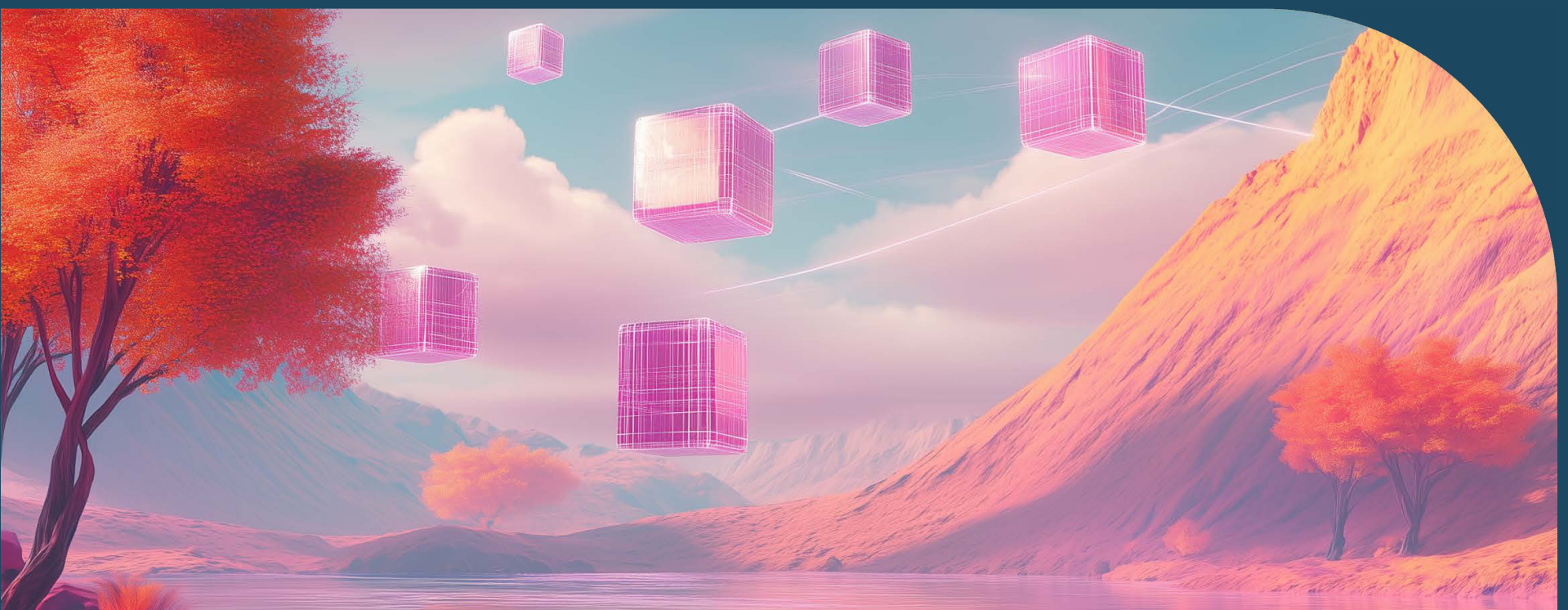




WHITEPAPER

AI Agents: How enterprises can unlock the next wave of productivity gains



ORI

ori.co

Contents

Introduction: Welcome to the age of AI agents	1
Decoding the basics	2
What are AI agents?	2
How are agents different from generative AI?	3
Designing AI agents	4
What does AI agent architecture look like?	4
Popular agent workflows	6
How do these workflows come together to form agents?	7
What is multi-agent orchestration?	7
AI agents for enterprises	8
The benefits of Agentic AI for enterprises	8
What use cases are ideal for agents?	9
What might not be ideal applications for agentic AI?	9
Crossing the AI adoption chasm	10
What are some pain points in adopting agentic AI and how to resolve them?	10
A detailed example of an AI agent and its entire workflow	11
What are the common agent platforms?	12
Build for success	12
Design for scaling, extensibility, and modularity	12
Enabling trust and collaboration	13
Partner with AI-native providers	13
Build Enterprise AI on Ori Global Cloud	14

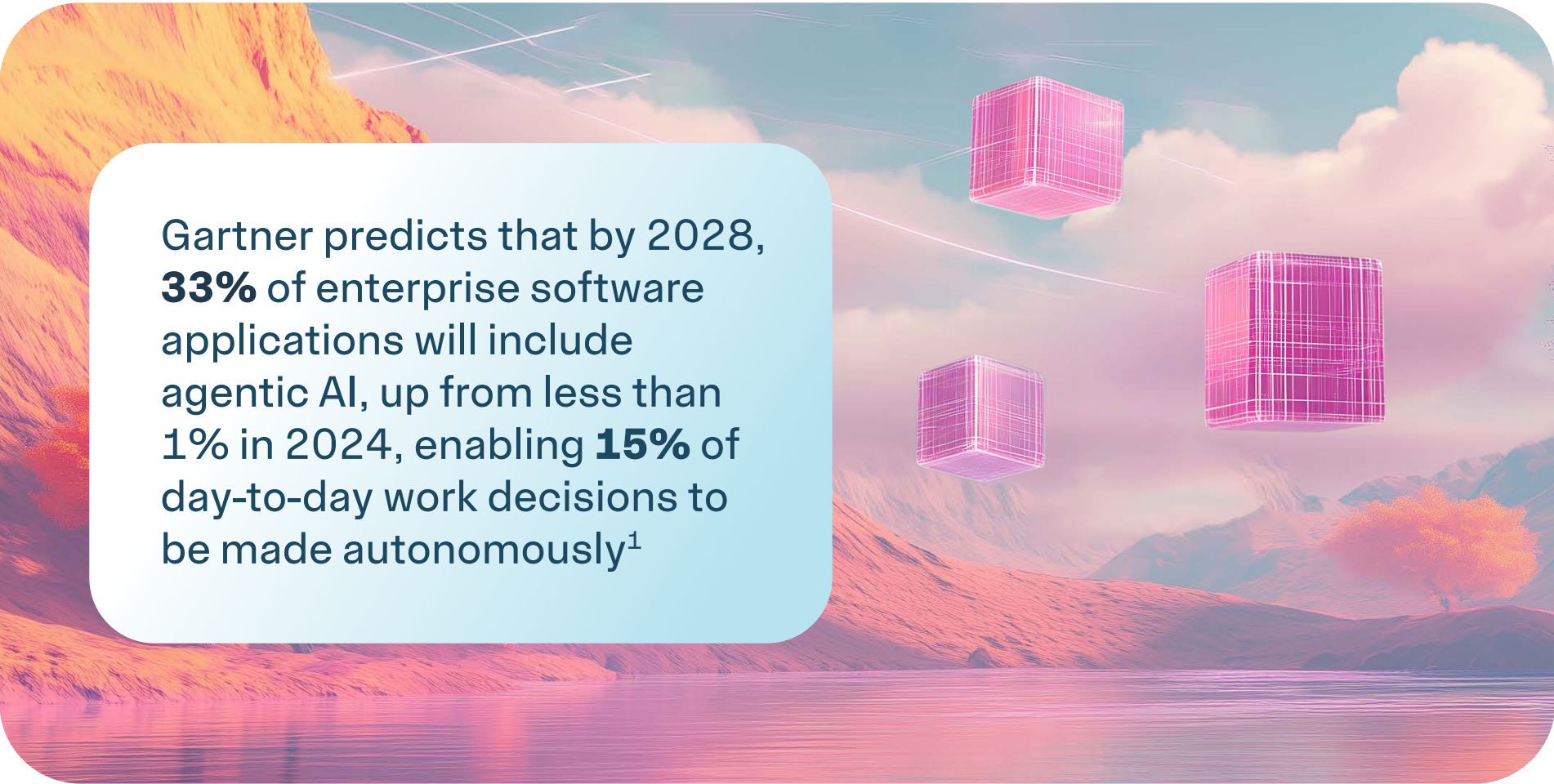
Introduction: Welcome to the age of *AI agents*

Today, employees across businesses are weighed down by administrative tasks that often divert their focus from showcasing their core skills and excelling in their primary responsibilities. For example, sales representatives dedicate a mere 28% of their workweek to actual selling activities, with the remaining time swallowed by administrative chores and manual data entry.

However, there is a seismic shift underway that is transforming the landscape of enterprise technology to help teams be more productive. Artificial Intelligence (AI) agents, with their ability to automate complex workflows, make informed decisions, and learn over time, are at the forefront of this revolution.

AI agents can boost efficiency across functions in the modern enterprise by assisting in administrative tasks, thereby helping employees focus on their core skills and high-value projects. For example, 83% of sales teams with AI saw revenue growth in the past year, and 75% of service teams say AI has helped them improve response times. The result is more efficient operations and happier customers.

Beyond sales, AI agents are cracking logistical puzzles, running customer interactions, optimizing delivery routes, creating software, assisting HR teams in employee onboarding, streamlining supply chains, and much more. Enterprises that embrace this new paradigm are poised to unlock unprecedented value!



Gartner predicts that by 2028, **33%** of enterprise software applications will include agentic AI, up from less than 1% in 2024, enabling **15%** of day-to-day work decisions to be made autonomously¹

1. Gartner (2025). Capitalize on the AI Agent Opportunity.



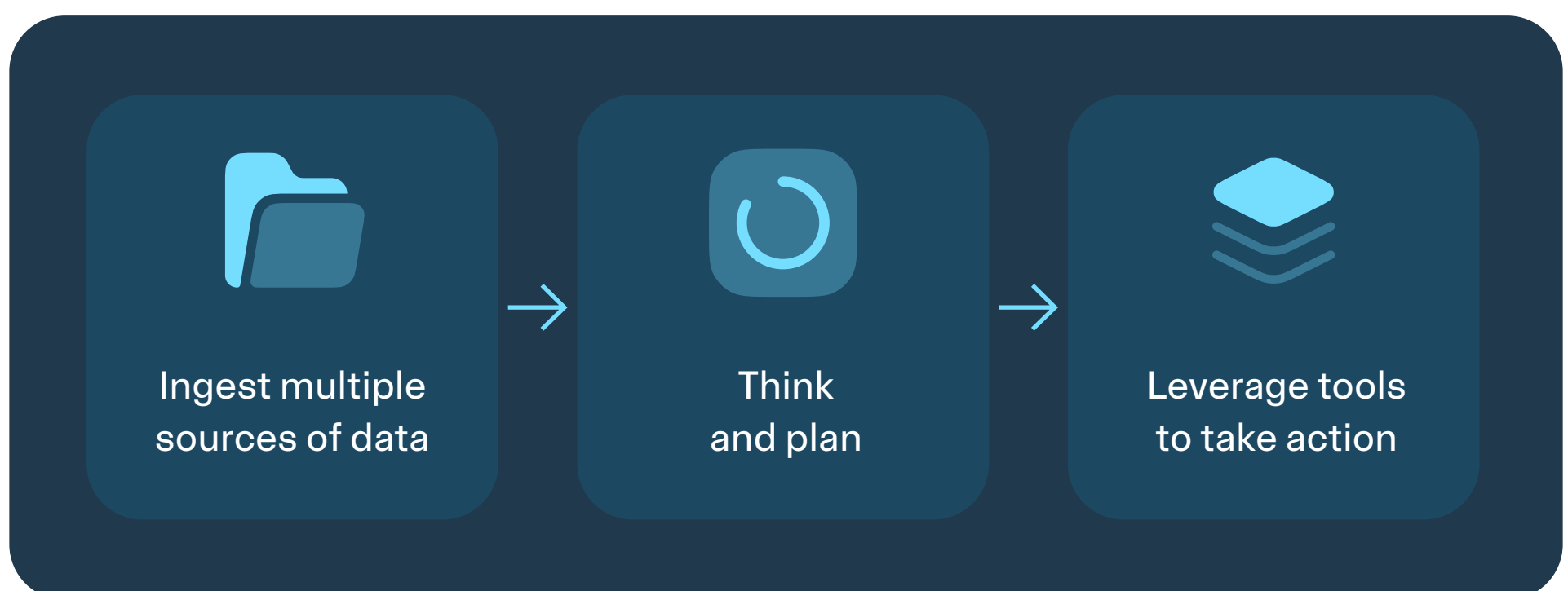
Decoding the *basics*

What are *AI agents*?

A Generative AI agent, at its core, is an application designed to achieve a specific goal by observing its environment and taking actions autonomously using the tools available to it. Unlike traditional automation tools, AI agents use advanced reasoning, natural language processing, and machine learning to achieve specific goals. They operate with minimal human intervention, adapting to new information and contexts. In contrast to conventional software where you have to manually dictate every click and keystroke, AI agents are autonomous within their designed parameters.

For example, an AI agent might sort and prioritize your emails, summarize key points, or generate a detailed monthly report without supervision. These agents can be particularly helpful for tasks that are mundane or labor-intensive, such as gathering agenda items or handling scheduling.

By digitizing or automating those tasks, AI agents free up employees to focus on work that involves more creativity or critical thought. One real-world example is an Employee Self-Service Agent that handles HR or IT-related questions. This makes processes such as resetting a password or locating a company's policies easier and more efficient, taking the burden off support teams.



How are agents different from *generative AI*?

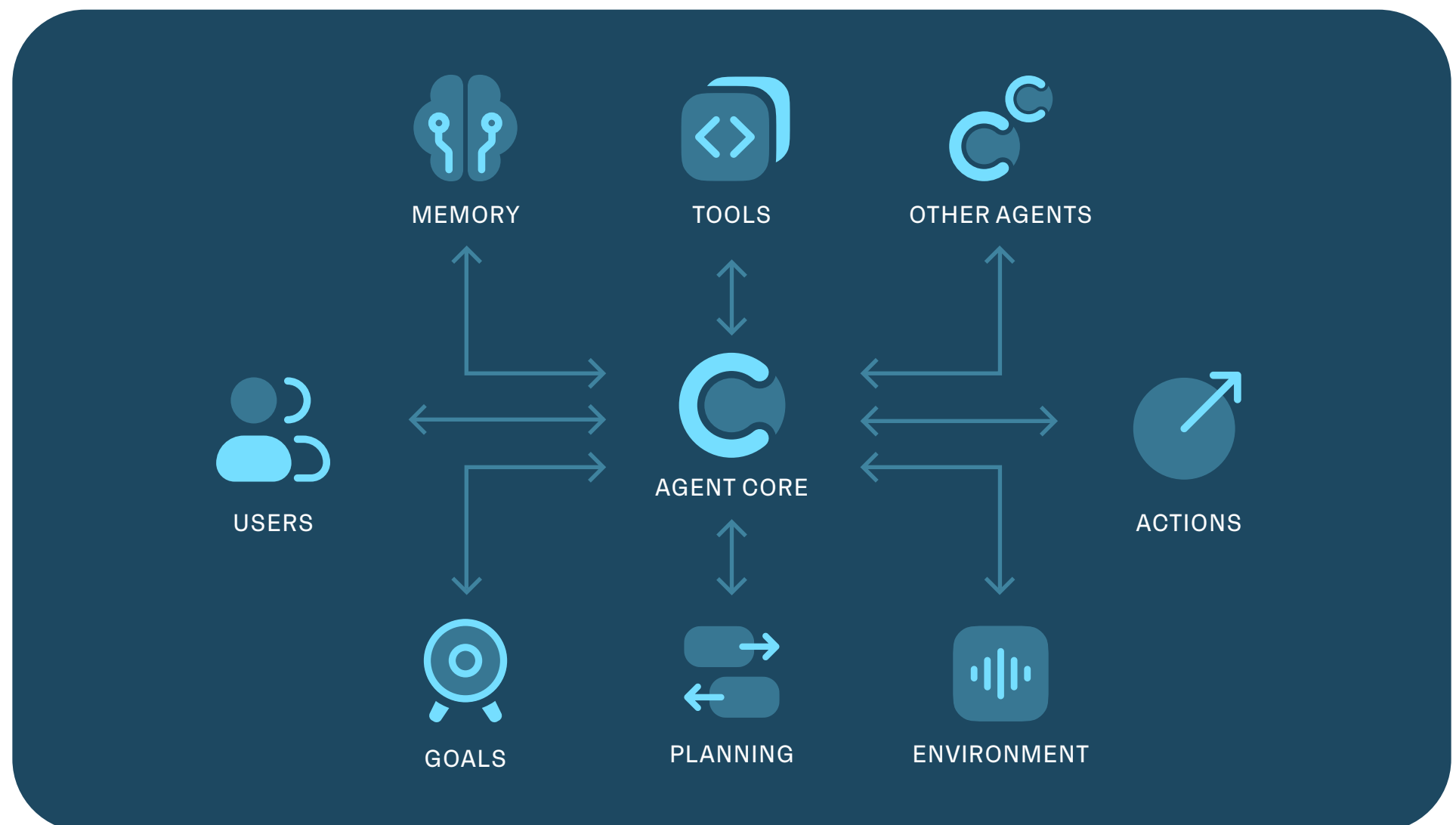
While generative AI creates content (e.g., text, images, code), AI agents take actions based on goals. For instance, a generative AI model might write an email draft, whereas an AI agent could send that email, schedule a follow-up, and analyze responses to optimize future communications.

AI agents integrate generative AI capabilities but extend functionality through autonomous orchestration. This distinction makes AI agents more versatile and impactful in enterprise settings.

Capability Area	Gen AI Models	Agents
KNOWLEDGE MANAGEMENT	Knowledge is limited to training data only	Knowledge extends through connections to external systems and tools
QUERY PROCESSING	Single inference/prediction based on individual user queries	Managed session history allowing multi-turn inference and contextual processing
TOOL INTEGRATION	No native tool implementation; requires custom development	Tools are natively implemented in agent architecture
SYSTEM INTEGRATION	External integrations must be custom-built	Built-in capabilities for connecting to external systems
REASONING FRAMEWORK	No native logic layer; relies on manually structured prompts and reasoning frameworks	Native cognitive architecture with built-in reasoning frameworks
PROCESSING MODEL	Static: processes each prompt independently	Dynamic: maintains state and learns from interaction flow
IMPLEMENTATION COMPLEXITY	Requires extensive prompt engineering for complex tasks	Comes with pre-built frameworks for handling complex interactions

Designing AI agents

What does AI agent *architecture* look like?



Agent Core: Acts as the brain of an agent, handling key decisions and actions. Usually built around a powerful AI/ML model such as an LLM, it leverages model knowledge to guide how the agent thinks and behaves. The Agent Core knows the agent's goals, provides a "user manual" for its tools, explains how to use planning modules, and adds relevant memory from past conversations based on user queries. It can also include a persona to shape the agent's behavior and response style.



Memory modules: Are vital for AI agents, acting like a diary of the agent's actions and user interactions. These memories are retrieved based on a mix of relevance, importance, recency, and other factors, ensuring the agent can provide accurate and contextually appropriate responses. There are two types of memory modules for AI agents:

- **Short-term memory:** Tracks the agent's thought process for answering a single question. This memory could be implemented using queue/FIFO data structure
- **Long-term memory:** Keeps a history of interactions over weeks or months. Local storage or a scalable vector database can be used to implement text embeddings or vectors to be used as long-term memory.





Tools: Are well-defined workflows that allow agents to perform specific tasks efficiently, much like specialized third-party APIs. For example, an agent might use a RAG pipeline to generate context-aware answers, a code interpreter for complex programming tasks, an API to search the internet for information, or services for tasks such as weather updates or instant messaging or calendar events.



Planning module: Tackling complex problems, like analyzing financial reports to answer layered business questions, requires nuanced approaches. An LLM-powered agent handles this by using task and question decomposition along with reflection or critique. For instance, answering “What were the three takeaways from a particular earnings call” involves breaking it down into smaller questions, such as identifying business headwinds, market shifts and financial results. Each sub-question can then be further divided, with a specialized AI agent guiding the entire process for clarity and precision.



Actions: Involve executing tasks based on cognitive decisions and can be physical actions in real-time such as actuators in case of robots or virtual actions such as calling APIs or updating content, and more.



Users: Can be human collaborators or other virtual AI agents who provide inputs ranging from prompts to objectives, deadlines, contextual knowledge and even feedback.



Environment: The agent core can also calibrate its response continuously based on external inputs such as text, camera, audio, and sensors which makes it particularly useful and more precise in physical environments.



Popular agent workflows

AI agent workflows are structured approaches that enable agents to efficiently complete complex tasks by leveraging distinct processes and components. Each workflow is designed to address specific challenges, such as task sequencing, resource allocation, or performance optimization. The table below provides an overview of common AI agent workflows, their key components, and how they function to achieve their goals.

Workflow	Short Description	Components	How It Works
PROMPT CHAINING	Sequentially connects multiple prompts or actions to complete a task	Prompts, AI models, task inputs, intermediate outputs	Each prompt generates an output that serves as input for the next prompt, creating a chain of tasks to achieve a goal
ROUTING	Directs tasks to the most suitable agent based on expertise or workload	Routing logic, task queue, agent pool	Incoming tasks are evaluated and matched to agents with the appropriate skills, ensuring optimal allocation of resources
PARALLELIZATION	Executes multiple workflows simultaneously for efficiency	Task splitters, distributed agents, aggregation mechanisms	Tasks are divided into sub-tasks and processed concurrently by agents, with results aggregated at the end
ORCHESTRATOR-WORKERS	Central orchestrator manages and assigns tasks to worker agents	Orchestrator, worker agents, task manager, coordination protocols	The orchestrator assigns specific tasks to worker agents, monitors progress, and integrates outputs for unified results
EVALUATOR-OPTIMIZER	Combines evaluative agents for monitoring and optimization agents for refinement	Evaluator agents, optimizer agents, feedback loop, performance metrics	Evaluator agents assess task outcomes, provide feedback, and optimizers adjust parameters to improve future performance



How do these workflows come together to *form agents*?

These workflows aren't siloed environments. Instead, these models produce text, images, sounds, or actions that cohere and form a unified AI agent.

In a healthcare use case, a digital worker overseeing a patient's data organizes records by routing them through the appropriate categories. It processes thousands of patient histories at once thanks to parallelization and prescribes treatment plans with an evaluator-optimizer. When these workflows are merged together, the agent is able to address complicated tasks that would typically require several human experts.

By combining workflows, enterprises can create robust agents capable of handling diverse tasks. For example, an orchestrator-worker setup can route customer support tickets, while an evaluator-optimizer duo ensures resolution quality. This modular approach allows enterprises to tailor agents to their specific needs, maximizing their utility.

What is *multi-agent* orchestration?

Multi-agent orchestration involves coordinating multiple agents to work collaboratively toward a shared objective. This approach mirrors human teamwork, where different agents assume specialized roles. For example, in an e-commerce setting, one agent might handle inventory checks while another manages customer queries, all within a unified workflow. The orchestrator ensures that these agents work together efficiently, enabling seamless execution of complex tasks.

Consider it like an orchestra of specialists working together on a project, each bringing their unique talent. That's what drives faster, cheaper, better solutions, creating the momentum needed for wider AI adoption and unlocking new ways to solve problems.



AI agents for *enterprises*

The benefits of Agentic AI for enterprises

Agentic AI empowers enterprises to operate smarter, faster, and more cost-effectively by automating repetitive tasks, enabling employees to focus on strategic initiatives. With real-time data insights, streamlined workflows, and seamless scalability, it drives better decision-making, reduces overheads, and ensures businesses can adapt to growing demands without escalating costs.



Enhanced efficiency

Automates repetitive tasks, freeing employees for strategic work.



Improved decision-making

Leverages real-time data for accurate insights and proactive actions.



Cost reduction

Reduces operational overheads by streamlining workflows and eliminating inefficiencies.



Scalability

Easily adapts to increased workloads without proportional increases in costs.



What use cases are ideal for agents?

- **Sales:** Agents enhance lead scoring by analyzing customer data, interactions, and engagement patterns, enabling sales teams to identify high-value prospects and concentrate their efforts on leads with the highest conversion potential.
- **Customer Service:** AI-powered customer service agents elevate the support experience by answering commonly asked questions instantly, search their knowledge base or Internet to refine their responses and route complex queries intelligently to human experts.
- **Software Development:** Coding agents enhance efficiency by providing AI-driven code suggestions, detecting bugs, and automating repetitive tasks, which accelerates the development process and elevates code quality.
- **Financial Services:** Fraud detection agents that identify patterns and anomalies that might elude human analysts, automatically block suspicious transactions and notify respective teams.
- **Human Resources:** Agents make it easier to streamline recruitment processes, employee onboarding, and performance management.

What might not be ideal applications for agentic AI?

AI agents do best in clearly defined spaces. When it comes to tasks requiring creativity, emotional intelligence, or complex human judgment, they are not very effective. For example, jobs that require long-term strategic planning or advanced negotiations are more appropriate for a human. AI agents really shine when they're able to identify patterns and be consistent.

Agents falter at processes that require dealing with variables that are difficult to quantify or make predictions about. Tasks requiring deep human judgment, creativity, or empathy (e.g., strategic visioning, complex negotiations) may not yet be suited for agentic AI. Additionally, scenarios with highly unstructured data or unpredictable variables may pose challenges.



Crossing the AI adoption chasm

What are some pain points in adopting agentic AI and how to resolve them?

Adopting AI agents isn't without its challenges. Individuals and organizations, particularly large legacy organizations, can be resistant to change. They are used to taking a much slower, more cautious approach. This is often accompanied by a corresponding lag when it comes to embracing innovation.

These firms must **prioritize identifying low-risk, high-value use cases**. For one thing, they can use AI tools that simplify and automate repetitive tasks such as data entry, collection, and appointment scheduling. By starting small and scaling gradually, you help build confidence across teams.

A second challenge is the very human desire to interact with humans. Even when AI has the potential to be more accurate or efficient, humans tend to favor human judgment. Bridging this gap will require **addressing transparency head-on**, providing plain language explanations for how AI agents function and demonstrating their reliability in real-world situations.

An AI-powered customer service tool can help agents behind the scenes in more effective ways. For example, it can analyze and summarize customer history prior to the agent entering a direct conversation.

Enterprise teams should not think of AI agents through the traditional software engineering lens where processes are very deterministic and rigid. Agents are more akin to designing a process with humans where outputs vary significantly, making it vital to include checks and balances within agentic workflows.

Finally, there's the matter of complexity. Many organizations struggle to fit AI agents into current workflows. Making this process easier with **intuitive platforms and robust technical assistance makes for a much more seamless adoption**.

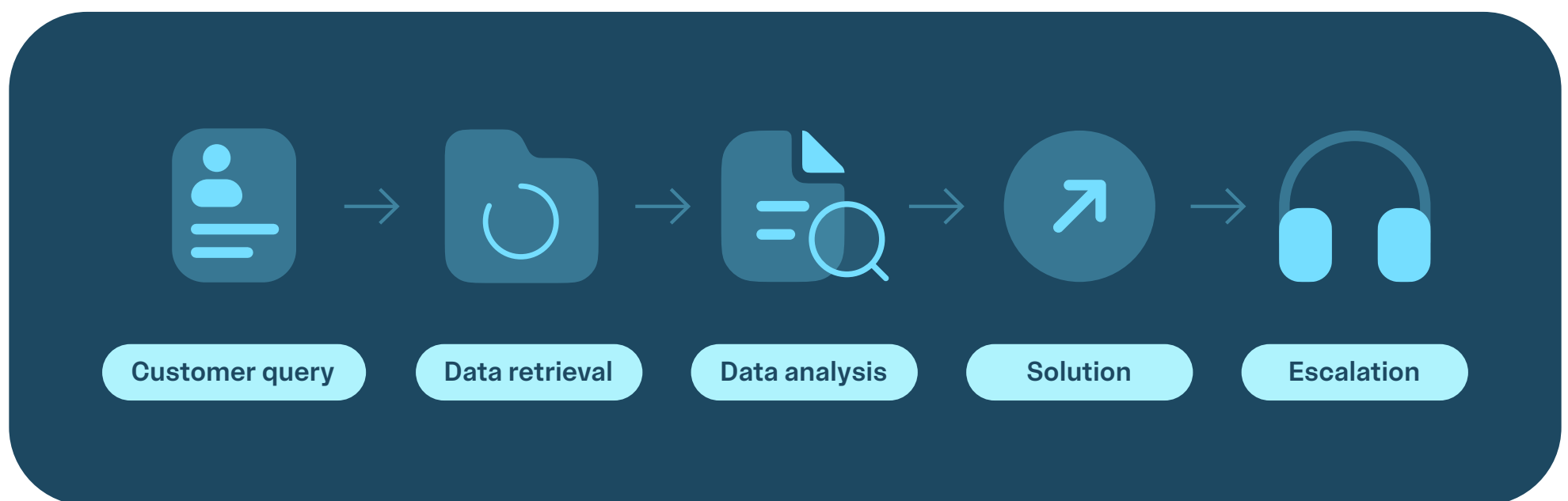


A detailed example of an AI agent and its entire workflow

Here's a quick step-by-step look at the workflow of a customer service AI agent. Say, for instance, a customer reaches out to support through chat looking for help with a billing question. Here's how the AI agent operates:

1. The agent collects the customer's query and account details in real-time.
2. It accesses the customer's billing history from the company database.
3. The AI identifies the issue, such as an overcharge, by comparing historical data against current charges.
4. It then automatically drafts a proposed resolution, like the issuance of a refund or an explanation of the charge.
5. If the problem is too complex, the agent forwards all details to a human representative, saving time by pre-filling the context.

Below is a simple diagram illustrating this process:



This new, streamlined workflow frees up time for human agents by automating repetitive tasks. Because of that, they can focus more of their efforts on the finer points of customer engagement.



What are the common agent platforms?

- LangGraph: Workflow optimization and automation for diverse use cases.
- Crew AI: Multi-agent collaboration designed for enterprise scalability.
- Beam AI: Enterprise platform to create and manage a team of AI agents.
- Rivet: Task-specific agent solutions with user-friendly interfaces.
- Vellum: Comprehensive orchestration tools enabling seamless integration with existing systems.

Build for success

Design for scaling, extensibility, and modularity

To meet evolving enterprise needs, IT leaders must prioritize architectures that support scalability, extensibility, and modularity from the outset.

- **Scaling** ensures the system can handle increased workloads as adoption grows, whether that involves deploying additional agents, processing larger datasets, or integrating with new applications.
- **Extensibility** allows the seamless addition of new capabilities, such as incorporating emerging AI models or integrating with future enterprise systems without the need for extensive rework.
- A **modular approach** is essential for enabling iterative development. By breaking the system into discrete, manageable components, enterprises can improve individual modules without disrupting overall operations. For instance, enhancing an agent's natural language processing capabilities should not require re-engineering its integration with CRM systems. IT teams should leverage containerization and microservices architecture to isolate functions, simplify testing, and streamline updates.

Tip: Use open standards and API-driven designs to maximize compatibility with other enterprise systems and technologies. This ensures flexibility for future integrations, minimizing vendor lock-in.



Enabling trust and collaboration

For AI agents to thrive in enterprise environments, trust and collaboration are non-negotiable. Start by ensuring transparency in AI operations—explain how decisions are made, which data is used, and how outcomes align with business goals. IT leaders should prioritize deploying agents with interpretable AI capabilities that allow users to understand, audit, and trust their recommendations.

Equally important is aligning AI agents with organizational objectives. This can be achieved by embedding organizational values into AI design and fostering cross-functional collaboration during the development and deployment phases. Encourage regular employee feedback to refine the performance and behavior of AI agents over time. Feedback loops help identify pain points, reveal unexpected use cases, and maintain alignment with end-user expectations.

Establishing a governance framework is another cornerstone of trust. This framework should define ethical AI principles, compliance measures, and accountability structures to ensure responsible AI usage. Regular audits and adherence to industry standards like ISO/IEC 38500 can reinforce ethical practices and build stakeholder confidence.

Tip: Create a dashboard for monitoring AI agent activities, providing end-users and managers with insights into performance, decisions, and areas for improvement.

Partner with AI-native providers

Collaborating with specialized AI providers offers enterprises unparalleled advantages, from leveraging cutting-edge technology to gaining access to expert support. These partnerships can significantly reduce the complexity of building AI systems in-house, offering cost-efficient, scalable solutions tailored to specific industry needs.

Choose providers with a proven track record in enterprise-grade AI solutions, emphasizing their ability to integrate with existing infrastructure, meet compliance standards, and scale alongside organizational growth. Look for partners that actively invest in R&D, ensuring your enterprise remains ahead of technological trends like advanced generative AI models or federated learning techniques.



Additionally, AI-native providers often have access to extensive libraries of pre-trained models, enabling faster time-to-value and more reliable outcomes. Their ongoing updates and optimizations can keep systems aligned with the latest best practices, allowing enterprises to focus on strategic objectives rather than technical maintenance.

Tip: Assess potential providers based on their ability to offer flexible deployment models—on-premises, cloud, or hybrid—to suit your organization’s infrastructure preferences and data governance requirements.

Build Enterprise AI on Ori Global Cloud

Learn more about Ori Private Cloud →

Ori is the first AI Infrastructure provider with the native expertise, comprehensive capabilities and end-to-endless flexibility to support any model, team, or scale. Our Private Cloud platform enables you to build AI with fully-customized infrastructure:



Accelerate AI development and market *readiness*



Develop AI in a secure, compliant *environment*



Optimize operational costs and *maximize* your ROI



About Ori

Ori is the first AI Infrastructure provider with the native expertise, comprehensive capabilities and end-to-endless flexibility to support any model, team, or scale. We're building the backbone of the AI era so that the technology of tomorrow can advance our world.

Ori believes that the promise of AI will be determined by how effectively AI teams can acquire and deploy the resources they need to train, serve, and scale their models. By delivering comprehensive, AI-native infrastructure that fundamentally improves how software interacts with hardware, Ori is driving the future of AI.

Learn more at www.ori.co →

ORI

ori.co