



# Chapter 8

## Large Language Models (LLMs)

**Gusti Muhamad Sardana**

 <http://orcid.org/0009-0001-4281-2432>  
*Universitas Esa Unggul, Indonesia*

**Binastya Anggara Sekti**

 <http://orcid.org/0000-0001-5489-4888>  
*Universitas Esa Unggul, Indonesia*

**Diah Aryani**

*Universitas Esa Unggul, Indonesia*

**Hani Dewi Ariessanti**

*Universitas Esa Unggul, Indonesia*

### ABSTRACT

*Large Language Models (LLMs) are transformative AI systems trained on vast text data, enabling natural language understanding and generation. Evolving from rule-based and statistical NLP, LLMs utilize transformer architectures, attention mechanisms, and tokenization strategies for high contextual comprehension. They support tasks from content creation to code generation, and find applications in education, healthcare, law, and creative industries. Despite their capabilities including emergent reasoning and multimodality, LLMs face challenges like bias, hallucination, high energy use, and data privacy risks. Ethical governance and sustainable development are critical as LLMs reshape digital interaction and approach Artificial General Intelligence (AGI). This article provides a comprehensive overview of their architecture, training processes, applications, and future trends.*

### 1. INTRODUCTION TO LARGE LANGUAGE MODELS (LLMs)

Large Language Models (LLMs), which are extremely complex and adaptable models, are a recent development in artificial intelligence (AI). This model serves as the foundation for a wide range of language-based applications, including chatbots, content creation, automatic translation, and intelligent support in the fields of medicine and education. Large-scale text datasets are used to train LLM language

DOI: 10.4018/404019

models, which use millions to billions of factors to predict, comprehend, and produce text that mimics human speech.

The distribution of words in natural language is the main emphasis of the probabilistic LLM model, which has a deep understanding of context. In contrast to traditional NLP models, LLM learns language patterns using vast volumes of data covering a variety of styles, domains, and cultural settings rather than depending just on linguistic rules. LLM incorporates sophisticated representation learning and attention techniques as deep learning technology develops, allowing contextual comprehension across phrases, paragraphs, and even lengthy texts (Hadi et al., 2023).

Prior to LLM, NLP development progressed through several stages, including rule-based methods, statistics, and basic neural models. Although they depended on local word associations, early models like n-grams and Hidden Markov Models (HMM) struggled to capture global context. NLP underwent a big revolution with the introduction of word embeddings like Word2Vec and GloVe, which led to notable advancements and their use in models like BERT and GPT. More intricate contextual linkages between text tokens and parallel processing are made possible by the Transformer design (Han et al., 2024).

In addition to enhancing technical performance, the switch from conventional language models to LLMs has altered how people communicate with computer systems. In some situations, LLMs like GPT-3, PaLM, and Claude have demonstrated the ability to tackle a variety of NLP problems with performance that is comparable to or even better than that of humans. The model can tackle new tasks without requiring extra retraining because to a feature known as zero-shot and few-shot learning. Because of this, LLMs may be applied with great flexibility in a variety of fields, including software development, journalism, law, and knowledge management (Mao et al., 2025).

The versatility of LLM is one of its key advantages. This indicates that the model is now designed to be able to generalize across domains rather than being trained for a single task type. For instance, LLM can serve as a virtual instructor in the classroom, providing interactive explanations of difficult ideas. LLM has been tested in the healthcare industry to assist in the analysis of medical material, the summarization of scientific publications, and even the provision of preliminary diagnosis recommendations based on symptoms (Petro anu et al., 2023). Economic decision support systems, consumer sentiment research, and automated legal report authoring are among more uses. Since models are trained on data that may not be free of social prejudice, there is also a danger of adding gender, racial, or political biases that are hard to eradicate. In interactions between humans and machines, this may systemically perpetuate preconceptions (Oralbekova et al., 2023).

Nevertheless, there are a number of difficulties associated with this outstanding advancement. The extremely high energy usage throughout the training process is one of the most basic problems. Large computational infrastructure and a substantial carbon footprint are necessary for models such as GPT-3. Additionally, LLMs are vulnerable to social biases, cultural prejudices, and even false information because they are trained on material that is routinely gathered from the internet. Particularly in public and professional uses, this presents ethical and accountability issues. Limitations in logical thinking, the interpretability of outcomes, and the possibility of being utilized maliciously for things like information manipulation or the creation of bogus material are further difficulties (Khan et al., 2023).

Researchers are advocating for a more responsible and moral approach to LLM development in order to solve these issues. The creation of models that reflect human values using a training procedure that incorporates reinforcement learning from human feedback (RLHF) is one suggested tactic. Furthermore, maintaining public openness and confidence in LLM also heavily relies on dataset and model parameter

27 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

[www.igi-global.com/chapter/large-language-models-llms/404019](http://www.igi-global.com/chapter/large-language-models-llms/404019)

## Related Content

---

### A Hybrid Model for Service Selection in Semantic Web Service Composition

Sandeep Kumar and R.B. Mishra (2008). *International Journal of Intelligent Information Technologies* (pp. 55-69).

[www.irma-international.org/article/hybrid-model-service-selection-semantic/2443](http://www.irma-international.org/article/hybrid-model-service-selection-semantic/2443)

### The Ethical Implications of Artificial Intelligence in Antimicrobial Resistance

Vineet Singh (2025). *AI-Driven Breakthroughs in Antimicrobial Resistance* (pp. 341-350).

[www.irma-international.org/chapter/the-ethical-implications-of-artificial-intelligence-in-antimicrobial-resistance/374870](http://www.irma-international.org/chapter/the-ethical-implications-of-artificial-intelligence-in-antimicrobial-resistance/374870)

### Agent-Based Modeling for Simulation of Complex Business Systems: Research Design and Validation Strategies

Christoph Schlueter Langdon (2005). *International Journal of Intelligent Information Technologies* (pp. 1-13).

[www.irma-international.org/article/agent-based-modeling-simulation-complex/2386](http://www.irma-international.org/article/agent-based-modeling-simulation-complex/2386)

### The Role of AI in Personalized English Language Learning

Muhammad Usman Tariq (2026). *AI-Powered English Teaching* (pp. 93-124).

[www.irma-international.org/chapter/the-role-of-ai-in-personalized-english-language-learning/384169](http://www.irma-international.org/chapter/the-role-of-ai-in-personalized-english-language-learning/384169)

### Deep Learning-based Framework for Smart Sustainable Cities: A Case-study in Protection from Air Pollution

Nagarathna Ravi, Vimala Rani P, Rajesh Alias Harinarayan R, Mercy Shalinie S, Karthick Seshadri and Pariventhan P (2019). *International Journal of Intelligent Information Technologies* (pp. 76-107).

[www.irma-international.org/article/deep-learning-based-framework-for-smart-sustainable-cities/237967](http://www.irma-international.org/article/deep-learning-based-framework-for-smart-sustainable-cities/237967)