

EXECUTIVE BRIEFING // 2026

Work in 2029

How people and agents work together, and the foundation a company needs

KRISTIAN ARVIDSSON, CHIEF AI OFFICER

Contents

Chapter	What it covers	Page
Culture and values	How a strong culture can guide agents	5
The brain	The Codex and the Knowledge Base	9
Agents	Harnesses, what agents may do, monitoring, communication and ownership	17
Access and rights	Who may see, change, approve and publish, for people and agents	24
Data	What makes information usable and trustworthy	29
Systems and tools	Connected systems, and when to build, buy or replace tools	36
Workstreams and processes	What work is done, who owns it, and how it runs with people and agents	41
Vision	The picture of work in 2029	45
Risks and open questions	What could go wrong, and what is not yet known	49
Terms	The key terms, in one line each	55
Sources	Research, ideas, regulation and standards referred to	56



INTRODUCTION

By 2029 the difference will be between companies where agents do whole pieces of work beside people, and companies where AI still waits to be asked.

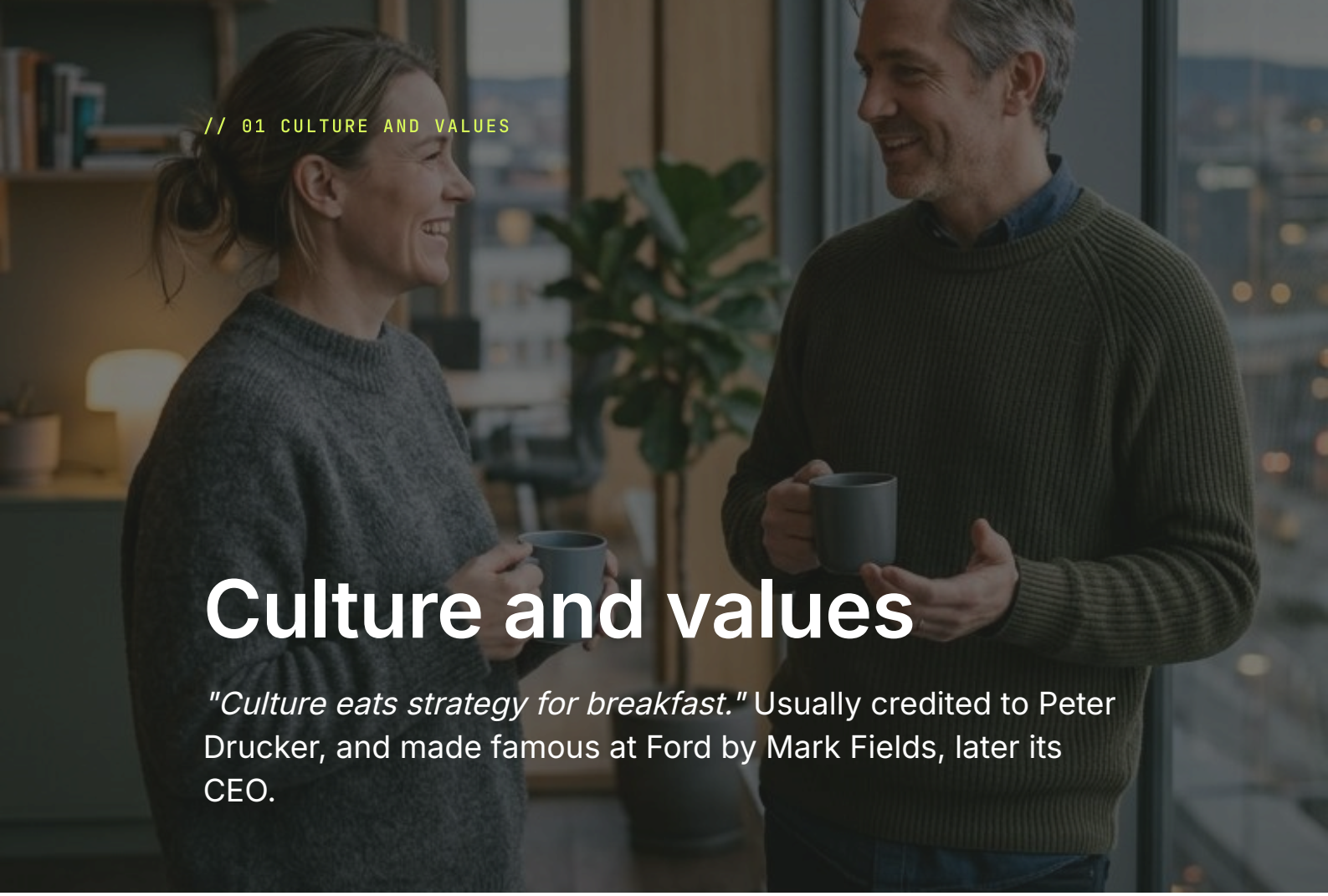
Most companies today are in the second group. They use AI as a copilot: a clever assistant that helps one person with one task at a time. It saves minutes. It does not change how the company works.

What holds them back is not the models. Models can write, analyse and plan in seconds. What they cannot do is remember. Every conversation starts from zero: they do not know a company's customers, its promises, its way of working or why it does things the way it does. Give an agent poor or missing knowledge and the result is confident work that is wrong. The old rule was rubbish in, rubbish out. With AI the rubbish comes out polished, and at scale.

Companies that stop there hit a ceiling: AI makes them quicker, but not better. The ones that pull ahead will not have better models, because the same models are available to everyone. They will have built a new kind of foundation: one made for people and agents working together, not for people alone.

At the centre of that foundation is the brain: a shared record of what the company knows, how it works and why, used by people and agents alike. Even if models learn to remember, a company should own what it knows. Around the brain come the culture that guides the work, the agents that do it, the rules for who may see and change what, and the way work is organised. With that in place, agents carry the routine. People set the direction, make the calls that matter and answer for the result.

Agents that do real work beside people are only a few years old, and nobody has a proven playbook yet, not even the companies that build them. None of this is a finished answer. It is a direction: a way to organise work so that agents make a company stronger rather than spread its mistakes faster. Some of it rests on research, cited where it is used. Much of it rests on experience of building and working with agents. Parts of it will change as practice shows what holds. What is not yet known is listed under risks and open questions.



// 01 CULTURE AND VALUES

Culture and values

"Culture eats strategy for breakfast." Usually credited to Peter Drucker, and made famous at Ford by Mark Fields, later its CEO.

Culture cannot be installed. It is built slowly, through what people do every day, and it fades as soon as nobody looks after it.

Agents are installed. So how does a culture reach them?

Why agents need the culture

An agent has no culture of its own. It has a default, set by whoever built it, and that default is often to agree. AI tools tend to affirm what people say rather than challenge it. Inside a company, that is not harmless: unchallenged ideas, and leaders who only hear agreement, have brought companies down before.

BT Group tested this. One of its values is to be a "Challenger". In a trial with 2,400 employees setting their goals for the year, one group used an AI that tended to agree, and another used an AI designed to challenge them: to ask for specifics and to push further on scale and timeline. The challenging AI produced three and a half times as many goals rated excellent, and fewer weak ones. People enjoyed it just as much, and felt just as safe.

3.5×

More goals rated excellent when the AI was built to challenge rather than agree (BT Group trial, 2,400 employees).

The lesson is that culture does not reach agents by itself. It has to be written down in the Codex and built into how agents work with people. A commitment to constructive challenge must guide the agents too.

Values that guide decisions

Values help when they change what people do. People and agents need guidance they can use when the answer is unclear.

Qualities like curiosity, constant learning, transparency, accountability, inclusion and diversity matter, but single words give neither people nor agents enough to act on. A value must be specific enough to decide something. These are values worth working by. For agents, each value comes with a test that matters most when the answer is inconvenient.

01

Everyone means everyone

Before settling on a solution, involve the people affected and seek perspectives missing from the discussion. Make it possible to question a proposal regardless of role or seniority.

Agent test: Does the agent name the perspective that is missing?

02

Make work compound

Leave the sources, reasoning and next steps so others can build on the work. Improve a shared process when the same problem returns, instead of repeating the workaround.

Agent test: Does it leave its sources and reasoning behind?

03

Show, don't assume

Make assumptions visible, check the source and test claims against evidence. Use maintained knowledge as the starting point; when evidence conflicts with it, raise the discrepancy with its owner.

Agent test: Does it flag the assumption it cannot support?

04

People decide, agents prepare

Agents gather evidence, compare options and challenge weak reasoning. People make the call when a decision reaches a customer, commits the company or cannot be undone, and they own the result.

Agent test: Does it hand over the decision it may not make?

05

Share what you learn, including what failed

Record what happened, what the evidence supports and what should change. Share it with the people who can use it. Propose useful lessons for review so they can improve shared guidance.

Agent test: Does it propose the lesson from what failed?

How the culture becomes practice

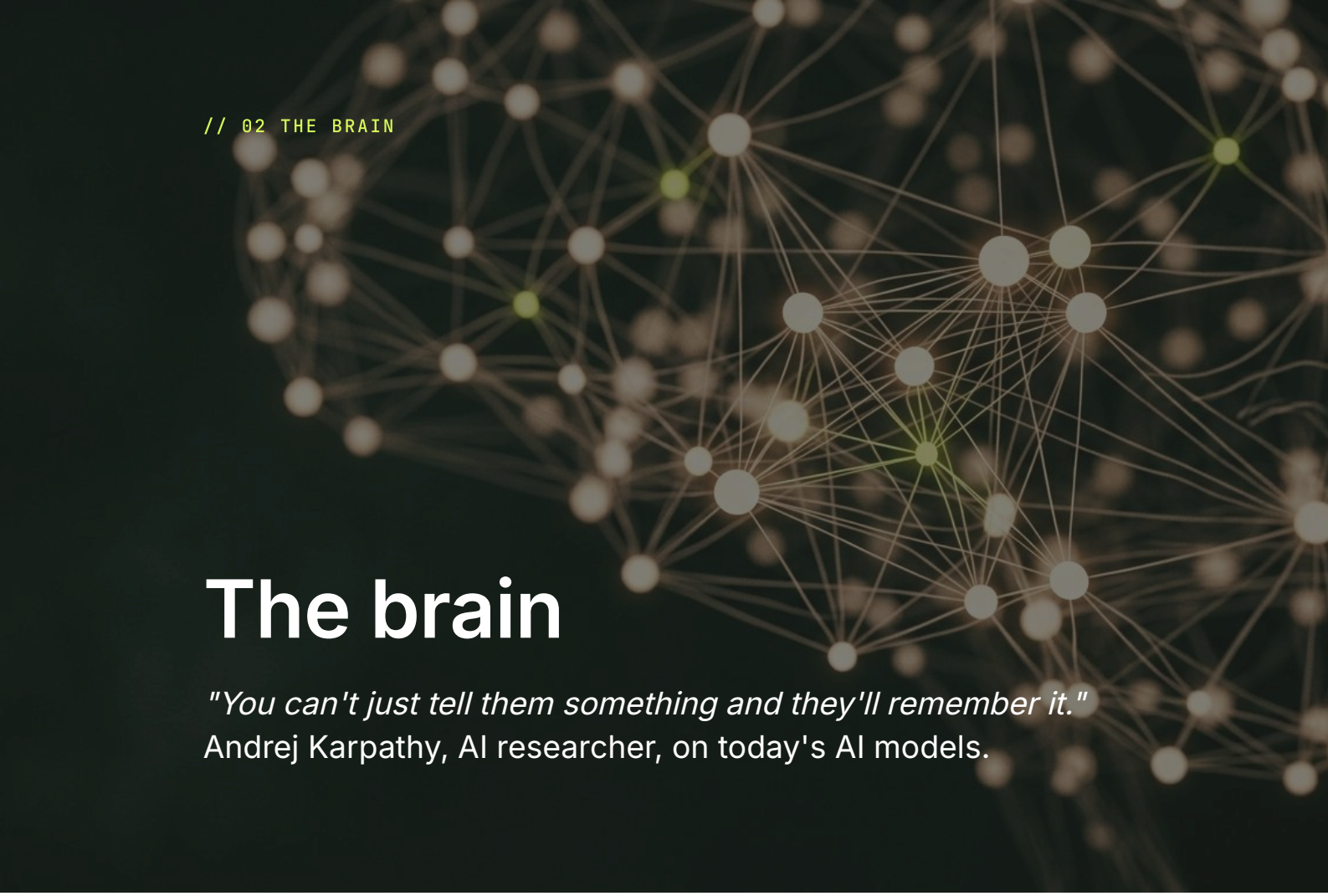
Culture becomes credible through what people do every day. Leaders ask what could make a proposal wrong, and welcome answers that change their minds. A company where people only hear agreement from each other will not notice when they only hear agreement from their AI. Teams make room to examine mistakes and improve the work, without hiding responsibility.

These expectations belong in the Codex and in agent instructions, so agents carry them too: they ask for specifics, point out weak spots and say when the evidence does not support a claim. Agents are checked on whether they follow them when the answer is inconvenient, not only when everyone agrees.

Habits for people and agents alike are the main defence against AI that agrees, and people who stop checking:

- **Say when an AI wrote it.** Passing on AI output is fine. Passing it off as your own judgment is not.
- **Check before you put your name to it.** If you send it, you own it.

Teams openly share how they work with agents, what worked and what failed, so good habits spread instead of staying with individuals.



// 02 THE BRAIN

The brain

"You can't just tell them something and they'll remember it."
Andrej Karpathy, AI researcher, on today's AI models.

The brain is a shared record of what the company knows, how it works and why. It is the central system for how the company works: it owns the company's information, and people and agents turn to it to understand the company and do the work.

A senior colleague once explained to a junior why a batch of old agreements was much older than its dates showed. After a disk failure, every recovered document had been stamped with the same date. The junior had no idea. An agent would not have known either. It would have read the date, built its whole analysis on it, and been confidently, silently wrong. Every company runs on knowledge like this, and the real question is: what does a ten-year employee know that is written down nowhere?

The fighter pilot John Boyd described every decision as a loop: observe, orient, decide, act. At its centre he put orientation: how a person reads what they see, shaped by experience and culture. Agents are already fast at observing and acting. Orientation is what they lack, and it is what the brain gives them.

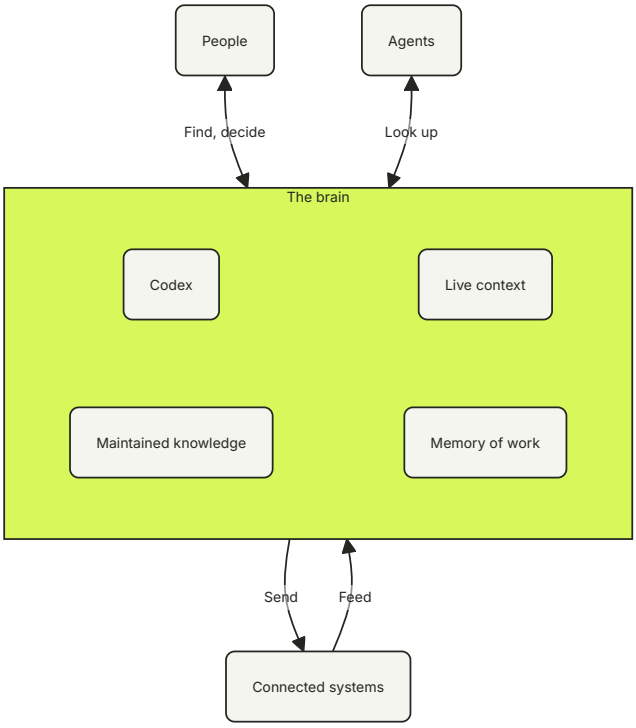
They also lack memory, as Karpathy says. Accenture reaches the same conclusion from its work with large enterprises: generic models "know a lot about the world, but little about your company." The brain is where they look it up, and where what the company learns stays.

The brain closes both gaps, for people and agents alike. An answer given once should be there for the next person or agent who needs it, and when the answer lives in someone's head, the brain should at least know who to ask. People put it there: they record what they know, question answers that look wrong and correct them at the source.

The brain has two parts: the Codex, which holds what guides the company and why, and the Knowledge Base, which holds three kinds of information: maintained knowledge, live context and memory of work.

The brain is not one product. It is built from connected tools, and its Codex, Knowledge Base, register and access rules are owned and run as one whole. Tools can be replaced; what the brain holds, and the rules around it, remain. One named executive answers for the brain as a whole: that its parts work together, and that gaps between owners are closed. Knowledge, agent, system and workstream owners keep authority over their own parts and bring conflicts across those parts to the brain's owner.

How the brain fits together



01

Codex: the why

- **Holds:** values and principles, the reasoning behind important decisions, how the company sees the world and wants to be seen, and what people expect of each other.
- **Read by:** everyone inside the company, people and agents. Nobody outside.
- **Changed by:** a named owner, slowly, with every version kept. Anyone inside may propose a change.
- **Not here:** personnel matters, deal details, restricted examples.

02

Maintained knowledge: what is approved

- **Holds:** agreements, approved facts, product knowledge, processes, playbooks, definitions.
- **Read by:** whoever each item is meant for.
- **Changed by:** anyone may propose a change; the named owner approves it. Reviewed when facts change and on set dates.
- **Not here:** unreviewed drafts presented as fact.

03

Live context: what is true right now

- **Holds:** current information from selected connected systems, such as customer activity, support cases and figures.
- **Read by:** whoever the brain's access rules allow, also in connected systems.
- **Changed by:** authorised people and agents, directly or through connected systems, continuously or at set times, with source and update time shown.
- **Not here:** copies that pretend to be current.

04

Memory of work: what the work leaves behind

- **Holds:** decisions made in the work, outputs, drafts, lessons and next steps, kept in each workstream's space.
- **Read by:** whoever may use it in the work.
- **Changed by:** the people and agents doing the work. A draft stays a draft; a lesson becomes guidance only after review.
- **Not here:** generated summaries treated as company truth.

Codex

The Codex matters most when there is no process. No company can write down every situation a person or an agent will meet. When a process is missing, the Codex is where they look to understand how to approach it, and to see when a situation needs a person to decide.

Everyone inside the company, people and agents, can read all of it, so only what everyone inside may read goes in. Some reasoning cannot be shared that widely, such as pricing logic or the details of a negotiation. The Codex then holds the principle and as much of the why as everyone may read, and the full reasoning lives in the Knowledge Base with its own audience. The rule is to share the why as widely as possible, and restrict only the details.

A principle in the Codex must be specific enough to decide something. *"We care about our customers"* decides nothing. *"Nothing reaches a customer without a person approving it"* decides a great deal. A useful Codex also says which trade-offs the company makes, such as speed or thoroughness and cost or quality, and where the hard boundaries are.

Without it, an agent can know every document and every fact, carry out its task perfectly, and still produce exactly the wrong thing, because it never understood what the company actually cares about. The result is work that is technically correct and strategically useless.

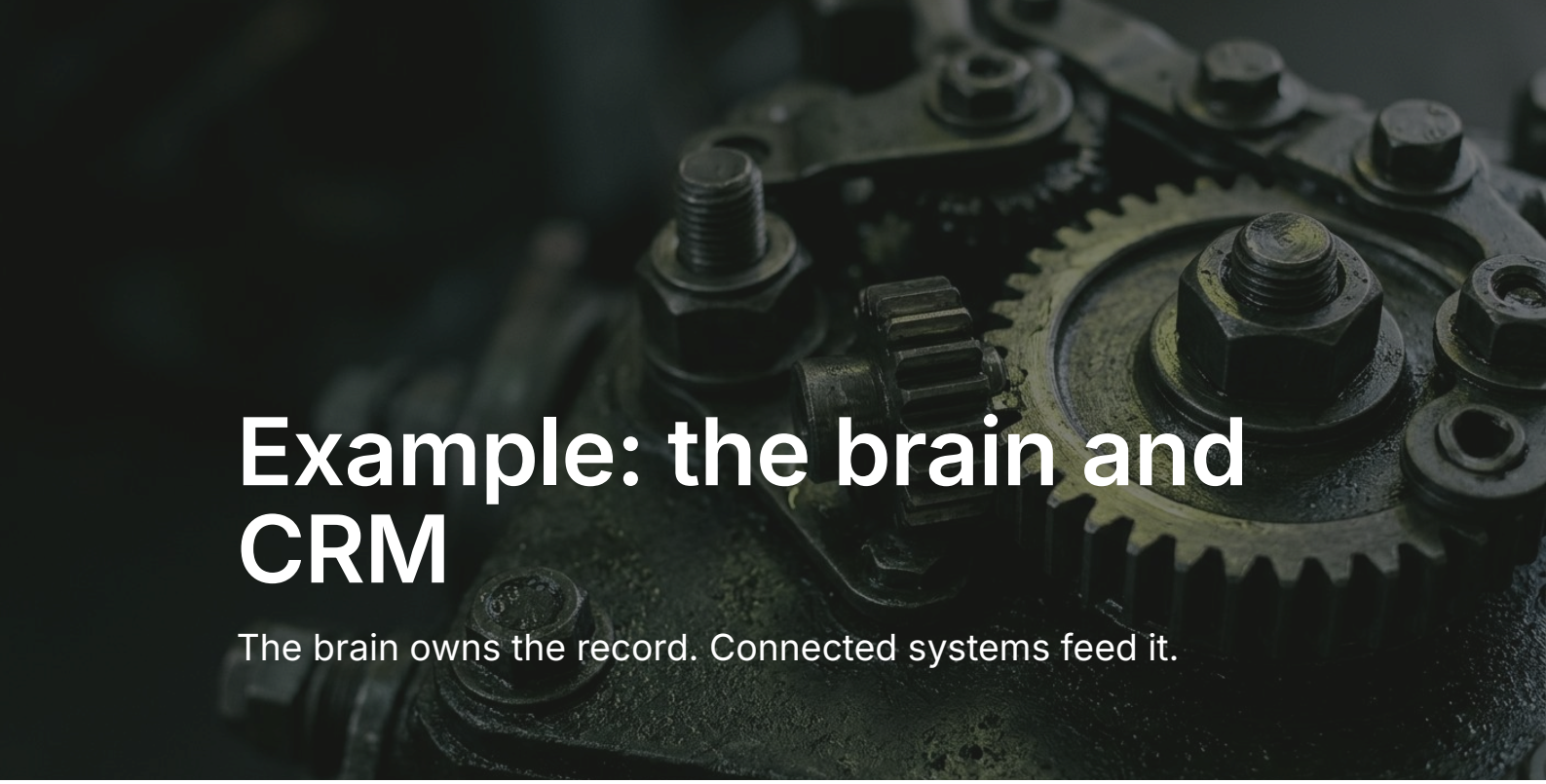
Knowledge Base

The Knowledge Base is where people and agents turn to understand a situation and continue work without collecting the same background again. Its three kinds of information are kept apart because each is trusted in a different way.

Maintained knowledge is approved: every item has a named owner, and nothing changes without that owner. Live context is current: every update shows where it came from and when. The source says where an update came from, not who owns it. The brain owns the information, and a named person answers for each part of it. Memory of work is a record: a draft stays a draft, useful lessons become maintained knowledge after review, and decisions that should guide everyone can be proposed for the Codex. A summary does not become company truth simply because an agent produced it.

Unlike the Codex, the Knowledge Base has access per item: an agreement, a personnel record and a product guide may each have a different audience. When something is wrong or missing, the correction goes to the owner of the information, who fixes it where it belongs, and the update reaches everyone and everything that uses it.

Together with the Codex, the Knowledge Base means that knowledge survives handovers, decisions can be understood later, and lessons from one task help the next one start further ahead.



Example: the brain and CRM

The brain owns the record. Connected systems feed it.

01

Agreements are owned by the brain, not by CRM. They are changed in the brain, by their owner, and CRM receives the details it needs, so salespeople see the terms where they work. CRM cannot change them: a change is proposed to the agreement's owner in the brain.

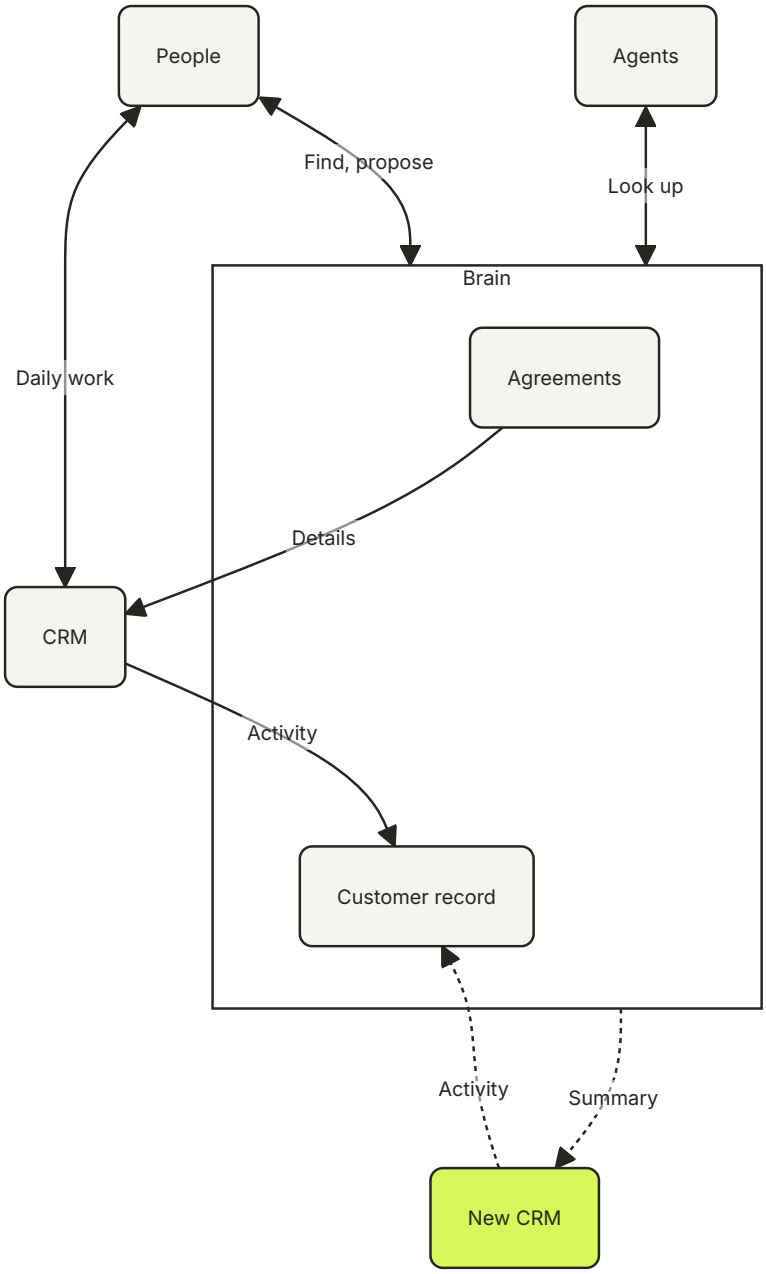
02

Customer activity is recorded in CRM, where salespeople do their daily work. CRM feeds that knowledge to the brain with its source and time, and the brain keeps the record. Agents and colleagues in other departments find it in the brain, not in CRM.

03

If CRM and the brain disagree about a fact, the conflict is shown and the owner decides which version is right.

The brain and CRM, today and after a switch



Solid lines: how the brain and CRM work together today. Dotted lines: what a new CRM receives and sends back when CRM is replaced.

Because the brain owns the agreements and the record, CRM can be replaced. A new CRM does not need every past transaction. It receives what fits: the details of each agreement and a summary of past actions with each customer. The summary is a working view, not the record itself: the full history stays in the brain. From then on, the new CRM feeds the brain just as the old one did.



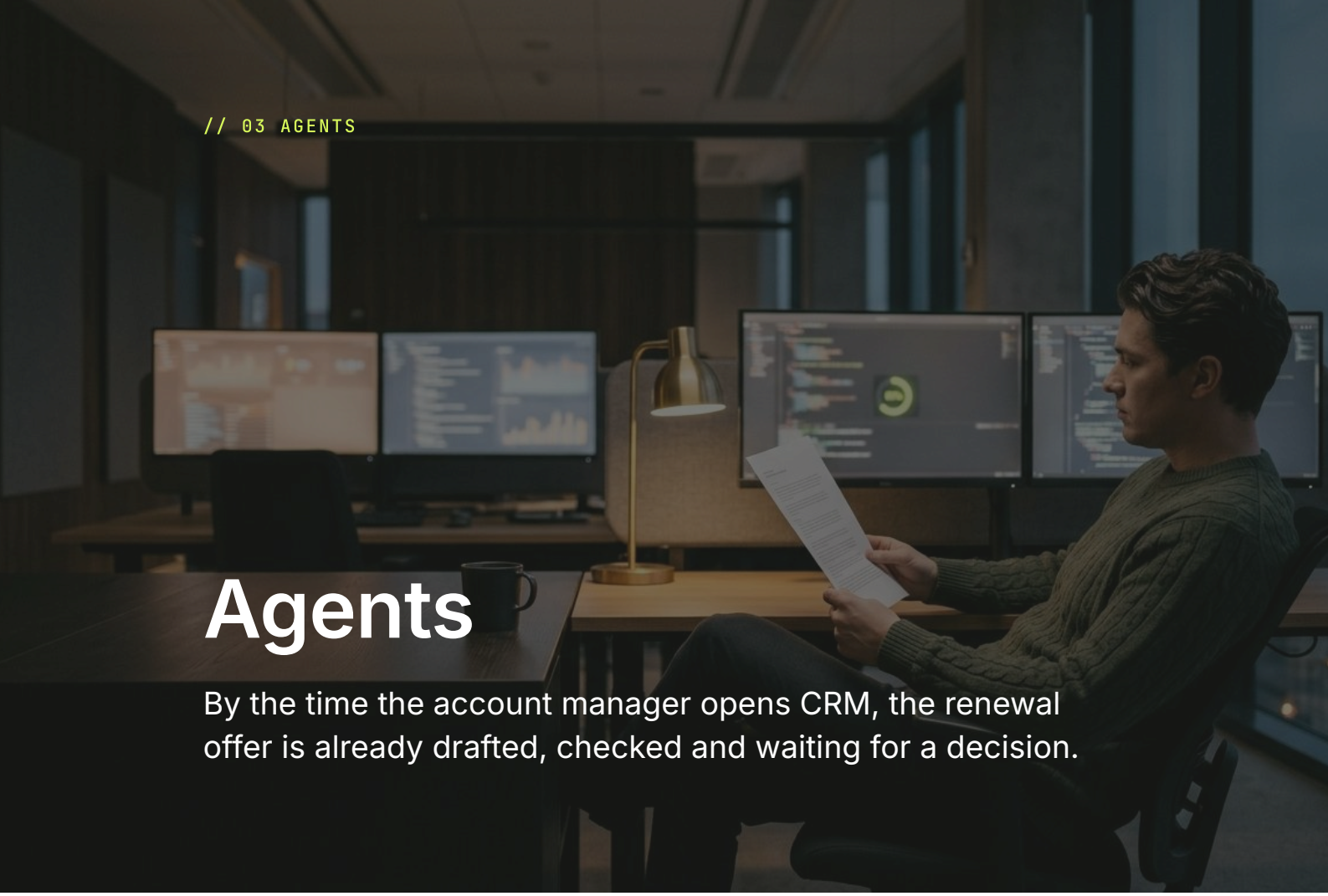
The same pattern applies to every system. Finance, contract management and e-signing work the same way: the signed agreement is kept in the brain, and each system receives the details it needs, such as prices and invoice terms for finance and renewal dates for CRM. No system changes the agreement itself. Because none of them owns it, any of them can be replaced, and a company can run more than one, for instance a different CRM in a new market.

What the brain cannot hold

Not everything can be written down. Some of the most valuable context is built by being in the room: hearing how sales explains a stalled deal, how finance talks about margins, how support describes a frustrated customer. That is absorbed by being close to the people who have it.

The brain does not replace that.

Instead, it makes sure that what can be written down is written down, so the time people spend together goes to what only people can share.



Agents

By the time the account manager opens CRM, the renewal offer is already drafted, checked and waiting for a decision.

An agent is AI that can take steps toward a task. It finds information, uses tools, drafts, checks its own work and hands over the result. In the future of work, every department works with agents beside its people.

Take a renewal agent in sales. Ninety days before an agreement ends, it notices the renewal, reads the agreement and the customer's history, and drafts an offer. The account manager decides what is sent. The head of sales, as its agent owner, answers for how the agent works.

Agents and harnesses

A capable model still needs a clear job and a way to carry it out. The harness provides the instructions, context and controls an agent needs to do useful work:

- **Instructions:** what the job is and what a good result looks like, set by the agent owner.

- **Knowledge:** the Codex first, then the parts of the Knowledge Base the job needs.
- **Tools:** the systems the agent can use and the actions it can take.
- **Access:** what the agent may see and do, and on whose behalf.
- **Rules for stopping:** when the agent must stop and ask a person.

Every department has a set of harnesses, all connected to the brain. An agent reads how the company works before it starts. It records its work and lessons in the brain, so the next task starts further ahead. Lessons become maintained knowledge only after the relevant owner has reviewed them.

The renewal agent's harness lets it draft in CRM but not send, and tells it to stop when a customer asks for terms outside the price list.

Reactive and proactive harnesses

Reactive harnesses start when someone asks. A person, or another agent working for a person, gives them a task: prepare a meeting brief, answer a question, draft a document, collect information. They are the day-to-day support for the work.

Proactive harnesses start on their own, on a schedule or when something happens. They work around the clock to find problems and opportunities before anyone asks: an agreement that is up for renewal, a pattern in support cases, a figure that moves in the wrong direction. They prepare the material and the suggested next step, so people start the day with the work already prepared.

Proactive does not mean free to act. A proactive harness works on behalf of its owner, with no more access than its owner has, and what it may do on its own follows the same rules as every other agent.

A proactive harness also has limits on what it raises: only what its owner has asked it to look for, a handful of items rather than a long list, and each with a reason it matters now. Suggestions nobody acts on are reviewed, and the harness is adjusted or stopped, so people do not learn to ignore it.

Who decides

Every action an agent can take belongs to one of three levels:

01

The agent does it

Examples: Finds information, summarises, drafts, compares, checks

Who decides: The agent, and the work is logged

02

The agent prepares, a person approves

Examples: A message to a customer, an offer on agreed terms, publishing content

Who decides: A person: the owner of what is affected, such as the agreement owner

03

A person does it

Examples: New commitments, changes to an agreement's terms, negotiations, decisions about people, strategy

Who decides: A person, with the agent preparing material if asked

The rule is simple: anything that reaches a customer, commits the company or cannot be undone needs a person to approve it, message by message or as a whole class of work.

Approval can cover one message or a whole class of work. For routine, low-risk replies, such as answers to common support questions, the owner of the customer process approves the wording, the limits and the cases that must go to a person. The agent works inside those limits, every reply is logged and samples are checked. Exceptions, new commitments and changes to agreed terms still need a person before anything is sent, and the agent cannot widen its own limits.

The renewal agent drafts the offer; it does not send it. The account manager reads it, changes what needs changing and sends it, or decides not to. The change and the reason are recorded.

When an agent has shown that it does a task reliably, its agent owner can explicitly grant it more autonomy for that task, and take it back if the results slip. The renewal agent might, after months of correct work, be allowed to update the renewal status in CRM on its own. The offer itself still goes to the account manager: it reaches a customer. Tasks that only a person may do never move. The decision stays within the Codex and the approval rules above. How much evidence is enough is still an open question. Good results do not give an agent permission to expand its own authority.

People in the loop

People are part of the loop because some decisions need judgment and accountability. That only works if their attention is used well.

- **Ask only when it matters.** A person is asked when their judgment is needed, not to confirm what the agent can safely do itself.
- **Do not overload people with decisions.** Someone asked to approve a hundred things a day stops reading and approves everything. That is worse than no approval at all, because it looks like control.
- **Show as little as needed to decide.** Present the decision, the recommendation and the sources, with the option to see more.
- **Make it easy to say no.** Changing or rejecting a proposal should be as easy as approving it.

Monitoring and logging

All harnesses are monitored, and the proactive ones most of all, because they run when nobody is watching. Monitoring answers three questions:

- **Is it working?** The harness runs as expected and does not fail without anyone noticing.

- **Is it right?** The results hold up when people check them, and errors that people report are followed up.
- **Is it worth it?** The cost of the work, including the time people spend checking it, is lower than the value it creates. This is judged at the three-month review.

For the renewal agent, *is it right?* means: would the account manager have sent the offer as drafted?

When something goes wrong, the agent owner is told. An agent can flag missing information in a draft and continue with the parts it can support. If uncertainty affects whether an action is permitted or whether a consequential decision is justified, it pauses that action and asks a person.

Every agent logs its work: what it did, on whose behalf, which sources it used, and what it decided and why. The logs make it possible to understand what happened, find problems and see who is accountable.

Logs are records of work, not automatically approved company knowledge. When they reveal missing or unclear guidance, the agent proposes a correction to the relevant knowledge owner. That owner reviews and approves changes to the Knowledge Base or the Codex, so later work can use the corrected guidance. Logs keep only what is needed to understand the work, with as little sensitive information as possible. Logs are protected from change, including by the agent that wrote them.

Communication

Between people and agents

- People can always reach their agents: ask what they are doing, change their direction or stop them.
- Agents report back where people already work, and they say clearly when they are stuck or need a decision.
- It is always clear when someone is talking to an agent and not to a person.

Between agents

- Agents need a shared space to hand over work, ask each other questions and collaborate.
- Conversations between agents are visible to the people responsible for them.
- The access rules apply to both sides of a handover: neither agent gains wider access by asking the other.

Handovers

The task's status, decisions, outputs and next steps are kept in the brain's memory of work. Another person or agent can pick up where the last one stopped, even when the agent or its harness is replaced.

Ownership and review

Every agent has:

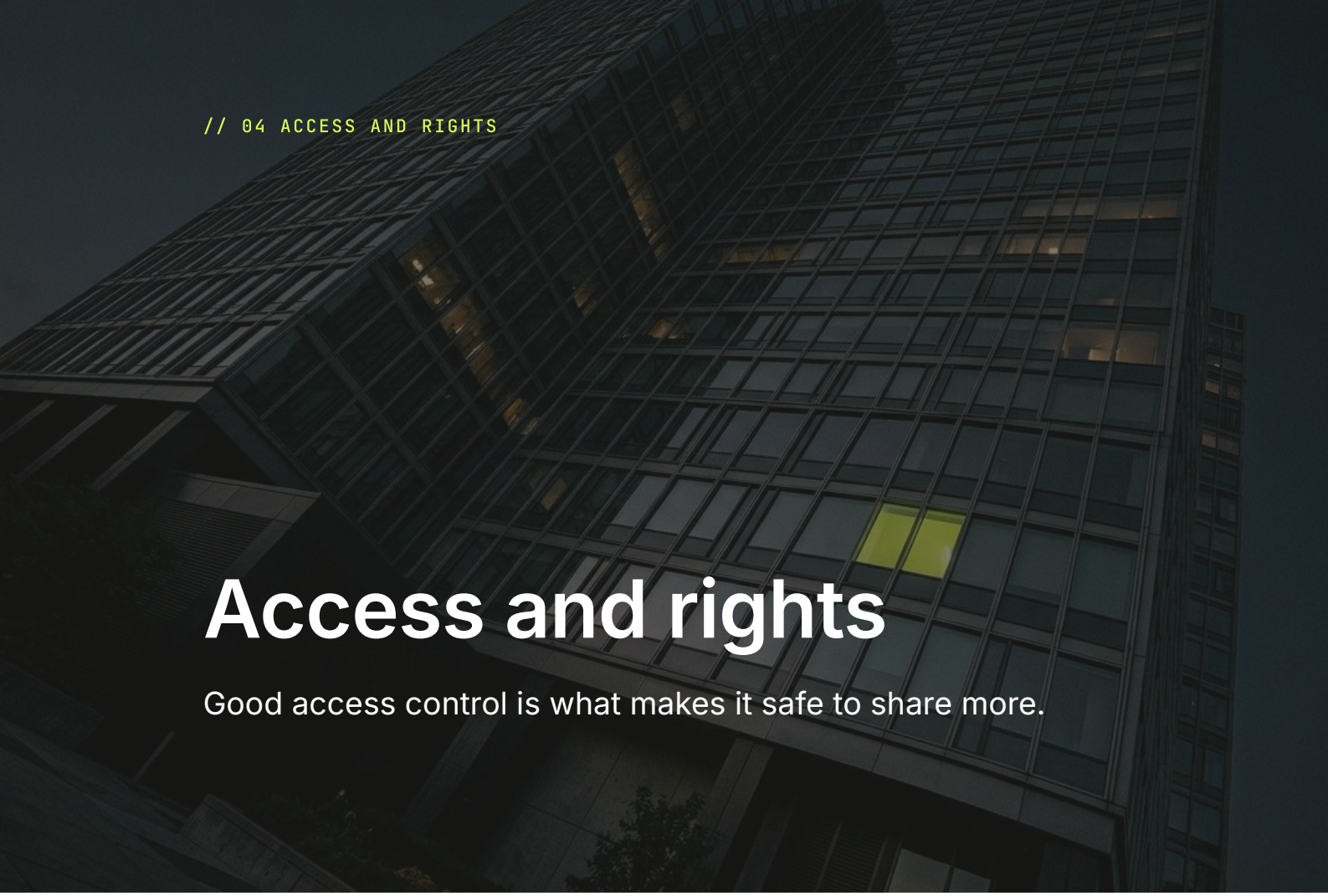
- **Its own identity.** It does not borrow a person's login.
- **Limited access.** It sees and does only what its job requires.
- **A named agent owner.** The owner answers for what the agent does, the same way a manager answers for the work of a team.

All agents are listed in the brain's shared register, with their purpose, their owner and what they may access. An agent that is not in the register should not be running.

Owning an agent is real work: following its results, answering for its mistakes and keeping it current. Agent owners need time for it, just as knowledge owners do.

Every three months, each department reviews its agents. Which tasks can agents now take over reliably? Which tasks stay with people, and why? What should the agents learn next?

For the renewal agent, the review starts with the offers account managers changed or turned down, and why. A change that keeps coming back is proposed as a lesson, and once its owner has approved it, the next renewal starts from it. Agent owners grant more autonomy where the evidence supports it and the approval rules allow it, and people move on to more valuable work.



Access and rights

Good access control is what makes it safe to share more.

When people and agents work from the same brain, access decides what each of them can see and do. Get it right, and a company can share far more than it dares to today. Get it wrong, and information reaches people it should never reach.

Why access matters more now

A document shared with the wrong person used to be read by one person, if at all. An agent reads everything it can reach and uses it in its work. Information shared too widely is no longer a small leak: it ends up in drafts, summaries and answers before anyone notices.

That is why access is not something added at the end. It is part of how the brain is built.

The brain owns the rules

Access rules are maintained in one place, in the brain, by the owners of the information, and apply everywhere: in the brain itself, in the systems connected to it, and for every agent. A person does not get one set of rights in the brain and another in CRM. When a connected system is replaced, the rules stay where they are.

Access is checked before information is retrieved, not after. An agent never sees something first and is then told it should not have.

Codex and Knowledge Base

The two parts of the brain follow different rules:

- **The Codex** is open to everyone inside the company, people and agents alike. There are no separate read permissions, because it only holds what everyone inside may read.
- **The Knowledge Base** has access per item. An agreement, a personnel record and a product guide may each have a different audience.

Most items get their audience from a few levels, such as company-wide, team, restricted and closed. The Codex is, in practice, the company-wide level.

Different rights for different things

Being able to read something is not the same as being allowed to do everything with it. The brain keeps these rights apart:

01

Read

See the information and use it in the work

02

Propose

Suggest a change; open to everyone who may read it

03

Change

Edit the content: the owner, or someone the owner has authorised

04

Approve

Decide that a change becomes the approved version: the owner

05

Publish

Share it outside the company

Permission to read something internally does not mean permission to publish it. A customer's story, for example, is theirs to share.

Agents act on someone's behalf

An agent never has more access than it needs for its job. When it works for a person, it never sees more than that person may see. When the renewal agent prepares an offer for an account manager, the only customer records it sees are that account manager's customers and agreements.

- **Every agent has its own identity.** It does not borrow a person's login, so it is always clear who did what.
- **Both sides of a handover must be allowed.** When one agent passes work to another, the receiving agent must be allowed to use the information for the task. Neither agent gains wider access by asking the other.

Closed spaces

Some work needs a room that only a few people can enter: a personnel matter, a sensitive negotiation, a confidential project. The brain supports closed spaces where only the people involved, and the agents working for them, have access. Content stays invisible outside the space. A small, named group knows the space exists, without the right to read what is in it, so its use can still be checked.

Opening up to partners and customers

When the rules are precise, the brain can open up beyond the company:

- **Partners** see what they need to sell and deliver with the company.
- **Customers** see what applies to them and their own solution.

They draw on the same maintained knowledge, so a fact is explained the same way to everyone. Only the access differs. A person with the Publish right approves every item shared outside the company, just as a person approves any message that reaches a customer or partner. Outside the company, it is always clear when an agent is speaking.

When things change

People change roles, leave the company or join a project. Agents are added and retired. Access must follow straight away, everywhere at once: in the brain, in connected systems and in every agent that works for that person.

When the brain is unsure whether someone may see something, the answer is no until the owner has checked. Access is also reviewed at set intervals, because some changes are missed. Who saw what, and who changed what, is recorded, so it can be followed up.

Oversight and incidents

Running the brain needs powerful permissions, but it does not give the people who run it the right to read everything. That access is limited to a few named people, logged, and reviewed by someone else. Nobody approves changes to their own access.

When information reaches someone it should not, or an agent acts on instructions from outside, it is handled as a security incident: reported at once, contained, and fixed at the cause in the brain, not only in the output.

The approved way must be the easiest way

If access is hard to get, people find their way around it. They copy documents into their own tools, paste information into services nobody has approved, and build their own shortcuts. That is how shadow tools appear.

The answer is not more rules, but an approved way that is easier than going around it. Asking for access should be simple, the answer should come quickly, and the reason for a no should be clear.



Data

An agent does not just use bad data. It spreads it.

A wrong figure that a person might have caught once ends up in a hundred briefs, drafts and answers before anyone notices. The more work agents do, the more everything depends on the information they start from.

People and agents need information they can trust, and clear decisions about what to keep. The brain owns that information, and connected systems help collect it and carry out the work.

Shared identities

A customer must be recognisable as the same customer across sales, support and finance. Shared identities connect the records, so information about one customer, product or agreement can be brought together wherever it was recorded. Matching records across systems is never finished: new systems and duplicates keep creating work. It is ongoing work with an owner, not a switch that is turned on once.

Example: preparing a renewal

When the renewal agent prepares an offer for an account manager, it brings together what the customer has bought, open support cases, earlier promises and unpaid invoices, within the account manager's access rights. If finance has not updated its figures, the gap is flagged for the account manager instead of hidden.

Shared definitions

Shared definitions make figures comparable. What counts as an active customer, or how a KPI is calculated, is maintained knowledge with a named owner. A figure means the same thing whether it appears in a sales report, a board summary or an agent's answer.

The same goes for the company's own words and for how things relate. The brain knows what the company's terms and abbreviations mean, and that a customer has agreements, an agreement covers products and a support case concerns one of them. Without that, an agent can know every fact and still miss how they connect.

Source and freshness

Information shows its source and when it was updated. Missing information stays visibly missing. An old figure does not become current because an agent repeats it.

How information was produced

How information was produced, and who reviewed it, is recorded when it is created or changed. Some information comes from people, some from agents, and much from both together. Knowing which, and whether someone has reviewed it, tells people and agents how far they can rely on it.

This is hard to reconstruct afterwards, so it is captured at the source. An agent-generated draft remains a draft until it is approved.

Change history

Important information has a history. It shows what changed, who changed it, or which agent did and for whom, when, and why where it matters. Maintained knowledge keeps its earlier versions. Information that changes all the time, such as customer activity, keeps a record of its changes or dated snapshots instead.

History answers a question that matters when a decision is reviewed later: what was known at the time? It also keeps that apart from what has been learned since about that time.

History also makes it possible to recover from mistakes, including mistakes made by agents. Restoring information follows the same approval rules as changing it. But putting information back does not undo what has already happened: an invoice that was sent, or a message a customer has read, needs its own correction.

History follows the same access and retention rules as the information itself. Old versions are kept out of everyday view, and when information must be deleted, such as personal data, it is deleted from its history too.

Some information is the original itself: a signed agreement, an approved policy, a decision that commits the company. The brain then keeps it the way the law and the company's own rules require: it cannot be changed once signed, every access is recorded, it is kept for as long as the rules say, and it can be restored from a tested backup. Amendments are added as new versions; the original stays as it was signed.

Fix errors where they start

When an error keeps coming back, the cause is fixed, not the symptom. If an import repeatedly assigns the wrong payment status, the fix goes into the import and the affected records, instead of correcting every brief that uses them. A recurring correction is a sign that the real work has not been done yet.

When sources disagree, the disagreement is made visible and goes to the owner for resolution. Corrections reach the connected systems and the work that depend on them.

Collect what the work needs

Trustworthy information still has to earn its place: more data is not better data. AI makes it tempting to keep everything: storage is cheap, and an agent can search anything. But an agent that searches everything also finds everything, including old versions, half-finished drafts and things that were never true.

Information is collected and kept for a clear purpose, with an owner responsible for keeping it relevant. A new source is added because a piece of work needs it, not because it is available.

What data is for

Data earns its place by helping people act, decide, learn, explain what happened or continue work:

01	**Act**	What the customer has bought and their open issues, needed for tomorrow's meeting
02	**Decide**	Figures that are compared, followed up and acted on
03	**Learn**	What was tried, what failed, and examples of good results
04	**Explain**	Why a decision was made, and whether a change actually helped
05	**Continue**	Status and next steps, so someone else can pick up the work

Some data is also kept because it must be, such as accounting records and signed agreements.

Why hoarding hurts

01

Noise

The more irrelevant material there is to search, the harder it is to find what matters. More data can make answers worse.

02

Stale facts

Data nobody uses is data nobody maintains, and it slowly becomes wrong while still looking current.

03

False confirmation

The same claim copied into three places looks like three sources agreeing, when it is really one. Each fact has one home, and everything else points to it.

04

Conflicts

Old material contradicts current guidance, and an agent may not know which one applies.

05

Risk

Every record about a person or a customer is something that can leak or be misused.

06

Upkeep

Storage is cheap. Owning, reviewing, interpreting and protecting data is not.

Keeping is not the same as making available

Some information should be kept but not used every day: history, earlier versions and records that must be kept. It is archived with its date and status, and looked up deliberately when it is needed. Current guidance stays in front, so people and agents do not stumble on what has been replaced.

When people and agents mark what helped, what was wrong and what led nowhere, the knowledge that proves useful comes first next time, and the dead ends sink. Usefulness only decides the order among what is current and approved. It never puts an old version ahead of current guidance, and it does not decide what must be kept.

What stays and what fades

Not everything should be kept for the same length of time, and not everything should be judged the same way.

- **Use keeps knowledge in view.** Information that helps someone stays easy to find. Information nobody uses moves out of everyday view, and its owner then decides whether to review, archive or delete it, within the company's retention rules. Low use alone never deletes anything.
- **Some information stays available regardless of use.** Current core facts and key procedures remain easy to find, and required records are kept for the period that applies.
- **Lessons do not fade from lack of use.** A lesson exists because something once went wrong and someone paid to find out why. A lesson nobody has needed for a year has not gone stale. It is a mistake that has not happened again yet. A lesson that is no longer true is corrected or replaced, and the history of what was learned is kept.

Usefulness is judged for the people and agents who need it, not by company-wide popularity. A definition finance uses every week matters, even if nobody else opens it. Use is followed for the work and the roles involved, never as a measure of individual people, and within the access rules, so a closed space is never revealed. If information that should be used is not, its owner finds out why before archiving it.

Collect on purpose, but not too little

Collecting less must not mean losing what matters. Some data is only valuable if collecting it starts early. A baseline has to be measured before a change, or nobody can show whether the change helped.

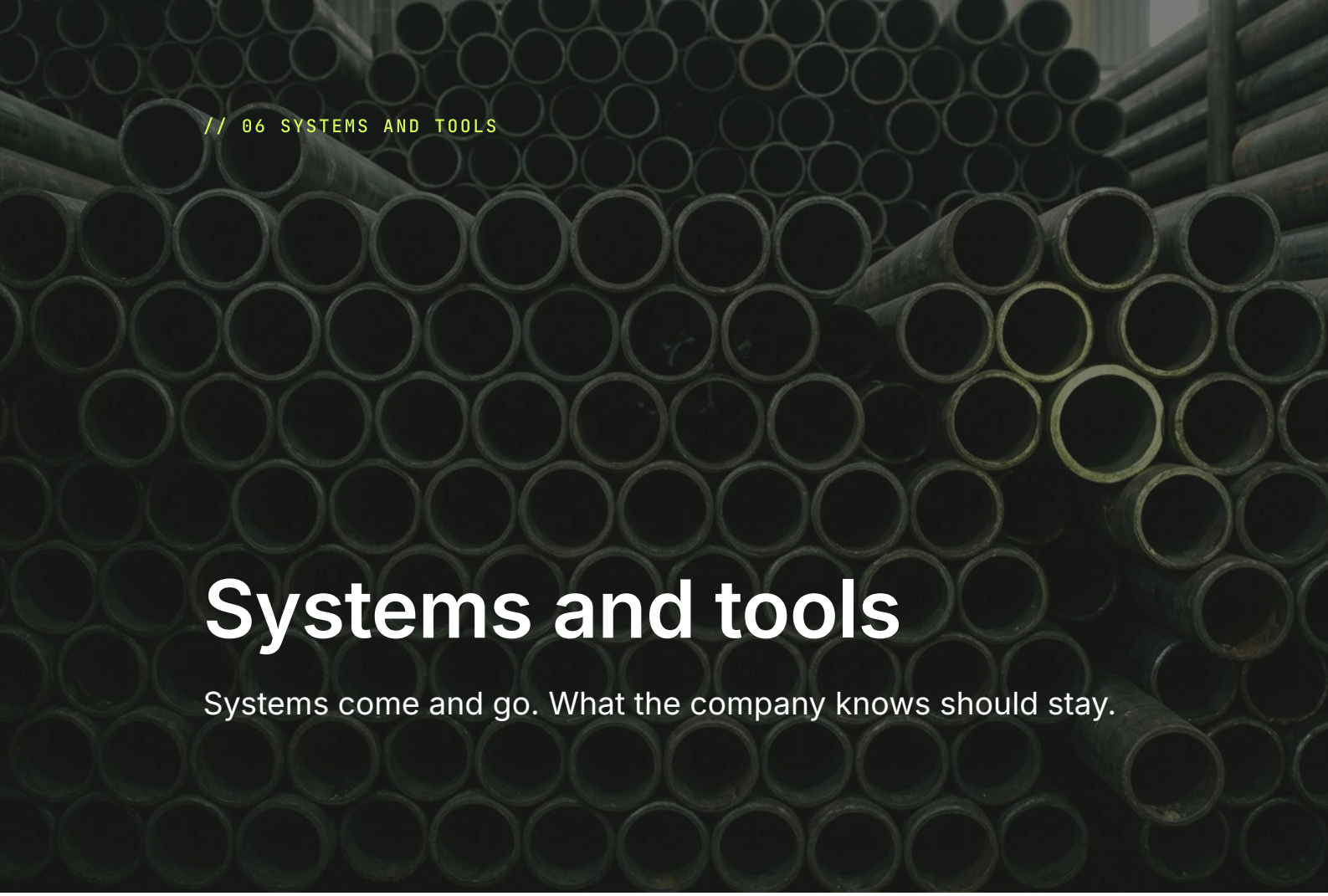
The failures are kept too, not only the successes, and the evidence someone would need to challenge a conclusion.

Four questions before keeping data

1. What work or question does it support?
2. Why keep this detail rather than a summary?
3. Who owns it?
4. When is it reviewed for keeping, archiving or deleting?

Ordinary information that nobody has used for a year is a signal for its owner to review it. Keeping data creates a responsibility: to preserve its meaning, and to decide when it has stopped being useful.

Reliable data lets people and agents build on each other's work without checking the same facts from scratch.



Systems and tools

Systems come and go. What the company knows should stay.

Every system and tool connects to the brain and has a named owner from the day it starts to the day it is retired.

A system here means software that many people share to do their work and record what happens, such as CRM or finance. The records themselves are owned by the brain. A tool means what people and agents use to do a piece of work. The line is not sharp: CRM is a system for the company and a tool for the salesperson. The same rules apply to both.

Register and monitoring

The brain keeps one register of every system, tool and agent the company uses, whether bought or built, and of the connections between them. Each entry shows its purpose, its owner, its status, who may use it, what information it handles, what it costs and when it is next reviewed. The register also shows what

depends on what, so an owner can see which work is affected when something fails.

Everything in the register is monitored. Is it working? Is it doing the job correctly? Whether it is worth what it costs, including the time people spend working around it, is judged at each review. The brain does not run every check itself; other monitoring feeds it the results. When something fails, or the checks themselves stop arriving, the owner is told, and automated work that depends on it can be paused.

The register follows the same access rules as everything else, so entries are visible only to people allowed to see that work, and a closed space is recorded without its content or purpose. It never holds passwords or sensitive content.

A system, tool or agent that is not in the register should not be running. The approved way to add one must be easy, so nobody needs to go around it.

Systems connect to the brain

People and agents use systems suited to their work. The brain owns the information and the rules that govern it. Connected systems receive the information they need and return what the work produces, so company knowledge stays available beyond any one application.

Each connection has a named owner and a clear purpose: what information moves, in which direction, and how current it needs to be. Selected systems update the brain continuously where possible, otherwise at set times.

Reliable updates

People and agents can see when information was last updated and whether a connection has failed. The owner is told when updates stop arriving as expected. Work that depends on missing or outdated information waits, or is clearly marked as incomplete. Before work continues on the basis of a change, it must be possible to see that the information it depends on has been updated. Conflicting changes go to the information owner for resolution.

Each connection is checked for more than whether it is running. Is information arriving when expected? Is it complete? Does it look reasonable compared with before? A connection that goes quiet is checked, not assumed to have nothing to send, and information that arrives but looks wrong is held back and flagged to its owner instead of being passed on.

For example, the finance system sends the brain each department's actual costs when the month is closed. If the figures are late, are missing for some departments, or suddenly show one department's costs ten times higher than usual, they are held back and the connection owner is alerted. An agent preparing next year's budget then shows finance the gaps, instead of drafting a budget on wrong numbers.

The same rules everywhere

Connected systems enforce the access and approval rules set in the brain. Their connections distinguish between reading information, proposing a change and carrying out an approved action. If a system cannot enforce the rules for some information or action, that part of the work cannot run through it until the gap is resolved.

Choosing tools

What good tools are like

- **Easy to use.** People can learn it quickly and do their work without workarounds.
- **Easy to connect.** Information flows to and from the brain in the directions the work needs, and access rules are enforced.
- **Usable by agents as well as people.** An agent can read from it and act through it, within its rights.
- **Replaceable.** The information belongs to the brain, not the tool.
- **A good fit for the work.** The capabilities in use justify its cost and complexity.

01

When to buy

Buy when the product meets the need today and the vendor is heading somewhere worth following. A roadmap worth believing in is worth paying for: the vendor carries the development, security and maintenance, while the company's named owner still answers for how it is used.

If a vendor is not heading somewhere worth following, that is a reason to look at the alternatives: building, another supplier, or changing how the work is done. It is not proof that building is right.

Model providers and other cloud services are suppliers like any other. What they may see, where it is stored, whether it may be used to train their models and how it is deleted is agreed in the contract by the service's named owner, and checked before company information reaches them.

02

When to build

Build when a tool shaped around how the company works would serve it better, especially where it pays for much more than it needs and a focused tool would do the job. AI has made it far cheaper to build focused tools for a company's own work, and paying for capabilities nobody needs can be a leak. Recreating the full depth and reliability of an established product is another matter.

Building is not free: what is built must also be run. Everything built:

- has a named owner who answers for running, maintaining and eventually retiring it, and is listed in the register
- is reviewed for security before it handles company information, however quickly it was built
- comes with checks that show it does its job, and clear rules for how it exchanges information
- can be run by an agent: it shows what happened, reports problems and follows written recovery procedures, within its rights

An agent may test a change against those checks; the owner decides whether it goes into use.

Replacing systems and tools

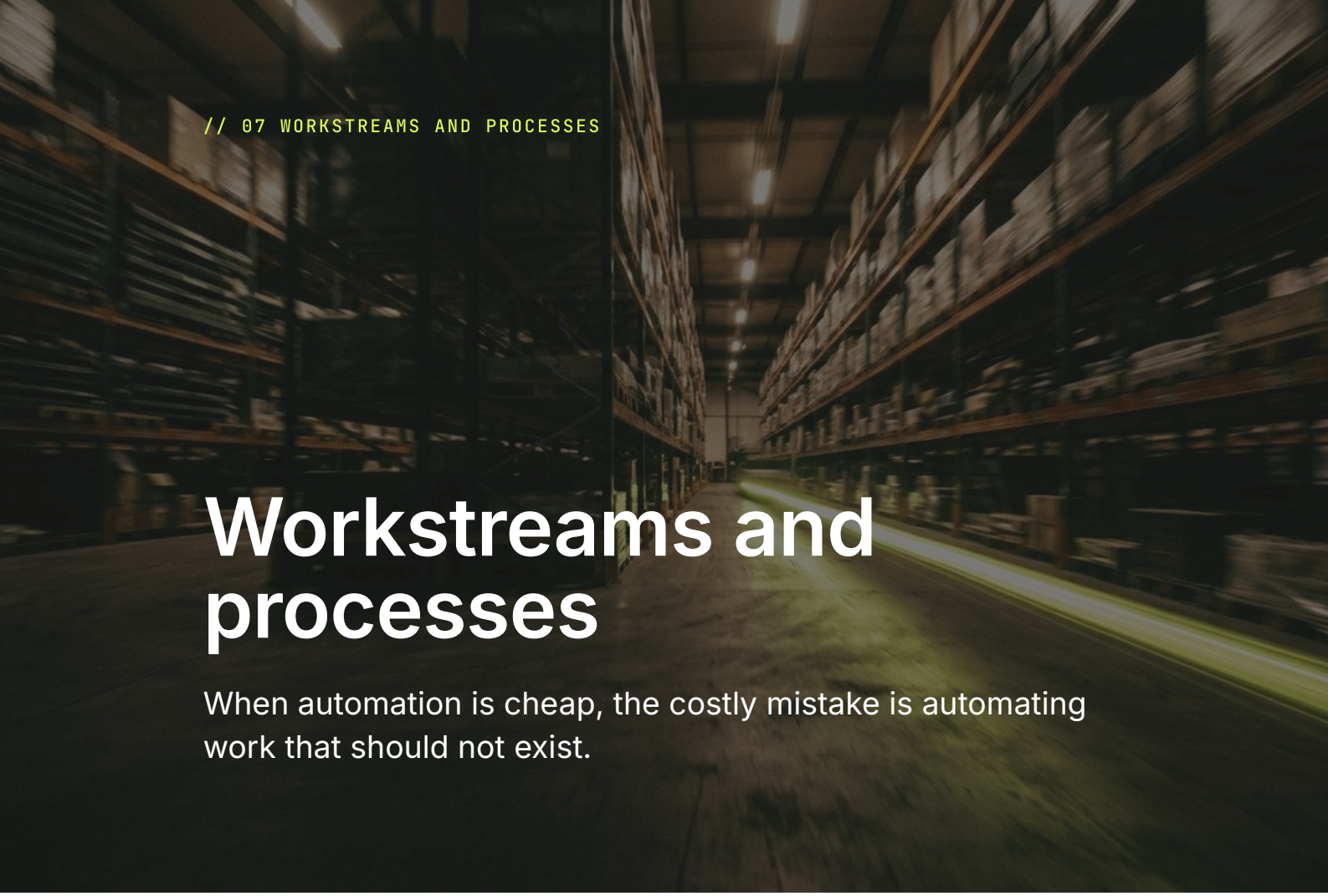
Signs that a system or tool should be reassessed: people work around it, only a small part of it is used, it does not connect to the brain, the company's information is locked inside it, or its cost grows faster than its value.

Using only a small part of a tool is a reason to compare, not proof of poor value: the part in use may be the part that matters most.

Before replacing, compare honestly: total cost over time, maintenance, reliability, what is lost by leaving the vendor, the cost of moving, and who will run the replacement.

Replacing a system or tool preserves the information, history and relationships needed to continue the work, within the retention rules. The replacement receives what it needs from the brain and follows the same access rules. Before the old one is retired, its owner checks that the work can continue in the replacement and that updates reach the brain correctly.

Because the brain owns the information, replacing a system or tool changes where the work happens, not what the company knows.



// 07 WORKSTREAMS AND PROCESSES

Workstreams and processes

When automation is cheap, the costly mistake is automating work that should not exist.

By 2029, much of the routine work will be done by agents, because automating it has become cheap. That makes one question more important, not less: should this step exist at all? An agent carries out a needless step perfectly, and at scale. So before a step is handed to an agent, it is questioned. Every requirement has a named person behind it who can say why it is there. What is not needed is removed, what is left is simplified, and then it is automated.

A workstream brings related work together around an outcome. It can be ongoing, such as answering customer requests, or time-limited, such as a project. A process describes how the work is carried out. Processes are maintained once in the brain and shared by the workstreams that use them.

Every process is documented in the brain, whether people or agents carry it out. A process that lives only in someone's head is hard to improve, hand over or give to an agent. Documenting does not mean writing a manual: it means the steps,

who does each one, where the work waits or changes hands, and what is still unclear. Agents can help. They interview the people who do the work, turn notes or a whiteboard photo into a draft, and write down what is uncertain as a question instead of guessing.

There are two roles. The workstream owner answers for the outcome and can pause or stop the work. The process owner approves the shared process and keeps it current. Often one person holds both.

Start with the work

Start with the work, not the technology. Look at how it runs today: where people wait, where work is done again, and where people have built their own workarounds, because workarounds show where the process breaks. Sometimes the answer is a simpler process rather than another agent. It does not take months of analysis: a short experiment on real work, compared with how it ran before, shows more than a long study.

Not every step is equally ready for an agent. Simon Wardley, who maps how business activities evolve, describes work moving from new and uncertain, through custom-built and product, to commodity. The more settled and well understood a step is, the more of it an agent can carry. The newer it is, the more it needs people's judgment, and it is documented by its purpose and boundaries until a pattern appears.

More agent activity is not a result, and time saved in one team can be work created in another. Suppose sales uses an agent to draft quotes. The drafts arrive faster, but the technical team now has to check integration requirements, correct prices and chase missing details. The result worth measuring is a correct quote reaching the customer sooner, at a cost that makes sense, counted across both teams.

Two measures reveal most of it: the time from request to result, waiting included, and the share of incoming work each step can use without first putting it right.

How work starts

Work starts with a request, an event or an agreed schedule that says what is needed, who is responsible and by when. A trigger starts work; it does not grant new authority.

01

A customer's request starts the quote. An agent gathers what the customer has bought before, the current prices and the product rules, with sources and gaps shown, as described in Data. The process says how a repeated or changed request is handled, so one request does not produce competing quotes. If something essential is missing, the agent prepares what it can and waits on the rest.

Steps and handovers

For each step, the process says who does it: an agent, an agent that prepares the work for a person to approve, or a person, following the levels in Agents.

Much of the time lost in work is lost between steps. A handover in the workstream's space in the brain carries what was produced, what was checked, what is still unknown and who takes it next, within the access rules. Preparing work for someone else to review is part of the job, for people and agents alike.

02

When the quote moves from sales to the technical team, the team gets the draft with its sources and a clear note: prices are checked against today's list, integration requirements are not. The technical team can focus its review where it is needed.

Decisions and approvals

Every process shows where decisions are made and who makes them, following the approval rule in Agents.

Oversight is not a click. It is the boundary set before the work starts: what agents may touch, what counts as irreversible, and what stays off-limits. Approvals then go where judgment is needed, with the evidence in front of the person. If more agents only make the approval queue longer, that is scaling the wrong part of the work.

03

The salesperson approves the quote before it is sent. A non-standard discount goes to the pricing owner named in the process before it is offered.

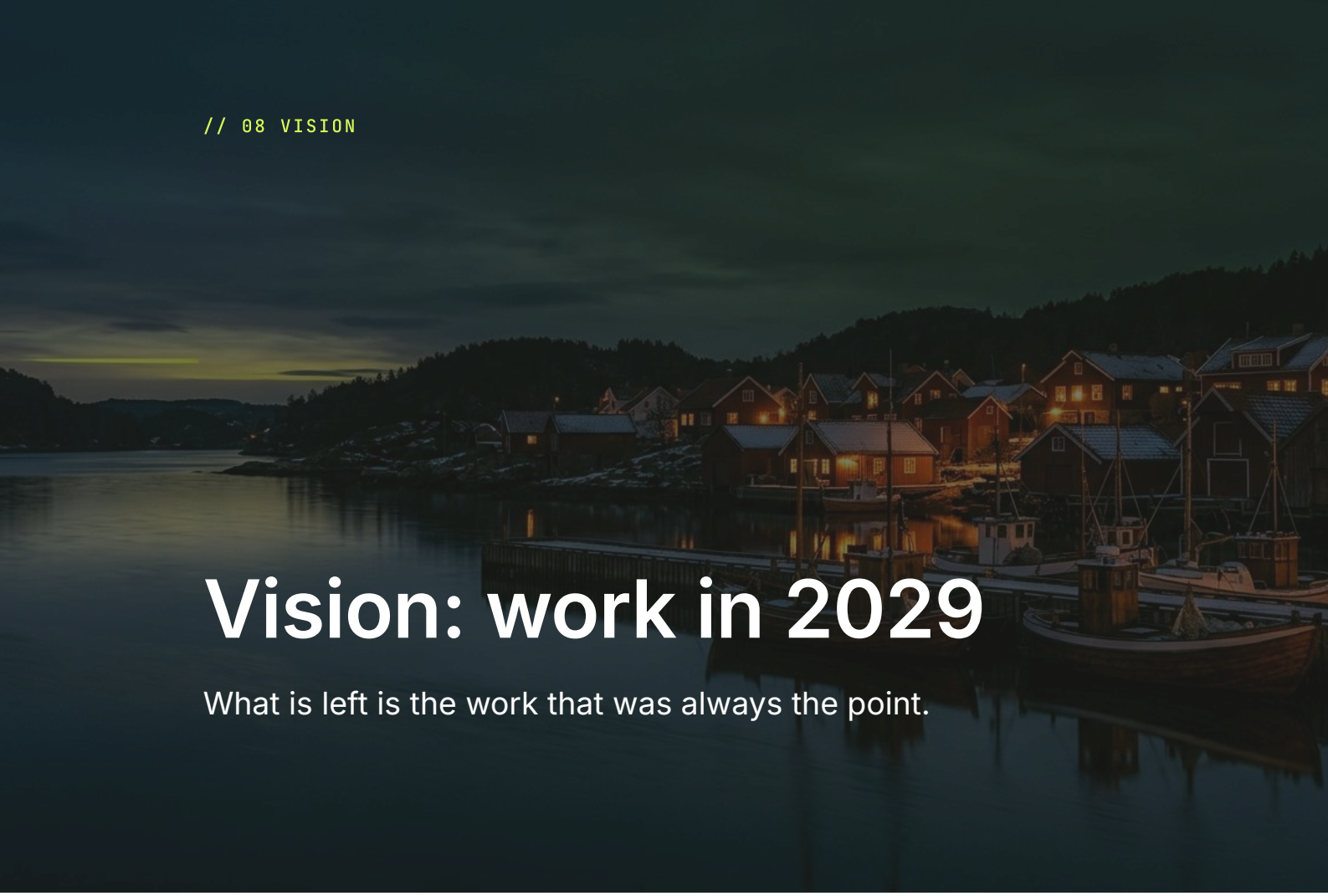
Memory and improvement

Each run leaves its decisions, results and lessons in the brain. When the same problem comes back, the fix goes into the process, not into another workaround. Lessons are proposed to the process owner, who approves changes to the shared process. Because it is kept once in the brain, the improvement reaches every workstream that uses it.

Sometimes the process itself turns out to be wrong in the middle of the work. Then the workstream owner pauses the affected work, or lets a person continue within the existing access and approval rules, and the process owner is told at once. The process is fixed before it runs again.

After a change, the work is measured again: the result, and the total effort including checking and rework. Then the workstream owner keeps, fixes or stops the change. If quotes keep failing the same integration check, the step or the product data is fixed, and the agent is not given to more people until the failure stops.

// 08 VISION



Vision: work in 2029

What is left is the work that was always the point.

It is a Tuesday in 2029. Nobody digs through old emails to find what was promised to a customer. Nobody asks the same colleague the same question for the third time this month, or rebuilds a report someone already made. That work happens quietly in the background, done by agents that look it up in the brain.

Time to really listen to a customer. Time to think a problem through before deciding. Time to sit next to a new colleague and show them how things are done, not just where they are written down. Time for the idea nobody has had yet, and for the conversation that leads to it.

Room for what people do best

01

Sales: ready before the meeting

Two days before a first meeting with a new customer, a slide deck arrives in the salesperson's inbox. The agent has built it from what the customer has published, the notes from the first call and the product knowledge in the brain. It covers the challenges the customer faces, how the company can help with each one, and a proposal ready to sign. Every claim links to its source. The salesperson spends twenty minutes on it instead of two days. She cuts a slide that answers a question the customer never asked, and adds what no document holds: the existing customer she knows they will want to talk to. She approves the proposal, and the meeting is about the customer, not the slides. The customer signs that week.

02

Marketing: the story was already there

Next week's blog post is due. Instead of chasing other teams for ideas, the content team turns to the agent, which already knows what the company is working on: a feature that ships on Thursday, the questions support hears again and again about the problem it solves, and the customer feedback that started it all. It also knows where the company is heading, because the direction the board has set for everyone to know is in the Codex. It drafts three angles, each linked to its sources and each pointing in that direction. The team picks the one built on the support questions, since that is what customers are really asking, and rewrites the opening in the company's own voice. A customer quote waits until the customer has said yes, and a person with the Publish right approves the post before it goes live.

03

Support: a case no process covers

A customer has just bought another company and wants to merge the two accounts, each with its own contract and its own users. No process covers it, and the Codex is clear: when no process covers a case that affects a customer, a person decides. The agent hands the case over, and the brain points the support specialist to the colleague who handled the last account merger. She decides how to merge the accounts, and the customer has an answer the same day. What she learned goes to the support process as a proposal.

04

Finance: approved in minutes

The sales proposal includes a discount above the standard level, so it waits for the pricing owner. She sees the request, the customer's usage and the margin side by side, and approves it in minutes. It is month end too. An agent drafts the commentary on recurring revenue and churn, and every figure means the same thing in every report, because each definition has one owner. Finance checks, corrects and signs off.

05

A new colleague: his own calls in three weeks

He joined three weeks ago. From his first day, he could find the processes, the reasons behind them and who to ask about the rest. He sees what his role needs, and nothing he has no reason to see. What the brain cannot teach him, he learned by sitting in on customer calls. Today he runs his first one alone.

06

Leadership: an agent that does not just agree

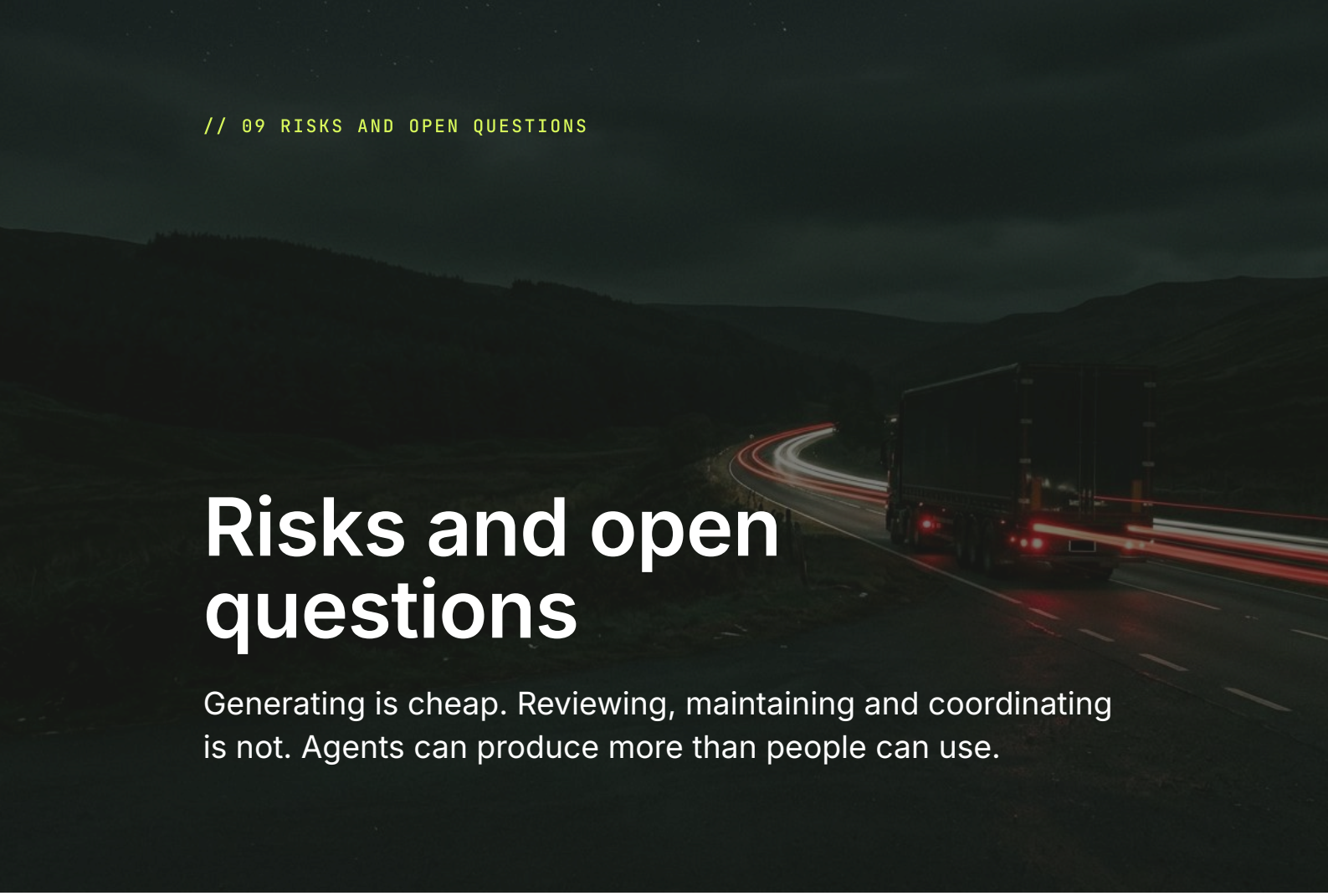
The management team is deciding where to grow next year. They ask an agent to check the plan, and it does not simply agree. The figures behind one assumption are from last year, and two customer segments are shrinking, not growing as the plan assumes. The team changes the plan. Six months later, when they review the decision, they can see what they knew at the time and why they chose as they did.

What changed

Nothing in this day needs magic. The agents prepare, check and draft. The people listen, decide, say no and take responsibility. The brain holds the notes from the first call, the questions support keeps hearing and who knows how an account merger works. People still approve what reaches the customer, and still own the result.

Some tasks disappear, some roles shrink or change, and new work becomes possible. Time saved does not by itself decide what happens to a job: leaders remain responsible for those choices and for helping people learn new work. Decisions about people stay with people.

Less of the day goes to rebuilding context. More goes to judgment, to customers and to each other. That is the work worth getting back to.



Risks and open questions

Generating is cheap. Reviewing, maintaining and coordinating is not. Agents can produce more than people can use.

Running a company on a shared brain changes what can go wrong. When people and agents work from the same knowledge, the same rules and the same connections, the brain's strengths become its risks: what is right spreads quickly, and so does what is wrong.

Responsibility can blur between people, agents and systems. Judgment can fade when the brain always has an answer. And the whole organisation can come to depend on something it does not fully control. None of this is a reason to stand still, but each of these risks needs an owner and a response.

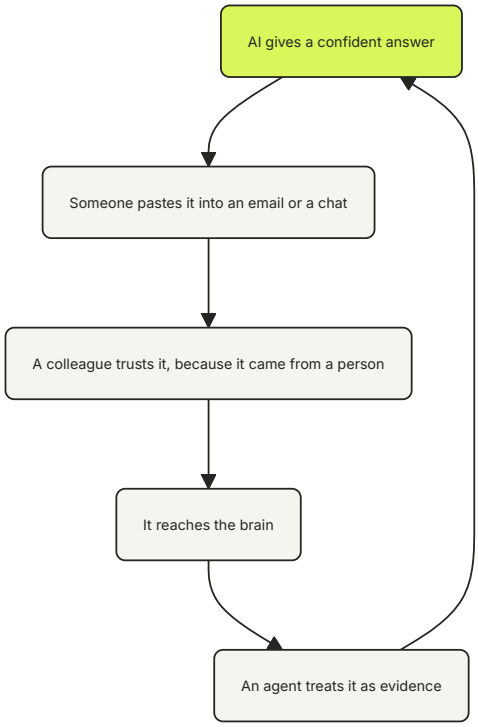
Risks

AI that agrees, and people who stop checking

Two risks feed each other. AI tools tend to agree with the people using them, and people tend to stop checking what agrees with them. Researchers at Wharton

call the second one cognitive surrender: adopting an AI's answer as one's own, instead of using it to think. In their experiments, when the AI was wrong, people still followed it four times out of five. They did worse than people with no AI at all, and felt more confident while doing it. Time pressure made it worse; feedback and accountability made it better.

It usually starts small. Someone asks an AI a question, gets a confident answer that matches what they already thought, and pastes it into an email, a chat or a document. Now it looks like their own view. A colleague trusts it because it came from a person, and sooner or later it reaches the brain, where an agent treats it as evidence. Something nobody checked has become company knowledge, and AI is now agreeing with AI.



Over time, judgment fades, not because anyone decided to stop thinking, but because the agent always did the first draft and it always looked fine. A company can drift this way without anyone noticing.

The goal is not less AI. It is to offload, not surrender: to hand work to the AI while people keep thinking. The response: a culture where challenge is normal, agents built to challenge, people who make the decisions that matter and own them, approvals that are real rather than routine, and feedback on what turned out to be right. The brain records whether a person or an agent produced a piece of information and whether someone reviewed it, and pasted or generated content stays a draft until it has been reviewed.

Watch for: approvals that are almost never changed, and answers accepted faster than they could have been read.

Too much depends on the brain

When knowledge, rules and connections all run through one brain, an outage stops many departments at once, and a breach exposes more than any single system could. Central ownership must not mean one fragile point.

The response: access checked before anything is retrieved, closed spaces for sensitive work, tested backups and recovery, held to the same standard as the finance system's because the brain keeps original agreements, and a clear decision about which work can continue without the brain. When one connection fails, only the work that depends on it is paused.

Watch for: work that stops entirely when the brain is slow, and recovery plans that have never been tested.

One mistake becomes many actions

A wrong instruction or assumption in shared guidance is not one error. Every agent that uses it repeats it.

The response: test changes on limited work first, check important outcomes independently, and pause the work that depends on something as soon as it fails.

Watch for: the same error turning up in several places at once.

Outside input steers the agents

Agents read emails, tickets and documents from outside the company. Hidden instructions in that text can try to make an agent do things nobody asked for, across tools and at speed.

The response: outside text is treated as information, never as instructions. Agents get only the access their task needs, what they do is logged, and anything that leaves the company or cannot be undone needs a person's approval.

Watch for: agents taking actions nobody asked for, or acting on text that came from outside.

Ideas converge on the average

When everyone works from the same guidance and the same models, strategies and texts start to sound alike. Two agents agreeing is not two pieces of evidence.

The response: agents gather evidence and challenge reasoning, while people own the judgment calls and the bets, and actively seek out the views that are missing.

Watch for: proposals that all look alike, and decisions where nobody argued for an alternative.

Replaceable in principle, dependent in practice

Owning the company's information does not make models, integrations and accumulated know-how easy to replace. Prices, terms and how a model behaves can change, sometimes overnight.

The response: the brain owns the company's knowledge, checks and procedures. A replacement is tested on real work before anyone depends on it, and the real cost of moving is compared, not the promised one. See Systems and tools.

Watch for: work that only one model or vendor can do, and procedures that exist only inside a tool.

Visibility becomes surveillance

The brain records what is done, on whose behalf and how it went. If that turns into personal scores, people stop sharing mistakes and stop recording what failed, and the brain loses exactly the lessons it needs.

The response: say what monitoring is for and who sees it, follow use by work and role rather than by individual, treat errors as lessons for the system, and involve the people affected in setting the rules.

Watch for: fewer failures being recorded, and people asking who will see their logs.

More activity, less value

Generating is cheap. Reviewing, maintaining and coordinating is not. Agents can produce more than people can use.

The response: measure results and total effort, including review and rework, against how the work was done before, and retire work that does not justify itself.

Watch for: more output without better results, and review queues that keep growing.

The brain is bypassed

If keeping the brain current feels heavy, people go back to chats and local files, and the brain becomes an empty shell.

The response: connected systems feed it automatically, agents draft updates for owners to approve, owners are given time for the work, and unused material fades.

Watch for: answers given in chats that never reach the brain, and knowledge past its review date.

Responsibility falls between owners

Agent owners, system owners, knowledge owners and workstream owners can each do their part while nobody owns a harmful result of the whole.

The response: a named person is accountable for each workstream's outcome and can stop it, and the brain's owner coordinates when something goes wrong across systems.

Watch for: incidents where every owner says their part worked.

Open questions

- **How do people build judgment?** Experienced people learned by doing the groundwork. If agents prepare everything, how do newcomers develop the judgment to challenge and approve? *Next step: follow how newcomers learn during their first year, and keep some groundwork with people on purpose.*
- **Which friction is useful?** Writing from scratch often exposes weak thinking. Which work should stay manual, even when it is slower? *Next step: let each workstream owner name the work that stays manual, and review that choice.*
- **What replaces the human glue?** Fewer handover conversations save time. What replaces the informal trust they used to build? *Next step: notice where trust drops when a handover moves into the brain, and keep the conversations that matter.*

- **What happens when the brain is down?** Which work must continue without the brain or an AI provider, and with what limits? *Next step: decide it in advance for each workstream, and test it.*
- **When does an agent earn more autonomy?** What evidence is enough to give an agent more room, and what should take it back? *Next step: agree the evidence task by task in the three-month review, and record each decision.*
- **Local models or frontier models?** Models run inside the company's own environment keep information in-house, but give up some capability and must be maintained. Frontier models bought from vendors are the most capable, but information leaves the company and dependence on the vendor grows. Which work belongs where? *Next step: sort work by how sensitive its information is, and test both kinds of model on real tasks before deciding.*
- **Who changes shared guidance?** When departments disagree, who decides, and how can the decision be challenged? *Next step: name an owner for each piece of shared guidance, and a way to appeal.*

Terms

Term	What it means
The brain	The shared record of what the company knows, how it works and why. Built from connected tools and owned as one whole.
Codex	What guides the company and why: values, principles and the reasoning behind important decisions. Everyone inside may read all of it.
Knowledge Base	Maintained knowledge, live context and memory of work, each with its own audience.
Maintained knowledge	Approved information such as agreements, product knowledge and processes. Every item has a named owner.
Live context	Current information from connected systems, shown with its source and update time.
Memory of work	The decisions, outputs, drafts and lessons the work leaves behind. A draft stays a draft until it is reviewed.
Agent	AI that can take steps toward a task: it finds information, uses tools, drafts, checks its work and hands over the result.
Harness	The instructions, knowledge, tools, access and rules for stopping that let an agent do a job.
Agent owner	The named person who answers for what an agent does.
The approval rule	Anything that reaches a customer, commits the company or cannot be undone needs a person to approve it, message by message or as a whole class of work.
Closed space	A part of the brain that only the people involved, and the agents working for them, can access.
Publish right	The right to share information outside the company. A person with it approves every item that leaves.
Register	The brain's single list of every system, tool and agent the company uses, with its purpose, owner and status.
Connected system	Software such as CRM or finance that receives what it needs from the brain and returns what the work produces.
Workstream	Related work brought together around an outcome. The workstream owner answers for the outcome.
Process	How the work is carried out. Kept once in the brain and approved by the process owner.
The brain's owner	The named executive who answers for the brain as a whole.



// 10 SOURCES

Sources

Research

- **Wright, K., Fabian-Gill, P., Ellinger, T., Tobin, K. and Searson, H. "What Sycophantic AI Can Cost Your Organization."** Harvard Business Review, 4 November 2026. https://hbr.org/2026/11/what-sycophantic-ai-can-cost-your-organization
  - Describes BT Group's randomised trial with 2,400 employees, in which an AI designed to challenge produced three and a half times as many goals rated excellent as an AI that tended to agree.
  - Background to Culture and values ("Why agents need the culture").
- **Shaw, S. D. and Nave, G. "Thinking—Fast, Slow, and Artificial: How AI is Reshaping Human Reasoning and the Rise of Cognitive Surrender."** The Wharton School, University of Pennsylvania. Working paper, SSRN, version of January 2026. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=6097646
  - Introduces the term cognitive surrender. In three experiments with 1,372 participants, people followed an AI assistant even when it was wrong in about four out of five cases, and became more confident when using it. Time pressure increased surrender; incentives and feedback reduced it.
  - Background to Risks and open questions ("AI that agrees, and people who

stop checking").

Ideas

- **Boyd, J. R. "Patterns of Conflict."** Briefing, 1986. The briefing in which the US Air Force strategist John Boyd introduced the decision loop observe, orient, decide, act (the OODA loop). His later briefing "The Essence of Winning and Losing" gives its most formal definition.
- **Crowley, B. "The OODA Loop."** The Decision Lab, 30 August 2021. https://thedecisionlab.com/reference-guide/computer-science/the-ooda-loop
  - A popular summary of Boyd's loop, describing orientation, shaped by experience, culture and mental models, as central to it.
  - Background to The brain (introduction). A secondary source: Boyd's ideas are used here as a way of thinking, not as evidence.
- **Patel, D. "Andrej Karpathy — AGI is still a decade away."** Dwarkesh Podcast, 17 October 2025. https://www.dwarkesh.com/p/andrej-karpathy
  - Karpathy on why today's models do not remember what they are told ("You can't just tell them something and they'll remember it"), and on AI changing work gradually rather than in one leap.
  - Background to The brain (introduction).
- **Accenture. "Intelligent Digital Brain: Uplifting enterprise competitive advantage."** Technical white paper, 2026. https://www.accenture.com/en/insights/ai-data/intelligent-digital-brain
  - Describes an architecture that turns generic AI models into company-specific ones, on the grounds that generic models "know a lot about the world, but little about your company."
  - Background to The brain (introduction). A vendor's white paper: used here for its diagnosis, not as evidence.
- **Wardley, S. "Wardley Maps: Topographical intelligence in business."** Book published in chapters online, licensed CC BY-SA 4.0. https://medium.com/wardleymaps. Introduction at https://www.wardleymaps.com/learn.
  - Describes how business activities evolve from genesis through custom-built and product to commodity, and why each stage calls for a different way of working.
  - Background to Workstreams and processes ("Start with the work"). Used here as a way of thinking, not as evidence.

Regulation

- **EU AI Act, Article 50: Transparency obligations for providers and deployers of certain AI systems.** AI Act Service Desk, European Commission (DG CONNECT). https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-50
  - Requires that people are informed when they interact directly with an AI system, unless it is obvious, and that certain AI-generated content is marked as such.
  - Background to Agents ("It is always clear when someone is talking to an agent and not to a person") and to Data (recording how information was produced).

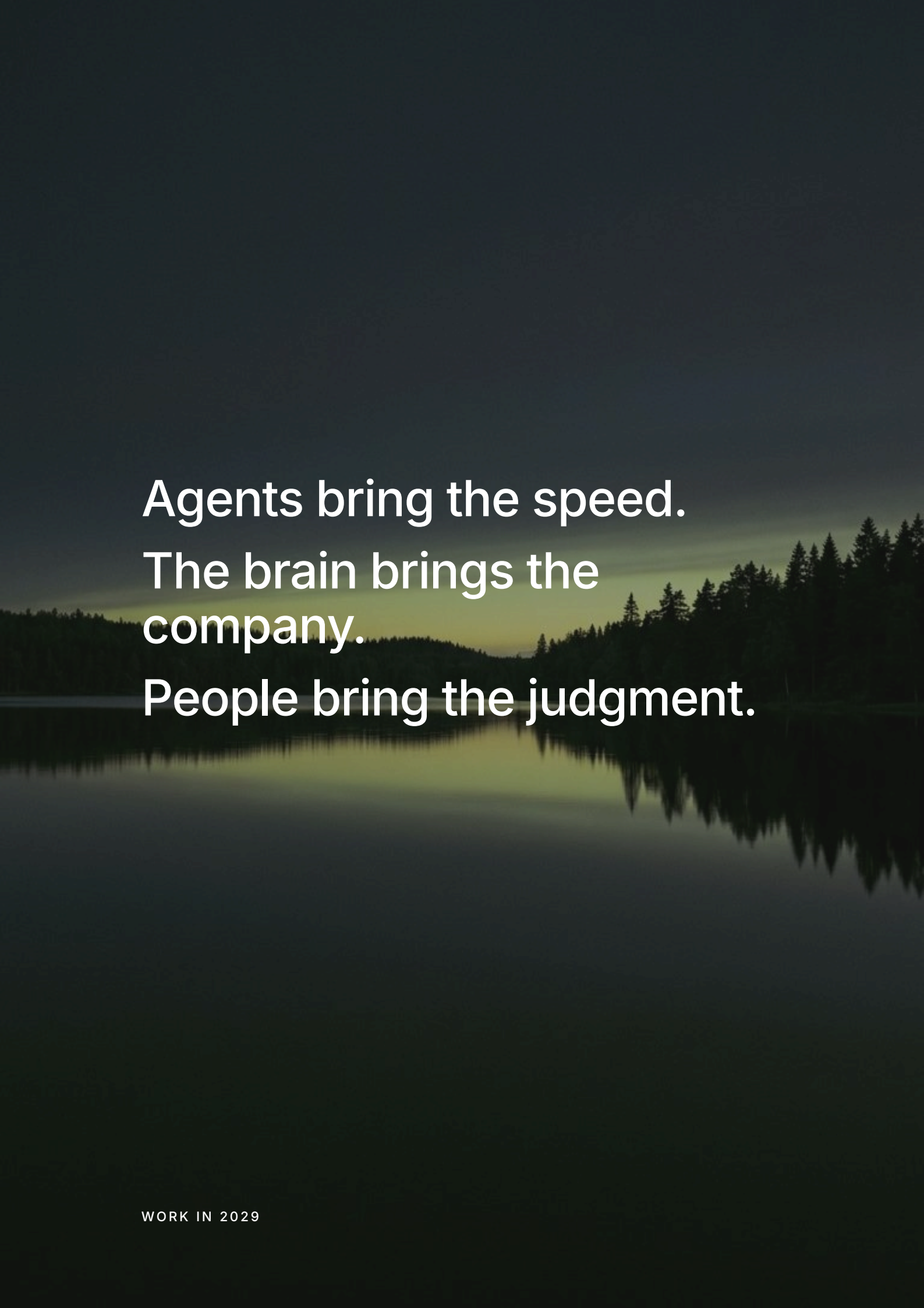
Standards

- **ISO/IEC 27001:2022. Information security, cybersecurity and privacy protection — Information security management systems — Requirements.** International Organization for Standardization and International Electrotechnical Commission, 2022. https://www.iso.org/standard/27001
  - The international standard for managing information security. Its Annex A controls cover access, supplier and cloud security, logging, incident handling and secure development.
  - Background to Access and rights (access, administrator access, incidents), Agents (logging) and Systems and tools (suppliers and building).

About the author

Kristian Arvidsson is Chief AI Officer at a Nasdaq-listed SaaS company, where he leads AI strategy and innovation. He has spent nearly two decades building commerce technology: as founder and CEO of Geins, an API-first commerce backend, as CEO of Carismar, and before that as engineer, product lead and COO. He started coding at 13 and still builds what he writes about, from SDKs and MCP servers to AI agents in daily use. He is based in Stockholm.

linkedin.com/in/kristian-arvidsson-7718b146


Agents bring the speed.
The brain brings the company.
People bring the judgment.