Software Architecture for Large Language Model Integration

Objectives:

This research focuses on addressing the challenges of integrating Large Language Models (LLMs) into software architectures, examining trade-offs in key quality factors such as performance, cost, and scalability. The aim is to provide best architectural practices for including LLM-based components into current systems.

Aims:

The Aims is to create a flexible framework for integrating LLMs into various software architecture models, such as cloud-native, edge, and hybrid environments. This research strives to offer practical insights into designing efficient, secure, and scalable systems that incorporate LLMs, while also upholding high standards for security, privacy, and governance.

Methods:

A mixed-methods approach will be used, combining a thorough literature review, case study analysis, and expert interviews. The literature review will cover existing research on AI integration in software systems, while case studies will explore real-world LLM applications. Interviews with experts will provide qualitative data to help identify challenges and gaps in current LLM integration practices.

Results:

The expected outcomes include identifying key issues with LLM integration, such as latency, resource consumption, and the trade-offs in model accuracy. The study will propose system architectures and design patterns for LLM-based solutions, addressing scalability, security, and privacy concerns. The results will highlight best practices for balancing performance and ethical considerations, along with guidelines for managing the lifecycle of LLM models in production environments.

Conclusion:

In conclusion, this research will make valuable contributions to the emerging field of LLM software architecture by providing frameworks and best practices. These contributions will fill existing gaps in research and offer practical guidance for developers and architects working with LLM-based systems. The findings will also set the stage for further studies on LLM integration in diverse software environments.

Keywords:
Large Language Models, Software Architecture, System Integration, Security, Scalability, Privacy, Ethical Governance, Cloud-native, Edge Computing, AI Integration.