prosus

The Coming Age of AI Colleagues

Prosus Agentic AI Research Report

2026

Table of Contents

prosus

Report preface: Introduction to Prosus and why we wrote this report

AI agents are one of the hottest topics in technology today, but there is a dearth of *independent* data on what AI agents actually do and the value they deliver. The scarce data that does exist is largely produced by model labs or tech companies selling agentic services, making it difficult to separate real insights from marketing content in disguise. This absence of *independent* data and best practices for implementing agentic AI makes corporate boardrooms more hesitant to fully invest in the technology.

To address this, Prosus set out to publish what we believe to be the first independent analysis of AI agents operated at scale. Prosus is a global technology company building AI-driven lifestyle ecommerce ecosystems across Europe, Latin America and India. It owns and operates leading brands such as Just Eat Takeaway.com in Europe, iFood in Brazil, Despegar in Argentina, and PayU in India. Prosus also holds large minority stakes in prominent tech companies, such as Tencent, Swiggy and Meesho, amongst others. Prosus is the fifth-largest publicly-traded technology company in the EU.[1]

Prosus has built several different proprietary AI products for its portfolio companies, most notably an internal AI agent platform called Toqan, which is for internal use only. **Prosus' 40,000 employees globally have built over 60,000 AI agents using Toqan.** Aside from increasing velocity, productivity and agility of our teams, this is the first step towards reimagining entire corporate departments from the ground up, and ultimately building autonomous AI-enabled organizations.

The first section of this report shares the data and insights we have gathered from building 60,000 AI agents. The second section explains how we accelerated AI adoption within our companies. We outline how to drive change management and use of AI at large organizations. In the conclusion we discuss the concept of autonomous AI-enabled organizations that reimagine entire corporate departments based on AI. The appendix draws on insights from our AI engineering team to explain how to address some of the thorniest technical challenges in implementing AI agents at scale. Whether you are an AI engineer, an executive, a change management specialist, or someone who is just curious to learn about how agentic AI actually works, this report will provide insights that are relevant to you.

[1] As of May 6th, 2026

Executive summary

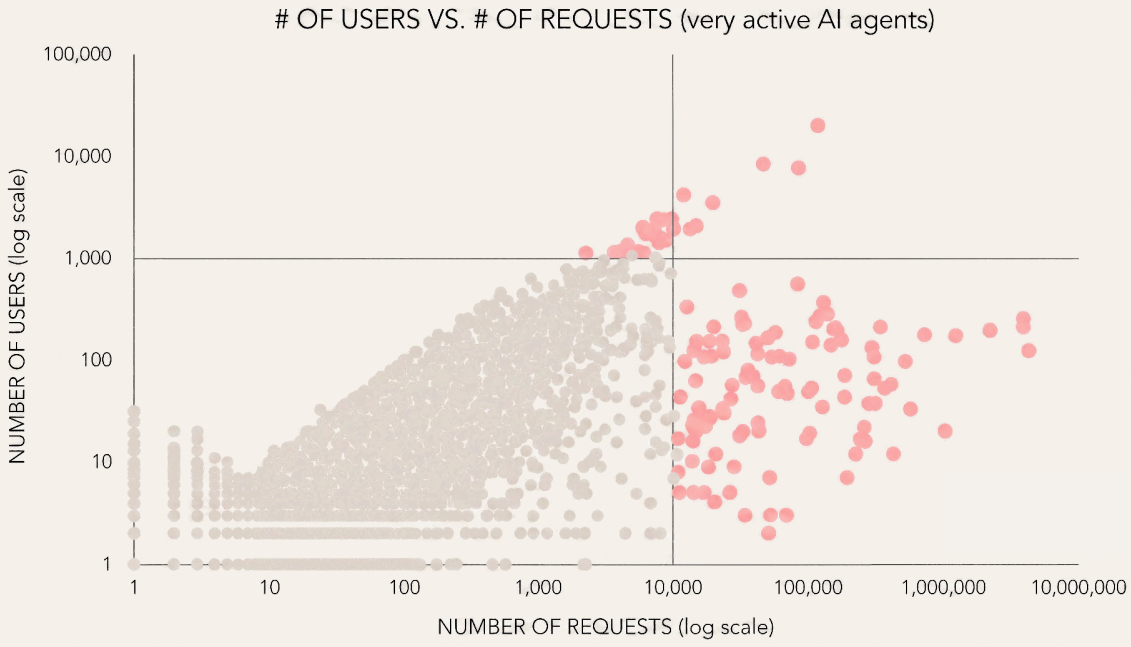
Much of the global conversation around AI remains fixated on model breakthroughs, timelines to AGI, and predictions about the future of organizations driven by autonomous AI. Despite all the headlines, most organizations are still relatively early in their AI adoption journey. With this in mind, we have set out to share our unique experience as a large global technology company that has transformed its business into an AI-first organization. Prosus has over 40,000 employees across our wholly-owned, global portfolio of companies. Together, these employees have built over 60,000 AI agents.

We believe Prosus is among the largest operators of AI agents globally whose core business is not the sale of AI services to other companies. *The Coming Age of AI Colleagues* report provides a hype-free guide to help organizations accelerate AI at scale, understand what to expect, discover what use cases drive ROI, and more.

Implementing AI agents on top of traditional corporate structures is, of course, just the first step. This report concludes by exploring a more radical horizon: autonomous AI-enabled organizations that reimagine entire corporate departments - sales, operations and beyond - from the ground up. Most companies, however, are still far from that reality. In the meantime, we hope that sharing our insights from over 60,000 AI agents deployed across traditional company structures will offer practical value to others navigating this journey.

Key findings

The classic business "power law" holds true for AI agents: approximately 2% of active AI agents drive a disproportionate share of business impact. Below is a chart of our AI agent workforce that provides a visualization of this phenomenon. The implication is clear – the first priority for all organizations is to double down on those agents. Organizations should also use data-driven analysis to find and then nurture the agents that are trending towards the "power law" group.



A clear pattern emerged: companies across the Prosus portfolio kept building the same 20 "power law" agent use cases. Across different industries, geographies, and languages – and with no mandate from Prosus HQ – companies consistently converged on the same 20 agentic AI use cases, each delivering a strong, immediate ROI. While not a comprehensive list, these 20 use cases are a good starting point that all organizations should implement. They also offer a window into the future: tasks that will likely become "default AI" for every business within the next few years. Examples include: reviewing and triaging high volumes of inbound messages, querying company datasets to generate custom reports, and agents designed to track customers at risk of churn.

AI agent complexity falls into four tiers that map closely to human employee seniority levels. Senior-level agents have the largest *total* number of users (below). When looking at *daily* usage, it splits almost evenly between senior and junior agents, revealing that simpler AI agents carry significant weight in daily tasks.

We identified 54 distinct AI agent tasks across corporate functions. Data analytics and market intelligence claimed the largest share at 18%, where AI's ability to surface business trends from large internal and external data sets is extremely valuable. Operations follows at 15%, for tasks such as forecasting supply or managing inventory. Notably, the third largest share of AI agents – 14% – sits outside any formal department, representing employees' personal AI assistants. A full breakdown by department and task is included in the report.

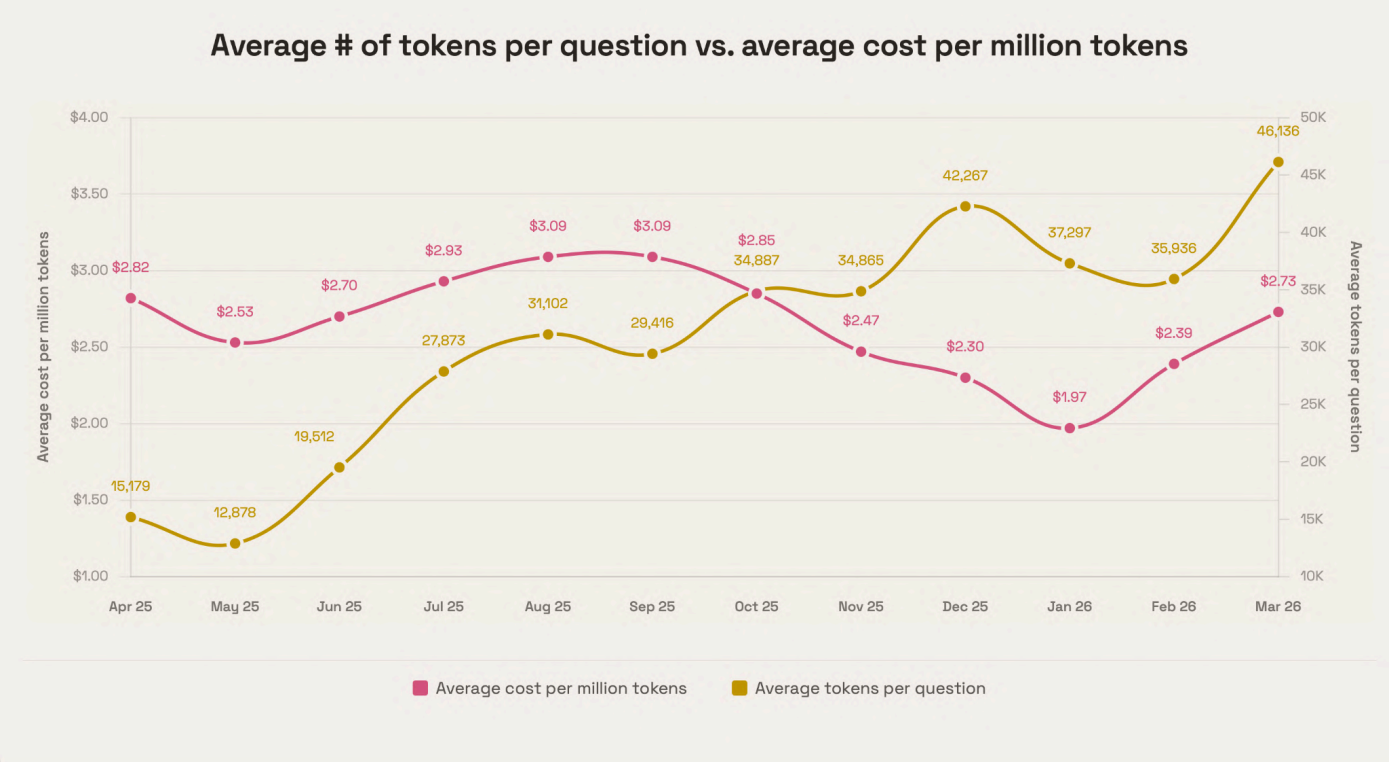
Productivity agents – where a clear ROI is measured in hours saved – can be divided into three distinct bands. The majority (82%) deliver modest but meaningful gains: under 20 hours saved per month, typically personal AI assistants for individual employees. The middle tier (17%) saves between 20 and 173 hours monthly – roughly up to one FTE – by automating tasks and redirecting human effort toward higher-value work. At the top, less than 1% of AI agents operate at a different scale entirely, delivering the equivalent of thousands of hours of monthly work.

Value agents are those that can be measured by revenue growth or cost reduction. A small group of outliers, value agents deliver outsized gains. Most deliver under $1M in annual value (e.g., reducing annual audit costs). A smaller middle tier generates $1M–$10M in value, such as a customer-facing assistant that answers niche vacation rental questions. A tiny number of outlier AI agents deliver in the tens of millions of dollars in annual value. For example, one portfolio company used AI agents to manage communications and onboarding for a new third-party affiliate marketplace. Projected annual revenue for this new marketplace is $83M.

Where ROI is difficult to quantify, we use the "delete it tonight" test. Business unit leaders are asked "what would happen to revenue or costs if the agent were permanently deleted?" When framed this way, most leaders can provide a tangible ROI estimate based on hours, revenue, or cost.

Across-the-board improvements in AI means that most models are now "good enough" for nearly all agent tasks. Despite this, users' reluctance to switch between models can create cost challenges. Our internal agentic AI platform, Toqan, offers ten different AI models to our 40,000+ employees. We observe that advancements in AI mean that the latest cutting-edge models are only needed for the most complex or difficult AI agent tasks. We have also found that users are reluctant to switch models once their agent works. As a result, introducing new models creates a trade-off: employees may use unnecessarily expensive, state-of-the-art models for tasks that could easily be handled by cheaper alternatives.

AI costs are volatile and unpredictable. We've found it's more effective to focus on optimization and empower individual business lines to make their own cost decisions. AI costs are difficult to predict at scale because AI agents rely on different combinations of models, infrastructure, and tool usage. For example, we have found that there is no straightforward relationship between the average number of tokens per question and the average cost per million tokens (pictured below). Our solution is a two-tier approach: anyone can use AI freely under 200 requests per hour. Above that threshold, employees need departmental approval for what we call Production Tier AI, as the business line itself is best placed to decide whether the benefits outweigh the costs. Separately, our central AI engineering team works to optimize the setup, and drive lower AI costs across the board.



We have developed a three-step process for driving adoption of agentic AI at large organizations. The need for such a framework is real, as most large organizations have struggled to move from experimentation to scale. In Part Two of this report, we share our framework and best practices that took us from zero to 60,000 AI agents in roughly eighteen months, covering everything from pricing structure and dedicated roles, to launching a “co-creation” AI agent, to the importance of prompt training.

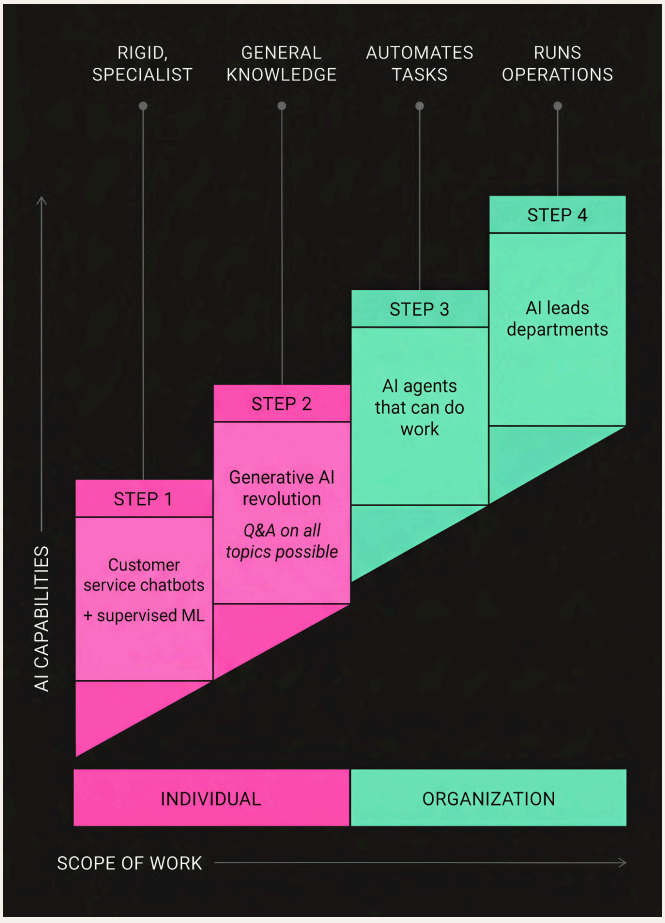
Looking ahead

After scaling to 60,000 AI agents, we now have a clear view of what an AI-driven organization looks like in practice and what the future holds.

What we observe across our portfolio is transformation rather than total disruption. We see three trends. First, individuals and small teams are building AI agents to work more efficiently. Second, every company is adopting the same "quick win" AI use cases that will likely become "default AI" in the future. Third, portfolio companies nurture the "power law" agents that create a competitive advantage.

This report focuses on implementing AI within existing ways of working — but that is just the starting point. The deeper transformation lies ahead: the rise of autonomous, AI-enabled organizations, where core business functions like sales, customer support, and operations are largely executed and coordinated by networks of AI systems, with minimal human oversight.

Comparisons are often made to how the invention of electricity impacted factories. Electrification delivered its largest productivity gains only after firms reorganized the entire production process around electricity. Simply substituting steam motors or water wheels with electric motors in existing factory designs offered only relatively smaller gains. The same framing is applied to AI. Proponents argue that it is only when companies start re-building core functions from the ground up with AI in mind that businesses will experience the biggest gains from AI.



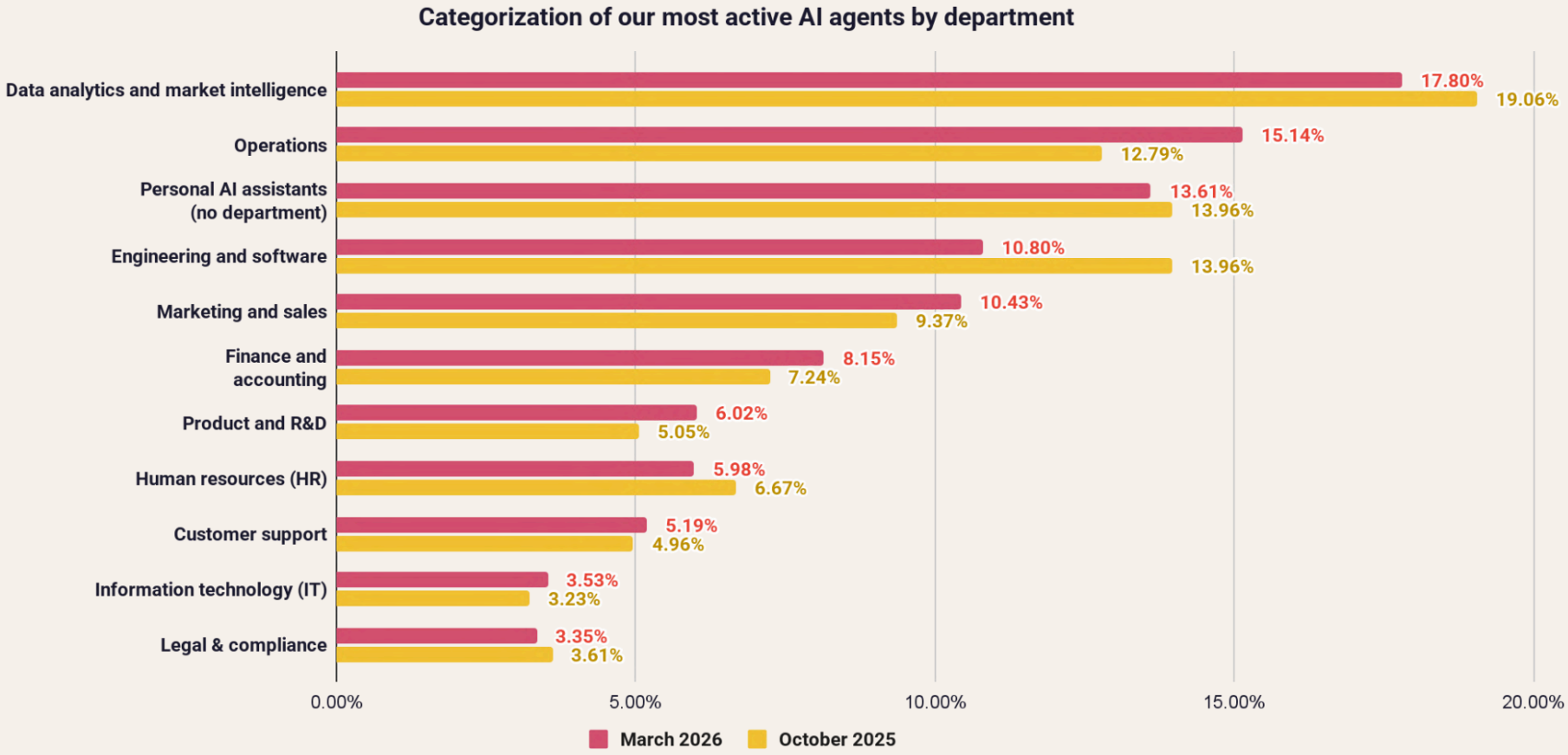
We are actively testing what happens when entire functions, departments or full organizations are led by AI, rather than simply layering AI agents on top of existing hierarchies and job descriptions. In this model, AI is oriented around desired *outcomes* of the workflow, rather than constrained by existing vendor arrangements, team structures, or approval chains. AI drives the desired outcome (like fulfilling all orders within two business days, sales growth over 10%/year, etc.) and only then do new kinds of human roles emerge that support that outcome.

There are significant challenges to work through before this concept can be implemented at scale. Stay tuned for a future Prosus research report, where we share the results of our work on autonomous AI-enabled organizations.

Part One: Insights from 60,000 AI agents around the globe

A quantitative evaluation: what do these AI agents actually do?

Below is a breakdown of all active agents by department. For comparison, we also included the data from October 2025, the month when Prosus' agentic AI initiative started exponential growth. 18% of all active AI agents are used by data analytics and market intelligence functions. While not every company has a distinct data analytics department, it is clear that employees are embracing AI's ability to crunch external data and internal data sets to look for business trends and actionable insights. Next, 15% of AI agents are used by the operations department for tasks like forecasting supply or inventory. The third most popular department is not a traditional corporate department. The third largest share of AI agents - 14% - serve as employees' personal AI assistants.



Across the 54 distinct AI tasks we have identified, what are the ten most common tasks?

Below is a list of the ten most common AI tasks. Of course, how you group tasks is extremely subjective (e.g., should *inventory forecasting* and *warehouse operations support* be distinct tasks?). Therefore, this top ten list is meant to provide an illustrative example of the kind of work AI agents perform. It's also important to recognize that the table below reflects the *total* number of AI agents, not ROI or business impact. We'll discuss agentic AI ROI in more detail later in this report.

AI Agent Use Case	Summary	Example	% of Agents
Review large data sets for business insights and actionable steps	Employees have built many AI agents to review huge data sets and uncover insights and possible opportunities.	The groceries team set up an AI to track different regions and to look for actionable improvements.	**9.6%**
Review and/or triage messages, reviews, and feedback	Teams built an AI designed to review large volumes of messages, online customer reviews, online comments, etc.	The car dealership division used an AI to monitor online reviews and comments to catch new issues right away.	**7.8%**
Marketing content generation	Marketing teams created many different AI agents to generate different types of marketing assets.	The small Bulgarian team set up an AI to help localize marketing assets while following the company's branding.	**4.3%**
Personal document and text creation assistant	Employees built custom AI assistants optimized to generate documents and text that meet their specific needs and desired tone.	A manager set up an AI to brainstorm how to frame their annual review feedback for each employee in the right way.	**3.5%**
Vendor-facing AI assistant	Teams created custom vendor-facing assistants to reduce admin work and/or to improve the experience of key partners.	A marketplace team set up a vendor AI to answer common questions, to free up more time to onboard new vendors.	**3.2%**
Product management lifecycle automation	Automates steps of the product management lifecycle, such as maintaining tickets, sprint reporting, drafting PRDs, etc.	A product manager set up an AI to help them monitor and maintain Jira tickets.	**3.2%**
Personal coaching and growth agent	Individual employees, specific departments, and HR departments created many AI agents that provide personal coaching.	An HR department created an AI to give coaching to employees who are looking to move to a different role.	**3.1%**
SQL and Databricks support for technical users	Helps technical users write, tune, and troubleshoot SQL and Databricks queries.	An engineer asks their AI assistant to generate a draft SQL query to determine the Q3 retention rate.	**3.1%**
Financial forecasting and scenario planning	Financial departments created AI assistants to crunch huge data sets to forecast future scenarios.	The finance team set up an AI to generate a first draft of the impact of five different interest rate scenarios.	**2.8%**
Schedule and meeting management	Employees use an AI assistant to help them manage their schedule and meetings.	Each morning the AI sends a daily meeting brief, based on reviewing emails and previous AI-generated notes.	**2.5%**

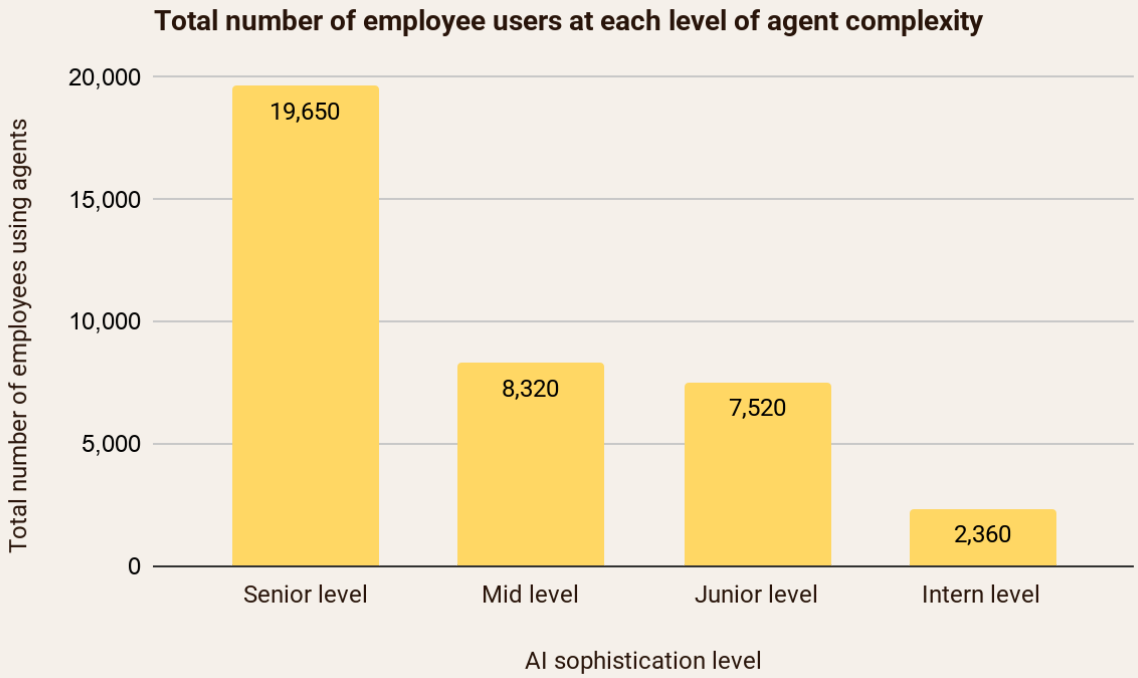
AI agents mirror the seniority levels of human employees

We ran an analysis of all our AI agents to determine how complex the agents are. We found that there are essentially four levels of AI agent complexity that roughly mirror the seniority levels of real-life employees. In other words, our framework that measures AI output emulates the level of output you expect from different levels of human employees. A simplified overview of these four AI seniority levels is outlined below:

- **Intern-level AI agents:** Agents that perform low complexity tasks and use 0-3 tools.
- **Junior-level AI agents**: Agents that generally handle low complexity tasks, but with more tool use than intern-level agents.
- **Mid-level AI agents**: Agents that handle moderate-to-high complexity tasks, using somewhere between 4-10 tools.
- **Senior-level AI agents**: Agents with extremely high tool use (11+ tools) and/or very complex tasks.

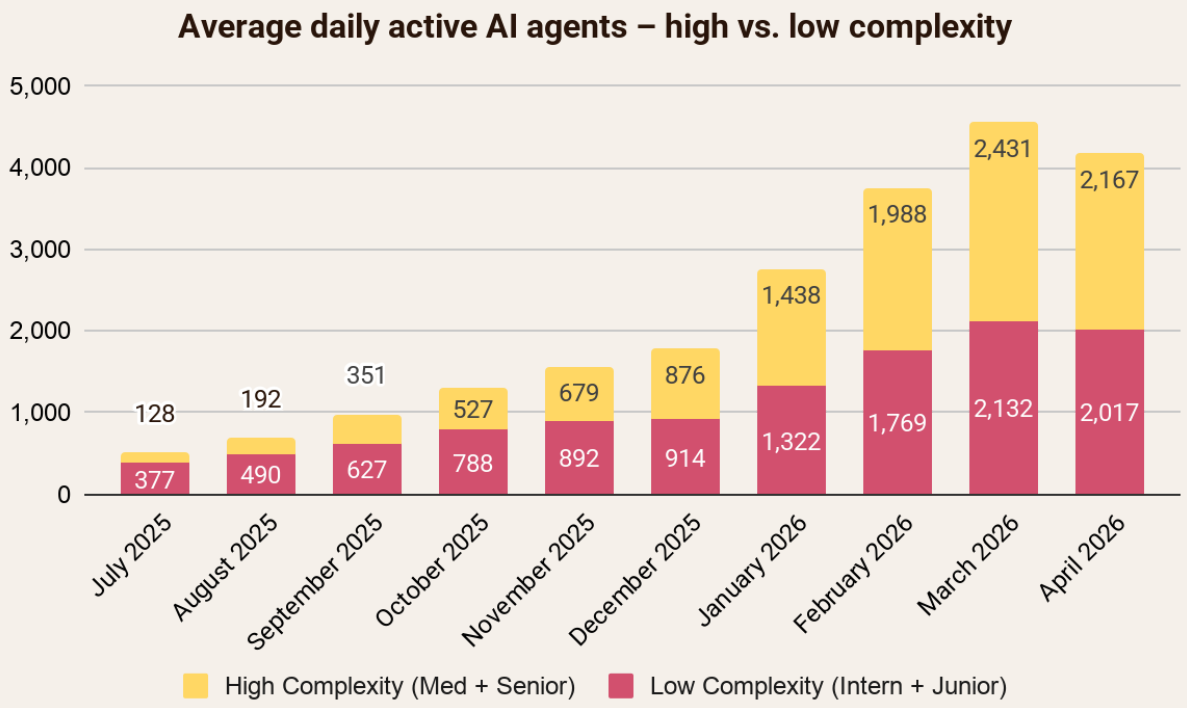
The classification is based on human and AI judgment, and is somewhat subjective. In addition, a higher number of tools is not a perfect proxy for sophistication. Nonetheless, we have found this four-tier framework to be helpful and directionally correct. It matches what we have heard from countless interviews with employees who use AI agents.

In fact, we have mapped the total number of employees using each level of AI agent seniority. Below is the total number of users by category as of March 2026. The more advanced AI agents have by far the highest number of users. This suggests that more complex agents that use many different tools provide the most valuable output to employees.



By focusing on daily usage, however, one can see the value of relatively simpler personal AI assistants

The chart below showcases average daily active agents broken down by High Complexity (Senior and Mid level) versus Low Complexity (Junior and Intern level). For the last six months, the daily usage split between High and Low Complexity agents has somewhat stabilized and hovers around a 50-50 split. This close-to-even split for daily active AI agents contrasts sharply with the dominance of High Complexity agents in the monthly view by the total number of employee users on the previous page. The daily average view showcases the value of these simpler agents for routine, day-to-day tasks. Relatively simpler agents are still valuable, just in a different way.

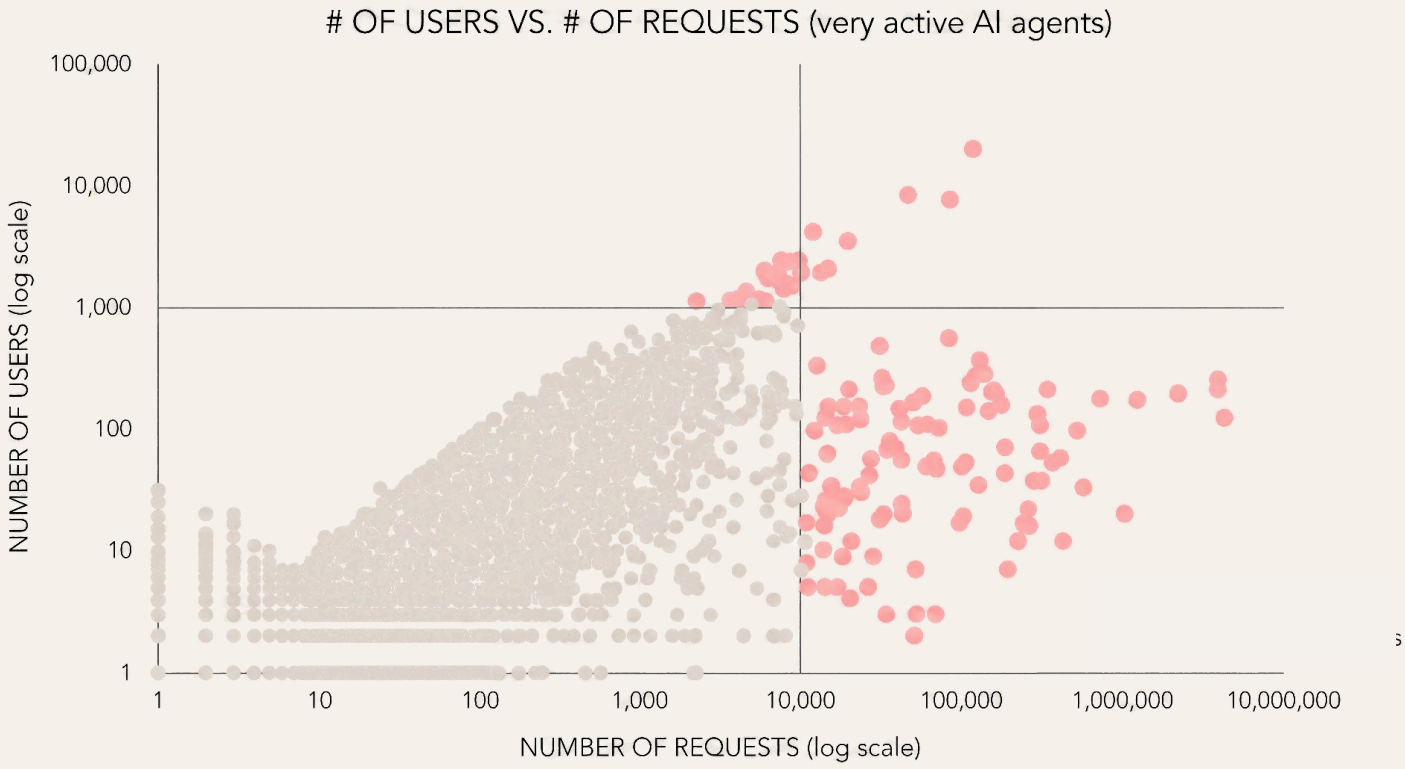


Our data set reveals an AI agent “power law”

After a year and a half of studying agentic AI growth across our portfolio companies, it is clear that the classic business “power law” effect can be seen in AI agents. We estimate that approximately 2% of the most active AI agents generate a hugely disproportionate impact for the business.[2]

You can see the AI agent power law when plotting agents on a chart

Below is a chart showing the most active AI agents in terms of the number of employees using the agent (y-axis) and the number of requests (x-axis). The vast majority of the active AI agents are clustered under 100 users and under 10,000 requests, but there is a smattering of outlier agents generating hundreds of thousands of requests and/or are being used by thousands of employees.



[2] Similar to the category-level breakdown mentioned earlier in this report, for this power law analysis we filtered out “dead” AI agents and AI agents that are infrequently used (e.g., only used once every three months or less).

What are some examples of power law AI agents at Prosus portfolio companies?

AI agents used to launch a new third-party affiliate marketplace

A set of AI agents handles onboarding, communication, and decision-making for a new marketplace for third-party affiliates to sell the company's products. Since the value of each individual affiliate partner is very low and onboarding can be cumbersome, this kind of marketplace was not economical before the rise of AI. The AI agents make decisions and escalate the unusual edge cases to a human employee as needed. **The portfolio company estimates this new third-party marketplace run by AI will increase the firm's revenue by $83M per year.**

An AI agent to help the "long tail" of small restaurants on the platform

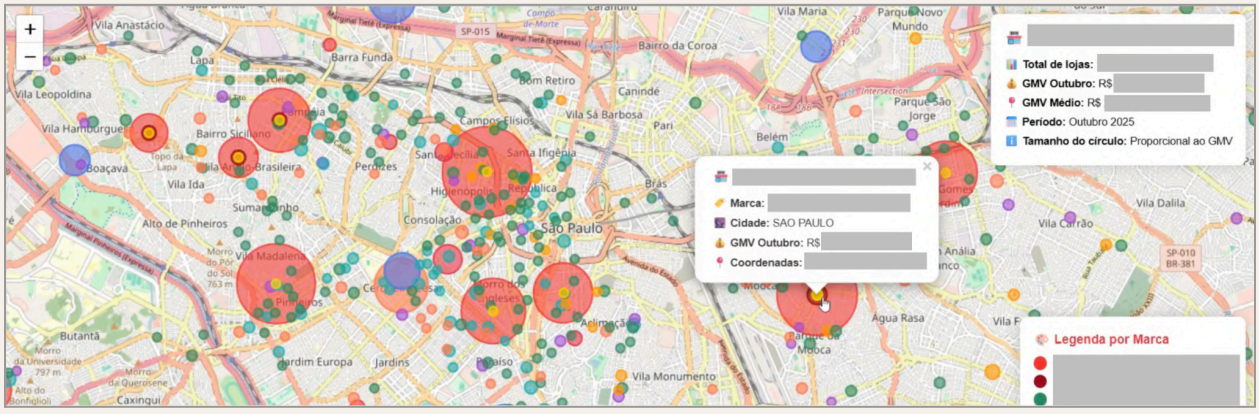
One of our food delivery portfolio companies has thousands of small "mom and pop" restaurants that could be positioning their offerings better. The low value of each of these small restaurants means it is not economical to assign human account managers. So the company built an AI agent to support the "long tail" small restaurants categories. **This AI drove a 119% increase in orders in the long tail category, and improved small restaurant retention by 73%.**

An AI agent that answers customers' specific questions about a vacation rental

A portfolio company created an AI agent designed to be able to answer customers' specific questions about a vacation rental. E.g., "does this beachside house have good Wi-Fi and a desk?" The agent reviews both internal data and external data/reviews from Google Maps and TripAdvisor to answer customer questions. On the flip side, the agent also sends insights to property managers. **Customers who chat with the AI have a 138.3% higher booking rate.**

An AI agent that supports the grocery division by generating custom insights/reports

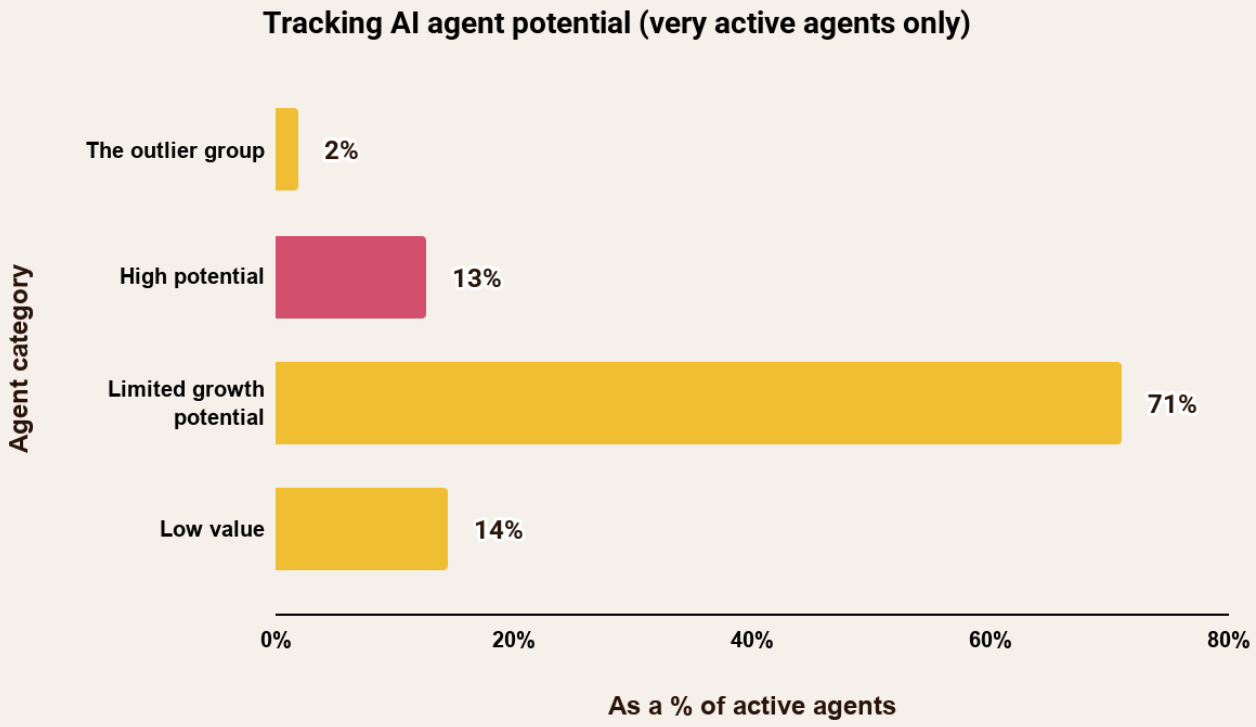
Employees ask the AI questions directly in Slack. The agent can answer questions about existing products and partners, generate reports and visualizations, evaluate performance, make future projections, etc. The groceries department estimates that this custom agent **saves the average employee 46 workdays per year** by removing the need to manually generate a custom report.



The AI agent for the groceries division can generate a heat map of grocery stores by revenue in a particular area.

It's important to track the high potential AI agents that could join the "power law" 2% group in the future

A separate but related finding to the "power law" trend is that your organization can also use data science to identify the high potential AI agents that could become a top 2% AI agent sometime in the future. Tracking AI agents over time reveals that there are tell-tale signs of potential. Roughly 13% of the most active AI agents are high potential. A breakdown is below.



Be sure to double check some "low value" agents that show promising adoption

Whether it's an AI or a human employee, evaluating "value" is extremely subjective, so as a second step, we also task the AI to identify agents that it determined to be "low value" despite having unusually high adoption. This exercise usually generates a small number of agents. For example, a recent run of this analysis identified 31 agents that the AI perceived to be "low value" despite having high adoption numbers. It's relatively easy for someone on the team to manually review this small group of agents and look for errors. More detail about our methodology and how we track and identify high potential agents can be found in the appendix.

20 “power law” AI agents that represent a good starting point for all organizations

In the previous section of this report, we noted how we have observed a “power law” effect in agentic AI. This power law observation comes from countless meetings with portfolio company leadership teams, business unit leaders, and the individual employees using our internal Toqan agentic AI platform. We have also conducted multiple surveys of top AI creators across Prosus’ 40,000 employees.

From our surveys and conversations with employees and leadership around the world, **we have noticed patterns across the various portfolio companies**. Despite differences between industries, geographies, languages, and no mandate from Prosus HQ to build agents that do X, Y, and Z, we noticed similar outcomes. Our portfolio companies almost always organically built the same 20 AI agents that solve a clear business problem. While these 20 use cases are certainly not a comprehensive list and they’re skewed towards the needs of a technology company, the list on the next page represents a good *starting point* for organizations looking to implement AI agents that can generate an immediate return.

These 20 use cases also provide a window into tasks that will likely be “default AI.” To encourage all Prosus employees to adopt these AI agents that will likely become “default AI” in the future, we built an AI marketplace into our internal Toqan agentic AI platform. The marketplace makes it easy for any employee to copy the template for top AI use cases. We’ll cover the importance of the “marketplace of top AI agents” in more detail later in this report.

20 agentic AI “power law” use cases that we recommend all organizations implement as a good starting point

AI Agent Use Case	Details	Example
Forecast product demand or shortages	AI crunches huge data sets to forecast product demand and/or flag potential supply/inventory shortages.	AI tries to predict when items will go out of stock, so the team can take it off the marketplace before customers complain.
Monitor churn risk and/or prevent churn	AI scans for customers who are showing signs of churn, then automatically launches outreach and a retention cadence.	The account team built an AI to look for signs of customers at risk of churning, and then proactively notify the team.
Customer loss analyzer	An AI that analyzes customer interactions and usage of the service to look for insights on why customers left.	The retention team set up an AI that reviews all losses over X timeframe to look for any patterns or insights.
AI assistant for sales reps	AI assistants can help gather intelligence, source opportunities, and close deals.	The sales team built an AI agent that can look for sales opportunities and generate a prospect briefing for the rep.
Review and prioritize prospect messages	Individual sales teams built AI assistants to review all inbound prospect messages for prioritization/triage and overall insights.	Leaders in the partnerships team use an AI to look for insights or trends on why prospects reach out to the firm.
Specialized customer AI assistant	An AI agent that can answer very specific questions about products in a conversational way, increasing the conversion rate.	The travel team set up an AI agent that can answer very specific questions about the vacation rental and the neighborhood.
Conversational AI for suppliers or partners	Conversational AI agents that can complete tasks to improve the vendor/partner experience and reduce costs.	The delivery team set up an AI assistant to onboard standard delivery riders faster, allowing them to focus on edge cases.
Scan for security issues and/or outages	AI to automatically scan the codebase to look for vulnerabilities or to understand the source of outages.	The IT team built an AI to continually scan for security vulnerabilities. The team reviews any issues identified by the AI.
Scan for anomalies and/or fraud	Business units use AI to monitor huge company data to look for anomalies, edge cases, and/or illegal activities.	The finance team uses an AI to review all payments to look for suspicious transactions or signs of fraud.
Supplier, vendor, or product risk assessment	AI helps large organizations track hundreds of suppliers and vendors, escalating issues to employees when needed.	The procurement team uses an AI to scan the internet and court cases to look for any red flags that the vendor failed to disclose.
Contract and invoice creation and/or review	AI can create first drafts of invoices and contracts and/or serves as the first reviewer of external invoices and contracts.	The finance team built an AI to review a large number of invoices for price changes vs. the previous month, quarter, and year.

Automating parts of product management	Automates elements of the product management lifecycle, such as maintaining tickets, sprint reporting, drafting PRDs, etc.	A product manager set up an AI that turns their brief summary into a first draft of a structured PRD.
Specialized news or event tracker	Monitors specific news topics across countries and/or scans for events relevant to the company.	The legal team uses AI to scan LATAM for any news related to freelancer regulations and potential trademark infringements.
Automated monitoring of competitors	AI can monitor competitor websites, social media accounts, court cases, and news headlines to surface important changes.	The product team set up a monitoring agent to track competitor changes to UI/UX, promotional offers, and social media activity.
Internal research assistant	Employees can ask an AI assistant to find the company's existing research and materials on any topic.	Employees enter a prompt and the AI scans company docs (with permissions) to look for existing research or decks.
"Self-serve" portal for data inquiries/reports	An internal AI assistant designed for non-technical employees. Lets them run a custom inquiry without knowing any SQL.	An AI assistant allows anyone in the restaurants division to run a custom report based on their question and interests.
Company-specific marketing/sales content	AI agents can be trained to produce high-quality sales and marketing material that follows the company brand and tone.	The central marketing team set up an AI to help different countries localize content.
Personal AI assistants	AI can manage personal calendars, provide custom meeting notes, send reminders, and create first drafts of content.	Individual employees built an AI to help them manage their work and create first drafts based on their needs.
Coding co-pilot	Arguably the best known AI use case that most companies have already adopted, it is still an essential productivity tool.	Engineers use personal AI agents to help them create, review and maintain code.
Specialized employee onboarding assistant	Large teams set up their own onboarding AI that can answer specialized questions and help new employees find things.	The engineering team set up an onboarding AI assistant that can answer new joiners' questions in a conversational way.

How do we measure agentic AI ROI?

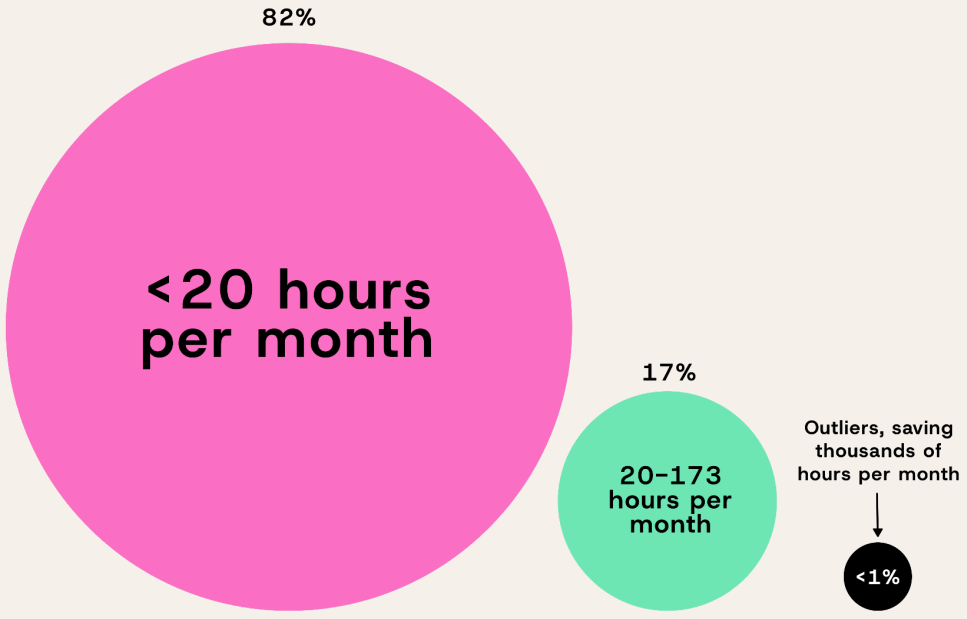
Measuring agentic AI ROI and tangible financial benefits is a hot topic. As companies spend more and more on AI – and worry that someday the leading frontier labs will be forced to significantly increase their costs to justify stratospheric valuations – business leaders are looking to justify the investment. At Prosus, we have found that bottom-up ROI estimates sourced from individual business lines are more effective than top-down estimates. Top-down estimates of the benefits of agentic AI generated by the Engineering or Strategy department are not very reliable.

In addition, we focus on quantifying the ROI of AI agents that are shared and used by multiple employees. Trying to quantify the value of every single personal AI assistant created by employees is not worthwhile.

With our bottom-up ROI approach focused on shared AI agents, we measure agents in three ways: time saved, increased revenue, or fixed cost reduction. Whenever AI ROI is difficult to quantify, we use the shockingly simple but effective "delete it tonight" test that we'll explain in more detail below.

Productivity agents generate clear ROI based on time saved. These agents typically fall into three tiers

These agents save a team a certain number of hours a month and potentially even the total headcount dedicated to manual and/or repetitive tasks. We have observed that productivity agents fall into three tiers: under 20 hours/month, between 20 hours/month and one FTE (173 hours/month), and a small "power law" group that does the work of dozens of FTEs.



The first tier – representing 82% of hours-based agents – is composed of agents with relatively modest time savings of fewer than 20 hours/month. This is usually a personal AI assistant that helps an individual employee be more productive. 17% of hours-based agents save somewhere between 20 hours/month and 173 hours/month (173 hours is the rough equivalent of one FTE). These are AI agents that automate tasks and free FTEs to spend more time on value-added tasks. Finally, there is the same "power law" effect in headcount-based agents, where a tiny percentage (<1%) perform the equivalent of thousands of hours of work. A simple sum of all these gains leads to over 1000 FTE time equivalents generated by these agents.

Value agents generate increased revenue and/or lower costs. We have noticed that value agents can also generally be grouped into tiers

AI agents that are tied to revenue or costs have a clear benefit. Financial benefits range from under $1M (e.g. reducing annual audit costs) to $1M–$10M (e.g. an enterprise churn reduction agent) to a small number of outlier agents in the tens of millions. An example of an outlier revenue scenario is a portfolio company using a set of AI agents to manage communications and onboarding for a new marketplace for third-party affiliates to sell the company's product. Since the value of each individual affiliate partner is very low and onboarding can be cumbersome, this kind of marketplace was not economical before the rise of AI. The company projects that the new marketplace for third-party affiliates will generate $83M/year in revenue.

Of course, depending on the size and nature of your organization, something like a churn reduction agent may be more valuable than a third-party affiliate marketplace. The specific examples are less relevant here - the more important finding is that Value agents generally fall into three different levels of financial benefits.

We use the "delete it tonight" test to determine the ROI for hard-to-assess AI agents

Unfortunately, many AI agent use cases represent some kind of new activity or task that cannot easily be tied to a headcount or revenue/cost-based estimate. We ultimately realized that the best way to determine the ROI of these AI agents is to ask the leader of that business unit a simple question: "what would happen if we deleted this AI agent tonight? And your team could never use this AI agent again? What would happen to our revenue and/or costs?" When you frame it this way, most business leaders can estimate some kind of tangible revenue, cost, or time saved benefit.

ROI is merely the first step – doubling down on what works is what really matters

We aggregate up the ROI estimates of the various shared AI agents. While the ROI quantification exercise is important (to make sure the benefits outweigh the costs), we view that as merely the first step. Calculating the ROI is just the starting point. The real work and value of agentic AI is to figure out how to "double down" on those "power law" AI agents that provide the most benefit. The 20 AI use cases highlighted earlier are likely worth "doubling down" in our organization, but your firm will undoubtedly have other "outlier" AI agents specific to your business.

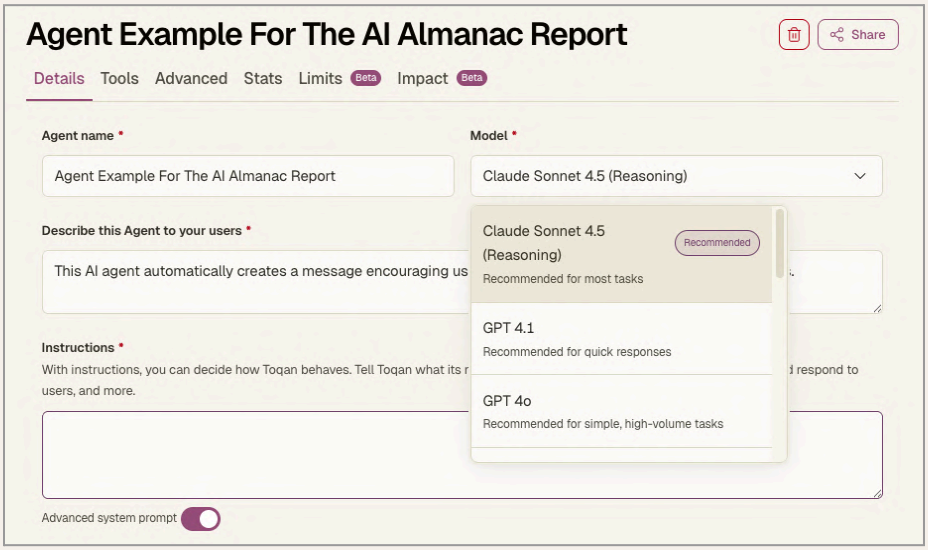
Our observations on whether AI models impact agent performance and ROI

The AI space moves very quickly. There are seemingly constant headlines about new cutting-edge AI models, models blowing past old benchmarks, etc. Given all of the attention on the various AI models, we wanted to share our experiences and how they impacted agent performance and ROI.

Adding support for Claude Sonnet 4 was a big breakthrough. Since the rollout of Claude Sonnet 4, however, we have noticed diminishing returns on new models.

The internal Toqan agentic AI service we offer to Prosus' portfolio companies and 40,000 employees around the globe lets users choose from a selection of AI models. At launch, Toqan offered access to one AI model, but that has now grown to a selection of ten different leading AI models. Over the course of continually adding models to the platform, we have noticed only one seeming "big bang" moment where adding a new AI model led to a sudden and meaningful improvement. When we launched support for Claude Sonnet 4 – with Sonnet's strong reasoning and tool calling capabilities – many of our users reported that they were able to create new (or at least more effective) AI agents that were not possible/reliable before Claude Sonnet 4.

Since the rollout of Claude Sonnet 4, we have noticed diminishing returns on new models. There are some very complex use cases where cutting-edge models are still needed, but most top AI models are now "good enough" for most agentic tasks. And open source models (like Kimi K2.5), are becoming increasingly competitive. Of course, there are still some differences between model capabilities. For example, as of this report writing, we have measured that Kimi K2.5 is on par with the cutting-edge models at tool chains, but it is worse at code execution. But these model nuances are not that relevant to most users.



Selecting AI models within the internal Toqan Agent Platform. Brief descriptions of the value of each model is provided to help users determine which model to pick.

Despite the constant release of new models, most users do not want to bother changing the model powering their AI agent(s)

We have found that once employees have set up an AI agent that works well, they no longer care about which model they use or the latest developments at frontier labs. Users prefer stability over experimenting with a new cutting-edge model to look for marginal improvements. New models are generally only relevant to employees who are still trying to solve a very complex agentic AI use case they have not managed to get to work yet. Of course, some future breakthrough in AI could see a relatively sudden mass migration to a new AI model, but at least for now, our observation is that most users rapidly lose interest in models once their agent(s) starts working properly.

Users' preference for stability means the main impact AI models may have on ROI is on cost

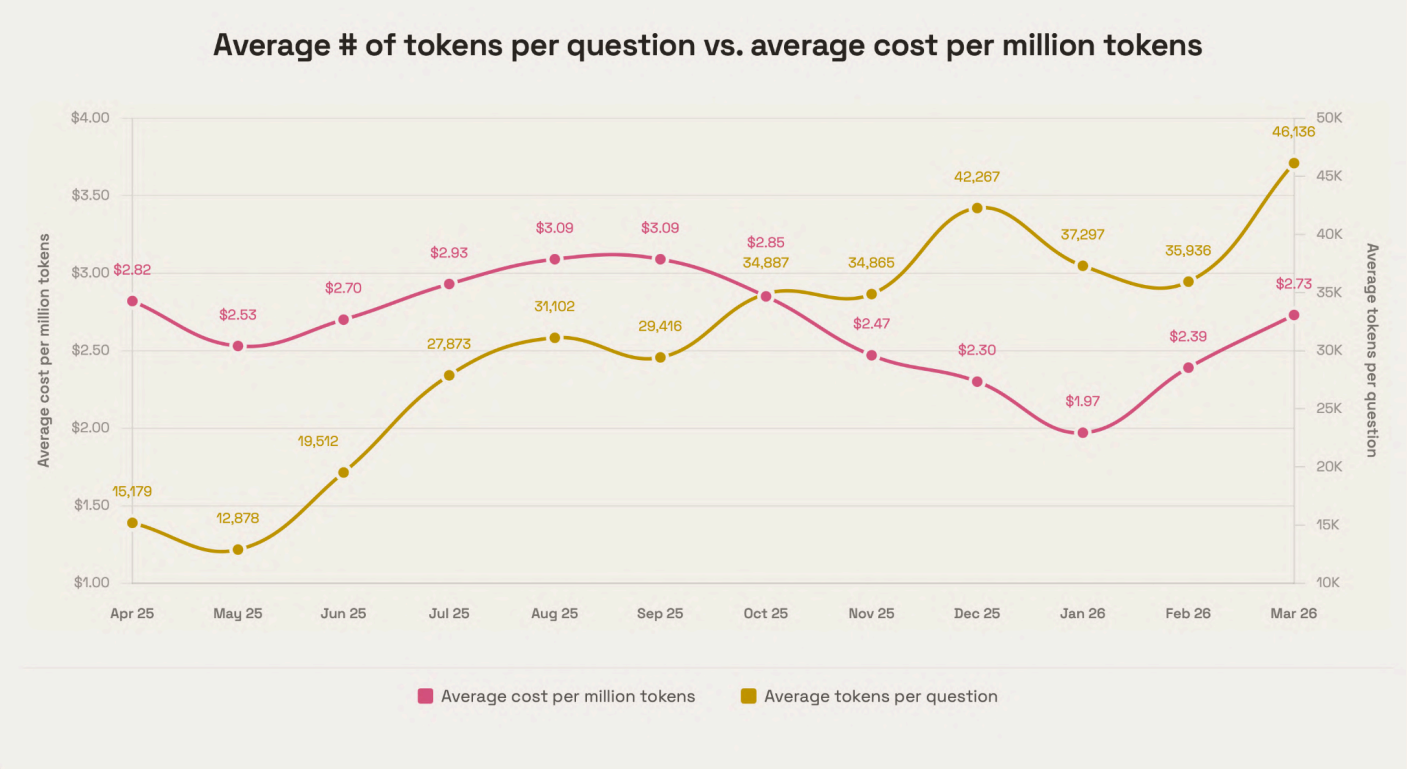
The downside of the above "if it ain't broke, don't fix it" mentality we see from most of our agentic AI creators is that it can be hard to get employees to think about cost tradeoffs. AI models continue to improve, to the point that many tasks can be accomplished with any of the leading AI models. Most agentic AI use cases do not need the most expensive cutting-edge model from OpenAI or Anthropic. That means the model race may impact cost more than anything else - and we'll discuss managing and minimizing cost in more detail on the next page.

AI costs are very hard to predict, so focus instead on empowering individual teams

We have found that managing and projecting AI costs at scale is difficult. There's a lot of complexity when there are tens of thousands of AI agents using different AI models (reasoning models are generally more expensive), different infrastructure setups, and different levels of tool use. In addition, any changes to the AI agent platform itself, user behavior, and/or the overall business priorities can suddenly change AI costs.

There have been many headlines in the industry about how AI costs are – or will – inevitably go down. In our experience, however, several factors mean we have not seen a clear downward cost trend. First, while open-source models are cheaper, the infrastructure setup behind them can be expensive at scale, especially when you need both 24/7 reliability and redundancy. Second, AI agents that handle complex tasks need sequential tool calls, and these kinds of agents need the best and most expensive models.

To illustrate these challenges, below is a breakdown of the average number of tokens per question (yellow line, right y-axis) and our average cost per a million tokens (pink line, left y-axis) from April 2025 to March 2026. There is no straightforward relationship between the average number of tokens per question and the average cost per million tokens. And token consumption is trending up, increasing overall costs.



We controlled costs by setting up a two-tier approach to AI usage and empowering each business line to make cost decisions

Given the complexity associated with predicting and modelling AI costs, we set up a two-tier structure that encourages experimentation while putting some guardrails in place. The internal Toqan agentic AI platform can be used by any of Prosus' over 40,000 employees around the globe free of charge as long as the user doesn't exceed 200 requests per hour via an API. If someone wants to exceed 200 requests per hour, then they are moved to the paid Production Tier. The business line approves Production Tier usage, and the individual business then pays for that cost. In addition, some kind of limit or cutoff prevents accidental spending. Sometimes people who are new to AI accidentally create an unnecessarily expensive AI with very intensive tool calls.

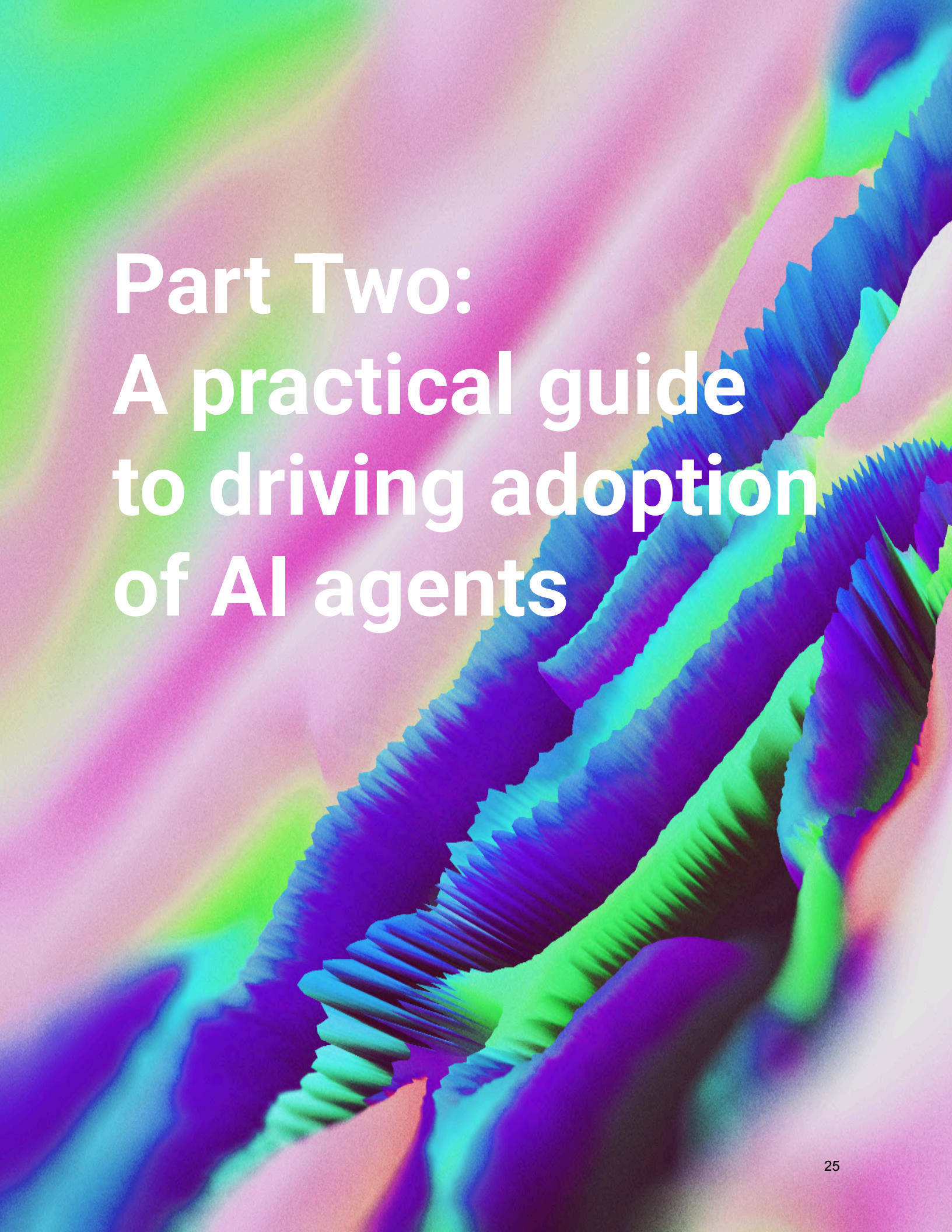
The exact cutoff is not important – your business may "draw the line" at a different number. The philosophy behind a tiered approach where the business line approves the cost is what matters. We have found that any centralized attempt to mandate X or Y cost saving approach across thousands of AI agents is not very effective. A central AI team does not have the nuanced understanding of a particular business line's needs, and the user is the one who is testing which model works best for their particular needs. Therefore, we empower each business unit to make their own decision about costs.

We also optimized our internal AI agent platform to control costs

Depending on the size of your company, your organization may or may not be building AI capabilities in-house. For those organizations that do have an in-house AI team and internally built AI services, a handful of optimizations can help you reduce costs. First, build an admin panel that makes it easier for teams to monitor AI agent usage, approve Production Tier agents, set limits, etc.

Second, allow the agent to parallelize tool calls when appropriate. For a simplified example, if the AI agent's task is to search the internet for public information on X, then that agent should perform five online searches at once, rather than five sequential searches performed one after another.

Third, focus on effective in-conversation caching to reduce costs. Caching can break relatively easily based on simple elements like a mention of the data or time. Finally, consider setting the threshold for compacting (the process of summarizing older context) much lower than the model's actual maximum input size. Our internal platform currently starts compacting at 60% of the model's maximum input. This helps decrease costs with the added benefit of improving speed and reliability.



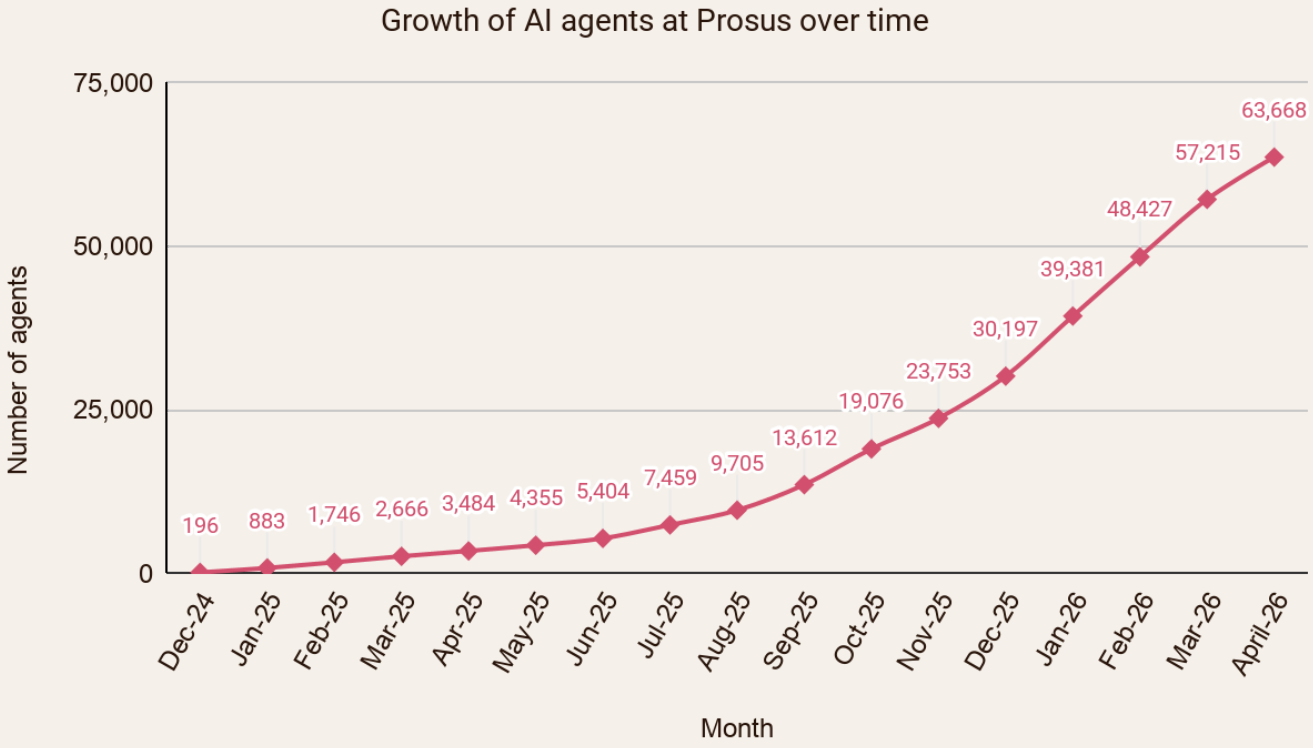
Part Two: A practical guide to driving adoption of AI agents

Driving adoption of AI agents within a traditional corporate structure is a slow and difficult

At Prosus, the journey to over 60,000 AI agents has been a long and bumpy one. While the overall trend line (pictured below) has been “up and to the right,” we learned many things along the way. We eventually realized the correct approach to getting a workforce to adopt agentic AI en masse is a three-step process.

The first phase is the groundwork that needs to happen to set the agentic AI initiative up for success. Think of it as everything that needs to happen before the first month on the chart below. The second phase is focused on driving the initial adopters. That’s the first 10 months in the chart below. The third phase is the scaling phase. At Prosus, scaling up started in fall 2025. We’ll conclude Part Two of this report with a high-level checklist that summarizes all of these steps in a digestible format.

Driving employees to build and use AI agents is the first step towards the ultimate goal of organizations where each department (e.g., sales, operations, etc.) is led by AI and managed by a radically different type of human workforce. The real potential of AI lies in reimagining entire corporate departments from the ground up based on AI. Realistically, however, most organizations are not ready or willing to jump straight from their current low levels of AI adoption into redesigning entire departments to be run by autonomous AI agents. Therefore, we encourage all companies to follow this simple guide in order to drive meaningful adoption of agentic AI and take the first step towards the future.



Phase One: Establishing the agentic AI initiative

The first step of transitioning your company to an AI-driven organization is to lay the groundwork for success. Without the right setup, agentic AI initiatives will fail thanks to inertia, competing priorities, cost concerns, and more.

Executive-level sponsorship of the initiative, and AI adoption becomes an OKR

The highest level of the company – the CEO and the c-suite – need to be the ones pushing and prioritizing agentic AI adoption. Without executives pushing the initiative, their deputies will not prioritize agentic AI and the needed integrations and standardization efforts (see below) to make agentic AI successful. Agentic AI adoption also needs to become one of the key OKRs that drives bonuses and promotions. Across Prosus' many portfolio companies, we have seen that agentic AI initiatives that do not have executive-level buy-in tend to fail, or at least progress very slowly.

The needed integration and standardization work must be prioritized

AI agents that are not connected to internal company systems cannot provide much value. The business and engineering teams need to work together to determine which integrations are needed, how to prioritize the various possible integrations, and when "read" vs. "read and write" access will be given, etc. Beyond integrations, AI adoption also requires departments to undertake some standardization work. In many departments, employees are producing slightly different versions of outputs based on their personal preferences. Therefore, automating parts of their workflow requires all team members to agree on a standardized output that satisfies everyone.

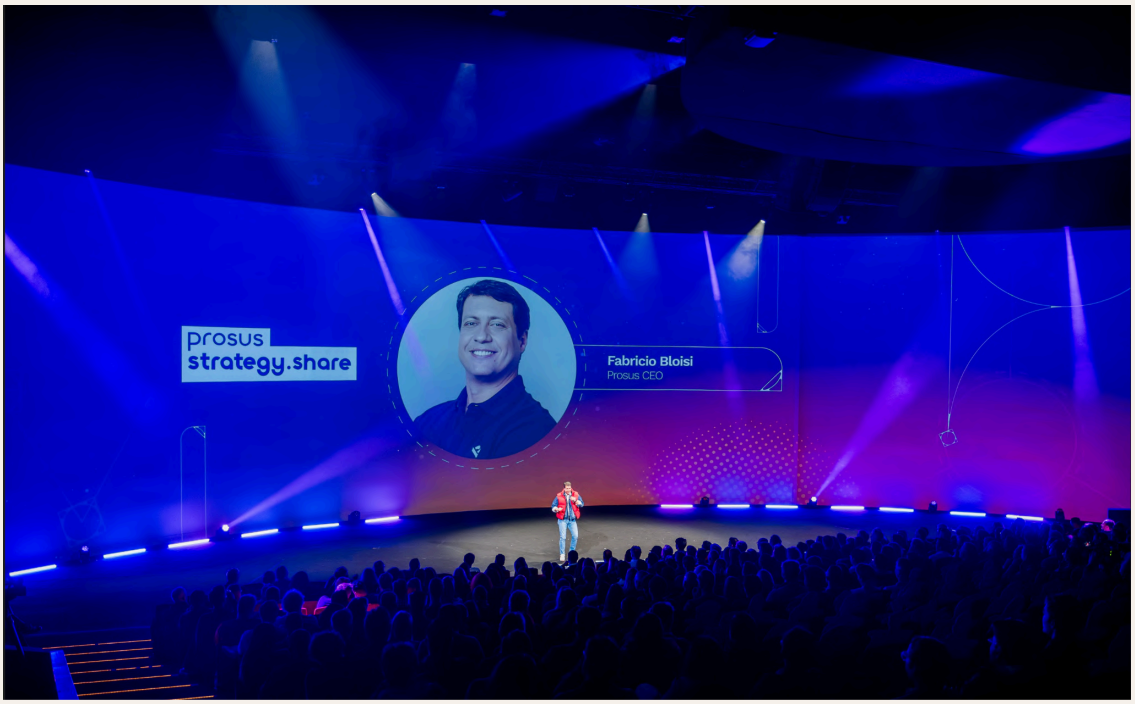
To start, offer free (or at least heavily discounted) pricing to encourage experimentation

AI costs can be a deterrent to adoption. Managers may be reluctant to sign up to new costs with an uncertain financial gain. Therefore, it is important that the company's agentic AI initiative offers free or discounted costs below a certain usage threshold. You want to encourage employees to freely experiment with AI, and not be discouraged from testing complex tasks that burn a lot of tokens. Once your company is more mature and has AI adoption at scale, then the firm can slowly phase in costs and a production tier (discussed earlier in the report) to reduce waste.

Set up the full-time dedicated roles needed to support agentic AI growth

Driving AI adoption across a large organization requires resources with the capacity to provide training and technical support. We have found that two types of roles are needed. First, each department needs dedicated ambassadors. Ambassadors are existing employees in each department who are trained on agentic AI, and then use their knowledge of both their department and agentic AI to help encourage adoption.

And second, depending on the size of your organization, you will need to hire 1-3 full time technical support specialists. The need for full-time technical specialists is self-explanatory. Your Engineering department is likely already using AI. The benefits of agentic AI will come from getting non-technical employees to use AI agents. You will need technical experts who can spend time holding training sessions and troubleshooting with employees and ambassadors.



The CEO of Prosus – Fabricio Bloisi – about to give a presentation on the importance of embracing agentic AI at Prosus' bi-annual internal company meeting

The AI Ambassador program at Prosus Portfolio company JustEat Takeaway (Amsterdam)

Phase Two: Driving the initial adopters

Find and identify the early adopters who will become the power users

AI Adoption will be relatively slow in the beginning. We have found that the key to kickstarting the growth of AI adoption is to find and nurture the early adopters in each department. You will need small, high-touch workshops and internal hackathons to really train these early adopters on how to use agentic AI. Said another way, you need to "lock them in a room" for a day or two, have them test out dozens of potential use cases, and find the ten use cases they can start implementing relatively fast.

Another key step to drive early adopters is to set up a dedicated Slack channel (or equivalent) for power users so they can meet each other, swap tips, and get inspired. The organization should give gifts or incentives to the early adopter community. For example, as a reward for their efforts, Prosus paid for some of the top early adopters to travel to Paris for an AI hackathon.

Vinicius is an example of an early adopter at Prosus portfolio company iFood in Brazil. Left, he is filming training videos for iFood. Right, he was selected for a free trip to Paris to attend an AI hackathon.

What is the profile of the early adopters?

Power users are best described as those who are enthusiastic about AI and finding ways to automate the tedious parts of their job (they have continuous and high frequency of usage). Power users do not necessarily need a degree in computer science – and your corporate messaging needs to emphasize that AI agents are designed and optimized by writing prompts, and anyone can create AI agents. Each business unit (driven by the ambassadors) needs to send out communications to the team asking for early adopters and beta testers who want to join a department workshop or hackathon. These early adopters will eventually become power users that drive AI growth at the company.

The power user community and ambassadors will be the main driver of AI growth in the second phase of growth

One of the key learnings from our experience going from zero to 60,000 AI agents across the portfolio is that power users and ambassadors within the department are the key to getting the average employee to test out agentic AI. Executive buy-in is needed to get the business to prioritize AI, but what actually drives AI adoption is employees seeing their colleagues creating and using AI agents. We have heard a variation of this story many times: "*once my colleagues started seeing my AI agent automatically completing annoying tasks in Slack for me, without me lifting a finger, they started asking me lots of questions about how they can set up an AI agent.*"

Employees often find AI intimidating, assume they need a PhD in computer science to be able to create an AI agent, and are afraid of sounding "dumb" by asking a "basic" question about AI. Seeing their non-technical colleagues successfully using AI helps change attitudes and perceptions, and they are less nervous about asking "basic" AI questions to colleagues.

In particular, power users and ambassadors should be the ones leading internal workshops, training, and hackathons - not consultants. We have observed that employees react much better to their peers "that look and sound like them" explaining how they use agentic AI for very practical use cases related to their department and business lines. We have found that bringing in consultants or an external group to lead a workshop or training session generally does not lead to big uptake of AI agents among average employees.

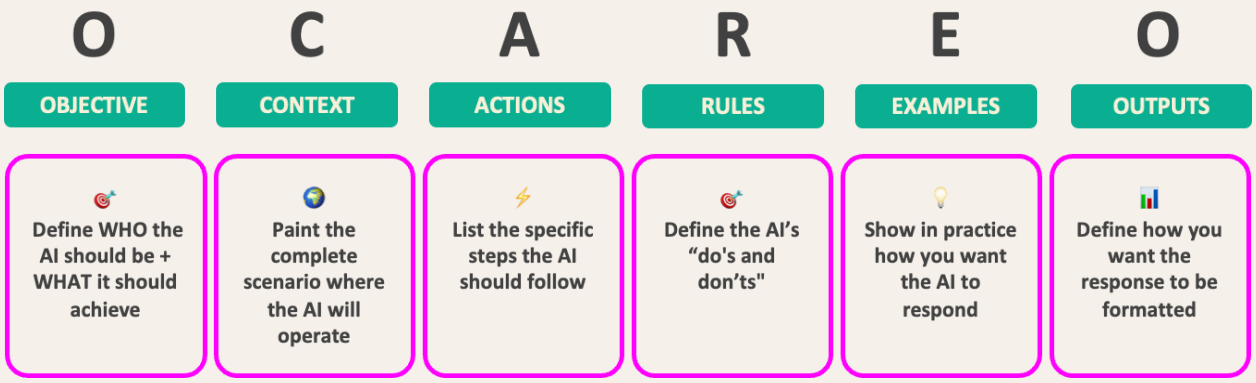
AI Agent workshops at Prosus portfolio companies PayU in India (left) and iFood in Brazil (right)

Workshops, trainings, and hackathons need to get prompt training right

As someone who is voluntarily reading a detailed report on agentic AI, you are likely already very familiar with how to write good AI prompts. But the average employee, in, say, the compliance department has no idea how to write a good prompt for an agentic AI. Workshops, seminars, and internal hackathons must reserve a good amount of time for how to write good AI prompts. You will need to provide training and guidance around how to communicate with AI agents and how to write good prompts.

Here are two best practices for training non-technical employees on how to think about writing a good prompt. First, all training and workshops should involve an exercise where employees partner up and write out what they would like an AI agent to do on a piece of paper. Employees must then show their prompt to their co-worker. The colleague then tries to summarize what an agent would do. The partner colleague in the exercise is usually *way off* in terms of their interpretation of what their co-worker would like the agent to do. This exercise helps everyone realize that whether it's a human or an AI, no one can read your mind. It helps drive home the message that AI prompts need to be very thoughtful, detailed, and specific.

Second, it helps to provide employees with a clear framework for how to write a good AI prompt. We have found that the OCAREO framework (below) is a particularly effective prompt-writing methodology, although there are certainly other frameworks you could use. Whichever framework you choose, it's important to provide a specific prompt-writing framework to employees.



Putting the time and effort into making good video training content is essential

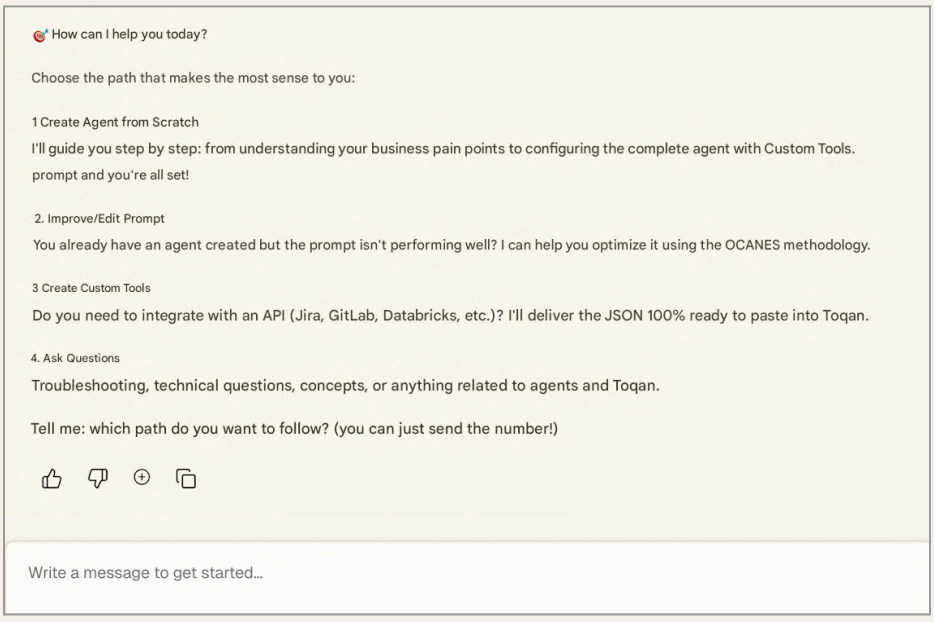
While this may seem like a very basic statement, we have found that creating a series of short video walkthroughs of how to use the agentic tool is extremely helpful. Video walkthroughs take more effort to create and maintain (as they need to be updated as the product and interface changes), but they are far more effective than any kind of text-based "how-to" guide. Your organization will need to create training videos on a wide range of topics – from connecting a Gmail account to troubleshooting integrations with ERP systems.

We see a noticeable difference between our portfolio companies that take the time to make high-quality video walkthroughs and our portfolio companies that only make a simple text-based FAQ-type page on how to create AI agents. Video guides where employees can watch how one of their peers set up an AI agent are incredibly helpful and lead to higher adoption.

Create a "setup assistant" AI agent that can help employees start using agentic AI

This section of the report on driving the initial adopters has mostly touched on many of the classic change management techniques, such as training and workshops. But AI can and should also be used to help solve the problem of most employees not being familiar with how to properly use AI. Your organization can develop an AI agent designed to help employees create the AI agent they have in mind.

Beyond the need to help the average employee with prompt engineering, non-technical employees may need help with topics like complex JSON configurations, conflicts (e.g., conflicting data in Slack vs. Google Sheets), which tools to use, troubleshooting, etc. An AI agent can "co-create" with employees and help them overcome these obstacles. Co-creation AI agents are also optimized based on the company's technology stack and are designed to follow a certain approach (like the OCAREO prompt framework mentioned earlier). In fact, the single most popular AI agent (in terms of number of users) at one of Prosus' biggest portfolio companies (the Brazilian delivery giant iFood) is a co-creation AI agent.



The iFood Optimiza AI helps employees create agents. The first step in the Optimiza AI flow pictured above (translated from Portuguese to English) is asking the user what they are looking to accomplish.

Over 1,300 iFood employees have used this co-creation AI agent to create over 10,000 AI agents for iFood. This demonstrates the value of co-creation AIs to help drive Agentic AI adoption at your organization.

Phase Three: Scaling up to thousands of agents

The third phase of driving agentic AI adoption is scaling up by incentivizing the entire organization to adopt AI technology. Broadly speaking, there are four key steps to scaling up.

Bi-weekly meeting with executives to review AI agent adoption data and efforts to "double down" on what works

At the scale up stage, there needs to be a weekly, bi-weekly, or monthly meeting to review the data and the numbers with executive leadership. How many agents were created over the period, are the average usage figures going up, etc. etc. The exact cadence and the exact executives that are involved in the meeting depends on your organization. But we have observed that our portfolio companies with a periodic data-based review of AI adoption with executives have far more success at scaling up and rapidly increasing AI adoption. Said another way, if there isn't a regular, dedicated meeting with executives focused on scaling up AI, it's unlikely it will happen as fast as you would like.

This regular meeting with leadership also needs to include progress reviews around your organization's efforts to nurture the high potential AI agents and to double down on the "power law" AI agents that deliver outsized benefits. Your organization needs to make sure time, attention, and resources are invested in the top AI agents that generate outsized returns.

Run internal contests with a big prize to drive adoption

One of the surprising learnings from our efforts to drive AI adoption across Prosus' 40,000 employees is that an internal contest was incredibly effective at getting employees around the company to finally "take the plunge" and carve out time to see if they could use agentic AI.

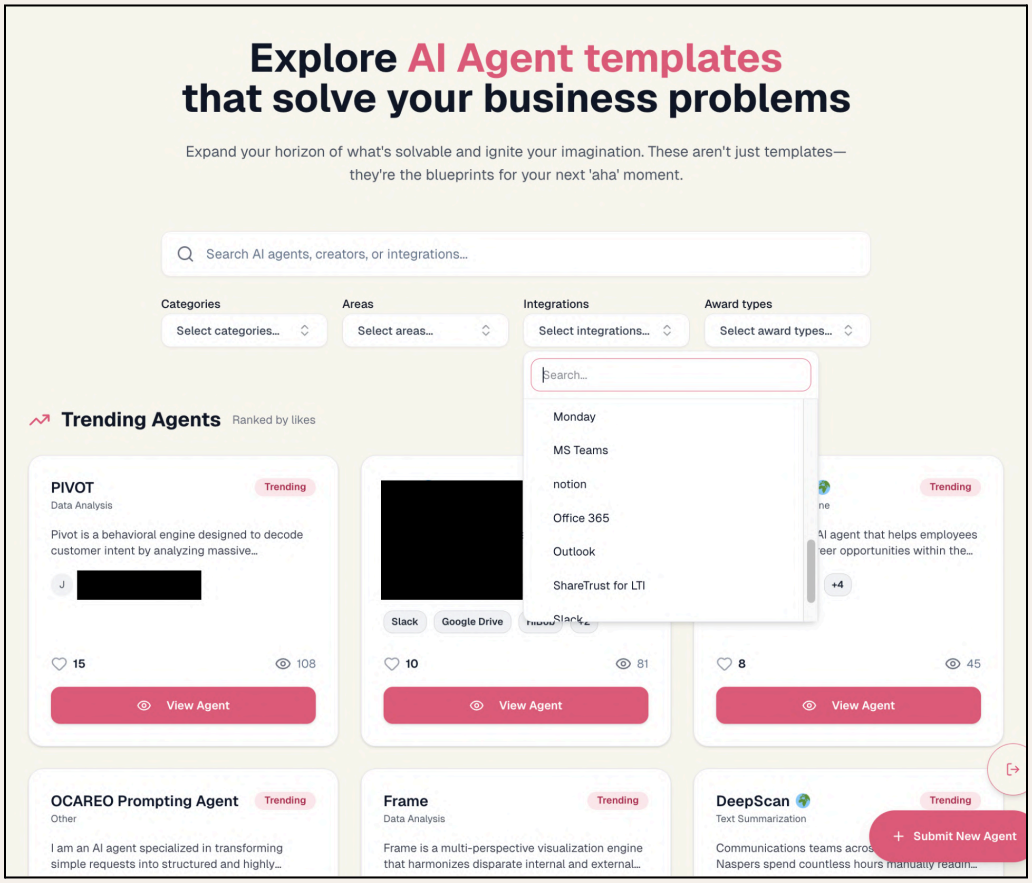
We held a two-step contest. First, each Prosus portfolio company held an internal competition to pick the best AI agents at the company. While there were some slight variations between companies, most companies picked 12 finalists for their company-wide grand finals. The top three winners from those 12 finalists received cash prizes ranging from $1,000 to $25,000. The winners were flown to China to compete against the winners from other Prosus portfolio companies in April 2026. We highly recommend any company that is serious about driving agentic AI run an internal contest. Some pictures from this competition can be found on the next page.

Pictures from Prosus' internal AI contests

Clockwise from left: A picture of the teams participating in the Prosus-wide grand finals (China), the finalist trophies for the Prosus home office (Amsterdam), the finalist trophies for Takealot's company contest (South Africa), the winners of the iFood AI agent contest (Brazil), and the finalist trophies for EMAG's company contest (Romania).

“Productize” the best AI agents by launching a marketplace of top AI agents that employees can access

We have hit on the importance of doubling down on the most valuable AI agents many times in this report. One of the best ways to amplify the value of the best AI agents is to “productize” top AI agents and make them easily available to other employees. Productizing top AI agents means creating a marketplace of agents that any employee can access. This marketplace offers “off the shelf” templates and prompts that can easily be customized. Users can search the marketplace, filter by categories, etc. An example of Prosus’ internal Toqan agentic AI platform is below. A searchable marketplace (including by department and by integration type) helps drive organic adoption of the best AI agents.



The Toqan marketplace of top AI agents

The summary checklist for driving agentic AI

To conclude Part Two of this report, below is a high-level checklist that summarizes the key steps involved in each phase of driving AI adoption.

Phase One: Establishing the agentic AI initiative

Executive-level sponsorship of the initiative, and AI adoption becomes an OKR	☐
Prioritize the needed integration and standardization work	☐
Agree on free (or at least heavily discounted) pricing to encourage experimentation	☐
Set up the full-time dedicated roles needed to support agentic AI growth	☐

Phase Two: Driving the initial adoption

Find and nurture the early adopters who will become the power users	☐
Trainings and hackathons are led by the power user community and ambassadors	☐
All workshops include the proper prompt training	☐
Create a series of video "how to" guides that explain agentic AI	☐
Launch a “co-creation” AI that can help employees build an agent	☐

Phase Three: Scaling up to thousands of agents

Cadence with executives to review adoption data and efforts to “double down"	☐
Set up an internal contest with a big prize to drive adoption	☐
Create a marketplace of top AI agents to "productize" the top agents	☐

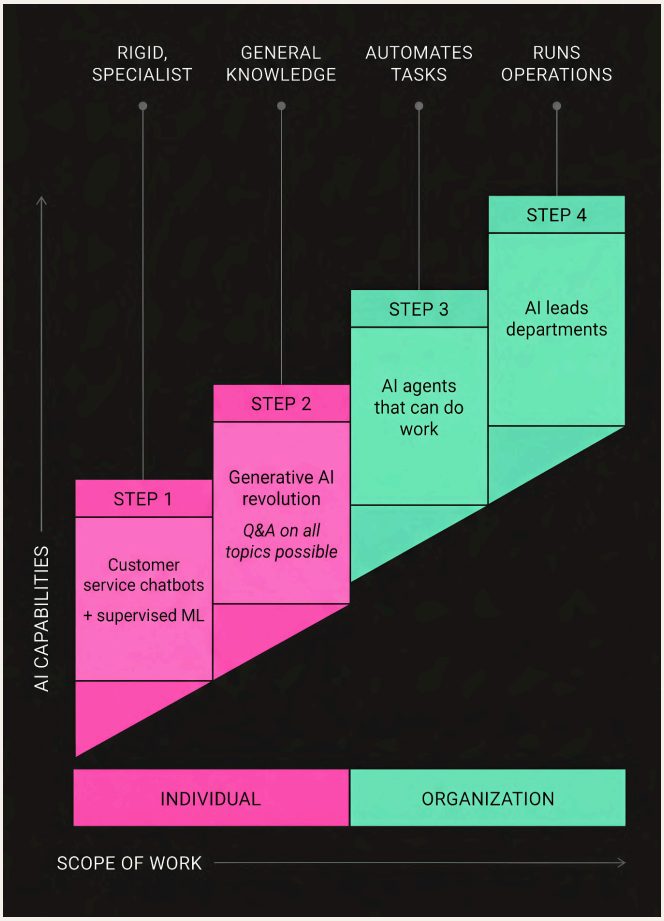
Conclusion and report authors

Conclusion: The AI-first company of the future and rebuilding organizations from the ground up

After scaling to 60,000 AI agents, we now have a clear view of what an AI-driven organization looks like in practice and what the future holds**.** What we observe across our portfolio is transformation rather than total disruption. We see three trends. First, individuals and small teams are building AI agents to work more efficiently. Second, every company is adopting the same "quick win" AI use cases that will likely become "default AI" in the future. Third, portfolio companies nurture the "power law" agents that create a competitive advantage.

This report focuses on implementing AI within existing ways of working — but that is just the starting point. The deeper transformation lies ahead: the rise of autonomous, AI-enabled organizations, where core business functions like sales, customer support, and operations are largely executed and coordinated by networks of AI systems, with minimal human oversight.

Comparisons are often made to how the invention of electricity impacted factories. Electrification delivered its largest productivity gains only after firms reorganized the entire production process around electricity. Simply substituting steam motors or water wheels with electric motors in existing factory designs offered only relatively smaller gains. The same framing is applied to AI. Proponents argue that it is only when companies start rebuilding core functions from the ground up with AI in mind that businesses will experience the biggest gains from AI.



We are actively testing what happens when entire functions, departments or full organizations are led by AI, rather than simply layering AI agents on top of existing hierarchies and job descriptions. In this model, AI is oriented around desired *outcomes* of the workflow, rather than constrained by existing vendor arrangements, team structures, or approval chains. AI drives the desired outcome (like fulfilling all orders within two business days, sales growth over 10%/year, etc.) and only then do new kinds of human roles emerge that support that outcome.

There are significant challenges to work through before this concept can be implemented at scale. Stay tuned for a future Prosus research report, where we share the results of our work on autonomous AI-enabled organizations.

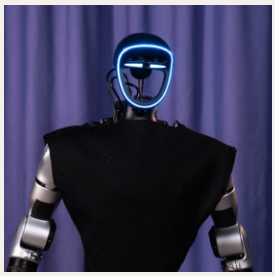
Meet the report authors

Euro Beinat
Lead author
EVP and Global Head of AI

Grant Easterbrook
Lead author
Entrepreneur in Residence

Asad Ismail
Senior AI Engineer

Bartosz Hawelka
Senior Data Engineer



Darwin The Robot
Training for the 2027 Beijing marathon

Donné Stevenson
Senior AI Engineer

Eduardo Corazzin
Director HR Products & Analytics

Ioannis Zempekakis
Director of AI and Data Science

Isha Argawal
AI Engineer

Jeroen van Dijk
Data Scientist

Maggie Konstanty
AI Product Engineer

Marino Capitanio
Head of Creative

Niek Naber
Senior AI Engineer

Pradyut Nair
AI Engineer Intern

Raiane Sampaio
Head of HR Strategy

Riccardo Bevilacqua
Senior Software Engineer

Sean Kenny
Principal Product Manager

Seb Vassen
Head of Strategy

Vitor Fernandez
Gen AI Intern

On behalf of the entire Prosus AI team in Amsterdam,

thanks for reading our report!

Check out the appendix below for more insights on how we did this from an engineering perspective

Report Appendix: Solutions for technical challenges

When it comes to multi-agent structures, we focus on sub-agents

True multi-agent structures are not needed for most agentic use cases, and can create many problems at significant scale

When working with AI agents at large scale (both in terms of the sheer number of AI agents and the size of the tasks they are doing), reliability is of the utmost importance. Therefore, we have found that multi-agents structures where agents are "stacked on top of each other" exponentially increases the chance for errors. And it's very hard to run Evals on a complex multi-agent structure to figure out what is wrong and what is causing the error.

Multi-agent structures also create operational challenges. For example, a true "agents talking to agents" multi-agent structure can overcome access management controls. An agent that's not supposed to have access to X can ask another agent for X and potentially get around controls. In summary, most users are better off trying to solve a problem by creating a single agent and using multi-agent structures on selective use cases.

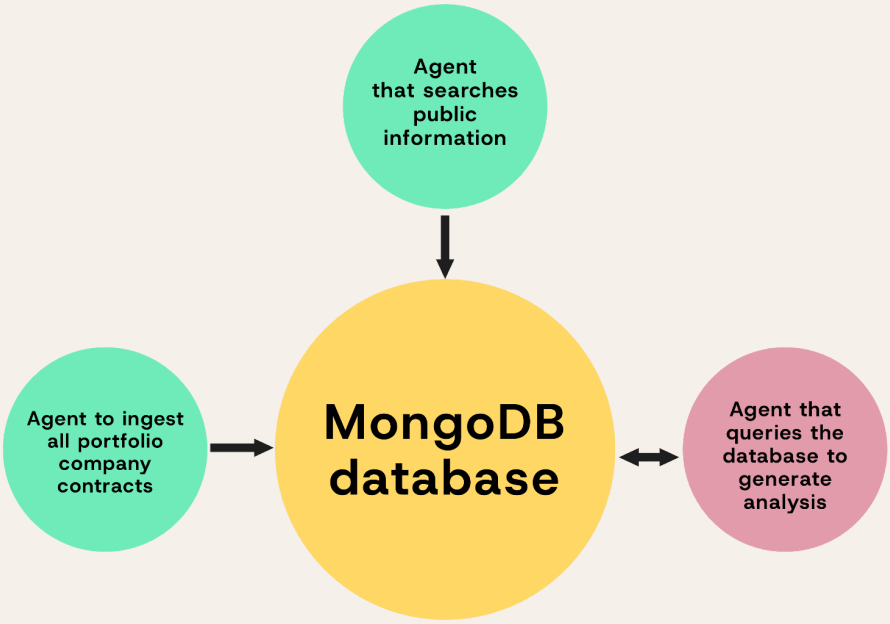
For complex use cases, we offer a sub-agent structure rather than multi-agent setups

Some tasks (very long length, high complexity, enormous volumes of data, etc.) do require more than one AI agent. In these situations, our approach is to focus on sub-agent structures and treating the sub-agents similar to the lead agent calling a tool. The main agent aggregates up all of the approvals. We have found this kind of sub-agent structure to be more reliable and less risky.

That said, sub-agent structures are still more of a "power user" configuration that takes an advanced understanding of AI. We have been gradually and carefully rolling out sub-agents on our internal Toqan agentic AI platform, and the sub-agent option includes guardrails (e.g., protection against loops). The UI/UX layout and language makes it clear that setting up a sub-agent structure is not the default setup or something that someone brand new to AI should be trying to do.

Sometimes agent orchestration can be done via a database at the center

To simplify, "agent orchestration" usually refers to infrastructure or logic that coordinates multiple agents. Our philosophy is that in many cases, the database or an ERP system can act as the "orchestration layer." Let's take an example from a hypothetical agent designed to review all of the various Prosus portfolio companies contracts to look for opportunities to lower pricing and save the firm money.



One AI agent ingests thousands of contracts and extracts the key data (e.g., vendor, contract date, price per user, etc.). A second AI agent looks for publicly available data points on prices and sends that information to the same MongoDB database. Finally, a third AI agent queries the data in the MongoDB database to generate the analysis. The third agent provides the report that outlines the opportunities to renegotiate and/or consolidate contracts (e.g., four portfolio companies all have contracts with ExampleCo, but JustEatTakeaway's seat license is 25% higher than the other three contracts).

After our first attempt failed, we are working on a new approach to long-term memory

Our first attempt at rolling out long-term agent memory was not successful

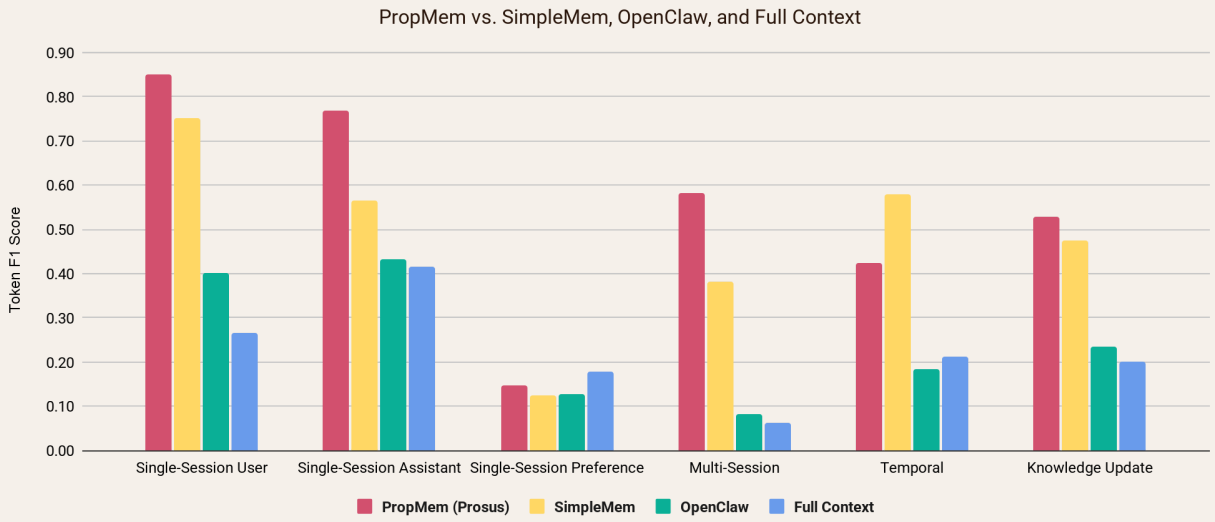
Our initial attempt at rolling out long-term AI memory on the Toqan platform was not successful. We originally used a Mem0 approach, but we found it stored too much as memory based on every interaction. As the number of “memories” grew, the size of the memory increased exponentially rather than in a linear way. Our initial attempt did not focus enough on the “durable facts” that the AI needed to maintain.

The failure didn’t happen right away. It started failing for many users after about a month as the “memory” started to really grow. It got to the point that our users started turning off the memory feature. Users found that updating the prompt and/or having a reference file worked better. So we went “back to the drawing board” and are now in the process of rolling out a new approach to memory.

Based on those learnings, we are now deploying our new PropMem approach

We have built a new approach to long-term AI memory focused on what to remember and what needs to be the durable facts. From a technical perspective, we extract what we think is the valuable information for memory and have another LLM verify this. In order to reduce token usage, we extract the whole session, rather than every single interaction. We have implemented very strong guardrails on what to extract for memory, as most conversations don’t have anything worth extracting. Our guardrails are focused on durable facts like “I prefer X format over Y format.”

There are some other additional details. For example, we make sure the memory system extracts the exact dates (to avoid the AI getting confused about “next week”). But that is the gist of how our PropMem system works. We have tested our PropMem system in different scenarios (below), and found that PropMem outperforms other benchmarks.



After some initial struggles with MCPs, we built an MCP creation engine

We initially struggled to use MCPs in a safe and reliable way

Like many others, we were drawn in by the excitement around MCPs. While we didn't "jump on the proverbial bandwagon" until fall 2025, we eventually launched MCP support on our internal Toqan agentic AI platform. We were surprised to discover the glaring shortcomings of the MCP landscape. Community-supported MCPs were (and still are) too risky to use for enterprise-level tasks. And the MCPs built by private companies suffer from misaligned incentives.

Private companies aim to make their MCPs as broadly applicable as possible. But users want their AI agent to do a very specific task and to not do anything outside of the agent's mandate. It's a clash. Some of these MCPs also include shockingly high-risk capabilities, like the ability to delete production data.

Of course, there are some exceptions and some private companies offer high-quality MCPs. As of this report writing, our internal Toqan agentic AI platform offers users both traditional integrations and the option to use four company MCPs (from Attio, Linear, Miro, and Monday.com). We will inevitably add additional company MCPs over time as the space matures. But we found we needed to create a new solution and a new approach (below) in order to use MCPs to their fullest extent.

Based on this experience, we built our own MCP creation engine

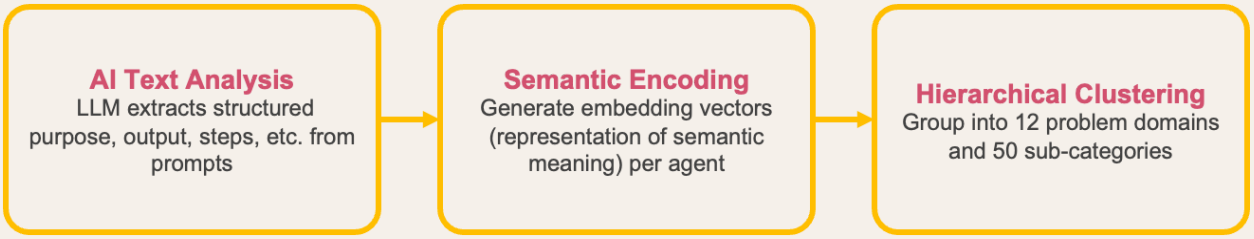
After struggling with MCPs for a few months, Prosus decided to build an MCP creation engine for the internal Toqan AI agent platform. This MCP engine turned out to be very valuable to users. We can now generate our own MCP for any platform the user wants to integrate with. In practice, the only input we need from the user is the platform's schema documentation, which allows us to spin up new MCPs quickly and at scale.

By generating MCPs ourselves, we found we have much better control over reliability, uptime, and overall quality. Users can also tailor the toolset for their precise needs. For example, if a platform exposes 200 tools, the user can restrict it to accessing only 3/200 tools. That level of control not only simplifies the agent's decision-making, but also reduces noise and potential failure points, making the entire system more robust.

Categorizing 60,000 AI agents required a blend of both AI and human insight

From a technical perspective, how did we perform this quantitative analysis?

We'll first explain how we ran the analysis of 60,000 agents, and then discuss the problems we ran into and how we solved it. After consolidating the data from different tables and removing inactive agents, three steps were performed. Those three sections are outlined below. For this section, we thought it would be helpful to give a detailed technical overview of how it was done.



Step One: LLM prompt decomposition (GPT-5-mini): Before extraction, prompts are cleaned by stripping common boilerplate patterns (e.g. "you are an AI assistant...") and normalising whitespace. Language is detected via langdetect. The cleaned prompt is sent to GPT-5-mini with a structured JSON schema enforcing four output keys: goal (the agent's objective), steps (its workflow), output (what it delivers), and constraints (rules it follows). Non-English prompts are translated during extraction. Calls run in parallel (50 concurrent, rate-limited at 500 RPM) and are incremental: previously processed agents are skipped.

Step Two: Embedding generation (OpenAI): A canonical text is assembled from extracted fields in priority order (problem, use-case,goal, steps, output) and embedded using Text-embedding-3-small (1536-dimensional vectors). A SHA-256 hash of the input text enables incremental re-embedding so only changed agents are reprocessed. Embedding calls are batched at 100 texts per request with 10 batches running in parallel.

Step Three: Hierarchical clustering (pyhercules): Embeddings are clustered in two levels: 60 fine-grained clusters, then merged into 12 broad domains. Clustering uses direct centroid representation. Cluster labels are generated by an LLM and refined to be mutually exclusive and collectively exhaustive (MECE). Each agent is assigned to its nearest cluster by cosine similarity; agents with a similarity score below 0.15 are flagged as weak fits for manual review.

We found that categorizing AI agents relies on humans and AI working together

After running the analysis with the framework above (and some other approaches that were even less effective), we found that AI struggled to properly categorize what these AI agents were doing. The AI often misplaced sub-categories and kept creating sub-categories that were very similar and not distinct enough. We also found that AI struggled with properly categorizing highly specialized department-specific tasks. While a human intuitively knows that, say, an agent that monitors interest rate offers at partner banks for the finance team is a highly specialized agent for the finance department, we found that LLMs often flagged these kinds of specialized use cases incorrectly.

After iterating through many failed attempts to have the AI categorize all 60,000 use cases, we took the "first draft" that AI came up with, met as a team, and came up with our own list of categories and subcategories. We then asked the AI to categorize all 60,000 agents within the exact category and subcategory list we had provided. That exercise was successful. AI returned an accurate breakdown that only required a few last tweaks to correct a few sub-categories.

We expect that most businesses will want to undertake a similar classification exercise of their agents, so we wanted to share the challenges we faced categorizing these agents, and how we ultimately solved this problem.

How to find and identify high potential agents

In order to find the agents that may join the "power law" group one day, we used an AI-based analysis of our most active AI agents to look for those AI agents that showed the most potential. To simplify, we define potential as a combination of value, adoption, and quality.

Our ranking is computed as a weighted linear model over three pillars: Value (45%), Adoption (35%), and Quality/Consistency (20%). The **Value** pillar assesses dimensions like the value it produces, who it benefits, and what decisions it informs. **Adoption** examines the total number of users, the number of requests per user, the growth rate over the last 30 days, etc. **Quality/Consistency** rewards consistent long-term usage and runs on a check on whether the perceived value is correct. A 60% anti-domination cap prevents any one pillar from overwhelming the final score before a dense global ranking is assigned. We used the GPT 5 mini model for this assessment, as we found a mini model was sufficient based on how we set up the classification task.

Of course, depending on the scale your organization has reached, you may need to further refine the high potential category. For example, at the scale Prosus has reached, there are 1,600 AI agents in the high potential group. That is too many for a human to review. So we apply even more selective filters and weights to the high potential category. The simplest way to reduce the number of agents is to increase the weight assigned to Adoption, but you can experiment with the weights to find the right number of high potential agents.