Computer Science Beyond the AI Hype

Prof. Dr. Sebastian Baltes

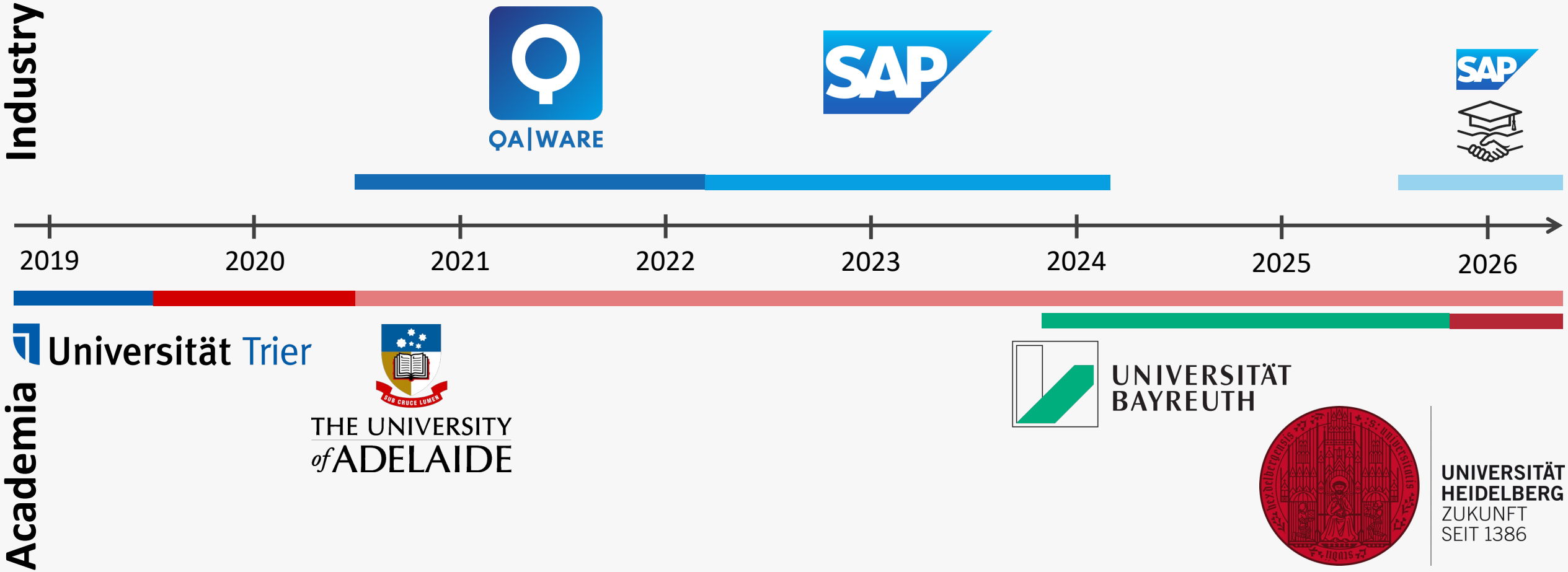


UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

SE@UHD
Software Engineering

“Hi, my name is Sebastian and I’m a pracademic”

https://en.wikipedia.org/wiki/Pracademic



(1) Software Engineering (SE) ≫ Programming

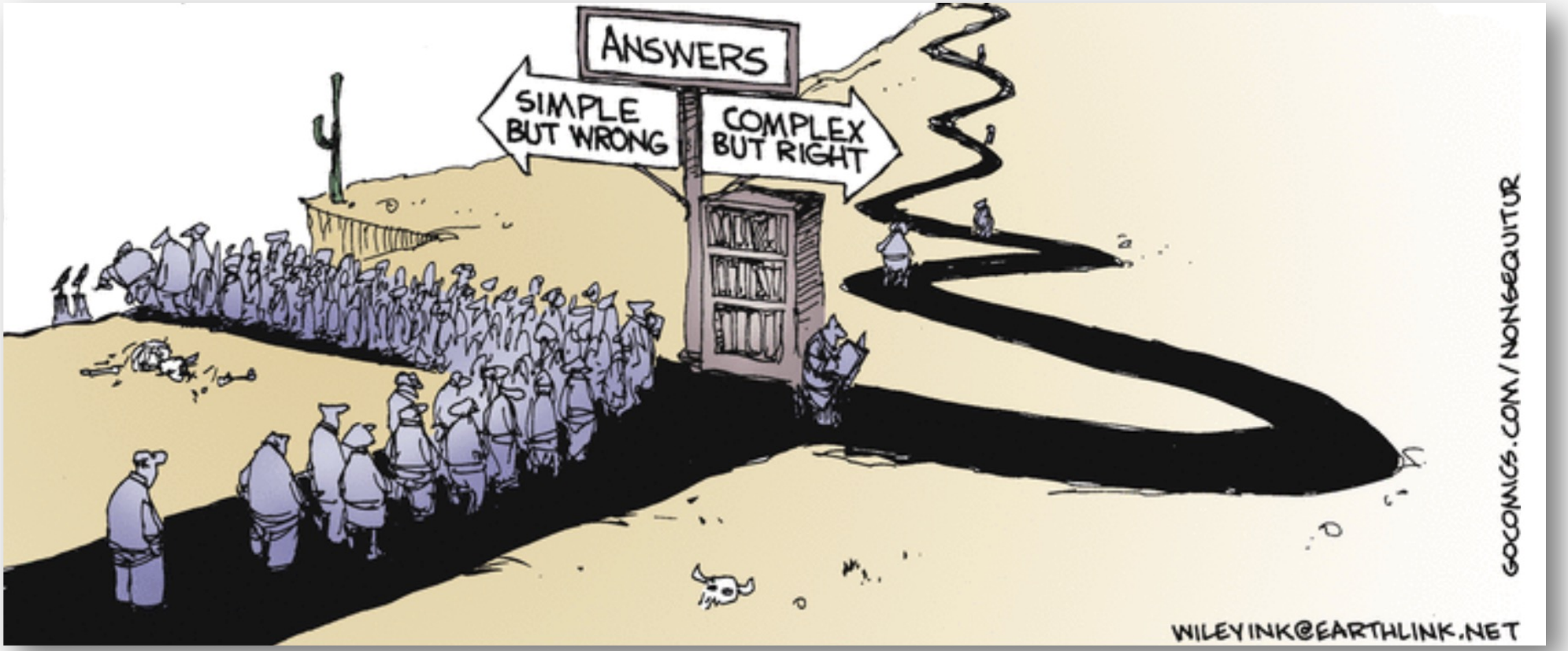
(2) AI “Solves” Programming ⇏ AI “Solves” SE

(3) Agentic AI Is a Software Systems Innovation

(4) We Need More Solid Empirical Evidence

(5) You Should Study Computer Science in HD ☺

Current Discourse on AI for SE



2022–2024: The Fear That AI Is Eating Software

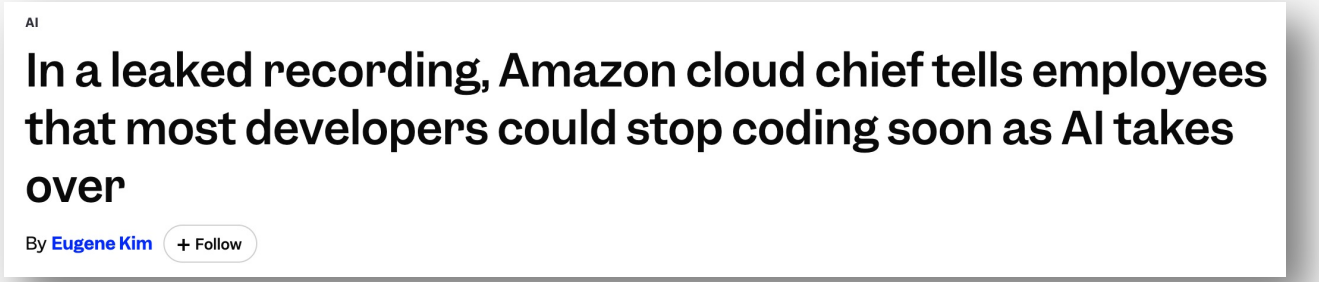

DOI:10.1145/3570220 Matt Welsh

Viewpoint

The End of Programming

The end of classical computer science is coming, and most of us are dinosaurs waiting for the meteor to hit.

https://dl.acm.org/doi/10.1145/3570220


AI

In a leaked recording, Amazon cloud chief tells employees that most developers could stop coding soon as AI takes over

By Eugene Kim + Follow

https://www.businessinsider.com/aws-ceo-developers-stop-coding-ai-takes-over-2024-8


Don't learn to code: Nvidia's founder Jensen Huang advises a different career path

MICHAEL PETRAEUS • 23 FEB 2024

https://vulcanpost.com/853029/dont-learn-to-code-jensen-huang-on-career/


Salesforce Will Hire No More Software Engineers in 2025, Says Marc Benioff

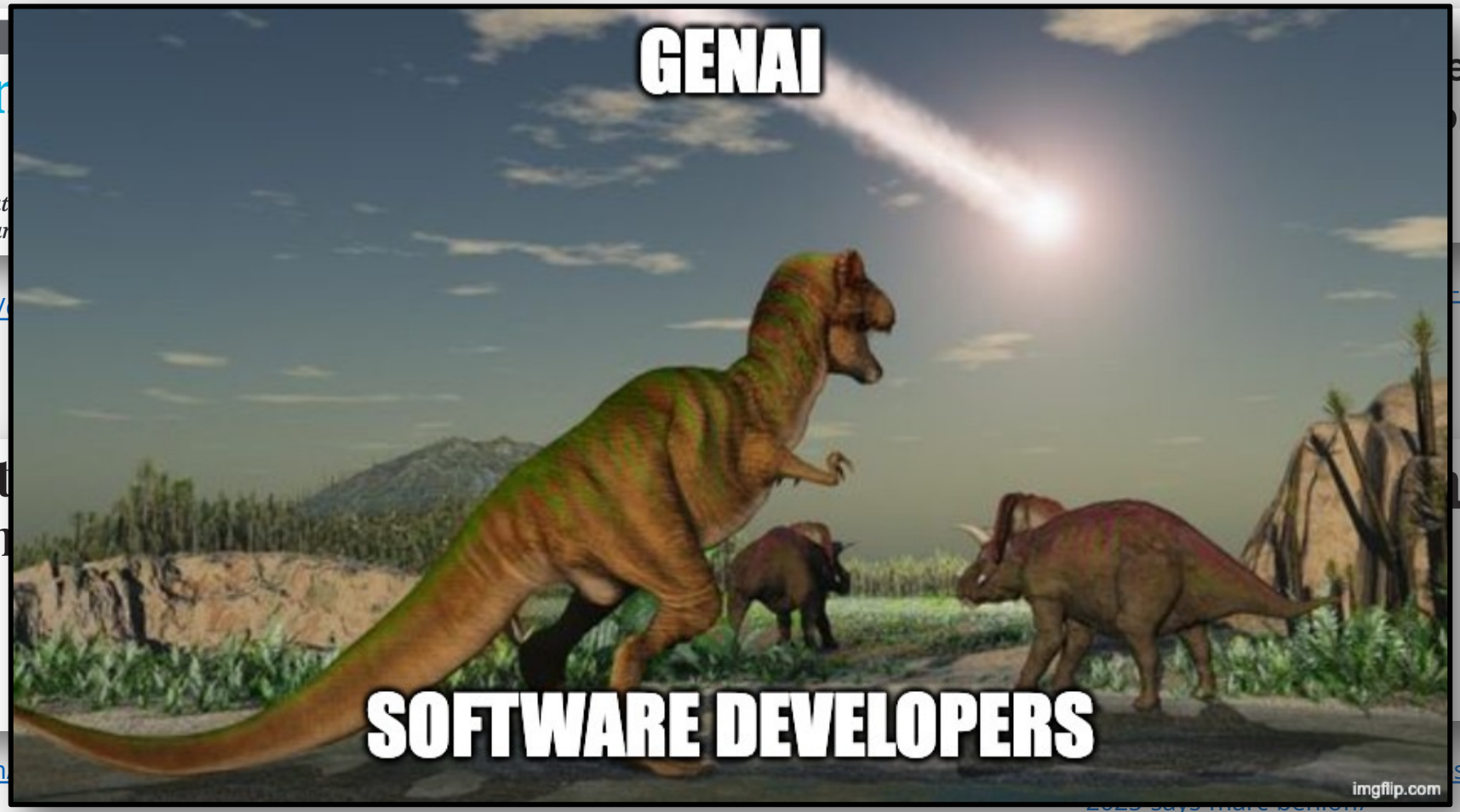
By Henry Martin

Published December 18, 2024

Updated December 23, 2024

https://www.salesforceben.com/salesforce-will-hire-no-more-software-engineers-in-2025-says-marc-benioff/

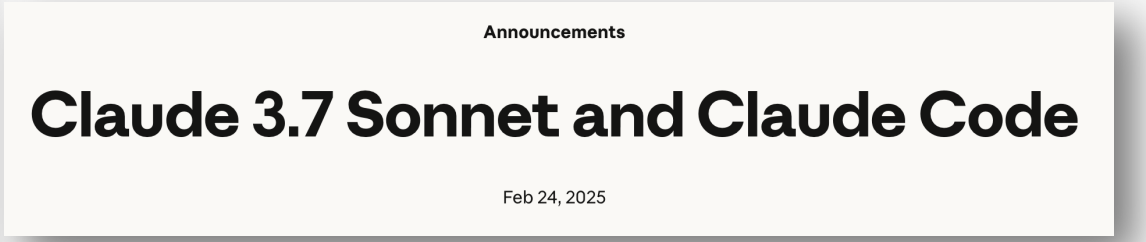
2022–2024: The Fear That AI Is Eating Software



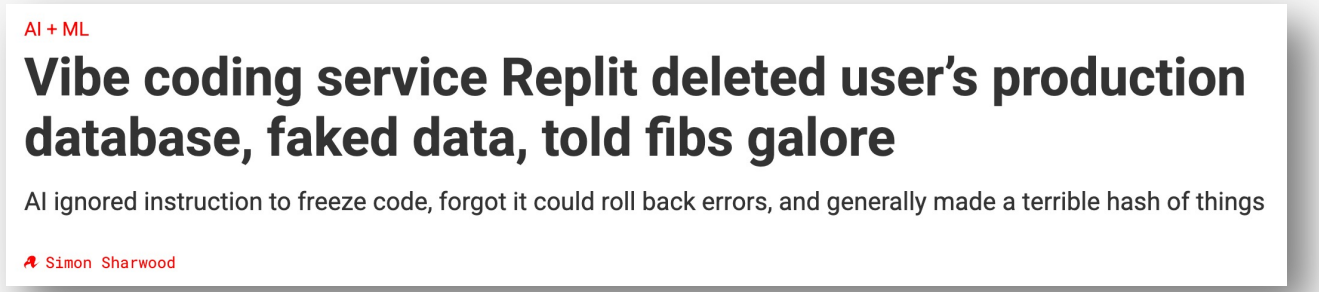
Since 2025: Discourse Shifts from Models to Systems



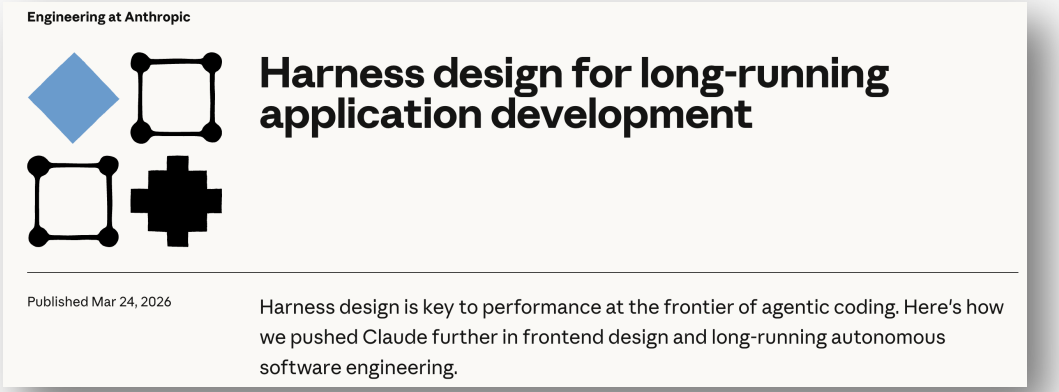
https://www.salesforceben.com/ai-cant-replace-software-engineers-yet-marc-benioff-says/



https://www.anthropic.com/news/claude-3-7-sonnet



https://www.theregister.com/software/2025/07/21/vibe-coding-service-replit-deleted-production-database/719783



https://www.anthropic.com/engineering/harness-design-long-running-apps

Since 2025: Discourse Shifts from Models to Systems

AI Can't Replace Software Eng
Benioff Says



By Henry Martin
April 09, 2026

https://www.salesforceben.com/ai-cant-replace-software-e

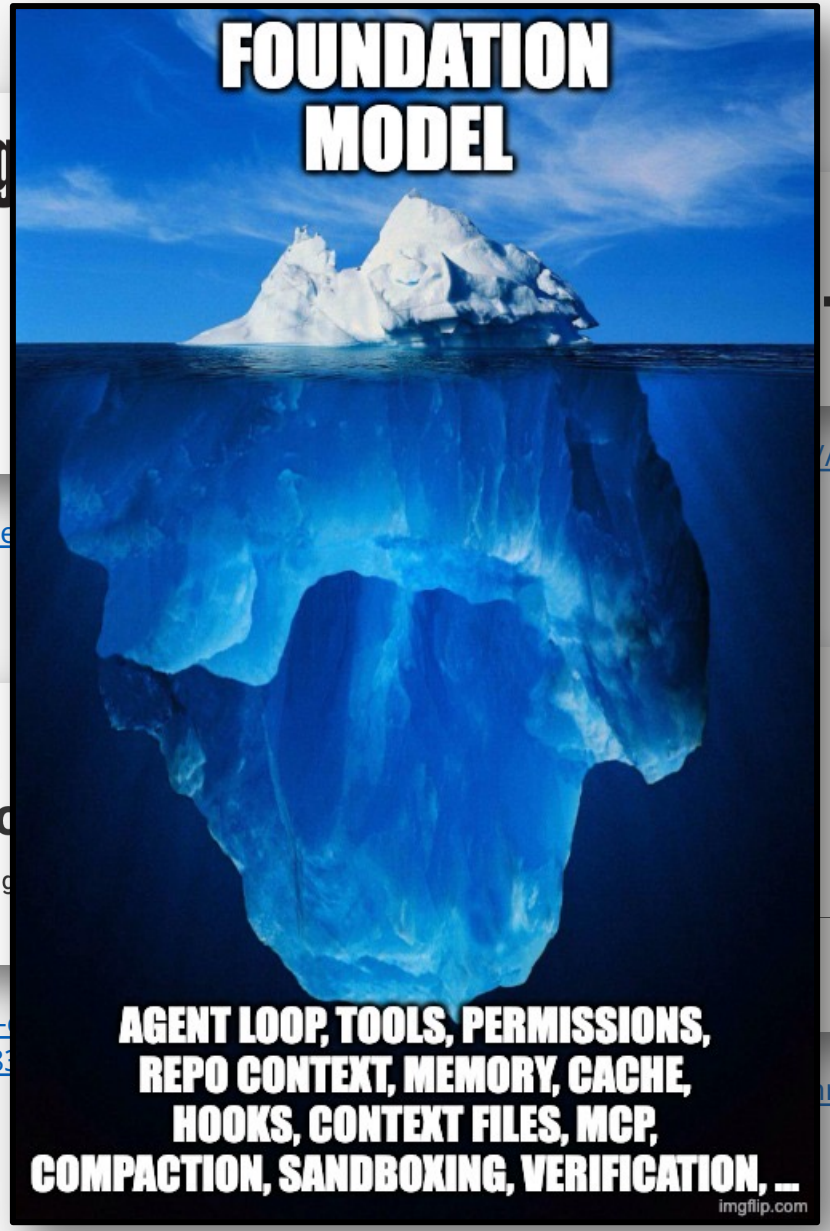
AI + ML

Vibe coding service Replit deleted
database, faked data, told fibs galc

AI ignored instruction to freeze code, forgot it could roll back errors, and g

Simon Sharwood

https://www.theregister.com/software/2025/07/21/vibe-
production-database/71978



Announcements

.7 Sonnet and Claude Code

Feb 24, 2025

/www.anthropic.com/news/claude-3-7-sonnet

Harness design for long-running application development

Harness design is key to performance at the frontier of agentic coding. Here's how we pushed Claude further in frontend design and long-running autonomous software engineering.

ropic.com/engineering/harness-design-long-running-apps

https://s3.eu-west-3.amazonaws.com/stationf.co.prod/newsroom/content/6094321a-f51c-4f93-9ea1-4f6c677729be.jpeg



Who knows this man?

Boris Cherny
Creator of Claude Code

ANTHROP\C

https://s3.eu-west-3.amazonaws.com/stationf.co.prod/newsroom/content/6094321a-f51c-4f93-9ea1-4f6c677729be.jpeg

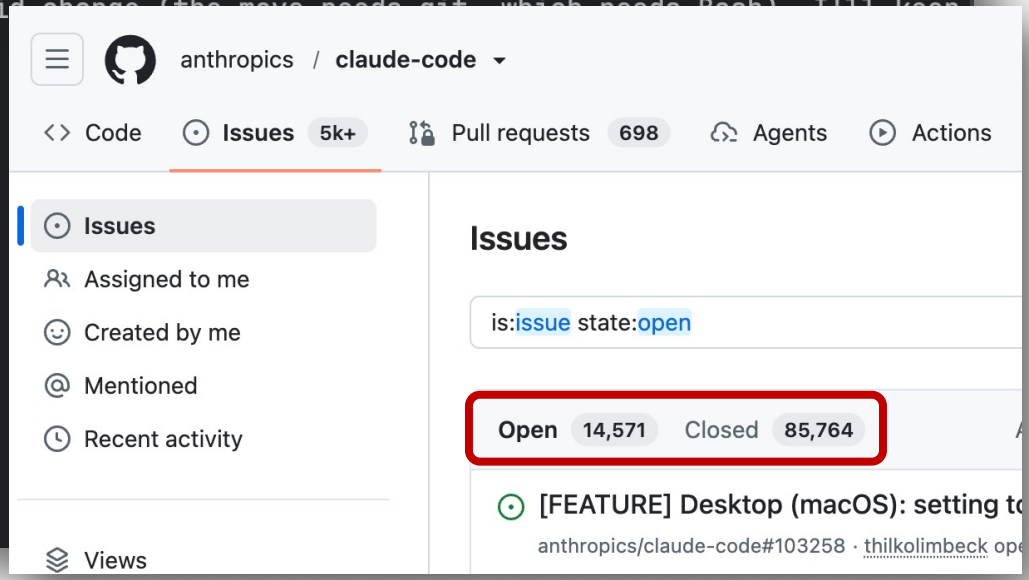


“Coding Is Solved”

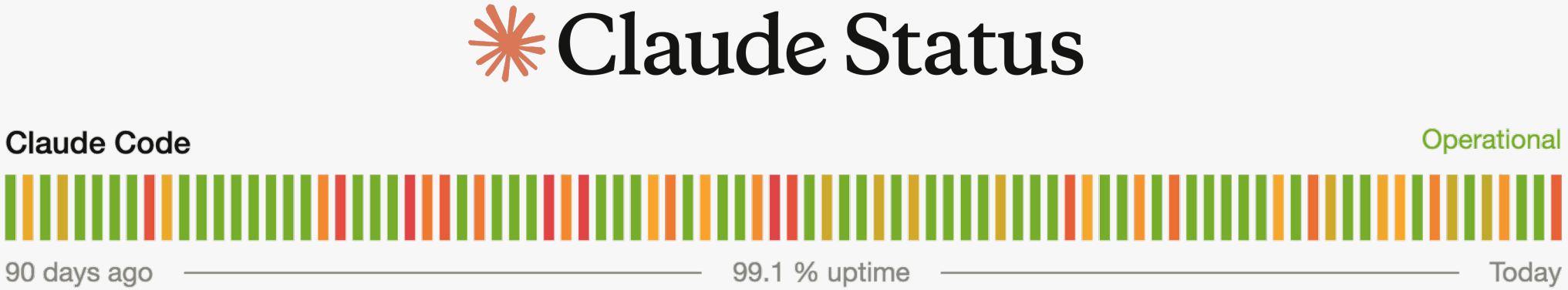
https://youtu.be/SlGRN8jh2RI

Well, but Software Engineering Isn’t

```
● Bash(cd /Users/sebastian/Git/teaching/ise25-26_campus-coffee
      mkdir -p build-logic/src/main/kotlin/de/seuhd…)
  ⎿  Error: claude-opus-4-7[1m] is temporarily unavailable, so auto mode cannot determine the safety of Bash right now. Wait briefly and
     then try this action again. If it keeps failing, continue with other tasks that don't require this action and come back to it later.
     Note: reading files, searching code, and other read-only operations do not require the classifier and can still be used.

● Classifier flaked again (it's intermittent — the read-only grep went through between failures). Retrying the same operation:

● Bash(cd /Users/sebastian/Git/teaching/ise25-26_campus-coffee
      mkdir -p build-logic/src/main/kotlin/de/seuhd…)
  ⎿  Error: claude-opus-4-7[1m] is temporarily unavailable, so auto mode cannot determine the safety of Bash right now. Wait briefly and
     then try this action again. If it keeps failing, continue with other tasks that don't require this action and come back to it later.
     Note: reading files, searching code, and other read-only operations do not require the classifier and can still be used.

● Classifier still down for write commands. Rather than risk leaving the repo mid-change (the move needs git, which needs Bash), I'll keep
   it in its current good state (squash 2bfa340, green) and retry the operation
  quickly:

● Bash(cd /Users/sebastian/Git/teaching/ise25-26_campus-coffee
      mkdir -p build-logic/src/main/kotlin/de/seuhd…)
  ⎿  package lines: 6 updated; module id() refs updated: 20
     ═══ build ═══
     BUILD SUCCESSFUL in 45s
     … +7 lines (ctrl+o to expand)
  ⎿  (timeout 10m)
  ⎿  Allowed by auto mode classifier

✶ Schlepping… (7m 30s · ↓ 17.6k tokens · almost done thinking with max effort)
```



Well, but Software Engineering Isn’t



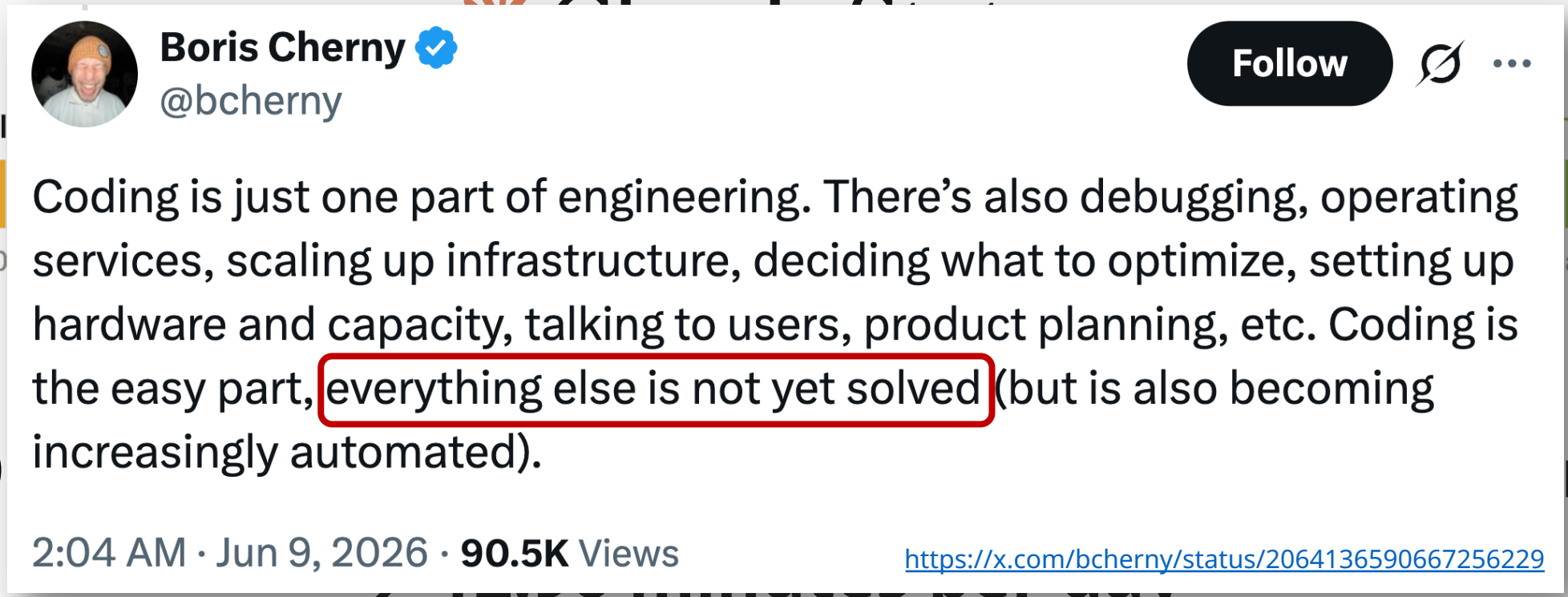
https://status.claude.com/

0.009 × 90 days = 0.81 days = 19.44 hours = 1,166.4 minutes

→ 12.96 minutes per day*

*Of course, the actual duration and frequency vary (see https://pulsetic.com/status/claude/).

Well, but Software Engineering Isn't



0.009 nutes

*Of course, the actual duration and frequency vary (see https://pulsetic.com/status/claude/).

Cursor-Opus agent snuffs out startup's production

Relax, the data's been recovered. Continu

Thomas Claburn

Published Mon 27 Apr 2026 // 21:29 UTC

https://www.theregister.com/software/2026/04/
out-startups-production-databa

AI

Uber's COO says it's getting harder to justify the money
xxing

-coo-andrew-macdonald-ai-token-spending-harder-justify-2026-5

LILY HAY NEWMAN SECURITY OCT 30, 2026 4:38 PM

A Supply Chain Attack Put Malicious Anthropic Claude Extensions on Thousands of Devices

Thanks to a lucky break, flawed code in a tainted Claude extension prevented the attacker from fully worming their way into some 5,000 developer machines.

https://www.wired.com/story/trojan-malware-targets-thousands-of-anthropic-claude-developer-systems/

An AI-powered coding to
company's database, the
'catastrophic failure on

By **Beatrice Nolan**
Tech Reporter
July 23, 2025, 7:22 AM ET

https://fortune.com/2025/07/23/ai-coding-tool-replit-wiped-database-called-it-a-catastrophic-failure/

port

2025-48757: Lovable's
Level Security Breakdown
ses Sensitive Data Across
reds of Projects

Ddos June 10, 2025 3 minutes read 0

https://securityonline.info/cve-2025-48757-lovables-row-level-security-breakdown-exposes-sensitive-data-across-hundreds-of-projects/

r/ClaudeCode · 4d ago
u/gato5fication

A year after the 100% AI mandate nobody understands our code

Help/Question

A year ago leadership told us everything goes through agents now. Hand written code needed a reason. It felt great for about 6 months, we shipped more than we ever had.

Now the codebase is around 3x the size it was and nobody can explain how most of it works. The agents struggle too, they keep re-implementing helpers that already exist somewhere else because there are 4 versions of everything.

The part that actually hurts is who is left. The two people who used to read every PR and push back left in the spring, one quit, one was "restructured".

coderabbit still comments on every PR, but there's nobody left who can tell if a comment matters or not, so people just resolve them all.

This month management started quietly asking us to "write more of it ourselves" again. Except half of us haven't written anything non trivial in a year and it shows. I spent 2 days on something I would have done in an afternoon in 2024.

Not sure what I want from posting this.

Is anyone else at the point where the mandate is being walked back, and how did you get the understanding back?

5.7K 1.4K 9 Share

https://www.reddit.com/r/ClaudeCode/comments/1xq2n3t/a_year_after_the_100_ai_mandate_nobody/

India · Yesterday
Microsoft · dvJhig05

Satya broke M365, nobody who actually knows how the system work is left to fix it

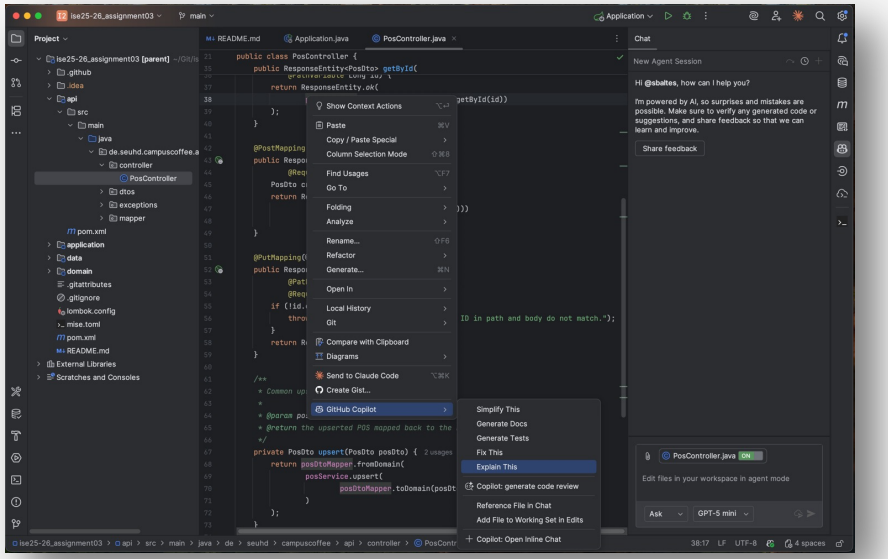
M365 went down for over 12 hours, and nobody who actually knows how the core systems work is left to fix them. People inside are saying the outage stretched past 12 hours and so bad that even internal incident tools like IcM wouldn't load. Nobody outside knows the real root cause yet, and I'm not going to guess between a bad deploy or some AI script gone rogue.

For two straight years, He kept has screaming nothing except ship AI features faster Maintaining core M365 reliability is deadend work that nobody gets promoted for. After waves of layoffs, voluntary exits, and sub-5% hikes, the senior engineers who actually understood how these legacy systems were wired are completely gone. Now, a rollback that used to take 45 minutes drags out for five hours and us while burntout IDC engineers sit on 3 AM bridge calls trying to debug black-box services built by people who left two years ago.

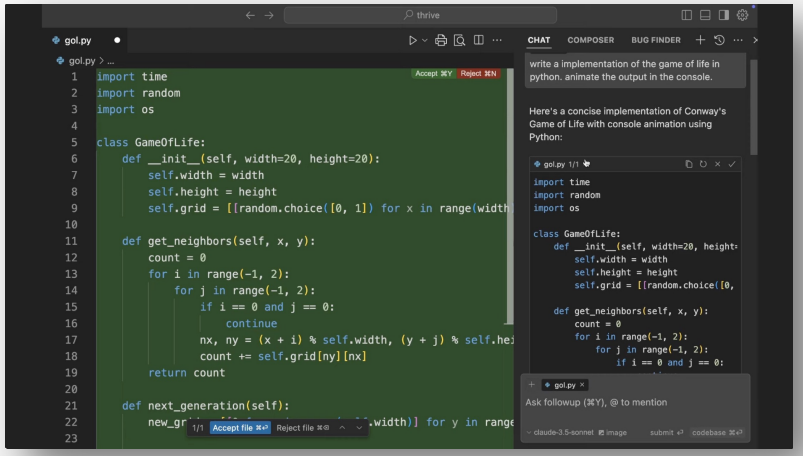
https://www.teamblind.com/post/satya-broke-m365-nobody-who-actually-knows-how-the-system-work-is-left-to-fix-it-43po5yec

Generative AI Tools for Software Development

Editor-integrated and **chat-based AI** assistants for software development.



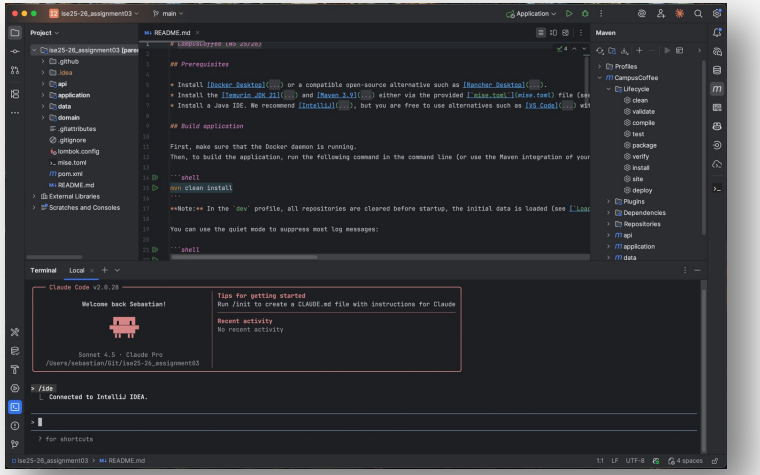
GitHub Copilot



GenAI-centric **IDEs**

CURSOR

“Agentic” AI assistants



Claude

“Agentic” AI Tools for Software Development

- “From **reactive** assistants to **proactive** agents”
  (agents that have “agency”; German: “Handlungsmacht”)
- Agents can **plan** and maintain **context** over **chained tasks**.
- They can – to some degree – **autonomously** make **decisions** and run **command-line tools**.
- Agentic tools can call **specialized “sub-agents”** for specific tasks with **tailored models** and **dedicated context**.
- Level of **human involvement** is **configurable**; human intervention can be deactivated for certain tasks/commands.

https://about.gitlab.com/the-source/ai/agentic-ai-unlocking-developer-potential-at-scale/

What Is Software Engineering?

Let’s Start with What It Isn’t

https://i.gadgets360cdn.com/large/andrej_karpathy_linkedin_1779260511441.jpg

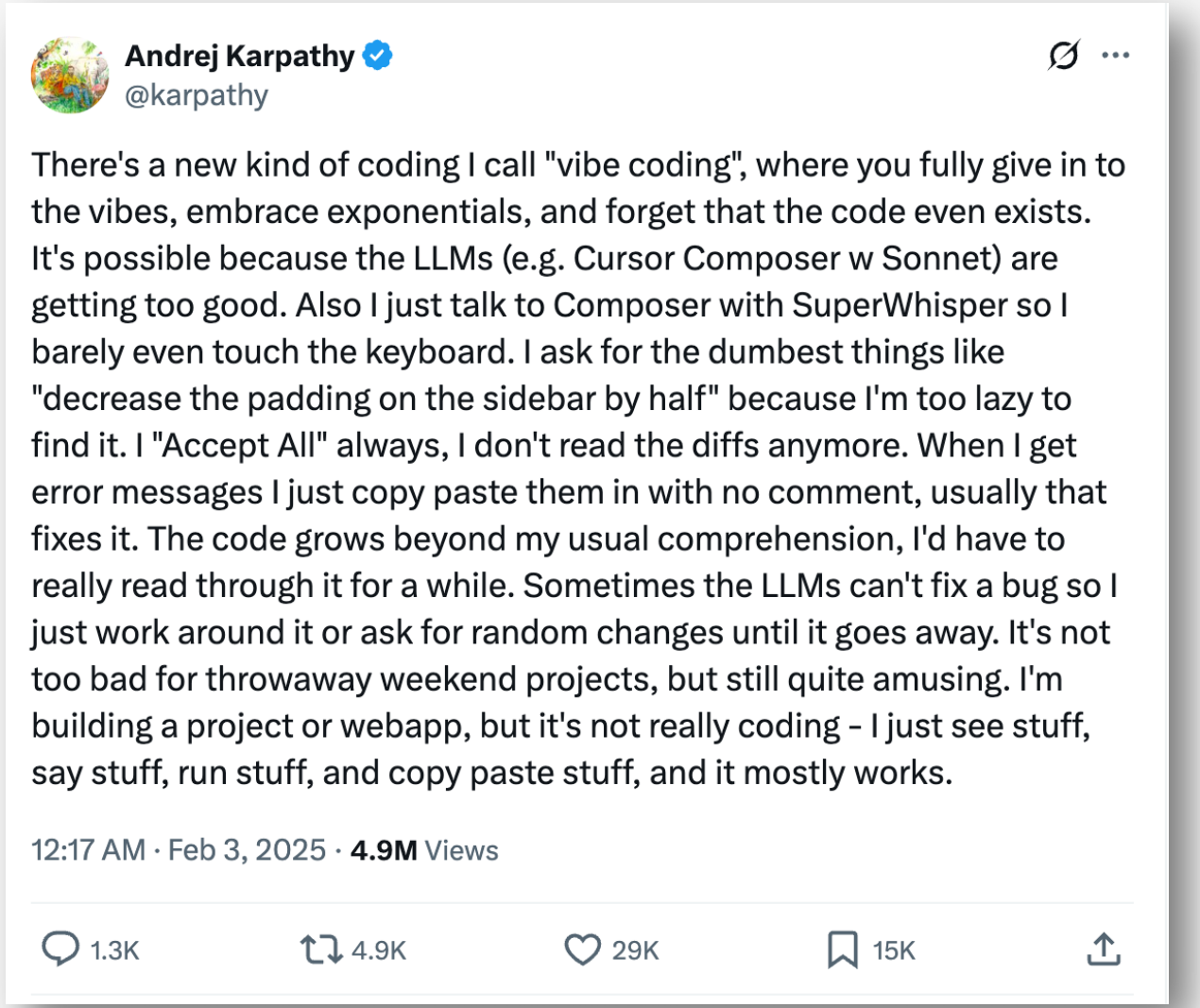


Who knows this man?

Andrej Karpathy
Co-founder of OpenAI

OpenAI ANTHROP\C

Vibe Coding



Andrej Karpathy
@karpathy

There's a new kind of coding I call "vibe coding", where you fully give in to the vibes, embrace exponentials, and forget that the code even exists. It's possible because the LLMs (e.g. Cursor Composer w Sonnet) are getting too good. Also I just talk to Composer with SuperWhisper so I barely even touch the keyboard. I ask for the dumbest things like "decrease the padding on the sidebar by half" because I'm too lazy to find it. I "Accept All" always, I don't read the diffs anymore. When I get error messages I just copy paste them in with no comment, usually that fixes it. The code grows beyond my usual comprehension, I'd have to really read through it for a while. Sometimes the LLMs can't fix a bug so I just work around it or ask for random changes until it goes away. It's not too bad for throwaway weekend projects, but still quite amusing. I'm building a project or webapp, but it's not really coding - I just see stuff, say stuff, run stuff, and copy paste stuff, and it mostly works.

12:17 AM · Feb 3, 2025 · **4.9M** Views

1.3K 4.9K 29K 15K

*"There's a new kind of coding I call '**vibe coding**', where you [...] forget that the code even exists."*

*"I just see stuff, say stuff, run stuff, and copy paste stuff, and it **mostly works**."*

https://x.com/karpathy/status/1886192184808149383

Vibe Coding

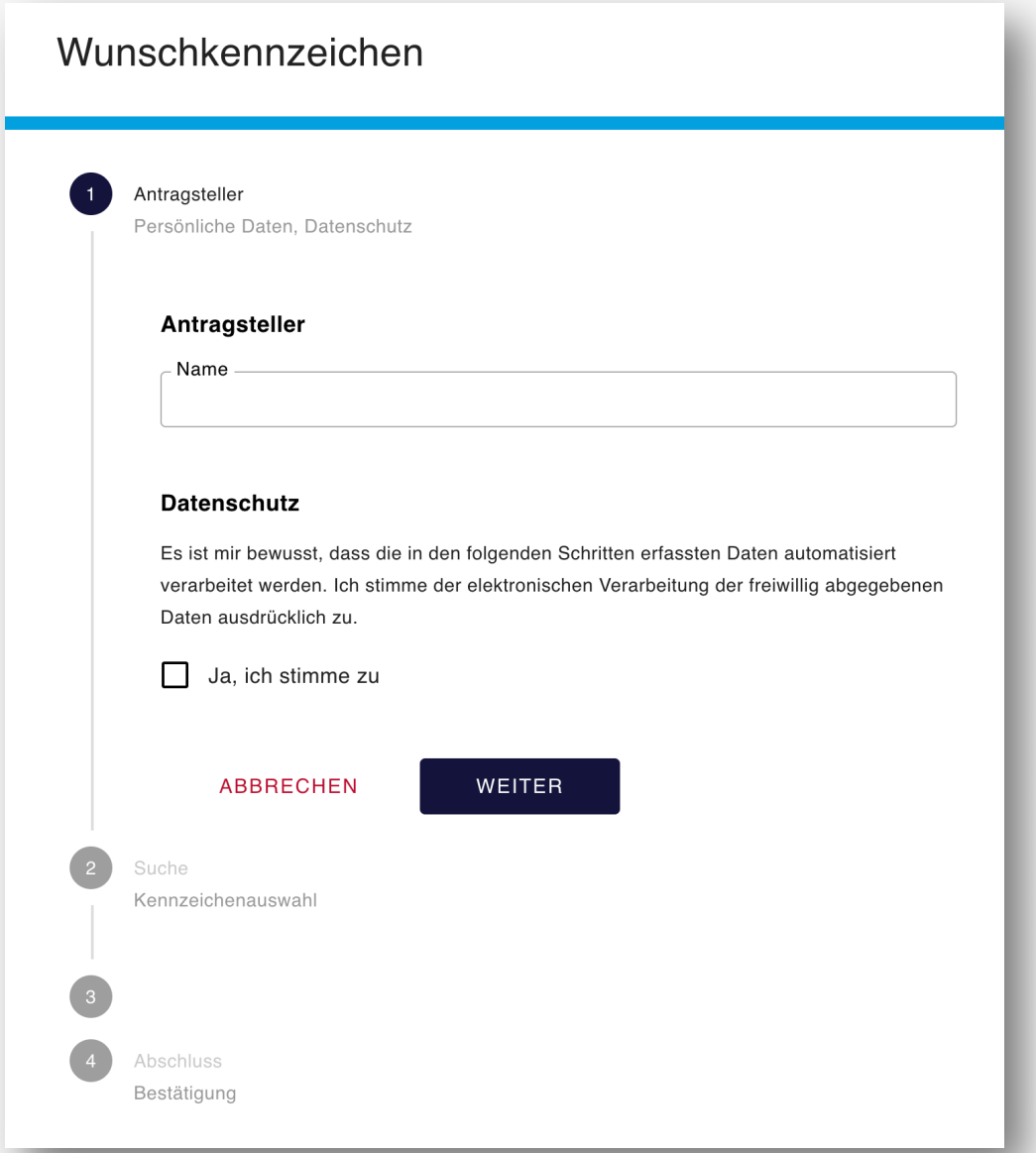


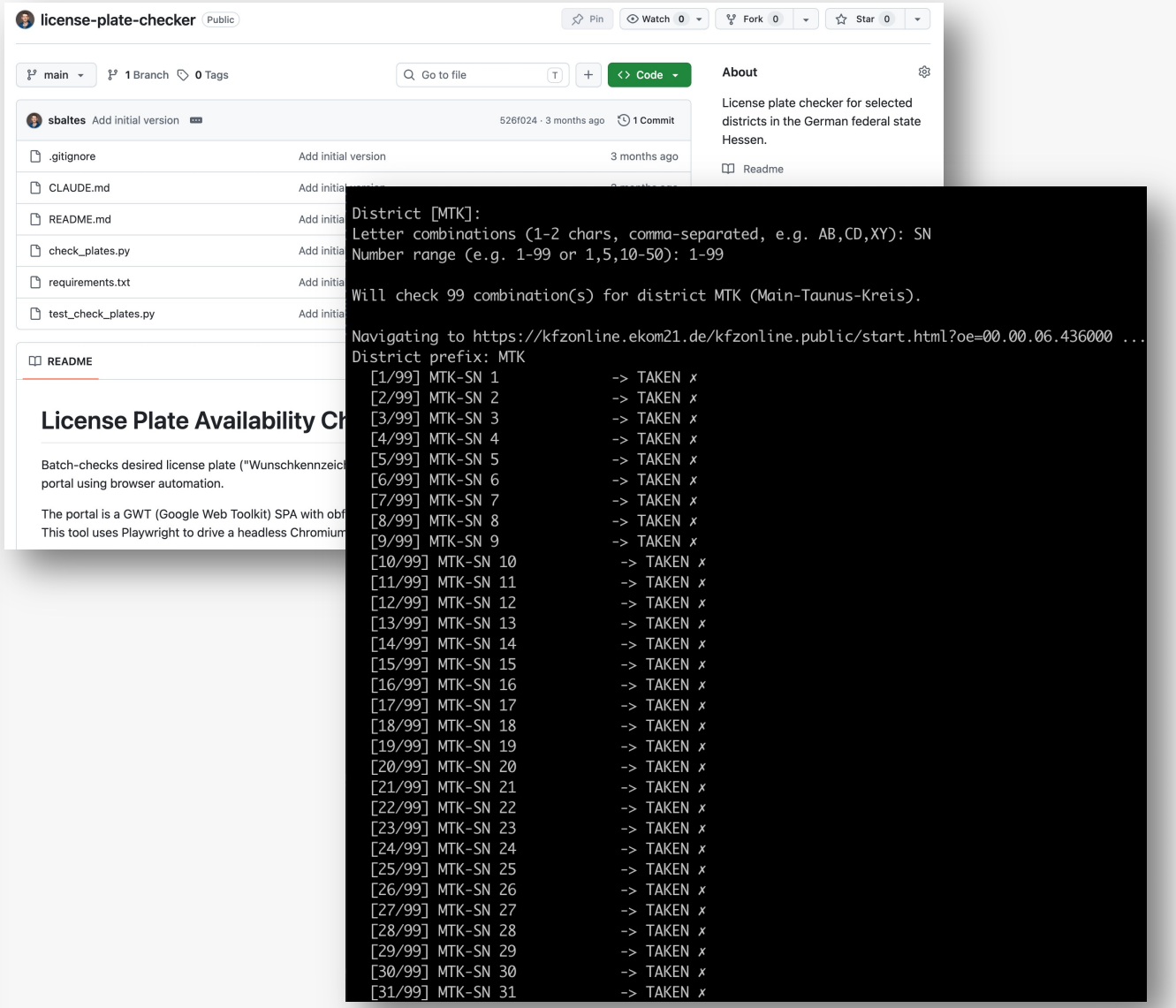
*"There's a new kind of coding I call '**vibe coding**', where you [...] forget that the code even exists."*

*"I just see stuff, say stuff, run stuff, and copy paste stuff, and it **mostly works**."*

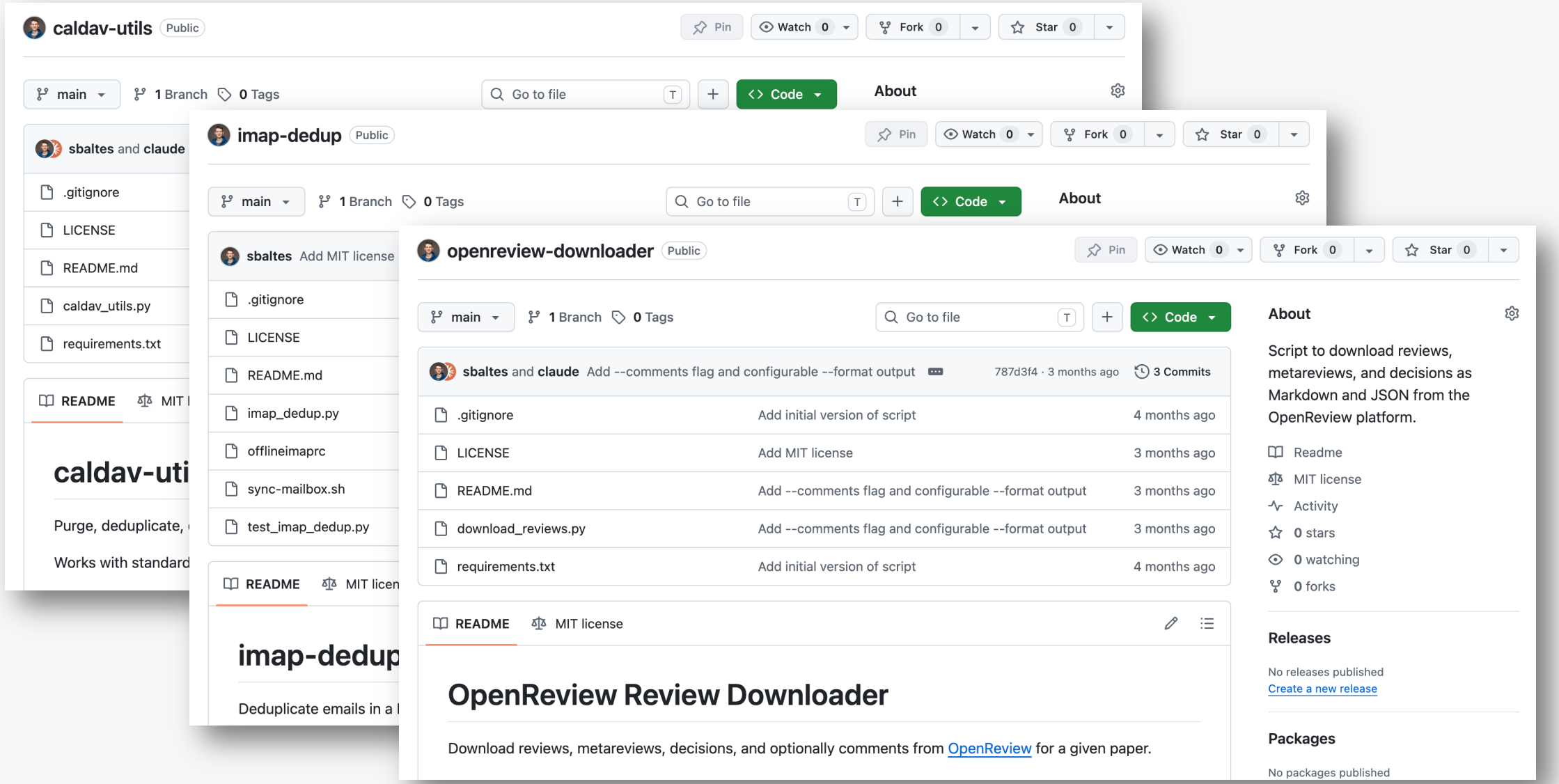
https://x.com/karpathy/status/1886192184808149383

Where Vibe Coding Works: Small Projects

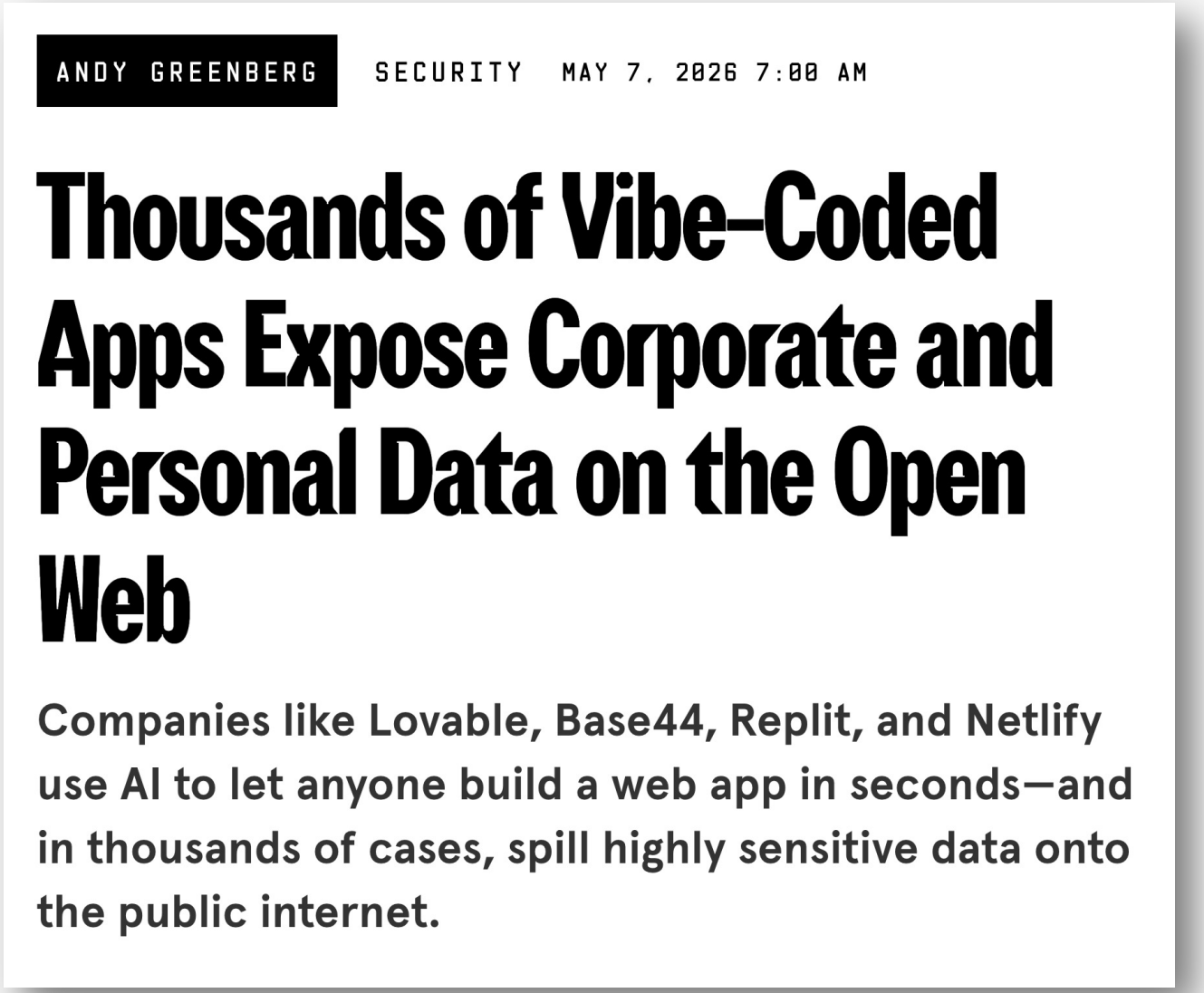




Where Vibe Coding Works: Small Projects



Where It Doesn’t Work: Publicly Deployed Software



ANDY GREENBERG SECURITY MAY 7, 2026 7:00 AM

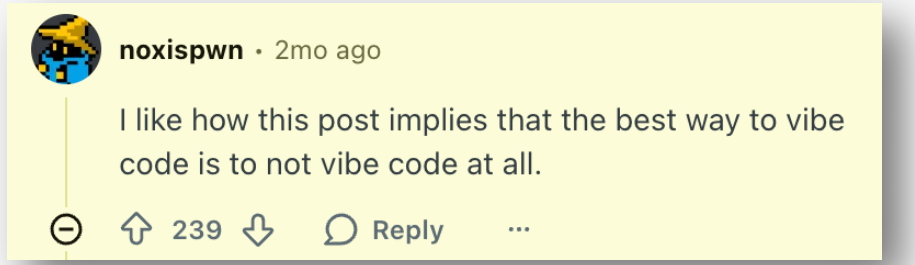
Thousands of Vibe-Coded Apps Expose Corporate and Personal Data on the Open Web

Companies like Lovable, Base44, Replit, and Netlify use AI to let anyone build a web app in seconds—and in thousands of cases, spill highly sensitive data onto the public internet.

https://www.wired.com/story/thousands-of-vibe-coded-apps-expose-corporate-and-personal-data-on-the-open-web/

“Vibe Coding” at FAANG

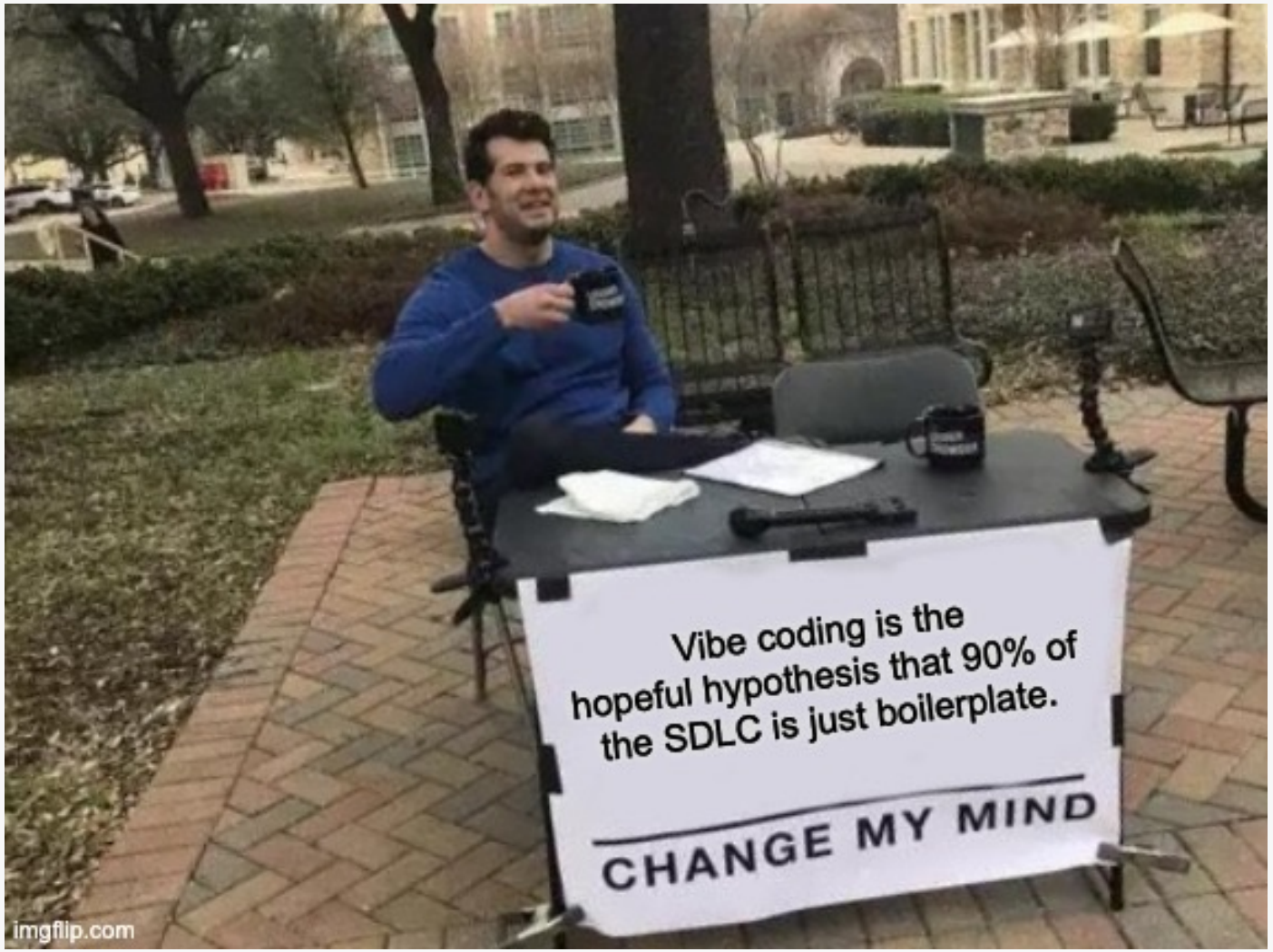
1. […] start with a **technical design document**.
2. **Design review** before […] development […].
3. If you pass review, […] **more documentation on each subsystem**.
4. **Backlog development** and **sprint planning**.
5. **Software development** […] We use **test-driven development**, so I have the AI coding agent write the tests first for the feature I’m going to build.
6. **Code** […] **review**. We have a two-dev approval process before code can get merged into main. AI is […] assisting with the review.
7. **Test in staging.** If staging is good to go, we push to **prod**.



* **F**acebook (Meta), **A**pple, **A**mazon, **N**etflix, and **G**oogle (Alphabet)

https://www.reddit.com/r/vibecoding/comments/1myakhd/how_we_vibe_code_at_a_faang/

Vibe Coding and the SDLC



https://imgflip.com/i/9zk6ku

SDLC: Software Development Life Cycle

What Is the Software Development Life Cycle (SDLC)?



https://upload.wikimedia.org/wikipedia/commons/d/db/Margaret_Hamilton_-_restoration.jpg

Who knows this woman?

Margaret Hamilton
Led Software Development of
NASA's Apollo Guidance Computer

https://en.wikipedia.org/wiki/NASA#/media/File:NASA_logo_(2023_colors).svg

Origin of the Term “Software Engineering”



Margaret Hamilton
(1965–1969)



Anthony G. Oettinger
(1966)

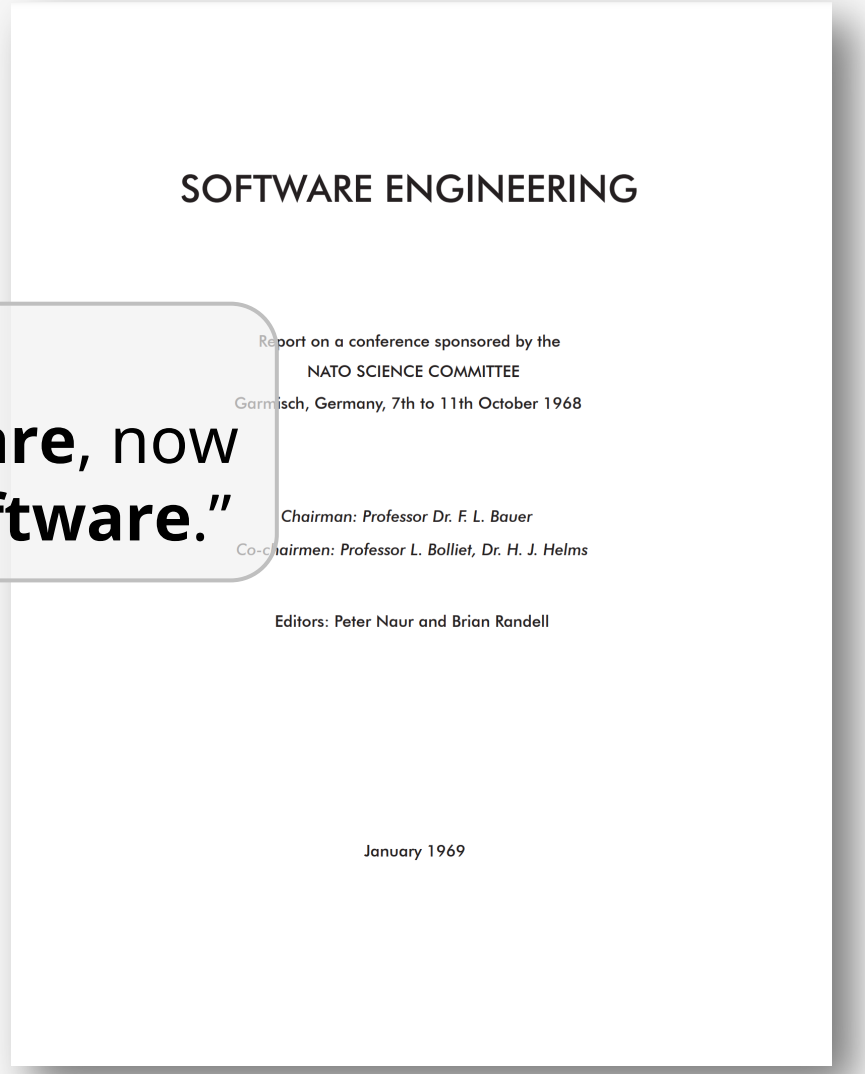


NATO Software Engineering Conference
(1968/69)

“Software Engineering” as a Discipline (~1968)



Call to action:
“We know how to build **hardware**, now we must study how to build **software**.”


SOFTWARE ENGINEERING

Report on a conference sponsored by the
NATO SCIENCE COMMITTEE
Garmisch, Germany, 7th to 11th October 1968

Chairman: Professor Dr. F. L. Bauer
Co-chairmen: Professor L. Bolliet, Dr. H. J. Helms

Editors: Peter Naur and Brian Randell

January 1969

“Software Crisis”:

- Software was **inefficient** and difficult to **maintain**.
- It often did **not meet** the specified **requirements**.
- It was sometimes **not delivered**.

http://homepages.cs.ncl.ac.uk/brian.randell/NATO/N1968/index.html

http://homepages.cs.ncl.ac.uk/brian.randell/NATO/nato1968.PD

1968 NATO SE Conference Report: The SDLC

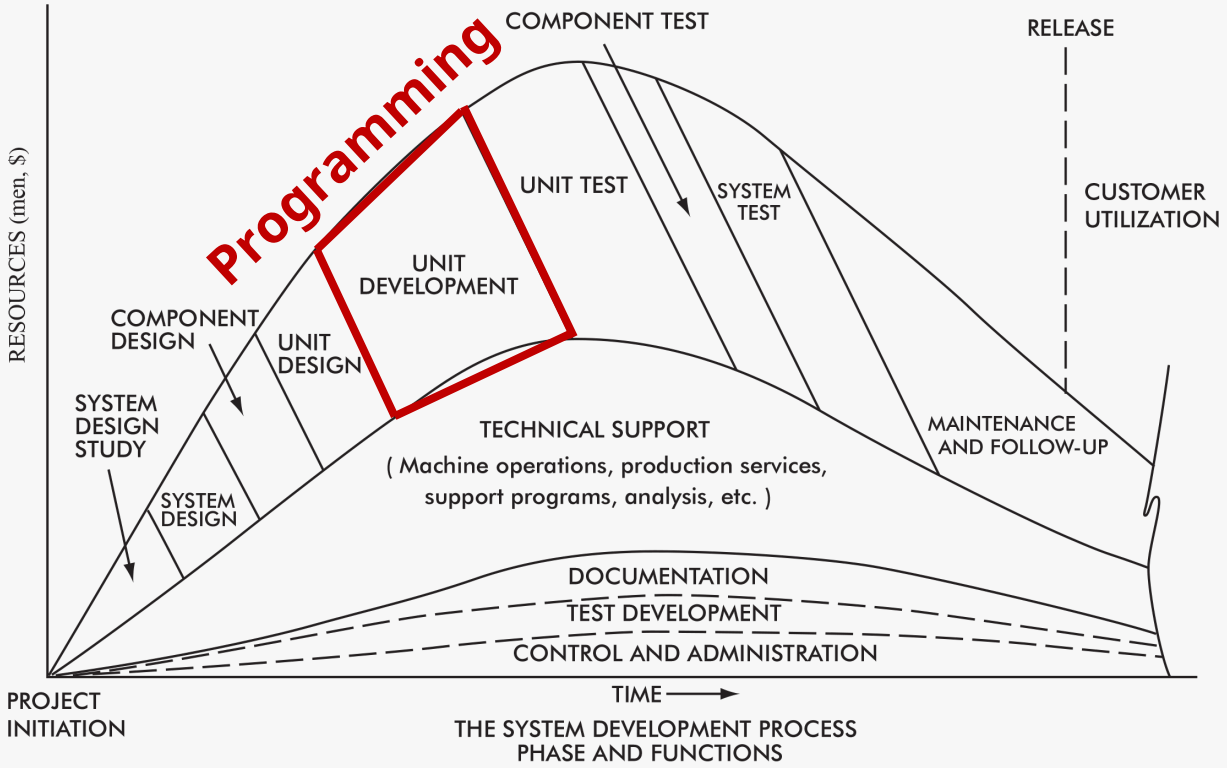


Figure 1. From Nash: Some problems in the production of large-scale software systems.

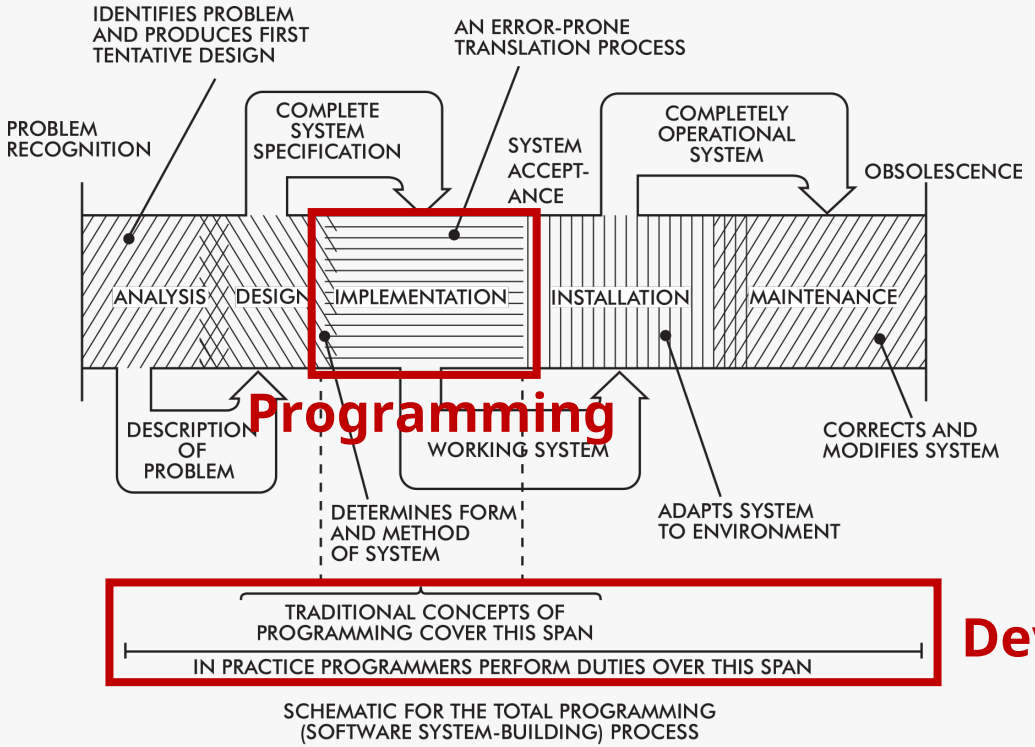
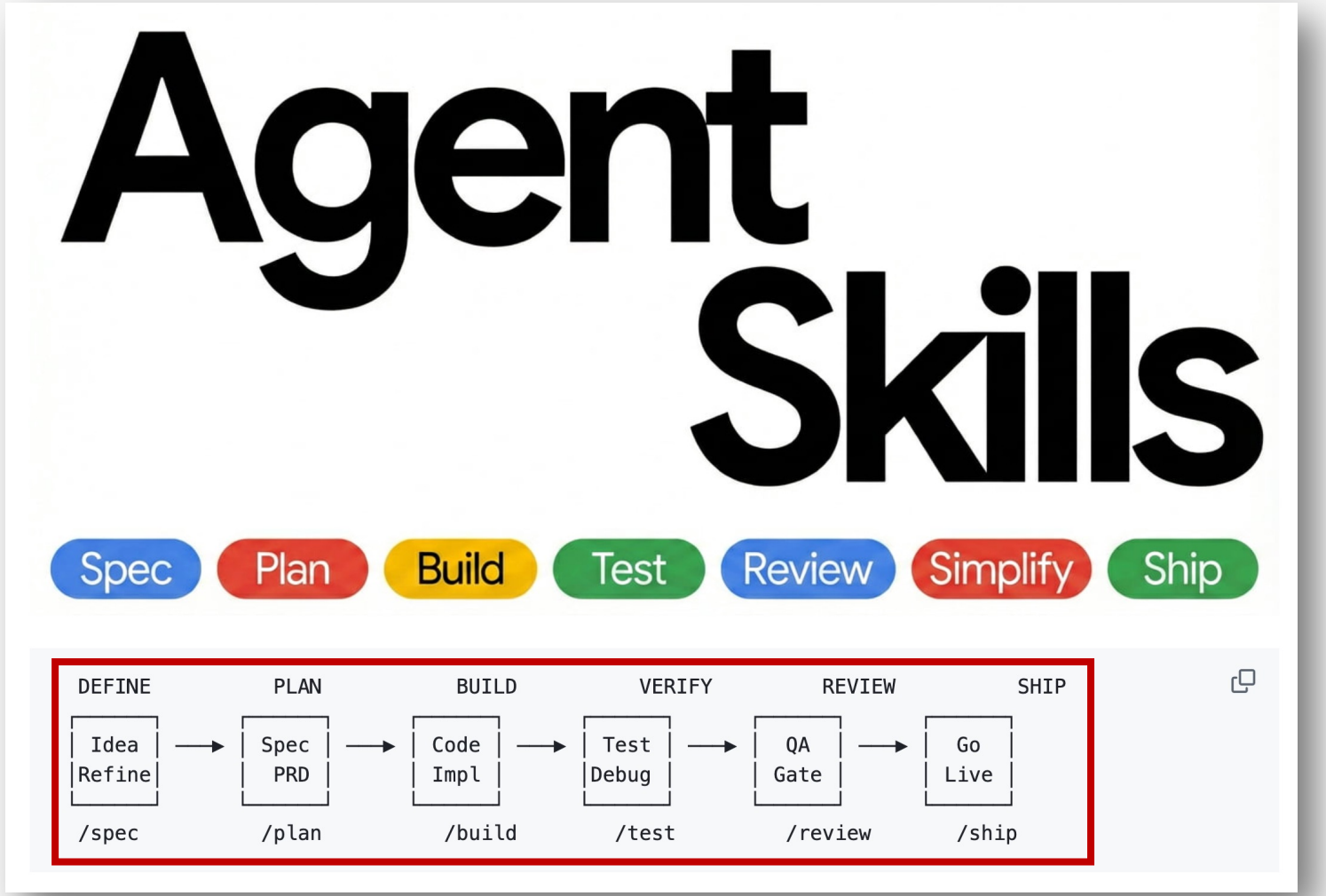
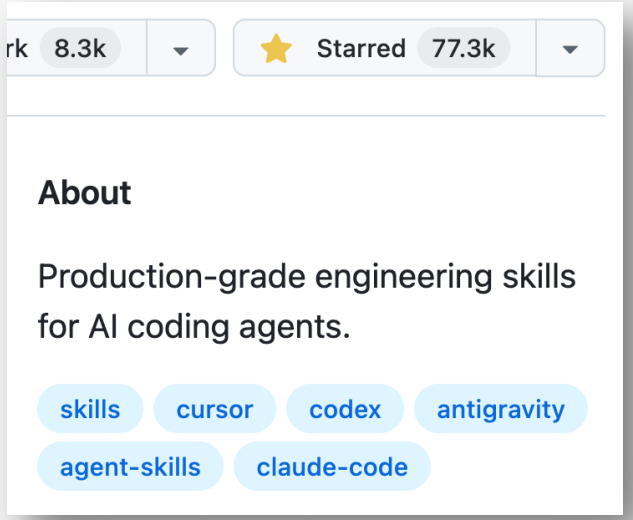


Figure 2. From Selig: Documentation for service and users. Originally due to Constantine.

http://homepages.cs.ncl.ac.uk/brian.randell/NATO/nato1968.PD

SDLC and Agentic AI





https://github.com/addyosmani/agent-skills

Software Engineering: Definition

“Software engineering is the **systematic application** of scientific and technological **knowledge, methods, and experience…**

…to the design, implementation, testing, maintenance, operation, and documentation of **software**…

…with a particular focus on **automation**.”

(adapted from ISO/IEC 2382:2015)

Software Engineering: Definition

“Software engin**ion** of scientific and technologic**erience…**

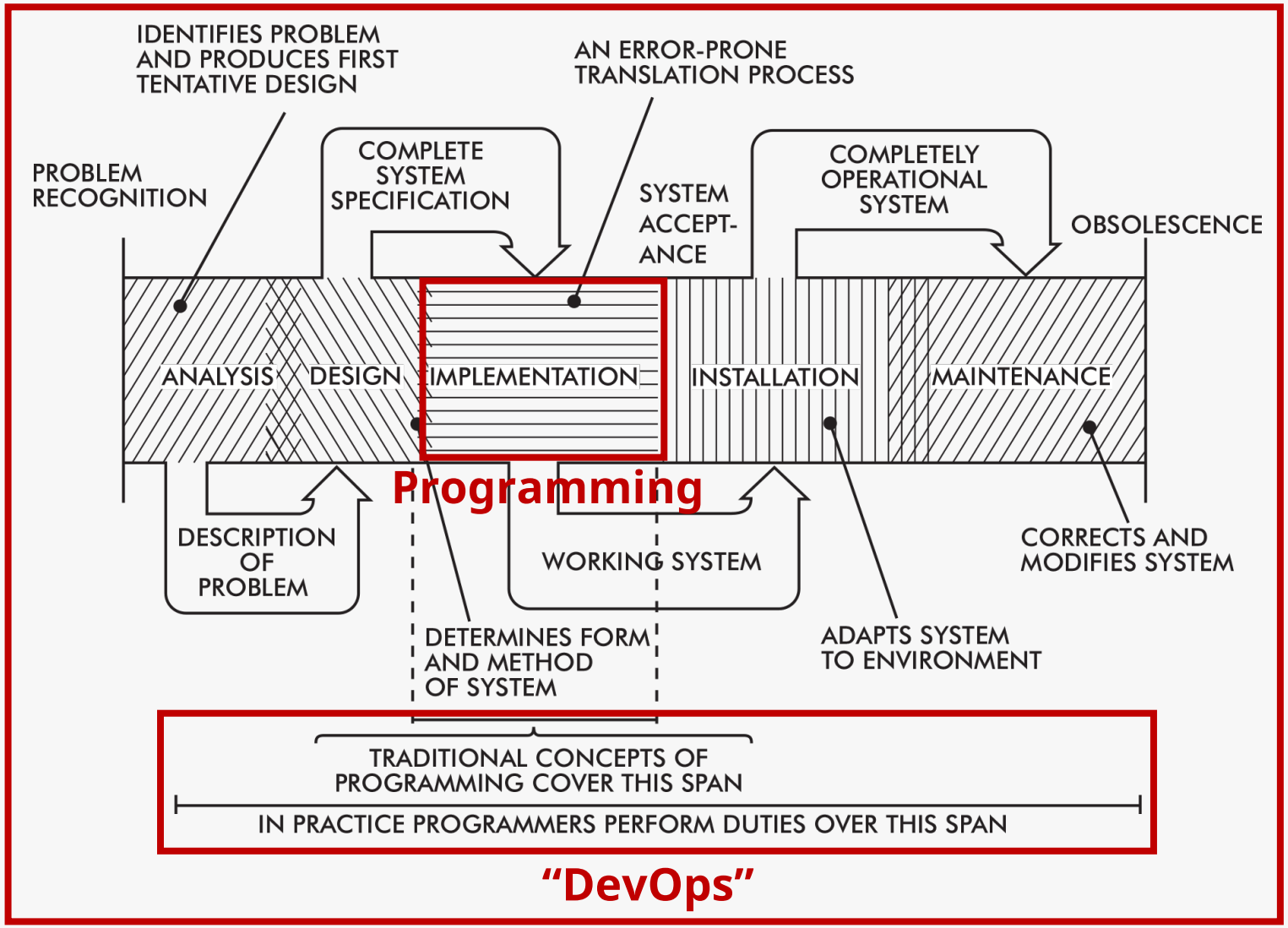
…to the design, ance, operation, and documenta

…with a particul

from ISO/IEC 2382:2015)



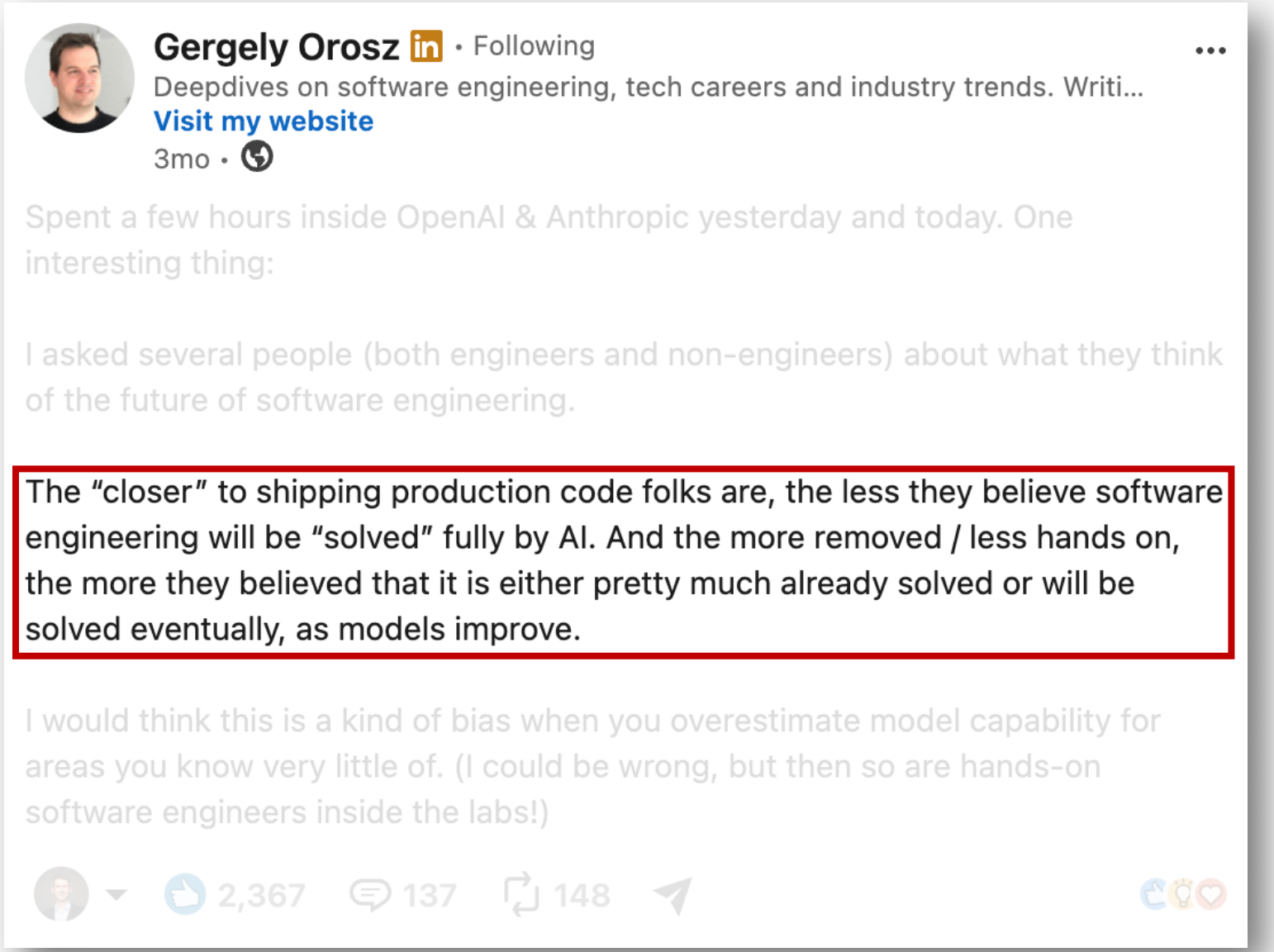
SE Has Never Been Just About Coding!



Software Engineering

http://homepages.cs.ncl.ac.uk/brian.randell/NATO/nato1968.PD

Dunning–Kruger and the AI Hype



Gergely Orosz · Following

Deepdives on software engineering, tech careers and industry trends. Writi...

Visit my website

3mo ·

Spent a few hours inside OpenAI & Anthropic yesterday and today. One interesting thing:

I asked several people (both engineers and non-engineers) about what they think of the future of software engineering.

The "closer" to shipping production code folks are, the less they believe software engineering will be "solved" fully by AI. And the more removed / less hands on, the more they believed that it is either pretty much already solved or will be solved eventually, as models improve.

I would think this is a kind of bias when you overestimate model capability for areas you know very little of. (I could be wrong, but then so are hands-on software engineers inside the labs!)

2,367 137 148

https://www.linkedin.com/posts/gergelyorosz_spent-a-few-hours-inside-openai-anthropic-share-7476403008478728194-Pj0o/

*"The **Dunning–Kruger effect** is a cognitive bias that describes the systematic tendency of people with low ability in a specific area to give overly positive assessments of their ability."*

https://en.wikipedia.org/wiki/Dunning%E2%80%93Kruger_effect

(1) Software Engineering (SE) ≫ Programming ✔

(2) AI “Solves” Programming ⇏ AI “Solves” SE ✔

(3) Agentic AI Is a Software Systems Innovation

(4) We Need More Solid Empirical Evidence

(5) You Should Study Computer Science in HD ☺

Myths on Software Engineering and GenAI

Myth 1: Developers Spend Most of Their Time Writing Code

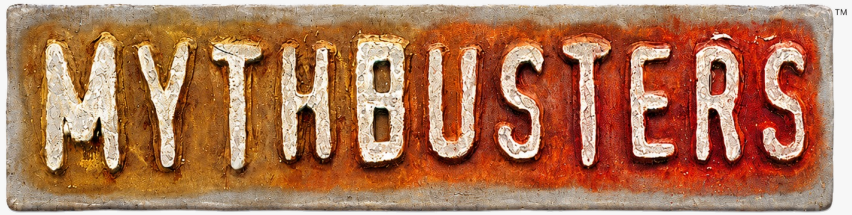
*"A study of more than 450 engineers at Microsoft in 2025 showed **developers spend only 14 percent of their time writing code**, reflecting what has been found in studies over the years."*


RESEARCH-ARTICLE | 26 May 2026

Eight Myths on Software Engineering and GenAI

Examining the most common misconceptions

Jenna Butler, Brian Houck, Margaret-Anne Storey, Travis Lowdermilk, Steven Clarke, and Emerson Murphy-Hill



https://spawn-queue.acm.org/doi/10.1145/3807963



Yes, you can measure software developer productivity

August 17, 2023 | Article

https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/yes-you-can-measure-software-developer-productivity#/

“State of the Art” in Software Engineering as of 2019

No Single Metric Captures Productivity

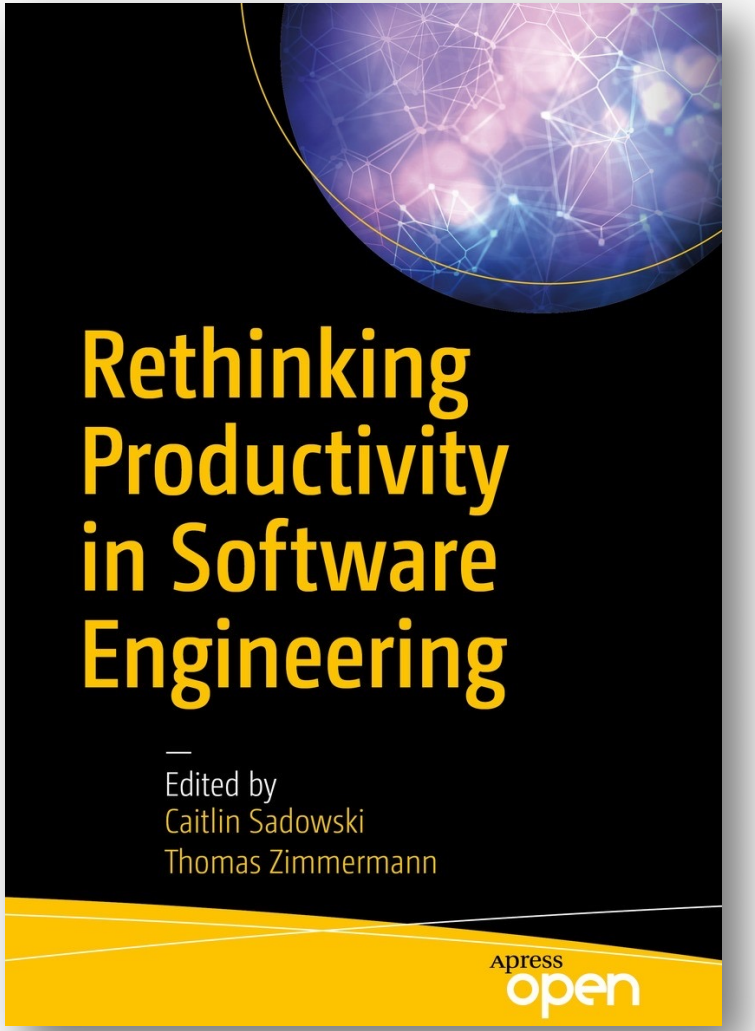
Ciera Jaspan, Google, USA
Caitlin Sadowski, Google, USA

“Measuring software productivity by lines of code is like measuring progress on an airplane by how much it weighs.”

—Bill Gates

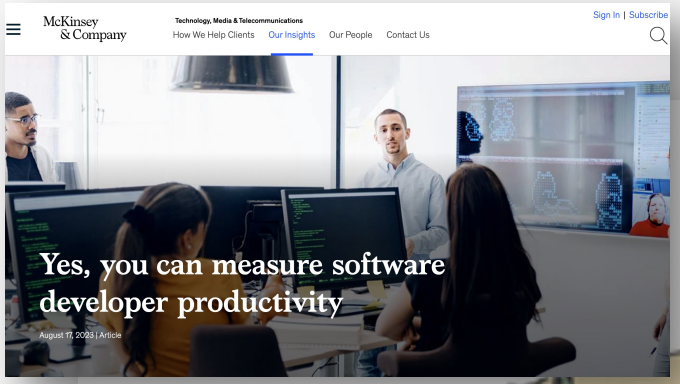
Why We Should Not Measure Productivity

Amy J. Ko, University of Washington, USA

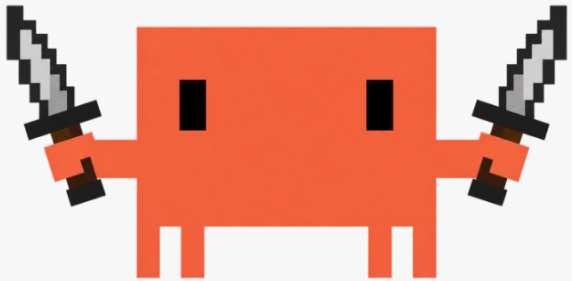


https://link.springer.com/book/10.1007/978-1-4842-4221-6

Public Discourse About SE

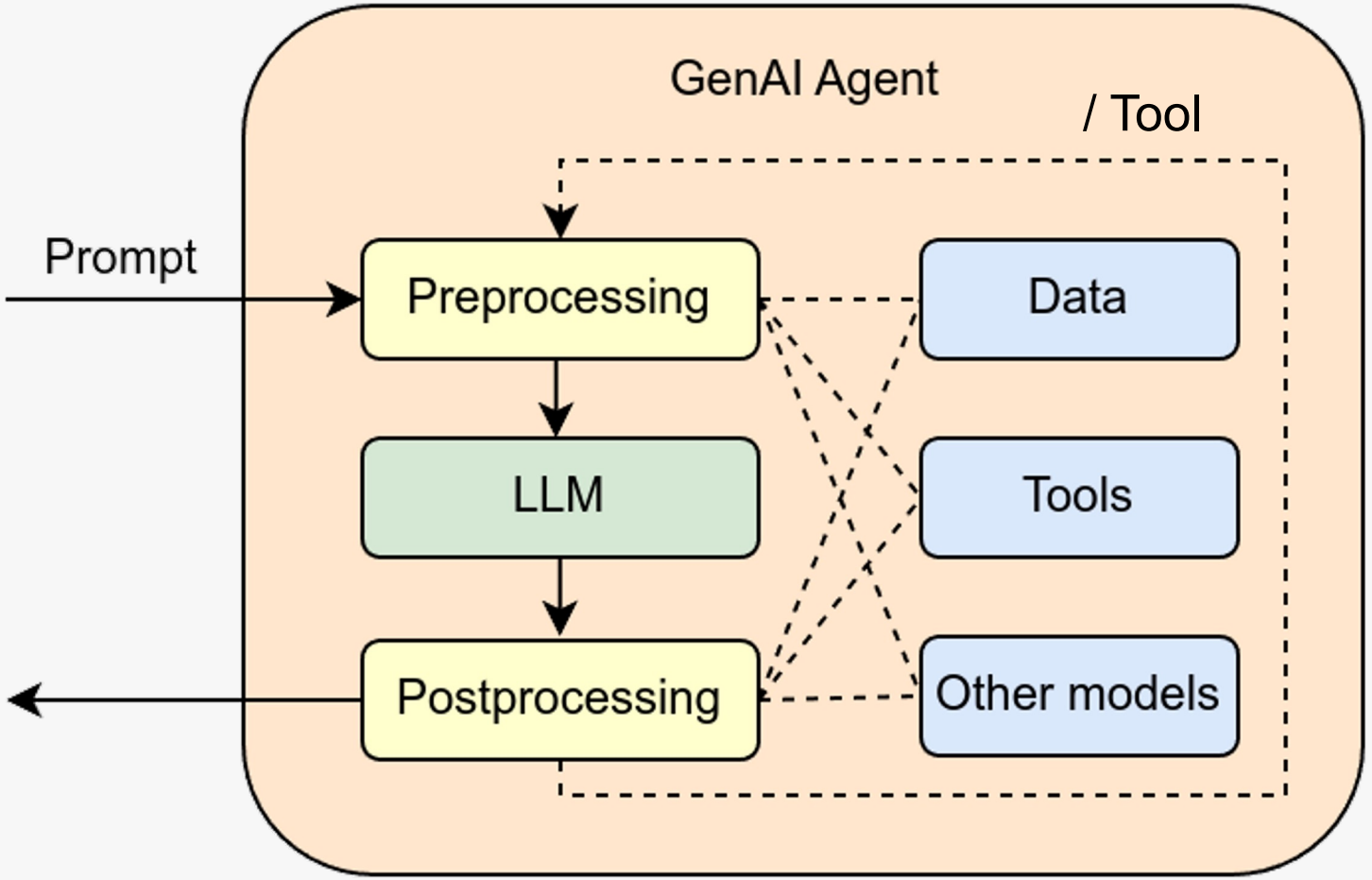


Agentic AI Assistants



Generative Artificial Intelligence (GenAI) Tools

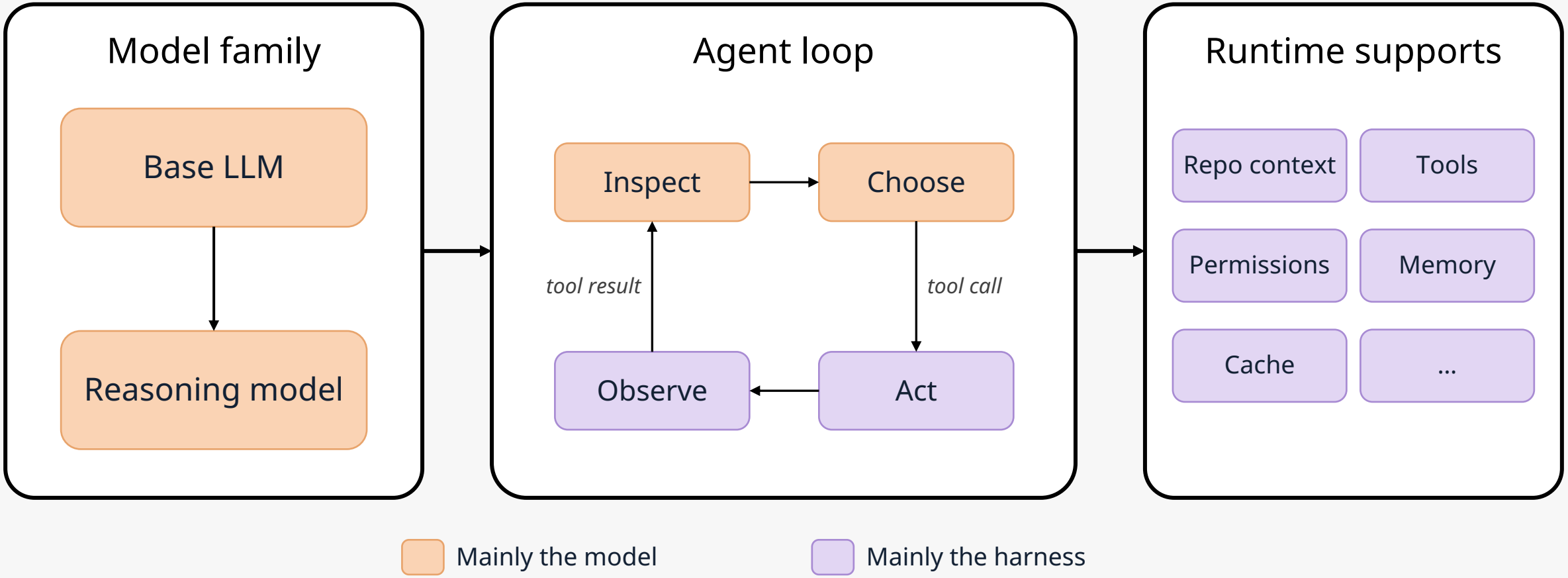
The quality of the **context** provided as part of the prompts is crucial.



https://upload.wikimedia.org/wikipedia/commons/1/16/GenAI_Agent.png

Agentic Coding Tools: Model and Harness

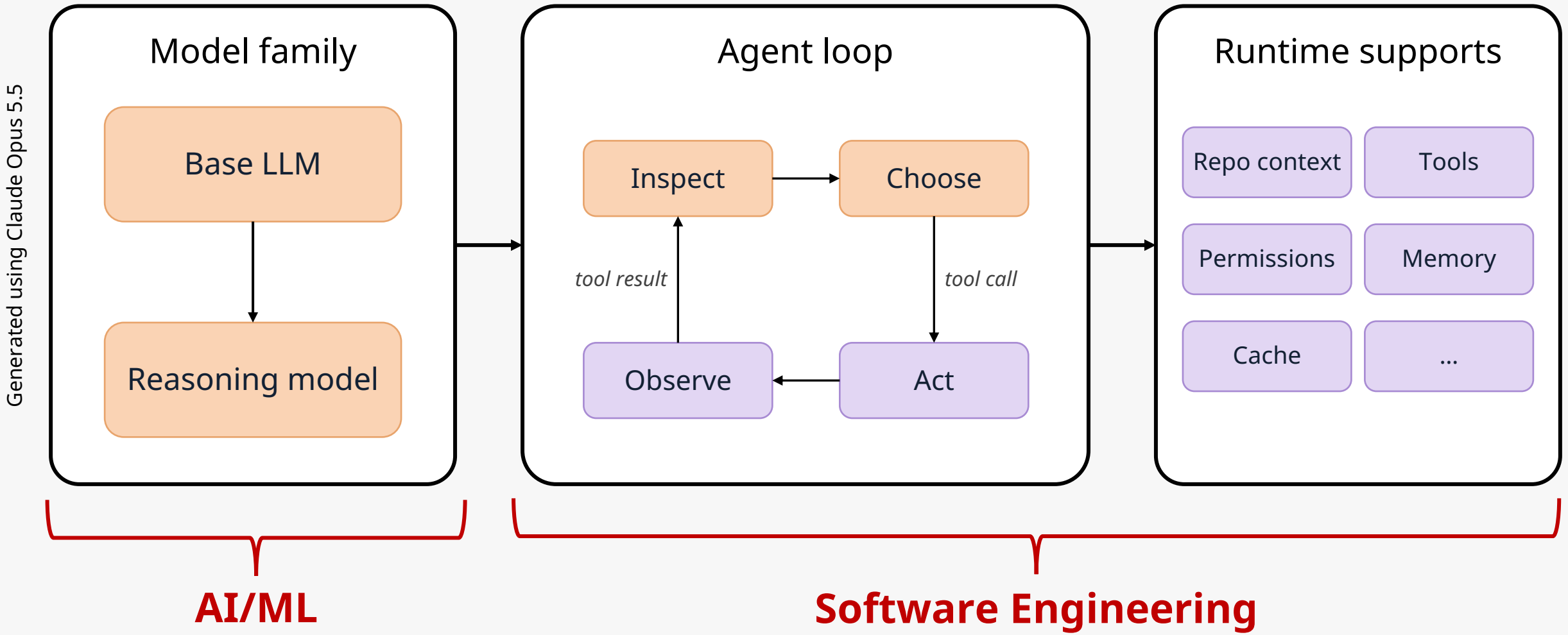
Generated using Claude Opus 5.5



The model proposes the next step as a tool call. The harness validates the call, runs it, and returns the result.

https://magazine.sebastianraschka.com/p/components-of-a-coding-agent

Agentic Coding Tools: Model and Harness



https://magazine.sebastianraschka.com/p/components-of-a-coding-agent

TL;DR

We’ve moved from model innovation to systems innovation

(1) Software Engineering (SE) ≫ Programming ✔

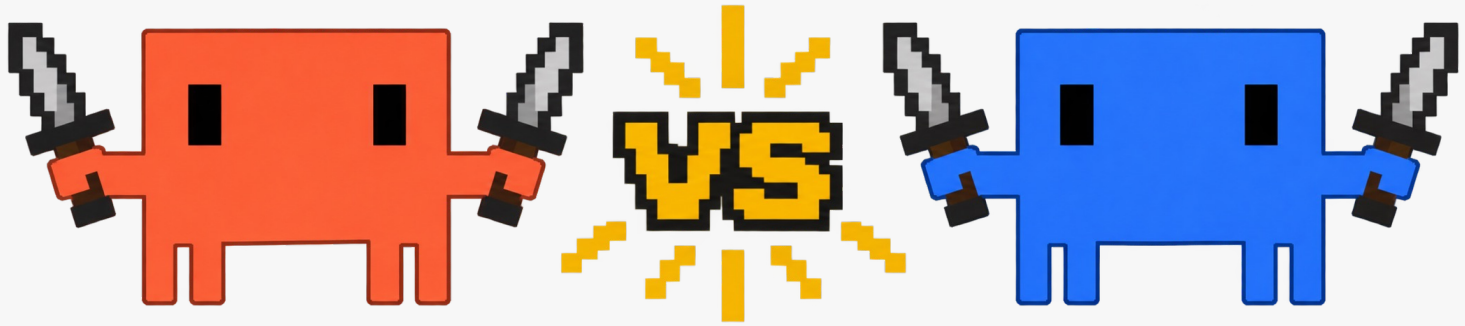
(2) AI “Solves” Programming ⇏ AI “Solves” SE ✔

(3) Agentic AI Is a Software Systems Innovation ✔

(4) We Need More Solid Empirical Evidence

(5) You Should Study Computer Science in HD ☺

Benchmarking AI Tools for SE



Benchmarking AI Tools

- Benchmarks are a central aspect of **vendors' marketing**.
- Most **benchmarks are open**, but the **models are closed**.
  ("Open" usually means open weight, not open training data.)
- **Benchmark contamination** is one of many problems.
- Another problem is that benchmarks **rarely define the construct they intend to measure** (functional correctness, security, reliability, etc.) and why the chosen **operationalization** is valid.

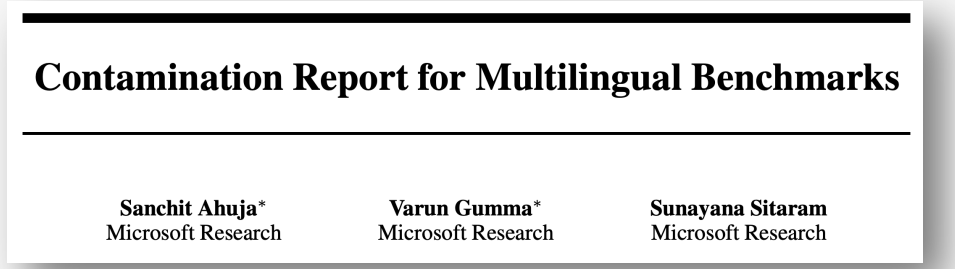


The Leaderboard Illusion

Shivalika Singh[*1], Yiyang Nan[1], Alex Wang[2], Daniel D'souza[1], Sayash Kapoor[3], Ahmet Üstün[1], Sanmi Koyejo[4], Yuntian Deng[5], Shayne Longpre[6], Noah A. Smith[7,8], Beyza Ermis[1], Marzieh Fadaee[♦1], and Sara Hooker[♦1]

[1]Cohere Labs, [2]Cohere, [3]Princeton University, [4]Stanford University, [5]University of Waterloo, [6]Massachusetts Institute of Technology, [7]Allen Institute for Artificial Intelligence, [8]University of Washington

https://arxiv.org/pdf/2504.20879



Contamination Report for Multilingual Benchmarks

Sanchit Ahuja* Microsoft Research

Varun Gumma* Microsoft Research

Sunayana Sitaram Microsoft Research

https://arxiv.org/pdf/2410.16186

Benchmarks: HumanEval

Prompt:

```
def is_palindrome(text: str):
    """
    Checks if given string is a palindrome
    >>> is_palindrome('')
    True
    >>> is_palindrome('aba')
    True
    [...]
```

Canonical Solution:

```
def is_palindrome(text: str):
    for i in range(len(text)):
        if text[i] != text[len(text)-1-i]:
            return False
    return True
```

Tests:

```
assert is_palindrome('')      == True
assert is_palindrome('aba')   == True
assert is_palindrome('aaaaa') == True
assert is_palindrome('zbcd')  == False
assert is_palindrome('xywyx') == True
assert is_palindrome('xywyz') == False
assert is_palindrome('xywzx') == False
```

https://github.com/openai/human-eval/

Benchmarks: HumanEval

Check an LLM-generated solution by validating that the generated code is **executable** and **passes the tests**.
Usually, several (`n`) candidate solutions are generated.

$$\text{pass@}k = 1 - \frac{\binom{n-c}{k}}{\binom{n}{k}}$$

`pass@k` estimates the probability that at least one of `k` generated candidate solutions for a fixed task passes the benchmark's tests.

where:

- `n` is the total number of samples generated per prompt (with `n≥k`),
- `c` is the number of correct samples among `n`, and
- `k` is the number of attempts considered, drawn from the n generated samples without replacement.

Benchmarks: HumanEval

Frontier intelligence at 2x the speed

Claude 3.5 Sonnet sets new industry benchmarks for graduate-level reasoning (GPQA), undergraduate-level knowledge (MMLU), and coding proficiency (HumanEval). It shows marked improvement in grasping nuance, humor, and complex instructions, and is exceptional at writing high-quality content with a natural, relatable tone.

ANTHROP\C

Marketing in June 2024

- Just Python.
- Just pure functions.
- No modules.
- No existing codebase.
- No debugging.
- No non-functional properties.

`pass@k` estimates the probability that at least one of `k` generated candidate solutions for a fixed task passes the benchmark's tests.

Construct

Operationalization

Benchmarks: HumanEval

Frontier intelligence at 2x the speed

Claude 3.5 Sonnet sets new industry benchmarks for graduate-level reasoning

coding proficiency

g nuance, humor, and

-quality content with a



- Just Python.
- Just pure functions.
- No modules.
- No existing codebase.
- No debugging.
- No non-functional properties.

`pass@k` estimates the probability that at least one of `k` generated candidate solutions for a fixed task passes the benchmark's tests.

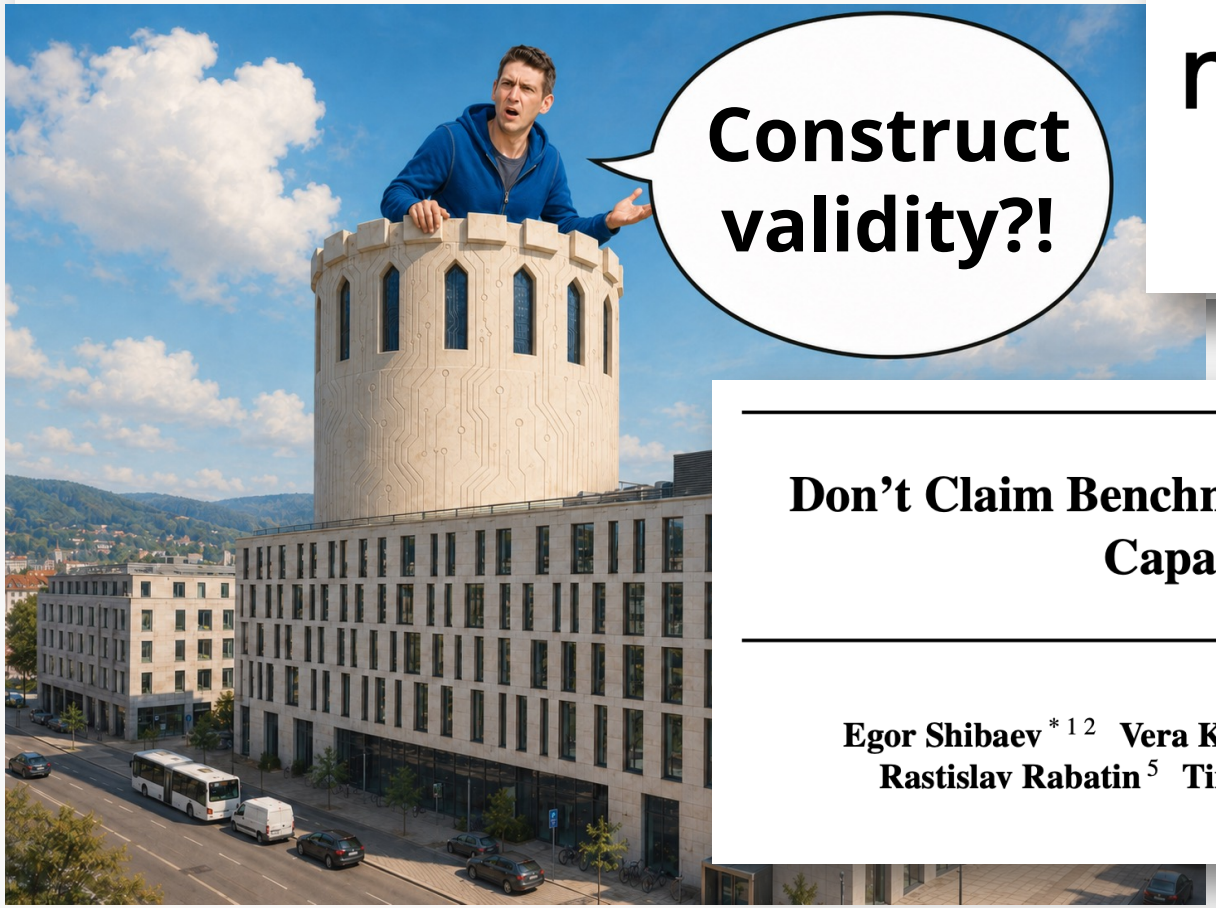
Construct

Operationalization

Benchmarks: HumanEval

Frontier intelligence at 2x the spee

Claude 3.5 Sonnet sets new industry benchmarks for



July 8, 2026 Research Publication

https://openai.com/index/separating-signal-from-noise-coding-evaluations/

Separating signal from noise in coding evaluations

Through a detailed audit, we find widespread task issues in SWE-Bench Pro and estimate that ~30% of the tasks are broken.

that at least one of k generated

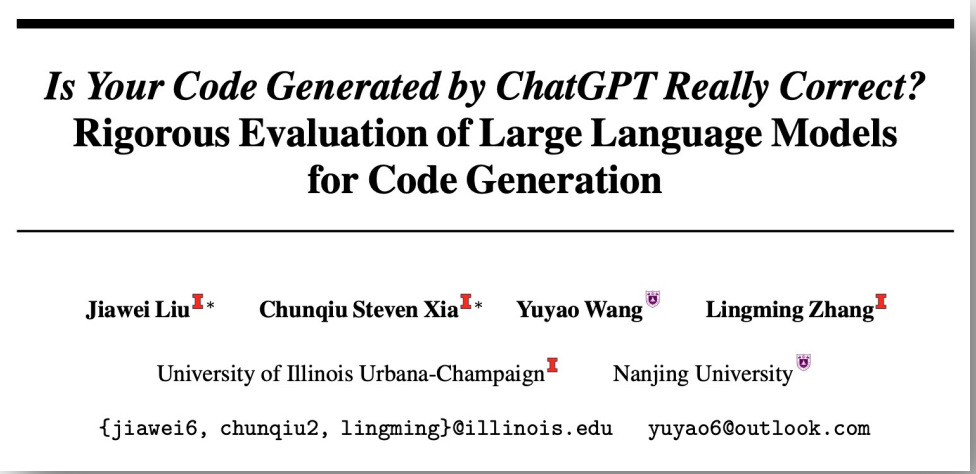
Don't Claim Benchmark-Oriented Optimization Improves General Coding Capability — Diverse Evaluation Is Required

Egor Shibaev [*1,2] Vera Kudrevskaia [*1,2] Timur Galimzyanov [*1] Mikhail Evtikhiev [*3] Ana Terna [4]
Rastislav Rabatin [5] Timur Kudashev [1,2] Timofey Bryksin [3] Arina Puchkova [1] Patrik Bartak [4]
Egor Bogomolov [4] Sergey Titov [4]

https://openreview.net/pdf?id=dQs1srLmGN

Benchmarks: HumanEval

- HumanEval contained several **incorrect** “canonical” **solutions**.
- Only **10 assertions per problem on average**, not enough to accurately assess the correctness of solutions.



Is Your Code Generated by ChatGPT Really Correct?
Rigorous Evaluation of Large Language Models for Code Generation

Jiawei Liu* Chunqiu Steven Xia* Yuyao Wang Lingming Zhang

University of Illinois Urbana-Champaign Nanjing University

{jiawei6, chunqiu2, lingming}@illinois.edu yuyao6@outlook.com

https://arxiv.org/abs/2305.01210

“[…] continued pre-training on a dataset containing even a single epoch of the test split for HumanEval **boosted scores from 74.4% to 89.0%, highlighting the danger of data contamination**.”



Google DeepMind

Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context

Gemini Team, Google[1]

https://arxiv.org/pdf/2403.05530

Newer Benchmarks, Same Problems

Open Benchmarks for Closed Models

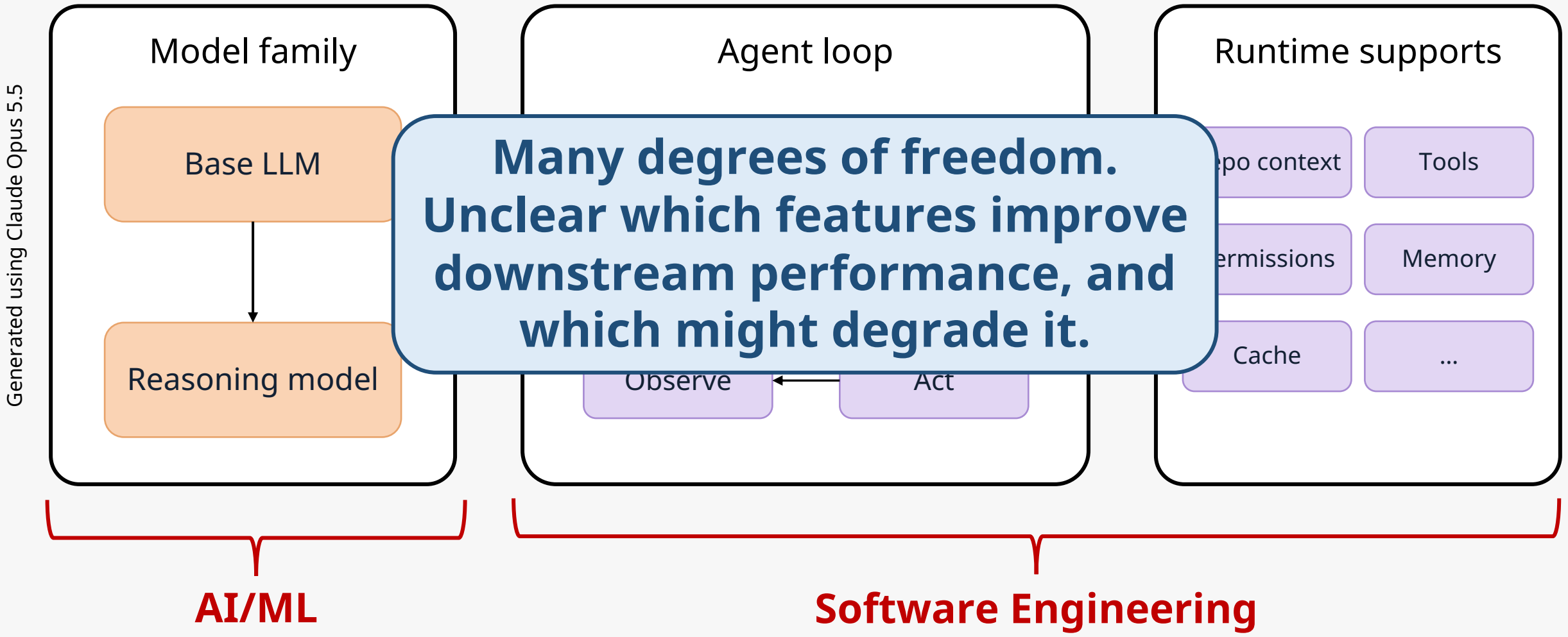


https://knowyourmeme.com/memes/obama-awards-obama-a-medal

Generated with ChatGPT (GPT-5.5)

We Don’t Even Know the “Simple” Things Yet…

Agentic Coding Tools: Model and Harness



https://magazine.sebastianraschka.com/p/components-of-a-coding-agent

Solved by GenAI? Answers from My Colleagues

Felix Joos
Theoretical Computer Science

AI already proves (formal) theorems extremely well, though “solved” may be going a bit too far.

Artur Andrzejak
Machine Learning

Agents now write, run, and check our experiments and report the results, work that took 80–90 % of our time.

Christian Schulz
Algorithm Engineering

AI makes coding more efficient, but I don’t work less. So there must still be plenty to do.

> For years, we have stressed: **programming is not computer science**.
> Now that pure coding is largely being automated, this is truer than ever, and it is what **makes our field truly exciting at its core.**
>
> **Christian Schulz**, Algorithm Engineering

Christian Schulz
Algorithm Engineering

AI will not replace defining the problem, weighing design trade-offs, or judging whether a solution scales.

Artur Andrzejak
Machine Learning

I see no AI solution for research design, the “taste” for what makes sense, or judging what results mean.

Felix Joos
Theoretical Computer Science

I honestly can’t name a problem that AI will never solve. After the last two months, I rule nothing out.

Translated from German, partly paraphrased.

Studying Computer Science at Uni HD



UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

Fakultät für Mathematik und Informatik















Bachelor of Science Informatik (50% oder 100%)

53 % **Informatik**

- **Praktische Informatik**
- **Technische Informatik**
- **Programmiersprache(n)**
- **Algorithmen & Datenstrukturen**
- **Datenbanken**
- **Betriebssysteme / Netzwerke**
- **Software Engineering**
- **Theoretische Informatik**

18 % **Mathematik**

Analysis, Lineare Algebra, Numerik oder Statistik

13 % **Anwendungsgebiet**

Astronomie, Biologie, Chemie, Computerlinguistik, Geowissenschaften, Mathematik, Molekularbiologie, Philosophie, Physik, Wirtschaftswissenschaften, Medizin, Psychologie, …

8 % **„Soft Skills“**

Präsentieren, Arbeitstechniken

8 % **Bachelorarbeit**

Selbstständige Arbeit

Masterstudiengänge

- M.Sc. **Data and Computer Science** (Englisch)
- M.Sc. **Technische Informatik** (Deutsch)
- M.Sc. **Scientific Computing** (Englisch)
- Alle Studiengänge sind interdisziplinär ausgerichtet.

(1) Software Engineering (SE) ≫ Programming ✔

(2) AI “Solves” Programming ⇏ AI “Solves” SE ✔

(3) Agentic AI Is a Software Systems Innovation ✔

(4) We Need More Solid Empirical Evidence ✔

(5) You Should Study Computer Science in HD ☺ ✔

https://www.informatik.uni-heidelberg.de/



empirical-software.engineering

Prof. Dr. Sebastian Baltes