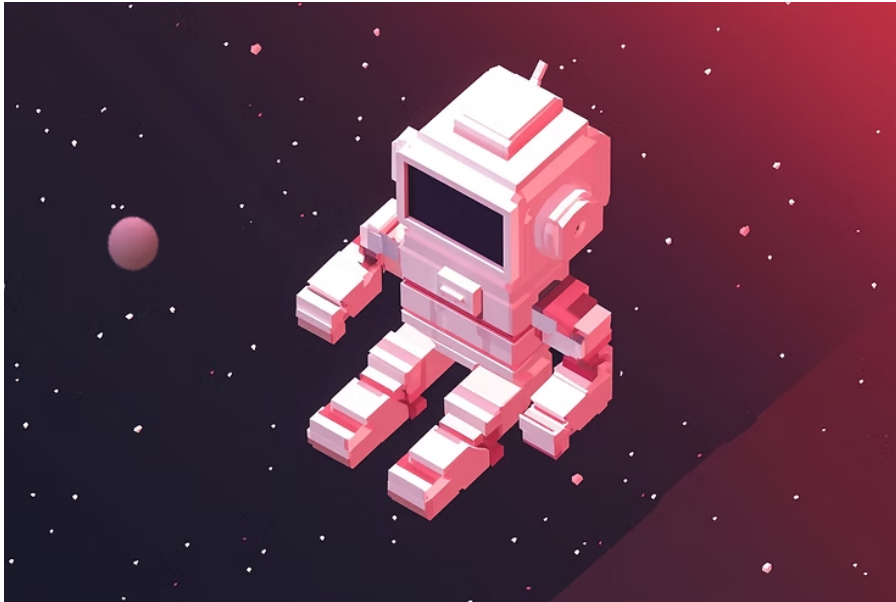


The New Kid on the Block: How OpenAI's Agent Stacks Up Against the Real World

July 23, 2025



Here's a confession that might undermine my credibility as someone who allegedly keeps up with technology: I've been quietly using AI agents for months, and nobody seemed to notice.

Whilst everyone was getting excited about ChatGPT's latest party tricks, I've been relying on Manus for the actual grunt work. You know, the tedious bits that make you question your life choices—coordinating stakeholders, synthesising research, keeping track of who said what in which meeting three weeks ago.

So when Sam Altman [announced ChatGPT Agent](#) last week with the kind of fanfare usually reserved for product launches that will “change everything forever,” my first thought wasn't excitement. It was curiosity about whether OpenAI had finally caught up to what some of us have been using all along.

The Reality Check Nobody's Talking About

Let me paint you a picture. I'm juggling three concurrent projects, each with stakeholders who seem to operate in parallel universes where deadlines are merely suggestions. My Manus agent has already sorted through overnight emails, flagged the urgent ones, drafted responses to the routine queries, and updated project timelines based on the latest delay notifications.

It's not particularly glamorous. There's no fanfare, no tweets about revolutionary breakthroughs. It's just... useful. Properly, mundanely useful in the way that good project management tools should be.

Which brings me to my slightly cynical question about OpenAI's latest offering: is this genuinely better, or just better marketed?

When Gold Medals Meet Monday Morning Reality

Fair play to OpenAI—their system achieving [gold medal performance on mathematical olympiads](#) is properly impressive. The kind of abstract reasoning required for those problems does suggest capabilities that could transform how we approach complex project challenges.

But here's what I've learned from months of using AI agents in real project environments: mathematical brilliance doesn't always translate to understanding why your stakeholder is suddenly silent on emails or recognising that "quick chat" in someone's calendar invite actually means "prepare for a complete scope change."

The messy reality of project delivery isn't a maths problem—it's a human problem with mathematical elements. And whilst I'm genuinely curious to see how OpenAI's reasoning capabilities handle the gloriously irrational world of stakeholder management, I'm reserving judgement until I've put it through its paces.

The Browser That Might Actually Matter

OpenAI's plans for [an AI-powered browser](#) intrigue me just as much as their agent, if I'm honest. Not because I think browsing needs revolutionising, but because information synthesis is where AI genuinely excels.

I spend embarrassing amounts of time trying to stay current with industry developments, client sector news, and the latest methodologies that might actually improve how we deliver projects. If an AI browser could pre-digest the relevant bits and present them in project-specific context, that might be genuinely transformative.

Though knowing my luck, it'll probably just serve up more articles about how AI is going to replace project managers whilst I'm trying to research construction industry regulations.

Context Engineering: The Unsexy Bit That Actually Matters

This is where companies like Manus have been quietly brilliant. Their [work on context engineering](#) tackles the unsexy but crucial challenge of helping AI understand what you actually mean, not just what you literally said.

Anyone who's worked on projects knows the pain of perfectly clear requirements being completely misunderstood because someone missed the implicit context. The organisational politics, the historical decisions, the unspoken assumptions that shape why we're approaching things this way.

Manus has spent considerable time solving this puzzle, and it shows in daily use. The system doesn't just follow instructions—it grasps the broader project context and adapts accordingly. It's the difference between having a diligent intern who needs everything spelled out and a seasoned colleague who

understands the subtext.

I'm curious whether OpenAI's agent will match this level of contextual sophistication or whether it'll be another case of impressive demos that fall apart when faced with real-world ambiguity.

The Creative Leap That Caught My Attention

Here's something that genuinely surprised me: recent observations about [imagination in AI systems](#) suggest we're moving beyond sophisticated pattern matching into something approaching genuine creativity.

In project delivery, creativity isn't optional—it's survival. Every project manager has faced that moment when conventional approaches fail spectacularly and you need to invent a solution that nobody's tried before. The ability to imagine alternative scenarios, spot unexpected connections, find elegant workarounds when everything seems impossible.

If AI agents can genuinely contribute to this creative problem-solving rather than just executing predetermined tasks, that could be genuinely significant. Though I'll believe it when I see it handle a client who's changed their mind about fundamental requirements two weeks before go-live.

The Honest Assessment

After months of using AI agents for actual project work, here's my decidedly unrevolutionary conclusion: they're extremely useful for the mundane bits and occasionally helpful for the complex bits, but they're not able to replacing human judgement on anything that matters....yet

The strategic thinking, stakeholder diplomacy, and creative problem-solving that define effective project delivery remain stubbornly human. But the administrative overhead, routine coordination, and information synthesis that consume disproportionate amounts of our energy? That's where AI agents genuinely excel.

OpenAI's entry into this space is interesting primarily because it might force the entire sector to improve. Competition tends to accelerate development, and if ChatGPT Agent pushes companies like Manus to enhance their offerings, everyone benefits.

What This Means for Your Next Project

My recommendation? Approach this with curiosity rather than either excitement or scepticism. The technology is useful enough to experiment with but not transformative enough to bet your career on.

Start small. Use AI agents for the bits that make you want to hide under your desk—the routine updates, basic research, initial draft communications. See what works, understand the limitations, and gradually expand usage as you build confidence in the technology's reliability.

Most importantly, don't feel pressured to adopt everything immediately just because it's new. Some of us have been quietly using these tools for months without fanfare. The real value lies in thoughtful

integration, not breathless adoption.

I'll be testing OpenAI's agent against my current Manus setup over the coming weeks. Not because I expect revolutionary changes, but because understanding the landscape helps make better decisions about which tools actually improve project delivery versus which ones just add complexity.

The future of project management might be arriving gradually, mundanely, and far less dramatically than the headlines suggest. And honestly? That sounds about right for our industry.

Based on practical experience using AI agents in real project environments and ongoing developments in the field. Results may vary, batteries not included, your mileage may differ.