

**ANTHROPIC**

# **Frontier AI for Semiconductors: Automated Verification, Design Agents, and Faster Cycles**

# Anthropic: Safety and Intelligence

# AI will have a profound impact on the world.

Anthropic's mission is to ensure that the world safely makes the transition through transformative AI. We do this by building reliable, interpretable, and steerable AI systems.



Safety



Science

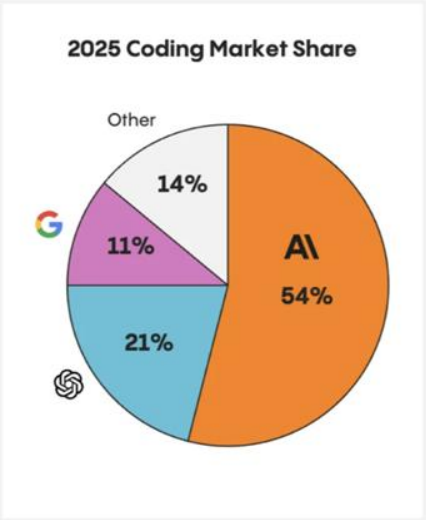
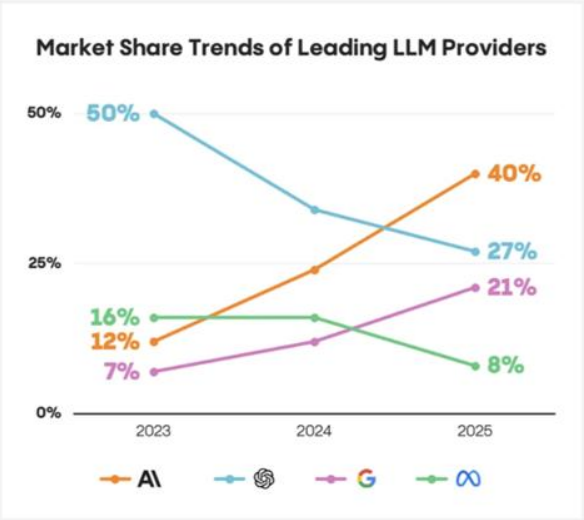
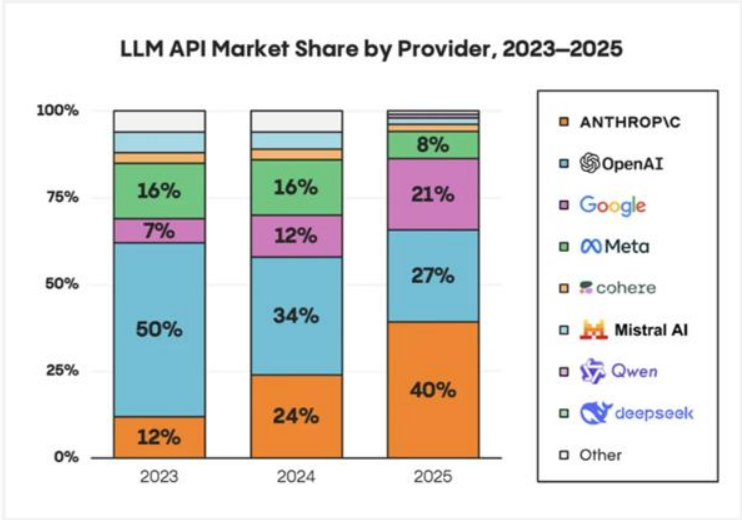


Stewardship



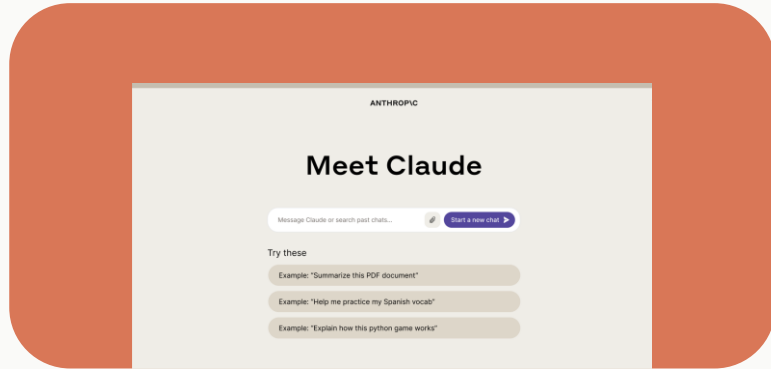
Simplicity

# Enterprise LLM API Market Share by Usage



# The evolution of AI at work

## Chat



Ask questions.  
Get answers.  
Do the work yourself.

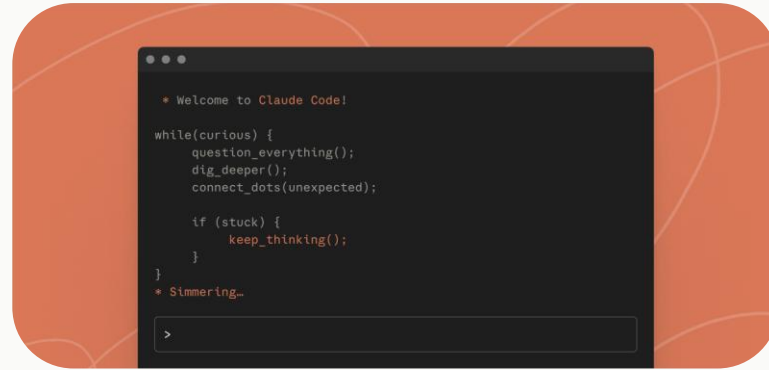
2023

# The evolution of AI at work

Developers got access to a different category.

AI that goes beyond answering questions.

## Code



Chat

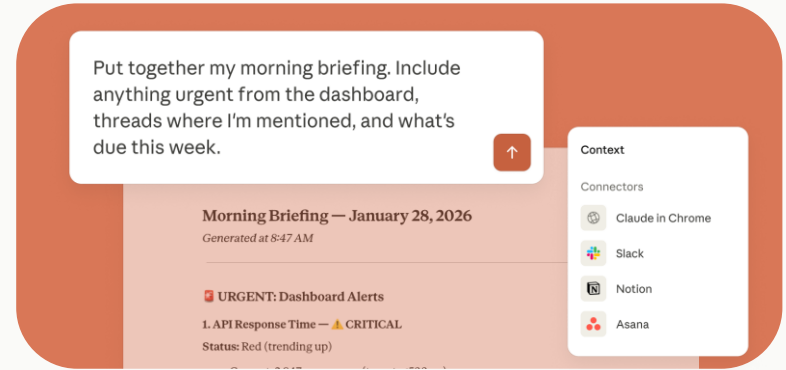
2023

2025

# The evolution of AI at work

## Cowork

Everyone else gets the same execution power. No terminal.



Chat

2023

Code

2025

2026

# The Intelligence Explosion

# The Artificial Intelligence Exponential

✳️ Claude  
Human + AI  
Workforce

We are here → ✳️ Claude Opus 4.6

Advanced coding, multi-step agents, office tasks

Google  
Nano Banana

Claude Sonnet 3.5

Coding assistant

Gemini

ChatGPT



Scaling Laws

Google Research

Transformer  
Architecture

2017


2020

2022

2024

2025





# Can Claude clone Claude.ai?

We asked a few models to try.

Engineering at Anthropic



# Building a C compiler with a team of parallel Claudes

Published Feb 05, 2026

We tasked Opus 4.6 using agent teams to build a C Compiler, and then (mostly) walked away. Here's what it taught us about the future of autonomous software development.

Enabling long-running Claudes

Running Claude in parallel

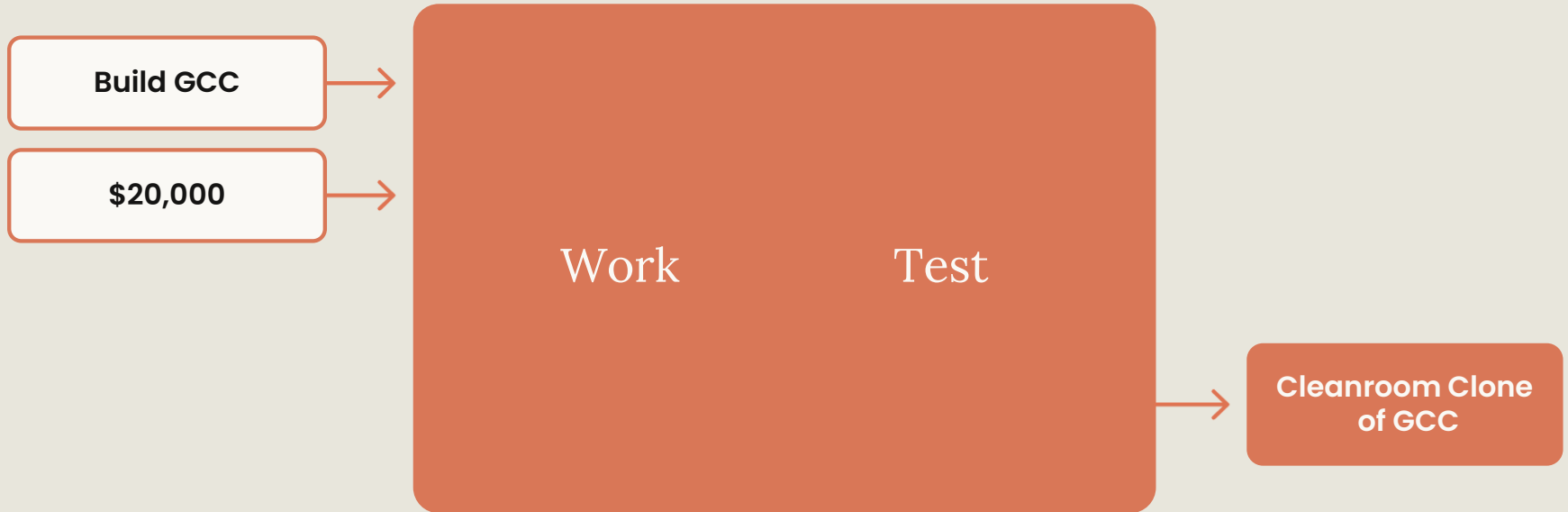
Lessons from programming with Claude agent teams

*Written by Nicholas Carlini, a researcher on our Safeguards team.*

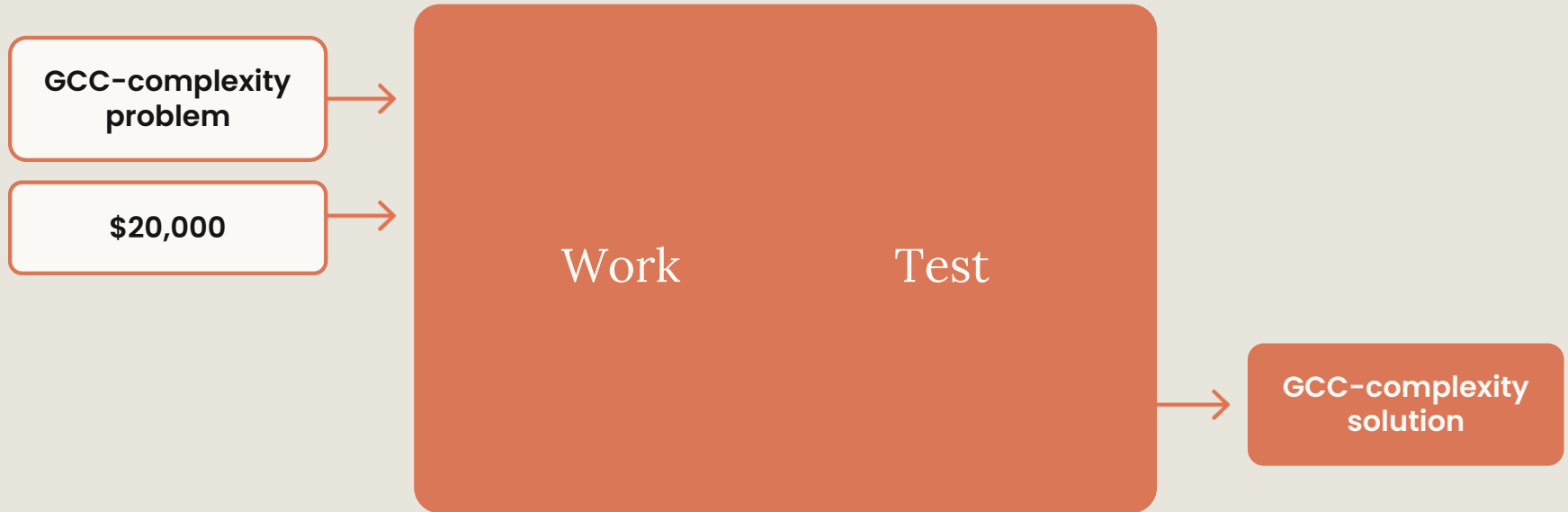
I've been experimenting with a new approach to supervising language models that we're calling "agent teams."

# Harnesses and Closed Loops

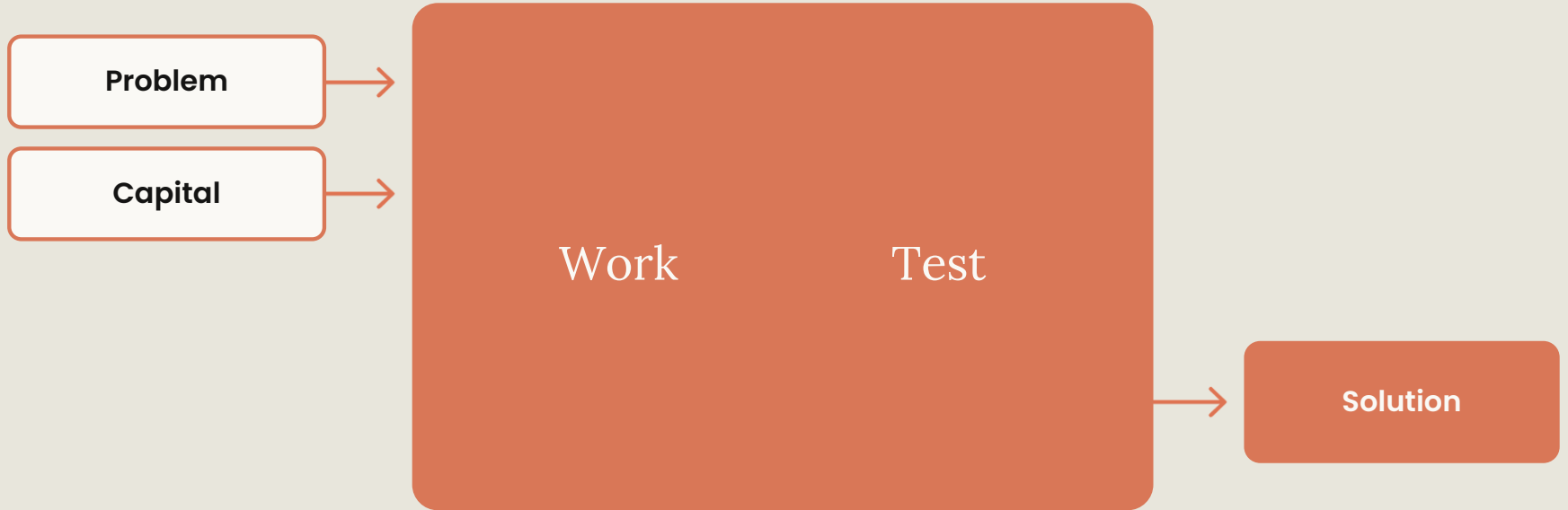
# Closed-loop iteration



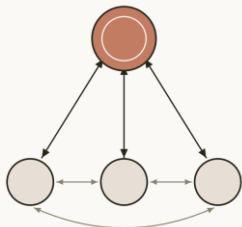
# Closed-loop iteration



# Closed-loop iteration

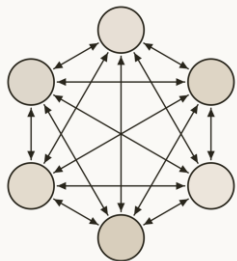


# Multi-agent research



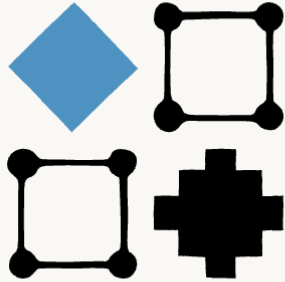
## Agent hierarchies

- Orchestrator controls a team and gives orders, stays high level CEO
- Core skills include delegation, communication
- Creator-Verifier loop ensures high quality results



## Fleets of peer agents

- Many agents with differing goals need to work together
- Agents may be cooperative or semi-adversarial
- Core skills include scaled communication, theory of mind, negotiation



# Harness design for long-running application development

Published Mar 24, 2026

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.

*Written by Prithvi Rajasekaran, a member of our [Labs](#) team.*

Over the past several months I've been working on two interconnected problems: getting Claude to produce high-quality frontend designs, and getting it to build complete applications without human intervention. This work originated with earlier efforts on our [frontend design skill](#) and [long-running coding agent harness](#), where my colleagues and I were able to improve Claude's performance well above baseline through prompt engineering and harness design—but both

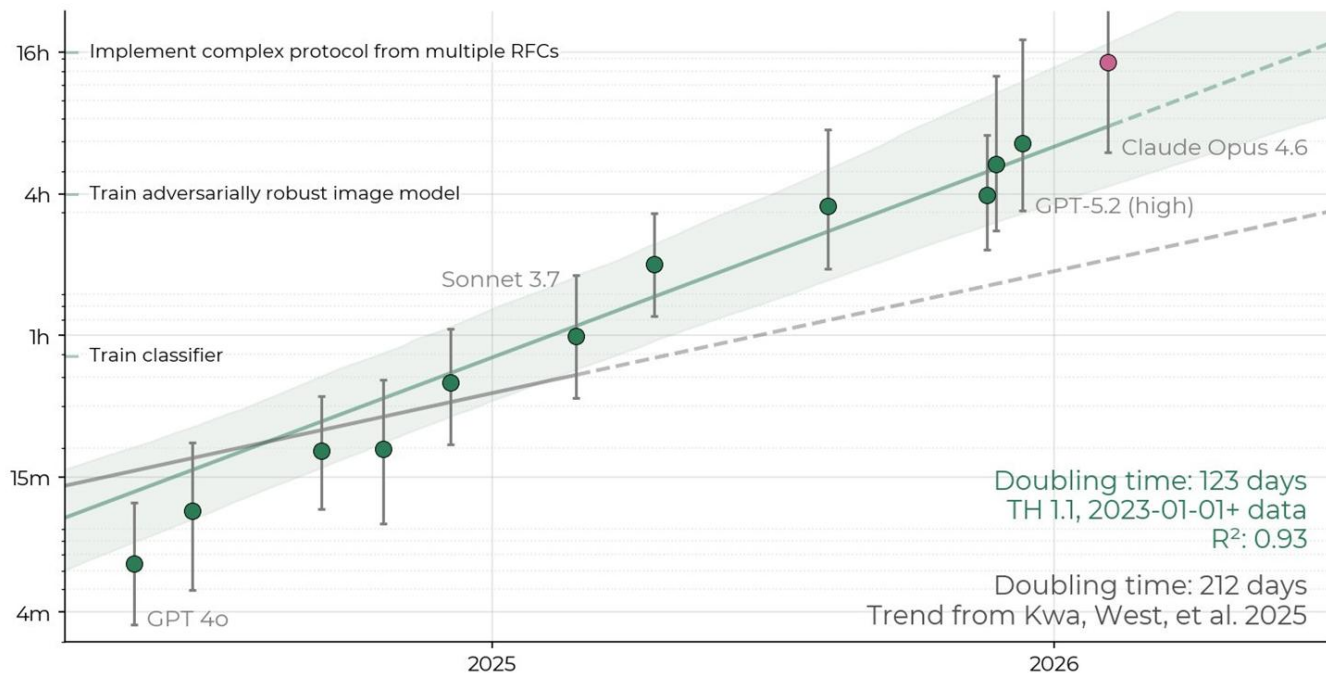


# Semiconductors on the Exponential

# Performance on Complex Tasks Doubles in 7 Months

Claude Opus 4.6 has a 50%-time-horizon of about 14.5 hours (95% CI: 6 hrs to 98 hrs)

Task length (at 50% success rate)



# Most models perform worse on debugging tasks with raw waveforms.

8

models *helped* by one-shot waveform debugging



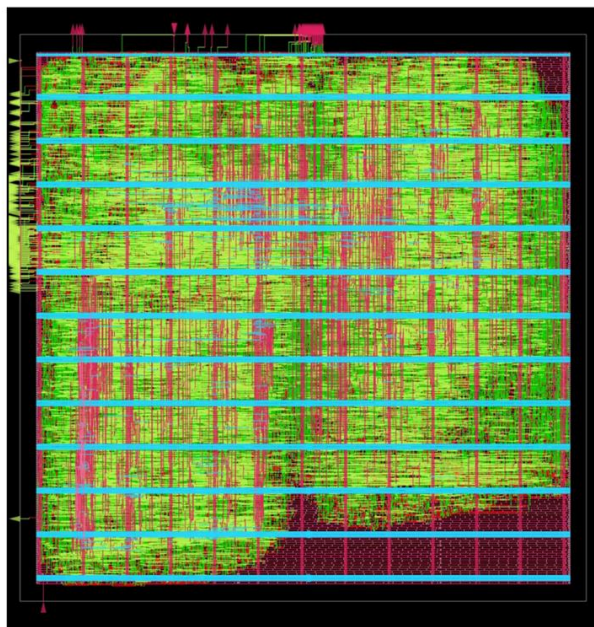
13

models *hurt* by one-shot waveform debugging

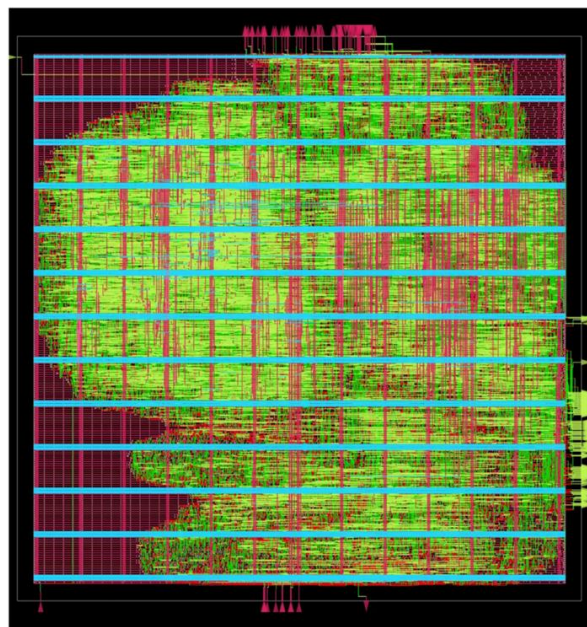


# Design Conductor: An agent autonomously builds a 1.5 GHz Linux-capable RISC-V CPU

Krishna, Krishna & Chin · Verkor · arXiv 2603.08716 · February 2026



(a) VerCore design A (higher performance)



(b) VerCore design B

Figure 1 GDSII plots of VerCore designs ( $70\ \mu\text{m} \times 70\ \mu\text{m}$ )

# From requirements to tape-out-ready in 12 hours

219

words of spec input

**RTL implementation** · five-stage in-order pipeline, register file in flops, early branch resolution, Booth-Wallace multiplier — all chosen by the agent.

12 h

total agent wall-time

**Functional verification** · Spike-driven cycle-by-cycle architectural-state checker; pass through MD5, CoreMark, and many regression programs.

3 variants

microarchitectures explored

**Physical design** · synthesis, place-and-route, parasitics, STA, power, DRC, LVS on OpenROAD with ASAP7.

3,261

CoreMark score

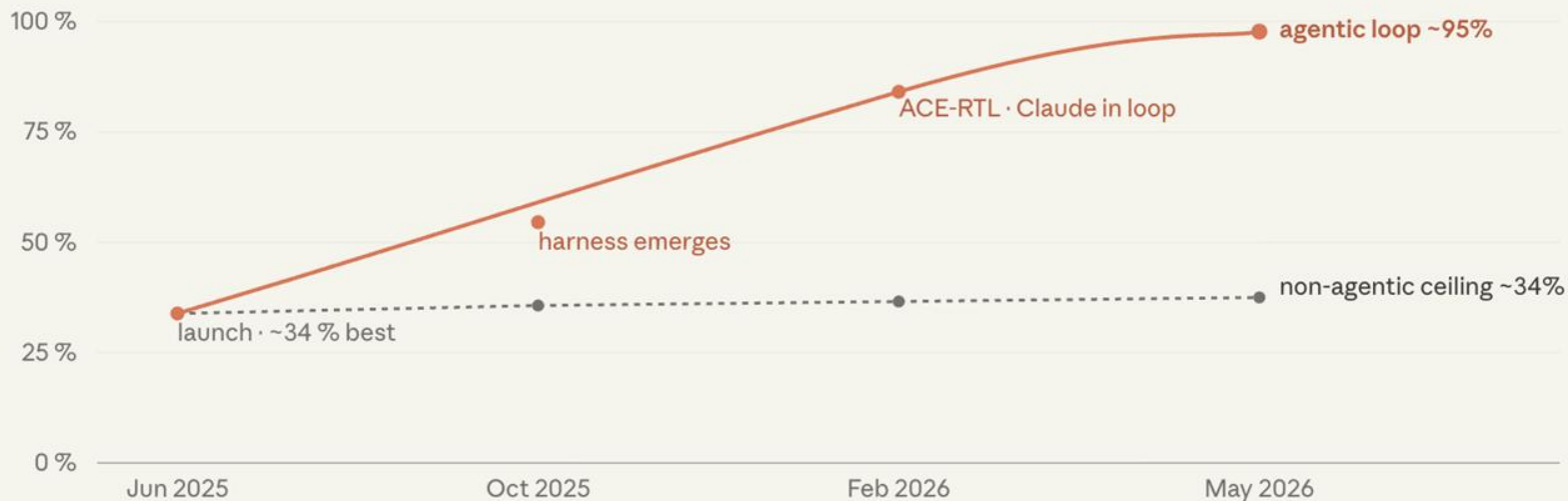
**PPA closure** · agent updated the RTL after reading STA reports, then re-ran the flow until timing closed at 1.48 GHz.

# NVIDIA's ACE-RTL agent reaches 80–96 % across Comprehensive Verilog Design Problems (CVDP) categories.

A generator / reflector / coordinator agent team, combining Claude Sonnet and domain-adapted RTL models, improved up to **+44.87 %** over prior state of the art

CVDP · NVIDIA's post-VerilogEval measuring stick

# Twelve months. Roughly the same models. Roughly 3× the pass rate.



CVDP · 783 problems, 13 categories · arXiv 2506.14074 · positions over time, code-gen subset · Best public agentic result · Best non-agentic single-shot.

# Catching a Ride

# Get the most from frontier models in design cycles

- 1 Make sure agents can access all technical information, designs, and rules

---

- 2 Build strong eval foundations: formalize rules, define “good”, use golden references

---

- 3 The verifier is the eval - iterate on your own sign-off flow

---

- 4 Translate raw checker outputs into useful feedback

---

- 5 Close the loop and pursue improved convergence

Treat the model like a strong but new engineer.

Give it the spec it needs.

Give it tools it can run.

Give it feedback structured enough to act on.

# Views from the Frontier

“

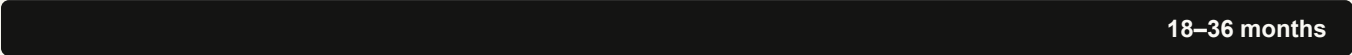
We observed instances in which models reason about Verilog, which is an event-driven language, as if it were sequential code.

We posit that this is due to the heavy presence of software code in both LLM pre-training and post-training, and we expect that, **as frontier labs recognize chip design as a more important application, this issue will be resolved.**

“Empowered with systems such as DC... teams will be able to tape out the most complex designs that exist in 3-6 months, instead of the current 18-36 months.”

## Today

~\$400M, hundreds of engineers, one design at a time



18–36 months

## Soon

Same headcount · many parallel explorations · closed-loop tooling



3–6 months

“...teams that now comprise 100 or more people working on a single design will be able to simultaneously explore many different designs, architecture, and product ideas, each one from concept to GDSII.”

“

We enumerate some of the “LLM hang-ups” we encountered in this work below.

“

We enumerate some of the “LLM hang-ups” we encountered in this work below.

For these reasons, **we believe it will remain critical to have experienced human architects guide systems** such as DC.

Thank you

**AI**

# Overview

- Introducing Anthropic
- The Intelligence Explosion
- Harnesses and Closed Loops
- Semiconductors on the Exponential
- Catching a Ride
- Views from the Frontier

# Five areas where frontier models fall short:

## 1 Read raw waveforms.

Pre-process; never hand it the .vcd.

---

## 2 Reason reliably about clock-domain crossings.

There isn't even a published benchmark for this yet.

---

## 3 Implement complex bus protocols end-to-end.

AXI, CHI — the model will sketch an interface and miss corner cases.

---

## 4 Lay out mixed-signal or analog blocks.

An EE-student vocabulary, not a designer's judgement.

---

## 5 Close timing autonomously on a hard advanced-node design.

Verkor: ~96 % of target on an open node. Hard nodes break the loop.

# Cowork:

## Delegate complex tasks to Claude



### Works with your files

Read, write, and organize files on your desktop



### Browses the web

Navigate pages, fill forms, and take action across sites with Claude in Chrome



### Connects your tools

Slack, Google Drive, Asana, and more via connectors

I have a data room for Tasselo Systems — can you analyze the documents and flag key risks?

### Revenue Reconciliation Analysis

Tasselo Systems, Inc. | Prepared January 2025 | Confidential

#### 1. Executive Summary

This analysis reconciles management reported revenue metrics against audited financials, contract level data, and sales system reports provided in the Tasselo Systems data room. Several material discrepancies were identified that warrant further diligence and may impact the valuation basis.

**Key Finding:** Reported ARR of \$22.1mm as of December 31, 2024 cannot be fully reconciled to underlying contract data. Adjusted ARR is estimated at \$20.3-\$21.4mm after normalizing for the Onstream management, Reported financials, and contracts with pending revenue.

#### 2. ARR Bridge — Reported vs. Adjusted

Item	Management Reported	Adjusted	Delta	Notes
Reported ARR (as stated)	\$22.1mm	\$22.1mm	—	As per Tasselo data
New Logo ARR	\$4.2mm	\$3.8mm	(\$0.4mm)	Excludes 2 unpledged Ch's covered by legal
Contract ARR	\$1.8mm	\$2.4mm	(\$0.6mm)	Contracts properly normalized to annual rate
Customer ARR	(\$1.1mm)	(\$1.1mm)	—	Contracted against short leg
Channel ARR	(\$1.4mm)	(\$1.8mm)	(\$0.4mm)	Includes Revenue partner discounts
Existing ARR (as stated)	\$20.5mm	\$22.2mm	(\$1.7mm)	Management Agns revenue includes items
Further adj. revenue (as stated)	(\$0.4mm)	(\$0.4mm)	—	Reconciles 2 contracts existing with revenue
Fully Adjusted ARR	\$21.4mm	\$21.4mm	—	Conservative basis for valuation

Source: Data Room — ARR Bridge & Contract Analysis, Meeting Pkgs., Customer Contract Book

#### 3. Revenue Quality Assessment

Revenue Type	Reported	Adjusted	Final	Final	Trend	Risk Level
Subscription Revenue	\$14.5mm (65%)	\$13.8mm (64%)	\$13.8mm (64%)	\$13.8mm (64%)	—	MODERATE
Professional Services	\$7.7mm (35%)	\$7.2mm (33%)	\$7.2mm (33%)	\$7.7mm (36%)	↑↑	ELLEVATED
Support & Maintenance	\$2.1mm (10%)	\$2.3mm (11%)	\$2.3mm (11%)	\$2.3mm (11%)	→	LOW
Other One-time	\$8.0mm (36%)	\$8.0mm (37%)	\$8.0mm (37%)	\$8.0mm (37%)	→	LOW
Total Revenue	\$22.1mm	\$21.4mm	\$21.4mm	\$21.4mm	→	—

**Observation:** Professional services revenue has nearly doubled as a percentage of total revenue, growing from 35% to 36% over two years. Management attributes this to implementation complexity of the v.8 platform release, but the

#### Working on it...

Documents

- Active deals
- Deal pipeline
- Tasselo Systems - data room

#### Progress

- Read and index 47 data room documents
- Cross-reference revenue claims against financials
- Flag risks and produce reconciliation report
- Synthesize findings into slide format
- Build IC-ready summary deck

#### Working folder

/Documents/outputs

- Revenue\_reconciliation.pdf
- Risk\_summary.xlsx
- IC\_deck\_draft.pptx

#### Context

##### Connectors

- S&P Global

##### Skills

- PDF
- XLSX
- PPTX

# Claude Data

Build, interrogate, explain, and publish models in plain English. Every employee gets a thought partner that knows the numbers.

Coming Soon

**Bongo** | Sample Project | model | DCF Overview | Data Feeds | Auto-refresh | Comment | Ask Claude | Export | 100%

### DCF Model Overview

Discounted Cash Flow — 5-Year Projection

ENTERPRISE VALUE	NPV	TERMINAL VALUE	INITIAL INVESTMENT
<b>\$1.55M</b>	<b>\$547.2K</b>	<b>\$1.82M</b>	<b>\$1.00M</b>

#### Cash Flow Projections

METRIC	YEAR 1	YEAR 2	YEAR 3	YEAR 4	YEAR 5	TERMINAL
Free Cash Flow	\$100.0K	\$105.0K	\$110.3K	\$115.8K	\$121.6K	—
Terminal Value	—	—	—	—	—	\$1.82M
Total Cash Flow	\$100.0K	\$105.0K	\$110.3K	\$115.8K	\$1.94M	—
Discount Factor	0.9091	0.8264	0.7513	0.6830	0.6209	—

#### Key Assumptions

Discount Rate	10.0%	Growth Rate	5.0%
Starting FCF	\$100.0K	Exit Multiple	15.0x
Investment	\$1.00M	Terminal Value	\$1.82M

#### Free Cash Flow Trend

Year	Free Cash Flow
Year 1	\$100.0K
Year 2	\$105.0K
Year 3	\$110.3K
Year 4	\$115.8K
Year 5	\$121.6K

#### Comparable Companies — EV/EBITDA

Company	EV/EBITDA Ratio
Alpha Corp	~8x
Beta Industries	~12x
Gamma Holdings	~6x
Delta Services	~7x
Epsilon Group	~10x

#### Key Takeaway

All four methods point to an enterprise value roughly in the \$1.5M–\$2.1M range, suggesting the \$1M investment is justified with meaningful upside. The conservative at ~\$1.55M, while precedent transactions suggest up to ~\$2.1M.

Describe what you want to create...

Ask questions  Auto verify

CONFIDENTIAL - NDA REQUIRED

# Claude Managed Agents

Claude Managed Agents handles agent infrastructure and execution so you can focus on your end-user experience.



## Sandboxing

Secure execution environments for agent code and tool use



## Long running agents

Long-running execution that survives crashes and disconnections, built in



## Multi-agent coordination

Agents spin up and coordinate other agents to parallelize work

## Claude Managed Agents architecture

