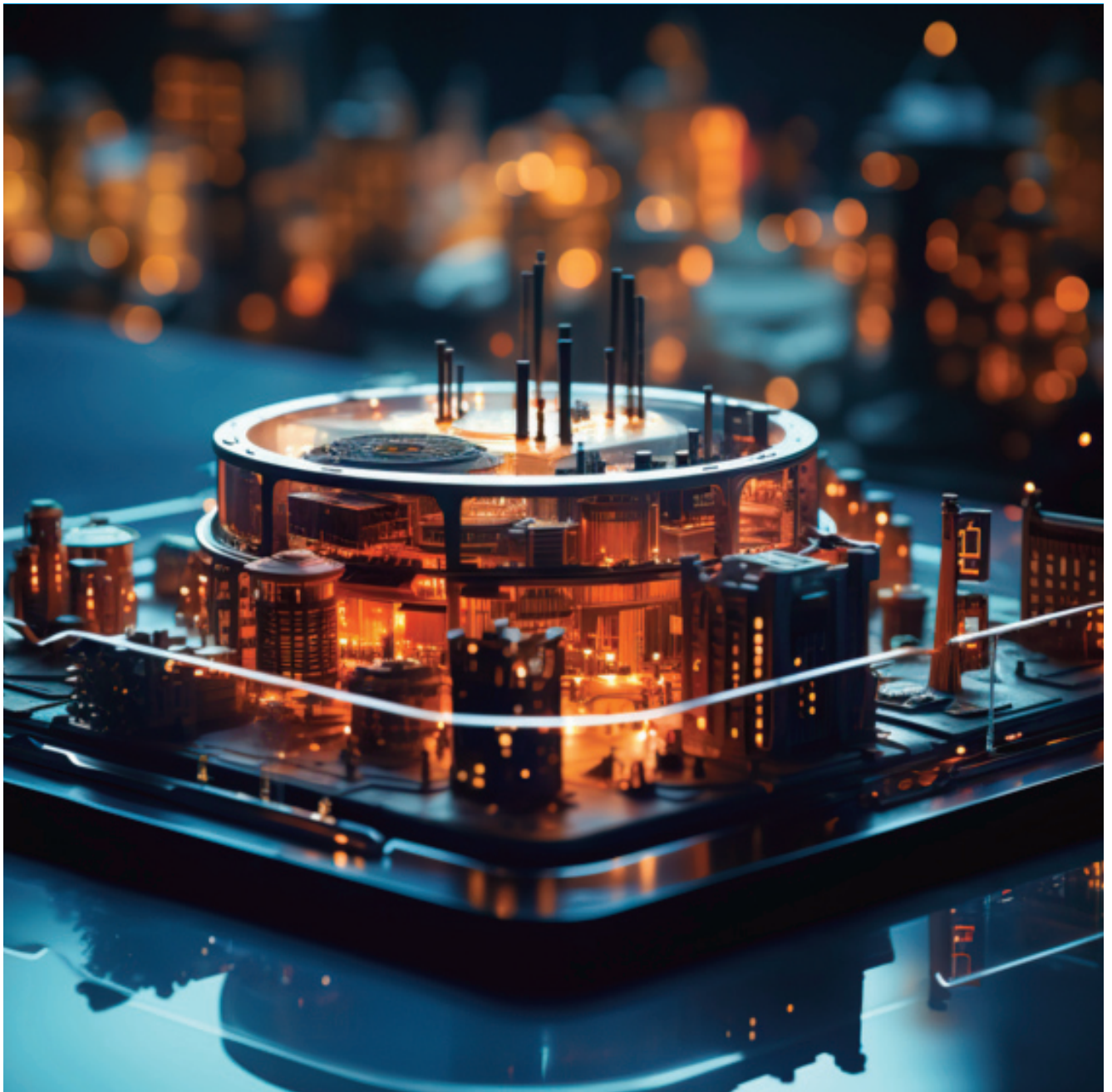


Conclusions

“LLMs are the only people who can write a novel, translate it into ten languages, and still not understand the plot.”
Perplexity¹³⁰



Large Language Models (LLMs) represent a significant advance in the field of artificial intelligence and are revolutionizing the way we interact with technology and leverage natural language processing. Their ability to process and generate coherent, contextualized text opens up a wide range of applications in a variety of industries, from content creation and sentiment analysis to task automation and improved user experience.

However, there are a number of important challenges and considerations in developing and deploying LLM. The presence of biases and hallucinations in their results, the lack of transparency and explainability of their decisions, the challenges of privacy and information security, and the high consumption of computational resources are some of the key challenges that must be addressed to ensure a responsible and ethical use of these systems.

To address these challenges, it is critical to establish a robust AI governance framework, especially in the area of generative AI and LLMs. This framework must encompass all key aspects, including strategy, risk appetite, governance, organization, control framework (policies and procedures), data, systems and reporting. Only a comprehensive and well-structured approach will ensure these technologies' responsible development and use.

Validation plays a crucial role within this governance framework. Adopting a multidimensional approach that covers all stages of the LLM lifecycle is desirable, from the quality of the input data and robustness of model design to thorough evaluation of results and appropriate implementation and use. This validation process should combine standardized quantitative metrics with human evaluation techniques tailored to the specific context of each use case.

In addition, it is necessary to keep abreast of the latest trends and advances in the field of LLM validation, such as the development of more advanced explainability techniques, the use of LLMs to explain the behavior of other LLMs, continuous validation and monitoring in production, and alignment with ethical principles and regulatory requirements.

The case study presented in this white paper illustrates how the application of a customized validation framework can help organizations identify strengths and areas for improvement in their LLM-based systems, and make informed decisions about their implementation and continuous improvement.

In short, LLMs have great potential to transform the way businesses and society at large benefit from artificial intelligence. However, to realize their full potential in a safe and responsible manner, it is imperative to establish a robust AI governance framework that addresses the challenges associated with their development and deployment, and includes a rigorous, multidimensional approach to validation. This is the only way to ensure that these systems are reliable, fair, and aligned with the values and goals of organizations and society.

¹³⁰Perplexity AI is an LLM-based conversation and research search engine founded in 2022 by Andy Konwinski, Denis Yarats, Johnny Ho, and Aravind Srinivas (formerly OpenAI) that answers queries using natural language predictive text.