Final Report: Evaluation of Various Software Development AI Assistants

Luna Sugiyama
Graduate School of Information Science and Technology
Information and Communication Engineering
lunagracesugiyama@g.ecc.u-tokyo.ac.jp

July 25, 2025

I have recently been interested in Devin, Claude Code, and Cline which are AI assistant tools to help software development. In this report I will choose these services and evaluate them on various tasks and aspects to see their strengths and weaknesses.

1. About Each Service

1.1. Devin [2] [1]

Devin, developed by Cognition Labs, is positioned as the world's first fully autonomous AI software engineer. It is designed to take a high-level development request and independently execute the entire project from start to finish. This includes tasks such as: Planning and Architecture: Devising a comprehensive plan to address the user's prompt. Coding: Writing the necessary code in various programming languages. Debugging: Identifying and resolving issues within the codebase. Testing and Deployment: Ensuring the software functions correctly and deploying it. Devin operates as a sandboxed application, meaning it has its own development environment where it can install dependencies, write and execute code, and access the internet for research. This autonomy allows it to tackle complex, long-running tasks that have traditionally required a human development team.

2. Evaluation

Describe your approach, data collection methods, assumptions, tools used, etc.

3. Results / Findings

Present key results, findings, or analysis. You can include figures, tables, and equations where appropriate.

4. Discussion

Interpret your results, compare with existing work if relevant, and explain their implications.

5. Conclusion

Summarize the main points, state limitations, and propose future work or recommendations.

References

[1] Kumar Madhukar. A comprehensive guide to vibe coding tools, 2025-03. Accessed: 2025-07-24.

[2] Salman Raza. Meet devin ai: The next generation of software engineering powered by autonomous ai, 2024. Accessed: 2025-07-24.