Mesh-AI

The Agentic AI Handbook: Everything an Enterprise Leader Needs to Know to Become an Early Adopter

Table of Contents

Introduction

The financial services sector is no stranger to harnessing technology to move faster. But faced with economic uncertainty, rising regulatory pressure, and growing customer expectations, institutions are being challenged to evolve and grow in different ways. Core operations and processes have become more complex and interconnected, demanding a new approach.

Against this backdrop, AI, machine learning, and now Agentic AI are emerging not just as enablers of efficiency, but as strategic levers for growth, innovation, and competitive differentiation. These technologies offer financial organisations the tools to build smarter products, make better decisions, and navigate volatility with confidence.

But this transformation isn't simply about plugging in new tools. It demands a fundamental rethink of how data is governed, how teams are structured, the skills you prioritise in new hires and training, and how technology integrates into these core operations. Financial institutions must dismantle legacy silos, adopt new operating models, and invest in modern data foundations that support intelligence at scale.

This ebook explores how data and AI are shaping the future of financial services, not only to optimise existing processes, but to utilise new business models, improve risk mitigation, and deliver deeper customer value. Along the way, we'll share key principles, frameworks, and real-world examples that can guide leaders toward meaningful, production-grade transformation.

Why Agentic AI Matters Now

AI is evolving rapidly, and one of the most significant advancements in this space is Agentic AI. Unlike traditional AI models, agentic systems are designed to act autonomously, make decisions, and deliver outcomes with minimal human intervention. These intelligent agents are poised to revolutionise business operations, driving efficiency, enhancing customer experiences, and improving decision-making at scale.

For data and business leaders, the challenge lies in understanding how to harness this technology effectively. While the concept is compelling, the journey from curiosity to implementation requires clarity on use cases, value generation, and strategic adoption.

Defining Agentic AI

Agentic AI refers to autonomous AI systems designed to pursue objectives through adaptive decision-making, real-time data insights, and proactive action. Agentic systems are:

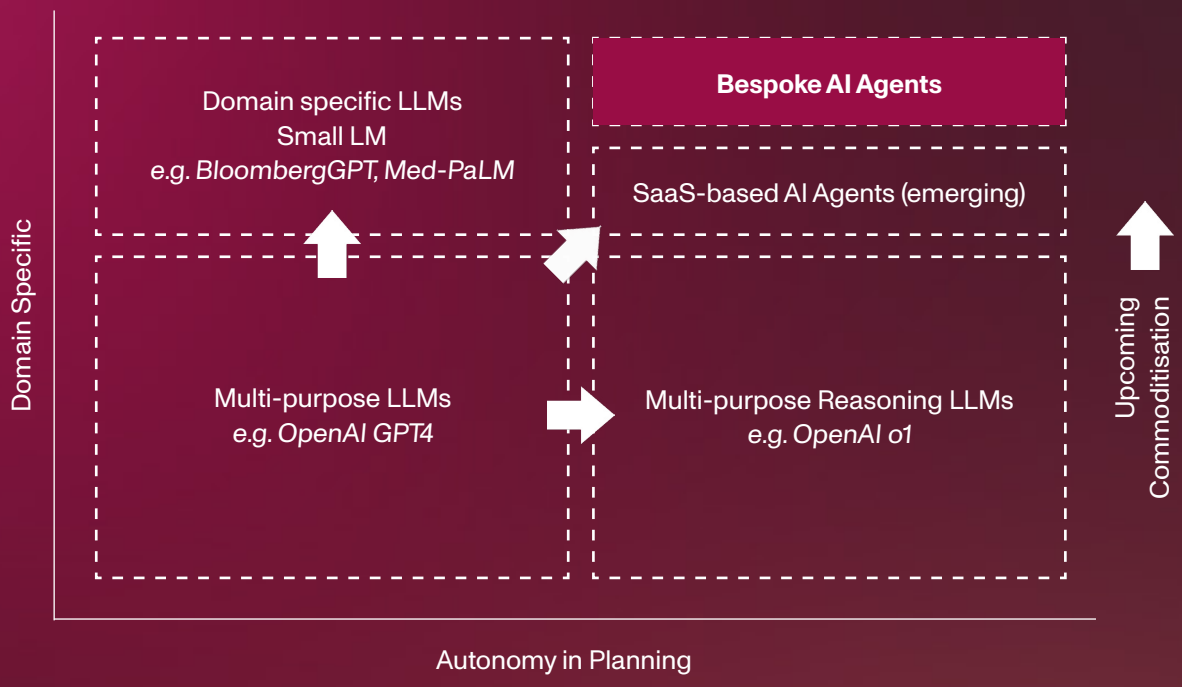
- **Autonomous:** Operating independently while adapting to changes in their environment.
- **Reasoning-Driven:** Capable of evaluating multiple options and choosing optimal solutions.
- **Outcome-Oriented:** Designed to achieve defined goals, often with minimal human oversight.

The advent of the reasoning model marks an evolutionary step in capability. Now LLMs can plan a sequence of tasks geared towards achieving a goal - and execute them. In contrast, over the last several years developers and AI engineers have focused on encouraging reasoning through other (more complex) techniques such as mixture of experts (MoE), retrieval augmented generation (RAG) and extended context windows.

A quick refresher. LLMs have two major vectors of advancement - domain specificity and task planning autonomy. Most enterprises will need some kind of balance between the two to maximise value and take advantage of the latest and greatest advancements. AI agents occupy that space.

The below shows where the opportunity for LLMs lies for the enterprise - where bespoke AI agents have the highest level of domain specificity and autonomy. Moreover, the use cases that actually achieve ROI tend not to be customer-facing chat interfaces for consumers. Rather they address highly specific problem statements that have some facet of repeatability - something that is a signal of a scaling challenge.

Building Cognitive Architectures



By embracing Agentic AI, enterprises can unlock powerful capabilities that enable new approaches to existing challenges, new routes to additional revenue and competitive advantage. The route to value has a number of levers that can be pulled on - but the north star is a bespoke AI Agent system. While there is an emerging pressure from commoditised SaaS-based agents, it is very likely that over the foreseeable future bespoke agents will be required to impart a competitive advantage and operate in a highly regulated environment.

Type of AI Agents

Zooming into the variations for bespoke agents we see three broad categories emerge.

Expert Agent

Multi-agent pattern with a structured flow, enabling LLM reflection.

This process, sometimes referred to as having an LLM "judge", can be combined with validation rules, which encourages reasoning.

Reasoning Agent

Pattern utilising reasoning models (over standard LLMs) without relying on a structured flow to process information.

Sequoia Capital – "Let's define an agent as any application that lets an LLM control the flow of the application."

Operator Agent

Pattern applying Reasoning Agent with ability to access external systems (e.g. call APIs, search the web).

OpenAI Operator is the most recent version of this, but offers a 'long-leash' view of the operator agent architecture.

Dispelling Common Myths and Misconceptions

Like any fast moving, impactful technology there is going to be hype, disappointment and then re-finement towards ROI. This comes hand in hand with myths about what Agentic AI can do, what it encapsulates and companies claiming they're harnessing agent-based systems when, in fact, they're not.

Here are some common misconceptions we've encountered from speaking to and working with enterprise leaders.

1. Myth: All AI assistants are agentic

Reality: Not all AI that responds or automates tasks qualifies as "agentic." Agentic AI refers specifically to systems that can reason, plan, and act towards a goal with some level of autonomy. AI that breaks down a user goal into steps, executes them across systems, and adapts based on results is agentic, but chatbot answering FAQs isn't.

2. Myth: Agentic AI replaces human decision-making

Reality: Agentic systems are designed to augment, not replace, human decision-making. The best use cases empower people with faster, richer, and more reliable decision support, especially in complex or repetitive processes. Trust, oversight, and human-in-the-loop governance remain essential.

3. Myth: Agentic AI is a black box

Reality: There's a perception that agents operate in an opaque and uncontrollable way. In fact, modern agentic frameworks increasingly emphasise transparency, traceability, and explainability, often using structured reasoning and logging mechanisms. Trust isn't optional – it's a design requirement. Because autonomous agents use reasoning models that provide step by step output, there is a kind of audit trail to the final output, or decision the model achieves. This can help increase the level of transparency and explainability over older LLM models.

4. Myth: You need cutting-edge AI models to build agentic systems

Reality: You don't need the latest proprietary LLM to build a useful agent. Agentic AI is more about orchestration – how models, APIs, data sources, and logic come together – than raw model power. Even smaller, open-source models can support agentic behaviour when paired with the right design, and may suit early experimentation better.

5. Myth: Agentic AI is just another hype cycle

Reality: While there's hype, the shift from passive systems to goal-directed, proactive automation represents a real evolution. We're seeing early deployments delivering real value in planning, customer support, ops automation, and more – especially where repeatability and structured reasoning are key.

6. Myth: All agentic AI involves complex self-reasoning

Reality: Agentic AI is often associated with deep, abstract reasoning – but in practice, not all agents need to "think like humans" to be useful. Many rely on basic logic chains, heuristics, or structured task decomposition. The key is purposeful reasoning aligned to the outcome, not philosophical depth.

7. Myth: Autonomy means full autonomy

Reality: Autonomy exists on a spectrum. An agent doesn't need to make end-to-end decisions across an entire business process to be valuable. Even something as simple as planning a few steps ahead, executing them, and adapting based on feedback can deliver meaningful gains. The most successful agents today operate with bounded autonomy, often within a defined set of APIs, rules, or oversight.

The Foundations of Agentic AI

Core Concepts Foundational to Agentic AI

- **Multi-Agent Systems:** Networks of intelligent agents that collaborate to achieve shared goals.
- **Reinforcement Learning:** Enabling agents to learn from experience and optimise outcomes over time.
- **Goal-Oriented Design:** Agents are programmed to achieve defined objectives while dynamically adjusting strategies based on environmental changes.

Three Key Measures of Agent-Based Systems

1. **Granularity:** How specific an agent's tasks should be to optimise performance. Highly granular agents excel in narrow, technical processes, while broader agents can manage complex goals.
2. **Reasoning:** The depth of an agent's decision-making. More sophisticated agents assess trade-offs, weigh risks, and make informed choices without human intervention.
3. **Autonomy:** The degree to which an agent can operate independently. More autonomous agents require less oversight, but they must align with governance frameworks.

Key Technologies Driving Agentic AI

- **Large Language Models (LLMs):** Powering natural language AI and decision-making capabilities.
- **Autonomous Agents and Orchestration Frameworks:** Systems that coordinate multiple agents to perform complex tasks.
- **Real-Time Data Pipelines:** Enabling agents to process data continuously and respond instantly to new insights.

LLMOps & the Feedback Loop

Governance of AI requires amendments to the way enterprises manage risk, and the advent of Agentic AI makes this even more pressing. The Open Worldwide Application Security Project has a convenient reference for AI Security and Privacy Guide as do many others, so this section won't emphasise governance but rather - LLMOps. We often find clients need to consider a couple of highly pragmatic steps to begin maturing their capabilities with LLMs.

LLMOps is a Practice and a Culture

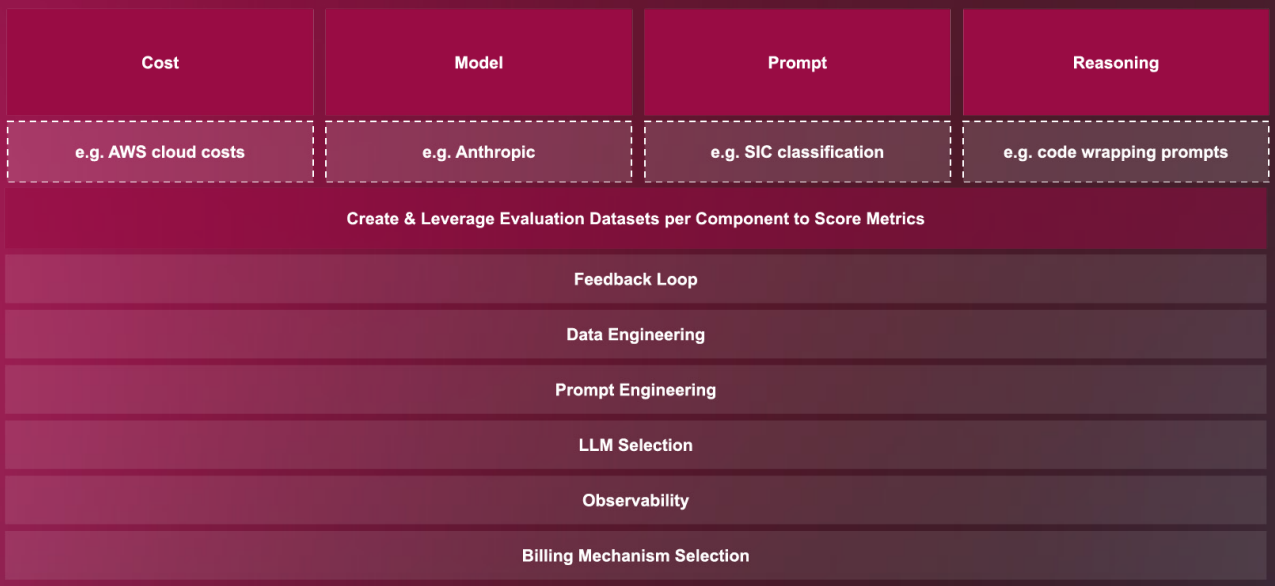
LLMOps is both a practice and a culture enabled through technology and data. It makes transparent the so-called 'black box' of AI, enabling teams to release to production and scale their application with confidence. Some readers may find this to be a different definition to what they are used to - LLMOps as a toolset.

The latter definition ignores the complexity of building, improving and monitoring LLMs, which are required to optimise the interdependent relationship between model, prompt, reasoning and cost.

Optimising to cost, for example, impacts your model selection, prompt smithing and choice in reasoning capability. The process involves facilitation of experiment tracking, which is what we advise all pursuing LLMOps to start with. The concept of the evaluation dataset is a core component of this facility. Given the nature of LLMs, it contains a specific input and testable output.

This allows the AI Engineer to test the combination of model and prompt for example, so that if they want to reduce cost by moving to a lower cost model they can assess if performance changes to an acceptable degree or not.

Start with Evaluation Datasets to Unlock Key Practices that Drive Technical Quality Improvement



Confidence in production also requires the right team shape and skills, which we discuss in more details further below.

The Feedback Loop is Needed... and has a Problem

The feedback loop is often one of the most understood elements of the AI engineering process when we engage with customers. We often see an assumption – there is a feedback loop to re-train the model. In the world of LLMs for the enterprise, 99% of the time this is not true.

When handling LLMs, the enterprise will rarely have the need (or ability) to re-train an LLM. This is similarly true for fine-tuning the model, it is generally not needed. Much of the finessing will be accomplished through the technical pillars we mentioned above – model selection, prompt optimisation and reasoning methodology.

With this in mind, the feedback loop is still essential. We cannot enhance cost and our technical pillars without advancing the evaluation dataset visa via this loop.

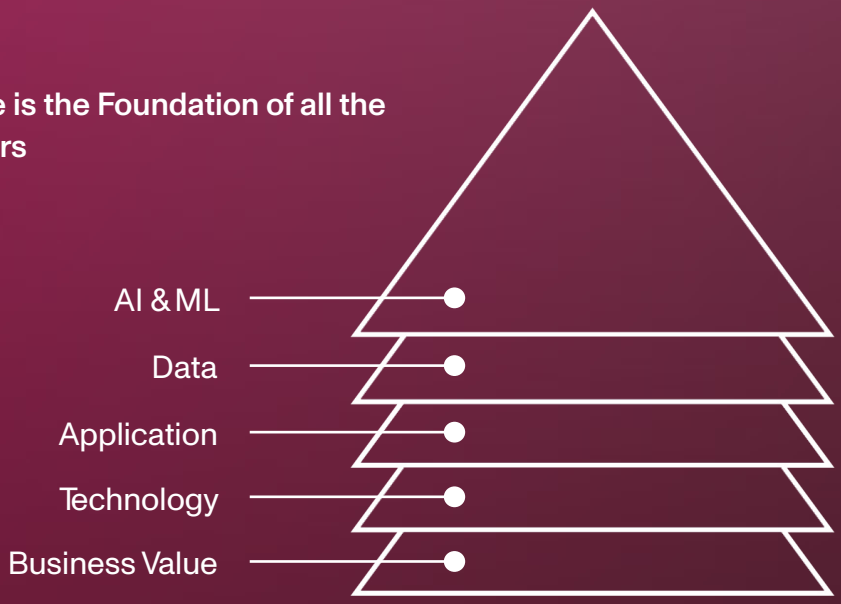
This introduces a non-trivial problem for many enterprises – particularly if you're an architect sitting on an architectural review board or equivalent within a highly regulated organisation. The problem? There must be production data that makes its way into a lower-level environment.

This often causes a stir as it breaks with the principle that those developing software should not have access to production-level data on an ongoing basis. Put simply, most organisations are used to a linear path for software development: development environment, to test, to pre-production, to production. There is nothing that completes the loop from production back to development as it would break with established governance practices.

However, to iteratively improve LLM-based applications, to enable LLMOps and get the most value from deploying them, a change in established practice is required.

The Road to Value: An Approach for Aligning with Strategic Goals

Business Value is the Foundation of all the Technical Layers



Business value underpins everything – it's the foundational layer to all technology and data investment and innovation. Initiatives that fail to ladder up to an organisational goal will not deliver value to the business. Solutions built in silos or on poor quality or inaccessible data run the risk of targeting business value but failing to deliver on their promise. All of these scenarios can dent confidence in AI initiatives.

Solutions should be built collaboratively between engineering, product and subject matter expert teams, delivering rapid proof of value through focused, outcome-driven engineering engagements.

From our experience building AI for highly-regulated enterprises, we make this happen by:

- Bringing together small, expert teams with the right mix of skills to move quickly from idea to value.
- Working closely with your teams to define the problem clearly, prioritise the most promising opportunities, and build strong business cases.
- Rapidly delivering working solutions that are tested in real business environments and measured against meaningful KPIs.
- Ensuring transparency and scalability from the start, so that your teams can confidently take AI into production and scale it across the organisation.
- This approach helps break out of "business as usual" and puts AI to work where it matters most – driving tangible, lasting value.

There are a number of ways to benefit from deploying Agentic AI, but several key decisions and considerations will define your route to value: and at scale.

Where to Start

Identify high-impact use cases that align with core business goals: The best AI doesn't just answer questions, it changes how people and organisations work.

Assess current data and platform readiness and invest in critical gaps: Invest in the discovery process, not just the tooling, and aligning with domain leaders from day one. Prioritise the big business problems that you're unable to solve and commit to utilising AI.

Modelling the Right Data

Ensure data is clean, accessible, and well-governed for optimal agent performance: AI can only be as effective as the data it draws from. That means moving beyond outdated data processing models and siloed data repositories to more dynamic, trusted, and accessible data platforms.

Develop scalable pipelines that provide agents with real-time insights: Focus on designing systems that give agents accurate data at the speed they need to operate effectively. This means rethinking traditional batch processes and building event-driven pipelines that prioritise freshness, context, and accessibility. Ensure agents have the data quality, structure and responsiveness required to reason and act with confidence.

Win Hearts and Minds by Proving Early Value Against Strategic Goals

Balance technical complexity with business impact, focussing on delivering incremental value: Start with simpler agents that solve real business problems, and scale in complexity as your internal capabilities mature. Work closely with both technical and business stakeholders to ensure each step in your roadmap drives tangible outcomes.

Prioritise initiatives that demonstrate early success, then expand agent capabilities over time: Quick wins help build trust and momentum, so identify high-value but manageable use cases where agentic AI can reduce effort, improve accuracy, or accelerate processes. Use these early examples to inform future investments, evolve internal skills, and scale deployment across the organisation.

Buy vs Build

Agentic AI is the buzz of the moment thanks to reasoning models such as OpenAI o1 and Anthropic's Sonnet 3.7. That is to say it is an extension of the current AI conversation and many of the prominent model suppliers will be part and parcel of this industry dialogue, which might suggest a buy first option - but that would be too hasty.

AI-first companies are introducing agentic AI as a new feature e.g. copilot, chat and 'agent' are all ways of using large language models (LLMs), and we also see recent agent-specific providers entering into the market that have yet to prove themselves as enterprise-grade.

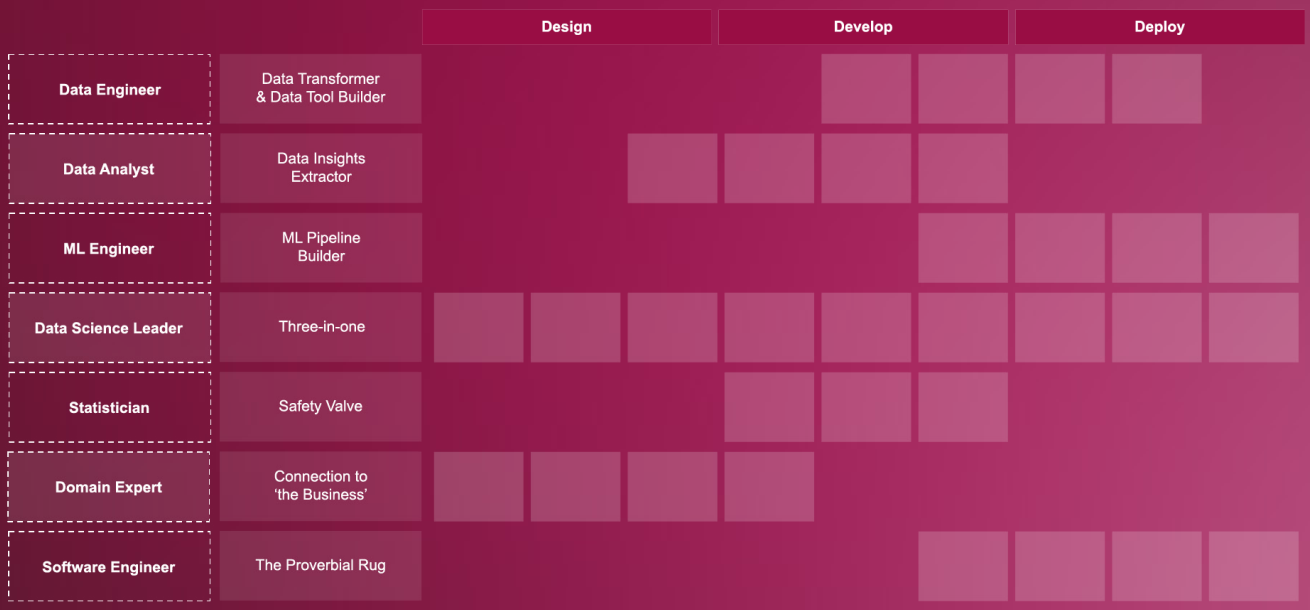
Additionally, it's worth taking a step back in reflecting on the point that leveraging Agentic AI is the biggest opportunity to re-think how you build new products. Why divorce yourself from the process that will help you learn how to navigate the nuances? To understand your levers and how to pull on them is key.

	BUILD	BUY
Customization	✅ Full control over features and tuning for specific needs	❌ Limited customization, must work within vendor constraints
Cost (Initial)	❌ Higher upfront investment (development, infrastructure, hiring experts)	✅ Lower upfront cost (subscription/licensing fees)
Cost (Long-Term)	✅ No recurring vendor fees after development	❌ Ongoing subscription or usage-based pricing can be costly
Data Privacy & Security	✅ Full control over data, reducing risks	❌ Data may be stored or processed by a third party
Competitive Advantage	✅ Unique AI model can provide differentiation	❌ Competitors may use the same AI tools, reducing uniqueness
Upskilling	✅ In-house teams are upskilled in emerging/ cutting-edge technologies	✅ In-house teams are upskilled in vendor platform
Innovation	✅ Freedom to innovate at your own pace (Or Mesh-AI's)	❌ Vendor-dependent

Invest in Skills to Supercharge Adoption

Equip teams with the knowledge and training needed to manage and enhance AI agent performance: You can't rely solely on vendor solutions or pre-packaged tools. You need to invest in internal capability: engineering, ops, AI, and cross-functional collaboration, to support AI use at scale. This isn't just limited to hiring new talent, focus on upskilling your workforce at all levels of seniority – including the exec.

Build cross-functional teams deployed across the product lifecycle:



How We Approach Building Agentic Systems

Mesh-AI proposes a pragmatic and agile approach grounded in several key principles. These form the key tenets of how we adopt any problem and build solutions that ensure ROI and build confidence in an organisation's ability to use technology to meet organisational goals.

- **Validate by Doing:** We emphasise rapid prototyping to test the feasibility and value of AI-driven solutions before full-scale development.
- **Adopt an Iterative Approach:** Start small, test outcomes, and scale effective strategies. We're focused on getting to value quickly whilst always keeping the bigger Northstar vision in mind.
- **Avoid Over-Engineering:** We focus on simplicity and cost-efficiency, ensuring that the solution meets needs without unnecessary complexity.
- **Hypothesis-Driven Framework:** We systematically test assumptions to refine models and focus on the factors that most influence energy supply and demand patterns.
- **Trustworthy AI:** We prioritise explainable AI, fairness, and robust governance to build trust among stakeholders and regulatory bodies.
- **Prioritise Explainability & Observability:** Ensure your agents' decision-making processes are transparent and understandable. We build data, AI and ML solutions with the precision of software engineering. Our approach ensures reliability, repeatability, and rapid deployment through automation-first practices

- **Balance Autonomy with Control:** Establish frameworks for when agents operate independently versus when human oversight is required.
- **Emphasise Human-centric Design:** Ensure agents enhance decision-making rather than replacing essential judgement. Maintaining human oversight ensures compliance with regulatory obligations and makes senior sponsorship easier to attain.

Case Studies: Early Stage Agents in Action

Agentic AI offers the potential to streamline operations, enhance customer experiences, and improve decision-making across the enterprise.

By automating complex processes and coordinating workflows in real time, agentic systems can significantly increase operational efficiency. From dynamic scheduling and supply chain optimisation to automated risk assessments and predictive maintenance, these AI agents enable more intelligent, autonomous value chains across industries like finance, energy, and logistics.

But as we mentioned above, the real value lies beyond efficiency gains, and in redefining entire processes.

Agentic AI also plays a critical role in enhancing customer engagement and strategic decision-making. Intelligent virtual assistants and personalised recommendations deliver more responsive and tailored user experiences. Meanwhile, AI agents equipped to analyse vast datasets can identify emerging patterns, assess risks, and surface timely insights to support human judgement.

The value lies not only in reducing costs and scaling efficiently, but also in uncovering new revenue opportunities and proactively managing risk in a constantly changing business environment.

Deploying Expert Agents in Financial Services

- **Reduced research time by 65% and unlocked $2.2m in cost savings**
- **Solution delivered in six months frees up expert capacity to focus on high-value analysis**

A leading provider of risk screening and due diligence solutions for financial institutions sought to modernise its research workflow and reduce reliance on manual processes. The goal: apply AI to alleviate tactical burdens on analysts, enabling them to focus on more strategic, high-impact work and improve the quality of flagship data products.

Mesh-AI developed and delivered an expert agent to automate the extraction, enrichment, and validation of information from media coverage and third-party sources – tasks that previously consumed up to a third of researchers' time. The result is a streamlined, intelligent curation tool that has reduced research time by over 65%, while maintaining the precision required for regulatory-grade outputs.

At the heart of this transformation is a single expert agent trained to identify entities involved in potentially risky activity and apply inclusion criteria with minimal human input. The system has already laid the foundation for $2.2 million in cost savings over three years, while boosting service-level adherence, client trust, and retention.

This solution, delivered in just six months, is now a cornerstone of innovation within the division – empowering the client's teams with new AI and ML capabilities and future-proofing the business against growing data volumes and rising customer expectations. Unlike more advanced agent-based systems, this expert agent doesn't collaborate with other agents but we have paved the way for a multi-agent system.

Agentic AI in Wealth Management

- **Faster identification of market trends, freeing up wealth manager resources.**
- **Improved portfolio performance through real-time adjustments, backed by human-in-the-loop.**

The current process for creating market commentary and investment opportunities reports to clients is costly and manually intensive, with limited room for personalisation. Wealth managers charged with creating these reports are taken away from higher value opportunities and tasks.

Mesh-AI has designed and implemented a solution which allows wealth managers to generate these market commentaries automatically based on various trusted news feeds; the solution also personalises deeper the commentary based on the assets owned by the customer.

The solution is based on an Agentic workflow of multiple LLMs which summarise articles at multiple levels (daily/weekly/monthly), then critiques and fact checks the content and finally iterates until the expected level of quality has been achieved. There is still a human in a loop before the commentary is sent to clients but the aim is to allow for more straight automation in the future while regularly comparing automated outputs with human-generated commentaries to avoid any drift in quality.

We can imagine expanding the scope of what this Agent does to more complex interactions for customers - for instance, proactively suggesting portfolio rebalancing based on signals from the market which are then carried out automatically once the customer has given consent.

Agentic AI in Construction

- **Driven 30% productivity uplift across the organisation**
- **Elevated the quality of reports, thereby reducing the risk of long-term contractual issues**

A multi-billion pound global construction and consultancy firm wanted to boost productivity through innovation and automation, targeting a 30% uplift across the business to increase its profit margin.

Mesh-AI developed an agent-based system to transform the organisation's high-volume bid response process, focusing on lower-value submissions that typically receive less oversight. Leveraging internal data and Microsoft Azure OpenAI, the tool generates first-draft answers directly within Microsoft Word, marking the company's first major step into AI-led transformation.

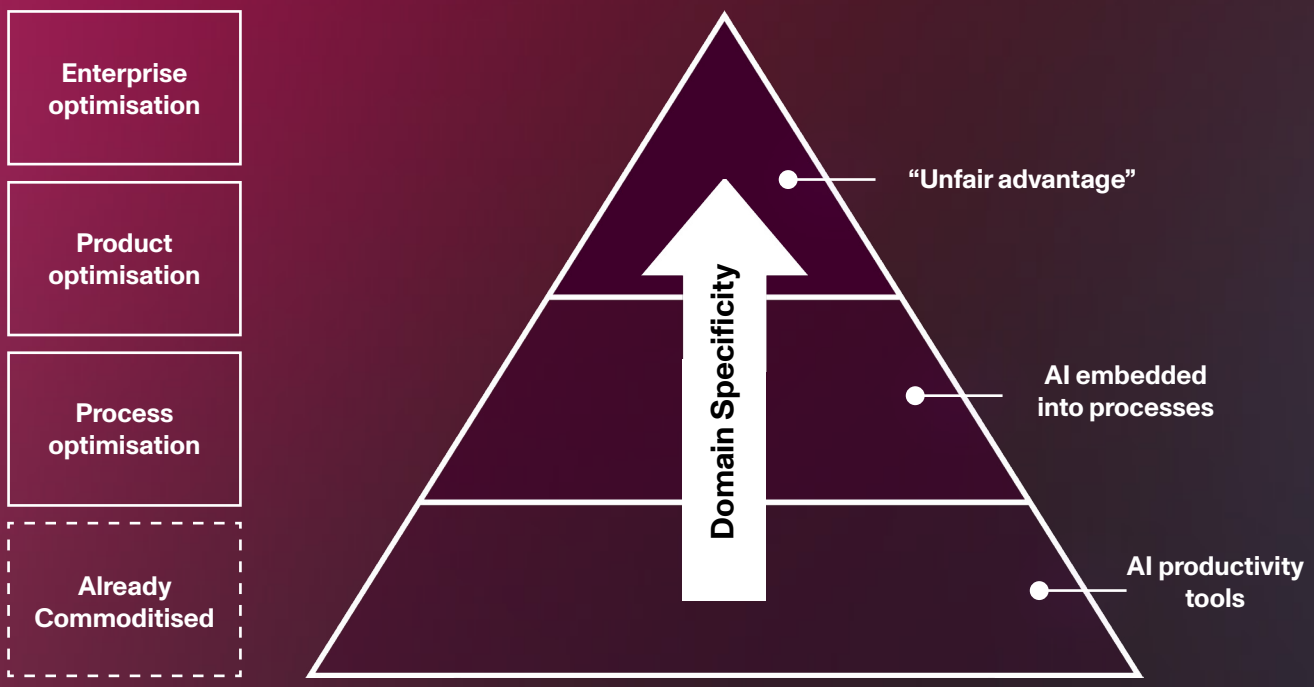
The tool's core feature is 'Generative Response', creating responses to questions contained within a bid document at the click of a button. This feature is grounded in a multi-agent framework that asks one agent to write and the other to review (i.e. give feedback), providing a 'check-and-balance system' to refine the output.

The agent-based system will increase bid writing efficiency for over 1,000 people in the organisation and save an estimated £500k a year. Bid Response AI will also elevate bid response quality, driving the organisation's competitive edge in the market, improving employee satisfaction and retention.

By demonstrating the impact of AI from PoC to production, our work is paving the way for the customer to meet its digital strategy and unlock the desired 30% productivity gains across the firm.

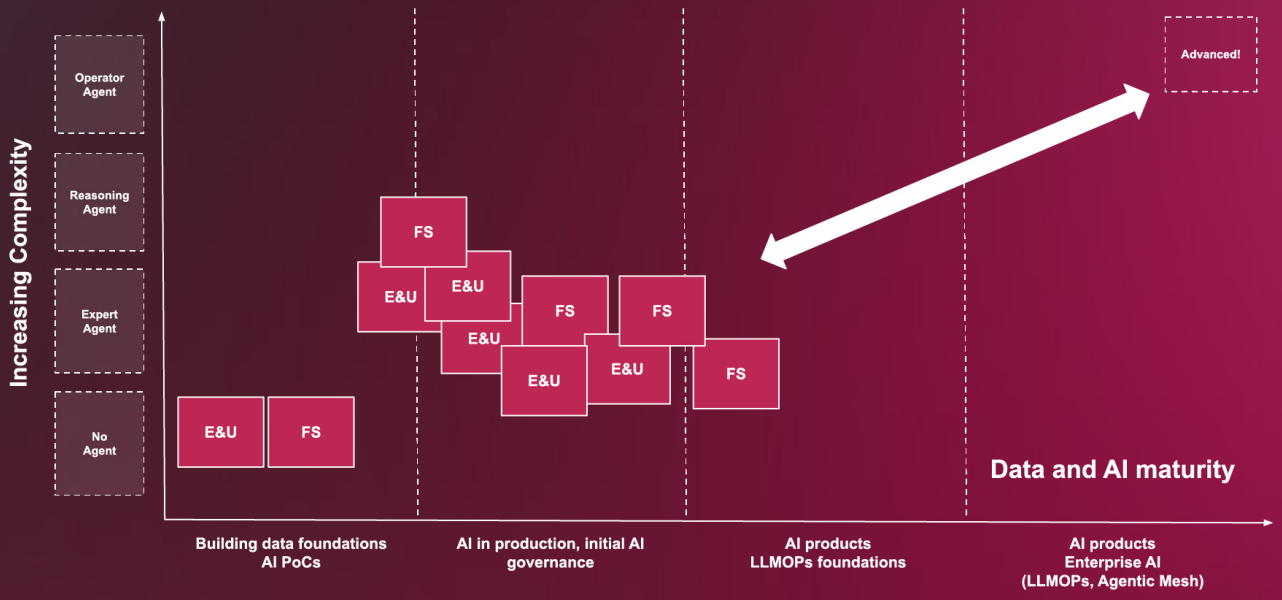
What Does the Future Hold for Agentic AI?

- **The Expanding Role of Multi-Agent Systems:** Agents collaborating across enterprise ecosystems will improve efficiency and productivity.
- **AI in the Enterprise Ecosystem:** Agents integrated into enterprise resource planning systems, IoT networks, and cloud platforms will enhance decision-making at scale.
- **Digital Workforces:** Expect a rise in digital workforces powered by intelligent agents performing diverse tasks with minimal supervision.
- **Granularity is the Next Wave:** Multiple agents with hyper narrowed focus on specific areas will deliver more impact, giving more knowledgeable responses.



Many of the opportunities over the last several years have been around operational efficiency related to process/product optimisation. In the market though, we are increasingly seeing clients ask themselves the question - is this our moment to re-think the opportunity in front of us? Is there an "unfair advantage" we can acquire? Is there a new revenue stream we can build instead of focusing on operational efficiency.

In subsequent months and years, applying domain specificity more broadly to many, more granular and more autonomous AIs will help clients move towards that advantage. At the moment the industry is still in an early state, where enterprises will likely spend 2025 getting their heads around placing AI products into production assisted by LLMOps. It's also the year they will experiment with reasoning models.

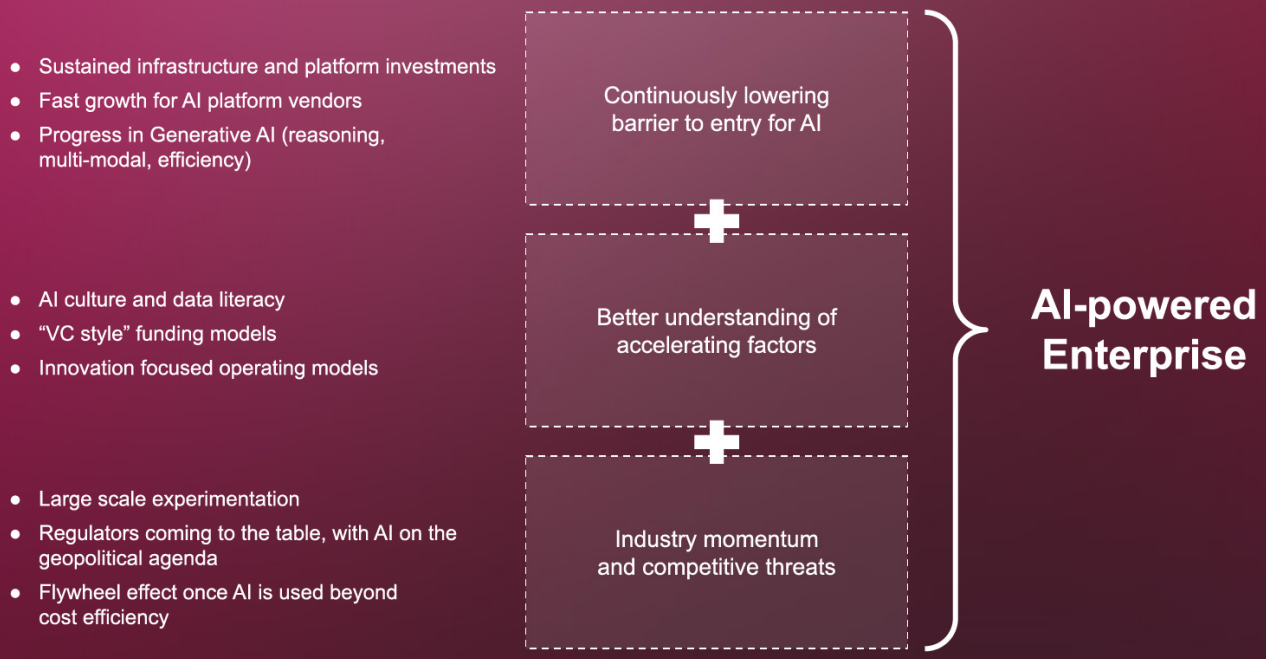


This does however set up 2026 as the year enterprises truly experiment with these concepts of autonomy and granularity. We believe these concepts will also begin to challenge the shape of at least one tightly held expectation of AI in the Enterprise.

There is a tightly held assumption that human in the loop (HITL) is the go to operating model to safely acquire the most performant AI-driven outcomes. In previous years, Mesh-AI would have without question agreed with this position.

However, there is evidence that the position of HITL is beginning to wiggle. For example, in a New York Times opinion piece regarding their paper Combining Human Expertise With Artificial Intelligence Experimental Evidence From Radiology, the author writes, "When A.I. worked independently to diagnose patients, it achieved 92 percent accuracy, while physicians using AI assistance were only 76% accurate – barely better than the 74% they achieved without AI."

As we see Agentic AI mature in the coming years, we believe Enterprises that find the greatest competitive advantage will be able to navigate the nuance of where and how to implement HITL along with granular, autonomous agents (amongst any other AI or ML models). The balance impacts not only the design of the system but also the safety of the product and performance of the organisation.



An Agentic Era

Agentic AI will undoubtedly play a pivotal role in reshaping business operations and accelerating innovation. Early adopters will swiftly realise the oppportunities lie beyond simply cost cutting, and will have developed the right foundations to take pivot with the inevitable developments as agent-based systems become more sophisticated, easier to embed and wid espread.

Those who see success will start by identifying key pain points where autonomous AI can add value, explore pilot projects and scalable implementations, and develop the AI literacy of their workforce at every level.

And just as we saw with the explosion of Generative AI tools, the agentic era we're entering will further emphasise the critical importance of data foundations. Without quality data to rely on, agent-based systems will hit a dead end, stifling internal appetite and confidence in operationalising innovation and, ultimately, holding an organisation back.

Mesh-AI is a transformation consultancy that exists to reimagine how enterprises operate, making data and AI their competitive advantage.

We turn enterprises into data-driven and AI enabled organisations, unleashing business growth and accelerating outcomes.

Mesh-AI

mesh-ai.com | hello@mesh-ai.com