March 2026

initiative for applied artificial intelligence

# AI Maturity 2026: The Year Agents Redefine Value Creation

**“Moving at the speed of 2025 in 2026 means falling behind.** Agents transformed AI from a tool we use into a workforce we lead. The question is no longer if you adopt – it’s whether you can rethink how work gets done. In 2026, adaptation speed is the new competitive moat.”

# The AI Landscape 2026: Four Hypotheses That Matter

The gap between AI's promise and its enterprise reality has never been wider, or more consequential. While foundation models have become commoditized and agent frameworks proliferate, most organizations still struggle to translate capability into sustained value. The companies that will lead in 2026 and beyond are not those with the most advanced technology, but those that understand the structural shifts underway and act on them decisively.

This chapter presents four hypotheses that define what matters in 2026. They are deliberately pointed. If you find yourself disagreeing, that disagreement is worth examining. It may reveal assumptions that no longer hold due to the technological developments between the time of writing and that of reading.

The game has shifted. The exercises that built AI capability in 2023 (prompt engineering, chatbot pilots, basic RAG implementations) are now table stakes. Staying competitive requires a new training regimen, and the organizations that fail to update their approach will find themselves falling behind, no matter how hard they work.

## 1. Agents Are a Step Change, Not an Increment

The shift from AI assistants to AI agents is not a feature upgrade. It is a fundamental change in what AI systems can do and what they require from organizations. More profoundly, it transforms AI from a tool to a user, from an instrument that humans operate to a workforce that operates alongside them.

Assistants respond to prompts. They augment human work by generating text, answering questions, or summarizing documents. The human remains in the loop for every decision and action. This is the paradigm most enterprises have operationalized: useful, but bounded.

Agents pursue goals. They decompose complex objectives into subtasks, use tools, interact with systems, and adapt their approach based on intermediate results, often across multiple steps without human intervention. The human defines the objective and constraints; the agent determines how to achieve it.

This distinction has profound implications:

- **Autonomy requires trust infrastructure.** When an agent can take actions (send emails, modify data, trigger workflows) the question shifts from "Is the output good?" to "Should this action be allowed?" Permissions, guardrails, and audit trails become operational necessities, not compliance exercises.
- **Evaluation becomes continuous.** A chatbot can be evaluated on response quality. An agent must be evaluated on goal achievement, action safety, cost efficiency, and behavioral consistency across contexts. This demands new metrics and new disciplines.
- **Failure modes are different.** Assistants fail visibly: wrong answers, poor formatting. Agents can fail silently, taking suboptimal paths, accumulating costs, or drifting from intent in ways that only surface later.

The evidence is compelling. Recent benchmarks like OpenAI's GDPval [1] suggest that agentic systems can automate 25-40% of knowledge work tasks with current capabilities, not in theory but in controlled enterprise deployments. The MIT "Iceberg Study" [2] documents that the visible productivity gains from GenAI represent only a fraction of the value unlocked when agents are integrated into core workflows.

The question is no longer whether agents work. It is whether your organization can learn to lead them.

*For a deep dive into agentic architectures, governance patterns, and implementation strategies, see the appliedAI whitepaper "AI Agents in Action."*

## 2. Context Engineering Becomes a Core Competency

The era of prompt engineering (crafting clever instructions to coax better outputs from language models) is giving way to something more fundamental: context engineering.

Context engineering is the systematic discipline of providing AI systems, especially agents, with the right information, at the right time, in the right structure, to perform optimally. It encompasses:

- **Data provisioning:** Which knowledge sources, documents, and data streams should an agent have access to? How is freshness ensured? How is sensitive information protected?
- **Process understanding:** What does the agent need to know about the workflow it operates in? What are the upstream dependencies and downstream consequences of its actions?
- **Tool access:** Which APIs, systems, and capabilities can the agent invoke? Under what conditions? With what permissions?
- **Boundary definition:** What is the agent explicitly not allowed to do? Where must it escalate to humans? How are these constraints encoded and enforced?

This is not prompt engineering at scale. It is systems design for AI-native operations.

The difference becomes clear when agents fail. Poor prompt engineering produces bad outputs. Poor context engineering produces agents that confidently take wrong actions because they were operating on incomplete, outdated, or inappropriate information. They hallucinate not from model weakness, but from context starvation.

Organizations that treat context as an afterthought will build agents that impress in demos and fail in production. Those that invest in context engineering as a discipline (with dedicated roles, tooling, and governance) will build agents that compound in value over time.

If you are not building context engineering as an organizational capability, you are preparing to deploy agents that will be impressively wrong and could cause real damage.

[1] Source: https://arxiv.org/abs/2510.04374

[2] Source: iceberg.mit.edu

## 3. End-to-End Process Redesign Is No Longer Optional

Most enterprises approach AI the same way they approached previous technology waves: identify existing processes, find friction points, apply AI to reduce friction. This is intuitive, low-risk, and almost always suboptimal.

The pattern is familiar:

- Document processing is slow → Add AI-powered OCR (optical character recognition) and extraction
- Customer inquiries take time → Deploy a chatbot for first-level triage
- Sales forecasting is manual → Build a predictive model on historical data

Each intervention delivers local improvement. But the process architecture remains unchanged. And process architecture is where the real constraints live.

This changes the game for use case identification. The bottom-up approach that served enterprises well (collecting use case ideas from business units, prioritizing by feasibility and value, building pilots) remains necessary but is no longer sufficient. When AI can transform entire value streams, organizations need a complementary top-down perspective: Which core processes should be fundamentally reconceived? Where does the current architecture create structural limitations that no amount of local optimization can overcome?

**AI becomes a CEO agenda item**. Process redesign at this level cannot be delegated to an AI Center of Excellence or buried in IT priorities. It requires executive mandate, cross-functional authority, and willingness to challenge assumptions that have shaped the organization for decades. The CEOs who recognize this will transform their companies. Those who treat AI as a technology initiative will optimize their way to irrelevance.

Consider a typical B2B sales process: lead qualification, needs assessment, proposal generation, negotiation, contract execution, handoff to delivery. Adding AI to each step (lead scoring, meeting summarization, proposal drafts) yields incremental gains. But the structure assumes human-to-human interaction at every stage, with handoffs, approvals, and waiting time built into the flow.

An AI-first redesign asks different questions: What if qualification, needs assessment, and initial proposal happened in a single continuous interaction? What if the agent had real-time access to pricing, inventory, and delivery constraints, eliminating the “let me check and get back to you” delays? What if the human entered only for high-stakes negotiation and relationship decisions?

This is not automation. It is process reconception.

The companies that augment existing processes with AI will see 10-20% efficiency gains. The companies that redesign processes around AI capabilities will see step-change improvements of 10x and more. This will be impossible to catch with incremental optimization.

Organizations that embed AI into broken processes are automating their inefficiencies at higher speed.

## 4. Adaptation Speed Is the New Competitive Moat

There is a persistent confusion in enterprise AI discussions between adoption and adaptation.

Adoption is acquiring and deploying AI capabilities: rolling out copilots, standing up a GenAI platform, training employees on prompt techniques. It is necessary but insufficient.

Adaptation is changing how the organization works in response to AI capabilities, and continuing to change as those capabilities evolve. It includes:

- Redesigning roles and responsibilities as agents take on tasks
- Updating governance frameworks as new risk patterns emerge
- Shifting investment from human capacity to human-agent orchestration
- Accelerating learning cycles to incorporate new model capabilities within weeks, not quarters

The half-life of AI best practices is now measured in months. An organization that took 18 months to operationalize GPT-4 patterns will find those patterns obsolete before they scale. The winners are not those with the best initial implementation, but those that can iterate fastest.

This requires a different organizational posture:

- **Governance that enables speed:** Lightweight approval processes for low-risk use cases, with clear escalation for high-risk ones. Not uniform heavyweight review for everything.
- **Experimentation infrastructure:** The ability to test new models, agents, and workflows in controlled environments and promote successes rapidly.
- **Continuous capability building:** Not one-time training programs, but ongoing upskilling that tracks the technology frontier.
- **Leadership fluency:** Executives who can make informed trade-offs on AI investments without relying solely on technical teams to frame options.

The traditional competitive moats (proprietary data, scale, brand, distribution) still matter. But they are increasingly necessary rather than sufficient. The new moat is the speed at which an organization can learn to work differently.

In 2026, the winners will not be those with the best AI. They will be those who learn fastest to work with imperfect AI productively.

## The Cost of Inaction

The risks of moving too fast with AI are well-documented: governance failures, reputational incidents, wasted investment, employee backlash. These risks are real and must be managed.

Equally consequential are the risks of moving too slowly. Companies that don't move ahead fall behind as the competition advances. Organizations that remain at the Experimenter or early Practitioner level face a compounding disadvantage:

- **Talent attrition:** High-performers increasingly expect to work with modern AI tools. Organizations perceived as AI-laggards will struggle to attract and retain the people who drive innovation.
- **Cost structure divergence:** As competitors automate knowledge work, their cost per transaction drops. The gap between AI-enabled and AI-limited cost structures will widen each year.
- **Customer expectation mismatch:** B2C companies have set a new bar for responsiveness and personalization. B2B customers increasingly expect the same, and AI-enabled competitors will deliver it.
- **Strategic optionality loss:** Many AI capabilities require foundational investments in data, platform, and skills. Organizations that defer these investments will find themselves unable to respond when competitive pressure intensifies.
- **Regulatory preparation gap:** The EU AI Act and similar regulations require documentation, risk assessment, and governance capabilities. Building these under time pressure is expensive and error-prone.

The cost of inaction is not static. It compounds.

# AI Maturity Redefined: Framework, Levels & Investment Fields

Knowing that AI matters is no longer the challenge. Knowing what to do (in what sequence, with what investment, measured against what benchmarks) is where most organizations struggle.

This chapter introduces the appliedAI Maturity Framework: a structured approach to assessing, planning, and advancing enterprise AI capabilities. It reflects our work with leading organizations across industries and our role in national AI maturity initiatives across Europe.

Think of AI maturity like organizational fitness. Just as physical fitness requires training across multiple dimensions (strength, endurance, flexibility, coordination) AI maturity requires building capability across strategy, technology, data, people, and execution. And just as a fitness program must be calibrated to current condition and goals, AI initiatives must be sequenced based on where an organization stands and where it needs to go.

The technology shifts described in Chapter 1 have changed what "fit" means. The training regimen that built capability in 2023 is no longer sufficient. Organizations must recalibrate their understanding of what each fitness level requires and adjust their training accordingly.

## The appliedAI Maturity Framework: A Brief Introduction

The appliedAI Maturity Framework provides a comprehensive view of the capabilities required to create sustained value from AI. It is built on two core structures: nine strategic dimensions of AI adoption and four maturity levels.

**Nine Dimensions** that cover the full scope of enterprise AI capability:

| Dimension | Scope |
|---|---|
| **AI Ambition & Steering** | Strategy, governance, compliance, value measurement |
| **Use Cases** | Discovery, prioritization, product thinking, process integration |
| **Organization** | Operating model, roles, accountability |
| **Expertise** | Skills, enablement, knowledge management |
| **Culture** | Adoption, responsibility, change readiness |
| **Data** | Strategy, quality, access, privacy |
| **Technology** | Platform, architecture, tooling, security by design |
| **AI Ecosystem** | Vendors, models, partners, IP, cost/FinOps |
| **Execution** | Delivery, lifecycle, LLMOps/MLOps, evaluation, operations |

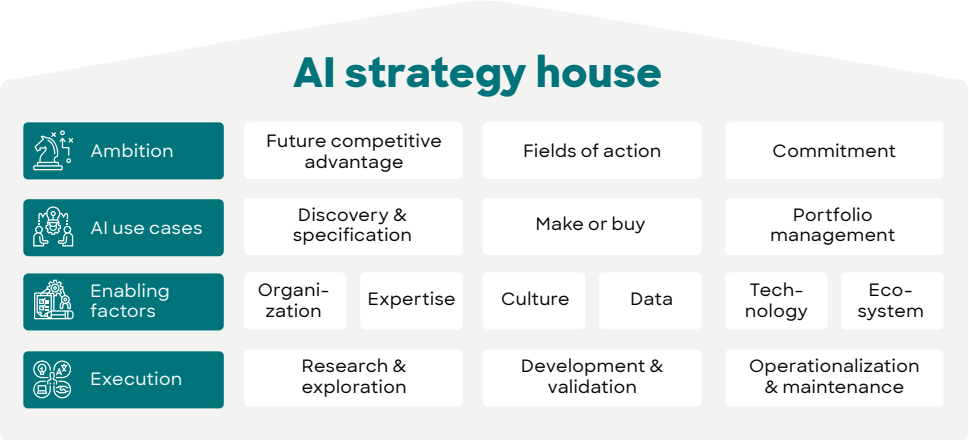



**Figure 1.** The elements of an AI strategy

These dimensions are not independent. Progress in one often requires, or enables, progress in others. Data quality constrains what Use Cases are feasible. Organizational structure determines how Execution scales. Ambition without Expertise produces failed pilots.

**Four Maturity Levels** that describe the progression from initial experimentation to industry leadership:

| Level | Name | Characteristic |
|---|---|---|
| L1 | Experimenter | Active experimentation; pilots underway; governance emergent |
| L2 | Practitioner | Operational use in selected areas; repeatable delivery emerging |
| L3 | Professional | Scaled enterprise capability; governance embedded; value measured |
| L4 | Shaper | AI as organizational DNA; continuous optimization; ecosystem influence |

The framework is not a maturity ladder to be climbed uniformly. Organizations may be at different levels across dimensions, and that is expected. What matters is understanding where you are, where you need to be, and what specific capabilities will close the gap.

This framework is used by leading national initiatives to assess AI maturity across entire industries in multiple European countries. It provides the foundation for benchmarking, strategic planning, and capability building at scale.

*For the complete framework, including detailed subdimensions and assessment methodology, see the appliedAI white-paper "Elements of a Comprehensive AI Strategy."*

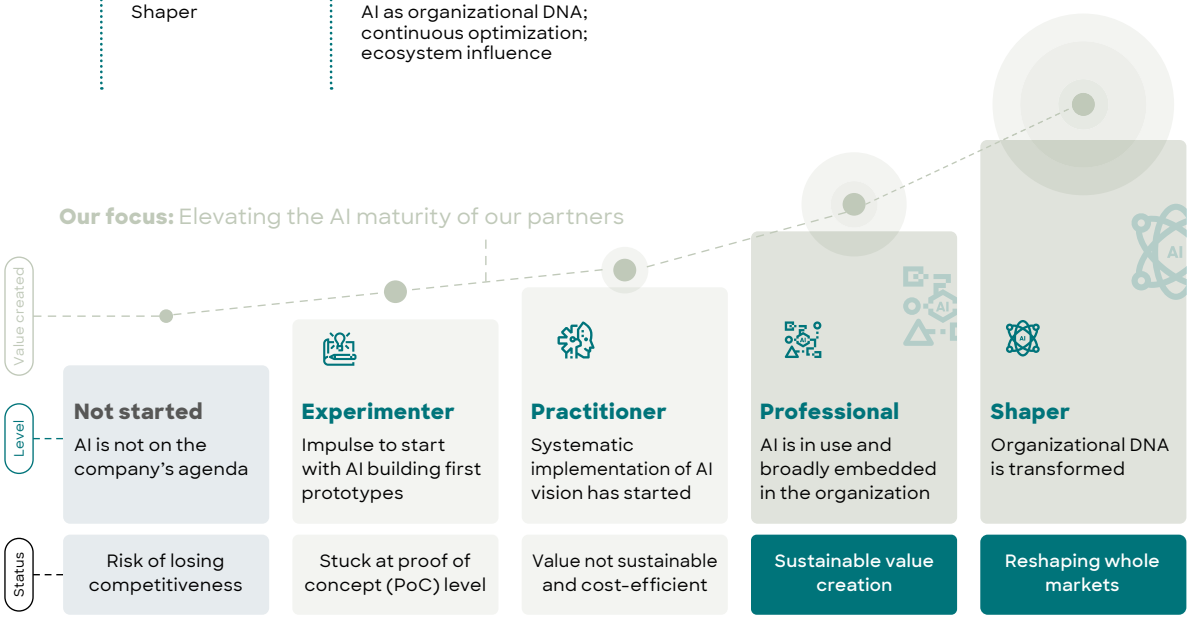
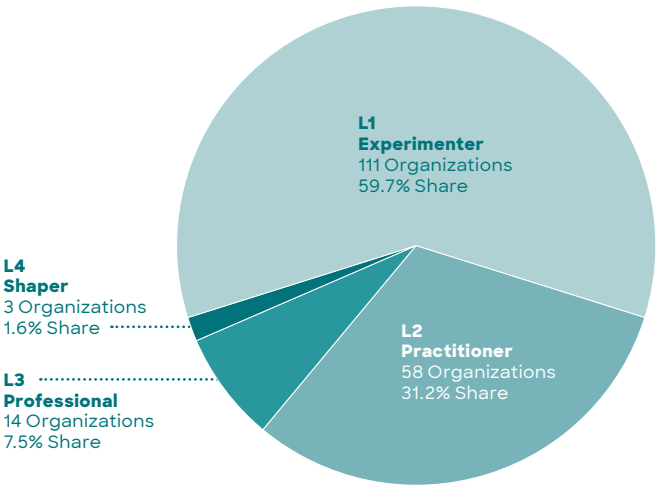



**Figure 2.** appliedAI Maturity Model

# Industry Benchmark: Where Companies Stand at the Time of Assessment

The following analysis is based on 187 maturity assessments conducted by appliedAI and partners across Europe in the past years. It provides a snapshot of where organizations actually stand, not where they believe they stand or where vendor marketing suggests they should be. The participants range from public institutions to large corporations. Smallest participants have several hundreds of employees. Largest participants are among the largest companies of Europe.

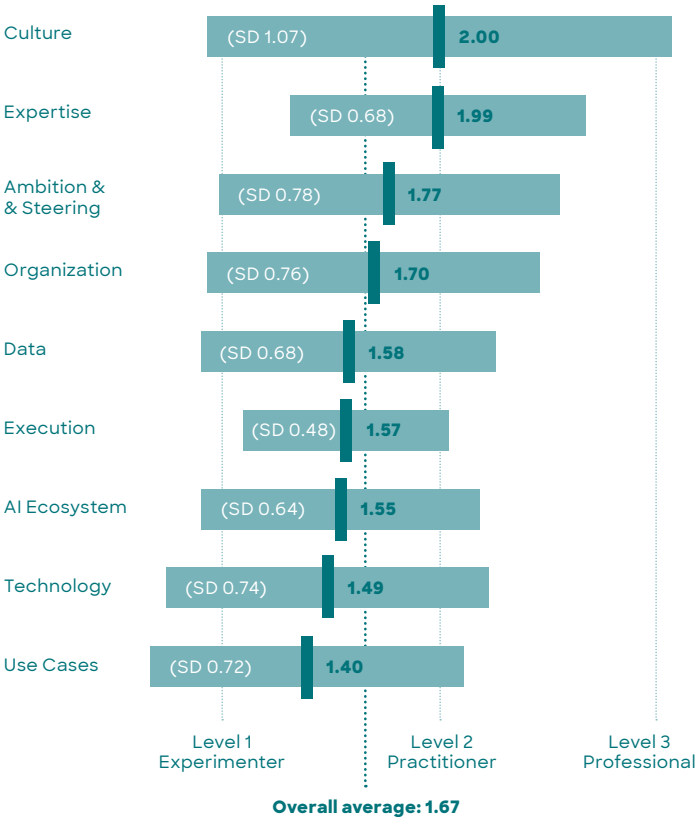
**The Distribution:**
**Most Organizations Are Experimenting**




The headline finding: **nearly 60% of organizations remain at Level 1**, active experimentation without reliable paths to production value. Only 9% have reached Professional or Shaper levels where AI is embedded as enterprise capability.

This is not a criticism nor do we consider this as statistically valid research. In particular, there might be selection biases on company and participant levels. It is a reflection of where companies are and where they struggle.

**The Dimension Gap:**
**Intent Outpaces Execution**




The pattern is consistent and telling:

**Strongest dimensions:**
**Culture, Expertise, Ambition & Steering (avg. 1.92)**
Organizations have intent. Leadership recognizes AI matters. Employees are curious and increasingly skilled. The "soft" foundations are being built.

**Weakest dimensions:**
**Use Cases, Technology, AI Ecosystem (avg. 1.48)**
Operationalization lags. Organizations struggle to move from ideas to production, from pilots to scaled value, from enthusiasm to governance. The "hard" capabilities that convert intent into results remain underdeveloped.

**The gap is 0.44 points**, nearly half a maturity level. This is the execution gap that separates organizations that talk about AI from those that generate value with it.

**What the Spread Reveals**

The standard deviation column tells its own story:

- **Culture shows the highest variance (1.07):** Some organizations have built genuine AI-positive cultures; others face resistance or indifference. There is no industry norm. Culture is a differentiator.
- **Execution shows the lowest variance (0.48):** Most organizations struggle here equally. This is not a capability some have cracked while others lag. It is a shared challenge, which suggests structural barriers rather than individual organizational failures.
- **Ambition & Steering varies significantly (0.78):** The spread between cautious experimenters and aggressive transformers is wide and strategic positioning on AI might further diverge, not converge.

**Key Interpretation**

Three patterns deserve executive attention:

1. **The intent-execution gap is structural.** Organizations are not failing because they lack ambition or awareness. They are failing because operationalization (Use Cases, Technology, Execution) requires capabilities that most have not yet built. The investment fields in Section 2.5 address exactly this gap.

2. **Level 1 is crowded, Level 3+ is not.** The competitive separation is happening now. Organizations that remain at Experimenter level for another year will find the gap to leaders increasingly difficult to close.

3. **Variance in Culture and Ambition & Steering signals strategic divergence.** Organizations are making different bets. Some are treating AI as transformational; others as incremental. The success of startups pursuing few-person unicorn approaches indicates that the more disruptive approaches might be the right path to follow.

The benchmark is not a leaderboard. It is a diagnostic. The question is not “where do we rank?” but “what specific capabilities must we build to move from where we are to where we need to be?”

# The Four Levels of AI Maturity in 2026

What does it mean to be at each maturity level in 2026, when GenAI is ubiquitous, agentic systems are production-ready, and the EU AI Act is in force? The following descriptions reflect the current technology landscape and regulatory environment.

The technology shifts of the past year have recalibrated what each fitness level means. An organization that was “Professional” by 2024 standards may find itself effectively at “Practitioner” level when measured against agentic AI capabilities, context engineering maturity, or end-to-end process transformation. The descriptions below reflect what each level requires now and what training is needed to advance.

# Level 1: Experimenter – The Beginning

This maturity level is characterized by active experimentation with AI, including GenAI tools used informally and first pilots initiated by motivated teams. There is a broad sense that AI matters, but the organization has not yet established a reliable path from ideas to repeatable production value. Controls for security, data protection, and compliance exist at best as initial guidance, but are not systematically embedded in delivery and operations.

In 2026, Level 1 is less about "awareness" and more about unmanaged adoption pressure: employees already have access to powerful GenAI capabilities, and the gap between speed and governance becomes visible early. Adoption is therefore real, but mostly uncoordinated. Some teams embrace AI enthusiastically while others remain skeptical or avoid it.

**What It Looks Like:**

- AI activity is fragmented; teams run pilots independently; "shadow AI" appears
- Use cases are chosen opportunistically; business value is discussed but rarely measured
- Basic principles exist (e.g., "don't paste confidential data into public tools"), yet adoption outpaces governance
- Early prototypes are convincing in demos, but operationalization (ownership, monitoring, change/adoption) is not in place
- Adoption is uneven: local power users emerge, while many employees lack clarity, confidence, or trust in AI outputs

**Technology Lens:**

- Traditional ML: POCs in notebooks; limited reproducibility; little or no monitoring or retraining plan
- GenAI: Ad-hoc prompting; pilots with public or semi-approved LLMs; early RAG attempts without strong source governance
- Agentic AI: Mostly absent; if present, narrowly scoped assistants without robust tool permissions or safety constraints

**Failure Mode:**

AI ambition becomes a collection of experiments with no strategic throughline.

**Priorities to Reach Level 2:**

- Define a minimum viable AI governance: roles, decision rights, intake, and a small set of enforceable rules
- Stand up an approved GenAI stack (identity, access control, logging, provider choices) to reduce shadow usage
- Establish a repeatable delivery path: from use case selection → build → test → release → operate
- Start measuring value for pilots (even simple KPIs) and retire low-value experiments fast
- Launch a baseline adoption and change package: clear communication, lightweight training, and a champion network

# Level 2: Practitioner – The Operationalization

Practitioners have moved from experimentation to operational use in selected areas. An AI/GenAI agenda is forming, and the organization is building repeatable ways to identify, prioritize, and deliver AI use cases. Basic governance and compliance processes exist and are increasingly applied, but scaling remains constrained by platform maturity, data readiness, and uneven adoption.

In 2026, Level 2 organizations typically deploy classic AI and GenAI in "safe" productivity scenarios and begin deploying constrained agents in low-risk workflows, while establishing the foundations for EU AI Act-ready governance. Adoption accelerates in pockets, but behavior change is not yet consistently anchored in incentives, workflows, and role expectations.

**What It Looks Like:**

- Use cases are prioritized using basic criteria (value, feasibility, risk, data availability)
- Initial centralized enablement exists (e.g., small AI team/CoE, playbooks, training)
- Security/legal/risk reviews exist, but depth varies; documentation is inconsistent
- Production AI exists in pockets; scaling across domains is still difficult
- Adoption is managed selectively: some teams integrate AI into daily routines, while others see it as an optional add-on; management support varies by unit

**Technology Lens:**

- Traditional ML: First MLOps pipelines emerge; monitoring begins; retraining is mostly manual
- GenAI: Approved tools/providers; early RAG with curated sources; baseline policies and human review for sensitive outputs
- Agentic AI: Limited deployment; agents operate under tight constraints (allowlisted tools/actions) and are monitored at a basic level

**Failure Mode:**

"Aspirational strategy" that does not translate into coherent strategic choices. Too broad, too many priorities.

**Priorities to Reach Level 3:**

- Establish standard release gates for AI (quality, security, privacy, compliance, human oversight)
- Build scalable platform patterns (reusable RAG/guard-rails/monitoring components, golden paths)
- Introduce systematic GenAI evaluation (offline/online metrics, regression testing, red-teaming where relevant)
- Make value real: connect AI products to outcome KPIs and benefits realization tracking
- Strengthen EU AI Act readiness: consistent risk classification, documentation, and evidence across the lifecycle
- Expand adoption manage-ment: role-based enablement, process integration, and adoption KPIs

# Level 3: Professional – The Embedding

Professionals operate AI as a scaled enterprise capability. Multiple AI solutions deliver measurable value across functions, and the organization has established robust end-to-end processes for building, evaluating, releasing, monitoring, and improving AI systems. Governance, security, privacy, and compliance are embedded "by design," enabling faster scaling rather than slowing it down.

In 2026, Level 3 is defined by operational reliability and evaluation discipline, especially for GenAI and agentic systems: systematic testing, monitoring, and controlled change management become non-negotiable. Adoption is no longer accidental: the enterprise manages AI as a workforce and operating model change, with measurable uptake and sustained usage.

**What It Looks Like:**

- AI products have clear ownership, roadmaps, and adoption/impact metrics
- A mature operating model (e.g., hub-and-spoke) works in practice across domains
- Platform and tooling enable reuse and consistency (observability, access control, logging, incident handling)
- Risk and compliance processes align with lifecycle thinking, with credible evidence trails
- Adoption is sustained: employees are enabled and expected to use AI appropriately; managers actively drive integration into processes and measure benefits and compliance in parallel

**Technology Lens:**

- Traditional ML: Mature MLOps; reproducibility, monitoring, drift management, and retraining are institutionalized
- GenAI: LLMOps is real with versioning, evaluation, regression testing, prompt/RAG governance, robust monitoring and feedback loops
- Agentic AI: Agents are deployed in meaningful workflows with strong permissions, sandboxing, escalation paths, and human oversight where required

**Failure Mode:**

Adoption plateau. Solutions are delivered, but incentives and workflow integration don't drive sustained usage.

**Priorities to Reach Level 4:**

- Move from "AI-enabled" to AI-first process redesign (end-to-end, measurable transformation)
- Automate assurance: policy-as-code, continuous evaluation, automated evidence capture
- Scale agentic capabilities safely: standardized tool APIs, permissioning, and enterprise-grade guardrails
- Optimize unit economics: systematic FinOps for GenAI/agents (routing, caching, right-sizing)
- Invest in differentiated capability: proprietary data/knowledge assets, reusable components, and internal methods
- Elevate change to transformation: redesign roles and decision rights, embed AI into performance routines

# Level 4: Shaper – The Thought Leadership

Shapers are leaders that make AI an essential part of their organizational DNA and competitive advantage. They have transformed multiple core value streams into AI-first operating models, including safe, enterprise-grade agentic automation. They continuously optimize value, cost, and risk using automated controls and measurement systems, and they influence the broader ecosystem through standards, reference architectures, partnerships, and sometimes open-source contributions.

In 2026, Level 4 is characterized by continuous optimization and ecosystem influence, not just breadth of use cases. Adoption and culture change are fully internalized: the organization operates as if AI capability were a core managerial and operational discipline, similar to finance or safety in industrial environments.

**What It Looks Like:**

- AI is embedded in strategy, operating model, and process design; "AI-first" is visible in how work gets done
- Continuous measurement loops exist across quality, safety, cost, and business outcomes
- Governance is highly effective and largely automated, enabling speed with confidence
- The organization attracts top talent and partners; it sets direction rather than following it
- Cultural adoption is durable: employees expect AI in workflows, managers run continuous improvement cycles, and responsible AI behavior is reinforced through routines, incentives, and clear accountability

**Technology Lens:**

- Traditional ML: Pervasive, industrialized lifecycle; advanced methods used selectively where they outperform; governance is deeply integrated
- GenAI: Multi-model orchestration, continuous evaluation at scale, provenance-driven knowledge, robust guardrails as reusable components
- Agentic AI: Enterprise agent platforms with standardized tools, strong permissions, action logging, escalation, and safe autonomy patterns across core processes

**Failure Mode:**

"AI-first pressure" turns into burnout or loss of trust. Pace increases without sufficient support, clarity, and safeguards for teams.

**Sustaining Leadership Beyond Level 4:**

- Keep leadership defensible through continuous improvement, ecosystem shaping, and disciplined governance as a competitive advantage
- Treat evaluation, assurance, and compliance as a product: always-on, automated, and continuously updated
- Continuously renew the culture: prevent change fatigue, maintain trust, and evolve roles responsibly as autonomy increases

## Dimension Deep-Dive: Two Examples

The maturity levels come to life through specific dimensions. Below we illustrate the progression for two dimensions that are particularly relevant for 2026: Ambition & Steering (typically the starting point for C-level discussions) and Execution (where AgentOps and LLMOps become critical)

| Dimension | L1: Experimenter | L2: Practitioner |
|---|---|---|
| **AI Ambition & Steering** (strategy, governance, compliance, value) | AI ambition is emerging and largely framed as "let's start using AI/GenAI" to explore benefits and avoid falling behind. The enterprise adopts what is readily available and learns what AI could mean for its context, without a sharp point of view on competitive advantage. Strategic decisions are driven by curiosity and early success stories; prioritization and risk stance are implicit rather than codified.<br>*Failure mode: AI ambition becomes a collection of experiments with no strategic throughline.* | AI ambition becomes explicit and directional: the enterprise defines where it wants to play and sets priorities linked to business objectives. Strategy increasingly reflects how the enterprise will leverage capabilities in targeted domains, still primarily adopting mainstream patterns, but with clearer choices and trade-offs. A baseline risk and compliance stance (including EU AI Act awareness) is translated into policies and decision forums.<br>*Failure mode: "Aspirational strategy" that does not translate into coherent choices.* |
| **Execution** (delivery, lifecycle, LLMOps/ MLOps, evaluation, operations) | Delivery is exploratory and project-like: prototypes appear quickly, but productionization is rare and relies on individual effort. Testing and evaluation are informal, monitoring is minimal, so reliability and repeatability remain low. GenAI is judged by qualitative demos, while agentic experiments are avoided or treated as novelty.<br>*Failure mode: "Pilot churn"; teams repeatedly rebuild prototypes with no pathway to stable operation.* | The enterprise can bring selected AI solutions into production and keep them running, but practices differ across teams and scale is constrained. Basic delivery steps exist from scoping to deployment, and initial operational routines appear. For GenAI, release checks are often manual and inconsistent; for agents, autonomy is limited and oversight is cautious.<br>*Failure mode: Reliability gap; solutions run, but quality issues and incidents erode trust.* |

## L3: Professional

AI ambition is differentiated and measurable: the enterprise articulates how AI will create competitive advantage, not just efficiency gains, and ties this to concrete strategic themes and investment horizons. The company moves from adopting generic patterns to building distinctive capabilities (e.g., proprietary knowledge assets, domain-specific copilots/agents as strategic bets). Risk appetite and regulatory readiness are defined at strategy level.

*Failure mode: Strategy is strong but not adaptive; ambition lags behind model/ market shifts.*

Execution is standardized and measurable: evaluation, release gates, versioning, regression testing, monitoring, and continuous improvement are consistently applied. GenAI is operated with LLMOps discipline (offline/ online evaluation, telemetry, prompt/RAG governance) while agentic systems run with clear oversight and action-level observability. Delivery scales across business units without sacrificing reliability.

*Failure mode: Metric overload; many signals exist, but teams lack clarity on which metrics trigger action.*

## L4: Shaper

AI ambition is transformative and externally oriented: the enterprise aims to shape its industry by redefining customer experiences, operating models, or value chains around AI-first principles. Strategy is dynamic and learning-driven: the enterprise continuously repositions as model capabilities change, and it actively influences ecosystem direction. The risk posture is strategic, with trust, safety, and transparency treated as competitive differentiators.

*Failure mode: Overreach; ambition to "shape the industry" outpaces ability to sustain focus.*

Execution becomes continuous and largely automated: evaluation and monitoring operate as always-on systems, and improvement cycles run frequently with high confidence. The enterprise safely operates agentic workflows at scale using robust constraints, rapid rollback, and continuous learning loops. Delivery excellence enables sustained AI-first transformation rather than periodic programs.

*Failure mode: Automation without governance; speed increases, but oversight fails to keep pace for higher-risk workflows.*

*The complete Dimension-Level Matrix covering all nine dimensions is available as a downloadable poster.*

**appliedAI Maturity Assessment**

Where does your organization stand? Think of this as your AI fitness check: a structured evaluation across all nine dimensions, benchmarked against industry peers, with clear recommendations for your training plan. Typical duration: 4 weeks from kick-off to results presentation.

# Three Critical Investment Fields for 2026

**A Note on Enterprise Context**

The productivity gains celebrated on LinkedIn ("I built this in 30 minutes with Claude") typically emerge from individual power users operating without governance constraints, integration requirements, or organizational complexity. Enterprise context is different. Deploying AI across thousands of users, integrating with legacy systems, managing regulatory obligations, and driving sustained adoption requires capabilities that go far beyond individual productivity hacks.

The following investment fields address exactly this gap.

**Why These Three Fields, Why Now**

The technological developments of the past twelve months have fundamentally shifted what is possible and what is required. Agents have moved from research curiosity to production reality. Context windows have expanded from thousands to millions of tokens. Multi-agent orchestration frameworks have matured. These advances create new opportunities, but they also create new capability requirements that most organizations have not yet built.

The three investment fields below emerge directly from this shift. They are not abstract maturity themes; they are the specific capabilities that, based on our assessment data and ongoing work with leading organizations, separate those who scale from those who stall. Companies we work with report the same pattern: progress from Experimenter to Practitioner is achievable with foundational governance and platform work. But moving from Practitioner to Professional, or sustaining Professional-level performance, increasingly depends on strength in these areas.

Organizations that build capability here will scale faster and more safely. Those that neglect them will find themselves stuck at Levels 1-2, watching competitors pull ahead.

### Investment Field 1: AgentOps & Observability

*What it is:*
The discipline of monitoring, debugging, governing, and continuously improving AI agents in production. AgentOps extends MLOps and LLMOps to address the unique challenges of autonomous systems: action traces, permission enforcement, cost tracking, escalation handling, and behavioral consistency over time.

*Why it is critical:*
Agents fail differently than traditional AI. They do not just produce wrong outputs; they take wrong actions. Without observability, organizations cannot detect drift, debug failures, or demonstrate compliance. Agents without AgentOps are black boxes with agency, a risk profile that no enterprise should accept.

**Concrete first steps:**

- Implement action-level logging for all agent interactions (what tools were called, what parameters were used, what results occurred)
- Define and enforce permission frameworks: which agents can take which actions under which conditions
- Establish escalation paths: when must an agent pause and involve a human?
- Build dashboards that track agent performance, cost, and behavioral patterns
- Create incident response playbooks specific to agent failures

*Primary dimensions affected:*
Execution, Technology, Data

### Investment Field 2: Leadership Enablement for Human-Agent Teams

*What it is:*
Building the capability for managers and leaders at all levels to effectively orchestrate teams that include both humans and AI agents. This goes beyond AI literacy to include practical skills in task delegation to agents, performance evaluation of human-agent collaboration, and organizational design for hybrid teams.

*Why it is critical:*
Agent adoption stalls without management support. Most managers have no framework for thinking about when to delegate to an agent versus a human, how to evaluate agent work, or how to redesign workflows

around agent capabilities. Without this capability, adoption remains fragmented and the promised productivity gains fail to materialize.

**Concrete first steps:**

- Develop decision frameworks for human versus agent task allocation (not everything should go to an agent)
- Train managers on agent capabilities and limitations specific to their domains
- Create feedback mechanisms where managers can evaluate and improve agent performance
- Redesign team goals and incentives to account for agent contributions
- Build communities of practice where managers share what works

*Primary dimensions affected:*
Expertise, Culture, Organization

### Investment Field 3: End-to-End Process Rethinking

*What it is:*
The systematic redesign of core business processes around AI capabilities. This is not augmentation of existing workflows but fundamental reconception of how work gets done. It requires cross-functional authority, clear outcome ownership, and willingness to challenge process assumptions that predate AI.

*Why it is critical:*
Local AI optimizations (better document processing, faster summarization) deliver incremental gains but leave process architecture unchanged. The step-change value comes from rethinking entire value streams: eliminating handoffs, collapsing decision loops, and enabling continuous rather than batch processing. Organizations that only augment will be outcompeted by those that redesign.

**Concrete first steps:**

- Identify 1-2 core processes where AI-first redesign could deliver transformative (not incremental) value
- Establish cross-functional task forces with mandate to challenge existing process assumptions
- Define outcome KPIs (not activity KPIs) that measure end-to-end value delivery
- Map current process architecture explicitly before redesigning; most organizations do not actually know how their processes work
- Start with constrained pilots, but design for scale from the beginning

*Primary dimensions affected:*
Use Cases, Organization, Ambition & Steering

## appliedAI Partnership

These investment fields are too important to navigate alone, and too complex to solve with generic consulting.

The appliedAI Partnership brings together leading organizations typically at Practitioner or Professional level to work on frontier challenges collaboratively. What this means in practice:

- **Tailored support from our AI experts:** Start with a maturity assessment, develop an individual roadmap, and benefit from continuous sparring. Access training, workshops, and expert sessions that address your specific context, not generic frameworks.
- **Peer exchange with established organizations:** Roundtables, working groups, and exclusive events where you learn from other pioneers and share experiences. The problems you're solving are not unique, but the people who've solved them are not easy to find.
- **Access to the appliedAI ecosystem and knowledge base:** Bundled expertise from conferences, workshops, and reports that help you make strategically sound decisions. This is curated insight, not information overload.

The Partnership is how we develop the knowledge that doesn't exist yet, together with the organizations facing frontier challenges. When established best practices don't exist, Partnership members help create them.

# The Horizon: What's Coming Next

The hypotheses and investment fields in the previous chapters address what organizations should act on now. This chapter looks further ahead, at developments that are emerging but not yet production-ready, and that may reshape enterprise AI in 2027 and beyond.

These are not predictions. They are hypotheses we are actively tracking. In the 2027 update of this paper, we will assess which materialized, which evolved differently than expected, and which proved premature.

## Agent Swarms: When Agents Create Their Own Subagents

*The concept:*
Current agent deployments typically involve a single agent (or a small, predefined set of agents) working on a task. Agent swarms go further: a primary agent dynamically creates and coordinates subordinate agents to handle subtasks, allocating work based on task requirements and available resources. The swarm can grow, shrink, and reconfigure itself as the problem demands.

*Current state:*
Swarm patterns are emerging in research and in platforms where agents spawn subagents to handle parallelizable workloads. The results are impressive in controlled environments but raise significant governance questions: How do you audit a system where the agents themselves were created by other agents? How do you enforce constraints that weren't anticipated when the primary agent was deployed?

*Enterprise implication:*
Swarm architectures promise massive scalability for complex tasks, but they also make governance exponentially harder. Organizations considering swarm patterns need enterprise-grade constraint propagation (subagents inherit the boundaries of their parents), comprehensive action logging across the entire hierarchy, and robust resource controls to prevent runaway costs.

*Watch signal:*
Major cloud providers begin offering native swarm primitives with built-in governance controls.

## Agent-to-Agent Communication: The Rise of AI Protocols

*The concept:*
Today, agents primarily interact with human-designed APIs and interfaces. Agent-to-agent communication envisions agents from different systems (potentially different organizations) interacting directly using standardized protocols. This would enable automated negotiation, coordination, and transaction across organizational boundaries.

*Current state:*
Multiple initiatives are emerging. New agent protocols aim to standardize how agents communicate intent and negotiate actions. Payment networks are exploring how agent-to-agent transactions could be authorized and settled. Early platforms demonstrate agent-only forums where AI systems exchange information without human intermediation.

*Enterprise implication:*
B2B processes could be fundamentally transformed if agents can negotiate, transact, and coordinate across company boundaries. But this requires solving authentication (how do you verify an agent's authority to act for an organization?), liability (who is responsible when agent-to-agent negotiations go wrong?), and interoperability (whose protocol wins?). Early movers will help shape these standards; late movers will comply with standards others wrote.

*Watch signal:*
First industry consortia form to define agent-to-agent protocols for specific domains (e.g., supply chain, financial services).

## Citizen Agent Development: Democratization or Risk?

*The concept*:
Just as low-code platforms democratized application development, emerging tools enable business users to create and deploy their own AI agents without deep technical expertise. Users can define agent goals, connect data sources, and specify workflows through natural language or visual interfaces.

*The opportunity:*
Citizen agent development could unlock massive value by enabling domain experts (who understand the problems best) to build solutions directly, without waiting in queue for central AI teams. It scales AI delivery beyond what any CoE could staff.

*The risk:*
Without guardrails, citizen agents become shadow AI at scale. Users may create agents with excessive permissions, connect sensitive data sources inappropriately, or deploy agents that take actions with unintended consequences. The governance challenges of shadow IT pale in comparison to the governance challenges of shadow agents.

*Enterprise implication:*
The question is not whether to enable citizen agent development, but how. Organizations need Guardrails-as-a-Platform: infrastructure that makes it easy for business users to build agents within safe boundaries, with automatic enforcement of permissions, data access controls, and action constraints. This is a platform capability, not a policy document.

*Watch signal:*
Low-code agent builders appear in mainstream enterprise software suites (Microsoft, Salesforce, ServiceNow) with built-in governance hooks.

# Perspectives:
# Top Leadership Questions to Ask

**Greenfield vs. Brownfield:**
**Where Should We Build New?**

Most enterprises operate in brownfield environments: existing systems, accumulated data, established processes. The temptation is to integrate AI into what exists. But sometimes the right answer is greenfield: building a new system or process from scratch, designed around AI capabilities from the beginning.

Three questions for your leadership team:

- Where is our existing process architecture so constrained by legacy that AI augmentation would require more effort than rebuilding?
- Which new products, markets, or customer segments could we only address with AI-first design, capabilities impossible to retrofit?
- Where can we run greenfield and brownfield in parallel during transition, and where would that create unacceptable complexity?

The answer is often "both": greenfield for new ventures, brownfield optimization for core operations that cannot be disrupted. The strategic question is which is which.

**The SME Roadmap:**
**How Should Mid-Sized Organizations Prioritize?**

Enterprises with dedicated AI teams and platform budgets face different challenges than mid-sized organizations with limited resources. For SMEs, three questions clarify the path:

*Which platforms can we leverage immediately?* The GenAI platform battle is being won by hyperscalers. SMEs should consume, not compete. But beyond the big names, several regional GenAI platform providers offer rapid integration at modest cost with immediate value creation. Automation and agent low-code platforms enable fast starts without enterprise budgets.

*What can we do now that creates value this quarter?* Enterprise AI has existed for nearly a decade. Many applications work reliably and require less than a quarter to integrate. The menu of proven solutions is extensive. Pick applications with demonstrated ROI, get started, and let adaptation follow from experience.

*Are we building for adaptation, not just adoption?* The constraint for SMEs is rarely technology; it's organizational capacity to change. Invest disproportionately in helping people work differently, not just in deploying tools. Adoption without adaptation is technology that sits unused.

**Rethinking Core Functions:**
**What Should Be on the CxO Agenda?**

*R&D: The CTO Question*
How will your R&D function operate in 2027 and beyond? Human-agent teams are rewriting the rules of software development, but the transformation extends further. Hardware specification, electronics design, testing, and prototyping will all work differently when agents participate in the engineering process.

The classic V-model for product development (sequential phases from requirements through validation) may be fundamentally challenged. Leaders should explore: synchronization loops that enable parallel workstreams, self-improving agentic workflows that learn from each development cycle, and new end-to-end development patterns that collapse traditional phase boundaries. The organizations that figure this out first will set the pace for their industries.

*Sales: The CRO Question*
Lead scoring and proposal drafting were the first wave. The next wave sees agents handling qualification conversations, objection handling, and routine negotiation, with humans entering for high-stakes relationship decisions and complex deal structuring. The quota-carrying rep becomes an agent orchestrator.

But the deeper question is architectural: How should sales workflows be designed so that through feedback loops and compounding effects, the system itself gets better at selling over time? Each customer interaction, each won and lost deal, each objection pattern becomes training data that improves future performance. The sales organization that builds this learning architecture will compound its advantage quarter after quarter.

*Production: The COO Question*
Predicting failures was step one. Step two is agents that not only predict but respond: adjusting parameters, scheduling maintenance, reordering supplies, and coordinating across systems, with human oversight for exceptions and strategic decisions. The plant manager oversees a hybrid workforce.

# Moving Forward with appliedAI

The shift to enterprise AI maturity is not a journey organizations should take alone. The challenges are too interconnected, the technology moves too fast, and the stakes are too high for isolated experimentation.

appliedAI is your organization's AI fitness trainer. We've worked with hundreds of organizations across Europe, we understand what it takes to build capability at each maturity level, and we know which exercises produce results versus which create the illusion of progress. Our frameworks (including the maturity model presented in this paper) serve as the foundation for national AI initiatives in multiple countries.

What shapes our approach: we are convinced that AI agents will fundamentally transform value creation, and that 10x improvements are achievable, not aspirational. We design processes "Agent-First" (optimized for agent capabilities from the start) and deploy "Agent-Ready" workflows (proven blueprints for agent-suitable work). And we practice what we advise: appliedAI operates with agentic AI across our own value chain.

## Engagement Options

**appliedAI Partnership**
For organizations that set standards, not just follow them. The Partnership is where we develop knowledge that doesn't exist yet, together with organizations at the frontier. Members get tailored expert support (maturity assessments, roadmaps, continuous sparring), peer exchange with leading organizations (roundtables, working groups, exclusive events), and access to the appliedAI ecosystem and curated knowledge base.

**appliedAI Professional Services**
Strategy, engineering, and transformation, because AI maturity requires all three. We meet you where you are and work across all dimensions to generate value and build lasting capability.

- **Strategy & Transformation:** From ambition to 10x reality. Roadmap, operating model, and governance designed to convert AI capabilities into competitive advantage.
- **Engineering:** GenAI and agents integrated into core processes with MLOps discipline and security-by-design. Measurable impact, not demo magic.
- **AI Agents Lighthouse:** Orchestrated multi-agent systems at the frontier of what's possible that redefine processes end-to-end, from concept to productive human-agent collaboration.

**appliedAI SME Program**
Measurable progress in 90 days. Strategy, validated use case, and team enablement delivered as one integrated program. Immediate wins first, strategic expansion second. Includes ongoing updates, monitoring, and benchmarking: a durable capability-building relationship, not a one-time project.

## Why appliedAI

**Thought Leadership:** Internationally recognized as a leading AI initiative. We help define the frontier, not follow it.

**Holistic Perspective:** AI is more than a model. We work across strategy, technology, data, organization, culture, and execution, not just the dimensions that fit a narrow offering.

**Technological Excellence:** State-of-the-art solutions for hard AI challenges. Built in-house, proven in production.

**Focus on Impact:** We don't stop at prototypes. The goal is value realized, not potential demonstrated.

**Enablement:** We build capability, not dependency. When we leave, you're stronger.

**Network:** Access to an ecosystem spanning industry leaders, research institutions, and policy makers. The goal is to strengthen Europe.

## Continue the Conversation

**Ready to move from AI ideas to real impact?**

If you'd like to discuss your company's AI maturity and the next practical steps, get in touch via our contact form: https://share-eu1.hsforms.com/1rTtHLKMNRDqZu6cp0xYEugfmlxv.

You can also find more information and a maturity assessment check on our website: https://www.appliedai.de/en/solutions-services/strategy-transformation/.

**Appendix: Full Dimension-Level Matrix**

The complete matrix showing all nine dimensions across all four maturity levels (including detailed descriptions, technology lens perspectives, failure modes, and transition priorities) is available as a downloadable resource.

## About appliedAI

appliedAI is Europe's largest initiative for the application of trusted AI technology. The initiative was established in 2017 by Dr. Andreas Liebl as a division of UnternehmerTUM Munich and transferred to a joint venture with Innovation Park Artificial Intelligence (IPAI) Heilbronn in 2022.

At the Munich and Heilbronn offices, more than 100 employees pursue the goal of making European businesses a shaper in the AI era in order to maintain Europe's competitiveness and actively shape the future.

appliedAI holistically supports international corporations, including BMW and Siemens, as well as medium-sized companies in their AI transformation. This is accomplished through partnership-based exchange and joint knowledge building, comprehensive accelerator programs, and specific solutions and services, such as strategy consulting and Use-Case development.

For more information, please visit
**www.appliedai.de/en/**

initiative for applied artificial intelligence

***AI Maturity 2026:*** *The Year Agents Redefine Value Creation*

Published by

appliedAI Initiative GmbH
August-Everding-Straße 25
81671 München
Germany
www.appliedai.de