

## TOWARDS TRULY OPEN, LANGUAGE-SPECIFIC, SAFE, FACTUAL, AND SPECIALIZED LARGE LANGUAGE MODELS

**Preslav Nakov**

*Mohamed bin Zayed University of Artificial Intelligence,  
Abu Dhabi, United Arab Emirates  
[preslav.nakov@gmail.com](mailto:preslav.nakov@gmail.com)*

**Abstract.** Large language models (LLMs) have rapidly advanced, but their development raises important challenges around transparency, inclusivity, safety, and reliability. At MBZUAI’s Institute on Foundation Models (IFM), we argue that addressing these challenges requires a multi-faceted approach that combines open-source innovation with targeted specialization. First, we emphasize the importance of transparency. Most existing LLMs remain opaque, limiting reproducibility and slowing scientific progress. Through the LLM360 initiative, we pioneer fully transparent open-source LLMs, providing the community with access to training data, model weights, and documentation. Second, we highlight the need for language-specific models. Current LLMs are disproportionately English-centric, often underperforming for low-resource languages. To bridge this gap, we develop a suite of open models including Jais, the world’s leading open Arabic-centric foundation and instruction-tuned LLM; Nanda, an open Hindi model; and Sherkala, an open Kazakh model. Third, we stress the critical importance of safety. LLMs must operate with robust safeguards to prevent harmful or inappropriate outputs. We introduce Do-Not-Answer, a benchmark dataset for evaluating guardrail performance, which is central to our safety mechanisms. Fourth, we address the issue of factuality. LLMs are prone to generating plausible but inaccurate information, creating barriers to real-world adoption. We discuss key factuality challenges and outline directions for building models that prioritize reliability. Finally, we argue for the need for specialized models,

## **Towards Truly Open, Language-Specific, Safe, Factual, and Specialized Large ...**

and we will present some other LLMs currently being developed at MBZUAI's IFM. Together, these initiatives form a roadmap for advancing open, inclusive, safe, and factual LLMs that better serve global communities.