

A Thought Leader's Perspective: Power and Responsibility

Of all the technology cycles I've witnessed over the past four decades—from the birth of the personal computer and the web, to the revolutions in mobile and cloud—none has felt quite like this one. For years, the discourse around Artificial Intelligence was a familiar rhythm of hype and disillusionment, the so-called “AI summers” followed by long, cold winters. But this time, something is different. The conversation has palpably shifted. If the last eighteen months were about the engine—the breathtaking, almost vertical ascent of Large Language Models (LLMs)—the next era will be about the car we build around it. It will be about the frameworks that harness this raw power, transforming it from a generator of plausible text into a true agent of action.

I admit, I began as a skeptic. Plausibility, I've found, is often inversely proportional to one's own knowledge of a subject. Early models, for all their fluency, felt like they were operating with a kind of impostor syndrome, optimized for credibility over correctness. But then came the inflection point, a step-change brought about by a new class of "reasoning" models. Suddenly, we weren't just conversing with a statistical machine that predicted the next word in a sequence; we were getting a peek into a nascent form of cognition.

The first time I experimented with one of the new agentic coding tools, I felt that familiar spark of magic. I tasked it with a personal project I'd never found the time for: migrating a charity website from a simple web builder to a proper, modern CI/CD environment. For the next twenty minutes, it went to work, asking clarifying questions, requesting credentials, and providing status updates. It felt less like using a tool and more like collaborating with a junior developer. When it presented me with a fully deployable package, complete with impeccable documentation and unit tests, I was floored.

Of course, it wasn't perfect. It made mistakes. It got stuck. It required my supervision and, crucially, my judgment to steer it back on course. The experience drove home a lesson I've learned the hard way over a long career: you cannot afford to trust blindly. Yet, the process was fascinating. Peeking into its "chain of thought" was like watching a mind at work—messy, non-linear, full of starts, stops, and self-corrections, not unlike our own human reasoning. It wasn't a straight line; it was a random walk toward a solution. Here was the kernel of something new: not just an intelligence that could generate content, but one that could generate a *plan*.

This is the promise of agentic frameworks. It's the difference between a static subway map and a dynamic GPS that reroutes you in real-time. A classic rules-based automaton follows a fixed path; when it encounters an unexpected obstacle, it breaks. An AI agent, powered by a reasoning model, has the potential to observe, adapt, and find another way. It possesses a form of digital common sense that allows it to navigate the countless edge cases of reality. It

represents a shift from simply telling a computer *what* to do, to explaining *why* we need something done and trusting it to figure out the *how*.

As exhilarating as this new frontier is, it brings a profound sense of responsibility, particularly from my vantage point as the CIO of a global financial institution. The stakes are immeasurably high. An agent that makes a mistake while creating a recipe for a "Chicken Salmon Fusion Pie" is a fun anecdote. An agent that makes a mistake while executing a trade, managing risk, or handling client data is a real problem. I've read the disclaimers and the cautionary tales: the web automation agent that, after failing a login, decided to email a member of parliament to complain about login walls. It's a darkly humorous reminder that we are dealing with a technology we don't fully understand.

This is where craft, culture, and a relentless focus on our principles become our essential guide. Our Engineering Tenets are not just words on a page; they are our compass. We must *Build with Purpose*, ensuring that every agent we design starts from a clear understanding of the client problem we are solving. We must *Look Around Corners*, anticipating failure modes and designing systems that are resilient by design. And above all, we must *Inspire Trust*, by being transparent about our methods and accountable for our outcomes.

In an agentic world, these tenets take on new urgency. The hard truth is that you cannot simply overlay these powerful new tools onto messy, inconsistent systems and expect good results. Messy systems plus agents are a recipe for disaster. An AI trained on "garbage" data doesn't just produce garbage-out; it produces plausible, confident garbage that can poison an entire process. Therefore, our first and most critical task is to prepare the ground. We must invest in clean data, consistent metadata, and well-defined APIs. We have to build the modern "interstate system" that allows these agents to operate safely and at high velocity. It is the hard, foundational work of building a programmable enterprise, an "enterprise as software," where our processes are as well-architected as our code.

Ultimately, this journey is not about replacing human ingenuity, but about augmenting it. It demands a new set of skills from all of us: the ability to explain a task with clarity, the wisdom to delegate, and the diligence to verify the quality of the output. It requires us to be humble, to acknowledge what we don't know, and to never stop learning. The pages that follow in this book offer a technical map for building these new frameworks. My hope is that you will use them not just to build what is possible, but to build what is right, what is robust, and what is responsible.

The world is asking every engineer to step up. I am confident we are ready for the challenge.

Enjoy the journey.

Marco Argenti, CIO, Goldman Sachs