The best book on AI Agents for the layman

27 March 2026 · agile, ai, banking, breaking things, bug, gothic

Alan J. Portis once quipped (https://dl.acm.org/doi/pdf/10.1145/947955.1083808) that: “the best book on programming for the layman is Alice in Wonderland, but that’s because it’s the best book on anything for the layman”. In the same theme, I’d like to suggest an addition to your library, the best non-technical book on AI Agents for the layman is “Frankenstein; or, The Modern Prometheus” by Mary Shelley. Why? Because it brings home the key facts about autonomous agent aspects like quality, autonomy & speed (I can see what you’re thing, i.e.: pick any two).

The book, unlike the ubiquitous screen adaptations and sequels, digs deep into the motivations & feelings of responsibility on Dr Frankenstein, the Monster’s creator. It’s clear Mary Shelly had an inkling of the sort of mayhem a spirited agent could get up to, and the anguish that would cause.

The parallels with modern software development and AI engineering are uncanny. For example, the doctor creates his ‘agent’ after a period of intense work and research into the field. No sooner has he created it, than he quickly feels dissatisfied, it really isn’t what he hoped for. Before he can get a true handle on his creation, the agent has fled from the lab. Later Dr Frankenstein learns the agent has killed the doctor’s brother (William) and friend (Henry) before going on to kill his bride & cousin (Elizabeth) - not only that - but the monster has also stitched up a close family friend for one murder (Justine is executed) and let the doctor be held for several months on suspicion of killing for the other!

I mean, who hasn’t been there right? it’s just another day in the modern dev-shop, can we all at least agree on that?

But seriously, Frankenstein’s monster isn’t the lumbering oaf of the TV renditions. He is a giant with superhuman strength. Shelly portrays the monster as a creature with great physical prowess, but with-out the emotional intelligence or temperament to match. So, unlike the Great Danes that lumber about the park near my home, whose mere size is imposing - but have a gentle and non-aggressive demeanour, Frankenstein’s giant monster has a hair trigger for violence.

One of these guys is a dangerous beast, the other loves a tummy rub & treat.

Similarly, our agents are powerful, with the tools we give them they can help us deliver more & faster. But they have little if any knowledge the problems they wrought accidentally & blindly. They also, by default, have poor stopping heuristics, requiring us to ensure we have guided them appropriately. While they don't lure you on a life draining trek to the high north [The book starts and ends in the arctic circle] – but sometimes I do get feeling I'm stuck in a listless pursuit. But a stern warning or feedback from an MCP server, for example, can help guide a coding agent towards an appropriate goal. Often a far better option than allowing it to YOLO its own definition of done.

Frankenstein, the book, is epistolary, formed from a series of letters, meaning we learn the events after the failures, second or third hand. This is how we often find ourselves working

with agents. we may plan, specify and spark them into existence with electricity (Or galvanism (https://en.wikipedia.org/wiki/Galvanism)as the book describes it, see what I mean bout similarities), but then we learn their actual behaviour from traces and logs, and the products they produce for us. There isn't the same hands-on debugging that was a key part of good 'trad' software development. I find myself reading and analysing considerably more than coding. But this delay between actions and review is a challenge and this has been demonstrated to be risky by the unsuspecting open claw early adopters. (who ran such agents locally without protections in place and found the agents deleting all their emails etc (https://www.businessinsider.com/meta-ai-alignment-director-openclaw-email-deletion-2026-2))

Thankfully the ecosystem is already adapting, with LangSmith for example, making behaviour review much easier. But you're following the agents' footsteps tracking what your monster did, rather than holding the reins.

In summary, don't make the same mistakes Dr Frankenstein made, "agentify responsibly" and don't give them any permissions you wouldn't give to a 7ft tall monster with the risk awareness and emotional maturity of a toddler. But most of all, don't forget the results could be doing things you and your employer might not like, and it's best to investigate and test what you produce. If you need help with that – get in touch (https://www.linkedin.com/in/peter-houghton-374a36/).