A Review of Methods for Reducing Hallucinations in Generative Artificial Intelligence to Enhance Knowledge Economy

Zahra Roozbahani[1*]

Article Type:
Research Article

Zahra Roozbahani
Corresponding Author, Assistant Professor, Department of Computer Engineering, National University of Skills (NUS), Tehran, Iran.
E-mail: zroozbahani@nus.ac.ir

ABSTRACT

Generative Artificial Intelligence (AI) models, such as large language models, are transforming various sectors of the knowledge economy, including education, research, development, and data analysis, due to their ability to generate new contents. However, hallucination, which refers to the generation of content that appears plausible but lacks scientific basis, presents a significant challenge to the safe adoption of this technology in critical applications such as financial analysis and market prediction. This study aims to explore the effects of hallucinations on productivity of the knowledge economy and proposes some approaches to mitigate them. Through conducting a systematic literature review and qualitative analysis, different types of hallucinations in generative AI models are identified, and their effects on trust and productivity in knowledge-based systems are examined. The review of hallucination reduction methods indicates that the approaches utilizing reinforcement learning with human feedback enhance the reliability of the generated content by correcting errors in the model's output through repeated adjustments. Finally, this study presents a hybrid approach to hallucination reduction in knowledge economy. This approach is based on three techniques of prompt engineering, retrieval-based generation, and self-improvement through feedback and reasoning. The results provide a foundation for future research on managing AI risks and enhancing the productivity of the knowledge economy.

KEYWORDS

Generative AI, Hallucination, Knowledge Economy, Large Language Models, Trust.

Spring & Summer (2025) 2(1): 21-34

Received: 15 January 2025
Revised: 17 February 2025
Accepted: 6 April 2025
Available Online: 7 April 2025

Cite this article: Roozbahani, Z., (2025). A Review of Methods for Reducing Hallucinations in Generative Artificial Intelligence to Enhance Knowledge Economy. *Journal of Knowledge Economy Studies (JKES)*, 2(1), 21-34. DOI: http//doi.org/10.22034/kes.2025.2049560.1042

Publisher: Hazrat-e Masoumeh University.
Online ISSN: 3060-7329

Introduction

Large Language Models (LLMs) such as GPT-3 (Brown et al., 2020) and ChatGPT have demonstrated remarkable capabilities in generating natural language texts. The ability of these models to provide fluent and comprehensive responses across a wide range of topics has surpassed earlier chatbots in terms of efficiency and performance. Generative AI, which includes deep learning models such as GANs and transformers, has made significant strides in recent years, particularly in the field of knowledge economy. The knowledge economy focuses on production, distribution, and use of knowledge in society. In this section, we highlight some of the impacts of generative AI on the knowledge economy. Generative AI can contribute to enhancing the quality of knowledge. These models can analyze the existing data and identify errors within it, improving the accuracy and credibility of knowledge. Furthermore, generative AI can foster collaboration in knowledge production by enabling people across the globe to work together and generate new knowledge.

Despite the remarkable capabilities of generative AI, these models have a significant flaw known as 'hallucination', which means they produce information that seems logical but is, in reality, inaccurate or meaningless (Gunjal et al., 2023). The encoding of knowledge in large language models is incomplete and the generalization of knowledge may lead to 'memory distortion'. This can result in the generation of incorrect or "hallucinated" information by these models. Hallucination arises when advanced models like GPT-4 and others may produce entirely inaccurate or baseless outputs (Ray, 2023). This problem occurs due to pattern some generation techniques during the training phase and the lack of real-time Internet updates, leading to inconsistencies in the output information (Ray, 2023).

Hallucination, or the 'fabrication of information', is a common failure mode for large language models in which they generate incorrect or meaningless content. In other words, AI hallucinations are also referred to as "self-made" outputs. AI cannot independently verify whether the text it generates is factually accurate or trustworthy. Hallucinations can lead to the spread of misinformation and in applications requiring critical decision-making, may cause harmful consequences or lead to a loss of trust in AI systems.

Generative AI can have various effects on the knowledge economy. These models can assist in the generation of new knowledge, improve the quality of knowledge, increase access to knowledge, and enhance collaboration in knowledge production. However, hallucination in AI presents multiple challenges, including the issues of truth detection, trust, security, and ethics. Therefore, responsible use of this technology is necessary and its challenges must be seriously addressed.

In this context, identifying hallucinations has become a significant concern and researchers have been working to identify the factors contributing to hallucination to propose effective mitigation methods. For instance, researchers have introduced the mFACT method as a tool for detecting hallucinations in summaries, expanding its applicability from English to other languages (Qiu et al., 2023). Furthermore, a framework

for identifying hallucinations based on contextual information has been proposed (Li et al., 2023). On the other hand, some researchers offer a different perspective on the causes of hallucinations, investigating contradictions as one of the factors contributing to this issue (Mündler et al., 2023). In terms of reducing hallucination, researchers have proposed effective methods. For example, some have explored retrieval-based generation approaches to address the key challenges of accuracy and timeliness in LLM outputs (Kang et al., 2023), while others have proposed prompt engineering techniques to reduce hallucinations (Farquhar et al., 2024; Peng et al., 2023).

This study aims to examine both the positive and negative effects of generative AI on the knowledge economy and identify the challenges associated with its hallucinations. The central question is how we can harness the capabilities of generative AI while mitigating the negative impacts of its hallucinations. The main questions of this research study are:

1. What effects does the hallucination of Artificial Intelligence have on the knowledge economy?
2. What methods are there to reduce the hallucination of Artificial Intelligence?
3. Which of the methods of reducing the hallucination of Artificial Intelligence is effective in the field of the knowledge economy?

In the following sections, we will first introduce the knowledge economy and its relationship with generative AI. In Section 3, we will discuss the issue of hallucination in generative AI and explore its various forms. Then, after presenting the problem of hallucination in generative AI, we will introduce methods for reducing hallucinations to improve the productivity of the knowledge economy. Finally, Section 4 will present the conclusion and future research directions.

Theoretical Literature

Strengths of Generative AI

Generative AI refers to computational techniques that can create seemingly new and meaningful content, such as text, image, music, or video, from training data. Generative AI systems can not only be used for artistic purposes, such as generating new texts, imitating authors, or creating new images that mimic the style of image makers, but also as intelligent response systems to assist humans (Feuerriegel et al., 2023). In other words, Generative AI refers to deep learning models that can generate original content (not simply copying from their own training data) such as long texts, high-quality images, videos, or realistic sounds, in response to user requests or prompts. Generally, Generative AI operates in three stages:

1. Training: to create a base model.
2. Fine-tuning: to adapt the model to a specific application.
3. Generation, evaluation, and further fine-tuning: to improve accuracy. (Lv, 2023)

Types of Hallucinations in Generative AI

1. Sentence Contradiction: In this case, a large language model produces a sentence that

is completely contradictory to the one requested and lacks a subject matter connection. For example:

- Requested sentence: "Write a birthday card for mom".
- Expected response: "Happy Birthday, Mom. I love you".
- AI response: "I am so happy we are celebrating our first anniversary! Looking forward to many more years. With love,"

2. Request Contradiction: This occurs when the language model generates a response that contradicts the user's request and does not align with the subject matter.

For example:

- User's question: "Tell me about the benefits of meditation".
- Expected answer: "Meditation has many benefits, including stress reduction, improved concentration, and emotional well-being".
- AI-generated response (on an unrelated topic): "Meditation goes beyond earthly concerns and opens dimensional portals where your thoughts are shining butterflies guiding you through the cosmic realm of inner peace".

3. Factual Contradiction: In this case, the large language model generates an incorrect response to the user's query and presents it as a fact, which could mislead the user and spread misinformation. For example:

- User's query: "What is the capital of France?"
- Expected answer: "The capital of France is Paris".
- Incorrect AI response: "The capital of France is Zagreb".

4. Irrelevant or Random Hallucinations: This occurs when the large language model generates random information based on the user's query that has no subject connection and is simply a guess. For example:

- User's query: "Can you suggest a good recipe for chocolate chip cookies?"
- Expected answer: "Here's a classic recipe for chocolate chip cookies with step-by-step instructions".
- Random AI response: "The temperature today in Toronto is -2 degrees".

Various methods have been proposed to reduce hallucinations (Kanaani et al., 2024; Köpf et al., 2024; Mardiansyah et al., 2024; Yan et al., 2024). Each of these methods can help generate more accurate and reliable knowledge, thus assisting in providing more precise analyses. In the following section, we will examine the methods of reducing hallucinations in AI in more detail.

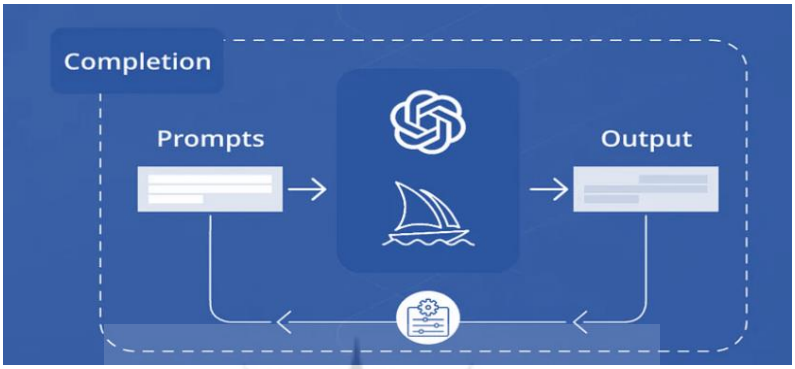
Methods for Reducing Hallucination

1. Prompt Engineering Methods

Prompt engineering is the art of crafting suitable inputs for AI models in such a way that they guide the models toward producing more accurate, relevant, and useful responses. This process combines logical thinking, creativity, language, and problem-solving skills. Prompt engineering is a field where, through experimentation and evaluation of different instructions, the goal is to obtain the best possible output from a text-based AI model (White et al., 2023). In terms of reducing hallucinations in model outputs, this process can

provide the context and expected results more precisely (Feldman et al., 2023). This process is very similar to coding, but instead of using structured instructions and vocabulary, natural language is employed. While the users have the freedom to use their own language, with that freedom comes great responsibility. A moderate request is likely to result in a moderate outcome. This structure is illustrated in Figure 1.

Figure 1.
Prompt Engineering Structure



(Source: Researcher's Findings)

By structuring the requests made to these models accurately, the likelihood of AI hallucination can be reduced. The main goal for users when utilizing AI is to obtain correct information. Therefore, the trust placed in acquiring information can greatly affect its use. To reduce the occurrence of AI hallucinations, users should consider three basic principles when making requests. These principles are:

1. Use clear and specific requests;
2. Provide examples and samples;
3. Adjust the LLM parameters.

Another prompt engineering-based method recently introduced is the LLM-Augmenter (Peng et al., 2023). This method proposes a system where an LLM is enhanced with a set of Plug-and-Play modules (Li et al., 2023). This system enables the LLM to generate prompts based on external knowledge. Additionally, using feedback generated by performance functions, it iteratively refines the LLM prompts to improve the model outputs. In this paper, the authors introduce the LLM-Augmenter to enhance the LLMs using external knowledge and automatic feedback.

2. Retrieval-Augmented Generation (RAG)

Retrieval-augmented Generation (RAG) systems combine the language model's capabilities with external information sources. The main idea is to retrieve relevant contexts from knowledge bases before generating text, so that the language model's outputs are based on real evidence. RAG enhances the responses of large language models by utilizing external, trusted knowledge bases instead of relying solely on potentially outdated training data or the model's internal knowledge. This approach addresses critical challenges related to the accuracy and timeliness of LLM outputs. (Kang et al., 2023).

A RAG system enhances the language model by integrating factual information obtained from various knowledge sources. This system reduces hallucination by ensuring that all the generated responses have supporting evidence and the model uses this evidence rather than making baseless guesses. If no evidence is found, the system can honestly declare its lack of knowledge. Here are a few examples to better understand how RAG requests work:

Example A: When was the first refrigerator invented?

Answer: 1854: The first working refrigerator was built by James Harrison, a Scottish inventor. His vapor compression refrigeration system paved the way for modern mechanical refrigeration technology.

Example B: Who was the lead actor in the movie Titanic?

Answer: Kate Winslet: She played the role of Rose DeWitt Bukater in the 1997 movie Titanic, directed by James Cameron. Leonardo DiCaprio also starred in the film as Jack Dawson.

In the above-mentioned examples, the system first retrieves relevant facts from a knowledge source, then generates the answer. This ensures that the output is based on the retrieved information, preventing the model from generating unsupported, incorrect facts. **Table 1** outlines the advantages and limitations of RAG. While RAG requests help reduce hallucination through retrieved facts, this technique is most effective when knowledge sources are large, high-quality, and regularly updated. Using multiple complementary knowledge bases can also offer greater flexibility.

Table 1.
Advantages and Limitations of RAG

Advantages	Limitations
Combines external information to reduce hallucination without basis.	Requires large datasets or knowledge bases for retrieving contextual information which can be costly.
Allows the model to return to an "uncertain" state if no contextual data is available.	If the knowledge source contains errors or real gaps, there is a risk of superficially repeating facts without a deep understanding.
Retrieved facts act as a strong signal to stabilize the generated text.	Not easily scalable compared to pure language model approaches.

(Source: Researcher's Findings)

3. Self-Improvement through Feedback and Reasoning

After a large language model generates an output for a specific prompt, appropriate feedback on that output can improve the quality and accuracy of the model's future outputs. In this context, specific techniques have been introduced to reduce errors caused by misconceptions. According to a study by Madaan et al. (2023), LLMs, especially GPT-3, have significant capabilities in few-shot prompting, which enhance their applications in real-world language tasks. However, the issue of improving GPT-3's reliability has not been fully explored. This study divides reliability into four key aspects—generalization, social biases, calibration, and reality—and introduces simple yet effective prompts to

improve each aspect. The research surpasses small-scale supervised models in all reliability metrics and provides practical solutions for enhancing GPT-3's performance.

4. CoVe (Chain-of-Verification Prompts)

Chain-of-Verification prompts (CoVe) explicitly ask models to provide step-by-step verification for their responses, citing credible external sources. These prompts are formulated as a series of logically verifiable inferences leading to the final answer. One type of chain-of-verification approach was presented by Si et al. (2022) in response to a user question. The LLM generates a base answer that may include inaccuracies, such as real hallucinations. They demonstrate an example of a question that GPT-3.5 failed to answer correctly.

Literature Review

Knowledge economy refers to an economy in which production and services are based on knowledge and information. In recent decades, the emergence of generative technologies, especially in the fields of AI and machine learning, has had profound and significant impacts on the knowledge economy. The knowledge economy, which is based on the production, distribution, and use of knowledge, is directly related to generative technologies. One of the prominent effects of generative technologies on the knowledge economy is the increase in efficiency and productivity in research and development processes. For example, using machine learning algorithms, companies can analyze large datasets and achieve faster results. Research shows that these technologies can significantly reduce the time required for the development of new products. Another application of generative technologies in knowledge economy is the automatic generation of knowledge through data processing. This can lead to increased efficiency and effectiveness in knowledge management (Alghanemi et al., 2022).

Next, we will examine the relationship between generative AI and knowledge economy more closely, as well as its impacts on various economic sectors.

Innovation and Creativity: Generative AI aids the innovation process by producing new ideas and content. In the knowledge economy, innovation is one of the key factors for economic growth and progress. For instance, designers can use generative AI to create new and innovative designs that help improve products and services (Teimuraz & Goderdzishvili, 2023). This capability not only increases competitive advantages but also fosters a culture of continuous improvement and adaptability in industries, ultimately contributing to economic growth in the knowledge economy.

Cost and Time Reduction: The use of generative AI in processes can significantly reduce the costs and time required for the content production. This means companies can manage their resources more efficiently and, as a result, have more capacity to invest in research and development (Yu & Lai, 2021).

Development of New Products and Services: Generative AI can assist companies in bringing new products and services to the market. For example, in the gaming industry, creating new content using AI can lead to the development of more engaging and diverse

games. This, in turn, can increase demand and foster the economic growth (Mayahi et al., 2022).

Personalization of Customer Experience: Using generative AI, companies can offer personalized experiences to their customers. This is due to the AI's ability to analyze the customer data and generate content that aligns with their preferences and needs. Such personalization can lead to increased customer loyalty and sales growth (Ramnarayan, 2021).

While generative AI can contribute to the development of the knowledge economy, it also brings risks of economic inequality. This technology may allow more advanced companies and countries with access to technological infrastructure and skilled labor to benefit more, which could further widen the economic gaps. Additionally, the use of generative AI introduces a series of ethical and legal challenges. For example, the creation of fake or misleading content can damage the credibility and legitimacy of information. Issues related to intellectual property and data usage must also be addressed. Furthermore, generative AI could lead to the reduction of traditional jobs. As AI increases productivity and efficiency, some traditional jobs might be affected and they will disappear. This creates a need for retraining and skill adaptation to help the workforce adjust to the new changes.

Another major challenge in the field of generative AI is the hallucination in large language models, which leads to the generation of seemingly correct but actually incorrect information, posing significant risks to the knowledge economy. We will review different types of hallucinations in AI.

Methodology

This article aims to investigate the impact of hallucinations in generative AI on the knowledge economy. Given the novelty of the field of hallucinations in generative AI, some relevant articles have been published in recent years. In this study, an attempt has been made to review all credible articles in this field systematically. Accordingly, Scopus, Google Scholar, Science Direct, and IEEE databases were searched using the keywords "artificial intelligence", "hallucination", and "knowledge economy". After an initial review, 50 articles were selected for a comprehensive review, and 15 articles were identified and selected for more in-depth analysis.

Findings

According to the reviews, the methods of illusion reduction that can be useful for productivity in the knowledge economy have been identified and the results are shown in Table 2. These studies show how different techniques can increase the accuracy and reliability of these models and, as a result, provide more accurate and practical analyses in the knowledge economy.

Table 2.
A Summary of Research Studies on Hallucinations in Large Language Models and their Impact on the Knowledge Economy

Authors	Insights	Conclusions
Yuji et al., 2024	The paper discusses how hallucinations in large language models, termed "knowledge overshadowing," can lead to over-generalization, resulting in inaccurate outputs. This undermines the reliability of knowledge-intensive tasks, affecting decision-making and trust in AI-generated information across various sectors.	• Knowledge overshadowing causes hallucination in language models. • Training data imbalance increases hallucination rates and impacts the model performance.
Ma et al., 2024	The paper does not specifically analyze the economic consequences of illusionary errors in large language models. It focuses on their error detection capabilities and the development of the Critical Calculation and Conclusion prompt to improve reliability in solving unreasonable math problems.	• LLMs can detect unreasonable errors but struggle with content generation. • CCC template improves the error detection and correction in LLMs.
Saxena et al., 2023	The paper highlights that hallucinations in large language models can lead to a loss of credibility and trust among users, particularly in critical fields like healthcare, education, and finance, ultimately affecting the economy of knowledge by undermining reliable information dissemination.	• The framework enhances the transparency and credibility of LLM responses. • The framework improves the accuracy and faithfulness of LLMs in drug-related inquiries.
Amayuelas et al., 2023	The paper does not specifically analyze the economic consequences of illusionary errors in large language models. Instead, it focuses on understanding LLMs' knowledge, measuring uncertainty, and addressing known-unknown questions to mitigate hallucinations.	• The study investigated capabilities of Large Language Models (LLMs) in understanding their own knowledge and measuring uncertainty. • It proposed a novel categorization scheme to elucidate the sources of uncertainty in known-unknown questions.
Bian et al., 2023	The paper highlights that false information in large language models (LLMs) can lead to global detrimental impacts, authority bias, and in-context pollution, ultimately challenging the reliability and safety of LLMs, which affects the economy of knowledge.	• False information spreads and contaminates memories in large language models. • Current large language models are susceptible to authority bias.

(Source: Researcher's Findings)

The review of hallucination reduction methods in the literature indicates that the approaches utilizing reinforcement learning with human feedback enhance the reliability of generated content by correcting errors in the model's output through repeated adjustments (Madaan et al., 2023; Tu et al., 2024; Vima et al., 2024). Based on the identification of various types of hallucinations in generative AI, this research attempted to propose a method that can reduce these types of hallucinations in the knowledge economy domain to produce and analyze accurate information. The proposed method is a hybrid approach based on three hallucination reduction techniques ofprompt engineering, retrieval-based generation, and self-improvement through feedback and reasoning. The proposed method is outlined step-by-step as follows:

1. Define the Question Clearly:

Formulate the question in the knowledge economy domain clearly and specifically.

a. Identify the key terms and concepts.
b. Define the scope and range of the response.

In the field of knowledge economics, it is crucial to precisely define specific terms and concepts to prevent misunderstandings and delineate the scope of inquiry. This involves clarifying the issues which should be addressed and excluded by the research question. For instance, instead of a broad question like "what is the role of technology in the economy?", a more focused question such as "what is the impact of information technology on economic growth in developing countries?" can be posed. In this refined question, key terms like "information technology," "economic growth," and "developing countries" are explicitly defined, thereby establishing clear boundaries for the response (Schulhoff et al., 2024).

2. Retrieve Relevant Information:

Use credible and reliable sources within the knowledge economy domain (e.g., scholarly articles, government reports, statistical databases) (Molina & Chicaíza, 2011)

a. Retrieve information related to the question using the identified key terms and concepts.
b. Evaluate the information based on the source credibility, publication date, and accuracy.

3. Prompt Engineering:

Design the prompt in such a way that guides the AI model to use the retrieved information (Lo, 2023). This involves crafting questions that are direct and focused, ensuring the AI can effectively utilize the data to generate relevant and insightful responses. In order to maximize the effectiveness of your inquiry, it is essential to establish a clear framework for analysis. This includes specifying the particular aspects of the knowledge economy you wish to explore, such as innovation, human capital, or digital transformation.

c. Use appropriate linguistic structures to direct the model (e.g., "based on the following information, ...").
d. Summarize and structure the retrieved information in the prompt and provide it to the model (Khamassi et al., 2024).

4. Generate the Response:

The AI model generates the response using the prompt and retrieved information. In this phase, the AI model leverages advanced algorithms and natural language processing techniques to analyze the retrieved information and generate a response that answers the research question (Feuerriegel et al., 2023)

5. Evaluation and Revision:

Evaluate the generated response for accuracy, completeness, and clarity. This iterative process allows for refining the inquiry and enhancing the quality of insights derived from

the AI model, ultimately leading to a more nuanced understanding of the complex dynamics within the knowledge economy. This iterative process enables the continuous improvement of inquiries, ensuring that insights gained from the AI model are not only accurate but also deeply informative and relevant to current trends in the knowledge economy. If necessary, revise the prompt and repeat steps 3 and 4 to obtain the desired response.

Discussion and Conclusion

As discussed in this research study, hallucination is recognized as the biggest obstacle to safely deploying powerful models in systems that impact people's lives. Therefore, the widespread adoption of large language models in practical environments depends heavily on resolving and mitigating hallucinations. To enhance the productivity of the knowledge economy using generative AI, it is essential to identify methods for reducing hallucinations. The knowledge economy involves optimizing the use of existing knowledge and information within a society or organization.

In this study, the issue of hallucinations in large language models, one of the tools of generative AI, was examined, and methods for reducing them were explored. Mitigating hallucinations in generative models for the knowledge economy leads to the production of accurate and up-to-date information. As a result, organizations can develop new products and services that better align with market needs and customer demands. Among the methods for reducing hallucinations, prompt engineering and self-improvement through feedback and reasoning are powerful tools that, by reducing hallucinations and enhancing the accuracy of information, can contribute to the productivity of the knowledge economy. These processes enable organizations and societies to make the best use of the existing knowledge while also generating new knowledge.

Accordingly, in this research study, a hybrid method based on prompt engineering, information retrieval, and feedback for generating accurate information using generative AI in the context of the knowledge economy was presented. Given the growing importance of generative AI and its role in knowledge production, it is recommended that researchers and developers pay greater attention to identifying and mitigating hallucinations in large language models. To this end, the following suggestions are noteworthy:

1. Further Research on Identifying Hallucinations: To improve the accuracy and reliability of large language models, more research is needed to identify and classify different types of hallucinations. Specifically, developing automated methods and algorithms to effectively detect hallucinations in model outputs should be prioritized.
2. Developing New Techniques to Address Hallucinations: Researchers should focus on developing advanced techniques to mitigate hallucinations. One proposed approach could involve using feedback systems that update language models based on the user inputs, thereby reducing errors in model outputs. Additionally, prompt engineering strategies can have a significant impact on reducing hallucinations.
3. Training Users to Use Generative Models Responsibly: Given that hallucinations

can lead to misleading information, users should be trained on how to use these models responsibly and how to identify incorrect or ambiguous outputs. This training could include warnings about the risks and challenges associated with using generative models.

4. Enhancing Transparency and Reliability of Models: An important recommendation is to enhance transparency in the design and training of the generative models. By improving transparency, users and researchers can gain a better understanding of the decision-making processes of models, thereby reducing the risk of hallucinations.

5. Further Study on the Impact of Hallucinations on Critical Decision-Making: It is essential to study the impact of hallucinations in large language models on critical decision-making, particularly in fields such as healthcare, law, and economics. These studies can help identify best practices for using models in these areas and mitigate risks arising from incorrect information.

6. Advances in AI Policy and Ethics: Researchers and policymakers should focus on developing ethical principles and legal frameworks for the use of generative AI. These principles could include responsible management of generated information, data ownership rights, and accountability for false information.

7. Future-Proofing and Continuous Model Updates: Continuous updating of large language models based on new data can help reduce hallucinations. Attention to up-to-date data and the use of appropriate scheduling methods for model updates can improve the accuracy and reliability of the generated results.

Finally, it is recommended that future researchers develop more comprehensive models in the area of hallucination reduction, specifically aimed at improving the knowledge economy. These models should be designed to reduce hallucinations in order to enhance the quality of knowledge produced for the knowledge economy. Additionally, creating regulatory frameworks for managing the generated content and developing assessment tools to detect and correct false information in this area is of critical importance.

REFERENCES

Alghanemi, J., & Al Mubarak, M. (2022). The role of artificial intelligence in knowledge management. In *Studies in Computational Intelligence*. *Future of Organizations and Work after the 4th Industrial Revolution* (pp. 359–373). 10.1007/978-3-030-99000-8_20.

Amayuelas, A., Wong, K., Pan, L., Chen, W., & Wang, W. (2023). Knowledge of knowledge: Exploring known-unknowns uncertainty with large language models. *ArXiv preprint arXiv:2305.13712.*

Bian, N., Liu, P., Han, X., Lin, H., Lu, Y., He, B., & Sun, L. (2023). A drop of ink may make a million think: The spread of false information in large language models. *ArXiv preprint arXiv:2305.04812*.

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., … Amodei, D. (2020). Language Models are Few-Shot Learners. Retrieved from http://arxiv.org/abs/2005.14165.

Farquhar, S., Kossen, J., Kuhn, L., & Gal, Y. (2024). Detecting hallucinations in large language models using semantic entropy. *Nature*, *630*(8017), 625–630. 10.1038/s41586-024-07421-0.

Feldman, P., Foulds, J. R., & Pan, S. (2023). Trapping LLM hallucinations using tagged context prompts. Retrieved from http://arxiv.org/abs/2306.06085.

Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2024). Generative AI. *Business & Information Systems Engineering*, *66*(1), 111-126.

Gunjal, A., Yin, J., & Bas, E. (2023). Detecting and preventing hallucinations in Large Vision Language Models. Retrieved from http://arxiv.org/abs/2308.06394.

Kanaani, M., Dadkhah, S., & Ghorbani, A. A. (2024, May). Triple-R: Automatic Reasoning for Fact Verification Using Language Models. In N. Calzolari, M.-Y. Kan, V. Hoste, A. Lenci, S. Sakti, & N. Xue (Eds.), *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)* (pp. 16831–16840). Retrieved from https://aclanthology.org/2024.lrec-main.1463.

Kang, H., Ni, J., & Yao, H. (2023). Ever: Mitigating hallucination in large Language Models through real-time Verification and rectification. Retrieved from http://arxiv.org/abs/2311.09114.

Khamassi, M., Nahon, M., & Chatila, R. (2024). Strong and weak alignment of large language models with human values. *Sci Rep*, *14*, 19399. https://doi.org/10.1038/s41598-024-70031-3.

Köpf, A., Kilcher, Y., von Rütte, D., Anagnostidis, S., Tam, Z.-R., Stevens, K., … Mattick, A. (2024). *OpenAssistant conversations - democratizing large language model alignment*. Article 2064. Presented at the Proceedings of the 37th International Conference on Neural Information Processing Systems, New Orleans, LA, USA. Curran Associates Inc.

Li, M., Peng, B., Galley, M., Gao, J., & Zhang, Z. (2023). Self-Checker: Plug-and-play modules for fact-checking with large language models. Retrieved from http://arxiv.org/abs/2305.14623.

Ma, J., Dai, D., Sha, L., & Sui, Z. (2024). Large language models are unconscious of unreasonability in math problems. *ArXiv preprint arXiv:2403.19346*.

Madaan, A., Tandon, N., Gupta, P., Hallinan, S., Gao, L., Wiegreffe, S., … Clark, P. (2023). Self-refine: Iterative refinement with Self-feedback. 10.48550/ARXIV.2303.17651.

Mardiansyah, K., & Surya, W. (2024). Comparative analysis of ChatGPT-4 and Google Gemini for spam detection on the SpamAssassin public mail corpus. 10.21203/rs.3.rs-4005702/v1.

Mayahi, S., & Vidrih, M. (2022). The impact of generative AI on the future of visual content marketing. 10.48550/ARXIV.2211.12660.

Molina, M. G., & Chicaíza, L. (2011). A Guide to Sources for Research in Economic Sciences. *Social Science Research Network*. https://doi.org/10.2139/SSRN.1766062.

Mündler, N., He, J., Jenko, S., & Vechev, M. (2023). Self-contradictory hallucinations of large language models: Evaluation, detection and mitigation. Retrieved from http://arxiv.org/abs/2305.15852.

Peng, B., Galley, M., He, P., Cheng, H., Xie, Y., Hu, Y., … Gao, J. (2023). Check your facts and try again: Improving large language models with external knowledge and automated feedback. Retrieved from http://arxiv.org/abs/2302.12813.

Qiu, Y., Ziser, Y., Korhonen, A., Ponti, E. M., & Cohen, S. B. (2023). Detecting and mitigating hallucinations in multilingual summarisation. Retrieved from http://arxiv.org/abs/2305.13632.

Ramnarayan, S. (2021). Marketing and artificial intelligence. In *Advances in Marketing, Customer Relationship Management, and E-Services* (pp. 75–95). 10.4018/978-1-7998-5077-9.ch005.

Ray, P. P. (2023). Chatgpt: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*, *3*, 121–154.

Saxena, S., Prasad, S., Prakash, M. V., Shankar, A., Vaddina, V., & Gopalakrishnan, S. (2023). Minimizing factual inconsistency and hallucination in large language models. *arXiv preprint arXiv:2311.13878.*

Schulhoff, S., Ilie, M., Balepur, N., Kahadze, K., Liu, A., Si, C., Li, Y., Gupta, A., Han, H., & Schulhoff, S. (n.d.), (2024). *The Prompt Report A Systematic Survey of Prompting Techniques*. https://arxiv.org/abs/2406.06608.

Si, C., Gan, Z., Yang, Z., Wang, S., Wang, J., Boyd-Graber, J., & Wang, L. (2022). Prompting GPT-3 to be reliable. Retrieved from http://arxiv.org/abs/2210.09150.

Teimuraz Goderdzishvili, T. G. (2023). Artificial intelligence and creative thinking, the future of idea generation. *Economics*, *105*(3–4), 63–73. 10.36962/ecs105/3-4/2023-63.

Tu, C.-H., Hsu, H.-J., & Chen, S.-W. (2024). Reinforcement learning for optimized information retrieval in LLaMA. 10.21203/rs.3.rs-3847100/v1.

Vima, C., Bosch, H., & Harringstone, J. (2024). Enhancing inference efficiency and accuracy in large language models through next-phrase prediction. 10.21203/rs.3.rs-4864441/v1.

White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., … Schmidt, D. C. (2023). A prompt pattern catalog to enhance prompt engineering with ChatGPT. Retrieved from http://arxiv.org/abs/2302.11382.

Yan, Y., Zheng, P., & Wang, Y. (2024). Enhancing large language model capabilities for rumor detection with Knowledge-Powered Prompting. *Engineering Applications of Artificial Intelligence*, *133*(108259), 108259. 10.1016/j.engappai.2024.108259.

Yin, Z. (2024). A review of methods for alleviating hallucination issues in large language models. *Applied and Computational Engineering*, *76*, 258-266.

Yu, L., & Lai, Z. (2021). The management of internet content products in the era of artificial intelligence. *Journal of Physics. Conference Series*, *1757*(1), 012017. 10.1088/1742-6596/1757/1/012017.

Yuji, Zhang., Sha, Li., Jiateng, Liu., Pengfei, Yu., Yi, R., Fung., Jing, Li., Manling, Li., Heng, Ji. (2024). Knowledge Overshadowing Causes Amalgamated Hallucination in Large Language Models. Retrieved from arXiv:2407.08039v1. https://doi.org/10.48550/arXiv.2407.08039.