

TRANSMISSION № 014

2026.06.07 ≈ 9 MIN READ

Welcome to the Orchestration Era

*The moment AI stopped being about prompts and started being
about workflows – and the three-year arc that got us here.*

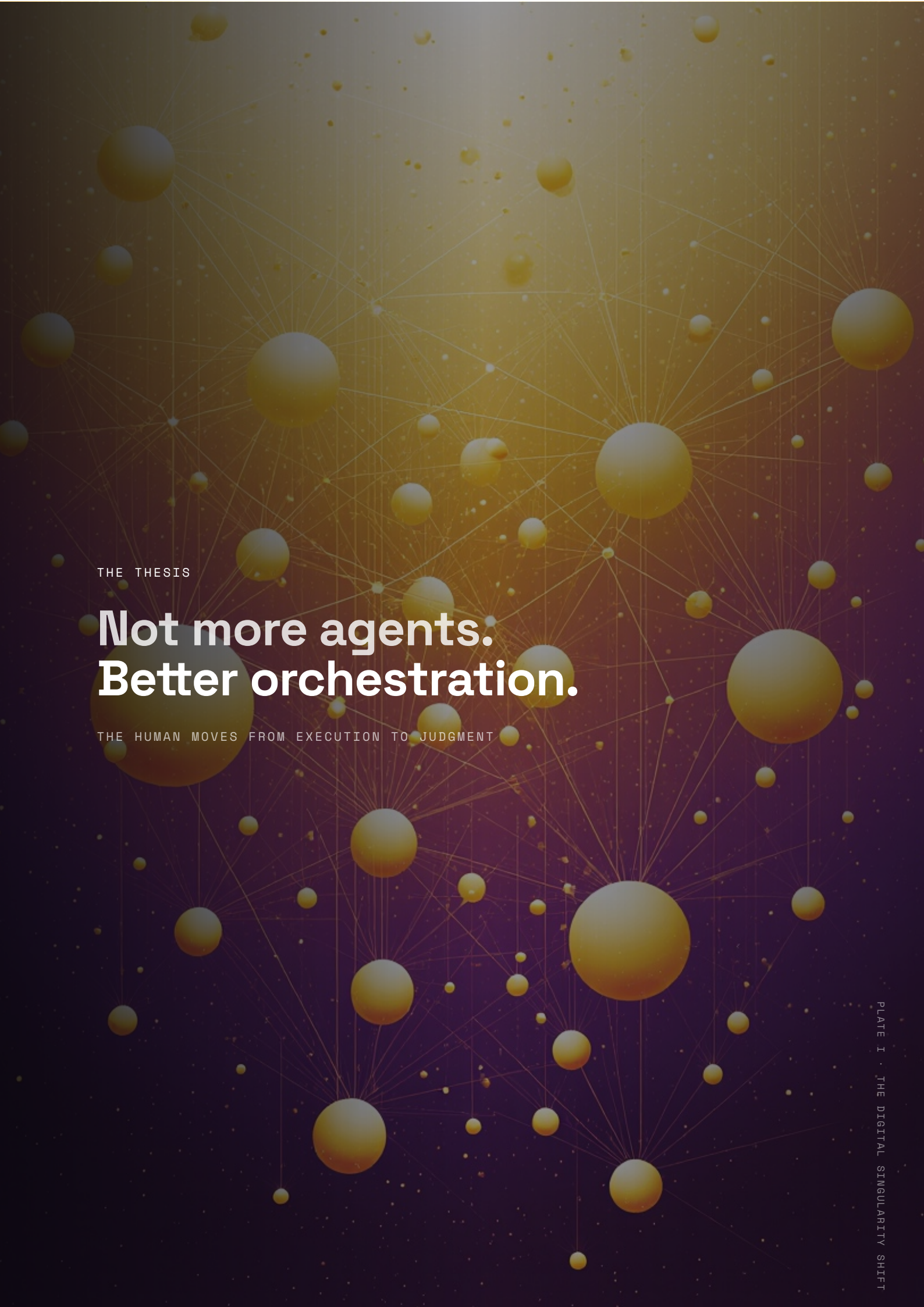
WORDS HERBERT CUBA GARCIA [CURAMANDO.COM](https://curamando.com)

On May 28, 2026, Anthropic shipped **Dynamic Workflows** alongside Claude Opus 4.8. The feature lets Claude write an orchestration plan, break it into subtasks, spin up hundreds of parallel subagents, and — before reporting anything back — have separate agents verify the results. It handles codebase-scale migrations across hundreds of thousands of lines of code, from kickoff to merge, with the existing test suite as its quality bar.

That release crystallised something. The question in the industry stopped being *can agents do real work?* — that was answered in H1. The new question: *how do we compose agents into repeatable, verifiable workflows that raise quality while reducing human involvement?*

This is the transition now underway: not more agents, but better orchestration. The human is becoming a workflow-author — the role that designs the choreography, defines what "good enough" looks like, and leaves the system to handle execution. It is a function closer to directing than to coding.

The shift didn't happen overnight. It took roughly three years. And each phase along the way introduced its own capabilities, its own vocabulary, and its own expertise. Looking back from here, the arc is surprisingly clean.



THE THESIS

Not more agents. Better orchestration.

THE HUMAN MOVES FROM EXECUTION TO JUDGMENT

Three years of AI in three axes

What follows is a retrospective – from the first Claude release to the present – with each era described along three axes.

01 What the AI can do

Capability – the frontier of what's technically possible.

02 What the human's job becomes

Role – where the person sits relative to the work.

03 Where the bottleneck sits

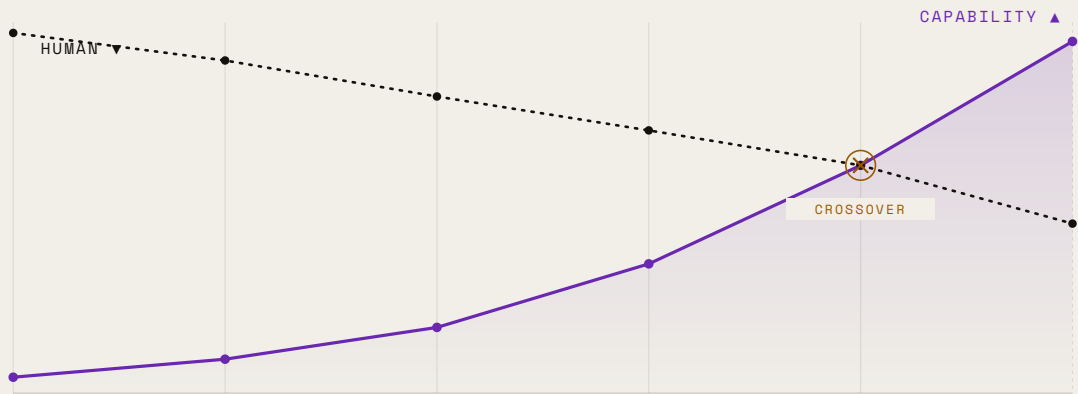
Constraint – the thing that actually limits output.

A human-role spine runs through it all: **prompter** → **context-provider** → **harness-builder** → **reviewer** → **workflow-author**. At each step, the human moved further from execution and closer to judgment.

Capability rises. Human involvement falls.

THE BOTTLENECK MIGRATES FROM CAPABILITY → TO ORCHESTRATION

CAPABILITY HUMAN INVOLVEMENT



2023	2024	2025	2026 · H1	2026 · H2	2027 →
PROMPTER	CONTEXT-PROVIDER	HARNESS-BUILDER	REVIEWER · STEERER	WORKFLOW-AUTHOR	OUTCOME-OWNER
BOTTLENECK CAPABILITY	BOTTLENECK CONTEXT	BOTTLENECK SCAFFOLDING	BOTTLENECK RELIABILITY	BOTTLENECK VERIFICATION	BOTTLENECK GOVERNANCE · ECONOMICS

BOTTLENECK MIGRATION CAPABILITY → CONTEXT → HARNESS → RELIABILITY → VERIFICATION → GOVERNANCE

2023

Era 01

The prompt as interface

CAPABILITY
INSTRUCT & RECEIVE

A model can be instructed in natural language and return useful output. GPT-4 lands in March. Claude follows. Teams build in a weekend what would once have taken months.

ROLE
PROMPTER

The practitioner sits at a chat window and types instructions. "Prompt engineering" enters the vocabulary – equal parts craft and superstition.

BOTTLENECK
RAW CAPABILITY

Models hallucinate, lose the thread, cannot use tools. RAG appears as a footnote – a means of inserting relevant documents into the context window – but it is a stopgap over a deeper problem. The models are simply not reliable enough for anything beyond drafts and prototypes.

prompt engineering

zero-shot

few-shot

chain-of-thought

hallucination

temperature

system prompt

This is the era of wonder. And wonder is useful – but it's not production.

2024

Era 02

Tools & context, **not just chat**

CAPABILITY

TOOL USE

Function calling and tool use become standard. Context windows expand to 128K tokens and beyond. Late in the year, reasoning models appear – OpenAI's o1 demonstrates that inference compute can be traded for better answers. AutoGPT generates enormous excitement, then reveals the gap between "autonomous" and "useful." The loop runs, but it runs off the rails.

ROLE

CONTEXT-PROVIDER

The job shifts from crafting perfect prompts to assembling the right information. What documents does the model need? What tools should it call? The work becomes curating inputs more than writing instructions.

BOTTLENECK

CONTEXT

The 128K window is impressive until it fills with noise. The real skill becomes knowing what to include and what to leave out.

function calling

tool use

reasoning models

context window

agent loop

rag

A lesson the industry keeps relearning: autonomy without structure produces chaos, not value.

CAPABILITY

THE SCAFFOLDING YEAR

Open-source LLMs explode – Qwen, DeepSeek and others close the gap at stunning speed. But the bigger shift is structural. Anthropic open-sources the Model Context Protocol – MCP, "USB-C for AI." By March, OpenAI adopts it; Google follows in April; by December it's donated to the Linux Foundation under the new Agentic AI Foundation. A protocol war that never happened. Instead: convergence.

ROLE

HARNESS-BUILDER

The discourse shifts too – Tobi Lütke frames the discipline as "context engineering," Karpathy amplifies, Willison endorses. The work becomes designing the system around the model: memory, tool selection, retrieval pipelines, guardrails. The model is one component. The harness is the product.

BOTTLENECK

SCAFFOLDING

The model is good enough. The protocol exists. But building a reliable harness – one that handles edge cases, manages state, recovers from errors – is genuinely hard engineering. Not prompt engineering. Software engineering.

context engineering

mcp

harness

scaffolding

agentic framework

tool orchestration

2025 · THE TURN

The expertise shifted from linguistics to systems design.

NOT PROMPT ENGINEERING — SOFTWARE ENGINEERING

CAPABILITY
REAL, UNSUPERVISED
WORK

Agents become reliable enough to deploy on real work. Not demos. Not prototypes. Production tasks with actual stakes – code review, migrations, data pipelines, content operations – handled end-to-end, with human oversight measured in minutes per day, not hours.

ROLE
REVIEWER & STEERER

Engineers run multiple parallel agents simultaneously – monitoring progress across workstreams, intervening when an agent gets stuck, steering priorities. The ratio of thinking to typing inverts.

BOTTLENECK
RELIABILITY &
INFRASTRUCTURE

Infrastructure, not intelligence, is now the constraint. Individual agents work – but they fail in ways that are hard to predict. They need the operational maturity that took traditional software two decades to develop. Agents need it now: authentication, state management, error recovery, audit trails, deployment pipelines.

agentic production

fleet management

agent reliability

human-in-the-loop

devops for agents

Dynamic Workflows is the canonical example of where we are now. An orchestrator plans the work, distributes it across specialised subagents running in parallel, and then – the critical piece – separate agents challenge and verify the results before anything gets reported back. Claude doesn't just do the work. It checks the work. Multiple times, from different angles.

Verification is the hidden protagonist of this entire three-year arc. Each era moved the bottleneck: capability → context → harness → reliability → and now, verification. Adversarial review and consensus between agents is what makes the jump from "production" to "workflow" possible. Without it, the result is chaos at scale. With it, quality at scale.

Evals become the moat. The teams that can define what "correct" looks like, measure it rigorously, and build it into their workflows will pull away from those that can't. The model is commodity. The eval is competitive advantage. Standards matter now too: MCP gave us agent-to-tool connectivity; what's emerging is agent-to-agent connectivity – protocols for handoff, verification and state across multi-agent workflows. And governance stops being a compliance checkbox: when agents modify a hundred files across a codebase, *who is accountable?*

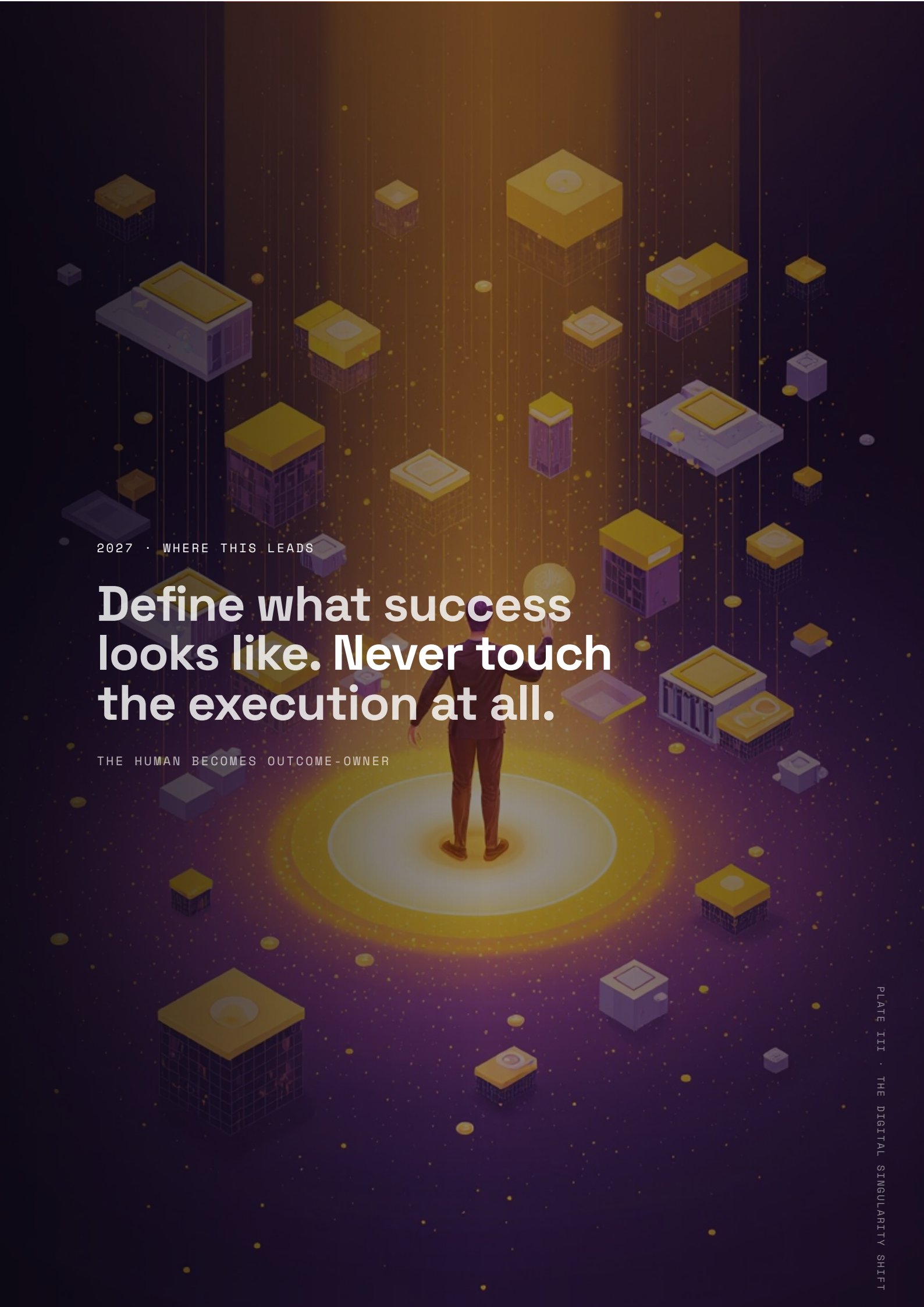
dynamic workflows

orchestration plan

adversarial verification

agent-to-agent protocol

workflow-author

A person in a dark suit stands with their back to the viewer on a glowing yellow circular platform. They are holding a glowing sphere in their right hand. The background is a dark purple space filled with numerous floating 3D cubes of various sizes and colors (yellow, purple, white) and small yellow dots, resembling a digital or futuristic environment.

2027 · WHERE THIS LEADS

Define what success looks like. Never touch the execution at all.

THE HUMAN BECOMES OUTCOME-OWNER

Where this leads

These are posed as open questions, not predictions – though the clock is already running on both.

The economics question

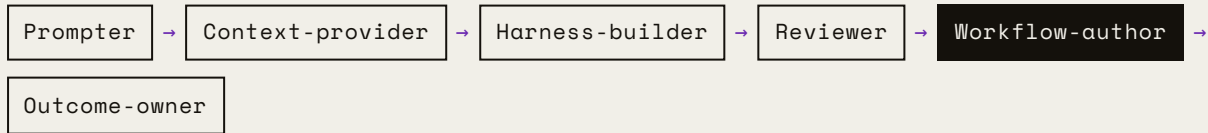
This is the question that decides whether the orchestration era scales or stalls. When a single workflow spins up hundreds of agents in parallel, compute stops being a line item and becomes the business model. One migration can burn more tokens in an afternoon than a developer costs in a month – and the meter never stops. Does agentic work create value faster than it consumes compute? It can. But only for the organisations that treat it as infrastructure to be engineered, not magic to be bought. Model routing becomes a core competency; local and on-prem execution come roaring back – for cost, but also for data sovereignty. The ones who never learn to measure it will watch their margins evaporate one invisible token at a time. **This is FinOps for agents – and the bill is already arriving.**

The autonomy question

Not “can agents do the task?” – that's settled. The real question, the one most organisations aren't ready for: can a single workflow swallow an entire function end-to-end – the work of a whole department – and when it does, *who signs?* Organisations were never just technology. They're accountability structures, legal entities, cultures of trust. An agent can ship the work; it cannot be held responsible for it. That gap doesn't close – it moves. The human role hardens into **outcome-owner**: a role that defines what “good” means, carries the consequences, and may never touch a single step of the execution. **The unsettling part isn't that the machines act without us – it's that we stay accountable for everything they do.**

Three years. Five eras. One pattern.

The human moved further from execution and closer to judgment at every step. Each era introduced new capabilities, new vocabulary, new expertise — and each one made the previous era's hard problems feel almost quaint.



Verification ties it all together. Every bottleneck in this timeline was, at its core, a verification problem in disguise. Can the model produce correct output? Can it use the right context? Can the harness catch errors? Can the workflow prove its results are sound?

The organisations that thrive in the orchestration era will not be the ones with the best models or the most agents. They will be the ones that became expert at defining, measuring, and verifying what "good" looks like.

That has always been the hard part. The tools have finally caught up.

The vocabulary index

Five eras, each minting its own literacy. A field guide to the words that defined the arc.

2023

THE PROMPT

NEW LITERACY, BUILT FROM
SCRATCH IN MONTHS

prompt engineering

zero-shot

few-shot

chain-of-thought

hallucination

temperature

system prompt

2024

TOOLS & CONTEXT

FROM LINGUISTICS TOWARD
INPUTS

function calling

tool use

reasoning models

context window

agent loop

retrieval-augmented generation

2025

THE HARNESS

LINGUISTICS → SYSTEMS
DESIGN

context engineering

MCP

harness

scaffolding

agentic framework

tool orchestration

2026 • H1

AGENTIC PRODUCTION

DEVOPS FOR AGENTS

agentic production

fleet management

agent reliability

human-in-the-loop

deployment pipelines

2026 • H2

ORCHESTRATION

THE EXPERTISE IS
CHOREOGRAPHY

dynamic workflows

orchestration plan

adversarial verification

agent-to-agent protocol

workflow-author

evals as moat