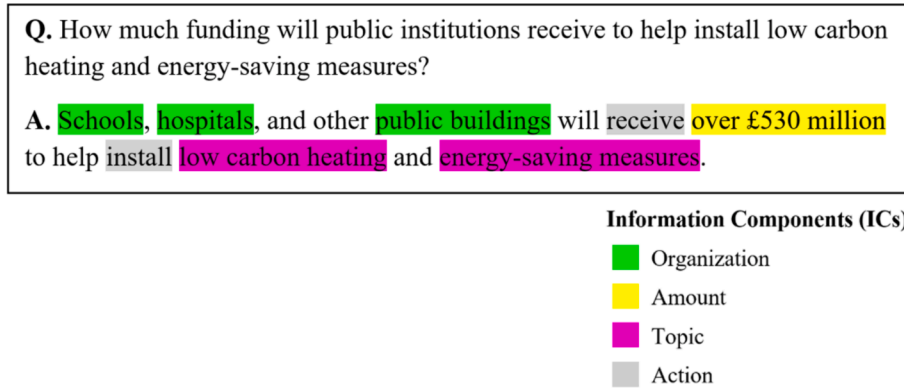
Fig. 4. Identifying ICs helps in determining the quality of the answers.

hospitals, and other public buildings will receive over £530 million to help install low carbon heating and energy-saving measures". Table 2 illustrates how variations in the ICs of different answers impact Precision and Recall metrics. These variations, in turn, contribute to the computation of Accuracy and F1 scores for the system. For example, the 1st case mentions only "schools and hospitals," resulting in high Precision (80 %) and Recall (83 %). In contrast, the 2nd case adds entities like community centers and libraries, which lowers Precision to 65 % but maintains a Recall of 83 %. The 3rd case, which perfectly matches the Gold Standard Answer, achieves perfect scores of 100 % in all metrics. The 4th case, with a more general statement about "schools and public buildings", shows the lowest performance with Precision at 50 %, and Recall at 33 %. The 5th case, which includes additional institutions and a detailed description, achieves medium Precision (65 %), and high Recall (83 %). This comparison highlights the impact of detail and relevance in the information provided on the evaluation metrics.

After detailing the use of ICs for calculating performance metrics, we applied this methodology to our study. For system testing, we constructed questions from the Energy dataset, as outlined in the existing literature, which often distinguishes between two types of test question sets: simple and complex. Simple questions are direct inquiries about information explicitly stated in the text related to the domain [100]. In contrast, complex questions involve compound queries requiring answers that are not directly present in the text but must be inferred through deeper analysis [100]. In our evaluation, we focused on simple, direct questions that involve multiple information components to ensure comprehensive answers. To maintain an unbiased evaluation process, we curated a diverse set of 100 questions from various media sources,

including news articles, government websites, and industry publications. This dataset was carefully constructed to cover a wide spectrum of Energy-related topics, such as Renewable Energy, Policy Regulations, Market Trends, Technological Advancements, and Environmental Impacts. We carefully varied the questions in terms of complexity and context to thoroughly test the system's capabilities.

Table 3 presents a representative subset of these questions, highlighting the dataset's extensive coverage across different Energy-related topics. This selection demonstrates the dataset's breadth and depth, providing a robust foundation for evaluating the system's performance. By including diverse questions from various information sources, we aimed to create a representative sample that accurately reflects the wide array of information in the Energy sector, enabling a rigorous assessment of the system's ability to handle varied and complex queries. To ensure accuracy and consistency, we configured the LLM with minimal creativity settings. This approach was designed to provide precise, focused, and conservative responses, which are crucial for our Energy domain analysis that requires reliable and factual information. While increasing creativity could potentially enhance the results, it might also introduce variability and unconventional responses, which could pose risks for SMEs that depend on consistent and accurate data from government organizations and regulatory bodies.

Our evaluation was conducted in two phases. First, we assessed the Chatbot's performance with the Llama models [90,101] without the RAG feature. The Llama models [90,101] come in various sizes, including 7B, 8B, 13B, 70B, and 405B parameters. For this study, we focused on the smaller variants; Llama2:7B, Llama3:8B, and Llama3.1:8B to provide a clear comparison while managing computational resources efficiently. However, the larger variants, such as Llama2 (13B, 70B), Llama3 (70B), and Llama3.1 (70B, 405B), would require significantly more memory and computational power, which were beyond the scope of this analysis. Table 4 presents a comparison of responses from Llama2:7B, Llama3:8B and Llama3.1:8B models against Gold Standard Answers for specific questions. Llama2:7B often struggles to provide explicit answers and acknowledge limitations in its training data, resulting in incomplete responses. Similarly, Llama3:8B and Llama3.1:8B also face challenges with the most recent data but recommend external sources for up-to-date information. Both models exhibit limitations in handling questions requiring current or specific data, revealing a gap between their responses and the Gold Standard Answers.

Table 5 presents a comparative evaluation of the Chatbot's performance without RAG integration using the Llama2:7B, Llama3:8B, and Llama3.1:8B models. The results show that the Llama3.1:8B model outperforms the other models in all key metrics. Specifically, the Chatbot with Llama3.1:8B achieves the highest Precision (40 %), Recall (37 %), Accuracy (38 %), and F1-score (38.4 %). The Llama3:8B model performs moderately, with a Precision of 35 %, Recall of 25 %, Accuracy of 30 %, and an F1-score of 29.2 %. In contrast, the Llama2:7B model exhibits the lowest performance, with a Precision of 18 %, Recall of 10

Table 2
Variations in information components (ICs) lead to differences in precision and recall values.

Case No.	Answer	Precision (%)	Recall (%)
1	Schools and hospitals will receive over £530 million to help install low carbon heating and energy-saving measures.	80	83
2	Schools, hospitals, community centers, libraries, and other public buildings will receive funding to help install low carbon heating and energy-saving measures.	65	83
3	Schools, hospitals, and other public buildings will receive over £530 million to help install low carbon heating and energy-saving measures.	100	100
4	Schools and public buildings will receive some funding for energy measures.	50	33
5	Schools, hospitals, and various other institutions will receive substantial funding to assist in installing various energy-saving measures, including low carbon heating systems and more, amounting to over £530 million.	65	83

Table 3

Sampe records of user queries and system generated answers for rag with llm evaluation.

No.	Question	Response
1	How much funding will public institutions receive to help install low carbon heating and energy-saving measures?	Schools, hospitals, and other public buildings will receive over £530 million to help install low carbon heating and energy-saving measures.
2	What energy-efficient upgrades will the government support in public buildings?	The specific energy-efficient upgrades supported by the over £557 million government investment in public buildings include the installation of heat pumps, solar panels, insulation, and low-energy lighting to reduce the use of fossil fuels.
3	How many projects have received funding through the Public Sector Decarbonisation Scheme since 2020?	Over 1,000 projects have received funding through the Public Sector Decarbonisation Scheme since 2020.
4	Which businesses will receive funding from the Industrial Energy Transformation Fund and for what projects?	The businesses announced to receive funding from the Industrial Energy Transformation Fund and their respective projects are: Pilkington UK, Sofidel, and Plastipak UK.
5	What is the estimated annual savings for taxpayers from decarbonising the public sector by 2037?	The estimated annual savings for taxpayers from decarbonising the public sector by 2037 is £650 million.
6	What types of installations does the Boiler Upgrade Scheme (BUS) support with upfront capital grants?	The BUS supports the installation of heat pumps and biomass boilers in homes and non-domestic buildings in England and Wales.
7	What are the new grant levels for air source heat pumps and ground source heat pumps under the BUS?	The new grant levels under the BUS are £7,500 off the cost and installation of an air source heat pump and £7,500 off the cost and installation of a ground source heat pump, including water source heat pumps.
8	What is the total BUS funding from 2022 to 2028?	The BUS has £450 million for 2022–2025 and an additional £1.5 billion for 2025–2028.
9	What is the purpose of the Non-Domestic Renewable Heat Incentive (NDRHI) and which regions does it cover?	The NDRHI is a government environmental programme designed to increase the uptake of renewable heat to help reduce carbon emissions and meet the UK's renewable energy targets. It operates within England, Scotland, and Wales.
10	How and for how long are accredited installations compensated under the NDRHI scheme?	Accredited installations receive quarterly payments over 20 years based on the amount of eligible heat generated.
11	Who is eligible for the NDRHI scheme and what are the installation criteria?	The NDRHI scheme supports businesses, public sector, and non-profit organisations. To be eligible, equipment must have been installed in England, Scotland, or Wales on or after 15 July 2009, along with meeting other specified criteria.
12	What steps is the UK government taking to protect the best agricultural land while promoting solar energy projects?	The government is ensuring that large solar projects avoid 'Best and Most Versatile' (BMV) land and are instead developed on brownfield land, contaminated land, industrial land, and lower quality agricultural land to protect food security.
13	How does the government plan to balance food security with the deployment of solar energy?	The government prioritizes high-quality land for food production, promotes rooftop and lower-quality land solar installations, and considers the cumulative impact on communities.
14	What measures are being taken to streamline the deployment of rooftop solar installations?	The government has made changes to permitted development rights to cut red tape and make it easier and

Table 3 (*continued*)

No.	Question	Response
		cheaper to install solar panels on rooftops of commercial buildings, including farm buildings, warehouses, factories, and carparks.
15	What is the purpose of the Renewable Energy Planning Database expansion announced by the Energy Security Secretary?	The expansion aims to track land use in solar projects, ensuring planning policy is informed by robust data on high-quality agricultural land.
16	What financial incentives and support are being provided to farmers to adopt solar energy?	The government launched the Improving Farm Productivity grant, offering £15-25 million for rooftop solar and equipment to reduce fossil fuel use, with a zero VAT rate on solar panels until March 2027.
17	What new regulations are being introduced to improve consumer protections for heat network customers?	New regulations will provide compensation for loss of heating, fairer pricing, clearer billing, and increased protections for vulnerable customers, bringing similar consumer rights as those for traditional gas and electricity connections.
18	Who will regulate the new protections for heat network customers once they are in place?	Ofgem will regulate new protections and intervene in cases of pricing, service, and reliability issues. The Energy Ombudsman will handle complaints, and Citizens Advice and Consumer Scotland will provide support.
19	How are vulnerable households expected to benefit from the new heat network regulations?	Vulnerable households, such as the elderly and those with health conditions, will receive increased protections under the new regulations, ensuring they have fairer pricing, clearer billing, and compensation for heating loss.
20	What support did the proposals for new heat network regulations receive from industry and consumer groups?	The proposals received strong support from industry bodies, consumer groups, and heat network customers who responded to the consultation in August 2023.
21	What projects have received funding from the Heat Network Efficiency Scheme, and how will they benefit consumers?	Newport City Homes received £3.7 million to upgrade the Duffryn Heating System, benefiting over 970 homes, a school, and businesses. Bristol Heat Networks received £746,582 to improve service for 740 residents.
22	What was the primary oil production level in 2023?	In 2023, primary oil production hit 34 million tonnes, the lowest since North Sea production began in the 1970 s.
23	How did the total net imports of oil in 2023 compare to previous years, and what factors contributed to this change?	Total net imports of oil in 2023 reached a record high of 30 million tonnes since the establishment of North Sea oil production. This increase was driven by a sharp reduction in exports, which were down by 10 per cent compared to 2022 and nearly 30 per cent compared to 2019.

%, Accuracy of 15 %, and an F1-score of 12.9 %. These results indicate that Llama3.1:8B provides a more balanced and effective response generation compared to its predecessors, particularly in terms of both Precision and Recall, which contribute to its higher overall Accuracy and F1-score.

To enhance Chatbot's performance in the Energy domain, we implemented the RAG method. To assess the effectiveness of this approach, we evaluated it using the Llama2:7B, Llama3:8B, and Llama3.1:8B models with the previously described test dataset.

Table 6 showcases a selection of 7 questions from the test dataset, along with their Gold Standard Answers and the responses generated by

Table 4

Analysis of model responses without RAG: LLAMA2 AND LLAMA3 vs. gold standard answers.

No.	Question	Gold Standard Answer	Response with Llama2:7B	Response with Llama3:8B	Response with Llama3.1:8B
1	What is the total Boiler Upgrade Scheme (BUS) funding from 2022 to 2028?	The BUS has £450 million for 2022–2025 and an additional £1.5 billion for 2025–2028.	The total BUS funding from 2022 to 2028 is not explicitly stated in the provided information.	According to the UK government's website, the total funding allocated for the Boiler Upgrade Scheme (BUS) from 2022 to 2028 is £450 million.	The total Boiler Upgrade Scheme (BUS) funding from 2022 to 2027, not 2028, is £450 million.
2	What was the primary oil production level in 2023?	In 2023, primary oil production hit 34 million tonnes, the lowest since North Sea production began in the 1970 s	I apologize, but I don't have access to real-time data or statistics. My training data only goes up to December 2022, and I cannot provide information beyond that date.	I don't have real-time data or information on current events, including the primary oil production level in 2023. However, I can suggest some reliable sources where you can find the latest data on oil production levels: IEA, EIA and OPEC.	I don't have access to real-time or specific data on the current (2023) oil production levels.
3	What is the estimated annual savings for taxpayers from decarbonising the public sector by 2037?	The estimated annual savings for taxpayers from decarbonising the public sector by 2037 is £650 million.	I don't have access to real-time data or specific estimates on the annual savings for taxpayers from decarbonizing the public sector by 2037.	I don't have real-time data or specific estimates on the annual savings for taxpayers from decarbonizing the public sector by 2037.	I don't have specific information on estimated annual savings for taxpayers from decarbonizing the public sector by 2037.

Table 5

Comparative evaluation of chatbot performance without RAG integration using Llama2, Llama3, and Llama3.1 models.

Model	Precision (%)	Recall (%)	Accuracy (%)	F1-score (%)
Chatbot without RAG with Llama2:7B	18	10	15	12.9
Chatbot without RAG with Llama3:8B	35	25	30	29.2
Chatbot without RAG with Llama3.1:8B	40	37	38	38.4

the Llama2:7B, Llama3:8B, and Llama3.1:8B models.

For Question 1 about the UK's oil production, Llama2:7B provides detailed historical context but with a long execution time of 15.42 s. Llama3:8B includes Q4 2023 data, leading to an even longer execution time of 24.38 s. In contrast, Llama3.1:8B offers a concise answer with the shortest execution time of 3.22 s. For Question 2, Llama2:7B lists both declining and record-high sources in 4.86 s, while Llama3:8B gives more detailed information, including natural gas, in 3.36 s. Llama3.1:8B provides the most comprehensive details, including wind and solar, with a slightly longer execution time of 3.92 s.

In Question 3, Llama2:7B focuses on the 2035 decarbonization goal but lacks context. Llama3:8B includes details on net zero and climate targets, referencing the Climate Change Act 2008. Llama3.1:8B adds information from the Strategy and Policy Statement but takes longer to execute. For Question 4, Llama2:7B gives a thorough overview of the strategic approach and SSEP, while Llama3:8B offers a similar but more concise response. Llama3.1:8B provides the briefest response with the shortest execution time, focusing on the strategic approach. Regarding Question 5, Llama2:7B accurately states the aim to meet 25 % of the UK's electricity demand with up to 24GW of civil nuclear power. Llama3:8B offers a similar response with slightly faster execution, and Llama3.1:8B provides the most succinct answer with the shortest execution time. For Question 6, Llama2:7B and Llama3:8B detail the UK government's £2.5 billion investment in Sizewell C, while Llama3.1:8B offers a concise summary with a slight reduction in detail but comparable execution times. Finally, for Question 7, Llama2:7B broadly addresses the goal of a decarbonized and secure electricity system at minimal cost. Llama3:8B gives a detailed description of market arrangements, and Llama3.1:8B provides a direct and succinct response, closely aligned with the strategic priority.

The initial observations from analyzing the responses of the Llama2:7B, Llama3:8B, and Llama3.1:8B models are summarized in Table 7. This table compares these models across three key performance factors: Execution Time, Response Detail, and Accuracy. Execution Time

measures the response speed of each model. Response Detail evaluates the depth of information provided by each model. Accuracy assesses how precisely the models convey information. Llama2:7B generally has longer execution times, indicating a slower response speed. Llama3:8B, in contrast, often results in the longest execution times for some questions due to its additional processing requirements. On the other hand, Llama3.1:8B excels with the shortest execution times, offering quicker responses. Llama2:7B tends to offer detailed responses, providing substantial context. Llama3:8B adds further context and extra details, making its answers comprehensive but sometimes lengthy. Llama3.1:8B focuses on delivering concise and efficient responses, which, while succinct, may omit some contextual depth. Llama2:7B produces thorough and mostly accurate answers. Llama3:8B provides accurate responses comparable to Llama2:7B but with added detail. Llama3.1:8B prioritizes delivering concise and accurate answers, maintaining precision while being more efficient. By understanding these nuances, analysts can better grasp the strengths and limitations of each Llama model, aiding in the selection of the most appropriate model for specific needs and contexts.

Table 8 presents a comprehensive comparative evaluation of the Energy Chatbot's performance with RAG integration across the Llama2:7B, Llama3:8B, and Llama3.1:8B models. The results show significant improvements across all metrics when RAG is utilized. Among the models, the Chatbot using Llama3.1:8B with RAG integration exhibits the highest performance, achieving a Precision of 94.2 %, Recall of 95 %, Accuracy of 94 %, and an F1-score of 94.6 %. The Chatbot with Llama3:8B and RAG also performs exceptionally well, with Precision at 88 %, Recall at 90 %, Accuracy at 89 %, and an F1-score of 89.0 %. The Llama2:7B model with RAG integration, while slightly trailing the others, still shows strong results with 85 % Precision, 87 % Recall, 85 % Accuracy, and an F1-score of 86.0 %. These findings indicate that the integration of RAG significantly enhances Chatbot's ability to deliver accurate, relevant, and consistent information, with Llama3.1:8B emerging as the most effective model in this setup.

4. Discussion

A wealth of information related to various components of SEIs, and the broader Energy landscape is dispersed across diverse media sources [102–104], which can effectively drive SETs. Each SEI necessitates up-to-date knowledge of government policies, funding, and support schemes. Additionally, understanding current market trends, available technologies, environmental impacts, economic analysis, and social acceptance is crucial [102]. Information on global and regional initiatives, success stories, and the risks involved in adopting new methods also plays a vital role in shaping these transitions. This comprehensive

Table 6

Comparison of model responses and execution times across different LLMs.

No.	Question	Gold Standard Answer	Llama2:7B		Llama3:8B		Llama3.1:8B	
			Response	Execution Time (sec)	Response	Execution Time (sec)	Response	Execution Time (sec)
1	What was the total production of oil equivalent in the UK in 2023, and how did it compare to previous years?	In 2023, the total production was 100.4 million tonnes of oil equivalent, which was 8.9 per cent lower than in 2022 and 66 per cent lower than in 1999 when UK production peaked. This represents a record low level in the published series.	In 2023, the UK's total oil equivalent production dropped to 100.4 million tonnes, marking an 8.9 % decrease from 2022 and a record low. Compared to the peak in 1999, oil production was 66 % lower. While coal, oil, and nuclear outputs reached record lows, wind, solar, and hydro electricity production increased by 1.5 % to a record high, driven by expanded capacity.	15.42	The total production of oil equivalent in the UK in 2023 was 100.4 million tonnes, which is: 8.9 % lower than in 2022 and 66 % lower than in 1999 when UK production peaked. Additionally, it's mentioned that in the fourth quarter of 2023, total production was 25.5 million tonnes of oil equivalent, which is 9.2 % lower than the same period in 2022.	24.38	The total production of oil equivalent in the UK in 2023 was 100.4 million tonnes, which was 8.9 per cent lower than in 2022 and at a record low level in the published series, and 66 per cent lower than in 1999 when UK production peaked.	3.22
2	Which energy sources saw a decline in production in 2023, and which sources reached record highs?	In 2023, production levels for all fuels except wind and solar declined. Coal, oil, and nuclear output reached record lows, while output from wind, solar, and hydro reached record highs.	The following energy sources saw a decline in production in 2023: Coal, Oil and Nuclear. On the other hand, the following energy sources reached record highs in 2023: Wind and solar (due to increased capacity).	4.86	The following energy sources saw a decline in production in 2023: Coal, Oil, Nuclear (due to reduced capacity and outages), Natural gas (fell by 10 %). On the other hand, the following energy sources reached record highs in 2023: Wind power, Solar power.	3.36	The following energy sources saw a decline in production in 2023: Coal, Oil, Nuclear (due to reduced capacity and outages), Natural gas, Onshore wind and hydro output (while total wind and solar output reached a record high due to increased offshore wind and solar capacity). On the other hand, the following energy sources reached record highs in 2023: Wind and solar output (total).	3.92
3	What is one of the UK Government's strategic priorities regarding its electricity system by 2035?	One of the strategic priorities is to achieve a decarbonised electricity system by 2035, subject to security of supply, in order to meet the UK's net zero and climate change targets.	One of the UK Government's strategic priorities regarding its electricity system by 2035 is: "Electricity market arrangements that meet our objectives for a decarbonised and secure electricity system at least possible cost to consumers".	3.18	One of the UK Government's strategic priorities regarding its electricity system by 2035 is to meet the UK's net zero and climate change targets, including the carbon budgets set under the Climate Change Act 2008 and the target for a decarbonised electricity system by 2035, subject to security of supply.	2.72	One of the UK Government's strategic priorities regarding its electricity system by 2035 is: "To have market arrangements that meet our objectives for a decarbonised and secure electricity system at least possible cost to consumers". (Section 3 of the Strategy and Policy Statement).	3.16
4	What approach does the UK Government emphasize for planning and building network infrastructure?	The UK Government emphasizes a strategic, whole-system approach to plan and build reliable, resilient, and sustainable network infrastructure that is appropriately connected to wider markets.	The UK Government emphasizes a "strategic, whole system approach" to plan and build reliable, resilient, and sustainable network infrastructure. This approach aims to provide certainty needed by industries to build with confidence, while controlling system costs by facilitating a smart, flexible, and digitalized energy	5.30	The UK Government promotes a whole-system strategy for sustainable network infrastructure, prioritizing proactive investment and cost control through a smart, flexible energy system. This approach supports the net zero transition and renewable targets. The government will also introduce the first Strategic Spatial Energy Plan (SSEP) to	3.78	The UK Government emphasizes a "strategic, whole system approach" to plan and build reliable, resilient, sustainable network infrastructure. This approach is mentioned in Section 1 of the Strategy and Policy Statement for Energy Policy in Great Britain, under the Government's Strategic Priorities.	2.37

Table 6 (*continued*)

No.	Question	Gold Standard Answer	Llama2:7B		Llama3:8B		Llama3.1:8B	
			Response	Execution Time (sec)	Response	Execution Time (sec)	Response	Execution Time (sec)
			system. Additionally, the government is committed to developing an interim Strategic Spatial Energy Plan (SSEP) to define the optimal location of generation and infrastructure required to meet forecast demand and 2050 targets.		identify optimal locations for generation and infrastructure, providing industries with certainty for 2050 goals.			
5	What percentage of the UK's anticipated electricity demand by 2050 is the government aiming to meet with civil nuclear power, according to the British Energy Security Strategy?	The government aims to meet around 25 % of the UK's anticipated electricity demand by 2050 with up to 24GW of civil nuclear power.	According to the British Energy Security Strategy, the government is aiming to deploy up to 24GW of civil nuclear power by 2050, meeting around 25 % of the UK's anticipated electricity demand by 2050.	2.57	The government aims to meet around 25 % of the UK's anticipated electricity demand by 2050 through the deployment of up to 24GW of civil nuclear power, as stated in the British Energy Security Strategy.	2.33	The British Energy Security Strategy aims to deploy up to 24GW of civil nuclear power by 2050, which will meet around 25 % of the UK's anticipated electricity demand.	2.08
6	What significant step did the UK government take in November 2022 regarding the Sizewell C nuclear project?	In November 2022, the UK government became a shareholder in the Sizewell C project alongside EDF and made available £2.5 billion of investment to support the project's development.	The significant step that the UK government took in November 2022 regarding the Sizewell C nuclear project was becoming a shareholder in the project alongside EDF and making available a total of £2.5bn of investment to support the project's development.	2.71	Since November 2022, the UK government became a shareholder in the Sizewell C project with EDF and made available a total of £2.5bn of investment to support the project's development.	2.09	The UK government made itself a shareholder in the Sizewell C project with EDF and made available a total of £2.5bn of investment to support the project's development.	2.07
7	What is one of the government's strategic priorities for electricity market arrangements by 2035?	One of the government's strategic priorities is to ensure electricity market arrangements that achieve a decarbonised and secure electricity system by 2035 at the lowest possible cost to consumers.	One of the government's strategic priorities for electricity market arrangements by 2035 is to have "electricity market arrangements that meet our objectives for a decarbonised and secure electricity system at least possible cost to consumers".	4.12	One of the government's strategic priorities for electricity market arrangements by 2035 is to ensure electricity markets that are competitive, transparent, and liquid. This means creating a market structure that promotes competition among energy generators and suppliers, provides transparency in the pricing and trading of electricity, and allows for easy entry and exit of participants.	2.80	One of the government's strategic priorities for electricity market arrangements by 2035 is to achieve a decarbonised and secure electricity system at least possible cost to consumers.	2.04

knowledge collectively drives SETs by offering detailed insights into the Energy landscape. However, SMEs leading Energy transition efforts often lack access to integrated ISs or Chatbots that consolidate this diverse data and provide the latest updates essential for driving SETs.

Implementing an IS involves several cost categories. Initial expenses include hardware purchases, software licensing, implementation, customization, and data migration [105,106]. Ongoing operating costs encompass maintenance, Information Technology (IT) personnel salaries, utility expenses, and security measures. Additional financial considerations include training, ongoing support, system upgrades, scalability, and potential downtime. Depreciation and replacement costs further influence financial planning. Existing Energy-related ISs often focus on aggregating news articles and social media data, lacking the depth and breadth of information needed for comprehensive decision-making. As noted in Section 2, many of these systems are also not open access, limiting their accessibility and utility.

To address this research gap, this study introduces an LLM with a multi-source RAG-based system in the form of an Energy Chatbot. This system acts as an integrated IS for SMEs seeking the latest information on Energy-related topics. These topics include Policy Updates, Market Trends, Technological Innovations, Environmental Impacts, Economic Analysis, Social Acceptance and Community Engagement, Global and Regional Initiatives, Case Studies and Success Stories, and Risk Assessment and Resilience. To gather this information, we utilized diverse

Table 7

Observing LLM performance in terms of execution time, response detail, and accuracy.

No.	Factor	Llama Model Variants		
		Llama2:7B	Llama3:8B	Llama3.1:8B
1	Execution Time	Mostly long execution times.	Results in the longest execution times for some questions.	The shortest execution times.
2	Response Detail	Generally provides detailed responses.	Often includes additional information	Tends to deliver concise and efficient responses. Answers are succinct but may lack some context or detail.
3	Accuracy	Produces thorough and mostly precise answers.	Produces accurate answers comparable to Llama2:7B.	Prioritizes concise and accurate answers.

Table 8

Comparative evaluation of Chatbot performance with RAG integration using Llama2, Llama3, and Llama3.1 models.

Model	Precision (%)	Recall (%)	Accuracy (%)	F1-score (%)
Energy Chatbot using RAG with Llama2:7B	85	87	85	86.0
Energy Chatbot using RAG with Llama3:8B	88	90	89	89.0
Energy Chatbot using RAG with Llama3.1:8B	94.2	95	94	94.6

media sources such as news websites, government reports, industry publications, academic research, and social media.

The scope of the Energy Chatbot is intentionally broad, encompassing a diverse range of Energy-related topics to provide valuable insights for SMEs seeking up-to-date information. This broad approach allows Chatbot to address a wide array of queries and integrate various sources of information through the RAG with LLM system. While this research focuses on demonstrating Chatbot's capabilities across above-mentioned topics, it acknowledges that other factors, not covered in this study, may also play a critical role in understanding the landscape of SETs. The expansive scope serves to illustrate how different types of information can be incorporated into the RAG with LLM-based ISs and accessed via the Chatbot interface. The scope of the Energy Chatbot is designed to be adaptable, allowing for adjustments to the information sources to better suit specific interests. While the content and focus can be customized, the underlying technology configuration will remain consistent to ensure that the Chatbot functions effectively across various contexts.

The implementation of the Energy Chatbot across three different Llama variants reveals distinct strengths and limitations. Llama2:7B generally provides detailed responses, which can be beneficial for users needing comprehensive information. However, it often includes extraneous details that may obscure the core answer, making it challenging for users to quickly grasp key points. Additionally, this model struggles with ambiguities related to names and locations, further complicating the clarity of its responses. Llama3:8B is notable for its thoroughness, offering highly detailed answers that cover a broad spectrum of information. While this detail enhances comprehensiveness, it can be overwhelming for users who prefer a concise summary, potentially detracting from its effectiveness for those seeking focused insights. Llama3.1:8B excels at delivering accurate and comprehensive responses that effectively address all critical aspects of a query. Its strength lies in its ability to balance detail with readability, ensuring that answers are both concise and clear. However, while the model maintains brevity and

clarity, this can sometimes result in responses that, though succinct, may lack some contextual depth or nuanced detail.

The proposed system offers several notable advantages over existing solutions. It excels in managing a diverse range of Energy-related information, effectively handling various media types and structures, whether semi-structured or unstructured. Its dynamic nature allows for easy integration of additional datasets, ensuring it remains current and relevant. The system's implementation is straightforward, leveraging RAG with LLM technology to seamlessly access the latest Energy-related datasets with minimal development effort. Updating the system involves simply replacing old data files with new ones, making the development and upgrade process efficient and requiring minimal time and resources. These attributes make the system a sustainable choice for SMEs aiming to invest responsibly and minimize the carbon footprint associated with new technologies.

By automating data retrieval and analysis, the RAG with LLM system reduces the need for extensive hardware and software investments, thereby mitigating many of the costs associated with traditional IS implementations. It provides accurate, up-to-date insights, lowering operational expenses related to personnel and training while enhancing adaptability to evolving data sources and reducing downtime risks. Furthermore, utilizing open-source LLMs with RAG enhances data security. Unlike proprietary systems that may be susceptible to data leaks or unauthorized access, open-source solutions offer greater control over data protection [107]. The accessible source code allows organizations to customize and manage their security measures effectively, ensuring responsible data handling and minimizing risks associated with sensitive information. This transparency and adaptability make open-source LLMs a more secure and flexible choice for managing information.

While LLMs demonstrate impressive language comprehension and generation capabilities, their performance is heavily dependent on the initial training data. As this data ages, the efficiency of LLMs in understanding and generating relevant context diminishes. LLMs trained on datasets that are even one day old cannot provide the latest media updates due to the inherent limitation of their training data. Therefore, integrating the LLM with RAG technology allows SMEs to remain upgradable and customizable, adapting to evolving organizational needs and interests over time with minimal cost. However, it is crucial to note that the quality of information generated by this integrated IS still relies on the external datasets linked to the LLM via the RAG functionality. There is an ongoing need to carefully extract accurate data from online media sources and add it to the RAG datasets. The presence of false information in the collected dataset could result in the generation of inaccurate information for SMEs.

Apart from the benefits, this research also has limitations. For the evaluation, we used a small set of manually selected direct questions, which may not comprehensively cover the vast range of topics related to Energy information. More extensive and detailed tests should be established to minimize bias in the testing process. Moreover, we used fundamental evaluation metrics (Precision, Recall, Accuracy, and F1-score) to assess the system responses against the Gold Standard Answers. However, there are also advanced evaluation metrics, such as faithfulness, context relevance, answer relevance, and answer correctness, recently introduced by various researchers [97,108]. These advanced metrics should be explored to achieve a more accurate and relevant assessment of the information provided by the system, compared to the approach used in this article.

Additionally, since the RAG technology retrieves information from the dataset, incorporating more data sources can result in delayed Chatbot responses. An evaluation in such scenarios is essential to identify advanced techniques for optimizing the IE process and minimizing delays. Another limitation is our reliance on a single model family, specifically the Llama-based models (Llama2, Llama3, and Llama3.1), for implementing the RAG with LLM integrated solution in our case study. Future research should explore the performance of other language models, such as GPT, to determine which performs best in the

Energy domain.

Furthermore, the Energy Chatbot is configured to provide answers based solely on the factual information present in the datasets by minimizing the LLM's creativity feature [109,110]. While this ensures responses are based on accurate data sources, it also limits the model's ability to infer and provide broader knowledge. Even slight variations in organization names, financial amounts, or other ICs can lead to inaccurate information. Further work is needed to explore different levels of LLM creativity to see how this integrated approach can function as an effective recommendation system for SMEs, understanding and leveraging historical Energy-related datasets.

5. Conclusions

SMEs are crucial for the UK's transition to a net-zero economy by 2050, focusing on upgrading homes by 2035 to reduce energy consumption and costs. However, SMEs often lack essential knowledge on SEIs, which can lead to inefficiencies and higher costs. Access to timely information on SEIs from government policies and technologies to market trends and environmental impacts is vital for effective SETs. Existing SME-focused systems frequently struggle to integrate diverse data sources effectively, hindering their ability to navigate SEIs comprehensively. Our study introduces an advanced LLM with a multi-source RAG-based Energy Chatbot, designed to provide SMEs with up-to-date insights across Policy Updates, Market Trends, Technological Innovations, and more. The comparative evaluation of Chatbot performance, as shown in our results, highlights a marked improvement with RAG integration. Specifically, the Energy Chatbot using RAG with Llama2, Llama3, and Llama3.1 models significantly outperformed those without RAG integration in terms of Precision, Recall, Accuracy, and F1-scores. This demonstrates the effectiveness of RAG in enhancing the quality and reliability of Chatbot responses. While promising, our approach requires further evaluation of accuracy and efficiency, considering potential delays in response times with increased data sources and the need to explore alternative LLM models for optimal performance. Future research should focus on addressing these challenges to better support SMEs in achieving their sustainable energy goals.

CRediT authorship contribution statement

Muhammad Arslan: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Lamine Mahdjoubi:** Resources. **Saba Munawar:** Data Curation, Validation, Writing- Review & Editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

References

[1] Green-Alliance (2021). Reinventing retrofit How to scale up home energy efficiency in the UK. Accessed: July 01, 2024. [Online]. Available: https://green-alliance.org.uk/wp-content/uploads/2021/11/reinventing_retrofit.pdf.

[2] Building-UK (2023). The battle SMEs face in reducing their emissions - and proving it. Accessed: July 01, 2024. [Online]. Available: https://www.building.co.uk/building-the-future-commission/the-battle-smes-face-in-reducing-their-emissions-and-proving-it/5124573.article.

[3] L.L.J. Meijer, J.C.C.M. Huijben, A. Van Boxstael, A.G.L. Romme, Barriers and drivers for technology commercialization by SMEs in the Dutch sustainable energy sector, Renew. Sustain. Energy Rev. 112 (2019) 114–126.

[4] F.R. Munro, P. Cairney, A systematic review of energy systems: the role of policymaking in sustainable transitions, Renew. Sustain. Energy Rev. 119 (2020) 109598.

[5] A. Kalair, N. Abas, M.S. Saleem, A.R. Kalair, N. Khan, Role of energy storage systems in energy transition from fossil fuels to renewables, Energy Storage 3 (1) (2021) e135.

[6] M. Amir, R.G. Deshmukh, H.M. Khalid, Z. Said, A. Raza, S.M. Muyeen, K. Sopian, Energy storage technologies: an integrated survey of developments, global economical/environmental effects, optimal scheduling model, and sustainable adaption policies, J. Storage Mater. 72 (2023) 108694.

[7] A. Shivakumar, A. Dobbins, U. Fahl, A. Singh, Drivers of renewable energy deployment in the EU: an analysis of past trends and projections, Energ. Strat. Rev. 26 (2019) 100402.

[8] O.A. Adelekan, B.S. Ilugbusi, O. Adisa, O.C. Obi, K.F. Awonuga, O.F. Asuzu, N. L. Ndubuisi, Energy transition policies: a global review of shifts towards renewable sources, Eng. Sci. Technol. J. 5 (2) (2024) 272–287.

[9] B. Lennon, N.P. Dunphy, E. Sanvicente, Community acceptability and the energy transition: a citizens' perspective, Energy, Sustainab. Soc. 9 (1) (2019) 1–18.

[10] F. Cappellaro, G. D'Agosta, P. De Sabbata, F. Barroco, C. Carani, A. Borghetti, C. A. Nucci, Implementing energy transition and SDGs targets throughout energy community schemes, J. Urban Ecol. 8 (1) (2022) juac023.

[11] S.V. Oprea, A. Bâra, Generative literature analysis on the rise of prosumers and their influence on the sustainable energy transition, Util. Policy 90 (2024) 101799.

[12] F. Neresini, P. Giardullo, E. Di Buccio, A. Cammozzo, Exploring socio-technical future scenarios in the media: the energy transition case in Italian daily newspapers, Qual. Quant. 54 (2020) 147–168.

[13] S.A. Qadir, H. Al-Motairi, F. Tahir, L. Al-Fagih, Incentives and strategies for financing the renewable energy transition: a review, Energy Rep. 7 (2021) 3590–3606.

[14] J. Markard, R. Raven, B. Truffer, Sustainability transitions: an emerging field of research and its prospects, Res. Policy 41 (6) (2012) 955–967.

[15] S. Hyysalo, Citizen activities in energy transition: user innovation, new communities, and the shaping of a sustainable future, Taylor & Francis, 2021, p. 190.

[16] U. Schmitt, Supporting the sustainable growth of SMEs with content-and collaboration-based personal knowledge management systems, J. Entrepreneur. Innovat. Emerg. Econom. 4 (1) (2018) 1–21.

[17] A. Mory-Alvarado, C. Juiz, B. Bermejo, M. Campoverde-Molina, Green IT in small and medium-sized enterprises: A systematic literature review, Sustainable Comput. Inf. Syst. 100891 (2023).

[18] C. Cavicchi, C. Oppi, E. Vagnoni, Energy management to foster circular economy business model for sustainable development in an agricultural SME, J. Clean. Prod. 368 (2022) 133188.

[19] M.J.B. Kabeyi, O.A. Olanrewaju, Sustainable energy transition for renewable and low carbon grid electricity generation and supply, Front. Energy Res. 9 (2022) 743114.

[20] Y. Chen, I.H. Rowlands, The socio-political context of energy storage transition: insights from a media analysis of Chinese newspapers, Energy Res. Soc. Sci. 84 (2022) 102348.

[21] C. Piselli, A.F. Colladon, L. Segneri, A.L. Pisello, Evaluating and improving social awareness of energy communities through semantic network analysis of online news, Renew. Sustain. Energy Rev. 167 (2022) 112792.

[22] J. Barnes, P. Hansen, T. Kamin, U. Golob, S. Darby, N.M. van der Grijp, D. Petrovics, Creating valuable outcomes: an exploration of value creation pathways in the business models of energy communities, Energy Res. Soc. Sci. 108 (2024) 103398.

[23] S. Ganowski, I.H. Rowlands, Read all about it! Comparing media discourse on energy storage in Canada and the United Kingdom in a transition era, Energy Res. Soc. Sci. 70 (2020) 101709.

[24] J. Dehler-Holland, K. Schumacher, W. Fichtner, Topic modeling uncovers shifts in media framing of the German renewable energy act, Patterns 2 (1) (2021).

[25] A. Marzouki, A. Chouikh, S. Mellouli, R. Haddad, From sustainable development goals to sustainable cities: a social media analysis for policy-making decision, Sustainability 13 (15) (2021) 8136.

[26] K. Rommetveit, I.F. Ballo, S. Sareen, Extracting users: regimes of engagement in Norwegian smart electricity transition, Sci. Technol. Hum. Values (2021), 01622439211052867.

[27] J. Krzywda, D. Krzywda, A. Androniceanu, Managing the energy transition through discourse. The Case of Poland, Energies 14 (20) (2021) 6471.

[28] R. Ibar-Alonso, R. Quiroga-García, M. Arenas-Parra, Opinion mining of green energy sentiment: a Russia-Ukraine conflict analysis, Mathematics 10 (14) (2022) 2532.

[29] J. Lyytimäki, Renewable energy in the news: Environmental, economic, policy and technology discussion of biogas, Sustain. Product. Consumpt. 15 (2018) 65–73.

[30] C. Walker, A. Alexander, M.B. Doucette, D. Lewis, H.T. Neufeld, D. Martin, H. Castleden, Are the pens working for justice? News media coverage of

renewable energy involving Indigenous Peoples in Canada, Energy Res. Soc. Sci. 57 (2019) 101230.

[31] D. Labonte, I.H. Rowlands, Tweets and transitions: Exploring Twitter-based political discourse regarding energy and electricity in Ontario, Canada, Energy Res. Soc. Sci. 72 (2021) 101870.

[32] J. Lyytimäki, N.A. Nygrén, A. Pulkka, S. Rantala, Energy transition looming behind the headlines? Newspaper coverage of biogas production in Finland, Energy, Sustainab. Soc. 8 (2018) 1–11.

[33] M. Antal, K. Karhunmaa, The German energy transition in the British, Finnish and Hungarian news media, Nat. Energy 3 (11) (2018) 994–1001.

[34] Gnewuch, U., Morana, S., Heckmann, C., & Maedche, A. (2018). Designing conversational agents for energy feedback. In Designing for a Digital and Globalized World: 13th International Conference, DESRIST 2018, Chennai, India, June 3–6, 2018, Proceedings 13 (pp. 18-33). Springer International Publishing.

[35] Suresan, A., Mohan, S. S., Arya, M. P., Anjana Gangadharan, V., & Bindu, P. V. (2021). A conversational AI chatbot in energy informatics. In Proceedings of International Conference on Intelligent Computing, Information and Control Systems: ICICCS 2020 (pp. 543-554). Springer Singapore.

[36] C.V.M. Rocha, A.A. Lima, P.H.C. Vieira, C.L.S. Cipriano, A.M.P. Paiva, I.F.S. da Silva, E.C. Fernandes, A chatbot solution for self-reading energy consumption via chatting applications, J. Control, Automat. Electric. Syst. 33 (2022) 229–240.

[37] Fontecha, J., González, I., & Salas-Seguín, A. (2019, November). Using Conversational Assistants and Connected Devices to Promote a Responsible Energy Consumption at Home. In Proceedings (Vol. 31, No. 1, p. 32). MDPI.

[38] Milano, F. (2023). EcoBot: study design and development of a persuasive FSM-free Chatbot for energy saving. Accessed: August 7, 2024. [Online] https://www.diva-portal.org/smash/get/diva2:1783085/FULLTEXT01.pdf.

[39] Onile, A. E., Belikov, J., Petlenkov, E., & Levron, Y. (2023, November). Leveraging Digital Twins and Demand Side Recommender Chatbot for Optimizing Smart Grid Energy Efficiency. In 2023 IEEE PES Innovative Smart Grid Technologies-Asia (ISGT Asia) (pp. 1-5). IEEE.

[40] T.A. Aminkeng, S. Huaming, N.B. Mukete, G.E. Mwalupaso, Impact of credit constraints on financial performance of small and medium size enterprises, Eur. J. Dev. Res. (2024) 1–29.

[41] D. Kindström, P. Carlborg, T. Nord, Challenges for growing SMEs: a managerial perspective, J. Small Bus. Manag. 62 (2) (2024) 700–723.

[42] S. Gawusu, Evolving energy landscapes: a computational analysis of the determinants of energy poverty, Renew. Sustain. Energy Rev. 202 (2024) 114705.

[43] S. Ben Sassi, N. Yanes, Data science with semantic technologies: application to information systems development, J. Comput. Inf. Syst. 64 (3) (2024) 388–407.

[44] D. Bianchini, V. De Antonellis, M. Garda, A semantics-enabled approach for personalised Data Lake exploration, Knowl. Inf. Syst. 66 (2) (2024) 1469–1502.

[45] Bechara, M. (2024). Understanding the Requirements of Data Spaces in the Energy Sector (Master's thesis, Hannover: Gottfried Wilhelm Leibniz Universität).

[46] V. Thomasi, J.C.M. Siluk, P.D. Rigo, C.A.D.O. Pappis, Challenges, improvements, and opportunities market with the liberalization of the residential electricity market, Energy Policy 192 (2024) 114253.

[47] M. Mountantonakis, Y. Tzitzikas, Large-scale semantic integration of linked data: A survey, ACM Computing Surveys (CSUR) 52 (5) (2019) 1–40.

[48] M. Pan, T. Li, Y. Liu, Q. Pei, E.A. Huang, J.X. Huang, A semantically enhanced text retrieval framework with abstractive summarization, Comput. Intell. 40 (1) (2024) e12603.

[49] M. Arslan, C. Cruz, Semantic Enrichment of Taxonomy for BI Applications using Multifaceted data sources through NLP techniques, Procedia Comput. Sci. 207 (2022) 2424–2433.

[50] Venkatakrishnan, R., Tanyildizi, E., & Canbaz, M. A. (2024, May). Semantic interlinking of Immigration Data using LLMs for Knowledge Graph Construction. In Companion Proceedings of the ACM on Web Conference 2024. pp. 605-608.

[51] Johnsen, M. (2024). Large Language Models (LLMs). Maria Johnsen.

[52] D. Tsirmpas, I. Gkionis, G.T. Papadopoulos, I. Mademlis, Neural natural language processing for long texts: a survey on classification and summarization, Eng. Appl. Artif. Intel. 133 (2024) 108231.

[53] Jiang, Z., Lin, H., Zhong, Y., Huang, Q., Chen, Y., Zhang, Z., ... & Liu, X. (2024). {MegaScale}: Scaling large language model training to more than 10,000 {GPUs}. In 21st USENIX Symposium on Networked Systems Design and Implementation (NSDI 24) (pp. 745-760).

[54] Ge, Y., Hua, W., Mei, K., Tan, J., Xu, S., Li, Z., & Zhang, Y. (2024). Openagi: When llm meets domain experts. Advances in Neural Information Processing Systems, 36.

[55] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, D. Kiela, Retrieval-augmented generation for knowledge-intensive nlp tasks, Adv. Neural Inf. Proces. Syst. 33 (2020) 9459–9474.

[56] 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.

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

[58] Zhang, B., Yang, H., Zhou, T., Ali Babar, M., & Liu, X. Y. (2023, November). Enhancing financial sentiment analysis via retrieval augmented large language models. In: Proceedings of the Fourth ACM International Conference on AI in Finance (pp. 349-356).

[59] 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.

[60] 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.

[61] Xia, M., Zhang, X., Couturier, C., Zheng, G., Rajmohan, S., & Ruhle, V. (2023). Hybrid retrieval-augmented generation for real-time composition assistance. arXiv preprint arXiv:2308.04215.

[62] Rackauckas, Z. (2024). RAG-Fusion: a New Take on Retrieval-Augmented Generation. arXiv preprint arXiv:2402.03367.

[63] Shi, E., Wang, Y., Tao, W., Du, L., Zhang, H., Han, S., & Sun, H. (2022). RACE: Retrieval-Augmented Commit Message Generation. arXiv preprint arXiv: 2203.02700.

[64] 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.

[65] I.K.F. Haugeland, A. Følstad, C. Taylor, C.A. Bjørkli, Understanding the user experience of customer service chatbots: an experimental study of chatbot interaction design, Int. J. Hum Comput Stud. 161 (2022) 102788.

[66] S.M. Suhaili, N. Salim, M.N. Jambli, Service chatbots: a systematic review, Expert Syst. Appl. 184 (2021) 115461.

[67] K.K. Nirala, N.K. Singh, V.S. Purani, A survey on providing customer and public administration based services using AI: chatbot, Multimed. Tools Appl. 81 (16) (2022) 22215–22246.

[68] GOV-UK (2024). Climate-change-energy. Accessed: July 09, 2024. [Online]. Available: https://www.gov.uk/environment/climate-change-energy.

[69] Ofgem (2024). Ofgem. Accessed: July 01, 2024. [Online]. Available: https://www.ofgem.gov.uk/.

[70] UKRI (2024). UKRI. Accessed: July 01, 2024. [Online]. Available: https://www.ukri.org/.

[71] InnovateUK (2024). InnovateUK. Accessed: July 09, 2024. [Online]. Available: https://www.ukri.org/councils/innovate-uk/.

[72] Environment-agency (2024). Environment-agency. Accessed: July 01, 2024. [Online]. Available: https://www.gov.uk/government/organisations/environment-agency.

[73] Carbontrust (2024). Carbontrust. Accessed: July 09, 2024. [Online]. Available: https://www.carbontrust.com/.

[74] ONS (2024). Census 2021 Data and analysis from Census 2021. Accessed: July 09, 2024. [Online]. Available: https://www.ons.gov.uk/.

[75] DUKES (2023). Department for Energy Security and Net Zero. Department for Energy Security and Net Zero. Accessed: July 09, 2024. [Online]. Available: https://assets.publishing.service.gov.uk/media/64f1fcba9ee0f2000db7bdd8/DUKES_2023_Chapters_1-7.pdf.

[76] Yougov (2024). Yougov. Accessed: July 09, 2024. [Online]. Available: https://yougov.co.uk/.

[77] Unfccc (2015). The Paris Agreement. Accessed: July 01, 2024. [Online]. Available: https://unfccc.int/process-and-meetings/the-paris-agreement.

[78] theccc (2024). Publication of the UK Progress Report. Accessed: July 09, 2024. [Online]. Available: https://www.theccc.org.uk/.

[79] European data (2019). Open energy data on the European Data Portal. Accessed: July 01, 2024. [Online]. Available: https://data.europa.eu/en/publications/datastories/open-energy-data-european-data-portal.

[80] Catapult (2024). Get safe access to India's booming market, at your pace. Accessed: July 01, 2024. [Online]. Available: https://es.catapult.co.uk/.

[81] Renewableuk (2024). Renewableuk. Accessed: July 09, 2024. [Online]. Available: https://www.renewableuk.com/.

[82] Nationalgrideso (2024). ESO. Accessed: July 09, 2024. [Online]. Available: https://www.nationalgrideso.com/.

[83] Metoffice (2024). UK Climate Projections (UKCP). Accessed: July 07, 2024. [Online]. Available: https://www.metoffice.gov.uk/research/approach/collaboration/ukcp.

[84] Ipsos (2024). Ipsos. Accessed: July 09, 2024. [Online]. Available: https://www.ipsos.com/en-uk.

[85] Greenpeace (2024). How the UK's climate majority found its voice. Accessed: July 07, 2024. [Online]. Available: https://www.greenpeace.org.uk/.

[86] Friends of the earth (2024). Join the fight for people and planet Donate today. Accessed: July 05, 2024. [Online]. Available: https://friendsoftheearth.uk/.

[87] 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.

[88] ZXhang, Y. X., Haxo, Y. M., & Mat, Y. X. (2023). Falcon llm: a new frontier in natural language processing. AC Investm. Res. J., 220(44).

[89] 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.

[90] 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.

[91] 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.

[92] 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.

[93] Langchain (2024). Applications that can reason. Powered by LangChain. Accessed: July 05, 2024. [Online]. Available: https://www.langchain.com/.

[94] 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.*

[95] llamaindex (2024). Turn your enterprise data into production-ready LLM applications. Accessed: July 05, 2024. [Online]. Available: https://www.llamaindex.ai/.

[96] Purwar, A. (2024). Evaluating the Efficacy of Open-Source LLMs in Enterprise-Specific RAG Systems: A Comparative Study of Performance and Scalability. arXiv preprint arXiv:2406.11424.

[97] Wang, Y., Hernandez, A. G., Kyslyi, R., & Kersting, N. (2024). Evaluating Quality of Answers for Retrieval-Augmented Generation: A Strong LLM Is All You Need. arXiv preprint arXiv:2406.18064.

[98] Alaofi, M., Arabzadeh, N., Clarke, C. L., & Sanderson, M. (2024). Generative Information Retrieval Evaluation. arXiv preprint arXiv:2404.08137.

[99] P. Jiang, X. Cai, A survey of text-matching techniques, Information 15 (6) (2024) 332.

[100] Fatehkia, M., Lucas, J. K., & Chawla, S. (2024). T-RAG: lessons from the LLM trenches. arXiv preprint arXiv:2402.07483.

[101] 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.

[102] M.O. Tetteh, E.B. Boateng, A. Darko, A.P. Chan, What are the general public's needs, concerns and views about energy efficiency retrofitting of existing building stock? A sentiment analysis of social media data, Energ. Buildings 301 (2023) 113721.

[103] H. Elsharkawy, P. Rutherford, Energy-efficient retrofit of social housing in the UK: Lessons learned from a Community Energy Saving Programme (CESP) in Nottingham, Energ. Buildings 172 (2018) 295–306.

[104] Z. Liu, C. Yu, Q.K. Qian, R. Huang, K. You, H. Visscher, G. Zhang, Incentive initiatives on energy-efficient renovation of existing buildings towards carbon-neutral blueprints in China: advancements, challenges and perspectives, Energ. Buildings 113343 (2023).

[105] Y. Pena, D.E. Correal, E. Miranda, O. González-Rojas, An innovative cost taxonomy: identifying and classifying costs of technology solutions, Int. J. Busin. Informat. Syst. 45 (3) (2024) 397–428.

[106] K.E. Pearlson, C.S. Saunders, D.F. Galletta, Managing and using information systems: a strategic approach, John Wiley & Sons, 2024.

[107] Y. Yao, J. Duan, K. Xu, Y. Cai, Z. Sun, Y. Zhang, A survey on large language model (llm) security and privacy: the good, the bad, and the ugly, High-Confidence Computing 100211 (2024).

[108] Es, S., James, J., Espinosa-Anke, L., & Schockaert, S. (2023). Ragas: Automated evaluation of retrieval augmented generation. arXiv preprint arXiv:2309.15217.

[109] Lu, L. C., Chen, S. J., Pai, T. M., Yu, C. H., Lee, H. Y., & Sun, S. H. (2024). LLM Discussion: Enhancing the Creativity of Large Language Models via Discussion Framework and Role-Play. arXiv preprint arXiv:2405.06373.

[110] Franceschelli, G., & Musolesi, M. (2023). On the creativity of large language models. arXiv preprint arXiv:2304.00008.