

Secure Software Engineering in a GenAI World

Bert Lagaisse (DistriNet, Groep T, KU Leuven)

Coen De Roover (SOFT, VUB)

Ingmar Malfait (EA-ICT, Groep T, KU Leuven)

Koen Handekyn (Nabutech)



Secure Software Engineering in a GenAI World

Kick-off Agenda – 19 May

- 14:00 Coffee & Welcome
- 14:30 Intro to LLM-based Code Assistants – Bert Lagaisse
- 14:50 Vibe Coding a game with Cursor and Claude – Ingmar Malfait
- 15:10 Agentic AI: Harness Engineering for Coding – Koen Handekyn
- 15:30 Code Smells in AI-supported workflows – Coen De Roover
- 16:00 Security Risks: Myths, Memes and Scientific Evidence – Bert Lagaisse
- 16:15 Recent Security Findings from empirical studies – Coen De Roover
- 16:45 Project Planning
- 17:00 Reception & Networking

GenAI ?

Generative models at work...

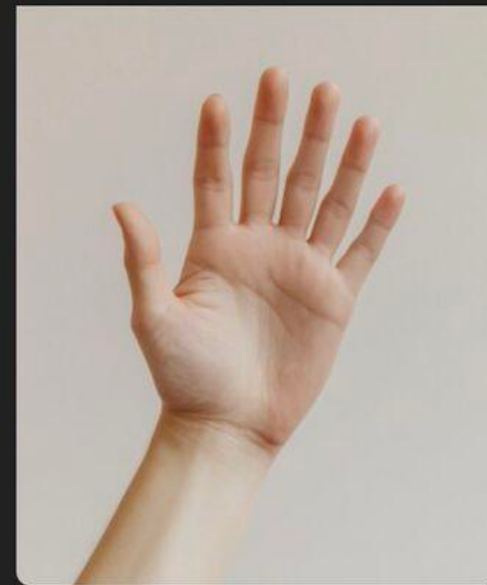
Generate a 2D game character which is an Italian plumber wearing a hat.

Sure, here you go.



ChatGPT o1

Share



Is there anything wrong with this photo?



Thought about photo analysis for a couple of seconds

It appears to be a fairly ordinary photograph of a human hand, presented palm-forward with fingers extended. The lighting seems natural and even, and the background is a neutral off-white. From a visual standpoint, there's nothing overtly "wrong" with the image: it's well-centered, in focus, and the hand looks normal with no immediately noticeable deformities or injuries.

Of course, without context or a specific standard you're measuring against (such as a photography quality standard, an artistic criterion, or a concern about some medical anomaly), it's hard to say if anything is amiss. Purely as a casual viewer, the photo appears completely fine.

Hallucinations

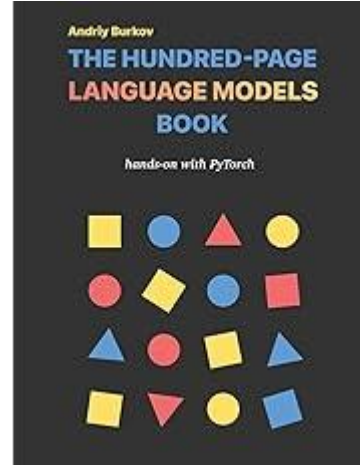
who is noble ackerson

 Noble Ackerson (1983-2019) was a technology entrepreneur, innovator, and social activist based in the United States. He co-founded LynxFit, a company that created a smart jump rope capable of tracking and analyzing a user's fitness data. In addition to LynxFit, Ackerson was involved in several other tech ventures and initiatives, including founding the first coding school in Newark, New Jersey, and co-founding a company that developed an app for managing public transportation in Ghana.

Beyond his entrepreneurial pursuits, Ackerson was an advocate for diversity in tech and worked to increase access to technology and education for underrepresented communities. Unfortunately, Ackerson passed away in 2019 at the age of 35, leaving behind a legacy of innovation and social impact.

 False information  Fact

Image Credit : Towards Data Science



Andriy Burkov: “The 100 page ML Book”. “The 100 page LM Book”.

The 100 Page LM Book.
PhD in AI
Major AI influencer and critic.

- Right now, hundreds of millions of people are learning technical and scientific concepts from general-purpose chat LLMs, which are ~40% wrong. This will start manifesting itself in concrete technogenic disasters and overall progress decay 10–15 years from now.

OpenAI admits AI hallucinations are mathematically inevitable, not just engineering flaws

News

Sep 18, 2025 • 6 mins

In a landmark study, OpenAI researchers reveal that large language models will always produce plausible but false outputs, even with perfect data, due to fundamental statistical and computational limits.

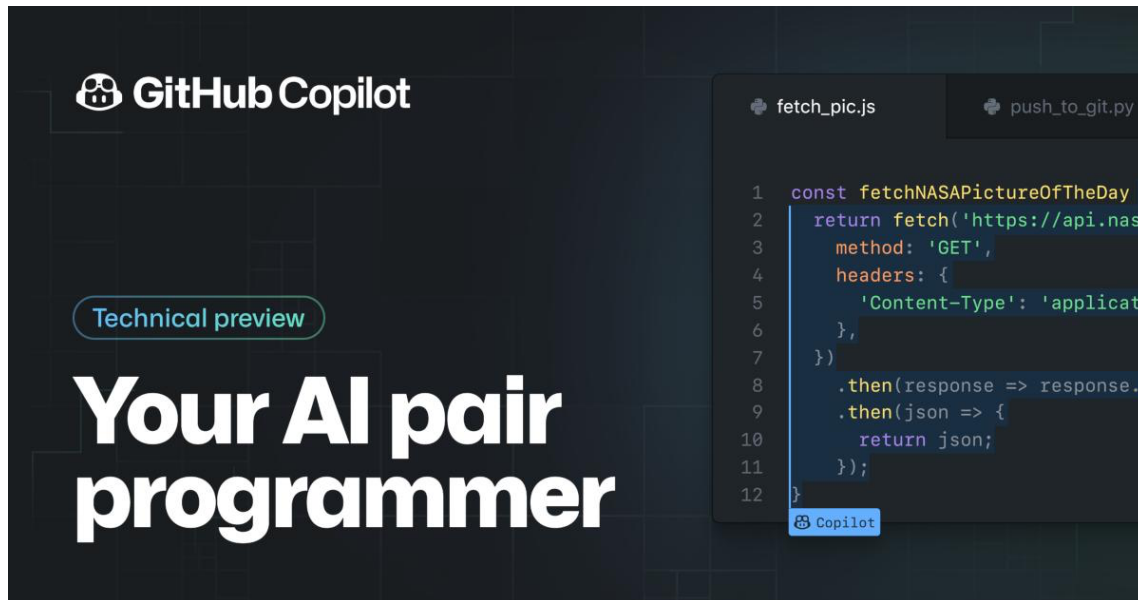
OpenAI, the creator of ChatGPT, acknowledged in its own research that large language models will always produce hallucinations due to fundamental mathematical constraints that cannot be solved through better engineering, marking a significant admission from one of the AI industry's leading companies.

“Like students facing hard exam questions, large language models sometimes guess when uncertain, producing plausible yet incorrect statements instead of admitting uncertainty,” the researchers wrote in the paper. “Such ‘hallucinations’ persist even in state-of-the-art systems and undermine trust.”

AI-assisted Software Engineering

GitHub Co-Pilot

- AI Pair programmer will not keep you away from procrastination ;-)



ChatGPT

- ChatGPT outperforms 85% of the 4m programmers assessed on Python on LinkedIn.



Tool	Category	Key Strengths (2026)
GitHub Copilot	AI pair programmer	Multi-model (Anthropic, OpenAI, Google), premium requests for chat/agent mode, strongest IDE ecosystem
ChatGPT (OpenAI)	General reasoning + coding	GPT-5.2/5.3 models, strong code interpreter, Pro tiers now differentiated for Codex usage
Claude (Anthropic)	Deep reasoning & multi-file coding	Claude Code, Projects, Opus 4.6, Max 5×/20× heavy-usage tiers
Cursor IDE	AI-first IDE	Multi-model (GPT/Claude/Gemini), credit-based usage, full agent mode, Composer multi-file edits
Tabnine	Privacy-first code completions	On-prem, customizable models, enterprise governance, strong secure deployments
Amazon Q Developer	AWS-integrated coding assistant	Agentic tasks, code transformation (Java upgrades), AWS resource Q&A
Gemini Code Assist	Google Cloud–integrated assistance	Agent mode, multi-file edits, deep GCP integrations (BigQuery, Firebase)
Replit AI / Replit Agent	Cloud IDE + agents	Autonomous feature building, app deployment, strong credits-based AI model pool
Continue.dev	Open-source IDE assistant	Bring-your-own model, multi-file edits, open-source control
Cline	Open-source autonomous agent	Full agentic workflows, plan/act mode, BYO API keys, free core extension

Category	Description	Representative Tools & Models (non-exhaustive)
1. Code-Completion Assistants	• Provide inline, real-time suggestions while typing. • Operate at the level of lines or blocks of code. • The developer fully controls structure and flow.	GitHub Copilot (completion mode), Amazon Q Developer / CodeWhisperer, JetBrains AI Assistant, Codestral, DeepSeek-Coder, CodeLlama, Continue.dev (autocomplete mode)
2. Chat-Based Coding Assistants	• Conversational tools that answer questions, explain code, generate snippets or debug logic on request. • Developer decides what to accept or apply.	Claude (chat), ChatGPT, GitHub Copilot Chat, Amazon Q Developer Chat, Google Gemini Code Assist, Continue.dev (chat mode)
3. Agentic / Autonomous Coding Assistants	• Agents plan and execute multi-step software engineering tasks. • Can modify multiple files, run tools, evaluate results and iterate. • Developer supervises rather than micromanages.	Cursor (agent / vibe-coding mode), Claude Code, OpenClaw, Copilot coding agent, Continue.dev + tools, Qwen3-Coder in agent frameworks,

What is Agentic AI?

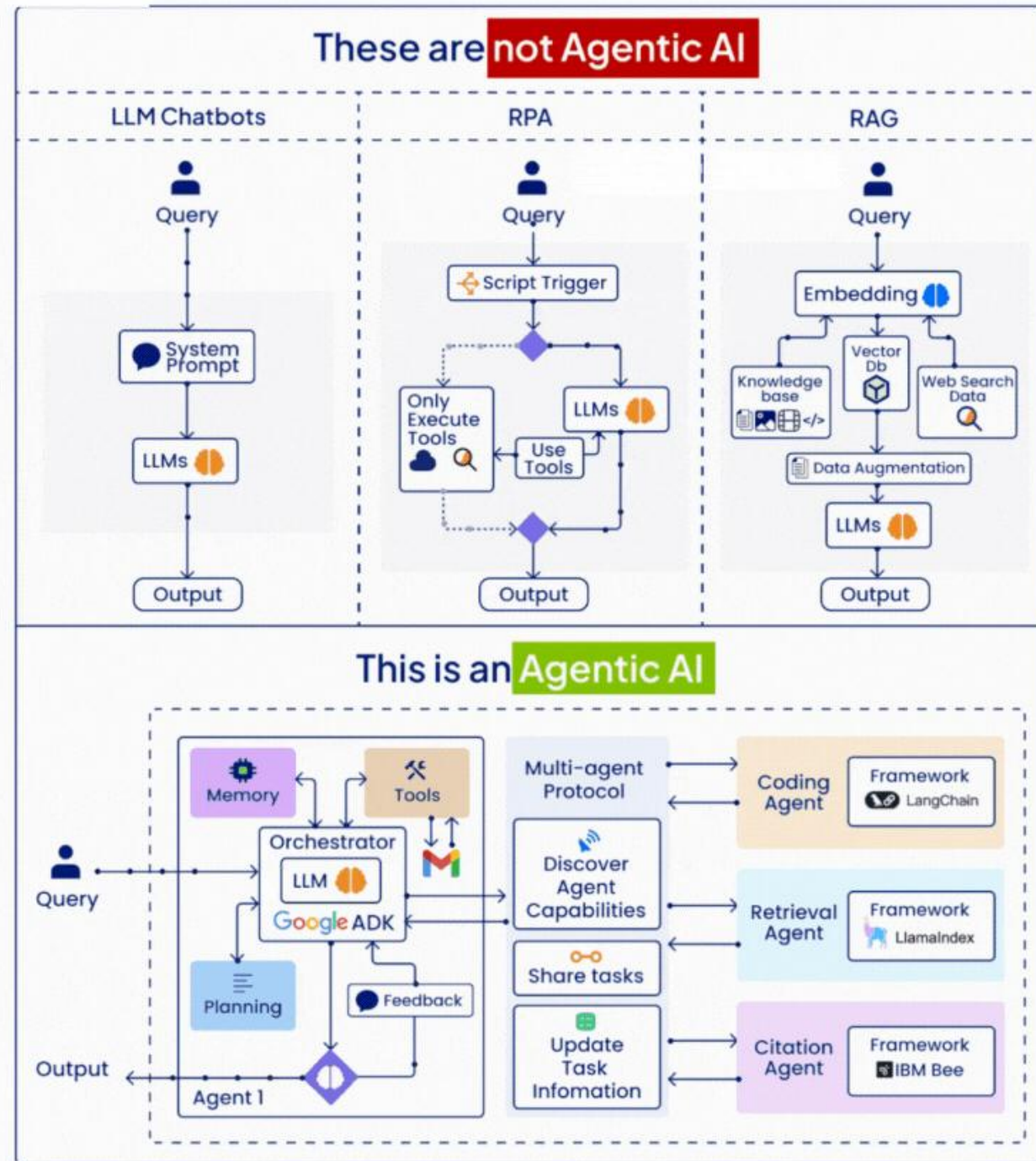
Autonomous Task Execution. Agentic AI agents perform complex tasks autonomously by reasoning about goals and breaking/planning them into subtasks.

Multi-Agent Collaboration. Specialized agents collaborate and orchestrate workflows to achieve a unified objective efficiently.

Workflow Automation. Agentic AI automates workflows including project setup, debugging, testing, and continuous integration.

Context Sharing Protocols. Agents share context using protocols like Model Context Protocol to coordinate actions effectively.

Agentic AI skills are reusable, domain-specific instructions. Teach AI agents how to execute complex, multi-step workflows. Instead of typing out long prompts repeatedly, containing procedural knowledge, scripts, and context



vibe cod·ing

/ˈvīb ,kōdɪŋ/ · noun, verb, vibe.

"You fully give in to the vibes, embrace exponentials, and forget that the code even exists."

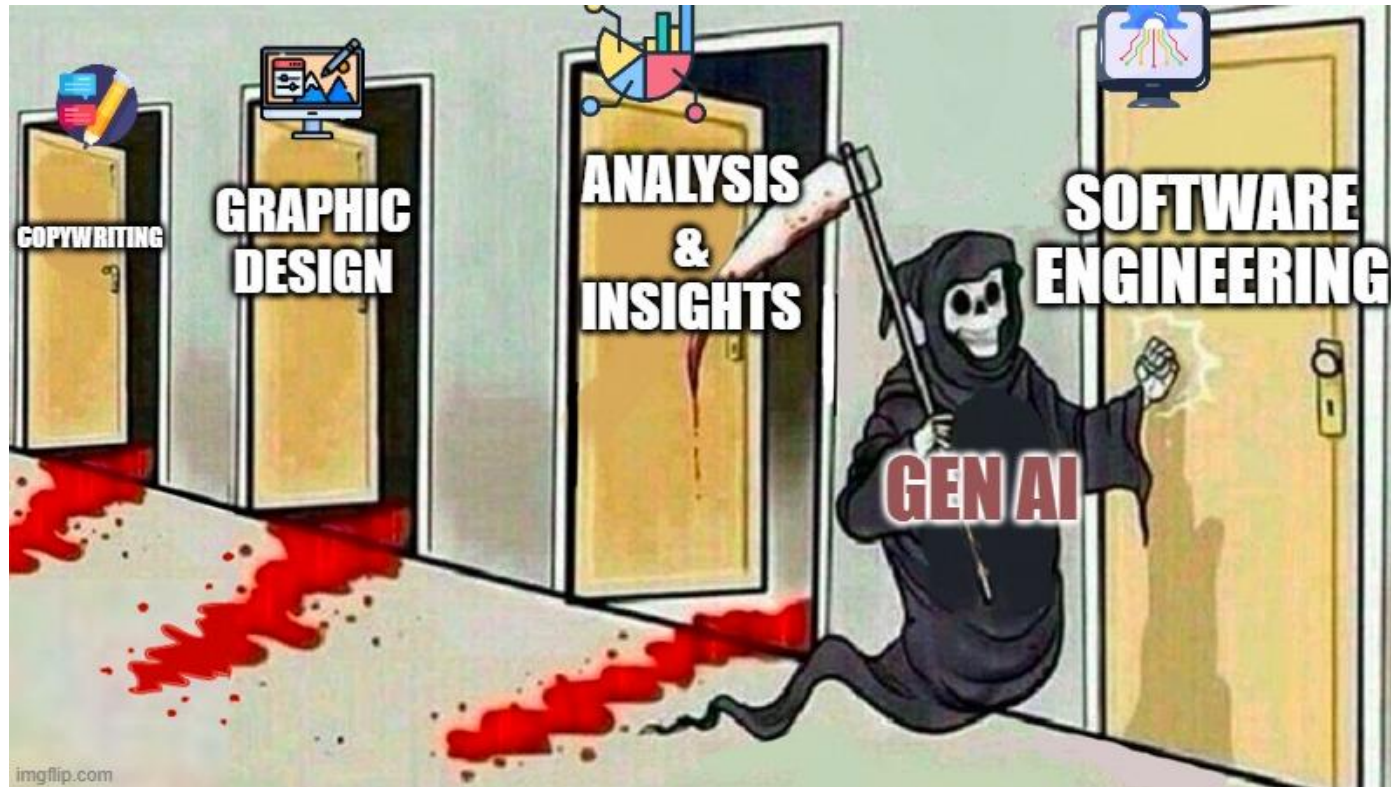
— Andrej Karpathy, February 2025

A software practice where you describe intent in plain English to an LLM, the model writes the code, and you nudge the output until it matches your vision.

TIMELINE

- Feb 2025** Karpathy tweets the phrase.
- Mar 2025** Merriam-Webster lists it as slang.
- Apr 2025** Cursor crosses 1M paid users.
- Q3 2025** Vibe coding hits enterprise pitches.
- Nov 2025** Collins names it Word of the Year.
- 2026** The post-mortem articles arrive.

"There is a world before and after Claude Code"



Disruption

Entire software development organisation fundamentally changes.

Steven Latre (columnist at De Tijd || IMEC VP AI):

"Software is the canary in the coal mine. With a revolution like this, various sectors will soon look completely different."

Andrew Ng (DeepLearning.AI):

"If you've never written code before, this is an incredible time to start. AI coding tools mean barriers that once stopped millions are now falling."

Low-Code is Losing the Battle

"Low-code is someone else's code that is not very flexible." - Maria Vechtomova

LOW-CODE (5 years ago)

Pro: Faster to build
Pro: Lower learning curve
Pro: Non-coders can ship

Con: Not flexible
Con: Boxed into abstractions
Con: Expensive license
Con: Vendor lock-in

Low-code is someone else's code that is not very flexible



AI CODING (today)

Pro: Faster to build
Pro: Lowest learning curve ever
Pro: Non-coders CAN code
Pro: Full flexibility
Pro: No license fee

"I don't see how low-code platforms survive."

Towards a synthetic bloated mess ?

Towards a new role: Vibe Coding Cleanup Specialist?

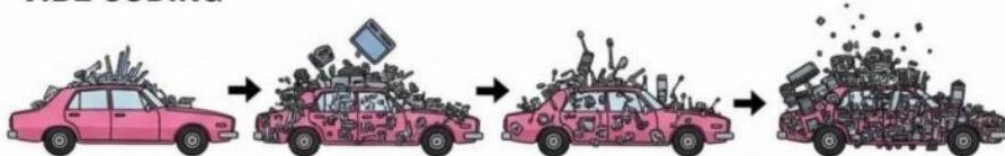
WATERFALL



AGILE



VIBE CODING



Jonathan Lin • 2°

Overengineering & Vibe Coding Cleanup Specialist

Territorio Federal de Kuala Lumpur, Malaysia

1 mil seguidores • Denis Lin es contacto en común



Hamza Gulraiz • 2°

Vibe coding cleanup specialist | React Native Developer | Mobil
Lahore

Ofrece servicios: Desarrollo de aplicaciones móviles, Desarrollo web,
de software, Desarrollo de Android



Nishant Gohel • 2°

Vibe coding cleanup specialist

Bengaluru

22 mil seguidores • Mani Kiran Nalabolu (He/Him) y Muhammad Fa
común



Wilmer Mauricio Herrera Renteria • 2°

Vibe Coding Fixer | Software Architect | MSc in AI

Quito

Anterior: SugarCRM Developer **Specialist** en Casabaca S.A.

Ofrece servicios: Desarrollo web



Victor Istrati • 2°

Senior Web Engineer Vibe Coding Cleanup Specialist

Chisinau, Moldavia

Actual: **Vibe Coding Cleanup Specialist** en Freelance

Victoria Negara es contacto en común

2026: The Year of Technical Debt?

A financial services firm went from 25,000 to 250,000 lines of code per month.

Result: 1 MILLION lines of code backlogged for review. Nobody has time.

What's happening:

- AI produces 10x more code - review capacity unchanged
- "Bad code - AI or human - can bloat and destabilize software and cause security flaws"
- New job emerging: **"Vibe Coding Cleanup Specialist"**
→ someone who fixes what the AI left behind
- This is becoming a major cost center for enterprises

The Effects of AI-Generated Code Tearing Through Corporations Is Actually Kind of Funny

Corporations are sinking in a sea of code that needs review



"Speed Can Kill" [1]

"Speed at the Cost of Quality" [2]

[1] Keynote by Alexander Serebrenik at International Conference on Software Engineering, 2026

[2] How Cursor AI Increases Short-Term Velocity and Long-Term Complexity in Open-Source Projects (He et al.), MSR2026

THE GOOD

- Statistically significant boost in dev velocity
- Cursor users ship more features faster
- Real, measurable productivity gains

THE BAD

- Substantial increase in static analysis warnings
- Persistent increase in code complexity
- Slowdown in the long run
- QA becomes the critical bottleneck

3–5×

lines added in month one

+30%

static analysis warnings

+41%

code complexity

QA must be a first-class concern in AI coding workflows - not an afterthought.