

Évaluation des assistants de programmation basés sur l'IA-LLM (GitHub Copilot, Amazon CodeWhisperer, ChatGPT)

Proposition de projet informatique – automne 2023

Marc Frappier, professeur, Département d'informatique, Faculté des sciences

Objectif

Évaluer de manière qualitative et quantitative les fonctions des assistants de programmation basés sur des LLM (*Large-scale language models*).

Méthodologie

Évaluer les limites des études existantes. Par exemple :

- [Is ChatGPT the Ultimate Programming Assistant - How far is it?](#)
- [Evaluating the Code Quality of AI-Assisted Code Generation Tools: An Empirical Study on GitHub Copilot, Amazon CodeWhisperer, and ChatGPT](#)
- [Evaluating Large Language Models Trained on Code](#)

Développer un nouveau barème de comparaison qui surpasse les limites identifiées, par exemple :

- Cas de test
- Métrique d'évaluation
- Forme d'assistance

Évaluer les assistants avec ce nouveau barème.