A werg guide · Seven chapters

Mastering your agentic transformation

How executives move from pilots to an agentic enterprise: the decision, the first agent, governance, value, data, the team and the scale-up.

Halil Aksu, Co-founder and CEO · November 2026

werg

About this guide

Most agentic AI programs stall between the first pilot and the first number a CFO will sign. This guide is written for the executives and managers who own that gap. Each chapter takes one decision on the path from a single agent to an agentic enterprise, says what tends to go wrong, and ends with where to start.

The guide draws on the werg founding team's work with enterprises and on the DAIMI 2026 benchmark, built by the werg founding team at Digitopia (1,575 assessments across 46 countries). Scenes introduced with "picture" or "consider" are illustrative, not customer cases.

1 Start with a decision, not a pilot

2 Where to point your first agent

3 Governance is a feature, not a brake

4 Value the CFO will sign off on

5 Your data is readier than you think

6 Who builds the agents?

7 From one agent to an agentic enterprise

Start with a decision, not a pilot

A pilot is designed to be abandoned. A decision has a process, a date, an owner and a number.

Most agentic AI programs fail before they begin, because they begin as pilots. A pilot is designed to be abandoned. A decision is designed to be kept. Organizational success starts with a leader deciding which process will run on agents, by when, owned by whom, and measured how.

Every mid-sized company has a pilot graveyard. An invoice-matching bot from 2019. A chatbot from 2023. A set of assistant licenses nobody renewed. Each one worked, more or less. None of them changed a number, because nobody had decided that it would.

Why pilots die

Pilots die for five reasons, and they are always the same five. No owner in the business, so IT runs it and the business watches. No number attached, so success is "we learned a lot." No path to production, because the pilot lives in a sandbox with no connection to SAP. No change in how people work, so the old way continues beside the new one. And success criteria set to "explore" rather than "move," so there is nothing to defend when budgets tighten.

A pharmaceutical distributor can run three such pilots a year for a decade. The DAIMI 2026 benchmark places pharmaceuticals at 2.84 against a cross-industry baseline of 2.93, and in our experience pilot culture is part of the reason.

What a decision looks like

"By March, the intercompany reconciliation runs on agents, exceptions only, owned by the controller, measured as cost per close." Four parts: a process, a date, an owner and a number. A fifth part matters as much: what stops. The old way is retired on a date, not kept as a safety net forever.

The decision is made at executive level and made public inside the company. People change their behavior when they see that the leadership has already changed its mind.

The organization around a decision

A business owner, never IT alone. A small delivery pod: a solution lead who knows the process, an AI engineer, and forward deployed engineers who sit with the team. A governance owner who sets the agent's permissions and reviews its log. A finance partner who certifies the number before and after. An executive sponsor who clears obstacles weekly, not quarterly.

The manager of the team sets the thresholds: what the agent decides alone, what it brings to a person. That one act does more for success than any steering committee.

Three signs you are doing it right

People on the team change how they work in week two, not month six. The exception log is reviewed by the business, in a standing meeting, with the agent's work on the agenda. And the second use case is proposed by a manager who saw the first one, not by IT or by the vendor.

A retailer that gets these three signs in the first quarter is on its way to a portfolio of agents within the year. One that gets none of them will have another pilot.

How we do it

At werg, delivery has three steps and none of them is called a pilot. Automated Discovery with our AI3R report, which ranks the company against its industry, lays out a roadmap and sizes the returns. Automated Development on wergOS with the Agent Builder. Augmented Deployment with forward deployed engineers, alongside the team that owns the process. The first deployment is contracted as a first deployment, with a number and a date, because that is the only kind we have seen survive.

Where to point your first agent

Score use cases on six criteria, map the volume, observe the person. The chatbot rarely survives.

The first agent should go where the pain is measurable, the rules are clear, the data is already in SAP and a manager wants it. Finding that spot is a design problem, and companies solve it best with two old disciplines they already own: process excellence and design thinking.

Picture the workshop. Forty use cases on sticky notes. Everyone's favorite is the chatbot for the intranet. The right answer is the deductions process in accounts receivable that nobody wants to talk about, because it is tedious, it is costing real money every week, and the person who runs it has been asking for help for two years.

Six criteria, one score

We score every candidate on six criteria, weighted. Pain intensity: how much it costs today, per unit, and how often. Buyer access: whether there is an owner in the business who can say yes. Solution maturity: whether agents have done this kind of work before, in this kind of system. Production proof: whether it can run on real transactions, not demonstrations. Time to value: weeks, not quarters. Competitive whitespace: whether doing it well will matter to customers or regulators.

Score each from one to five, weight them for your company, and the forty sticky notes become a ranked list of five. The chatbot rarely survives the first criterion.

Process excellence finds the volume

Lean and Six Sigma did not go away; they were waiting for better tools. Map the value stream. Count the transactions. Measure the cost per unit and the cycle time. Find the rework loops, the hand-offs, the queues. Agents are at their best exactly where process excellence finds waste: high volume, clear rules, repeated rework.

A consumer goods company that maps its order-to-cash flow will usually find the same three loops: price disputes, deductions and credit holds. Each one is an agent waiting to be built.

Design thinking finds the person

Numbers find the process. Observation finds the person. Sit with the planner, the buyer, the customs clerk for a day. Write down what they do, what they would hand over tomorrow, what they would never hand over. Ask what a good day looks like. Then prototype: an agent on one task, with that person reviewing every output, for a week.

The prototype is not a pilot. It is a conversation with a working example in the middle of it, and it tells you in days what a requirements document would not tell you in months.

From shortlist to roadmap

With five candidates scored, mapped and observed, the roadmap writes itself: one first deployment with the highest score and the most willing owner, two to follow in the next two quarters, two held for later. Avoid customer-facing processes first. Avoid anything that needs data you do not have yet.

At werg, Automated Discovery does this at speed. The AI3R report ranks a company against its industry on the DAIMI benchmark, which covers 1,575 assessments in 46 countries, then lays out a roadmap of prioritized projects by buying center and sizes the returns. It replaces the workshop with evidence, and it leaves the choice where it belongs, with the people who own the processes.

Pick one. Attach a number. Name the owner. The rest of this series is about what happens next.

Governance is a feature, not a brake

Governance lets a company run agents fast. Identity, permissions, oversight, cost and evidence fit on one page.

Governance is what lets a company run agents fast. Without it, every agent needs a meeting and every exception needs a committee. With it, agents get identities, permissions, limits and logs, and people get the confidence to let them run. Governance is not the brake. It is the steering.

Two companies, same industry. One has forty agents that nobody can list; they were set up by enthusiastic teams, run on personal accounts and share a password with a spreadsheet. The other has twelve agents in a registry, each with an owner, a permission set, a budget and a switch that turns it off. Guess which one's auditors sleep well, and which one deploys the next agent faster.

Five things to govern

Identity. Every agent is a principal of its own, with a name, an owner and a history, never a shared service account. Permissions. What it can read, what it can write, up to what amount, in which systems. Oversight. Who reviews the exceptions, how often, and which agents watch the other agents for drift and policy violations. Cost. A budget per agent and per process, visible to the business owner, not buried in the IT bill. Evidence. Every action logged with its reasoning, the rule applied and the version of the agent that applied it.

These five fit on one page. A company that writes that page before the first agent runs will never need the forty-agent cleanup.

Models, data and the law

Governance also covers what the agents are made of. Which model, which version, when it changed, and how its quality is measured. Where the data lives and where it may travel. Rules are tightening on both sides of the Atlantic: the European Union's AI Act is phasing in obligations through 2027, and in the United States sector rules such as SOX already reach any agent that touches financial reporting. The practical answer is the same everywhere: know what your agents do, prove it, and be able to stop them.

A listed manufacturer with plants in Europe and the United States needs both answers at once. So does any company that wants to sell to one.

What governance is not

It is not a committee that approves every agent; that is how companies end up with zero agents and a slide deck. It is not a policy document nobody reads. And it is not IT's job alone. The business owner of a process owns the agent's permissions and thresholds, because the business owner knows which error is expensive. IT and security own the platform those permissions run on.

Thirty days to a governed start

Week one: an agent registry, even if it is a spreadsheet, with an owner per agent. Week two: thresholds for the first process, written by its manager. Week three: the evidence format, agreed with internal and external auditors. Week four: a tested switch that turns an agent off, and a quarterly review on the calendar.

At werg, this page is built into wergOS. The Control Center holds identity, permissions and limits. The Quality Center measures accuracy and drift. The Financial Center shows cost per agent and per process. Oversight agents watch the working agents. The platform is SOC 2 and ISO 27001 certified and is agnostic on models and clouds, so the governance stays when the technology underneath changes. The rules are yours. The enforcement is built in.

Value the CFO will sign off on

Hours saved is the wrong number. Measure cost per unit, before and after, including the agents' own cost.

The value of agents is real and mostly invisible, unless it is measured the way a CFO measures: per unit, before and after, with evidence attached. Hours saved is the wrong number. Cost per transaction, cycle time, cash and error cost are the right ones, and they are the ones that get a program funded for a second year.

The scene repeats in many companies. A team reports that agents saved two thousand hours. The CFO asks whose hours, and what those people did instead. Silence. The program was real; the number was not. Six months later the budget line is gone.

Four kinds of value

We sort every benefit into four groups, and we insist that each one has a number and an owner. Growth: capacity to sell more without adding cost, faster quotes, promised dates that hold. Excellence: cycle time, quality, cost per unit, exceptions instead of samples. Cash: collections that happen on time, deductions resolved, customs refunds claimed, inventory released by better forecasts. Intangible: resilience, corporate memory that survives turnover, auditability, and the talent that stays because the work got better.

Three of the four can be measured in money. The fourth can be measured in events: audit findings, turnover, days to onboard a new controller. Intangible does not mean unmeasured.

Per unit, not per headcount

The honest unit is the transaction. Cost per close. Cost per purchase order. Cost per customs entry. Cost per control test. Cycle time from event to decision. Exception rate. Measure each before the agent arrives, during the parallel run and after, and include the agent's own cost in the number, because agents consume computing every time they work.

A pharmaceutical distributor that reports "cost per order fell from X to Y, including the agents" has a number a CFO can audit. One that reports "we freed up the team" has a feeling.

Proving it

Baseline first, with finance in the room. Then a parallel run, agent beside the person, for one cycle. Then the switch, and a before-and-after comparison that finance certifies. Publish the number monthly, inside the company, next to the exception count.

The business case itself fits on one page: the process, the baseline, the target, the cost of the agents, the timeline, the owner, the risks, and a section most companies skip, which is what the agent will not do. That last section is what makes the rest credible.

The questions a CFO will ask

What happens to the people? They move to the exceptions and the improvements, and their new work is on the same page. What if the agent is wrong? The thresholds, the review loop and the switch are in the governance plan. What does it cost next year? The agent's running cost is in the Financial Center, per process, and it is reviewed with headcount and capital.

At werg, the Returns section of our AI3R report is written to these rules: leader practices, shortfalls, and business cases per buying center in cost-per-unit terms. We write our own commitments the same way and expect customers to hold us to them. The DAIMI 2026 benchmark says where an industry stands. The business case says what one company will do about it, and by when.

Your data is readier than you think

Agents do not need perfect data. They need SAP data plus the rules nobody wrote down.

"Our data is not ready" is the most common reason for waiting, and it is usually wrong. Agents do not need perfect data. They need the data the process already runs on, which is in SAP, plus the knowledge in people's heads, which is the part nobody has ever captured. Both are closer than they look.

Picture a CIO at a manufacturer who explains that agents will have to wait for the data platform. Two years later the platform is still a program and the invoices are still being matched, by people, every day, with exactly the data the agent would have used. The data was ready. The decision was not.

What an agent actually needs

Four things, and only the fourth is usually missing. Transactional data: orders, invoices, entries, journals. It is in SAP, where it has been for twenty years. Master data: customers, suppliers, materials. It is good enough, and where it is not, an agent will tell you within a week. Documents: contracts, invoices, declarations, emails. Agents read them now, which was not true three years ago.

The fourth is the rules. How this company treats a deduction, when a planner overrides a forecast, which broker gets which lane. Those rules live in people's heads, in email threads and in the spreadsheet on one controller's desktop. Capturing them is the real data project, and it takes weeks, not years.

Corporate memory, the missing layer

Capturing the rules is only the start. A company also needs its agents to know what its words mean: what counts as a customer here, how a plant relates to a profit center, why a particular supplier is always paid early. That is an ontology, and the relationships between the things in it form a knowledge graph. Add the decisions made and the reasons behind them, and you have corporate memory.

At werg, that layer is The Library. It is what makes an agent yours rather than generic, and it is what stays when a person leaves. The ontology, the graph, the decisions and the reasons: nothing is forgotten when someone moves on.

Integration without a rebuild

Companies also wait for the migration. Do not. Agents connect to SAP on any version and any deployment, ECC or S/4HANA, on-premise, private cloud or public cloud, and they carry across when the migration happens. The same applies to the Microsoft 365 estate where most of the documents live. Standard connectors and open protocols such as MCP and A2A mean the integration is configured, not built.

A retailer with three SAP instances after acquisitions can run one set of agents across all three, through one Connection Builder, before it ever consolidates them.

Data quality improves by use

The surprise in many deployments is that data quality rises after the agents arrive, not before. An agent that reads every vendor master record every day finds the duplicates and the missing bank details that a quarterly report hid for years. The fix takes weeks because someone is finally looking.

Three steps for this month

List the five data objects your first process touches, and nothing else. Run the agent in read-only mode for two weeks and let it report what it finds. Write down the rules as the people explain them to the agent. That is the data project. The platform can come later, and it will be better for knowing what the agents needed.

Who builds the agents?

The first agents are built by a pod with the business. The later ones by your own people, on a platform you govern.

Not IT alone, not a vendor alone, not the business alone. The first agents are built by a small pod with the business in the room. The later ones are built by your own people, on a platform you govern. The operating model decides whether a company ends up with a capability or with a dependency.

Three models fail reliably. Outsourcing everything to a vendor produces agents nobody inside the company understands. IT building for the business produces requirements documents and six-month cycles. The business buying bots on its own produces forty ungoverned agents and an audit finding. The model that works borrows from all three and belongs to none of them.

The pod

A werg pod has four people. A solution lead who knows the process and the buying center it belongs to: finance, procurement, customs, whichever it is. An AI engineer who builds the agents on the platform. Two forward deployed engineers who sit with the customer's team, connect the agents to SAP and the documents, and stay until the agents are running on real transactions with the team reviewing them.

Forward deployed engineering is a discipline, not a job title. It means working from the customer's desk, in the customer's process, with the customer's data, and leaving behind something the customer's people can run. The first weeks are spent with the manager, not with the system.

The handover

The pod's job is to make itself unnecessary. Within the first deployment, three roles appear on the customer side. A process owner who writes the rules and sets the thresholds. Agent reviewers in the team who work the exception queue and feed corrections back. And, by the second or third deployment, an internal AI engineer or product owner who composes the next agents with the Agent Builder rather than waiting for a vendor.

We designed our model so that the need for our engineers declines with each deployment. A customer who still needs the same amount of our time on the fifth agent as on the first has been given a service, not a capability.

Skills to grow, and where

Process owners who can write a rule precisely are the scarcest skill, and they already work for you; they are the people who currently hold the rules in their heads. Reviewers are made in weeks, from the team itself. One internal AI engineer per two or three buying centers is enough to start. A

governance owner, usually in finance or internal control, who runs the registry and the quarterly review.

What you do not need: a center of excellence that owns every agent and becomes a queue. A data science team hired before a process is chosen. A platform selection that takes longer than the first deployment would have.

Where to start

Pick the process and the pod together. Name the process owner before the pod arrives. Put two team members on review from day one and give them the time to do it. Decide, in writing, which agent your own people will build second.

An automotive supplier, a pharmaceutical secondary manufacturer and a grocery chain have different processes and can share the same operating model. At werg, the pods, the Agent Builder and the Connection Builder exist so that the model transfers: we build the first with you, you build the rest on wergOS, and the governance stays the same throughout. The agents are yours. So is the capability.

From one agent to an agentic enterprise

Scaling is a rhythm problem: deliver, digest, prove, repeat, on one platform with one set of rules.

The first agent proves it works. The second proves you can repeat it. The tenth is when the company changes. Scaling is a rhythm problem, not a technology problem: one process at a time, one month to deliver, one month to digest, value proven in the third, the next one contracted in the fourth.

Picture a consumer goods company two years in. Finance closes on agents with exceptions only. Customs audits every entry. Planning runs daily. Controls are continuous. Procurement handles contract-covered orders without a buyer touching them. None of this arrived at once. It arrived in waves, each one smaller and faster than the last, on one platform with one set of rules.

The rhythm

Deliver: the agent runs on real transactions within a month, with the team reviewing every output. Digest: the team lives with it for a month, changes the rules, finds the limits. Prove: finance certifies the number against the baseline in the third month. Repeat: the next process is contracted in the fourth, with an owner who has watched the first one work.

Three waves fit in a year. Each wave is cheaper than the one before it, because the platform is in place, the governance is written, the reviewers know their job and the corporate memory is already there. By the second year, waves overlap.

One platform, not forty bots

The decision that makes scaling possible is made early: one platform with shared identity, governance, memory and cost, rather than a bot per department. Forty point solutions each need their own permissions, their own logs, their own integrations and their own vendor. One platform means the second agent inherits what the first one built.

At werg, that platform is wergOS: solution packs on top, the Agent Builder and the Connection Builder for composing and connecting, the Engine underneath, and the Control Center, Quality Center, Financial Center and Oversight agents watching every agent the same way. The Library holds the company's memory for all of them.

Managing the portfolio

A portfolio of agents is managed like a portfolio of anything else. A roadmap per buying center: finance, procurement, supply chain, trading, customs, internal control, IT and ERP. Prioritization with the same six criteria used for the first one. A quarterly review that adds, widens and also retires

agents, because an agent whose process has changed is a liability. A cost line per process, reviewed with headcount and capital.

The organization at scale

Process owners who write rules. Reviewers in every team. A small platform team, often two or three people, who own the registry and the integrations. A governance owner in finance or internal control. And the management rhythm that daily facts make possible: shorter meetings, decisions closer to the exception, a plan revised as reality reports in.

Signs you have arrived

A manager proposes the next agent without asking IT. The audit committee asks for the agent log before the sample. New hires are onboarded to the agents on their first day. The DAIMI score, which the werg founding team at Digitopia has measured across 1,575 assessments in 46 countries, moves from the industry average toward the leaders, and the company can say exactly which processes moved it.

Start with one. Keep the rhythm. The enterprise follows.

Where to go from here

werg builds enterprise AI that does the work: proven agentic solutions for finance, procurement, supply chain, trading, customs, internal control and IT and ERP teams, running on wergOS, our governed operating system for agents, in your cloud or on your premises.

Talk to us. A 20-minute conversation about the one process you would point your first agent at.
werg.ai/contact

Estimate the value first. werg.ai/roi

Read the rest of the series. werg.ai/resources

Follow werg. linkedin.com/company/werg-ai

© 2026 werg. Enterprise work, done.