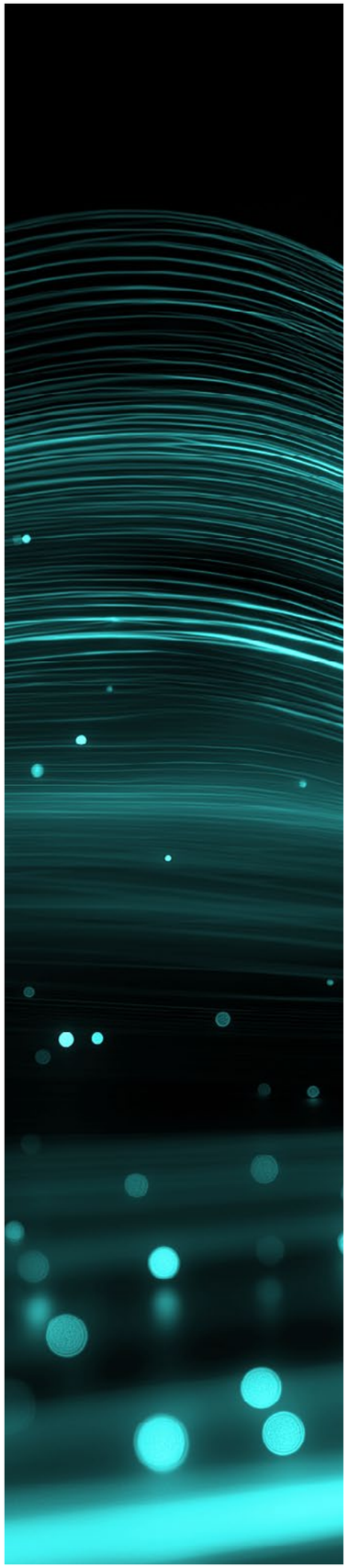
zenitech

Agentic AI

The rise and future of autonomous decision-makers

Dr. Bertalan Forstner
Emerging Technologies Strategist
bertalan.forstner@zenitech.co.uk

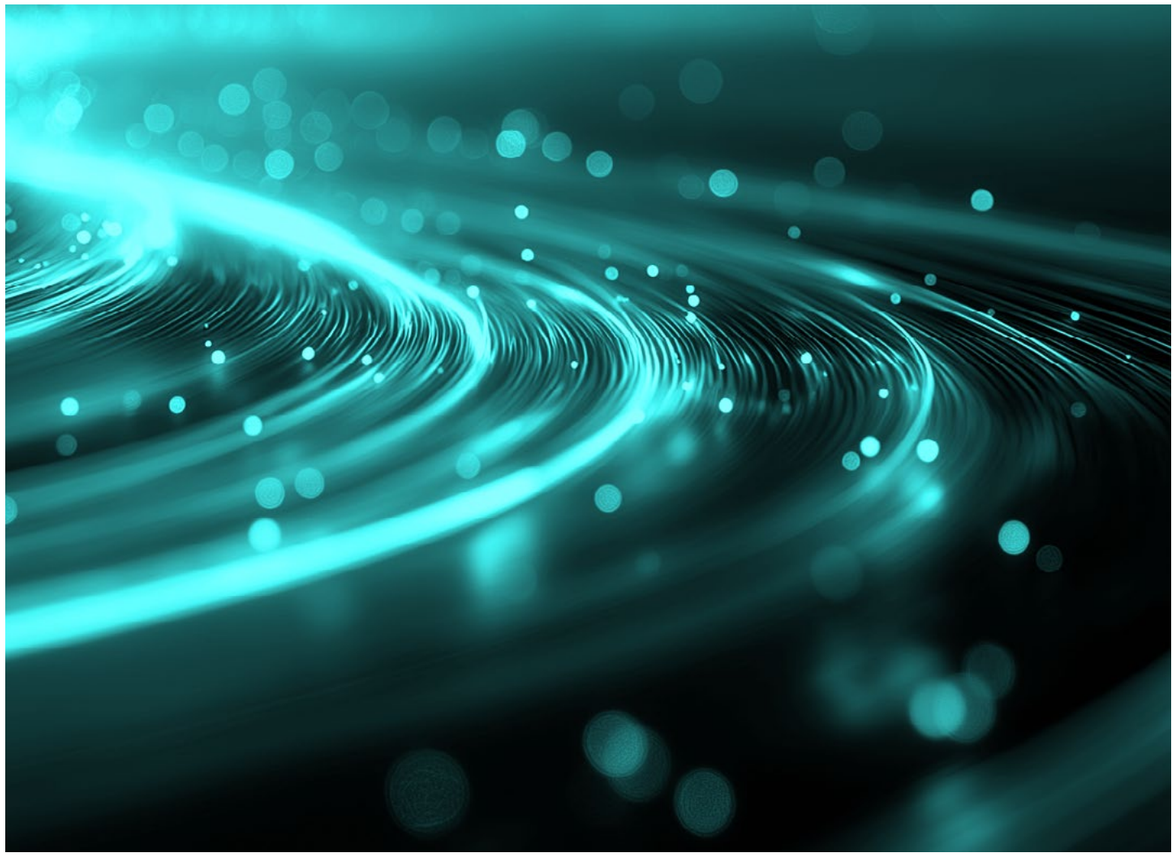
October 2025



CONTENTS

In recent years, many people have judged artificial intelligence by the “chatbot”. A polite conversation partner, someone who offers quick answers, produces summaries, helps a bit with searching and simple tasks. By 2024/2025, however, a new era has developed. Agentic AI has entered the stage, where we no longer think in terms of dialogue experience, but in autonomous digital actors that structure themselves around goals, use tools, prepare decisions and execute them. Software no longer asks the user, but undertakes the task itself.

At Zenitech we placed this shift at the core of our forward-looking advisory programme, and through our collaboration with the Budapest University of Technology and Economics we are strengthening research in this direction, because this is where the next decade of enterprise systems will be built.



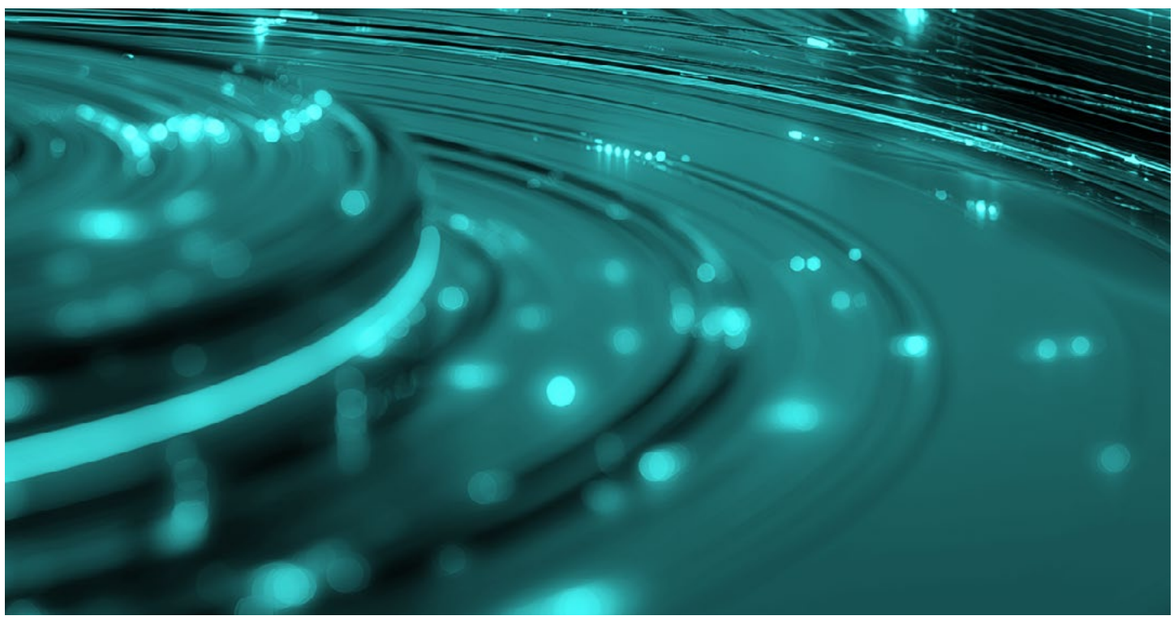
How we arrived here, and what was missing

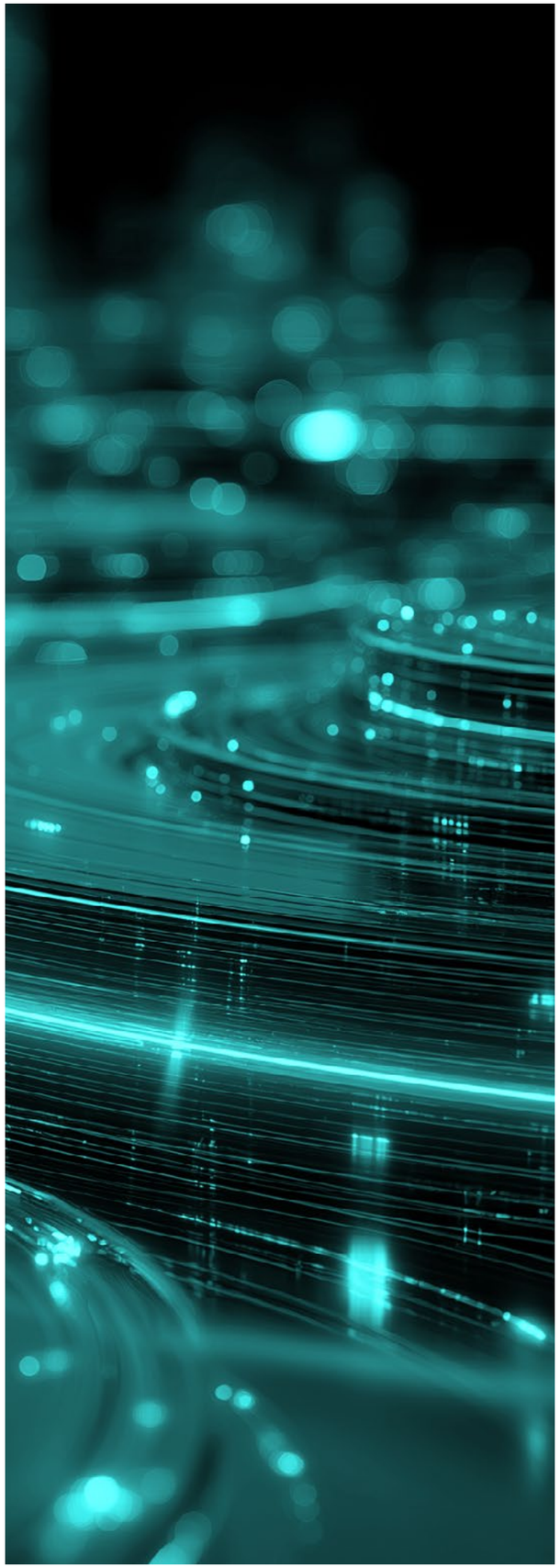
In the beginning our focus on “prompt tricks” was not because we believed they were the end goal, but because they were the quickest way to explore how far language models could be pushed into reasoning or structured tasks. We wanted to see if simple phrasing and clever instructions could stretch their capabilities toward acting more like problem-solvers than autocomplete machines. Then came the visible reasoning chains, experiments in structured thought, and the first “I think, then I act” patterns. AutoGPT and BabyAGI demonstrated impressively that software can break down goals into tasks and invoke tools, but in reality they often became stuck, fell into loops, or burned through cost budgets.

The real turn started with protocol-thinking. The Model Context Protocol (MCP) and similar initiatives finally unified the way an agent communicates with its environment. Around this grew agent frameworks, the newer iterations of function-calling APIs, and the realisation that an “agent” is not a single model but an entire architecture.

What is missing here, is that early “prompt magic” was only ever a bridge. It proved useful for enthusiasts, but enterprises quickly saw that scaling requires structure, not tricks. The missing element was discipline: protocols, governance, and architectures that could carry value into production.

Meanwhile, market data shows that enterprises are no longer just experimenting. According to McKinsey’s 2025 survey, 78% of respondents reported that their organisation already uses AI. The same material also shows that less than one third follow most of the practices required for broad adoption and scaling, and less than one fifth measure on KPI-level the outcomes of AI solutions. This duality is both promising and cautionary. Usage spreads, yet methodical discipline is not universal.





The agent as an enterprise component

A viable agentic system is not one big box. It is better understood as five-layered worlds.

- The first is the linguistic-reasoning component, which perceives and breaks down goals into steps, handles uncertainty and prepares decision alternatives.
- The second is memory, where short and long-term context coexist, with vector-based retrieval and curated logs of important moments.
- The third is the layer of tool usage, integrating APIs, databases, code generation and execution, and internal services.
- The fourth is the protocol layer, where the agent connects through standardised interfaces rather than hardwired cables. This is where MCP and other interoperability schemes come in.
- The fifth is governance, where permissions, audit, reversibility and approval chains are orchestrated. This is the point where technology becomes business, because in an enterprise environment it is not enough for the agent to be clever. It becomes usable when it fulfils compliance and risk expectations without creating wide-impact incidents.

Zenitech considers this system view essential. An agent is not a pilot for the sake of marketing departments. An agent is a product, just as a microservice is. We version it, approve changes, ensure rollback, and measure how much value it creates. That is how a “wow” demo turns into operational capability.

Why the moment has matured now

Companies already see the tangible benefits, and this accelerates decisions. In PwC's 28th Annual Global CEO Survey, more than half of executives reported productivity gains in employee time use due to GenerativeAI (GenAI) investments in the past 12 months, and around one third also experienced revenue or profit increases. Nearly half expect further profit increases from GenAI in the coming year. The same document notes that many leaders would integrate AI into technology platforms and business processes, while workforce and skills strategies lag behind. We consider this a risk, because the value of agents ultimately depends on how teams use them.

Deloitte's 2025 early-year survey (Q4 2024) shows that 78 percent of organisations plan to increase spending on GenAI in the next year. Respondents reported the most visible returns in time saving and efficiency, quality improvement, and cost reduction, and many are already exploring how these benefits can be extended through the agentic approach. This indicates clearly that agent-based operations have moved from experimentation into the phase of planned adoption.

Industrial scenes where it works already

In customer service, for instance, we do not employ a conversationalist but an actor who identifies the issue, checks the history, verifies contract status, creates the necessary records in the CRM, and closes the case. At the end the client receives a resolution summary, not another question. This does not replace the human agents, it takes monotonous moves away from them and frees time for the more complex cases.

In DevOps supervision the agent looks into logs, detects patterns, builds a hypothesis, reproduces the error in isolation and prepares a repair plan. If the risk is low, it performs the rollback, logging the effects along the way. The outcome is a faster recovery, and each incident becomes a learning object for the future.

In R&D several agents collaborate: one identifies sources, another analyses data quality and methodological compliance, a third coordinates experimental design and documentation. Each step is logged traceably, so when leaders give the green light, what launches is not a black box but a transparent experiment.

In financial planning the agent fills spreadsheets, integrates data from multiple sources, invokes statistical or ML models, builds scenarios and proposes decision options based on cost-risk profiles. Gradually monthly closing transforms into continuous financial capability, where the system keeps working while the human team addresses strategic themes.

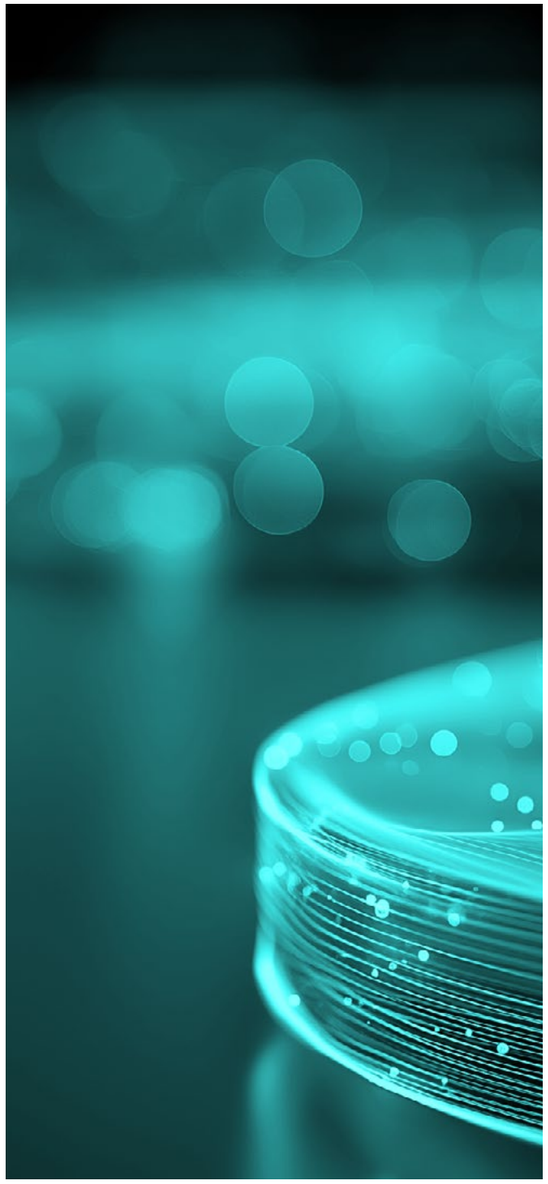
Four ideas, soon in need of standards

The real advance will come from four less-discussed areas, which are still being planned and tested in labs or are in early stage pilots. Zenitech, with our university partners, are involved at both stages. Without these less discussed areas, enterprise AI will remain like a marble statue in the hall: beautiful but untouchable.

The first is **capability-leasing** with temporary permissions. Today most agents access systems through fixed API keys. Once granted, they will remain unless they are manually revoked. A much healthier pattern is for a “capability broker” to issue a task-bound key that expires automatically. Permission for a change in ERP might last 20 minutes, apply to a narrow record set and defined functions, and be deactivated once the task ends. This principle of least privilege is not extra caution but a baseline of acceptable risk. Without it, every incident could be broad and unauditable.

The second is the **transactive agent** that never executes without dry-run and reversibility. A corporate agent should first calculate what will change, then commit only if policy deems the result safe. If not, it seeks approval and creates a change log from which complete rollback is possible. Such designs are being piloted, including undo-logs spanning multiple systems and dry-run outputs formatted for human readability, so approvers see consequences not code.

The third is **market-based multi-agent orchestration.** Today’s multi-agent systems, often allocate by fixed roles. We prefer reviving the contract-net view, closer to enterprise reality. Specialist agents bid for sub-tasks, quoting expected accuracy, time, cost and risk. The coordinator selects the “price,” which includes latency, energy, model fees, and human approval requirements. With this formula



efficiency grows and rare but expert agents are called only when worthwhile.

The fourth **is context-inventory and signed receipts.** AI decisions are often black boxes, but enterprises need traceability: which document snippets, which tool calls, which model versions produced the outcome. If listed and saved as digitally signed inference receipts, the entire decision path becomes auditable. This matters strongly in regulated sectors, where acceptance or rejection of AI depends on it.

What we risk if the four ideas are not adopted

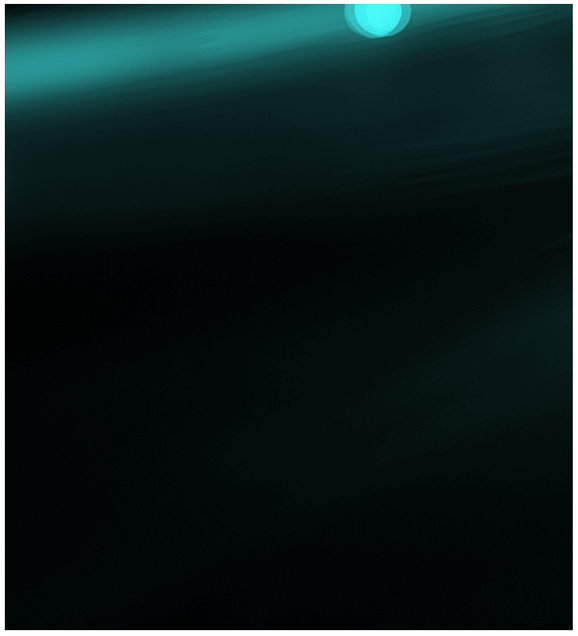
If capability-leasing is absent, accesses decay and leak. Permissions tie to persons or service accounts, enlarging the blast radius. Without transactive execution, agents risk irreversible states or silent inconsistencies. Without market-based organisation, systems become either over-centralised or spent wrongly. Without context inventory and signed receipts, audits end in shrugged shoulders: “the model said so”. Reputation and time are lost, and high-value use cases never receive approval.

PwC’s CEO survey also shows trust remains unsettled. Many leaders only moderately trust AI in core processes. Trust is not built by declarations but by engineering practices: dry-run, rollback, temporary credentials and traceable context.

How a team should start an agent that survives the pilot

Successful adoption does not depend only on the first use case, but on how we define returns, how we integrate the agent into the business rhythm, and how disciplined we are with controls and audits. Teams must think from day one in a triangle of cost-time-risk. The agent should not run on test data but on real data, but in monitored environments where rollback and logging are live. It should not “run away”, but have a state-machine. This way, unexpected situations trigger controlled reset instead of compulsive loops.

These steps matter early, because true operational value emerges there. McKinsey’s latest report shows more companies perceive unit-level revenue and cost effects from GenAI, but few yet feel impact on enterprise EBIT. That suggests the bottleneck is not the model but the workflow and the discipline of introduction. Where KPI measurement, workflow redesign and leadership engagement are present, the impact grows likelier.



Another angle rarely considered

Many speak of agents as mere automata that code better or close tickets faster. We prefer the metaphor of the digital colleague who finds its own economics. In contract-net view tasks are not handed down by a manager, but assigned through a local market. Agents use an internal currency where cost includes waiting, reliability and auditability. The version controller is not the final bottleneck, but the actor who maintains market rules. Cooperation becomes flexible bargaining under supervision, not rigid choreography.

Think about the software supply chain. Our teams already version, cache, reuse. We argue the same is needed for agents. They should not rediscover every task from zero. Successful workflows become plan-reuse libraries with metrics, context and risk tags. Without this, every agent project becomes too expensive and scaling derails over time.

Financial reality: objectives that are not single-dimension

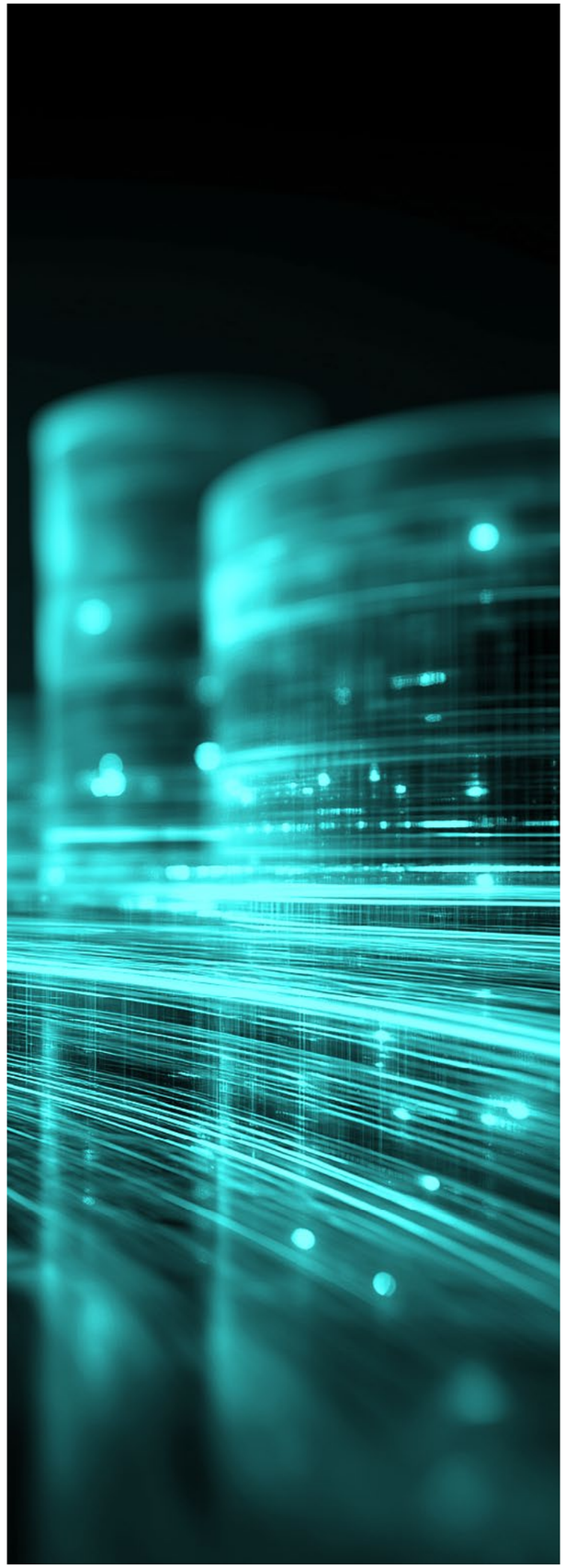
Enterprise AI often optimises along one metric: accuracy, latency or cost. For agents this creates bad incentives. We need objective functions that jointly handle cost, delay, risk and expected value, seeking Pareto efficiency. A service agent may choose not to finish a call with the biggest model, because a medium one is only slightly worse, yet faster and cheaper. The difference is determined by context and policy. Thus we avoid hidden cost-spirals, and engineers gain a language to explain why the agent chose as it did.

Deloitte's year-end survey confirms cost and quality move together. Many respondents already saw quantified time savings, error reduction and tangible cost benefits. That signals that agents are not only "big model = big bill", but that balanced portfolios bring the strongest enterprise results.

The story we tell teams

Imagine a new colleague. On day one, they do not get unlimited access to every door. They receive a workstation, temporary credentials for their task, and a mentor who observes their first steps. Agents are the same. In the first week they observe and act only where state can be restored. In the second they attempt automated approvals, yet still ask at critical junctures. By the third, "ask or act" is no longer binary. Policy tunes itself to organisational rhythm, and metrics reflect less the demo and more the repeatable value.

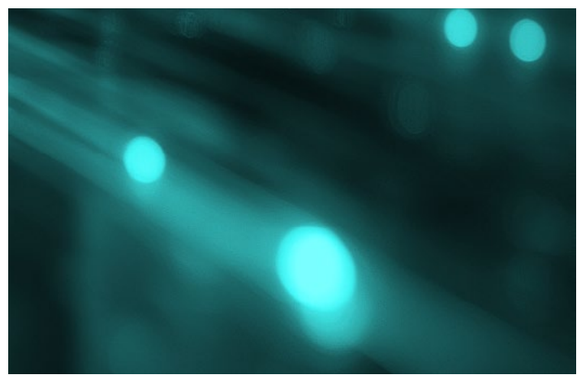
When Zenitech launches a new project, in the first week we rarely talk about models. We start with domain, process, and where the team feels friction. From there we trace back: which tools to admit into the protocol layer, where human approvals sit, where temporary keys are placed, which steps require dry-run. Control is not a barrier but the safety belt of speed.



Cost-conscious agent building that can start today

Enterprises have two options on how they can begin. The first is to introduce cost-aware objectives from the pilot onwards: every agentic step should have a short decision note explaining why that model, that tool, that context was chosen. Very soon you see where cheaper, faster or more reliable choices exist for the same outcome. The second is plan reuse: store successful runs in normalised form with metrics, context and risk tags; otherwise every new agent is like reinventing the wheel again, and budgets can explode.

To set up an agent for your team, treat it as launching a new service. Start with a business-critical, but controlled process where impact is measurable. Build in policies for dry-run and rollback, and use temporary capability-leasing. Record decision contexts with signed receipts, add training and communication, because as PwC's data also shows, many organisations lose momentum by neglecting workforce and skill strategy at the first stage. The McKinsey figures send a clear message: enterprise impact grows where workflows are truly reshaped and KPIs are measured consistently. Deloitte shows that investments do not only bring hope but already tangible profit, provided there is discipline and method in the rollout.



Closing: Why you should work with us

At Zenitech we keep both engineering discipline and future-seeking curiosity in focus. That is why we experiment with solutions we believe will soon become unavoidable. These are not all ready-made products. But organisations that start embedding these ideas into their planning today will find it far easier tomorrow to secure approval for initiatives with the highest business value.

If you want your next agentic project to survive the demo-room and stay in operation, send us your challenge. We will understand the process, assess the risk, and model how it turns into sustainable, measurable value.

Connect with us

www.zenitech.co.uk

info@zenitech.co.uk

linkedin.com/company/zenitechteam

facebook.com/zenitechteam

instagram.com/zenitech