

Building Trust in Conversational AI: A Comprehensive Review and Solution Architecture for Explainable, Privacy-Aware Systems using LLMs and Knowledge Graph

Ahtsham Zafar, Venkatesh Balavadhani Parthasarathy, Chan Le Van, Saad Shahid, Aafaq Iqbal khan, Arsalan Shahid

13 Aug 2023
arXiv:2308.13534v1 [cs.CL]

Abstract—Conversational AI systems have emerged as key enablers of human-like interactions across diverse sectors. Nevertheless, the balance between linguistic nuance and factual accuracy has proven elusive. In this paper, we first introduce LLMXplorer, a comprehensive tool that provides an in-depth review of over 150 Large Language Models (LLMs), elucidating their myriad implications ranging from social and ethical to regulatory, as well as their applicability across industries. Building on this foundation, we propose a novel functional architecture that seamlessly integrates the structured dynamics of Knowledge Graphs with the linguistic capabilities of LLMs. Validated using real-world AI news data, our architecture adeptly blends linguistic sophistication with factual rigour, and further strengthens data security through Role-Based Access Control. This research provides insights into the evolving landscape of conversational AI, emphasizing the imperative for systems that are efficient, transparent, and trustworthy.

Index Terms—Knowledge Graphs, Large Language Models, LLMXplorer, Role-Based Access Control, Trustworthiness, Neo4j

I. INTRODUCTION

The ability to comprehend and engage in natural language dialogue is an intrinsic aspect of human intelligence. In contrast, machines inherently lack this capability, necessitating methods to translate natural language into machine-readable formats and vice versa. Conversational AI, which has evolved since the mid-1960s, addresses this gap. Initial developments were based on pattern matching, while more recent approaches leverage Generative AI technologies for coherent, context-aware responses. This evolution has captured the interest of both researchers and the wider community, leading to applications across various domains such as healthcare [1], [2], [3], retail [4], [5], journalism [6], [7], finance [8], [9], and education [10], [11].

A. History and Evolution of Conversational AI

The journey of Conversational AI can be traced back to the mid-1960s, marked by seminal technologies like Eliza [12], PARRY [13], and subsequent innovations such as Watson [14], ALICE [15], and GPT models [16]. Eliza's keyword identification approach was succeeded by PARRY's conceptual modeling of mental disorders. A.L.I.C.E. in 1995 marked a significant shift by utilizing heuristic pattern matching with XML-based rules. The progression continued with SmarterChild's [17] natural language processing in 2001, IBM Watson's semantic analysis, and further personal assistants like Siri [18], Google Assistant [19], Cortana [20], and Alexa [21]. These developments culminated with OpenAI's ChatGPT [22], grounded in the GPT-3.5 architecture, which harnessed vast datasets for more nuanced conversational capabilities. Figure 1 provides a timeline encapsulating these critical advancements.

All authors are with CeADAR - Ireland's Centre for Applied AI, University College Dublin, Dublin 4, Belfield, Ireland, Email: arsalan.shahid@ucd.ie

B. Development and Advancement of LLMs

Large Language Models (LLMs) are critical components of machine learning, enabling the prediction of a word or character sequence within a given language. This can be formalised as predicting the $N(i)$ -th word given the preceding sequence up to $N(i-1)$ as illustrated in Figure 2.

Historically, language models were constrained by limited training data, typically tailored for specific tasks like classification. The introduction of Generative Pre-trained Transformers (GPT) led to the evolution of Large Language Models (LLMs), expanding the modeling capabilities to encompass zero-shot [23] and few-shot learning [24]. LLMs have proven to be versatile, with applications in various Natural Language Processing (NLP) tasks such as text summarization, machine translation, sentiment analysis, named entity recognition, etc.

The transformation began with BERT [25], a pioneer in utilizing bidirectional encoding, simultaneously interpreting text from both directions. Following BERT, OpenAI's GPT-1 laid the groundwork for Generative AI but struggled with coherence. GPT-2 [26] marked significant advancements, introducing "Prompt Engineering" and achieving improvements in text generation. T5 [27] further expanded the field, leveraging an 11-billion-parameter model.

OpenAI's GPT-3 [28] became a landmark achievement, with 175 billion parameters and multilingual text generation. Google's PaLM [29], featuring 540 billion parameters, and Meta's open-source LLaMA [30] continued this trend. OpenAI's ChatGPT, fine-tuned from GPT-3.5, specialized in dialogue generation through Reinforcement Learning with Human Feedback (RLHF).

Recent innovations include GPT-4 [31], a multi-modal model capable of processing both text and images, although details remain scarce. The rapid advancements in LLMs have paved the way for more robust and versatile natural language understanding and generation, laying a solid foundation for the future of Conversational AI and related fields.

Despite the substantial advancements in LLMs and their growing prevalence in various applications, several inherent limitations and challenges remain. These include:

- **Hallucination:** LLMs may generate information that is coherent but factually incorrect or misaligned with the underlying data.
- **Trustworthiness:** Ensuring the reliability and integrity of generated content poses a complex challenge.
- **Explainability:** The vast number of parameters and complex structures in LLMs often hinder clear understanding and interpretation of the models' decision-making processes.
- **Untraining Ability:** Modifying or retracting specific knowledge in an LLM without retraining the entire model remains an open challenge.
- **Privacy and Ethical Awareness:** Addressing concerns related to data privacy and ethical considerations in the application of

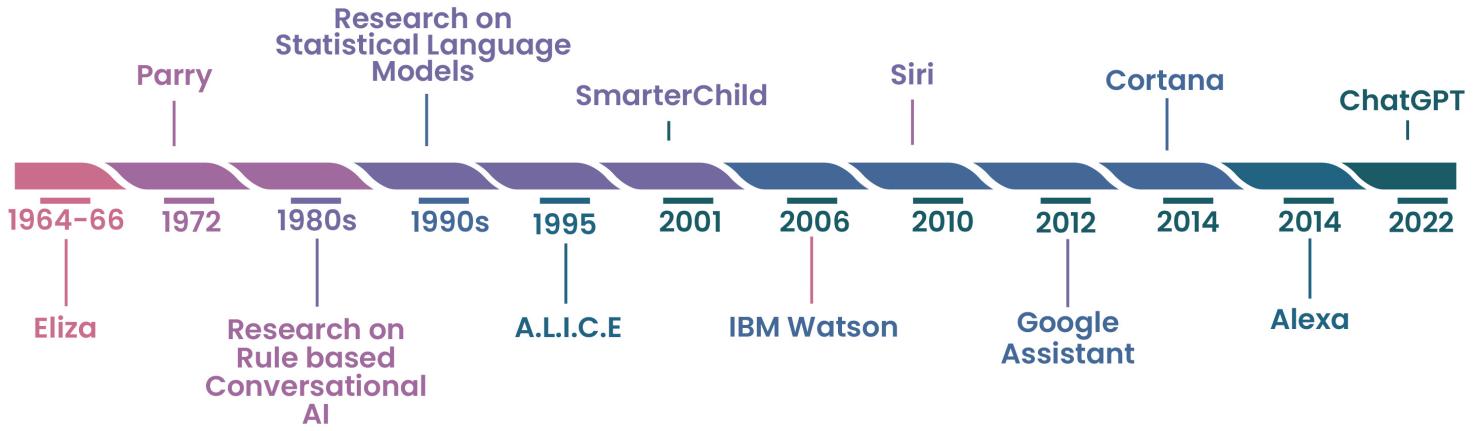


Fig. 1: Timeline summarising the key product developments in conversational systems

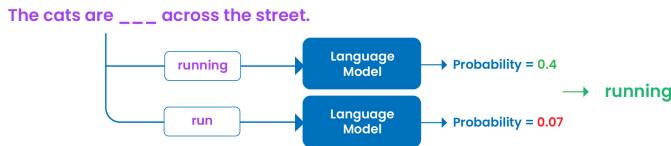


Fig. 2: Illustration of LLM's prediction based on previous inputs

LLMs require careful scrutiny and novel approaches.

These challenges motivate the need for innovative and comprehensive solutions. This paper proposes a novel architecture that combines the strengths of Large Language Models (LLMs), Knowledge Graphs (KG), and Role-Based Access Control (RBAC). By integrating LLMs' natural language understanding capabilities with KG's structured representation of real-world knowledge and RBAC's access control and privacy mechanisms, the proposed architecture aims to address the aforementioned limitations.

C. Main Contributions and Paper Organisation

The main contributions of this paper include:

- A Comprehensive Review of Large Language Models (LLMs):** We present a Large Language Model Explorer (LLMXplorer), an Excel-based tool that encapsulates over 150 key open and closed source LLMs. It provides a systematic overview of 16 key features, including parameters, training hardware, applications, industry relevance, country of origin, and more, offering insights into the state and evolution of the field.
- Applied Analysis of LLMs in Various Industries:** An extensive investigation into the practical use cases, limitations, and challenges in technology adoption of LLMs across different sectors. This analysis contributes to a better understanding of the real-world applications and potentials for trustworthy conversational AI.
- Functional Solution Architecture:** The development of a state-of-the-art, privacy-aware, explainable, and trustworthy conversational AI architecture. This design uniquely integrates LLMs, knowledge graph, and RBAC, and is empirically tested on the curated AI news dataset.

Section II delves into the methodologies and training processes tailored for Large Language Models (LLMs), shedding light on key aspects such as architecture, training methodologies, and accessibility. In Section III, an exhaustive review of prominent LLMs is provided and LLMXplorer is introduced. Sections IV and V discuss the practical implications, technological impacts, market trends, and industry-specific applications of LLMs. Section VI introduces a functional solution architecture integrating LLMs with knowledge graphs, aiming for a trustworthy conversational AI system. Section VII presents the discussions and future implications. Finally, Section VIII concludes the paper.

II. METHODS AND TRAINING PROCESS OF LLMs

Training a large language model is typically a multi-stage process. It consists of two main stages: pretraining and supervised fine-tuning. Recent advancements have introduced a third stage, i.e., Reinforcement Learning from Human Feedback (RLHF). This stage has been observed to substantially enhance effectiveness for certain applications [32]. Figure 3 depicts this tri-phase approach to LLM training.

A. Pretraining

During pretraining:

- A vast amount of unlabeled data is gathered from diverse online sources, which might encompass trillions of tokens from articles, web pages, and books.
- Preprocessing procedures, such as tokenization, lowercasing, and special character removal, are applied. Tokenization might operate at the word or subword level.
- The popular choice for LLM training is the transformer architecture. This phase is computationally demanding, often requiring robust GPUs or TPUs and incurring significant time and cost.

B. Supervised Fine-tuning

Post pretraining, LLMs are fine-tuned using labeled datasets. They build upon the patterns discerned from pretraining, and refining for specific tasks. Notably, fine-tuning requires less data and computational power than building models anew. It has exhibited

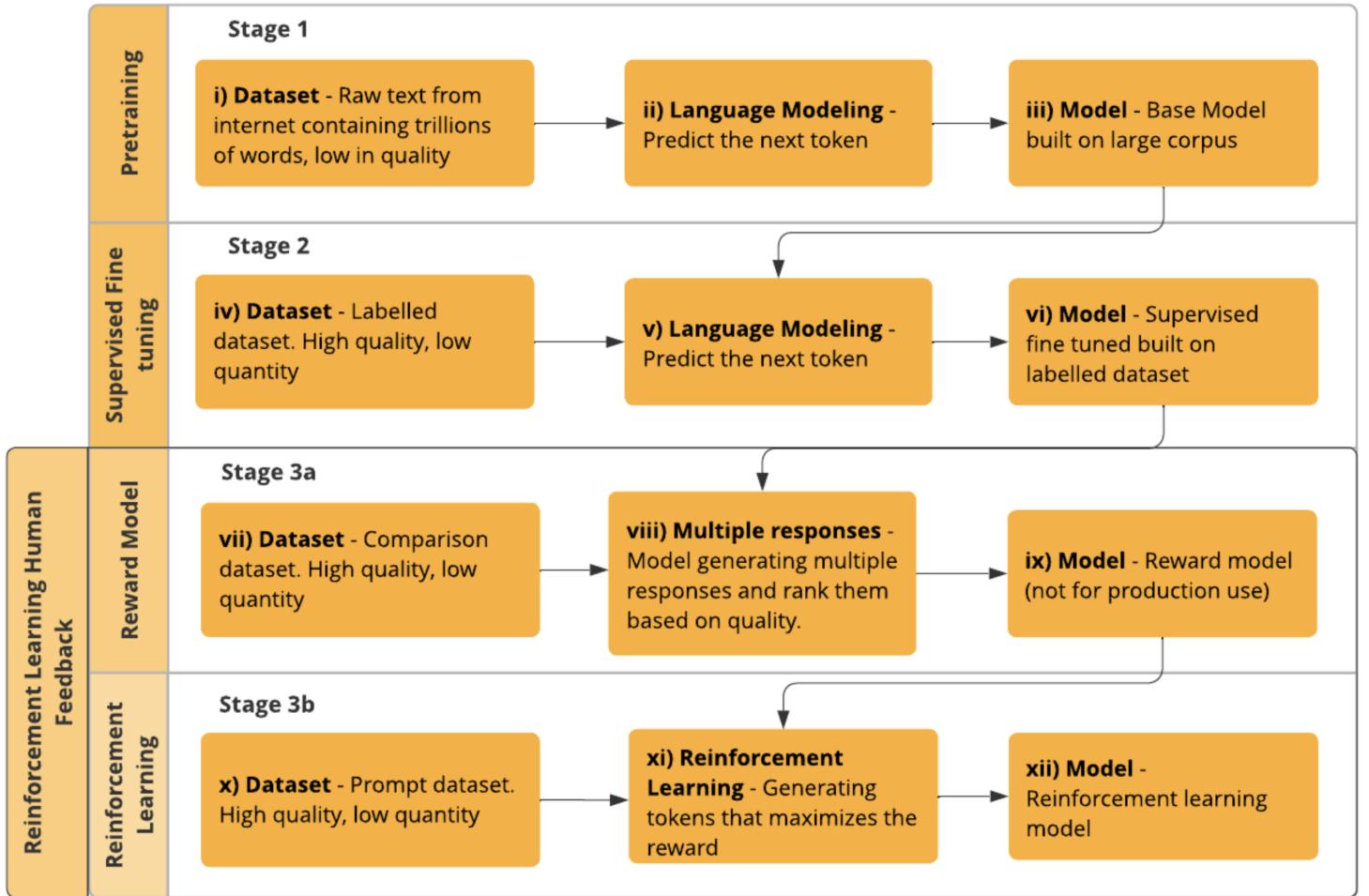


Fig. 3: Three-stage training process of Large Language Models (LLMs): Starting from an expansive pretraining on diverse data sources and utilizing the transformer architecture, transitioning into supervised fine-tuning with labeled datasets tailored for specific tasks, and culminating in dialogue optimization to refine AI-user interactions.

proficiency across a multitude of NLP tasks [33], [34], [35], owing to the marriage of broad linguistic knowledge from pretraining with task-focused adjustments.

C. Dialogue Optimization

The essence of dialogue optimization is refining the interaction between AI and users. The optimization has several facets:

- **Language Understanding:** Amplifying the AI's prowess in grasping user inputs, encompassing intent detection and semantic interpretation.
- **Response Generation:** Enhancing the model's ability to craft context-aware, fluent replies, which might be diverse or user-tailored.
- **Context Management:** Upholding a consistent conversational context, encompassing dialogue history tracking and coherence maintenance.
- **Error Management:** Implementing mechanisms to detect and rectify mistakes or to address ambiguous inputs.
- **User Feedback:** Integrating user feedback loops to finetune and uplift the overall user experience.

D. Reward Model

Subsequent to supervised fine-tuning, the model produces an array of possible replies for a prompt. Human evaluators rank these

based on quality. For further refinement, each response might be appended with a reward token. Although this process refines model outputs, it is labor-intensive and demands meticulous data curation.

E. Reinforcement Learning from Human Feedback (RLHF)

RLHF marries three steps: LLM pretraining, human preference-based reward model training, and reinforcement learning-guided LLM fine-tuning. Despite its efficacy, RLHF has its challenges, including the cost of obtaining human feedback and potential risks associated with producing detrimental or false content.

III. COMPREHENSIVE REVIEW OF STATE-OF-THE-ART LLMs

The October 2022 release of OpenAI's GPT3.5 [36] marked a significant milestone in the evolution of Large Language Models (LLMs). Subsequently, there was a noticeable uptick in the development of both open and closed-source models. Notably, a slew of open-source initiatives sought to replicate or surpass GPT3.5's capabilities. Renowned entities such as Google, Meta, Nvidia, Huawei, and Microsoft have unveiled LLM-based services, harnessing generative AI to augment their portfolios. The primary goal of these endeavours is the broadened accessibility of LLMs. This section gleans insights from an exhaustive review of recent literature on LLMs.

A. LLMXplorer: Large Language Model Explorer

With the proliferation of open-source LLMs, there are now over 16,000 distinct variations [37]. The Large Language Model Explorer (LLMXplorer)¹ [38] has been introduced as an Excel-based tool, providing insights on a concise catalogue of over 150 LLMs based on strict inclusion criteria such as public accessibility, detailed documentation, practical applications, recognizable releasing entities, and comprehensive training details. It classifies them based on capabilities, applications, industry sectors, and more. The LLMXplorer captures 17 key attributes for both open-source and proprietary models including:

- *Release Date*: Official launch date of the LLM.
- *Number of Parameters*: Indication of model complexity.
- *Base Model*: Foundational pre-trained model for the LLM.
- *Training Data Size*: Dataset volume used in training.
- *Training Hardware*: Infrastructure used in model training.
- *Training Time*: Duration of model training.
- *Context length*: Length of input text the model can process in one go.
- *Target Application*: Primary use-case of the LLM.
- *Model License Type*: Governing license of the LLM.
- *Modalities*: Supported data types (text, image, video).
- *Releasing Company*: Entity releasing the LLM.
- *Industry*: Targeted domain or sector of the LLM.
- *Origin Country*: Base country of the releasing entity.
- *Estimated User Base*: Approximate adoption rate.
- *Privacy Awareness*: Model's emphasis on data protection.
- *Ethical Awareness*: Model's ethical design and use considerations.
- *Reference Link*: Source link for more LLM details.

B. Insights from LLMXplorer

- Figures 5 and 6 detail LLM development timelines for open and closed-source categories, respectively. From Figure 4, in the past six years, open-source LLMs averaged 14 annual releases with a standard deviation of 15.4. The annual maximum was 34 models. For closed-source LLMs, over four years, the average is 15.5 releases per year with a standard deviation of 9.15. The range spans from 3 to 25 models annually. Both categories exhibit a rising trend in annual releases, signaling increased LLM development interests. Notably, closed-source models show a slower growth rate, possibly indicating varied community involvement or resource allocation.
- Figures 11 depict parameter distributions for open and closed-source models. Open-source LLMs range from 0.035 billion to 198 billion parameters, averaging 27.77 billion. The models CPM2, BLOOM, BLOOMZ, OPT, and BlenderBot3 exceed 175 billion parameters. Closed-source LLMs range from 0.06 billion to 1200 billion, averaging 182.94 billion. Notably, GLaM and PanGu-E surpass 1 trillion parameters.
- Figures 9 and 10 show LLM development distribution by country. For open-source models, the USA dominates, followed by the UK, Canada, and China. A similar distribution is observed for closed-source models.
- Figures 7 and 8 represent LLM development frequencies by leading companies. META dominates open-source development, whereas Google Research leads in closed-source LLMs.

¹LLMXplorer is freely available. Access requests can be sent via: <https://forms.gle/TNUbqHiCBsinD4Bu8>

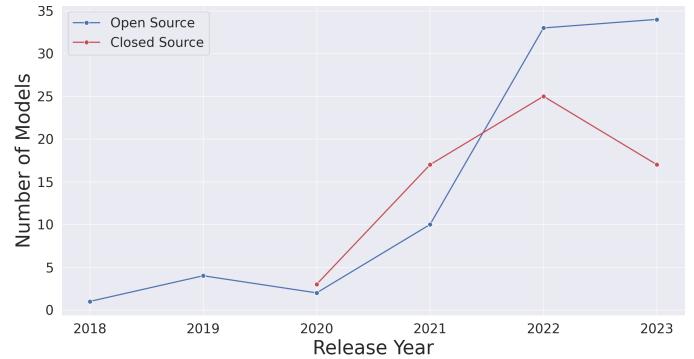


Fig. 4: Comparative analysis of annual releases: Open-source (n=89) vs. Closed-source (n=64) LLMs.

IV. APPLIED AND TECHNOLOGY IMPLICATIONS FOR LLMs

This section highlights the social, ethical, legal, privacy, regulatory, and physiological implications of LLMs.

A. Social and Ethical Implications

Democratization of Information: LLMs' open-source availability has broadened AI access, spurring innovation among diverse users [39]. Despite the slight quality lead of tech giants, open-source models present compelling alternatives, shifting the industry's landscape. Notably, personal LLMs, like LoRA [40], allow efficient internal data handling and quick model fine-tuning.

Bias and Fairness: Complexity in LLMs introduces misuse and bias risks. Addressing these necessitates diverse development teams and thorough bias audits to ensure fairness.

Digital Divide: Disparities in LLM access can widen societal digital divides. Addressing this requires promoting technology access and LLM education [41].

Responsible AI Development: Balancing LLM benefits against risks demands collaboration among varied stakeholders. Transparency, robust data governance, and educational initiatives can shape responsible LLM use.

Algorithmic Transparency: The "black box" nature of LLMs challenges trust and understanding. Research into LLM explainability and user-friendly output interpretations is essential.

Mitigating Harmful Content: Ensuring diverse training data and implementing watermarking can curb harmful LLM outputs. Oversight mechanisms, both algorithmic and human, can bolster responsible LLM deployment [42].

B. Legal, Privacy, and Regulatory Perspective

Data Protection and Privacy Laws: LLMs must adhere to regulations like GDPR while preventing unintentional data disclosure. Recent regulatory shifts, like the "EU AI ACT" [43] and proposed Chinese regulations, emphasize responsible generative AI use.

Intellectual Property Considerations: LLMs' capacity to generate human-like text challenges copyright norms. Addressing this requires clear intellectual property guidelines and plagiarism detection tools [44].

C. Physiological Perspectives

Human-Computer Interaction: LLM-based Conversational Agents carry potential risks, including undue trust, psychological vulnerabilities, and design biases. Promoting responsible design and prioritizing user privacy is paramount.

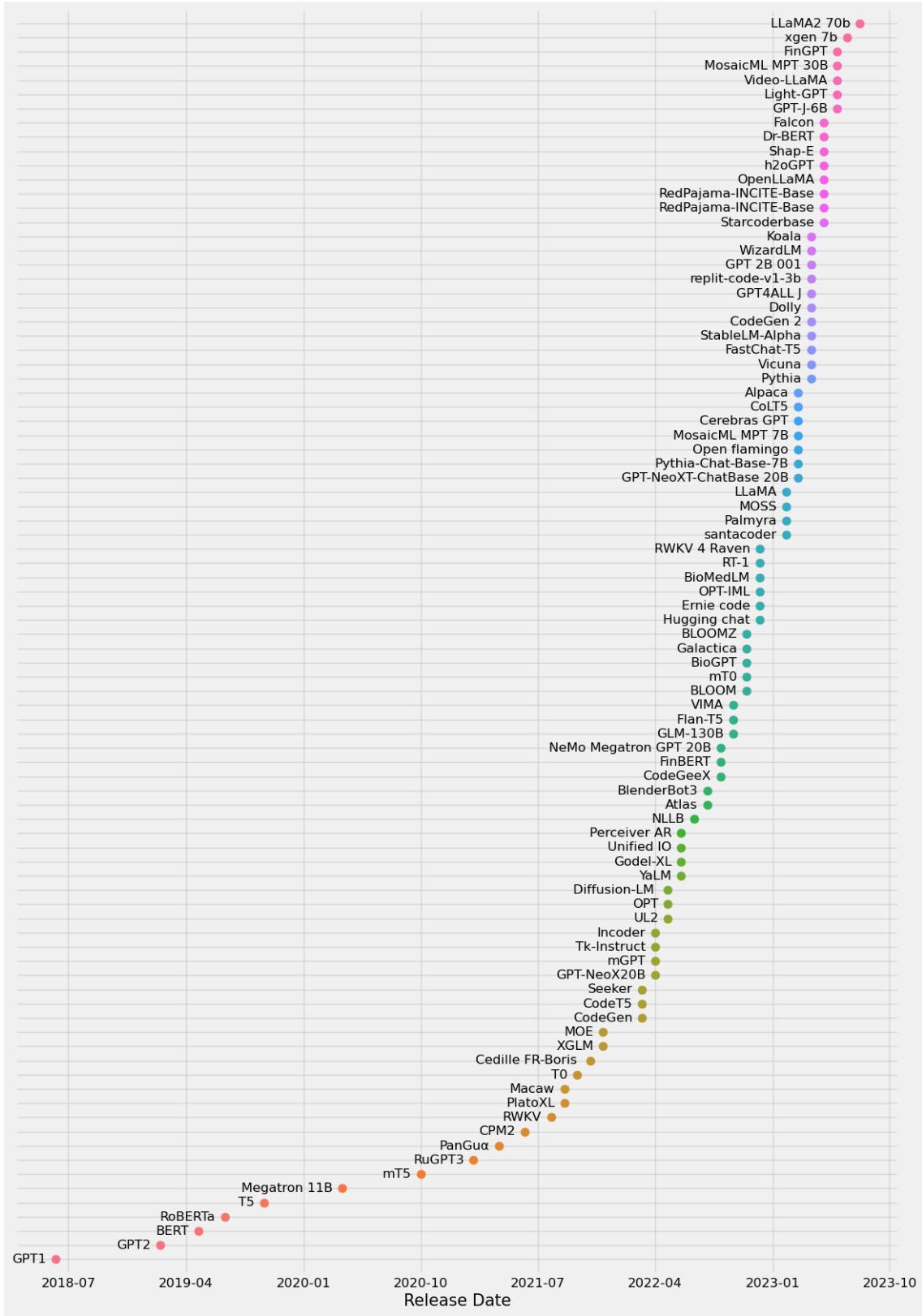
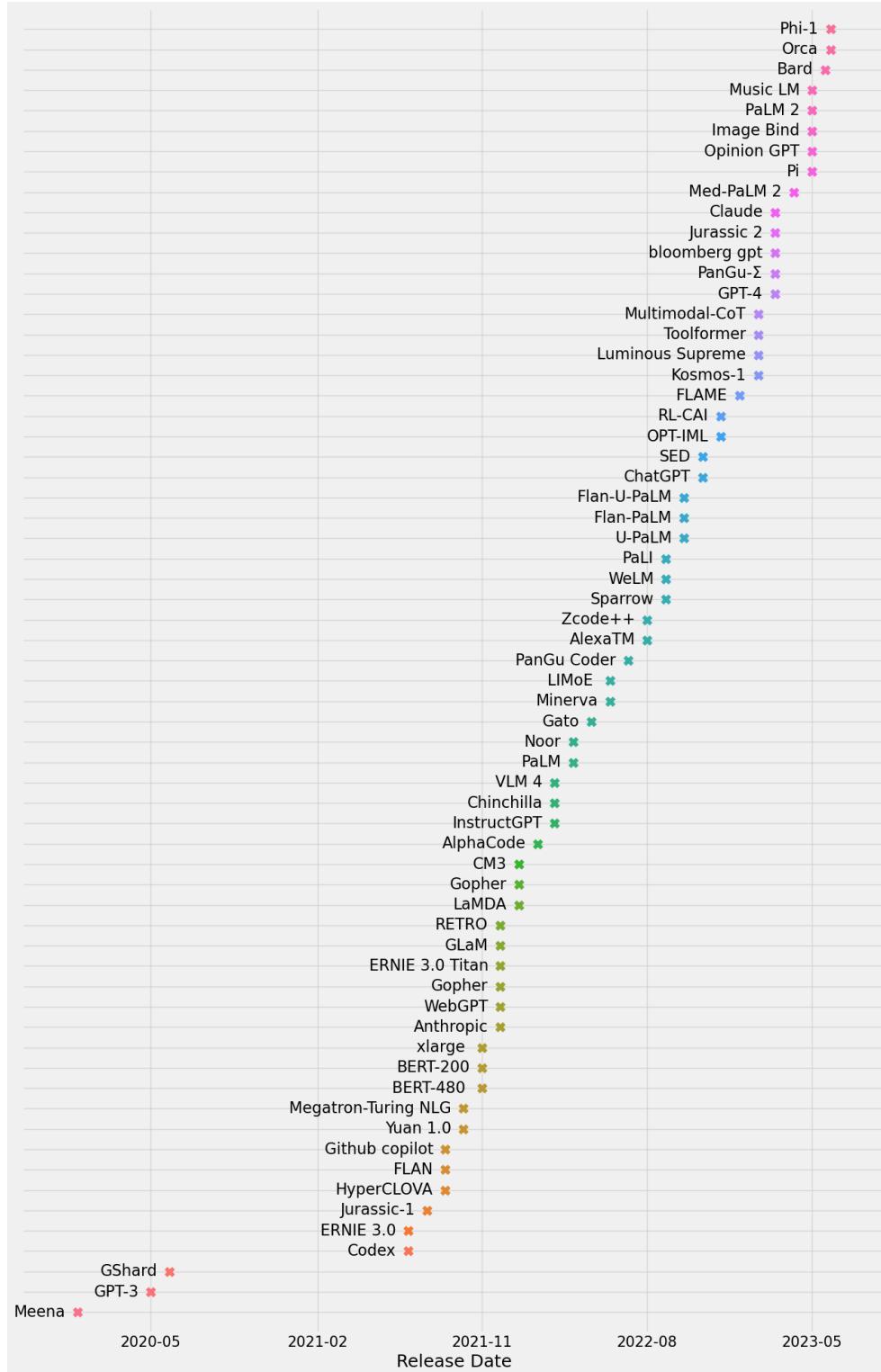


Fig. 5: Timeline of Open Source LLMs (n=89). The x-axis displays year-month, while the y-axis shows vertical stacks depicting the total number of models (in each stack) released within one month.

- Anthropomorphized systems risk user overreliance, potentially yielding unsafe outcomes or psychological harm. Emphasizing oversight and responsible use can counteract these threats.
- Users might inadvertently trust conversational agents with private information. Protecting against misuse demands robust safeguards and user awareness campaigns.
- Design choices in conversational agents can inadvertently perpetuate stereotypes. Addressing these requires design sensitivity and unbiased principles [45].



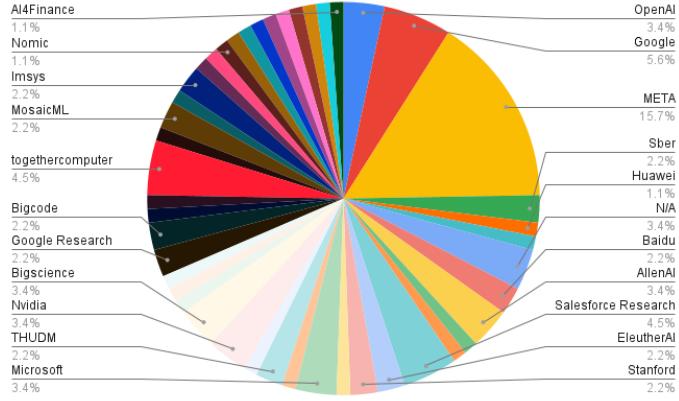


Fig. 7: Open Source LLMs developed by companies (n=89). The chart illustrates the percentage distribution of models developed by various companies.

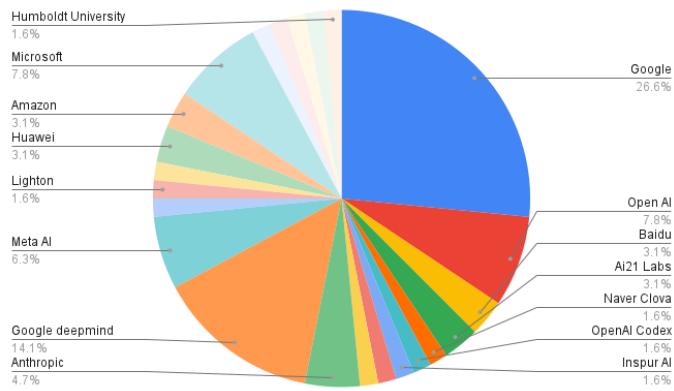


Fig. 8: Closed Source LLMs developed by companies (n=64). The chart illustrates the percentage distribution of models developed by various companies.

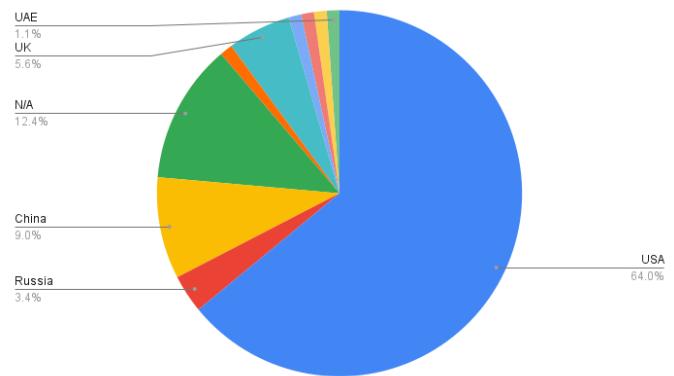


Fig. 9: Open Source LLMs categorized by the originating country of the developer company (n=89). The pie chart segments are labeled with the respective country names and their corresponding percentage contributions.

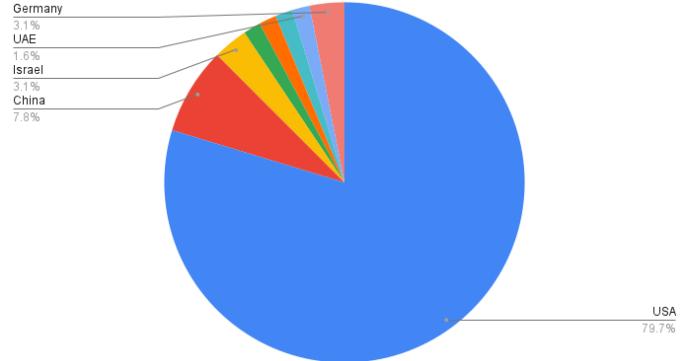


Fig. 10: Closed Source LLMs categorized by the originating country of the developer company (n=64). The pie chart segments are labeled with the respective country names and their corresponding percentage contributions.

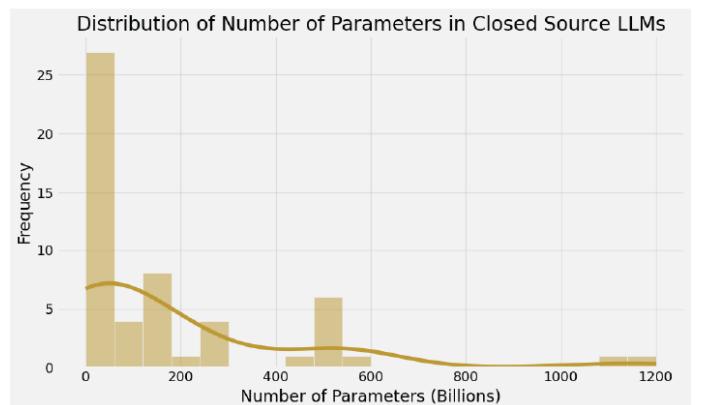
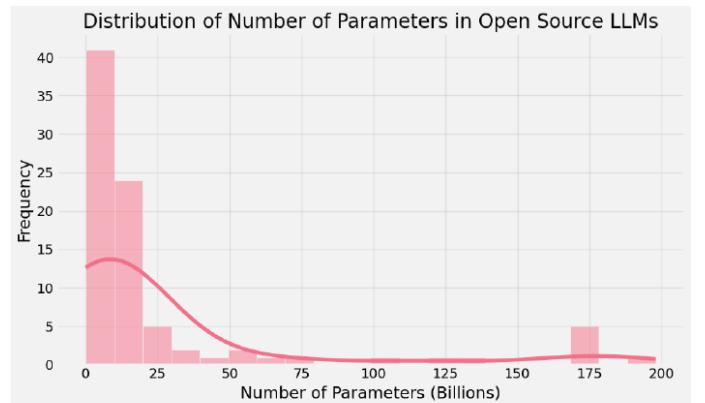


Fig. 11: Distribution of Number of Parameters (in billions) in Open and Closed Source LLMs (n=89 for Open Source, n=64 for Closed Source). The x-axis depicts the parameter count in billions (B), while the y-axis represents the frequency (i.e., the number of models within each histogram bin). The solid line (KDE) illustrates the kernel density estimate, offering a smoothed distribution shape insight.

- *Deepfakes Creation:* LLMs can produce sophisticated manipulated media, posing threats of deception and public manipulation.
- *Privacy and Data Concerns:* Handling vast training data introduces privacy vulnerabilities, potentially culminating in breaches and unintentional sensitive data exposure.
- *Bias Amplification:* LLMs might reflect and amplify biases present in training datasets, potentially reinforcing societal prejudices.
- *Malicious Usage:* LLMs can be tailored to produce harmful, offensive, or abusive content, potentially instigating hate or targeted assaults.
- *Unintended Outcomes:* Outputs from LLMs can occasionally have unexpected ramifications or ethical concerns, underscoring the need for careful oversight and guidelines.

TABLE I: Factors driving the LLM market growth

Factors	Description
Enterprise Adoption	Organizations utilize LLMs to optimize customer service, automate operations, and enhance efficiency. Sectors such as e-commerce, healthcare, and finance are primary beneficiaries.
AI Industry Growth	LLMs, central to AI initiatives, benefit from the holistic expansion of the artificial intelligence sector, accentuated by deep learning advancements.
Language-as-a-Service (LaaS) Emergence	LaaS platforms provide ready-to-use language models and APIs, enabling enterprises to harness language capabilities without intensive model training.
R&D Momentum	Continuous advancements in NLP and machine learning propel the market. Innovations in model design, training techniques, and performance optimization catalyze industrial interest.

Recognizing the challenges associated with LLMs, there exists a spectrum of opportunities for refining and evolving these models (Table II).

TABLE II: Opportunities in LLM Evolution

Opportunity	Description
Bias and Fairness	Develop techniques to detect and neutralize biases in LLMs, ensuring outputs that respect diverse perspectives.
Ethical Guidelines	Forge ethical standards and frameworks for the responsible creation and deployment of LLMs.
Explainability	Innovate methods to augment LLM transparency and interpretability, clarifying decision-making processes for users.
Contextual Accuracy	Enhance LLMs' contextual comprehension to produce more relevant and precise responses.
Collaborative Efforts	Foster collaborations amongst stakeholders, like researchers and policymakers, to address LLM-related societal and ethical implications.

C. Technology Applications and Use Cases in Diverse Industrial Sectors

At present, the broad application potential of LLMs remains partially untapped due to concerns regarding data sensitivity and model ethics. However, the evolution of privacy-aware LLM models could redefine numerous industrial landscapes. A selection of potential applications across various sectors is elucidated in Appendix A. These include finance and risk management (Table III), healthcare and diagnostics (Table IV), education and e-learning (Table V), customer service and support (Table VI), marketing and sales (Table VII), human resources and recruitment (Table VIII), legal and compliance (Table IX) and supply chain management (Table X). Notably, these listed applications represent a mere fraction of the extensive possibilities domain-specific LLMs can offer.

VI. SOLUTION ARCHITECTURE FOR PRIVACY-AWARE AND TRUSTWORTHY CONVERSATIONAL AI

With the advancement of conversational AI platforms powered by LLMs, there arises a need for addressing issues of explainability, data privacy, and ethical use of data to ensure their safe usage in industry and practical applications. LLMs, despite their significant

natural language understanding capabilities, often operate as black-box models, making their decisions difficult to interpret. Integrating Knowledge Graphs (KGs) [47] with LLMs offers a solution to these challenges by coupling the structured knowledge representation of KGs with the linguistic proficiency of LLMs. Furthermore, the incorporation of Role-Based Access Control (RBAC) into the architecture ensures that access to the AI system aligns with organizational policies and is granted only to authorized roles.

Advantages of Key Components:

- **LLMs:**

- Offer deep linguistic understanding, facilitating nuanced interactions.
- Capable of continuous adaptation based on user interactions.
- Provide a foundational layer for conversational platforms.

- **KGs:**

- Deliver structured and validated domain-specific knowledge.
- Enhance system explainability by tracing the origin of information.
- Complement LLMs by addressing gaps in domain-specific expertise.

- **RBAC:**

- Ensure data privacy through controlled access mechanisms.
- Align AI system usage with organizational policies and compliance mandates.
- Offer a methodical approach to managing data access, advancing security.

Several architectural paradigms can merge KG and LLM. As detailed by Shirui Pan et al. in [48], the primary techniques include KG-enhanced LLMs, LLM-augmented KGs, and a synergistic LLMs+KGs approach. Our proposition of a closed-loop architecture is in line with this synergistic design, underscoring the advantages of combining LLMs and KGs.

A. System Components and Design Choices

To elucidate the practical implementation of our architectural design, we anchor our exposition using media and journalism, specifically leveraging news data aggregated from the AI NewsHub platform (<https://ainewshub.ie>).

1) *AI NewsHub Dataset Description:* The AI NewsHub dataset is a systematically aggregated collection of news articles, procured daily from web crawling and search engines, and stored within a relational database system. Each record within this dataset represents an article and encapsulates the following attributes: identifier (ID), title, article content, published date, publisher name, and associated country of origin. Upon acquisition, articles undergo an analytical phase wherein they are classified based on their relevant topics, affiliated industry sectors, and originating publishers' categories.

2) *Construction of the AI NewsHub-based KG::* The structural organization of KG facilitates the extraction of nuanced insights, supports decision-making processes, and aids in the identification of concealed patterns within the data.

The KG derived from the AI NewsHub dataset comprises two primary node classes: articles and topics. The 'article' node encompasses attributes such as ID, content, sentiment analysis results, and a multi-dimensional matrix or vector array.

The former two attributes are directly extracted from the article dataset, while the sentiment is derived via the Vader sentiment

analysis tool[49], post preprocessing of the article content. The vector array is algorithmically generated leveraging the FastText library[50], contingent upon the content of the article.

Simultaneously, an 'article' node establishes a directed relationship with the 'topic' node when it is categorized under a specific topic. This 'topic' node bears the topic ID and the topic name, extracted from the designated topic table in the dataset. Storage and management of this KG are facilitated through Neo4j, with queries being structured using the Cypher language.

Neo4j and Cypher:

Neo4j [51], a premier graph database management platform, is adept at administering and querying intricate KGs. Given our emphasis on discerning the intricate interconnectedness of data, Neo4j is an optimal and open-source choice. In tandem with Cypher [52] — its dedicated query language — Neo4j empowers us to model multifaceted relationships, navigate patterns, and distil valuable insights from the KG. Cypher's design emphasizes human-readable syntax, ensuring that users can structure and interpret complex queries with relative ease.

3) *Llama-2 LLM*: In July 2023, Meta introduced the Llama-2 series [53], an advancement in its Large Language Model lineup. The models in this series are characterized by parameters ranging from 7 billion to 70 billion. When compared to the preceding Llama-1 models, Llama-2 variants have been trained on an expanded dataset with an increase in token count by 40% and an extended context length of 4,000 tokens. The 70B model incorporates the grouped-query attention mechanism [54], which is designed for efficient inference.

A specialized variant within the Llama-2 series is the Llama-2-Chat, fine-tuned for dialogue applications using Reinforcement Learning from Human Feedback (RLHF). Benchmark evaluations indicate that the Llama-2-Chat model offers improvements in areas such as user assistance and system security. Performance metrics suggest that this model's capabilities align closely with those of ChatGPT, as per human evaluators. Additionally, the 70B version of Llama-2 has been recognized on HuggingFace's OpenLLM leaderboard, exhibiting strong performance on several benchmarks, including but not limited to ARC [55], HellaSwag [56], MMLU [57], and TruthfulQA (MC) [58].

With the details of the three main components provided, we will now describe the proposed system architecture in the next section.

B. Architecture workflow

Figures 12 and 13 present the architectural workflow of our proposed system. The former offers a more general overview, elucidating the synergized connection between these three components. The latter is a detailed view of the intricate interactions between Knowledge Graphs (KG), Large Language Models (LLM), and Role-Based Access Control (RBAC).

The detailed sequence of operations in the system is described as follows:

- **Step 1:** The User (U) communicates a specific request to the RBAC Service (S).
- **Step 2:** The RBAC Service (S) evaluates the permissions associated with the user, forwarding the request to the Access Control (AC). AC determines the user's data access rights. Upon granting access, the process proceeds to *Step 3*. If denied, it transitions to *Step 2.2*.
- **Step 3:** The Prompt Analysis module refines and refactors the user's prompt (if necessary) and identifies the key capabilities required to put together an appropriate response. In our case

of journalism, the key capabilities include natural language understanding and generic output response or specialised capabilities such as similar article finder, sentiment analysis, fact-checking, and prediction for article topics and relevant industry sectors.

- **Step 4:** The Llama-2 LLM processes the user request based on the identified capabilities. If a generic response is required, the LLM responds to the user directly (*Step 4.1*). If specialised features from KG are required, the process moves to *Step 4.2*.
- **Step 5:** Llama-2 generates or invokes relevant Cypher instructions for Neo4j based on required capabilities.
- **Step 6:** A Cypher Validation Layer (CVL) ensures the integrity and safety of these instructions. Validated queries are executed on the Neo4j Knowledge Graph.
- **Step 7:** KG processes the queries, extracting the pertinent data. An Error Handling (EH) mechanism inspects the extracted insights for anomalies.
- **Step 8:** Error-free insights are compiled to be returned back to the LLM.
- **Step 9 and 10:** LLM formats the insights for user-friendly presentation and provides a response to the user. The User (U) receives the curated data, ensuring only permitted information is accessed. Users have the option to offer feedback through a Feedback Loop (FB), which might guide the LLM's subsequent interactions.

Based on the presented workflow, we now elucidate the decision-making process of the Language Learning Model (LLM) based on two primary scenarios: one where the LLM interacts with the Knowledge Graph (KG) to deduce answers and another where the LLM operates autonomously.

1) *Interactions between LLM and KG*: The LLM's reliance on KG is made manifest in the following application scenarios:

1. Similar Article Finder:

For implementing recommendation systems in media and journalism services, it is imperative to identify articles akin to a given one. Using the Cosine similarity metric, we ascertain the resemblance by evaluating the cosine value between the focal article and others. The topmost articles in terms of similarity scores are then selected, as represented in Algorithm 1.

Data: Graph database comprising articles

Result: Five articles most similar to Article 100

Procedure:

begin

Query Execution:

```

MATCH (a1:Article {article_id: 100}), (a2:Article)
WHERE a1 <> a2
WITH a1, a2, gds.similarity.cosine
(a1.content_vector,
a2.content_vector) AS
similarity_score
RETURN similarity_score,
a1.article_id, a2.article_id
ORDER BY similarity_score DESC LIMIT 5;
end

```

Algorithm 1: Identify Top 5 Articles resembling Article 100

2. Article Sentiment Analysis:

In situations necessitating sentiment extraction from a specific article or specific articles with the same set of sentiments, the

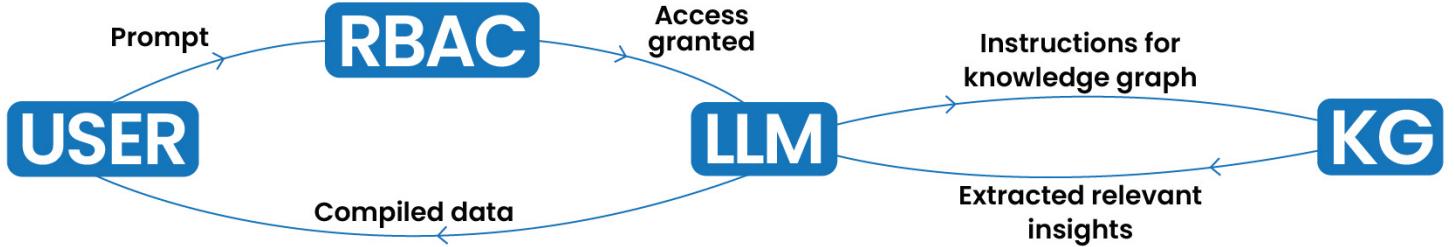


Fig. 12: Simplified functional workflow for combined Knowledge Graph and LLM

model retrieves the sentiment, pre-recorded in the article node within the KG. This retrieval is achieved by executing a MATCH Cypher query based on the article's unique identifier, as delineated in Algorithm 2.

Data: Graph database containing articles
Result: Sentiment associated with article_id 100

Initialization:
 Establish a connection to the graph database.

Procedure:
begin
Query Execution:
`MATCH (n:Article) WHERE n.article_id = 100
 RETURN n.sentiment`
Result Display:
 Display the fetched sentiment value.
end

Termination:
 Terminate the connection to the graph database.

Algorithm 2: Sentiment Extraction for Article with article_id 100

3. Article Topic Prediction:

Articles devoid of any designated topic can be assigned one based on the topics of closely resembling articles. This is realized by associating the given article with others and then linking the most similar one's topic. Algorithm 3 elucidates the related Cypher query.

2) *Autonomous LLM Operations:* The LLM, trained on vast datasets, can address certain user queries independently of external databases like the KG. This capability is evident in scenarios requiring comprehension, summarization, or responses that don't necessitate explicit fact-checking against a structured knowledge source.

1. Text Summarization:

For article summarization, the LLM utilizes its training to extract and condense key points without needing KG interaction.

2. Generic Journalism Queries:

LLM can handle queries related to journalistic standards, writing styles, or general media ethics, drawing from its extensive training on journalism-related topics.

3. Contextual Interpretations:

The LLM can interpret and respond to journalism-related queries based on previous context without extracting explicit facts, leveraging its inherent understanding of relational data.

VII. DISCUSSIONS

Conversational AI, while advancing rapidly, faces hurdles in balancing linguistic depth with accurate information representation.

Data: Graph database inclusive of articles

Result: Anticipated topic for article 100 by drawing from its nearest counterpart

Procedure:

```

begin
  Query Execution:
  MATCH (a1:Article {article_id: 100}), (a2:Article)
  WHERE a1 <> a2
  WITH a1, a2, gds.similarity.cosine(
  a1.content_vector,
  a2.content_vector) AS
  similarity_score
  WHERE similarity_score > 0.97
  RETURN a1.article_id, a2.article_id
  AS similar_article, t.name AS
  predicted_topic, similarity_score
  LIMIT 1
end
  
```

Algorithm 3: Topic Inference for Article 100 from its analogous article

As highlighted in our comprehensive review through the Large Language Model Explorer (LLMxplorer), which provides a systemic overview of numerous LLMs, linguistic proficiency often comes at the expense of transparency. Knowledge Graphs, while offering factual precision, might fall short in mimicking human conversational fluidity.

Our exhaustive applied analysis of the practical use cases, challenges, and limitations of LLMs across industries underscores their vast potential. Still, it also highlights the need for architectures that can bridge the aforementioned gaps. In response, our work introduces a functional solution architecture that uniquely integrates KGs and LLMs. This system not only stands out in its linguistic capabilities but also ensures factual consistency. With the incorporation of RBAC, we further the cause of data security, restricting users to role-specific information, and thereby fostering trust for real-world and industry-wide use cases.

The application domain of media and journalism, exemplified using rich data from the real-world product (AI NewsHub) platform, serves as both a case study and a validation point. It is a testament to the architecture's robustness and efficiency. Notably, the adaptability of the proposed architecture means it can seamlessly cater to a myriad of use cases and applications, emphasizing its cross-industry-wide relevance. Some pivotal aspects and design choices to underline the significance of the proposed architecture include:

- **Usage of Specialized LLM (Llama-2-Chat):** Our choice of the LLM variant, specifically fine-tuned for dialogue, caters to

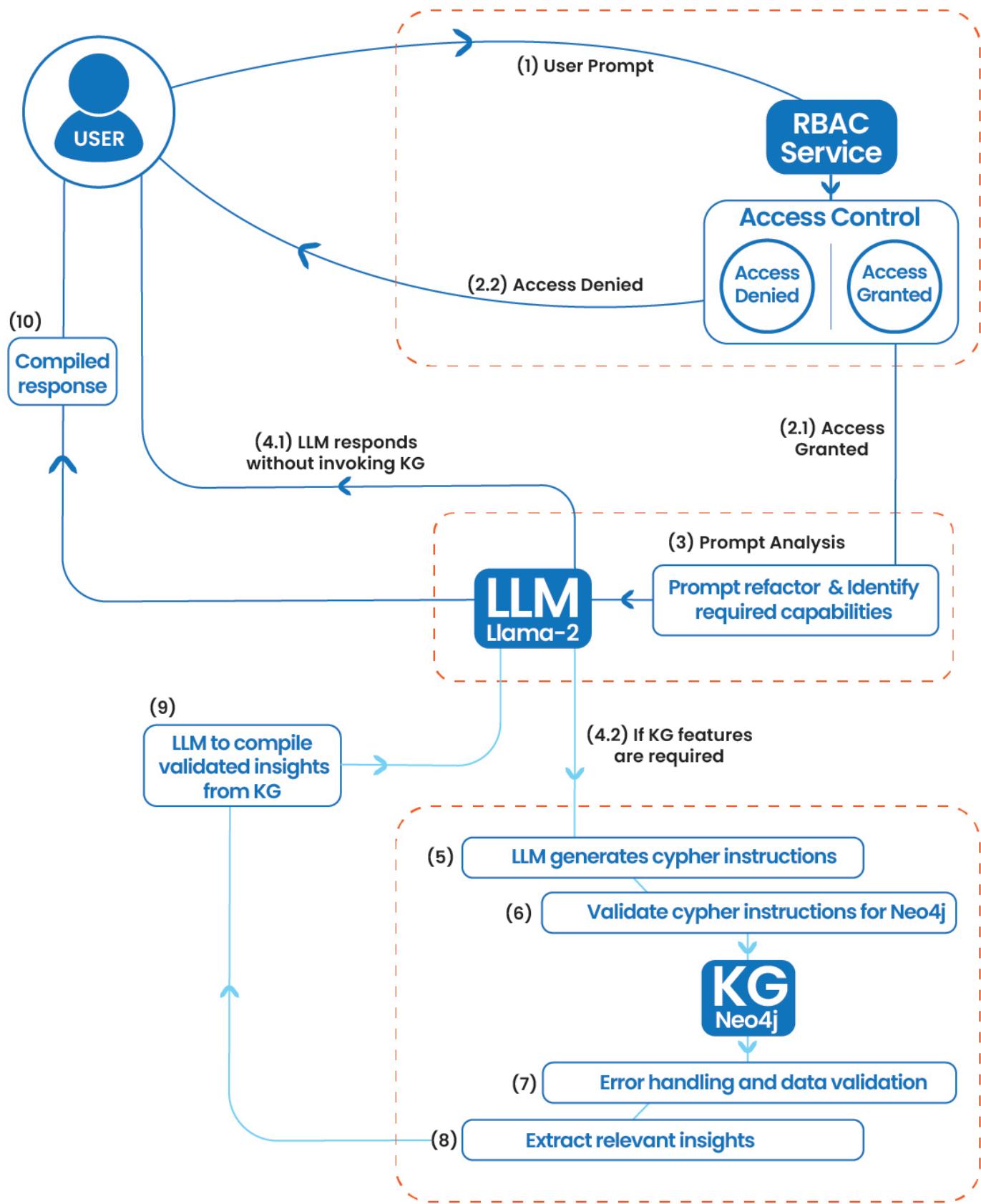


Fig. 13: Functional architecture illustrating the integration of Large Language Models (LLMs), Knowledge Graphs (KG), and Role-Based Access Control (RBAC).

enhanced conversational accuracy and quality.

- **Neo4j for Knowledge Graph Navigation:** The architecture uses Neo4j based on its open-sourceness, industry-wide adoption, proficiency in navigating knowledge structures and efficient data extraction via a query language, Cypher.
- **Cypher Validation Layer:** Located between the LLM and Neo4j, this layer validates Cypher queries from the LLM, mitigating unintended data accesses and system vulnerabilities.
- **Feedback Mechanism:** A feedback loop is integrated for iterative refinement of system responses and continuous LLM learning.
- **Algorithmic Approach:** The architecture employs specific algorithms for tasks such as article similarity detection, sentiment analysis, and topic prediction, ensuring transparency and maintainability.

A. Future Implications

The converging point of linguistic depth and factual accuracy, as epitomized by our system, is a promising indication of where conversational AI is headed. It sets a precedent for building architectures that are not only sophisticated in their response generation but also trustworthy in the information they convey. By addressing key challenges and providing tangible solutions, this work paves the way for future technologies that prioritize both efficiency and trustworthiness.

VIII. CONCLUSION

This research elucidates the advancements and challenges in conversational AI, emphasizing the imperative balance between linguistic richness and factual accuracy. By introducing a novel architecture that synergistically integrates Large Language Models and Knowledge Graphs, we provide a tangible solution to the current transparency and trust issues. Our comprehensive analysis of LLMs and their practical applications across various industries further contributes to the understanding and potential of trustworthy conversational AI. As we look ahead, this work not only addresses present challenges but also charts a promising trajectory for future conversational systems that prioritize efficiency, security, and user trust.

REFERENCES

- [1] Som S Biswas. Role of chat gpt in public health. *Annals of biomedical engineering*, 51(5):868–869, 2023.
- [2] Marco Cascella, Jonathan Montomoli, Valentina Bellini, and Elena Bignami. Evaluating the feasibility of chatgpt in healthcare: an analysis of multiple clinical and research scenarios. *Journal of Medical Systems*, 47(1):33, 2023.
- [3] Abdullah Bin Sawad, Bhuvan Narayan, Ahlam Alnefaie, Ashwaj Maqbool, Indra Mckie, Jemma Smith, Berkam Yuksel, Deepak Puthal, Mukesh Prasad, and A Baki Kocaballi. A systematic review on healthcare artificial intelligent conversational agents for chronic conditions. *Sensors*, 22(7):2625, 2022.
- [4] Simon Moore, Sandy Bulmer, and Jonathan Elms. The social significance of ai in retail on customer experience and shopping practices. *Journal of Retailing and Consumer Services*, 64:102755, 2022.
- [5] Chi Hong Leung and Winslet Ting Yan Chan. Retail chatbots: The challenges and opportunities of conversational commerce. *Journal of Digital & Social Media Marketing*, 8(1):68–84, 2020.
- [6] Seth C Lewis, Andrea L Guzman, and Thomas R Schmidt. Automation, journalism, and human-machine communication: Rethinking roles and relationships of humans and machines in news. *Digital journalism*, 7(4):409–427, 2019.
- [7] Andreas Veglis, Theodora A Maniou, et al. Chatbots on the rise: A new narrative in journalism. *Studies in Media and Communication*, 7(1):1–6, 2019.
- [8] Thomas Yue, David Au, Chi Chung Au, and Kwan Yuen Iu. Democratizing financial knowledge with chatgpt by openai: Unleashing the power of technology. Available at SSRN 4346152, 2023.
- [9] Xiao Ping Steven Zhang and David Kedmey. A budding romance: Finance and ai. *IEEE MultiMedia*, 25(4):79–83, 2018.
- [10] Stephen Atlas. Chatgpt for higher education and professional development: A guide to conversational ai. 2023.
- [11] Yogesh K Dwivedi, Nir Kshetri, Laurie Hughes, Emma Louise Slade, Anand Jeyaraj, Arpan Kumar Kar, Abdullah M Baabdullah, Alex Koohang, Vishnupriya Raghavan, Manju Ahuja, et al. “so what if chatgpt wrote it?” multidisciplinary perspectives on opportunities, challenges and implications of generative conversational ai for research, practice and policy. *International Journal of Information Management*, 71:102642, 2023.
- [12] Joseph Weizenbaum. Eliza—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1):36–45, 1966.
- [13] Kenneth Mark Colby. Modeling a paranoid mind. *Behavioral and Brain Sciences*, 4(4):515–534, 1981.
- [14] David Ferrucci, Eric Brown, Jennifer Chu-Carroll, James Fan, David Gondek, Aditya A Kalyanpur, Adam Lally, J William Murdock, Eric Nyberg, John Prager, et al. Building watson: An overview of the deepqa project. *AI magazine*, 31(3):59–79, 2010.
- [15] Richard S Wallace. *The anatomy of ALICE*. Springer, 2009.
- [16] Alec Radford, Karthik Narasimhan, Tim Salimans, and Ilya Sutskever. Improving language understanding with unsupervised learning. 2018.
- [17] Smarterchild was a chatbot available on aol instant messenger and windows live messenger. *Wikipedia*, Apr 2023. Page Version ID: 1151827590.
- [18] Siri is an easy way to make calls, send texts, use apps, and get things done with just your voice.: <https://www.apple.com/siri/>. *Apple*.
- [19] Meet your google assistant. ask it questions. tell it to do things.: <https://assistant.google.com/>.
- [20] Cortana helps you achieve more with less effort.: <https://www.microsoft.com/en-us/cortana>. *Cortana - Your personal productivity assistant*.
- [21] Amazon alexa: <https://alexa.amazon.com/>.
- [22] Open AI. <https://openai.com/chatgpt>. 2022.
- [23] Yongqin Xian, Christoph H Lampert, Bernt Schiele, and Zeynep Akata. Zero-shot learning—a comprehensive evaluation of the good, the bad and the ugly. *IEEE transactions on pattern analysis and machine intelligence*, 41(9):2251–2265, 2018.
- [24] Yaqing Wang, Quanming Yao, James T Kwok, and Lionel M Ni. Generalizing from a few examples: A survey on few-shot learning. *ACM computing surveys (csur)*, 53(3):1–34, 2020.
- [25] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.
- [26] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, et al. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8):9, 2019.
- [27] Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J Liu. Exploring the limits of transfer learning with a unified text-to-text transformer. *The Journal of Machine Learning Research*, 21(1):5485–5551, 2020.
- [28] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901, 2020.
- [29] Aakanksha Chowdhery, Sharan Narang, Jacob Devlin, Maarten Bosma, Gaurav Mishra, Adam Roberts, Paul Barham, Hyung Won Chung, Charles Sutton, Sebastian Gehrmann, et al. Palm: Scaling language modeling with pathways. *arXiv preprint arXiv:2204.02311*, 2022.
- [30] Hugo Touvron, Thibaut Lavigil, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*, 2023.
- [31] Open ai gpt-4: <https://openai.com/gpt-4>.
- [32] Yuntao Bai, Andy Jones, Kamal Ndousse, Amanda Askell, Anna Chen, Nova DasSarma, Dawn Drain, Stanislav Fort, Deep Ganguli, Tom Henighan, et al. Training a helpful and harmless assistant with reinforcement learning from human feedback. *arXiv preprint arXiv:2204.05862*, 2022.
- [33] Matthew E. Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. Deep contextualized word representations. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 2227–2237, New Orleans, Louisiana, June 2018. Association for Computational Linguistics.
- [34] Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research*, 21(140):1–67, 2020.
- [35] Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*, 2019.

- [36] Open AI. <https://platform.openai.com/docs/models/gpt-3-5>. 2022.
- [37] Llm explorer: Explore 18,000+ large language models and compare their parameters, url=<https://llm.extractum.io/>.
- [38] Ahtsham Zafar. Llmxplorer: Large language model explorer. *CeADAR Connect Group*, July 2023.
- [39] Guido Maciocci. Google is afraid of lamas: How open source is democratizing large language models, May 2023.
- [40] Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. Lora: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685*, 2021.
- [41] Matthias C. Rillig, Marlene Ågerstrand, Mohan Bi, Kenneth A. Gould, and Uli Sauerland. Risks and benefits of large language models for the environment. *Environmental Science and Technology*, 57(9):3464–3466, Mar 2023.
- [42] Azamat Abdoullaev. The dark side of large language models: Mitigating the risks of ai text generation.
- [43] The artificial intelligence act: <https://artificialintelligenceact.eu/>. *The Artificial Intelligence Act*, Sep 2021.
- [44] Intellectual property considerations: <https://www.linkedin.com/pulse/unleashing-power-large-language-models-omer-ali-phd>.
- [45] Laura Weidinger, John Mellor, Maribeth Rauh, Conor Griffin, Jonathan Uesato, Po-Sen Huang, Myra Cheng, Mia Glaese, Borja Balle, Atoosa Kasirzadeh, et al. Ethical and social risks of harm from language models. *arXiv preprint arXiv:2112.04359*, 2021.
- [46] Straits Research. Natural language processing market size is projected to reach usd 91 billion by 2030: Straits research, Aug 2022.
- [47] Dieter Fensel, Ümutcan Şimşek, Kevin Angele, Elwin Huaman, Elias Kärle, Oleksandra Panasiuk, Ioan Toma, Jürgen Umbrich, and Alexander Wahler. *Introduction: What Is a Knowledge Graph?*, pages 1–10. Springer International Publishing, Cham, 2020.
- [48] Shirui Pan, Linhao Luo, Yufei Wang, Chen Chen, Jiapu Wang, and Xindong Wu. Unifying large language models and knowledge graphs: A roadmap. *arXiv preprint arXiv:2306.08302*, 2023.
- [49] C. Hutto and Eric Gilbert. Vader: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1):216–225, May 2014.
- [50] Armand Joulin, Edouard Grave, Piotr Bojanowski, and Tomas Mikolov. Bag of tricks for efficient text classification. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, pages 427–431. Association for Computational Linguistics, April 2017.
- [51] Neo4j Graph Database & Analytics – The Leader in Graph Databases.
- [52] Cypher Query Language - Developer Guides.
- [53] Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, et al. Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv:2307.09288*, 2023.
- [54] Joshua Ainslie, James Lee-Thorp, Michiel de Jong, Yury Zemlyanskiy, Federico Lebrón, and Sumit Sanghai. Gqa: Training generalized multi-query transformer models from multi-head checkpoints. *arXiv preprint arXiv:2305.13245*, 2023.
- [55] Arseny Moskvichev, Victor Vikram Odonard, and Melanie Mitchell. The conceptarc benchmark: Evaluating understanding and generalization in the arc domain. *arXiv preprint arXiv:2305.07141*, 2023.
- [56] Rowan Zellers, Ari Holtzman, Yonatan Bisk, Ali Farhadi, and Yejin Choi. Hellaswag: Can a machine really finish your sentence? *arXiv preprint arXiv:1905.07830*, 2019.
- [57] Dan Hendrycks, Collin Burns, Steven Basart, Andy Zou, Mantas Mazeika, Dawn Song, and Jacob Steinhardt. Measuring massive multitask language understanding. *arXiv preprint arXiv:2009.03300*, 2020.
- [58] Stephanie Lin, Jacob Hilton, and Owain Evans. Truthfulqa: Measuring how models mimic human falsehoods. *arXiv preprint arXiv:2109.07958*, 2021.

APPENDIX A INDUSTRY-WIDE LLM USECASES

TABLE III: Applications of LLMs in Finance and Risk Management

Personalized Financial Advising:	Enhancing customer interactions by offering bespoke advice and assistance.
Intelligent Support Chatbots:	Addressing customer queries and assisting in real-time.
NLP in Financial Analysis:	Parsing unstructured financial data for insights.
Market Insights and Sentiment Analysis:	Gleaning consumer behaviour and market trends from diverse sources.
Risk Assessment and Portfolio Optimization:	Informing financial decisions through historical data analysis and prediction.
Regulatory Adherence:	Assisting institutions in maintaining compliance by analyzing regulatory documentation.
Fraud Prevention:	Identifying anomalies in transactions to detect potential fraud.
Real-time Market Analytics:	Providing current market sentiments and predictions.

TABLE IV: Applications of LLMs in Healthcare

Medical Data Synthesis:	Extracting insights from diverse medical datasets for informed decisions.
Clinical Decision Assistance:	Offering patient-centric insights for diagnostic and therapeutic interventions.
Medical Imaging Insights:	Detecting anomalies in radiological images.
NLP in Medical Documentation:	Assisting with transcription, language translation, and information retrieval.
Predictive Health Analytics:	Identifying disease risks and facilitating early interventions.
Drug Discovery Support:	Assisting in potential drug target identification and efficacy prediction.
Patient Education and Virtual Assistance:	Personalizing medical education and support for patients.
Real-time Patient Monitoring:	Providing alerts based on patient data analysis.

TABLE V: Applications of LLMs in Education and e-Learning

Customized Learning Paths:	Adapting content to cater to unique student requirements.
Intelligent Tutoring Systems:	Offering interactive guidance and adaptive feedback.
Content Generation:	Creating and adapting educational resources.
Language Learning Support:	Facilitating exercises, translation, and pronunciation guides.
Automated Evaluation:	Grading assignments with detailed feedback.
Dynamic Assessments:	Generating assessments tailored to student performance.
Resource Recommendations:	Suggesting pertinent educational materials.
Gamified Learning Experiences:	Creating engaging, interactive scenarios and simulations.

TABLE VI: Applications of LLMs in Customer Service and Support

Personalised Recommendations and Upselling: Analysing customer data to provide personalized product recommendations, suggest complementary items, and support upselling and cross-selling efforts.
Sentiment Analysis and Customer Insights: Analysing customer feedback and social media conversations to extract sentiment, provide actionable insights, and enhance products, services, and customer experience.
Self-Service Support and Knowledge Base: Assisting in creating and maintaining self-service support systems and knowledge bases, generating FAQs, troubleshooting, and offering step-by-step instructions for customers.
Voice and Speech Recognition: Contributing to voice and speech recognition technologies, accurately transcribing and interpreting customer voice inputs, enabling voice-controlled self-service support and interactions.
Call Center Support: Aiding call centre agents by providing real-time suggestions, relevant information, and access to knowledge bases during customer calls, enhancing efficiency and customer satisfaction.
Language Translation and Multilingual Support: Facilitating language translation in customer service interactions, overcoming language barriers, and supporting multilingual customer service.
Email and Ticket Management: Assisting in managing customer support emails and tickets, categorizing inquiries, prioritizing urgent cases, and suggesting appropriate responses.
Customer Feedback Analysis: Analyzing customer feedback surveys, reviews, and ratings to extract insights, understand preferences, and make data-driven decisions for product or service enhancements.
Complaint Resolution and Escalation: Assisting in complaint resolution by providing guidelines, suggested responses, or escalation procedures to customer service agents, ensuring consistent and appropriate handling.

TABLE VII: Applications of LLMs in Marketing and Sales

Content Generation: Assisting in generating marketing content, suggesting ideas, and automating parts of the content creation process.
Personalised Marketing Campaigns: Analysing customer data to create personalised marketing campaigns and tailored messages based on demographics and preferences.
Market Research and Trend Analysis: Assisting in market research by analysing industry reports, customer reviews, and social media conversations to identify trends, preferences, and competitive insights.
Customer Profiling and Segmentation: Analysing customer data to generate detailed customer profiles and segment customers based on various attributes.
Social Media Listening and Engagement: Monitoring social media platforms, tracking brand mentions, engaging with customers, and responding to inquiries or comments.
Lead Generation and Qualification: Analysing customer data to identify potential leads, assess lead quality, predict purchase intent, and prioritise leads for follow-up.
Sales Support and Product Information: Providing sales teams with real-time product information, pricing details, and assistance in answering customer queries and objections.
Customer Journey Mapping and Optimization: Analysing customer interactions, feedback, and behavioural data to map the customer journey, identify pain points, and optimise the customer experience.
Marketing Automation and Personalization: Powering marketing automation platforms, segmenting audiences, crafting personalised campaigns, and delivering dynamic content based on user preferences.

TABLE VIII: Applications of LLMs in Human Resources and Recruitment

Resume Screening and Candidate Evaluation: Analysing and evaluating candidate qualifications, skills, and experience against job requirements, shortlisting candidates based on predefined criteria.
Job Description Generation: Generating comprehensive job descriptions based on industry trends, suggesting appropriate job titles, responsibilities, and qualifications.
Candidate Sourcing and Talent Pooling: Searching through candidate data to identify potential candidates based on specific skills, experience, or qualifications, assisting in building talent pools.
Interview Assistance: Providing interview guidance, suggesting questions, evaluation criteria, and offering insights into candidate background information.
Employee Onboarding and Orientation: Supporting employee onboarding by providing resources, and policies, and generating onboarding materials such as welcome emails and orientation guides.
HR Policy and Compliance Guidance: Assisting with HR policy guidance, answering HR-related questions, interpreting employment laws, and ensuring compliance with regulations and standards.
Employee Engagement and Performance Management: Aiding in employee engagement and performance management by generating surveys, collecting feedback, and providing insights for improvement.
HR Knowledge Base and Self-Service Support: Contributing to HR knowledge bases, generating responses to HR questions, and offering guidance on policies, procedures, benefits, and programs.
Employee Performance Analysis: Analysing employee performance data, identifying patterns, strengths, areas for improvement, and development opportunities, aiding in performance evaluations.
Employee Satisfaction Surveys: Generating employee satisfaction surveys, analysing collected data, and providing insights to improve employee engagement, satisfaction, and retention.
Employee Training and Development: Contributing to employee training initiatives, generating training materials, courses, and resources to enhance effectiveness and scalability.
Diversity and Inclusion Initiatives: Supporting diversity and inclusion efforts by analysing employee demographics, suggesting strategies, and generating diversity reports.
Employee Exit Interviews and Feedback Analysis: Automating exit interviews, analysing employee feedback, identifying areas of improvement, and providing insights for addressing concerns and retention.
HR Analytics and Predictive Modelling: Analysing HR data to derive insights, predict trends, and assist in workforce planning, succession planning, and talent management strategies.
Employee Benefits and Wellness Programs: Providing information and guidance on employee benefits, wellness initiatives, and resources to promote employee well-being.
HR Compliance and Ethics: Assisting with HR compliance, providing information on employment laws, anti-discrimination policies, and ethical guidelines.

TABLE IX: Applications of LLMs in Legal and Compliance

Legal Research and Case Analysis: Assisting in legal research, extracting insights from legal documents, case law, statutes, and regulatory materials, and supporting attorneys in case analysis and preparation.
Contract Analysis and Due Diligence: Reviewing contracts, extracting key terms, conditions, obligations, and risks, aiding in due diligence processes, and identifying potential legal issues.
Compliance Monitoring and Risk Assessment: Monitoring for regulatory compliance, risk assessment, and suggesting mitigative actions based on analysis of internal processes and external regulations.
Legal Document Generation: Assisting in the generation of legal documents, contracts, agreements, and other legal materials by suggesting standardized language and clauses.
E-Discovery and Document Review: Aiding in e-discovery processes, categorizing, and tagging legal documents, identifying relevant information, and optimizing the document review process.
Legal Knowledge Management: Assisting in organizing and accessing legal knowledge resources, suggesting appropriate references, and facilitating the retrieval of information.
Legal Compliance Training and Support: Generating training materials for legal compliance, offering guidance on regulatory requirements, and assisting in creating compliance programs.
Legal Language Translation and Interpretation: Assisting in translating legal documents and terminology across languages, supporting multilingual legal practices, and ensuring accurate interpretation.
Legal Chatbots and Virtual Assistants: Providing general legal guidance, answering common legal queries, and offering initial consultation through virtual platforms and chatbots.
Dispute Resolution and Mediation Support: Assisting legal professionals in dispute resolution and mediation processes, suggesting strategies, providing relevant case law, and offering insights.

TABLE X: Applications of LLMs in Supply Chain and Manufacturing

Supply Chain Optimization: Assisting in supply chain optimization efforts, suggesting strategies, analysing logistics and distribution networks, and providing insights for streamlined operations.
Quality Control and Defect Detection: Assisting in quality control processes by detecting anomalies in production data, identifying potential product defects, and suggesting corrective actions.
Predictive Maintenance: Analysing equipment and machinery data, predicting potential failures, suggesting maintenance schedules, and optimizing equipment uptime.
Demand Forecasting and Inventory Management: Analysing sales data, historical demand, and other external factors to predict product demand, aiding in inventory optimization, and reducing stock-outs or overstock situations.
Supplier Relationship Management: Evaluating supplier performance, identifying risks or inefficiencies, and providing insights for better supplier collaboration and relationship management.
Product Design and Development: Offering insights for product design based on market trends, customer feedback, and competitive analysis, aiding in product innovation and development.
Sustainability and Environmental Impact: Analysing supply chain data to suggest environmental sustainability initiatives, reduce carbon footprint, and align operations with sustainability goals.
Supply Chain Risk Management: Identifying potential risks in the supply chain, suggesting risk-mitigative strategies, and ensuring consistent supply chain operations.
Production Planning and Scheduling: Assisting in production planning, suggesting optimized schedules based on demand forecasts, resource availability, and operational constraints.
Lean Manufacturing and Process Improvement: Analysing manufacturing processes, identifying inefficiencies, suggesting lean strategies, and providing insights for process improvement and optimization.