

How to Align Human Intent and AI Agent for Long-term Tasks

Dongjae Lee

2025.11.24

Abstract

Despite the overwhelming performance of recent models like Gemini 3.0, building production-grade software solely with AI remains a challenge. This essay identifies the “alignment problem” in long-term tasks like building a large-scale software as the root cause. Then, we propose a structured requirement writing language as the necessary solution.

AI capabilities are improving exponentially ¹, with recent benchmarks showing models outperforming human developers. Yet, we still struggle to construct large-scale, high-quality software using AI. Attributing this solely to a lack of specific training (e.g., DevOps) overlooks the fundamental issue.

The core challenge is alignment: ensuring the AI executes the user’s true intent. As task duration increases, alignment becomes increasingly fragile. A microscopic initial deviation results in a massive miss over a long distance. To leverage AI for tasks replacing hours of human labor, we must provide requirements of proportional precision.

Furthermore, the risk of deviation grows with task length, proportionally increasing the cost of verification. We have merely traded implementation time for review time; the total engineering cost remains high.

Writing rigorous requirements for massive tasks is inherently difficult. Relying on unstructured “prompts” is unsustainable for scaling software engineering. The solution to this complexity is abstraction. To effectively align human intent with AI agents for long-term tasks, we must transition from simple prompts to a well-defined requirement writing language.

¹Measuring AI Ability to Complete Long Tasks