

# システム開発業務への生成 AI 適用技法の研究 (クラス 1) ー初心者による対話型 LLM コーディングの課題抽出と QCD 評価の試みー

## アブストラクト:

本研究は、初心者開発者が対話型大規模言語モデル (LLM) を用いて設計書から包括的にコードを生成する際の、QCD (品質、コスト、納期) への影響と、それに伴う課題を実証的に調査した。Java、C#、COBOL の 3 つのプログラミング言語を対象に、LLM を使用しない人間中心の開発チームと、LLM を使用する開発チーム (初心者コーダーと熟練レビューアで構成) を比較する実験を行った。実験は、複雑度の異なる 2 つのタスク (実験 1: 低難度、実験 2: 高難度) と、ガイドラインを導入した実験 3 (高難度) の 3 回に分けて実施した。結果として、現時点の対話型 LLM を用いた包括的コード生成では、人間開発を上回る QCD は達成できず、特にメインフレーム COBOL では LLM によるコード生成が困難であることが示された。また、初心者開発者のプロンプトの不備、LLM の一貫性の欠如、設計書の不備などが複合的に作用し、品質低下や手戻りを招くことが明らかになった。さらに、ガイドラインの導入は、一定の効果を示すものの、初心者による誤用や思考停止、LLM の技術的限界により、効果は限定的であった。これらの結果から、対話型 LLM を用いた包括的コード生成の実現には、LLM の技術的改善、プロンプトエンジニアリング教育、設計書の品質向上、人間と LLM の明確な役割分担と協調を促す開発プロセス・ツールの整備が不可欠であると結論づける。

## キーワード:

対話型 LLM (Large Language Models)、初心者コーディング、QCD (品質・コスト・納期)、ヒューマンファクター