fluree

# The Six Levels of the Autonomous Enterprise

**Where your organization actually sits on the AI maturity curve** — and what it will take to move up.

fluree

WHY WE NEED A NEW MAP

# Record investment. Record activity. Record disappointment.

**Three numbers, three independent sources, one condition:** enterprise AI is widely adopted and rarely producing measurable business outcomes.

## 95%

of enterprise AI pilots delivered no measurable P&L impact.

MIT NANDA · THE GENAI DIVIDE · 2025

## 40%+

of agentic AI projects projected to be canceled by end of 2027.

GARTNER FORECAST

## 90%

of workers use personal AI tools; only 40% of employers sanction subscriptions.

SHADOW AI ECONOMY · MIT

fluree

THE REAL DIVIDE

# Most organizations are running higher-autonomy AI on lower-autonomy infrastructure.

**The ceiling on your AI program is almost never your AI.** It's the substrate underneath it.

fluree

HOW TO READ THE LEVELS

# Each level is read across five dimensions.

The SAE self-driving taxonomy describes what the car can do. This framework inverts that — the levels describe what the enterprise can **support**. A high-autonomy agent on a low-autonomy substrate is a liability, not a capability.

01 / Dimension

## Human role

What humans actually do — operator, reviewer, supervisor, strategist, policy-setter.

02 / Dimension

## Substrate

What the AI retrieves from — prompt context, indexed chunks, semantic graph.

03 / Dimension

## Agent capability

The scope of action the organization can safely extend — read, write, cross-domain.

04 / Dimension

## Governance

Where policy lives — email memo, app-layer filter, prompt guardrail, data-layer enforcement.

05 / Dimension

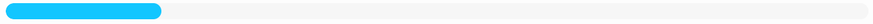
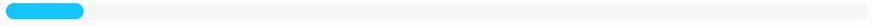
## ROI profile

Whether AI is producing measurable outcomes or masked costs hidden in review overhead.

fluree

THE LADDER

# Two-thirds of organizations sit below the line where AI begins to compound.

| LEVEL | STAGE | DISTRIBUTION | HUMAN ROLE | % OF ORGS |
|---|---|---|---|---|
| 1 | **Shadow**<br>Unmanaged, ad-hoc use |  | Individual employees | **25–30%** |
| 2 | **Assisted**<br>Sanctioned copilots, task-level lift |  | Operator | **40–45%** |
| 3 | **Integrated**<br>RAG over internal documents |  | Reviewer | **~20%** |
| 4 | **Contextual**<br>Semantic graph, grounded answers |  | Supervisor | **8–10%** |
| 5 | **Autonomous**<br>Bounded cross-domain action |  | Strategist | **3–5%** |

fluree

FRAMEWORK AT A GLANCE

# Six levels read across five dimensions.

| DIMENSION | 01 Shadow | 02 Assisted | 03 Integrated | 04 Contextual | 05 Autonomous | 06 Self-Govng |
|---|---|---|---|---|---|---|
| **HUMAN ROLE** | Individual employees | Operator | Reviewer | Supervisor | Strategist | Policy-setter |
| **SUBSTRATE** | None | Prompt context only | Vector embeddings | Semantic graph | Governed knowledge graph | Living ontology |
| **AGENT CAPABILITY** | Single-turn completions | Single-task assistance | Retrieval, single domain | Cross-silo reasoning, read | Bounded cross-domain action | Multi-agent coordination |
| **GOVERNANCE** | Policy-by-email | App-layer filters, DLP | Prompt guardrails | Data-layer access controls | Policy at the data layer | Provenance-native, auditable |
| **ROI PROFILE** | Mostly negative | Productivity lift, not P&L | Narrow savings, review cost | Process savings, P&L visible | Revenue-influencing | Business-model-altering |

fluree

# Shadow

**25–30%**
OF ORGANIZATIONS

**The technology is present — it's just unmanaged.** The CIO either doesn't know it's happening, or has decided not to ask.

| | |
|---|---|
| HUMAN ROLE | Individual employees making ad-hoc tool choices |
| SUBSTRATE | None — AI sees only what an employee pastes in |
| AGENT CAPABILITY | Single-turn completions; personal productivity only |
| GOVERNANCE | Policy-by-email, unenforced |
| ROI PROFILE | Mostly negative when accounting for risk exposure |

### TYPICAL USE

Drafting emails, summarizing meetings, first-pass code, rephrasing slide content.

### BLOCKER TO NEXT LEVEL

No executive accountability. No AI owner, no budget line, no mandate.

### FAILURE MODE

Data leakage. A sales leader pastes a customer list into a consumer chatbot — and creates a regulatory incident no one can trace.

fluree

# Assisted

**40–45%**
OF ORGANIZATIONS

**The AI augments individual tasks but doesn't know the business.** It can rephrase your email; it can't tell you whether the customer you're emailing is past due.

| | |
|---|---|
| HUMAN ROLE | Operator — human does the work, AI assists the step |
| SUBSTRATE | None at the enterprise level; prompts carry context |
| AGENT CAPABILITY | Single-task assistance inside one application |
| GOVERNANCE | Application-layer — tenant isolation, DLP on prompts |
| ROI PROFILE | Productivity lift, rarely P&L-visible at enterprise scale |

### TYPICAL USE

Email drafting, code completion, document summarization, meeting transcription, template generation.

### BLOCKER TO NEXT LEVEL

The AI has no memory of the business. Every prompt starts from zero; users compensate by pasting more context.

### FAILURE MODE

Productivity theater. Usage high, P&L unchanged after 18 months. The heart of the GenAI Divide.

fluree

# Integrated

**~20%**

OF ORGANIZATIONS

**For the first time, the AI knows something about the business.** It also, for the first time, can be confidently wrong about it. Accuracy reaches 70–80% — impressive in a demo, intolerable in production.

| | |
|---|---|
| HUMAN ROLE | Reviewer — AI proposes, human verifies before action |
| SUBSTRATE | Indexed documents, vector embeddings, chunked text |
| AGENT CAPABILITY | Retrieval-grounded answers within a single domain |
| GOVERNANCE | Prompt-layer filters and output guardrails |
| ROI PROFILE | First measurable savings — often offset by review overhead |

### TYPICAL USE

Internal KB search, customer-support assist, contract summarization, first-pass research, HR policy Q&A.

### BLOCKER TO NEXT LEVEL

Vector-only retrieval loses the structure and meaning of enterprise data the moment it hits a vector store.

### FAILURE MODE

"Two good answers, then a hallucination." Pilots succeed in demos and fail when users go off-script. Most budgets get cut here.

fluree

# 4 Contextual

**8–10%**
OF ORGANIZATIONS

**A semantic layer appears in the architecture.** Ontologies replace folder structures. The AI receives pre-connected context — and answers become auditable because every fact has a source.

| | |
|---|---|
| HUMAN ROLE | Supervisor — AI executes within domain, human handles exceptions |
| SUBSTRATE | Semantic knowledge graph, ontology, unified entity model |
| AGENT CAPABILITY | Grounded cross-silo reasoning; read-heavy actions |
| GOVERNANCE | Data-layer access controls; queries enforce policy automatically |
| ROI PROFILE | Process-level savings; first genuine P&L visibility |

### TYPICAL USE

Conversational analytics, customer-360 interactions, regulated research copilots, portfolio risk, supplier-360.

### BLOCKER TO NEXT LEVEL

Policy must travel with the data — not the application — so rules apply equally to humans and agents.

### FAILURE MODE

The read-only plateau. Trusted for answering questions. The moment leadership asks it to act, legal freezes the program.

fluree

# Autonomous

**3–5%**

OF ORGANIZATIONS

**Agents don't just retrieve and recommend — they act.** The data itself enforces policy. The enterprise stops teaching the agent one API call at a time and starts giving the agent a model of the business to operate on.

| | |
|---|---|
| HUMAN ROLE | Strategist — humans define goals, policies, exceptions |
| SUBSTRATE | Governed knowledge graph with read/write semantics and provenance |
| AGENT CAPABILITY | Bounded cross-domain action with embedded policy enforcement |
| GOVERNANCE | Data-layer policy; provenance and audit native to every transaction |
| ROI PROFILE | Revenue-influencing; automation of full processes, not tasks |

### TYPICAL USE

End-to-end customer service resolution, reconciliation, supply-chain re-routing, underwriting agents, regulatory reporting.

### BLOCKER TO NEXT LEVEL

Organizational readiness — not technology. Tech closes faster than the muscle for policy-only oversight.

### FAILURE MODE

Runaway cost and trust-loss incidents. Without mature governance, agents make plausible but policy-violating decisions.

fluree

# Self-Governing

<1%

TODAY · PLAUSIBLE BY 2030

**Humans don't manage workflows — they set the rules the system operates by.** Agents coordinate across functions and escalate only when policy demands it. The knowledge graph becomes corporate memory — surviving turnover, exposing one consistent truth to every system that queries it.

| | |
|---|---|
| HUMAN ROLE | Policy-setter — humans govern the system, not the cases |
| SUBSTRATE | Living ontology; the data model is the business model |
| AGENT CAPABILITY | Multi-agent coordination across the enterprise and beyond it |
| GOVERNANCE | Provenance-native, continuously verifiable, auditable by design |
| ROI PROFILE | Business-model-altering; AI becomes a factor of production |

fluree

WHY PILOTS STALL

# A high-autonomy agent on low-autonomy infrastructure is a liability.

**Stop evaluating agents.** Start evaluating substrates.

## MISMATCHED · PILOT STALLS

AGENT
**Level 05 — autonomous, cross-domain**

LOAD > CAPACITY

GOVERNANCE
**App-layer filters, policy-by-email**

SUBSTRATE
**Level 02 — prompts & pasted context**

The agent acts on data it cannot understand, secure, or trust. The pilot stalls in QA.

## ALIGNED · PILOT COMPOUNDS

AGENT
**Level 05 — autonomous, cross-domain**

GOVERNANCE
**Policy at the data layer**

SUBSTRATE
**Level 05 — governed knowledge graph**

Every read, write, and action checks role, classification, and provenance before completing.

fluree

LEVEL 03 → LEVEL 04

# From vector retrieval to semantic graph — where accuracy breaks the 80% ceiling.

LEVEL 03 / VECTOR-ONLY RAG

## Retrieving chunks of text without structure or meaning.

**~80%**

accuracy ceiling on complex questions

- Chunked documents in a vector store
- Model infers relationships it can't verify
- Plausible-sounding answers, subtly wrong
- Scope narrows until the use case stops mattering

LEVEL 04 / SEMANTIC GRAPH · GRAPHRAG

## Pre-connected context. Every fact has a source.

**95%+**

accuracy on the same workloads

- Ontology and unified entity model
- "ACME Corp" in CRM = "Acme Corporation" in ERP
- Cross-silo queries resolve in a single pass
- Every answer auditable to its source

fluree

WHAT THE FRAMEWORK ARGUES

# Three implications for the next two years of AI spending.

01

## Stop evaluating agents. Start evaluating substrates.

**"How good is this agent?"** gets answered in a curated demo. **"What level of autonomy can our data actually support?"** forces an honest answer. An organization at Level 03 buying the best Level 05 agent stays at Level 03.

02

## The Level 03 → 04 jump is the one that matters most.

Where accuracy breaks 80%, where cross-silo reasoning becomes possible, where ROI stops being a productivity story and starts being a business story. Programs that have made it are compounding; programs that haven't are being reviewed by the CFO.

03

## The question in 2026 is structural, not technical.

**What level is our substrate at, and what would it take to move it up one?** That question has a concrete answer, a budget line, and a direct correlation with whether the next round of AI investment produces outcomes or stalled pilots.

fluree

THE FASTEST DIAGNOSTIC QUESTION

# What level of autonomy could our current data infrastructure safely support tomorrow, with no new models added?

### WHAT MOST LEADERS FIND

When they answer truthfully, their organization sits **one level lower** than they thought.

### WHAT CHANGES WHEN YOU MOVE UP

The next round of AI investment produces measurable outcomes — instead of another round of stalled pilots.