May 1, 2026

Context Engineering Architecture 2.0

From service design to context design

What this article is about

Navigation Plan

Analogy → enterprise strategy

- Part 1: ’LLM service’ lens.
- Part 2: Architecture 1.0 vs 2.0.
- Part 3: signals 2025–2026.
- Parts 4–5: flexibility cost and strategy.

Part 1

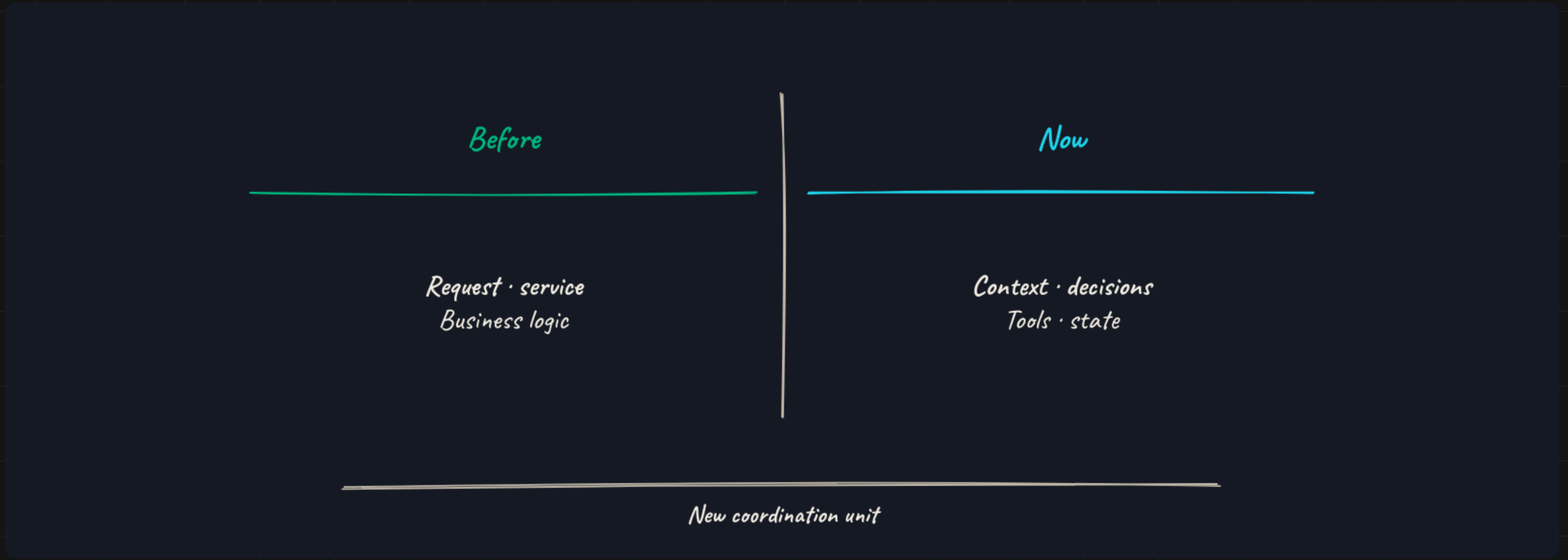
01

Historical Analogy

Why 'just add an LLM service' is too narrow a lens

Architectural shifts are underestimated

Analogy History



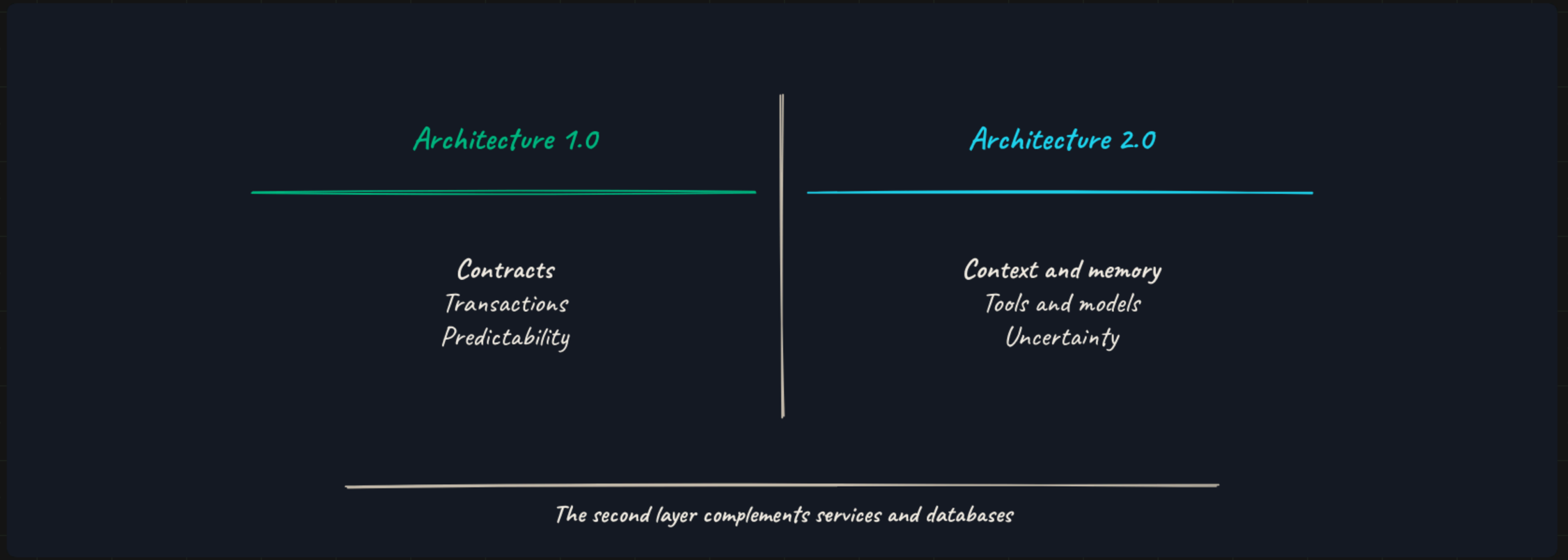
Part 2

02

Two Architecture Modes

Architecture 1.0 — deterministic systems, Architecture 2.0 — managing uncertainty

Architecture 1.0 vs 2.0



Part 3

03

Industry Signals 2025–2026

7 trends already changing architecture

Context and durable execution

Context → separate layer

From prompt engineering
to context engineering
More tokens ≠ better

Durable execution

Minutes, hours, weeks
Background mode + checkpoints

The two most underestimated shifts in AI-native architecture

Memory and tool layer

Memory

Short and long term

Facts · experience · procedures

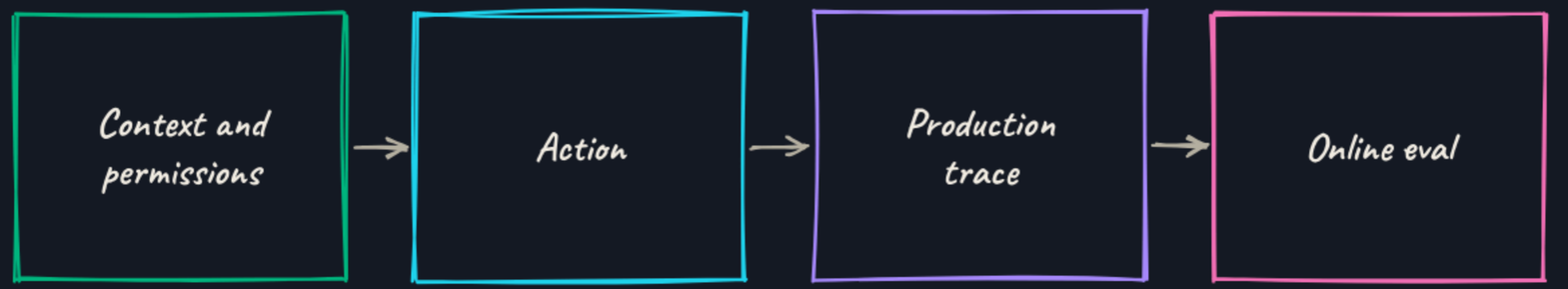
Tools

Name, schema, contract

Discovery by context

Evaluation and security

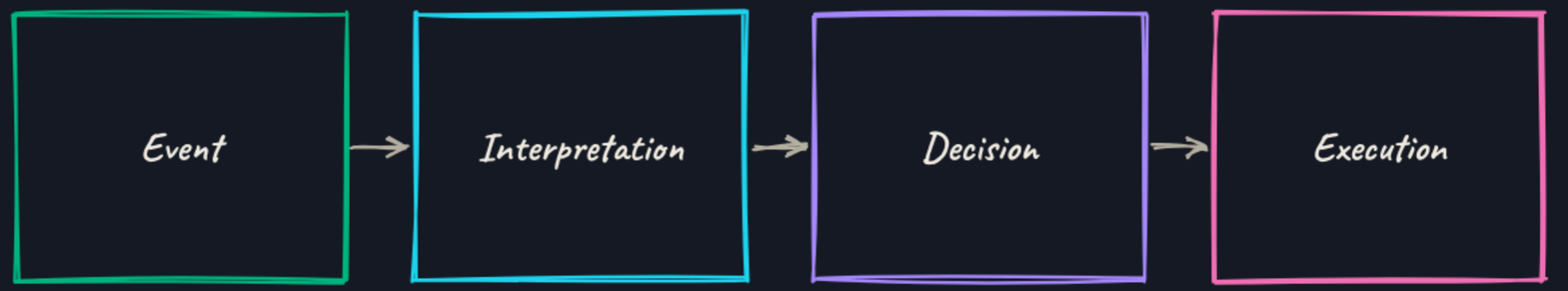
Trends 5-6 Signals



Least privilege limits tools and scopes

Event-driven + cognitive and scale in tokens

Trend 7 Signals



Compaction and prompt caching manage context cost

Part 4

04

Architecture 2.0 is not free

The trends bring new risks: four failure modes

Flexibility is not free

Cost Risks

Uncontrolled context	Stateless runtime	Excessive permissions	No online eval
Recall, latency, cost	Lost work	Injection risk	Blind operations

Part 5

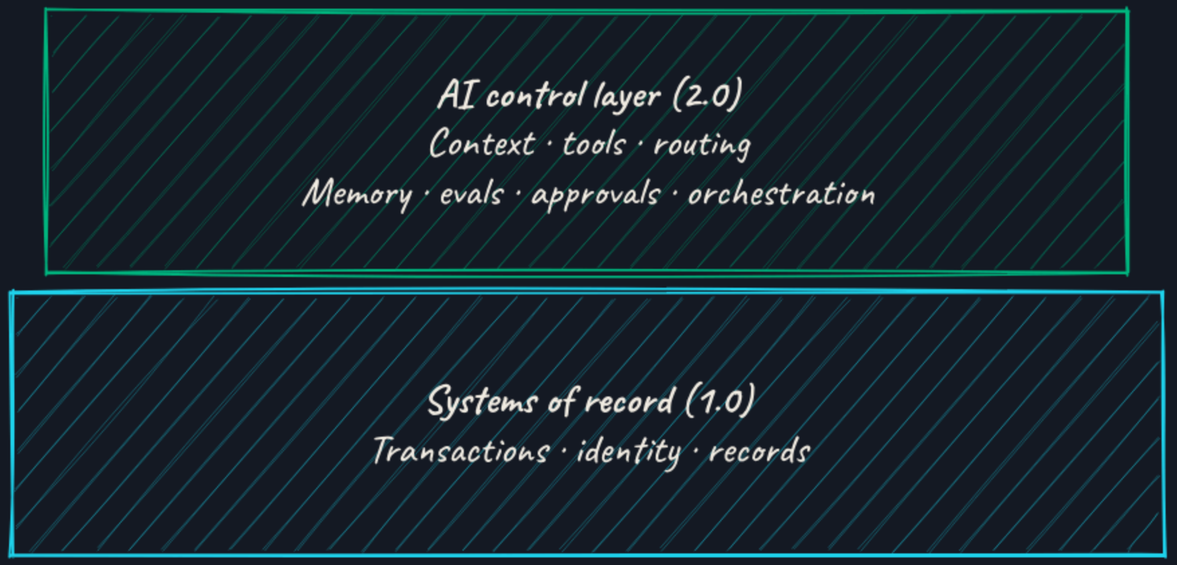
05

Enterprise Strategy

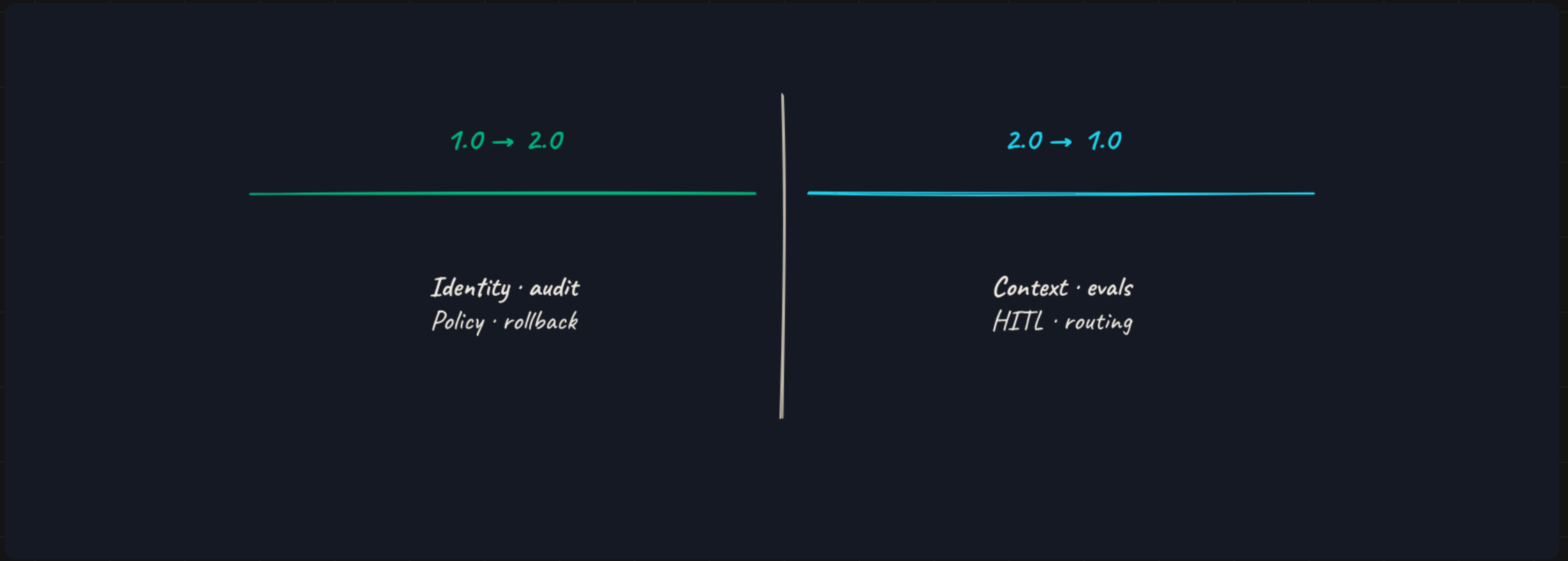
AI as a control plane over legacy, platform, and knowledge systems

Control layer over the landscape

Strategy Enterprise



What to borrow from where



Verdict

From fixed paths to context

- Design unit: context, decisions, execution.
- Architecture 2.0 manages uncertainty over 1.0.
- Enterprise: AI = control plane over legacy.

- ↗tellmeabout.tech · ↗ai4sdlc-research.space

References and Materials

Sources References

Agentic architecture, context, security

- Anthropic: Building effective agents.
- MCP and tool integration.
- OpenAI Agents SDK and Responses API.
- LangGraph agents + OWASP LLM Top 10.

Architecture 1.0 → 2.0

Thank you!

Materials and links are in the ”Book Cube” channel

Alexander Polomodov

Technical Director & Fellow,
T-Technologies

