

AI Maturity of French SMEs 2025-2026: MATIA Method™ · June 2026 Edition

The reference white paper for business leaders, June 2026 Edition

Paul-Antoine TUAL

June 2026

AI Maturity of French SMEs 2025-2026

The reference white paper for business leaders, June 2026 Edition.

About the MATIA Method™

The **MATIA Method™** is the French framework for assessing and supporting the AI maturity of SMEs, developed by Paul-Antoine TUAL (AI Transformation Leader). Designed on the basis of more than thirty documented engagements with French SMEs and mid-caps since 2024, it offers a five-level framework (**Spectateur, Artisan, Orchestre, Architecte, Pionnier**) and a five-phase approach over twelve months. It stands apart from global frameworks (McKinsey, BCG, Gartner, Microsoft, PwC) through its specialisation in French SMEs and the golden rule that underpins it: *no level can be skipped*. This white paper constitutes its canonical reference. The articles published on the MATIA portfolio are in-depth elaborations of it.

Preamble: the narrow window

In the second quarter of 2026, **58% of French micro-businesses and SMEs report using AI** (almost half of them daily), compared with 55% at the end of 2025 [1]. This figure, drawn from the Bpifrance Le Lab / Rexecode barometers, marks a historic turning point. In the space of three years, artificial intelligence has gone from an R&D topic for large corporates to a daily tool of the real economy.

However, behind this massive adoption, another statistic attracts far less attention: **only around 11% of organisations are “AI leaders” that derive clear business value from AI**. This is the share measured by the *KPMG Global AI Pulse* for the first quarter of 2026 (2,110 business leaders, 20 countries), where **82% of these leaders report significant value, compared with 62% of the others** [2]. The *Stanford AI Index 2026* confirms this from the other end

of the spectrum: **88% of organisations use AI in at least one function, but fewer than 10% have genuinely scaled it** in even a single one of them [29]. Out of 100 equipped SMEs, the overwhelming majority merely dabble; one in ten reaps the rewards; a handful genuinely transform. The window is narrow, but it is still open.

This white paper is addressed to business leaders who are not content to adopt AI merely to say they have done so. It offers a five-level framework, a method, and field data, drawn from more than thirty documented engagements with French SMEs and mid-caps since 2024 [8].

The thesis is simple: **AI in SMEs is not won through technology. It is won through method.** The 11% that succeed are not better equipped. They are better organised. The <0.5% that define their sector are not more innovative. They are better governed.

The 2026 edition incorporates four new elements compared with the initial edition of April 2026:

1. **The extension of the MATIA Method™ Scale to 5 levels:** Spectateur and Pionnier added at the top and bottom of the scale, in line with the 2025-2026 “golden standard” frameworks (Gartner 5 levels, PwC AI-Native, Microsoft Agentic L100 → L500).
2. **The updated AI Act timetable:** delay following the Digital Omnibus of 7 May 2026.
3. **Two new chapters:** sovereignty & infrastructure (Chapter 4), mastering the cost of AI & FinOps (Chapter 5).
4. **The integration of context engineering:** “directing AI” rather than “writing prompts”, within the 5-phase methodology (Chapter 6).

Chapter 1: Where French SMEs stand, 2025-2026

1.1 Surface adoption, depth of value

The Bpifrance Le Lab barometer from late 2025 indicates **55% reported use of generative AI** among micro-businesses and SMEs, up by around 30 points in eighteen months [1]. Three uses dominate: drafting marketing content, summarising documents, and help with drafting emails [1]. Three uses that share one characteristic: they are individual, outside any process, and unmeasured.

The France Num 2025 barometer (DGE, September 2025, 11,021 businesses: the most recent published data, with the 7th edition expected in autumn 2026) refines the picture: **26% of micro-businesses and SMEs report using at least one AI tool** in 2025 (double the 2024 figure), but the proportion rises with size: 23% for businesses with 1-4 employees, 35% for those with 20-49, **rising to 42% for those with 50-249 employees** [12]. *Regular* use, however, remains a minority.

Digging deeper into the results yields another figure, a far less flattering one: **only around 11% of organisations are “AI leaders”** drawing clear business value from AI. This is the finding of the *KPMG Global AI Pulse* for the first quarter of 2026 (2,110 business leaders, 20 countries); **82% of these leaders report significant value, compared with 62% of the rest** [2]. That is, a visible effect on the profit and loss account, on revenue, on costs, or on productivity. This international finding is borne out on the ground: according to the **Stanford AI Index 2026**, 88% of organisations use AI in at least one function, but **fewer than 10% have genuinely scaled it** in even a single

one [29]; and the OECD notes that **only 29% of SMEs using generative AI deploy it in their core business** [30]. The remainder, the yawning gap between reported use and measured value, is the zone of scattergun adoption.

Deloitte goes further in its *State of AI in the Enterprise 2026* (January 2026, 3,235 business leaders, 24 countries): **only 30% of organisations are redesigning their key processes around AI**, while **37% use it superficially**, without changing anything fundamental. And barely a quarter report a genuinely transformative effect [20]. This is the “industrialisation gap”: the dividing line between the minority that genuinely transforms its operations and the majority that remains stuck in permanent pilot mode.

1.2 The executive committee’s blind spots

The **Bpifrance Le Lab “AI in French SMEs and mid-caps”** study (fieldwork late 2024, 1,209 business leaders) brings to light two blind spots in the executive committee [3]:

- **Nearly six in ten business leaders see AI as a matter of survival, yet an equivalent proportion has no formalised AI strategy.** AI is perceived as a matter of tools, not as a subject for senior management.
- **43% of businesses carry out no analysis of their data at all.** No cross-functional dashboard, no data team, no governance, and therefore no raw material for a credible AI transformation. (At national level, INSEE confirms the scale of the task: only 10% of businesses with 10 or more employees were using an AI technology in 2024 [3].)

To this is added the phenomenon of **executive Shadow AI**: the Microsoft Work Trend Index 2026 documents a persistent gap between business leaders and employees in AI usage (67% of business leaders familiar with agents, against 40% of employees), with business leaders often adopting AI in a personal capacity before the business has formalised its use [13]. This phenomenon is not in itself a problem. It becomes a risk when it takes hold over time without a framework: exfiltration of sensitive data, dependence on non-auditable tools, opaque decisions.

Underlying use, however, remains cautious in France. According to ADP (*People at Work 2026*, close to 40,000 workers across 36 countries), **only 11% of French employees use AI on an almost daily basis, and 29% never do**, one of the lowest rates among the countries studied, against a global average of 20% for near-daily use [11]. Deep diffusion, not surface adoption, remains the real task at hand.

The most telling figure now concerns execution: **43% of major AI initiatives are judged doomed to fail**, owing more to a lack of organisational readiness than to technology (*HCLTech, May 2026*, a survey of 467 business leaders at companies with revenue above \$1 billion) [5]. Gartner confirms the trend in its *Hype Cycle for Agentic AI* from April 2026: at least half of generative AI projects are abandoned after the pilot phase, and **more than 40% of agentic AI projects will be by 2027** [4]. It is not AI that fails to work. It is projects that are not carried through.

1.3 The cost of waiting

Leaders, for their part, document a **median ROI of 159.8% measured over 24 months** on their well-scoped AI projects [7]. The figure, drawn from Denis Atlan’s IA & ROI PME France 2022-2025 barometer (more than 200 B2B deployments analysed, a 73% success rate), corroborates the findings of more than thirty documented engagements in France [8].

The question is therefore no longer “should we do it”. It has become “should we do it now or in two years”. And the answer to that latter question is simple: the laggards of 2026 will be the casualties of 2030. Not through sudden disruption: through margin erosion and loss of competitiveness.

The **Osez l'IA** plan, launched by Bpifrance in July 2025 (an initial envelope of €200 million, within an overall drive of €10 billion in AI financing), sets the target: **80% of French SMEs and mid-caps at Orchestre level or above by 2030** [14]. The diagnosis is shared at the very highest levels of government: what separates France from the world's leaders is not technology; it is deep diffusion throughout the real economy.

1.4 Regulatory context: the AI Act after the Digital Omnibus

On 7 May 2026, the European Council and Parliament adopted the political agreement on the **Digital Omnibus** [15]. This digital regulatory simplification package reorganises the timetable for several major pieces of legislation: the GDPR, NIS2, the Data Act, and notably the AI Act (EU Regulation 2024/1689 [9]).

For SMEs, three concrete implications:

1. **Deferral of obligations for high-risk systems.** Application of the main obligations for AI systems classed as high-risk, initially scheduled for 2 August 2026, has been pushed back to **end of 2027 / 2 August 2028, depending on category**. The timetable in the April 2026 edition of this white paper is therefore out of date on this point.
2. **Administrative relief announced.** The Digital Omnibus provides for a **35% reduction in the administrative burden for SMEs** by 2029, in particular through a one-stop shop for digital compliance and the consolidation of several registers.
3. **The AI literacy obligation remains in place.** Article 4 of the AI Act, which requires any organisation deploying an AI system to ensure a sufficient level of **user proficiency**, has remained applicable since February 2025, unchanged, and with no threshold.

For an SME, the practical reading of the post-Digital Omnibus situation is as follows: the compliance window widens slightly, but the issue does not disappear. A business at Orchestre level (Chapter 2) or above must maintain a simple register of the AI systems in use, document its data policy, and ensure that users have received adequate training. The cost of compliance is low when governance is built in from phase 3 of the methodology onwards (Chapter 6); it becomes onerous when tackled after the fact.

Key takeaway. *The deferral of the high-risk obligations does not change the trajectory: it changes the pace. SMEs that have put Architecte-level governance (MATIA-4) or above in place before 2 August 2028 will be compliant by design. Others will have to catch up urgently, amid an already well-documented skills shortage.*

Chapter 2: The MATIA Method™ Scale, the 5 levels

2.1 Why a framework dedicated to French SMEs

Most AI maturity models borrow their framework from large corporates. CMMI has its own AI version. MIT publishes its own every year. Gartner, Microsoft, PwC, BCG, McKinsey: each has its own dial [16][17][18][19][20]. They all share one major flaw for an SME business leader: they were designed for large corporates. They measure

technological sophistication, model governance, data culture: concepts that do not carry the same meaning in an SME with 80 staff as they do at Saint-Gobain.

In practice, an SME business leader who self-assesses using these frameworks systematically obtains the same score: “level 1 out of 5”. The framework tells them they need to appoint a Chief AI Officer, set up an AI ethics committee, and build an MLOps platform. None of these recommendations makes sense in their context. The business leader walks away frustrated and postpones the transformation.

The **MATIA Method™ Scale** corrects this flaw. It was designed on the basis of more than thirty documented engagements with French SMEs and mid-caps since 2024 [8]. It places a company on five rungs, each associated with a memorable metaphor and a simple question. The aim is for a business leader to position themselves in under ten minutes, understand where they stand, where they need to go, and why they cannot skip steps.

Positioning note. The 5-level framework is consistent with the “golden standard” 2025-2026 frameworks (Gartner 5 levels, PwC AI-Native, Microsoft Agentic Maturity L100-L500 [16][17][18]) while remaining specific to French SMEs. No official French 5-level framework exists to date; the Baromètre France Num 2025 measures adoption as a percentage, and Bpifrance’s Diag Data IA operates as an open-ended diagnostic with no formal level [12][14]. The MATIA Method™ fills this gap.

2.2 The five stages: overview

#	Name	Business leader question	% French SMEs	Dominant risk
1	Spectateur	“AI, we’re watching from a distance. Are we missing something?”	~50%	Gradual falling behind on the sector’s productivity standards
2	Artisan	“Are my people using AI, each in their own corner?”	~30%	Shadow AI, data exfiltration, undetected hallucinations
3	Orchestre	“Is AI embedded in our processes, steered and measured?”	~13-15%	Plateauing at 2-3 use cases, lack of architecture
4	Architecte	“Is AI a structural competitive advantage?”	~2-3%	Orchestration complexity, cost drift
5	Pionnier	“Are we setting the standard for our sector?”	<0.5%	Innovation risks, not the risk of lagging behind

Total for levels 3 and above: ~16-18%, consistent with the Stanford AI Index 2026 (fewer than 10% have scaled AI within a function) [29] and with the share of “AI leaders” measured by KPMG in Q1 2026 (~11%) [2]. The

consolidated narrative: fewer than one in five French SMEs has crossed the threshold of measurable value; fewer than one in fifty sets its sector's standard.

2.3 Level 1: Spectateur

“AI, we’re watching from a distance. Are we missing something?”

The company is aware that generative AI exists. The business leader has heard about it at the Chamber of Commerce, in the business press, or over dinner with industry contacts. A few staff members may have tried ChatGPT in a personal capacity, once. But **no official tool is in place, no business use case has been identified, and no budget line exists.**

This level is not a moral failing: it is a state of watchfulness that has not yet turned into action. The danger is not being at this level in May 2026; it is still being at this level in May 2028. Meanwhile, competitors are climbing the rungs, accumulating use cases and gaining a few points of productivity every quarter. The gap widens silently.

Indicators that you are at the Spectateur level:

- No AI tool has been formally contracted or trialled through a structured process.
- The question “how many of our staff use AI?” has no answer, not even an approximate one.
- No AI budget is planned for the current financial year.
- At executive committee level, the topic is treated as one for monitoring, never for action.

% of French SMEs: ~50%. A MATIA estimate anchored on the Baromètre France Num 2025 (74% of micro-businesses and SMEs declare no AI tool at all), cross-read with Bpifrance’s usage measurements: we distinguish those genuinely not using AI (~50%) from those who already have sporadic individual uses without declaring them as tools (~24%, who move into the Artisan level) [12].

2.4 Level 2: Artisan

“Are my people using AI, each in their own corner?”

The company harbours individual, uncoordinated uses. The salesperson who drafts follow-up emails with ChatGPT. The HR manager who pre-screens applications with Claude. The technical director who tests Mistral on ad hoc questions. Even the business leader, who uses an AI assistant to prepare their presentations. Each has subscribed to their own tool, sometimes on a personal account, without management being aware of it or formalising it.

This is the territory of **Shadow AI**. These uses are not a failure: they are a signal of appetite from teams, a positive signal, one that should not be stifled. The problem is not that they exist; it is that they are **private, fragile, invisible on the profit-and-loss statement, and ungoverned.**

Three concrete risks characterise this level:

- **Data exfiltration.** An employee pastes a client contract, an HR file, or a pricing formula into a consumer AI assistant. The tool's terms of use allow it to use this data to train its models. The leak has already happened, often without management's knowledge.
- **Undetected hallucinations.** Without a usage framework or training, users take incorrect outputs at face value. An incorrect amount in a quote. An invented legal reference in a letter. Meeting minutes that fabricate decisions.
- **Personal dependency.** Know-how with the tool belongs to the employee, not to the company. When they leave, they take their practice with them.

Indicators that you are at the Artisan level:

- You know that your staff use AI, but you do not know how many, or for exactly what.
- No AI usage charter has been approved at executive committee level.
- No business use case is steered using a defined success indicator.
- No AI budget is a line item in your P&L.
- No internal champion has been designated.

% of French SMEs: ~30%. This figure is the differential between Bpifrance's adoption rate (58% reported usage in Q2 2026 [1]) and the value actually measured ($\approx 11\%$ "AI leaders", KPMG Q1 2026 [2]), refined using the Microsoft Work Trend Index 2026 on Shadow AI [13].

2.5 Level 3: Orchestre

"Is AI embedded in our processes, steered and measured?"

The company has identified 2 to 5 priority business use cases, deployed them through a genuine project-based approach, and measures their effects. An assistant for drafting quotes. A structured product-sheet generator. A customer follow-up tool. Each with an end user identified from the outset, success indicators defined before launch, and a monthly review reported to the executive committee.

This level is not an end state. It is a **state of governance**. The company has recognised that AI is a matter for senior management, not for the IT department. It has appointed two to four internal champions, operational business staff, not technicians. It has written a usage charter that distinguishes what is permitted from what is not (in particular regarding customer data and intellectual property). It measures benefits monthly, not straight off the production line but once reintegrated into operational processes.

This is where the documented median ROI of 159.8% over 24 months for well-scoped AI projects resides [7]. It is also at this level that competitive advantage starts to become visible: a salesperson who churns out 30 quotes a month instead of 18 represents 12 additional opportunities that competitors are not addressing.

Indicators that you are at the Orchestre level:

- 2 to 5 business use cases in production, measured.
- AI usage charter approved and communicated.
- Documented data policy (classification, access, retention).
- Quarterly AI committee at a minimum.
- AI budget as a P&L line item, with ROI tracked.
- AI systems register maintained (in preparation for the AI Act).
- 2 to 4 business champions identified and briefed.

% of French SMEs: ~13-15%. Anchored on the share of “AI leaders” (~11%, KPMG Q1 2026 [2]) and Bpifrance Le Lab (15-20% with structured steering [1]), adjusted for the clear-cut boundary between Orchestre and Architecte.

2.6 Level 4: Architecte

“Is AI a structural competitive advantage?”

AI is no longer a project: it is an **attribute of the business**. The organisation has structured a **proprietary knowledge base** (typically 5 to 10 years of business archives, cleaned, structured and indexed) which becomes the fuel for its AI agents. These agents are no longer simple assistants: they make operational decisions under human supervision, within defined scopes (quote generation with managerial validation, preliminary review of case files, drafting of visit reports, continuous regulatory synthesis).

The advantage is **structural**: a competitor seeking to replicate the capability would need 18 to 24 months: the time needed to build an equivalent knowledge base, clean it, structure it, and train its teams to use it [8]. This time gap becomes a measurable commercial differentiator: better tender response rates, higher-quality deliverables, increased execution speed.

The Architecte is also characterised by **documented governance**: an AI Act register maintained in real time [9][15], monthly AI committees, actively managed risk indicators (hallucinations, bias, data leaks, consumption), ISO 42001 initiated or in preparation [21]. This is the first level at which AI FinOps (Chapter 5) becomes a mandatory discipline: without an LLM gateway and without budgetary governance, costs spiral out of control.

Indicators that you are at Architecte level:

- Operational proprietary knowledge base.
- Agents in production on at least one business scope, with documented supervision.
- Up-to-date AI Act register, monthly AI committee steered by a member of the executive committee.
- Internal LLM gateway (or equivalent) in place for routing, logging and cost control.
- ISO 42001 policy being implemented.
- Data/AI architecture designed at enterprise scale.

% of French SMEs: ~2-3%. Reference: BCG *AI Radar 2026*: 15% of “Trailblazer” business leaders globally [19], adjusted downward for French SMEs (agent and governance density being rarer). Consistent with the ~16% of “Frontier” professionals orchestrating agents (Microsoft, *Work Trend Index 2026*) [13], stricter for the SME segment.

2.7 Level 5: Pionnier

“Are we setting the standard for our sector?”

The Pionnier is the business that no longer follows: it is followed. Its processes are redesigned around AI (**AI-first**) and it operates **hybrid human-plus-agent teams in production**: multi-step autonomous agents work over durations measured in hours and then days, under asynchronous supervision, with token budgets and contractual quality indicators, **orchestrated** amongst themselves and with staff. AI is here a **subject of governance at the highest level** (sponsorship from senior management, explicit accountability, **certified ISO 42001 governance** / Responsible AI) and **generates demonstrated value on the income statement**, not just pilots. The business capitalises on **proprietary knowledge and data** (“Owned Intelligence”) that compound over time and differentiate it durably, and maintains a culture of “AI literacy” spanning business leaders, managers and operational staff, together with the capability to publish its own sector benchmarks.

The Pionnier is not a large group. It is an SME or a mid-cap that made the early choice of deep integration and is reaping the competitive differential today. Microsoft, in its **Work Trend Index 2026**, speaks of “**Frontier Firms**” [13]; PwC of “AI-Native” [17]; BCG, in its *AI Radar 2026*, of “**Trailblazer**” business leaders [19]. The vocabulary varies; the reality is the same: organisations that have made AI a constitutive attribute. Two markers often presented as “golden standard” deserve an anti-hype clarification. **Data sovereignty** (sovereign cloud, SecNumCloud, or on-premise) is a **deployment choice dependent on sector-specific constraints**, not a maturity tier: no global 2026 maturity framework (Gartner, McKinsey) places hosting location at the summit. Genuine differentiation through AI comes from proprietary data and knowledge (“Owned Intelligence”), not from on-premise deployment. As for the “**agent-to-agent**” (A2A) business, it plays out across **two layers that must be distinguished**. The **interoperability layer**, the **A2A** protocol (v1.0 “production-ready”, 150+ organisations, *Linux Foundation, April 2026*), **MCP**, exposure of processes via API, is **already achievable**: making one’s information system *agent-ready* (processes properly exposed, both to one’s own agents and to those of one’s customers and suppliers) is a **2026-2027 action objective**, not merely a watching brief. The **transactional layer “on both sides”** (widespread inter-agent commerce and agentic payment) remains, for its part, an **emerging frontier, 2028 horizon**: the first vertical deployments (ad-tech, June 2026), niche volumes, no established SME use case. The Pionnier therefore does not simply undergo A2A: it is **already preparing its interoperability** and experimenting with agentic transactions as the field matures.

Indicators that you are at Pionnier level:

- **AI-first** processes and **multi-agent orchestration in production**, with a measured share of critical workflows steered by agents.
- Autonomous agents supervised asynchronously (hours/days), token budgets and **contractual quality indicators**.
- **Widespread rollout of assistants** to almost all eligible roles, measured adoption, with a core of staff **orchestrating agents** (human-agent teams), *a stated target, never presented as a state achieved at 100%*.
- **AI governance steered by senior management**: explicit oversight (leadership sponsorship, an AI steering function, a committee reporting to senior management) and accountability. It is the **function** that is measured, not the title.
- **Certified Responsible AI / ISO 42001 governance** and AI Act compliance (Article 4 literacy, post-Digital Omnibus timetable).
- **Demonstrated value on the income statement** (agent ROI measured, not merely deployed) and **mature AI FinOps** (cost per task, managed token budget, internal LLM gateway).
- **“Owned Intelligence”**: proprietary data and knowledge capital, continuous organisational learning, published sector benchmarks.
- **Sovereignty & control of models** (2030 target): mastered sovereign AI, domain-adapted models (RAG + targeted fine-tuning where justified), **local/on-premise option for critical data**, full reversibility.
- **Resilience & imperviousness to the splinternet** (2030 target): multi-cloud redundancy and on-premise failover capability guaranteeing continuity even in the event of Internet fragmentation or a hyperscaler failure.
- **~100% cultural adoption** (2030 target): virtually all staff are “augmented” by a supervised AI assistant, within a culture of continuous AI literacy.
- **Technical debt < 1%** (2030 target): continuous purging via ultracoding (multi-agent orchestration + adversarial review), AI debt being contained rather than accumulated.
- **Security: score > 99.5%** (2030 target): NIS2/PQC-ready, agentic observability, secret management, zero untracked incidents.
- **Agentic interoperability** (2026-2027 objective): *agent-ready* information system, processes exposed via **MCP** and the **A2A** protocol to its own agents as well as to those of its partners; inter-agent transactions and agentic payment “on both sides” remain a frontier to be explored (2028 horizon).

These markers are read across **nine graduated dimensions** from Spectateur to Pionnier. The full grid appears in §2.10. Level 5 there describes the **“golden standard 2030” target** towards which the Pionnier is heading; the thresholds (100% augmented, debt < 1%, security > 99.5%) are apex targets, achieved today by fewer than 0.5% of SMEs. This is precisely what makes them rare.

% of French SMEs: <0.5%. Virtually non-existent to date. McKinsey, in its *State of AI Trust 2026* (March 2026), measures average maturity at **2.3 out of 4**, with **only one organisation in three at level 3 or above** [6]; set against the fabric of French SMEs, the fully “pioneering” fraction is smaller still. The Pionnier of 2026 will be the sector benchmark of 2030.

2.8 The golden rule: no level can be skipped

This rule is the most counter-intuitive and the most frequently broken. An ambitious business leader at Artisan level sees an impressive demonstration of autonomous agents at a conference. They return to the office and launch a €200,000 project to build an agent platform for their business. Six months later, the project is halted. Not because the technology does not work: it works very well at Saint-Gobain. Because the foundations are not there.

In concrete terms, skipping the Orchestre level means investing in a sophisticated platform while:

- business data is neither structured nor accessible;
- end users have no working habits with AI;
- risk governance does not exist;
- no internal champion is able to drive adoption within their team;
- success indicators have not been defined prior to launch.

In **9 cases out of 10**, this project fails. Not for lack of technology: for lack of foundations. This is exactly the definition of the “**POC graveyard**” that Gartner documents [4]. The *CIO Playbook 2026* (IDC-Lenovo, January 2026) quantifies it: **only 46% of POCs reach production**. Half remain stuck at the pilot stage [27].

Irreducible durations, as measured in engagements [8]:

Transition	Typical duration	Investment (SME 50-500)
Spectateur → Artisan	3 to 6 months	€5,000 to €15,000
Artisan → Orchestre	6 to 12 months	€30,000 to €80,000
Orchestre → Architecte	12 to 24 months	€80,000 to €250,000
Architecte → Pionnier	24 to 36 months	Structural investment, partly recognised outside the P&L

A **Spectateur → Architecte transformation takes 3 to 4 years**. Spectateur → Pionnier takes 5 to 7 years. This is the irreducible duration of organisational maturation, independent of the speed of the technology. Tools accelerate; people, processes and data follow their own pace.

Key takeaway. *No level can be skipped. An Artisan-level SME that invests €200,000 in an agent platform fails in 9 cases out of 10, not for lack of technology, but for lack of foundations [8]. Spectateur → Architecte takes 3 to 4 years: this is the pace of people, processes and data, not that of tools.*

2.9 How to position yourself in 10 minutes

Here is a quick test, derived from the 12-point MATIA Method™ grid used in client engagements. Count your “yes” answers across the following 12 statements:

1. At least one business AI use case is in production with its indicators measured.
2. An AI usage charter has been approved by the executive committee.
3. An internal AI champion (not from the IT department) has been appointed and briefed.
4. A data classification policy is documented.
5. An AI budget is a line item in the current year’s P&L.
6. An AI committee meets at least quarterly with a member of the executive committee.
7. The AI Act register of AI systems in use is kept up to date.
8. Structured AI training has been delivered to more than 30% of staff.
9. A proprietary knowledge base has been built or is under construction.
10. At least one AI agent operates under supervision on a business process.
11. An LLM gateway (or equivalent) centralises calls and logging.
12. The ISO 42001 process is under review.

Reading your results:

Yes	Level	First objective
0	Spectateur	60-90 minute diagnostic, minimal charter, first use case scoped
1-2	Artisan	Structure the foundations (charter, champion, first use case run as a project)
3-5	Transition Artisan → Orchestre	Maintain the discipline (scoping before deployment, monthly measurement)
6-8	Stabilised Orchestre	You are among the 13-15%. First target: median ROI of ~160% (24 months)
9-10	Architecte	Structural competitive advantage; FinOps and ISO 42001 become mandatory
11-12	Pionnier	The standard for your sector: capitalise and publish

2.10 The nine dimensions of maturity, level by level

Overall positioning breaks down into **nine dimensions**. Each progresses from Spectateur to Pionnier. Level 5 describes the **“golden standard 2030” target** towards which the Pionnier is heading, achieved today by fewer than

0.5% of SMEs, which is what makes it rare. The numerical thresholds ($\approx 100\%$, $< 1\%$, $> 99.5\%$) are apex targets, not market averages.

1. Sovereignty & control of models. *Spectateur*: consumer-grade tools, data not controlled. *Artisan*: personal accounts, exposed data. *Orchestre*: classified data, EU provider / SecNumCloud for sensitive data. *Architecte*: sovereign cloud + RAG on proprietary data, targeted fine-tuning where justified. **Pionnier (2030 target)**: mastered sovereign AI, domain-adapted models (RAG + targeted fine-tuning), **local/on-premise option** for critical data, full reversibility.

2. Resilience & imperviousness to the splinternet. *Spectateur*: total dependence on a single provider, risk ignored. *Artisan*: single provider, no continuity plan. *Orchestre*: basic disaster recovery/business continuity planning, data replicated within the EU. *Architecte*: multiple providers, proven reversibility. **Pionnier (2030 target)**: multi-cloud redundancy + on-premise failover capability, **continuity guaranteed even in the event of Internet fragmentation (splinternet)** or a hyperscaler outage.

3. Augmented staff. *Spectateur*: 0% official usage. *Artisan*: $\sim 10\text{-}30\%$, individual and unsupervised. *Orchestre*: $\sim 30\text{-}50\%$, priority roles equipped and measured. *Architecte*: $\sim 70\%$ of eligible roles. **Pionnier (2030 target)**: **$\sim 100\%$ of eligible staff “augmented”** by a supervised AI assistant, with a core orchestrating agents.

4. Cultural adoption & AI literacy. *Spectateur*: none. *Artisan*: self-taught, uneven. *Orchestre*: three-tier training launched ($> 30\%$ trained). *Architecte*: literacy rolled out organisation-wide, role-specific pathways. **Pionnier (2030 target)**: **$\sim 100\%$ culturally adopted**, continuous organisational learning, business leaders themselves “augmented”.

5. AI governance. *Spectateur*: none. *Artisan*: none (Shadow AI). *Orchestre*: usage charter, quarterly committee, AI Act register, prior consultation of the works council (CSE) for internal AI (businesses with ≥ 50 employees, Art. L.2312-8 of the French Labour Code). *Architecte*: monthly committee steered by the executive committee, ISO 42001 initiated, risk indicators tracked, structured dialogue with staff representative bodies. **Pionnier (2030 target)**: **golden standard governance**, certified ISO 42001, Responsible AI, senior management oversight, published sector benchmarks, permanent dialogue with staff representative bodies.

6. Technical debt. *Spectateur*: not applicable. *Artisan*: hidden debt (ad hoc scripts and unsanctioned integrations). *Orchestre*: debt tracked, backlog identified. *Architecte*: debt under control ($< 10\%$), continuous review. **Pionnier (2030 target)**: **technical debt $< 1\%$** , continuous purging via ultracoding (multi-agent orchestration + adversarial review).

7. Security (security score). *Spectateur*: attack surface not controlled. *Artisan*: possible data leaks (Shadow AI). *Orchestre*: access control, secret management, data policy. *Architecte*: NIS2-ready, encryption, regular audits (score $> 95\%$). **Pionnier (2030 target)**: **score $> 99.5\%$** , PQC-ready, agentic observability, zero untracked incidents.

8. Workflows & agents. *Spectateur*: none. *Artisan*: one-off assistants, outside processes. *Orchestre*: 2 to 5 use cases in production, measured. *Architecte*: agents under supervision on defined scopes. **Pionnier (2030 target)**: **AI-first workflows**, multi-agent orchestration in production (durations in hours and then days), **an agent-ready information system** (A2A/MCP interoperability, 2026-2027 objective).

9. Value & FinOps. *Spectateur*: no value measured. *Artisan*: costs not measured. *Orchestre*: ROI tracked, AI budget entered in the P&L. *Architecte*: mature FinOps (LLM gateway, cost per task). **Pionnier (2030 target)**: **demonstrated value on the income statement**, optimised FinOps, “Owned Intelligence” capital that compounds over time.

Key takeaway. *These nine targets do not describe the market: they describe the summit. A Pionnier does not “tick” nine boxes overnight: it advances each dimension, level by level, without skipping any. It is the conjunction of all nine, and their rarity, that defines the sector standard of 2030.*

Chapter 3: The 5 fatal mistakes

43% of major AI initiatives are deemed doomed to failure, for lack of execution (*HCLTech, May 2026*) [5]. Gartner, Deloitte and McKinsey converge on similar findings [4][20][6]. And the *CIO Playbook 2026* (IDC-Lenovo) drives the point home: only **46% of POCs reach production** [27]. Five mistakes are responsible for this [8]. None is technological. All are mistakes of execution discipline.

Mistake 1: The gadget syndrome

Choosing a tool before formulating a problem. “We will go with Microsoft Copilot.” “We will deploy Mistral in-house.” The tool arrives before the question. Six months later, nobody has a clear business use for it, and the subscription costs more than it delivers.

Antidote: never buy a tool before naming the process, the target user(s), and the expected, quantified benefit. The MATIA rule is explicit: *process audit precedes tool selection*.

Mistake 2: The perpetual-POC trap

More than half of AI POCs remain dead letters: only 46% reach production (*CIO Playbook 2026, IDC-Lenovo*) [27]. The pilot phase lasts six months, delivers encouraging results, and then... nothing. No transition to production, no change management plan, no deployment budget. The “POC graveyard” has become a structural feature of poorly scoped AI transformations [8].

Antidote: define the criteria for the transition to production before the POC begins. No POC without a production plan. If a use case does not deliver on its promises by sprint 1, it is stopped, and another takes its place. This discipline costs pride; it earns ROI.

Mistake 3: The lone-leader illusion

The business leader carries the matter alone. He grows enthusiastic, he prototypes, he evangelises. But nobody in the organisation takes up the torch. Six months later, the tool is used by him and three close colleagues. The others carry on “as before”. This is the most common configuration of leadership-driven Shadow AI [13].

Antidote: from the outset, appoint two to four **internal champions** (not technicians, but business operations staff) who champion usage within their teams and report to the executive committee. The business leader sets the trajectory; the champions hold the ground.

Mistake 4: The mirage of instant ROI

AI does not create value straight out of the box. It creates value within **redesigned processes**. Bolting an AI assistant onto a poorly defined process merely amplifies the noise. ROI comes when the process has been simplified *before* AI is added to it.

Antidote: every use case begins with a simplification of the process, not the addition of a tool. The *process audit* always precedes the *tool selection*. One hour of scoping saves twenty hours of deployment [8].

Mistake 5: Data blindness

Without reliable fuel, the most sophisticated engine will not start. AI plugged into poorly structured, poorly cleaned or poorly accessible data produces hallucinations, contradictions and, ultimately, distrust. This is the main cause of abandonment during the pilot phase.

Antidote: audit data quality *before* choosing the use case. If the data is absent or too noisy, the use case does not go ahead. Another one is found. An SME can begin structuring its last 5 years of business archives while deploying its first use cases: this is partly what makes the difference between Orchestre and Architecte.

Chapter 4: Sovereignty & infrastructure

The adoption of generative AI in SMEs raises three questions of sovereignty that were absent (or marginal) from the debate two years ago. In 2026, they have become central to the choice of architecture.

4.1 The CLOUD Act and what it means for an SME

The **Clarifying Lawful Overseas Use of Data Act** (CLOUD Act, United States, 2018, §2713) [22] allows US authorities to compel the disclosure of data held by a US-based service provider, even where that data is physically stored in Europe. Case law is now well established: a US-law entity hosting European data is required to cooperate with a US disclosure order, unless a specific bilateral agreement provides otherwise.

Practical implications for a French SME:

- Data sent to a generative AI provider under US law (OpenAI, Anthropic via its US entity, Google Cloud, Microsoft Azure where the configuration is not strictly European) may be accessible to US courts without notification to the French client.
- For an SME handling confidential B2B data, pricing, margins, intellectual property or sensitive personal data, this accessibility creates a legal and commercial risk, particularly in the healthcare, defence, energy and finance sectors, and more broadly for any business exposed to contracts that impose data sovereignty requirements (public administrations, OIV, OSE).

The practical MATIA rule: classify data *before* choosing the provider. Data for which exposure to the CLOUD Act is unacceptable must remain with a provider under European law or on-premise. Other data may legitimately pass through a US provider, provided the contractual terms allow it.

4.2 On-premise vs sovereign cloud: decision matrix

Three options are available to an SME, depending on its sensitivity and maturity level:

Option	Who it is for	Indicative cost (50-500 employees)	Maturity required
Hyperscaler cloud (US): OpenAI, Anthropic, Microsoft Azure, Google Cloud	Low-sensitivity data, unregulated sector	€5,000 to €30,000/year (licences + calls)	Artisan / Orchestre
European sovereign cloud: OVHcloud, Scaleway, Outscale (self-hosted Mistral, Llama models, or sovereign API)	Sensitive data, regulated sectors, commercial sovereignty requirements	€10,000 to €50,000/year	Orchestre / Architecte
On-premise: on-site GPU, self-hosted Mistral or Llama models	Critical data, OIV/OSE, strategic IP	€50,000 to €250,000 initial investment, then €30,000 to €80,000/year	Architecte / Pionnier

On-premise ROI: according to MATIA field feedback [8], an industrial SME that shifts 80% of its AI calls to on-premise (Mistral or Llama-class models hosted on internal infrastructure) recovers its initial investment in 18 to 30 months if volumes exceed 5 million tokens per month. Below that threshold, European sovereign cloud remains more cost-effective. Quantization (section 4.4) reduces the capex component and can bring this payback closer to the lower bound for narrow, high-volume use cases, but the determining factor remains adoption, not hardware.

4.3 Open-weight has caught up with the frontier: what this means for sovereignty

For two years, the “sovereignty” argument ran into a technical objection: open models were significantly behind proprietary US models. In 2026, that objection fell away, and the gap has continued to narrow month after month.

- **The capability gap has narrowed to a few points.** A year ago, the best open-weight model (DeepSeek V3 0324) topped out at **22** on the Artificial Analysis Intelligence Index, 13 points behind the best proprietary model; by 30 April 2026, the best open models (Kimi K2.6, MiMo V2.5 Pro) had reached **54 against 60 for GPT-5.5, a gap of 6 points** [33]. And the momentum has not slackened: on 16 June 2026, **GLM-5.2** (Z.ai lab, MIT licence) became the first open-weight model in the index and matched GPT-5.5 on the GDPval-AA agentic sub-benchmark [33].
- **A French laboratory remains in the frontier race under a permissive licence. Mistral Large 3** (2 December 2025) is a Mixture-of-Experts model with 675 billion parameters (41 billion active), released under the **Apache 2.0 licence** [31]: an SME may legally download it, fine-tune it and run it on its own hardware, with no royalty and with no data leaving its premises. Mistral has since released **Mistral Medium 3.5** (30 April 2026, 128 billion, open-weight), a mid-range model under a modified MIT licence, but useful for sovereign deployments at controlled cost [31].
- **The hardware floor has collapsed.** Google released **Gemma 4** in April 2026 (Apache 2.0 licence): the unquantised weights fit on **a single 80 GB GPU**, and quantised versions run **offline on a phone, a Raspberry Pi or an NVIDIA Jetson board** [32]. The “open versus closed” debate is no longer ideological.
- **Usage prices are collapsing too.** In hosted API form, a leading open model such as DeepSeek V4-Flash is priced at **~\$0.14 / \$0.28 per million tokens** (input/output) [43]. *A caveat worth bearing in mind:* open models remain weaker on factual recall (SimpleQA test), hence the importance of grounding (RAG) and verification, which are matters of **method**, not model.

Business leader takeaway. Model weights have become a *commodity*: raw capability can be bought, rented or downloaded. From this point on, an SME’s lasting advantage can no longer come from the model itself: it comes from how the business organises itself around it (proprietary data, redesigned processes, governance). Sovereignty, once a costly compromise on performance, is now a **genuinely available option**.

4.4 Hardware no longer justifies a wait-and-see approach: quantization

Two long-standing objections to on-premise deployment, quality loss and hardware cost, have been largely resolved in 2025-2026 by **quantization** (reducing the numerical precision of the model’s weights).

- **Quality holds up at low precision.** The 2026 measurements (versus full BF16 precision) are unambiguous: at 4 bits, degradation falls to **only –1.2% to –2.5%** on the major benchmarks (MMLU, HumanEval, MATH), for retained quality of around **95% in INT4 (AWQ/GPTQ), 97-98% in INT8 and ~99% in FP8** [34]. A 70B model thus goes from ~140 GB to **36-40 GB** in INT4 [35]. Degradation is noticeable on complex reasoning, acceptable on summarisation or classification: a trade-off that, once again, comes down to method.
- **Hosting costs have fallen sharply.** A 70B model quantised to 4 bits fits on ~48 GB of VRAM, that is, a single-card professional-grade workstation (in the order of €3,000-3,500, per vendor sources). The argument “we cannot afford a data centre” no longer holds.

The challenge therefore shifts from hardware to **integration, security and change management**, which confirms, rather than contradicts, the thesis of this white paper: transformation is won through method.

4.5 Where European sovereign infrastructure stands

Let us be clear-eyed about the balance of power: Europe will not win the race for raw computing power. The United States accounts for **~74.5%** of AI supercomputer performance, China for **~14.1%**, and the European Union as a whole for **~4.8%** (Epoch AI, May 2025, latest available measurement) [36]; in April 2026, the CNAS *Sovereign AI Index* summarised the same concentration: **the United States and China together control ~90% of the computing power required for frontier AI** [36]. On the market side, US hyperscalers control **more than 85% of the European cloud market** [37]. This is precisely why an SME's advantage is organisational, and why choosing a European building block is an act of risk management, not activism.

This act is now practicable, and the sovereign market is scaling up fast: according to Gartner (February 2026), global spending on **sovereign cloud (IaaS) will reach \$80 billion in 2026 (+35.6%)**, with **Europe surging by +83%** (from \$6.9 billion to \$12.6 billion), to the point of overtaking North America as early as 2027 [48].

- **Commercial providers available now.** OVHcloud AI Endpoints offers more than 40 open-weight models via API, hosted in Europe, with zero data retention and reversibility; Scaleway operates NVIDIA H100 clusters and OpenAI-compatible generative APIs, with data stored in Europe [38]. OpenAI compatibility keeps migration costs low.
- **A funded infrastructure trajectory.** Mistral is targeting 200 MW of compute capacity in 2027 and **1 GW in 2030** (a €4 billion plan), with a sovereign data centre at Bruyères-le-Châtel (44 MW, ~13,800 NVIDIA GB300 GPUs, coming online mid-2026) backed by an \$830 million debt raise led by Bpifrance (March 2026) [39]. At European level, the **InvestAI** initiative is mobilising ~€200 billion towards 2030, including **€20 billion** for 4 to 5 “AI Gigafactories”; in June 2026, a French consortium, **AION** (Scaleway, EDF, Orange, Iliad, Capgemini...), was formed to bid for one of them [40][49].

Business leader takeaway. Sovereign capacity that can be used *immediately* comes from commercial providers (OVHcloud, Scaleway, Mistral); gigafactories are a 2027+ horizon. An Orchestre-level SME can therefore start today on a European sovereign cloud, while retaining reversibility to on-premise for its most critical data.

4.6 ANSSI's PQC timeline (2027-2035)

ANSSI published its guidance on migration to **post-quantum cryptography (PQC)** as early as 2022, updated in 2024 [23]. The timeline is now explicit:

- **2027** (ANSSI): end of security certifications and qualifications for products lacking post-quantum cryptography, confirmed in June 2026.
- **2030** (ANSSI): public administrations and businesses should only acquire products that incorporate PQC.
- **2035**: full transition expected, a deadline set by the **EU's coordinated PQC roadmap** (June 2025), which ANSSI's national position anticipates [23].

For an SME, the implication is not immediate, but it is structural: the AI architecture choices made today (encryption of knowledge bases, of the channels used to call models, of logs) must be compatible with the PQC migration. In concrete terms, this means favouring providers that publish a documented PQC trajectory and disclose their timeline (OVHcloud, Scaleway, the main open-source building blocks). Inaction is paid for in 2027-2030 through an emergency migration, always more costly.

4.7 NIS2 and the Data Act: concrete obligations

NIS2 (EU Directive 2022/2555) [24] extends cybersecurity obligations to a broader range of businesses. For an SME, the practical question is whether it is:

- An **essential or important entity** under NIS2 (energy, healthcare, transport, finance, water, essential digital service sectors, etc.). If so, enhanced obligations apply: a risk management policy, incident notification within 24 hours, and training for business leaders.
- A **supply-chain provider** to an essential entity. If so, potential audits and enhanced contractual obligations apply, even without being directly regulated.

The **Data Act** (EU Regulation 2023/2854) [25], whose main provisions have applied since **12 September 2025**, re-organises rights over data generated by connected devices and imposes interoperability and portability obligations. For industrial SMEs or providers of connected services, the Data Act creates both opportunities (contractual access to new data) and obligations (making their own data available to customers).

4.8 Practical recommendation: the 2×2 matrix

A two-dimensional decision matrix (**data sensitivity × infrastructure budget**) clarifies the choice:

	Constrained infra budget (< €30,000/year)	Available infra budget (> €30,000/year)
Low-sensitivity data	Hyperscaler cloud with strict access control	European sovereign cloud (MATIA recommendation, best security/performance ratio)
Sensitive data	European sovereign cloud, excluding US hyperscalers for sensitive data	On-premise for critical data, sovereign cloud for the rest

The MATIA rule: **start simple, segment progressively**. An Orchestre-level SME can start with a single European sovereign cloud and progressively migrate its critical data to on-premise as volume and criticality justify it.

Chapter 5: Mastering the cost of AI: FinOps & governance

The cost of generative AI is not linear. Without governance, it spirals out of control: proliferation of accounts, redundant calls, oversized models, absence of caching. An SME at Artisan level can consume €2,000 to €10,000 per month in subscriptions and API calls without measuring the value obtained. An SME at Architecte level that has mastered its FinOps consumes €1,500 to €3,000 per month, and derives measurably more value from it. The difference does not lie in the technology; it lies in the governance.

The 2026 FinOps paradox. The price of the token is collapsing: at constant performance, a study revised in March 2026 (Gundlach et al., *The Price of Progress*) measures a decline of **~5 to 10× per year** for frontier models (and up to **31× per year** for the highest-performing segment), with algorithmic efficiency alone contributing ~3× [41]. In June 2026, the rate actually paid by companies stood at around **\$0.72 per million tokens on a weighted average ba-**

sis, with the lightest models falling to **\$0.07** [42]. And yet AI bills are exploding: between January 2025 and April 2026, **token consumption jumped by +1,001% while spending rose by +497%** (Ramp, June 2026) [42], because agentic workflows multiply the number of calls (an agentic task consumes 5 to 30× more tokens than a chatbot request). The symptom is well known: in early 2026, **Uber had exhausted its entire annual AI budget by April** [42]. **Cheaper tokens do not produce a lower bill: they produce a bill that is harder to steer.** The lever is therefore no longer the unit price: it is governance.

5.1 Why AI costs are exploding in SMEs

Three mechanisms fuel the spiral:

1. **Proliferation of individual accounts.** Each employee subscribes to their own tool (ChatGPT Plus, Claude Pro, Mistral, etc.). No consolidated visibility. No negotiating power.
2. **Redundant calls.** Without caching, the same prompt is billed on every execution. Without intelligent routing, simple requests are routed through the most expensive models (“overusing the Ferrari to fetch the bread”).
3. **Absence of a baseline and thresholds.** No one knows how much the company is spending on AI in month M; so no one knows whether month M+1 is drifting off course.

All of this set against a backdrop of **Shadow AI**: cumulative personal usage (Chapter 1.2) often amounts to more than the official budget.

5.2 LLM Gateway: the primary lever

An **LLM Gateway** (a centralised proxy between users/applications and AI models) is the number one operational lever for cost control. Several options coexist as of May 2026:

- **Open-source solutions:** LiteLLM, Portkey (free edition), Helicone, Langfuse, OpenLLMetry. Marginal infrastructure cost, deployment in 1 to 3 days.
- **Commercial SaaS solutions:** Portkey (paid edition), TrueFoundry, Cloudflare AI Gateway, Mistral and OpenAI Enterprise integrations.
- **In-house build:** relevant for Architectes / Pionniers wanting tight integration with their stack.

Benefits observed on MATIA engagements [8]:

- **Cost reduction of 30 to 60%** on gross AI billing, through the combined effect of caching (10-25%), intelligent routing to suitable models (10-20%), and rationalisation of calls (5-15%).
- **Centralised logging:** for the first time, the company knows *who* is calling *which model* on *what subject* and at *what cost*. This is the basis of all steering.
- **Rate limiting:** protection against usage drift (a looping script that consumes \$10,000 by accident).
- **Access control and secret management:** API keys are no longer embedded in employees’ code; they are centralised and auditable.

The right reflex for 2026: route to small models. The 2026 consensus favours a heterogeneous architecture: simple or well-scoped tasks are assigned to a **small language model** (SLM), and only complex requests are escalated to a large model. For high-volume repetitive tasks, this routing **reduces inference cost by up to 90%** while delivering near-instant latency, to the point that NVIDIA now describes SLMs as the “future of agentic AI” (*InfoWorld*, May

2026) [44]. The discipline consists of reserving the (expensive) frontier model for difficult reasoning, and assigning the rest to a cost-effective SLM, often open-weight and hostable in-house (Chapter 4.3). This is orchestration, and therefore a matter of method.

Deploying an LLM gateway is one of the first projects to undertake as soon as the transition from Orchestre to Architecte begins. Below that level, an Artisan-level SME can start with a lightweight open-source solution.

5.3 ISO 42001: the governance framework

The **ISO/IEC 42001:2023** standard [21] is the first international **AI Management System** (AIMS) framework. It is to AI systems what ISO 27001 is to information security: a certifiable governance framework.

For an SME, **ISO 42001 certification is not a prerequisite**: it is an 18-36 month target for Architecte and Pionnier levels. However, **the principles** are immediately applicable and structure governance from Orchestre level onwards:

1. **AI policy** documented and signed off by management.
2. **Register of AI systems** in use, with risk classification (in alignment with the AI Act [9][15]).
3. **Risk assessment** documented per system (impact, likelihood, mitigation measures).
4. **Management review**, at least every six months.

These four minimum elements can be put in place in 6 to 12 weeks at marginal cost for an SME at Orchestre level. They constitute a commercial asset **as of today**: large contracting authorities and public administrations are starting to include ISO 42001 in their selection criteria, around 40% of AI tenders in mid-2026, according to field feedback [8].

5.4 TCO and ROI: the right equation

The **TCO (Total Cost of Ownership)** of an AI initiative in an SME breaks down into six cost items:

Cost item	Typical share of Year 1 TCO	Change in Years 2-3
Support (transition manager, consulting)	35-50%	Falls sharply
Licences and API calls	10-25%	Stable or growing
Infrastructure (cloud, GPU)	5-15%	Grows with maturity
Integration and development	10-20%	Falls after implementation
Skills (training, recruitment)	10-20%	Stable, a structural dimension
Governance and compliance	3-10%	Grows with maturity

ROI breaks down into three levers:

1. **Measured productivity gains** (hours saved × fully loaded hourly cost).
2. **Attributable additional revenue** (revenue generated by freed-up capacity, increased conversion rate).
3. **Quantifiable risks avoided** (non-compliance, data leakage, customer loss).

The MATIA rule: **measure the three levers monthly, aggregate quarterly, never quote a ROI figure without the methodology underpinning it.** The median ROI of 159.8% [7] is credible because it is measured, over 24 months and more than 200 deployments; an ROI figure announced without a methodology is a marketing number.

5.5 Monthly budget governance

Four rituals structure AI FinOps in an SME at Orchestre level or above:

1. **Consolidated monthly dashboard:** total cost, cost per use case, cost per active user, cost/value-generated ratio.
2. **Threshold alerts:** automatic triggering beyond a monthly budget defined per use case.
3. **Quarterly review** by the AI committee: model optimisation, account rationalisation, supplier negotiation.
4. **Annual review** incorporated into the following year's budget, with a development roadmap.

These rituals cost little (2 to 5 hours per month for the AI champion) and prevent overruns that can reach €100,000 to €200,000/year in a mid-market SME.

Key takeaway. *The postponement of AI Act obligations following the Digital Omnibus (Chapter 1.4) does not mean postponing internal governance, quite the opposite. It is an 18 to 24-month window to embed FinOps and ISO 42001 practices before compliance becomes binding. SMEs that use this window will gain a governance advantage that latecomers will only be able to catch up on in an emergency.*

Chapter 6: The 5-phase methodology

The MATIA Method™ approach unfolds over five phases across 12 months. It is designed for an SME and mid-cap company of 50 to 500 employees, supported by an external transition manager throughout. The total budget (support + licences + tools) is between €30,000 and €80,000 for an Artisan → Orchestre transition [8].

Phase 1: 360° diagnostic (weeks 1 to 3)

Mapping of critical business processes, an audit of data quality and accessibility, the stance of the business leader and the executive committee on AI, and identification of likely resistance and potential champions. The deliverable is a 15- to 25-page report, presented to the executive committee, which concludes with 3 to 5 priority use cases ranked by impact/effort ratio.

The analysis grid is based on the 12 points of the MATIA Method™ Scale (Chapter 2.9): process maturity, data maturity, tools maturity, management stance, governance, training, security, compliance, measurement, change, infrastructure, budget. Each point is scored from 0 to 5.

Phase 2: Scoping the use cases (weeks 4 to 6)

For each case selected in phase 1, a 2- to 4-page scoping document: precise scope, named target users, data involved, candidate tools with selection criteria, costed budget, quantified expected ROI, identified risks, production go-live plan, defined success indicators.

One hour of scoping saves twenty hours of deployment. This is the most consistent lesson from field practice [8]. Use cases that skip this step always resurface in phase 4, in a degraded form and with a delayed schedule.

Phase 3: Preparing the foundations (weeks 7 to 12)

In parallel with the first two use cases getting under way, the company builds its foundations. An AI usage charter approved by the executive committee. A documented data policy (classification, access, retention). A three-tier training plan: executive committee, managers, operational staff. A target architecture defined (cloud, hosting, orchestration building blocks, LLM gateway). Structuring of priority data. Initial ISO 42001 elements (policy, register, risk assessment).

By the end of this phase, the company has become capable of making its deployments succeed. This is the invisible investment that makes the difference between the ~11% of “AI leaders” who reap the rewards and the majority who merely spread it thinly [2][29].

Phase 4: Pilot deployments (weeks 13 to 26)

Sprints of 4 to 6 weeks, one use case per sprint. For each sprint:

- An end user identified from the outset and committed to the outcome.
- Success indicators defined before launch.
- A mid-point review and a go-live review.
- A monthly executive committee review that aggregates the sprints.

The rule is strict: no go-live without the success indicators being met. If a use case fails to deliver on its promise, it is stopped, and another takes its place. This discipline costs pride; it earns ROI.

Phase 5: Consolidation and governance (months 7 to 12)

Industrialisation of the use cases that have gone live: optimisation, integration into in-house tools, a user upskilling plan, continuous measurement of benefits. Establishment of lasting governance: an up-to-date AI Act register, a quarterly AI committee chaired by an executive committee member, a monthly FinOps dashboard, a documented year 2 roadmap.

At the end of the 12 months, the company is solidly established at Orchestre level. The trajectory towards Architecte is mapped out and budgeted for the following 12 to 24 months.

Commanding AI: the shift from prompting to context engineering

A cross-cutting skill spanning all five phases emerged in 2025-2026 and now structures the practice: **context engineering**, or “commanding AI”. Where 2023 required learning to write prompts, 2026 requires **building the context**

in which AI operates: choice of model, structure of the request, integration of business knowledge, quality controls, guardrails, measurement of output.

For an SME at Orchestre level or above, context engineering is not an individual skill. It is a **team discipline**:

- **For business leaders:** “AI literacy”, understanding the principles well enough to steer, without needing to prompt themselves.
- **For managers:** the ability to scope a use case in terms of context (data involved, guardrails, indicators).
- **For operational staff and AI champions:** mastery of tools, structured prompts, and edge cases.

This transition, from *prompt* to *context*, is as structurally significant as the shift from typewriting to office computing in the 1990s. Companies that train for it now will gain a lasting advantage. See the portfolio article “The end of prompt engineering” for an in-depth treatment (28 sources) [26].

Chapter 7: INDUSTEC case study

INDUSTEC (an anonymised name) is a French industrial SME with 78 staff and €23.5 million in revenue [10], specialising in the design and manufacture of technical equipment for the process industries. Engagement launched in April 2025, measured at 9 months (January 2026) [10].

Starting point: Artisan level

At the time of the April 2025 diagnostic, AI was present in the company in three forms, all of them individual: a sales representative who used ChatGPT to pre-draft follow-up emails, the communications manager who used it for LinkedIn posts, and the technical director who tested Mistral on ad hoc questions. No shared tool. No process affected. No benefit measured [10]. A typical Artisan-level profile (Chapter 2.4).

Priority use case: drafting technical quotes

The diagnostic identified the drafting of technical quotes as the priority target. Volume: 600 to 800 quotes per year, each requiring 2 to 4 hours of an application engineer’s time for the drafting component (technical description, product references, warranty terms, commercial terms). That is approximately 1,800 to 2,400 hours per year [10].

Phase 2 scoping: an internal assistant drawing on a structured product repository, in-house quote templates, and the history of 1,800 quotes issued since 2022. Success indicator: **a reduction of at least 50% in drafting time per quote** [10].

Results at 9 months

Measured in January 2026, covering the last 6 months of production use [10]:

- **275 hours saved per month** [10] on quote drafting. This is equivalent to 1.7 FTE freed up, redeployed to after-sales account management and upstream client visits.
- **+18% in revenue across the commercial scope concerned** [10]. The additional capacity freed up made it possible to respond to more tenders and carry out more upstream client visits. The quote → order conversion rate rose from 38% to 44% [10].
- **ROI of 182% documented at 9 months into the engagement, consolidated over 12 months** [10]. Total engagement cost (€32,000 in advisory support + €8,000 in licences + €4,000 in internal structuring) recouped by the 5th month.

Confirmed shift to a stabilised Orchestre

Beyond the figures, the lesson lies elsewhere: INDUSTEC has moved from a situation where AI was a matter of personal curiosity (Artisan) to one where AI is embedded in the commercial process, measured monthly by the executive committee, and extendable (Orchestre). Two further use cases have been scoped for 2026: competitive intelligence, and assistance with drafting client visit reports.

The trajectory towards Architecte is mapped out for 2027-2028, involving the construction of a proprietary technical knowledge base (10 years of cleansed operational archives) and supervised automation of tender responses.

Anonymised case study. Any resemblance to an existing company would be coincidental. The results presented derive from a specific engagement and do not constitute a guarantee of equivalent results for other organisations. The return on investment of an AI transformation depends on the sector context, the company's initial maturity, team commitment and quality of execution. Past performance is not indicative of future results.

Chapter 8: 12-month roadmap

The roadmap below is the standard template for an Artisan → Orchestre engagement. The use-case labels are to be replaced with the specific cases selected during phase 1.

Quarter 1: Diagnostic and foundations

- 360° diagnostic and scoping of the 3 to 5 priority use cases.
- AI usage charter approved by the executive committee.
- Appointment of 2 to 4 internal champions.
- Three-tier training plan approved.
- Initial ISO 42001 elements (policy, minimal register).

Completion indicator: diagnostic report delivered, charter signed, champions appointed and briefed, initial AI register established.

Quarter 2: First pilot deployment

- Sprint 1 on the priority use case (4 to 6 weeks).
- Sprint 2 on the second use case.
- Training of operational staff on the selected tools.
- Rollout of the LLM gateway and the initial FinOps dashboard.
- First monthly executive committee review, including adoption measurement.

Completion indicator: at least one use case in production with its success indicators met; LLM gateway operational; FinOps dashboard in place.

Quarter 3: Industrialisation and expansion

- Sprint 3 and sprint 4 on the following use cases.
- Industrialisation of the Quarter 2 use case (tool integration, support).
- Implementation of the AI Act register in its full format.
- First gains documented at executive committee level, first FinOps optimisations.

Completion indicator: 3 use cases in production, gains measured over 2 consecutive months, measurable FinOps optimisations (-15% to -25% vs. initial trajectory).

Quarter 4: Consolidation and year 2 roadmap

- Optimisation of cases in production.
- Consolidated measurement of benefits over 12 months.
- Quarterly AI committee established and first meeting held.
- Year 2 roadmap approved by the executive committee, including the ISO 42001 trajectory and the Architecte plan.

Completion indicator: engagement review documented, ROI measured, sustainable governance in place, Architecte trajectory scoped for year 2.

Chapter 9: Conclusion & taking action

The window is still open

AI transformation is not a matter of technology. It is a matter of execution discipline. The tools are available, the methods are documented. What is lacking in most SMEs is not resources: it is the framework that turns those resources into results.

The 5-level MATIA Method™ Scale and the 5-phase methodology presented in this white paper are a response to this gap. They are not magic: they do not do away with the need for scoping work, executive committee commitment, or investment in training. But they chart a path that is practicable, measured and repeatable.

The post-Digital Omnibus timetable (Chapter 1.4) offers a window (18 to 24 months) to put governance properly in place before high-risk obligations become binding in late 2027 / mid-2028. SMEs that use this window to move from Artisan to Orchestre, and then from Orchestre to Architecte, will secure a governance and discipline advantage that latecomers will only be able to close in a rush, always costlier, always riskier.

Where is the frontier heading? What those building it are saying

A business leader is entitled to wonder whether all this will not be swept away by the next wave. Here is what those building the frontier, with no guarantee as to timing, **state**, as of June 2026:

- **Demis Hassabis** (Google DeepMind, May 2026) has narrowed his window for human-level AI (“AGI”) to **2029-2030**, from 2030-2035 a year earlier [45].
- **Dario Amodei** (Anthropic, January 2026) believes that “powerful AI” could be just **one or two years away**, “although it could also be considerably further off”, he himself notes [46].
- On the measurement side, the **Stanford AI Index 2026** documents a genuine acceleration: performance on the SWE-bench coding benchmark rose from ~60% to close to 100% of human level in a single year [29].

But the same edition of the AI Index provides the **honest counterweight**: these models, which win gold at the International Mathematical Olympiad, correctly read an **analogue clock only 50.1% of the time** [29]. Capability is “jagged”: superhuman on one task, failing on the next.

What conclusion should a business leader draw? The date of AGI is disputed; its trajectory is not: raw capability is accelerating and becoming commonplace. Yet the more the technology becomes a commodity that anyone can buy, **the more the only lasting advantage becomes organisational**. BCG puts a figure on it: a company with a **clear AI strategy** captures **+25 points** of impact, against only **+5 points** for one that merely buys better tools [47]. The right response to uncertainty is not to wait. It is to be ready: build maturity now, methodically.

What comes next? From diagnosis to industrialisation

This white paper answers the first question: *where do you stand?* The next one follows quickly: *how do you move from the first pilots to daily use, measured and governed?* That is the subject of the second white paper in the series, **“From PoC to Industrialisation: Leading AI Change Management in European SMEs”** (June 2026 edition, FR + EN). It documents the seven frictions of the “last mile”, the five levers of change management and AgentOps, because only 46% of AI POCs reach production [27], and this wall is 70% human, not technological [28].

→ Read white paper no. 2: paulantoinetual.fr/en/white-paper-industrialization

AI Express Audit & Roadmap: the first step

To position your company on the MATIA Method™ Scale and chart a 90-day roadmap, the **AI Express Audit & Roadmap** is designed as a calibrated first step:

- 60 minutes by video conference, with no obligation.
- Positioning of your SME on the 5 levels.
- Identification of 3 priority use cases for your context.
- An initial 90-day roadmap, delivered in writing within the week.

It is this first step, short but committing, that often makes the difference between projects that get off the ground and those that remain at the intention stage. The approach complies with the **Osez l'IA** framework (Bpifrance) [14] and may form part of a public funding pathway.

→ Book an AI Express Audit & Roadmap: junyr.fr/diagnostic

SOURCES: verifiable, June 2026

[1] **Bpifrance Le Lab / Rexecode**, biannual business-climate barometers for micro-businesses and SMEs. 82nd biannual barometer (January 2026): 55% reported AI use at end of 2025; Q2 2026 barometer (fieldwork 13-26 April 2026, 1,135 respondents, published 19 May 2026): **58% reported use, of which almost half daily**; 43% of business leaders report a productivity gain; leading barrier = difficulty identifying relevant use cases (~54%). <https://presse.bpifrance.fr> : dominant uses, growth over 18 months.

[2] **KPMG**, *Global AI Pulse, Q1 2026* (survey 17 February to 17 March 2026, 2,110 business leaders, 20 countries), 31 March 2026. <https://kpmg.com/xx/en/media/press-releases/2026/03/kpmg-global-ai-pulse-survey.html> : **11%** of organisations are “AI leaders”; **82%** of leaders report significant business value versus **62%** of others; **32%** are deploying or scaling AI agents. Source of the “~11%” share of organisations deriving measured value from AI used in this white paper.

[3] **Bpifrance Le Lab**, *AI in French SMEs and mid-caps* (fieldwork survey, October-December 2024, 1,209 business leaders): seen as an “existential challenge”, no formalised AI strategy in place, data not analysed. Supplemented by **INSEE Première n°2061, ICT in businesses in 2024** (July 2025): 10% of businesses with 10 or more employees used AI in 2024. <https://www.insee.fr/fr/statistiques/8604126>

[4] **Gartner**, *Hype Cycle for Agentic AI* (April 2026): at least 50% of generative AI projects abandoned after the pilot phase; **more than 40% of agentic AI projects abandoned by 2027** (costs, unclear value, risk control). *AI Maturity Model Toolkit* (5-level model), 2025-2026. <https://www.gartner.com/en/articles/hype-cycle-for-agentic-ai>

[5] **HCLTech**, *AI Impact Imperatives 2026* (20 May 2026, survey of 467 business leaders at companies with more than one billion dollars in revenue): **43% of major AI initiatives are judged likely to fail**, mainly due to an organi-

sational “execution gap” rather than a technological one. <https://www.hcltech.com/press-releases/hcltech-report-warns-43-enterprise-ai-initiatives-may-fail-leaders-face-shrinking>

[6] **McKinsey**, *The State of AI Trust in 2026* (25 March 2026, fieldwork Dec. 2025 to Jan. 2026, ~500 organisations): a 5-dimension AI trust maturity model; average score of **2.3 out of 4**, with **only one organisation in three at level 3 or above**; senior management oversight of AI ranks among the factors most strongly correlated with bottom-line impact. <https://www.mckinsey.com/capabilities/quantumblack/our-insights>

[7] **Denis Atlan**, *AI & ROI Barometer for French SMEs 2022-2025*, 2026. An analysis of more than 200 B2B AI deployments in French SMEs: median ROI of 159.8% measured over 24 months for well-scoped projects, with a 73% success rate. <https://www.denisatlan.fr/barometre-ia-pme> (DOI: 10.5281/zenodo.17795133).

[8] **Croissance et Transitions**, consolidated field feedback: more than 30 documented AI engagements in French SMEs and mid-caps since 2024; observed budgets, scoping-to-deployment ratios, recurring causes of failure, competitive replication timeframes.

[9] **Regulation (EU) 2024/1689** of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (the *Artificial Intelligence Act*). <https://eur-lex.europa.eu/eli/reg/2024/1689/oj> : entered into force 1 August 2024, phased application 2025-2027, timeline updated by the Digital Omnibus [15].

[10] **INDUSTEC case study** (anonymised name): MATIA Method™ engagement, April 2025 to January 2026, figures consolidated over the last 6 months of live production use (July 2025 to December 2025). Data drawn from the client’s internal reporting, validated by the monthly executive committee.

[11] **ADP Research**, *People at Work 2026* (nearly 40,000 workers, 36 countries; France component n = 1,051, released June 2026): in France, **11%** of employees use AI almost daily, 24% several times a week, **29% never**, compared with 20% near-daily use globally. <https://www.adpresearch.com/research/ai-powers-into-the-workplace>

[12] **DGE / France Num**, *Baromètre France Num 2025*, 6th edition, September 2025, 11,021 businesses. <https://www.francenum.gouv.fr/guides-et-conseils/strategie-numerique/comprendre-le-numerique/barometre-france-num-2025-le> : 26% of micro-businesses and SMEs used at least one AI tool in 2025 (×2 versus 2024); by size: 23% (1-4), 31% (10-19), 35% (20-49), 42% (50-249).

[13] **Microsoft**, *Work Trend Index 2026: Annual Report*, 5 May 2026. <https://www.microsoft.com/en-us/worklab/work-trend-index> : “Frontier Firms” (19% of users at Frontier organisations, 16% “Frontier Professionals” orchestrating agents), “agent boss” / human agency, the gap between business leaders and employees on agents (67% versus 40%), Owned Intelligence.

[14] **Bpifrance**, *Plan Osez l’IA*, July 2025; *Diag Data IA*. <https://www.bpifrance.fr/catalogue-offres/diag-data-ia> ; <https://bpifrance-creation.fr/entrepreneur/actualites/lancement-du-plan-national-osez-lia> : initial envelope of €200 million, overall momentum of €10 billion, targeting 80% of SMEs and mid-caps at Orchestre+ level by 2030. Since 1 January 2026, the Diag Data IA has been a flat-fee diagnostic of €10,000 excluding VAT, subsidised at 25% (mid-caps not eligible); the IA Booster funds up to 80%.

[15] **Council of the European Union**, *Digital Omnibus: political agreement on simplifying digital rules*, press release of 7 May 2026. <https://www.consilium.europa.eu/en/press/press-releases/2026/05/07/> : postponement of high-risk AI Act requirements to end of 2027 / 2 August 2028, a -35% reduction in the administrative burden on SMEs by 2029.

- [16] **Gartner**, *AI Maturity Model: 5 levels* (Awareness, Active, Operational, Systemic, Transformational), 2025-2026. <https://www.gartner.com/en/chief-information-officer/research/ai-maturity-model-toolkit>
- [17] **PwC**, *The AI-Native Enterprise: 5-level maturity model*, 2026. <https://www.pwc.com.au/services/artificial-intelligence/the-ai-native-enterprise.html>
- [18] **Microsoft**, *Agentic AI Adoption Maturity Model (Copilot Studio): L100 to L500*, 2025-2026. <https://learn.microsoft.com/en-us/microsoft-copilot-studio/guidance/maturity-model-overview>
- [19] **BCG**, *AI Radar 2026: As AI Investments Surge, CEOs Take the Lead* (15 January 2026, 2,360 business leaders including 640 CEOs, 16 markets). <https://www.bcg.com/publications/2026/as-ai-investments-surge-ceos-take-the-lead> : leadership archetypes **Trailblazers ~15% / Pragmatists ~70% / Followers ~15%**; 72% of CEOs describe themselves as the primary AI decision-maker (×2 in one year); > 30% of the AI budget allocated to agentic AI.
- [20] **Deloitte**, *State of AI in the Enterprise 2026: The Untapped Edge* (21 January 2026, 3,235 business leaders, 24 countries). <https://www.deloitte.com/us/en/about/press-room/state-of-ai-report-2026.html> : **only 30%** of organisations are redesigning their key processes around AI, **37%** use it superficially, ~25% report a transformative effect; 23% use agentic AI at least moderately, ~74% plan to do so within two years, **only 21%** have mature agent governance.
- [21] **ISO/IEC 42001:2023**, *Artificial Intelligence Management Systems: Requirements*, ISO, December 2023. <https://www.iso.org/standard/81230.html> : the first certifiable international AIMS standard.
- [22] **CLOUD Act** (United States, 2018), §2713: *Clarifying Lawful Overseas Use of Data Act*. <https://www.congress.gov/bill/115th-congress/house-bill/4943>
- [23] **ANSSI**, *Guidance on migration to post-quantum cryptography*, first published 2022, updated in 2024 and again in 2026 (certifications without PQC ending in 2027). <https://cyber.gouv.fr/> : national timeline 2027 / 2030. The 2035 deadline falls under the **EU's coordinated PQC roadmap** (NIS Cooperation Group, June 2025).
- [24] **Directive (EU) 2022/2555 (NIS2)** of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of cybersecurity across the Union. <https://eur-lex.europa.eu/eli/dir/2022/2555/oj>
- [25] **Regulation (EU) 2023/2854 (Data Act)** of 13 December 2023 on harmonised rules on fair access to and use of data. <https://eur-lex.europa.eu/eli/reg/2023/2854/oj> : main provisions applicable since **12 September 2025**.
- [26] **Croissance et Transitions / MATIA Method™**, article *The end of prompt engineering: from prompt to context engineering*, May 2026 (28 sources). <https://paulantoinetual.fr/en/blog/la-fin-du-prompt-engineering-pourquoi>
- [27] **IDC / Lenovo**, *CIO Playbook 2026: The Race for Enterprise AI* (January 2026, n = 800, Europe & Middle East): **only 46% of POCs reach production**; 94% of organisations anticipate a positive ROI, with **\$2.78 of expected value per dollar invested**. <https://news.lenovo.com/pressroom/press-releases/research-reveals-ai-is-paying-off-cios-arent-ready/>
- [28] **BCG**, *AI Transformation Is a Workforce Transformation* (2026). Reaffirming the **10-20-70** rule: 10% algorithms, 20% technology and data, **70% redesigning the human factor and processes**. <https://www.bcg.com/publications/2026/ai-transformation-is-a-workforce-transformation>
- [29] **Stanford HAI**, *The 2026 AI Index Report*, 20 April 2026: 88% of organisations use AI in ≥1 function (versus 78% the year before), <10% at scale; SWE-bench ~60% → ~100% of human-level performance in one year;

“jagged” capability (gold at the IMO, yet analogue clocks read correctly only 50.1% of the time). <https://hai.stanford.edu/ai-index/2026-ai-index-report>

[30] **OECD**, *AI adoption by small and medium-sized enterprises* (December 2025): AI adoption across the OECD from 5.6% (2020) to 14% (2024); only 29% of SMEs using generative AI apply it to their core business. See also *Compendium of Productivity Indicators 2025* (productivity gap between SMEs and large enterprises). <https://www.oecd.org/>

[31] **Mistral AI**, *Mistral Large 3* (2 December 2025, MoE 675B/41B, Apache 2.0 licence): the most recent frontier open-weight model from the French AI lab. <https://mistral.ai/news/mistral-3/> ; since then: *Mistral Medium 3.5* (30 April 2026, 128B, open-weight, modified MIT licence, mid-range tier).

[32] **Google DeepMind**, *Gemma 4* (2 April 2026, Apache 2.0 licence): unquantised weights fitting on a single 80 GB GPU; quantised versions runnable offline on a phone, Raspberry Pi or NVIDIA Jetson. <https://blog.google/innovation-and-ai/technology/developers-tools/gemma-4/>

[33] **Artificial Analysis**, *Intelligence Index*: as of 30 April 2026, the best open models (Kimi K2.6, MiMo V2.5 Pro) scored **54** against **60** for GPT-5.5 (a 6-point gap), versus ~22 a year earlier; on 16 June 2026, **GLM-5.2** (Z.ai, MIT) became the leading open-weight model on the index, matching GPT-5.5 on the GDPval-AA agentic sub-benchmark. Factual recall (SimpleQA) still lags behind for open models. <https://artificialanalysis.ai/articles/glm-5-2-is-the-new-leading-open-weights-model-on-the-artificial-analysis-intelligence-index>

[34] **Sesame Disk**, *Quantization Techniques for AI Inference in 2026* (14 May 2026). Quality retained versus BF16: ~99% (FP8), ~97-98% (INT8), ~95% (INT4 AWQ/GPTQ); 4-bit degradation measured at -1.2 to -2.5% (MMLU/HumanEval/MATH). <https://sesamedisk.com/quantization-techniques-ai-inference-2026/>

[35] **VRLA Tech**, *LLM Quantization Explained: INT4, INT8, FP8, AWQ and GPTQ in 2026* (April 2026): quality/memory scale by format (orders of magnitude). <https://vrlatech.com/llm-quantization-explained-int4-int8-fp8-awq-and-gptq-in-2026/>

[36] **Epoch AI**, *AI supercomputers performance share by country* (May 2025, latest available measurement): US ~74.5% / China ~14.1% / EU ~4.8%. <https://epoch.ai/data-insights/ai-supercomputers-performance-share-by-country> ; concentration confirmed by **CNAS**, *Sovereign AI Index* (April 2026): the United States and China hold ~90% of frontier computing power. <https://interactives.cnas.org/reports/sovereign-ai-index/>

[37] **EU Perspectives**, *Europe bets on EuroStack* (February 2026): US hyperscalers control >85% of the EU cloud market; EuroStack ~€300 billion; EU sovereign cloud ×3 → \$23 billion by 2027 (Gartner). <https://euperspectives.eu/2026/02/europe-bets-on-eurostack/>

[38] **OVHcloud**, *AI Endpoints* (40+ open models, zero data retention, EU hosting) <https://www.ovhcloud.com/fr/public-cloud/ai-endpoints/> ; **Scaleway**, *Generative APIs* (H100 clusters, OpenAI-compatible, EU data) <https://www.scaleway.com/fr/generative-apis/>

[39] **Maddyness**, *Mistral AI targets one gigawatt of computing capacity by 2030* (28 May 2026): 200 MW in 2027, 1 GW in 2030, €4 billion; sovereign data centre at Bruyères-le-Châtel (\$830 million debt financing round led by Bpifrance, March 2026). <https://www.maddyness.com/2026/05/28/mistral-ai-vise-une-capacite-de-calcul-dun-gigawatt-a-lhorizon-2030/>

[40] **European Commission**, *InvestAI / AI Gigafactories*: ~€200 billion towards 2030, of which €20 billion for 4-5 gigafactories. <https://digital-strategy.ec.europa.eu/>

[41] **Gundlach, Lynch, Mertens & Thompson**, *The Price of Progress: Price Performance and the Future of AI* (arXiv:2511.23455, revised version of 23 March 2026): inference cost at constant performance divided by **~5 to 10× per year** (up to 31×/year for the highest-performing segment); algorithmic efficiency **~3×/year**. <https://arxiv.org/abs/2511.23455>

[42] **Ramp**, *The cost of AI tokens for businesses* (8 June 2026): between January 2025 and April 2026, token consumption **+1,001%** versus spend **+497%**; average observed price **\$0.72/M tokens** (as low as \$0.07 for lightweight models). <https://ramp.com/blog/ai-token-cost-for-businesses> ; see also **TechCrunch**, *The token bill comes due* (5 June 2026): Uber exhausted its 2026 AI budget as early as April; an agentic task consumes 5-30× more tokens (Gartner). <https://techcrunch.com/2026/06/05/the-token-bill-comes-due-inside-the-industry-scramble-to-manage-ai-runaway-costs/>

[43] **DeepSeek**, API pricing (2026): DeepSeek V4-Flash at **~\$0.14 / \$0.28 per million tokens** (input/output), MIT licence. https://api-docs.deepseek.com/quick_start/pricing

[44] **InfoWorld**, *Small language models: rethinking enterprise AI architecture* (4 May 2026, citing NVIDIA, Gartner and Info-Tech): routing to SLMs reduces inference cost **by up to 90%** on high-volume repetitive tasks; NVIDIA now describes SLMs as the “future of agentic AI”. <https://www.infoworld.com/article/4160404/small-language-models-rethinking-enterprise-ai-architecture.html>

[45] **Demis Hassabis** (Google DeepMind), AGI window narrowed to 2029-2030 (statements made May 2026, Google I/O / Axios interview).

[46] **Dario Amodei** (Anthropic), *The Adolescence of Technology* (January 2026): “powerful AI” possible within 1-2 years, with an explicit caveat. <https://darioamodei.com/essay/the-adolescence-of-technology>

[47] **BCG**, *AI at Work 2026: Why Strategy Matters More Than Tools* (3 June 2026, 11,749 employees, 14 markets): a clear AI strategy = +25 points of impact, against ~+5 points for better tools alone; 74% of frontline white-collar workers are regular users. <https://www.bcg.com/publications/2026/ai-at-work-why-strategy-matters-more-than-tools>

[48] **Gartner**, *Worldwide Sovereign Cloud IaaS Spending* (press release of 9 February 2026): \$80 billion in 2026 (+35.6%); Europe from \$6.9 to \$12.6 billion (+83%), overtaking North America in 2027. <https://www.gartner.com/en/newsroom/press-releases/2026-02-09-gartner-says-worldwide-sovereign-cloud-iaas-spending-will-total-us-dollars-80-billion-in-2026>

[49] **AION** (Ardian, Artefact, Bull, EDF, Capgemini, iliad, Orange, Scaleway), bid for a French AI *Gigafactory* under the InvestAI initiative (5 June 2026). <https://www.storagenewsletter.com/2026/06/05/>

*White paper “AI Maturity of French SMEs 2025-2026: MATIA Method™”, by Paul-Antoine TUAL, AI Transformation Leader · Croissance et Transitions. Initial edition May 2026 (golden standard, v2.0). Published: 23 May 2026. Revision v2.1: 12 June 2026. Revision v2.2: 22 June 2026 (deep research: correction of attributions, enrichment of sovereignty & FinOps content, AGI horizon section). Last revision: 23 June 2026. **June 2026 Edition, v2.3:** strict freshness pass (all AI market sources older than six months replaced with their 2026 equivalents: KPMG Global AI Pulse Q1 2026, HCLTech, Deloitte / Gartner / BCG 2026, IDC-Lenovo CIO Playbook 2026, ADP People at Work 2026, the open-weight and FinOps landscape as of spring 2026; regulations and foundational evidence retained, dated). Next revision scheduled: November 2026.*

Master document: maintained in Markdown (`content/livre-blanc-maturite-ia-pme-EN.md`), generated as canonical HTML at paulantoinetual.fr/en/white-paper-maturity and distributed as PDF.