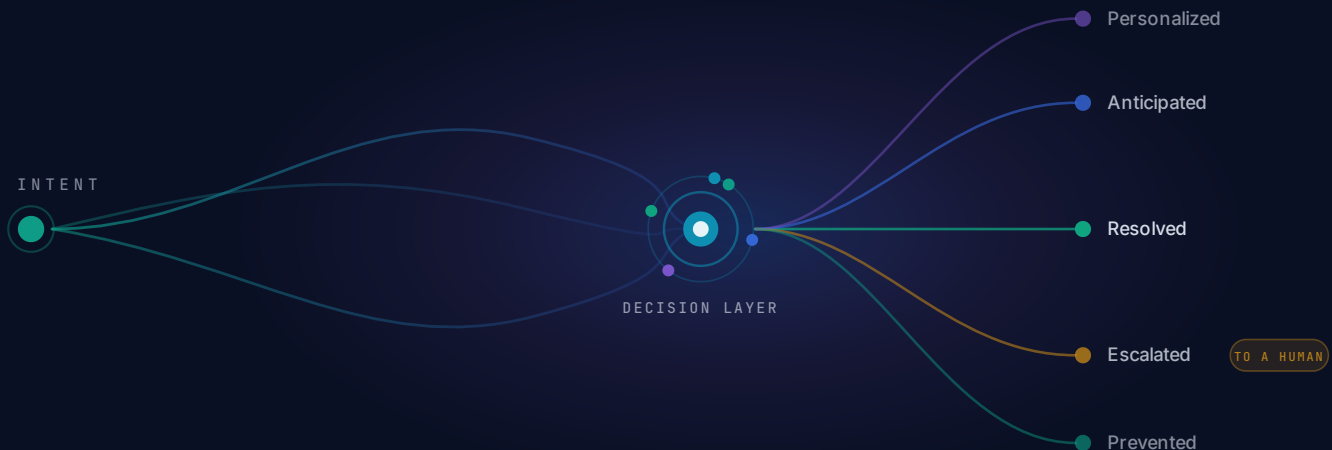


CUSTOMER EXPERIENCE DESIGN & INNOVATION

Agentic AI Customer Experience & Innovation

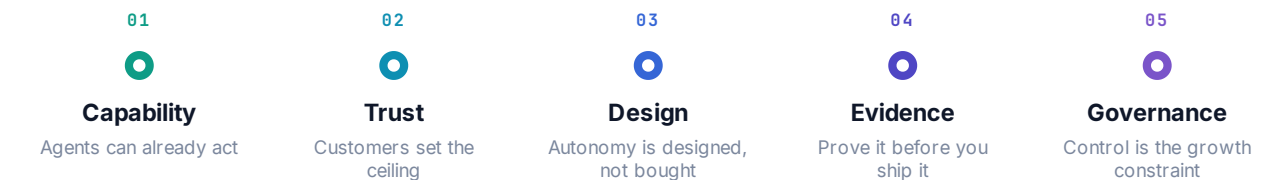
Your customers are about to meet a system that decides. This is how to design it, prove it, and stay accountable for what it does.



Five things are true at once

Agentic AI is not arriving in customer experience. Two-thirds of service organizations already run agents, and in production they resolve most routine conversations without a person. What is missing is not capability. It is the design, the evidence and the control to let a system act on a customer's behalf without spending the trust that took years to build.

FIGURE 1 • THE SPINE OF THIS PAPER



THE EXPERIENCE IS THE PRODUCT — AND AGENTS NOW DELIVER IT

Read it as a sequence. Capability without trust stalls. Trust without design is an accident. Design without evidence is an opinion. And none of it scales past a pilot without governance.

What this paper is for

It is written for the leader who owns the customer relationship and is now being asked to put an autonomous system inside it — the chief customer, experience, or marketing officer, and the design and service leaders who report to them.

It does three things: it separates what has actually been proven from what is being sold, it gives you the design models that make autonomy safe, and it names the measures that tell you whether it is working before your customers tell you it is not.

What it is not

It is not a technology survey. Which model, which orchestration framework, which vendor — those choices matter and they are not the constraint. The constraint is experience design and accountability.

Every figure is original. Every external number is attributed on the page and listed at the back. Where a claim is contested, the paper says so.

66%

of customer service organizations now use AI agents — up from 39% a year earlier

SALESFORCE, 2026

5%

of task-specific enterprise AI tools reach production

MIT NANDA, 2025

54%

of people would trust an agent with simple issues only

INTERCOM, 2026

41%

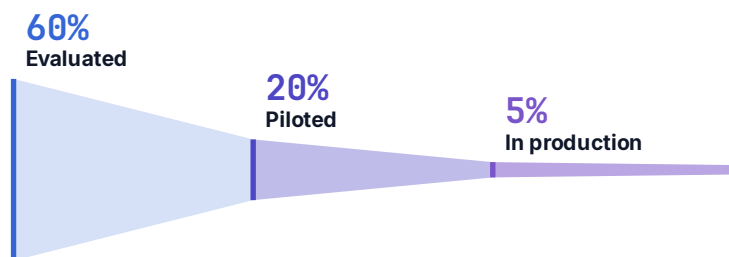
of customer-facing AI deployments fully scale — 3.5× the rate elsewhere

MCKINSEY, 2026

Two gaps, not one

The industry talks about an adoption gap. There are two, and they pull in opposite directions. One is internal: organizations cannot get agents out of the lab. The other is external: customers will not grant agents the room the business plan assumes.

FIGURE 2 • THE FIRST GAP — INTEREST DOES NOT SURVIVE CONTACT WITH PRODUCTION



-55

points between interest and impact — a 92% relative decline

McKinsey puts the same figure a second way: fewer than 10% of deployed use cases get past pilot.

TASK-SPECIFIC ENTERPRISE AI TOOLS • MIT NANDA, 2025

The collapse is not technical. Tools that clear evaluation die in pilot for unglamorous reasons: no owner, no data the agent can trust, no policy on what it may do alone, and no way to show anyone what it did. Gartner expects more than 40% of agentic AI projects to be canceled by the end of 2027.

The pilots that survive share one trait

They redesigned the process instead of decorating it. McKinsey's numbers are stark on this point: agents dropped into existing workflows return five to ten percent productivity gains. Where the process itself was reinvented around what an agent can do, up to 80% of common incidents resolve autonomously and resolution time falls 60 to 90 percent.

That is not a technology finding. It is a design finding, and it is why this belongs to experience leadership rather than to IT.

WORTH NOTICING

Customer-facing work is the exception to the stall. It is where agents most often reach full scale — because the feedback is immediate, the volume justifies the build, and the outcome is legible to everyone.

FIGURE 3 • WHERE AGENTS ACTUALLY SCALE



Share of AI deployments fully scaled, by function. McKinsey, 2026.

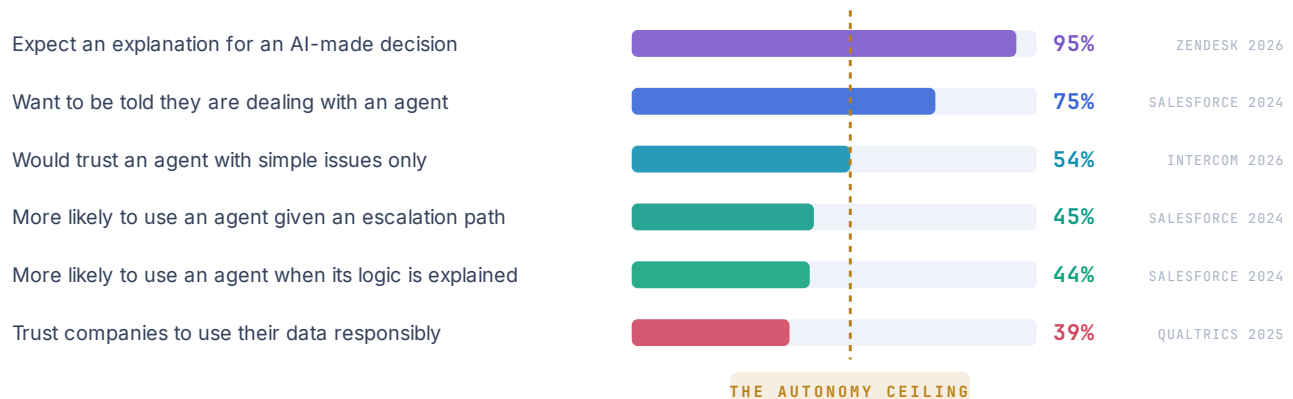
FIGURE 4 • WHAT SCALED PROGRAMS REPORT



Named enterprise programs. McKinsey and Salesforce, 2026.

The second gap is the one that decides your ceiling. It does not appear in a roadmap because it does not belong to you. It belongs to the customer, and it is measured.

FIGURE 5 • THE SECOND GAP — WHAT CUSTOMERS WILL AND WILL NOT HAND OVER



Read the bars as specifications, not preferences. Almost everyone expects an explanation. Three in four want to know they are dealing with an agent. Barely half will grant autonomy beyond simple matters. The ceiling moves only when disclosure, explanation and escalation are built in — which makes trust an engineering budget line.

The failure rate customers actually experience

Nearly one in five people who have used AI for customer service report no benefit at all — roughly four times the failure rate they report for AI in general. Concern is rising, not falling: over half now fear their data will be misused, and half worry AI will remove human contact altogether.

Forrester expects a third of companies to actively damage experiences with frustrating AI self-service. Gartner goes further and prices it: regulatory change and customer opt-outs will push assisted service interactions up 30% by 2028. The customers who refuse the agent do not disappear. They arrive somewhere more expensive.

Two things raise a customer's comfort with an agent more than anything else: an easy path to a human, and being told upfront that it is not one.

INTERCOM AI SENTIMENT REPORT, 2026

Which makes the strategy simple to state

You do not get to choose between autonomy and trust. Trust is the input that buys autonomy, and it is bought with three specific, buildable things: disclosure, explanation, and a way out.

The rest of this paper is about how to build them, and how to know they are working.

THE RISK IF YOU IGNORE IT

Containment theater

The agent hits its resolution target by making escape difficult. The number looks excellent. Retention quietly falls.

THE RISK IF YOU OVERCORRECT

Consent fatigue

Approval prompts on everything. People click through without reading and the oversight you paid for stops working — the most consistently exploited weakness Microsoft's red team found.

THE POSITION TO HOLD

Earned autonomy

Autonomy granted per moment, by consequence, with disclosure and redress shipped alongside it. Raised on evidence, never at launch.

From helping people act to helping them judge

For thirty years, experience design has been about crossing a gap between what someone wants and what a system will let them do. Agents cross that gap themselves. What is left for the customer is a harder question — was it done well? — and almost nothing in our craft was built to answer it.

FIGURE 6 • THE SHIFT IN WHAT DESIGN IS FOR



AFTER JOHN MAEDA, DESIGN IN TECH REPORT 2026 • GULFS AFTER DONALD NORMAN

John Maeda frames the change as a move between Donald Norman's two gulfs. Interface design closed the gulf of execution. Agentic design has to close the gulf of evaluation — and the artifacts that do it are not screens.

Three consequences you will feel immediately

- **The interface stops being permanent.** A dashboard the customer must monitor is a failure of delegation. What replaces it is situational: surfaces that appear only when a decision needs a person, then get out of the way.
- **The deliverable stops being a flow.** You are specifying behavior under uncertainty — what the agent may assume, when it must ask, what it does when a tool fails. That is a policy document with a picture of it.
- **Quality stops being measurable in the moment.** A single interaction can look flawless and still be wrong. The unit of quality becomes the loop: act, observe the outcome, take the correction, adjust.

FIGURE 7 • THE UNIT OF DESIGN IS NOW A LOOP



Six stages the agent runs continuously. The human checkpoint sits on **verify** because that is where consequence lands — never on every step.

FIGURE 8 • THE SAME JOURNEY, TWO ARCHITECTURES

AUTOMATION • ONE PATH, DECIDED IN ADVANCE



Anything unplanned becomes a dead end — a transfer, a queue, a repeated question.

AUTONOMY • THE PATH DECIDED EACH TURN



The same intent can be served many ways. The system picks one, then explains it.

Automation executes a path chosen in advance, so anything unplanned becomes a dead end. Autonomy chooses the path each turn — which is why it needs a decision layer, and why that layer needs governing.

The demo question

Ask the vendor to change the customer's mind halfway through the scenario — a different date, a new constraint, a complaint mid-transaction. An automation will restart or transfer. An agent will re-plan and tell you what it changed.

Ask the same question of your own build before it goes anywhere near a customer.

The distinction is not speed. It is permission.

Every automation wave since the mainframe reduced human effort. Only this one moves the decision. That single difference is what turns a technology program into an experience and accountability program.

It also explains why so much of what is marketed as agentic is not. Gartner has a word for it — *agent washing* — and counts roughly 130 genuine agentic vendors among the thousands making the claim. The test is simple and you can apply it in a demo: **can it decide something that was not decided for it, and can it act on that decision?** If not, it is a chatbot with better prose.

THE QUESTION THAT SEPARATES THE TWO

Stop asking “how can we automate this process?”
Start asking “what outcome should this system be continuously optimizing, and what is it allowed to do to get there?”

Answering the second half of that question is design work. It is also the part nobody can outsource to a platform, because it encodes what your organization is willing to be responsible for.

ONE MORE SHIFT

Your customer may not be a person

Consumer agents are beginning to shop, compare and transact on their owners' behalf. Interfaces that fail an agent tend to fail the human behind it too — which finally gives structured, accessible, machine-legible experience a hard commercial case.

WHAT THIS DOES TO SERVICE DESIGN

Differentiation moves

When basic execution is commoditized, advantage shifts to judgment, recovery and the quality of the explanation. The blueprint gains a new line — not just what is visible to the customer, but what is legible.

AND TO THE TEAM

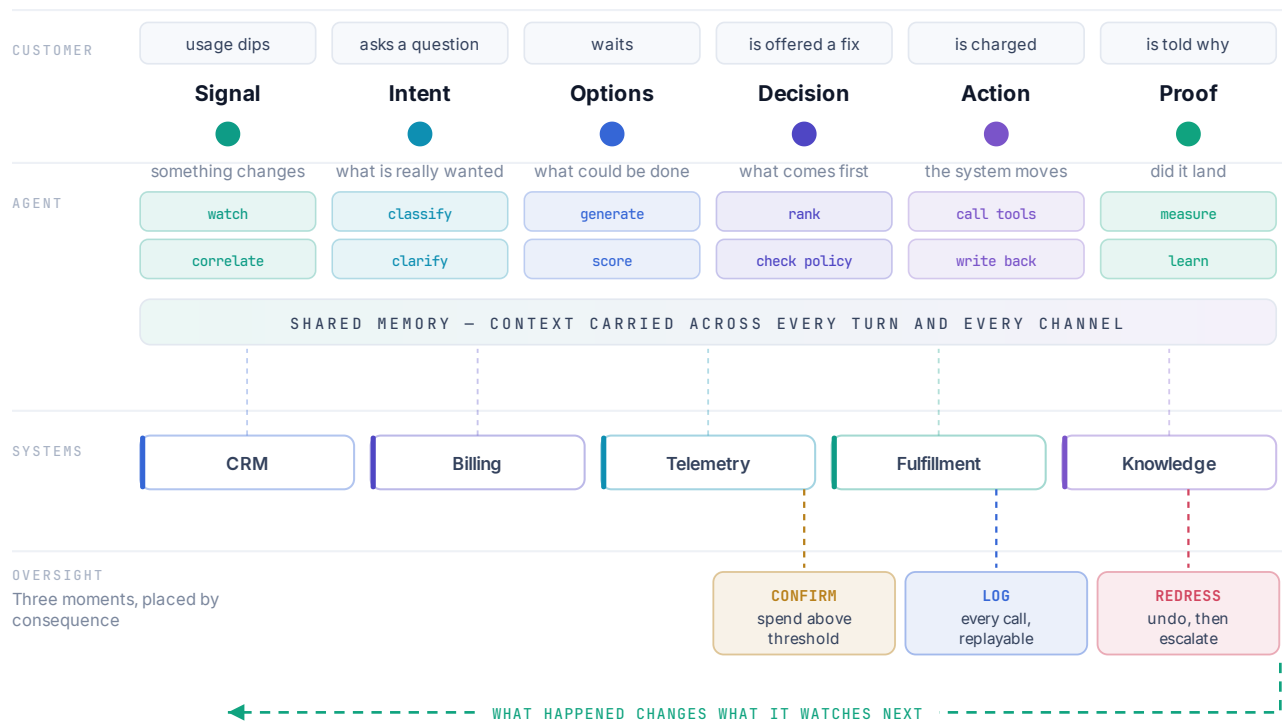
New artifacts, not new tools

Intent maps. Authority tiers. Confidence and escalation rules. Golden conversations that define good behavior. Exclusion lists. These are the working documents of agentic experience design.

Where the agent actually sits

A journey map drawn as customer steps and touchpoints cannot hold an agentic system, because the interesting behavior happens between the steps. This is the same journey drawn as four layers — what the customer experiences, what the agent does, what it reaches into, and where a person stays in the loop.

FIGURE 9 • ANATOMY OF AN AGENTIC CUSTOMER JOURNEY



Read down, not across. The customer layer is the only one they see. The agent layer is where the work moved. The systems layer is why most programs stall — an agent that cannot write back can only advise. The oversight layer is what makes the whole thing defensible, and it has exactly three moments in it, placed by consequence.

Four things in that diagram are worth stating plainly, because they are where programs are won and lost.

- 01 **Signal replaces request as the trigger.** The most valuable agentic work starts before the customer contacts you. That reframes service from a queue to be cleared into a set of conditions to be watched.
- 02 **Memory is the product, not a feature.** Three-quarters of people are frustrated at repeating themselves and four in five expect the conversation to continue rather than restart. Shared context is what makes an agent feel like your company instead of a widget.
- 03 **Write access is the dividing line.** An agent that can read your systems answers questions. An agent that can write to them resolves them. Data readiness is the blocker most cited by service operations teams, and it is a prerequisite, not a phase two.
- 04 **Proof closes the loop.** Confirming the outcome — the booking reference, the credit, the calendar entry — is the moment the customer can evaluate, which is the only moment trust is actually transacted.

Autonomy is a design decision, made per moment

The common mistake is treating autonomy as a property of the technology — how capable is the agent. It is a property of the deployment: how much room you granted it, in which moment, with what oversight. That is yours to decide, and it should be decided once for the whole estate rather than use case by use case.

FIGURE 10 • FIVE LEVELS, NAMED BY THE ROLE THE PERSON KEEPS

HUMAN INVOLVEMENT FALLS →

← WHAT MUST BE DESIGNED RISES



Naming levels by the human's role rather than the machine's cleverness keeps the conversation honest. After the Knight First Amendment Institute at Columbia University, 2025. The step from Consultant to Approver is where most of the design cost sits — it is the first level at which the system acts without being watched.

Do not set a level for the company

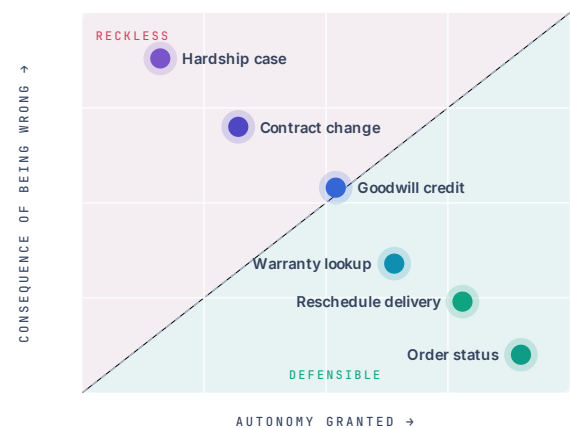
"We are moving to level four" is a sentence that has never survived contact with a real customer base. Level belongs to the moment, and the variable that sets it is not customer sophistication or channel — it is the cost of being wrong.

Anthropic's framing is the more useful instrument here: score autonomy and consequence separately, then look at the pair. Autonomy is emergent from the deployment, not fixed by the model — which means it can be dialed, and must be reviewed.

THE FAILURE MODE THIS PREVENTS

A single global threshold — "confirm anything over \$100" — invites slicing. Six sequential \$95 approvals pass a per-transaction rule and drain the same budget. Limits must hold across a session.

FIGURE 11 • PLOT THE MOMENT, NOT THE COMPANY



Autonomy against consequence, with illustrative service moments placed. Anything above the diagonal is an unfunded risk.

FIGURE 12 • THREE TIERS, APPLIED EVERYWHERE

TIER 1

Act

Low consequence, reversible, policy is clear.

- Status and tracking
- Address change
- Send confirmation

TIER 2

Confirm first

Money moves, or a commitment is made.

- Refund or credit
- Book and pay
- Change a plan

TIER 3

Escalate

Judgment, discretion or vulnerability.

- Waive a policy
- Complaint with harm
- Distressed customer

One authority model across the estate. Tiering by consequence — not by team, channel or product — is what keeps an agent's behavior predictable to customers and auditable to you.

Why three tiers beat a policy per team

Consistency is the point. When authority is decided locally, a customer gets a refund instantly in one channel and a three-day wait in another, and both experiences are technically compliant. Tiers make the rule legible to the customer and enforceable by you.

Two rules make tiers hold under pressure.

- **The tier is enforced in code, not in the prompt.** A model asked to respect a budget can be argued out of it. A backend that refuses the call cannot. Never let the system that decides also be the system that polices the decision.
- **Tier three is a real destination, not a fallback.** If escalation lands in the same queue as everything else, the tier is decorative. The people who receive tier-three work need the context, the authority and the time that tier implies.

Done properly, tiers are also how you raise autonomy safely: you promote specific moments from tier two to tier one on evidence, one at a time, and you can always name what changed.

The distinction that matters is not how clever the agent is. It is whether the organization has decided, in advance and in writing, what it is permitted to do alone.

22%

of executives trust agents for autonomous customer interactions

PWC, 2025

91%

of service leaders are under pressure to implement AI this year

GARTNER, 2026

32%

escalation rate held steady across 170x more agent conversations

SALESFORCE, 2026

80%

of common service issues Gartner expects agents to resolve alone by 2029

GARTNER, 2025

HOW A TIER GETS PROMOTED

Evidence, then a widening

A moment moves from confirm-first to act-alone when the confirmations were almost always granted, the reversions were rare, and nothing in the record surprised you. Say which of those three moved.

HOW A TIER GETS DEMOTED

The same evidence, in reverse

Rising reversions, rising complaints about a specific action type, or a single high-severity incident. Demotion should be as routine and as unremarkable as promotion.

WHO OWNS THE MODEL

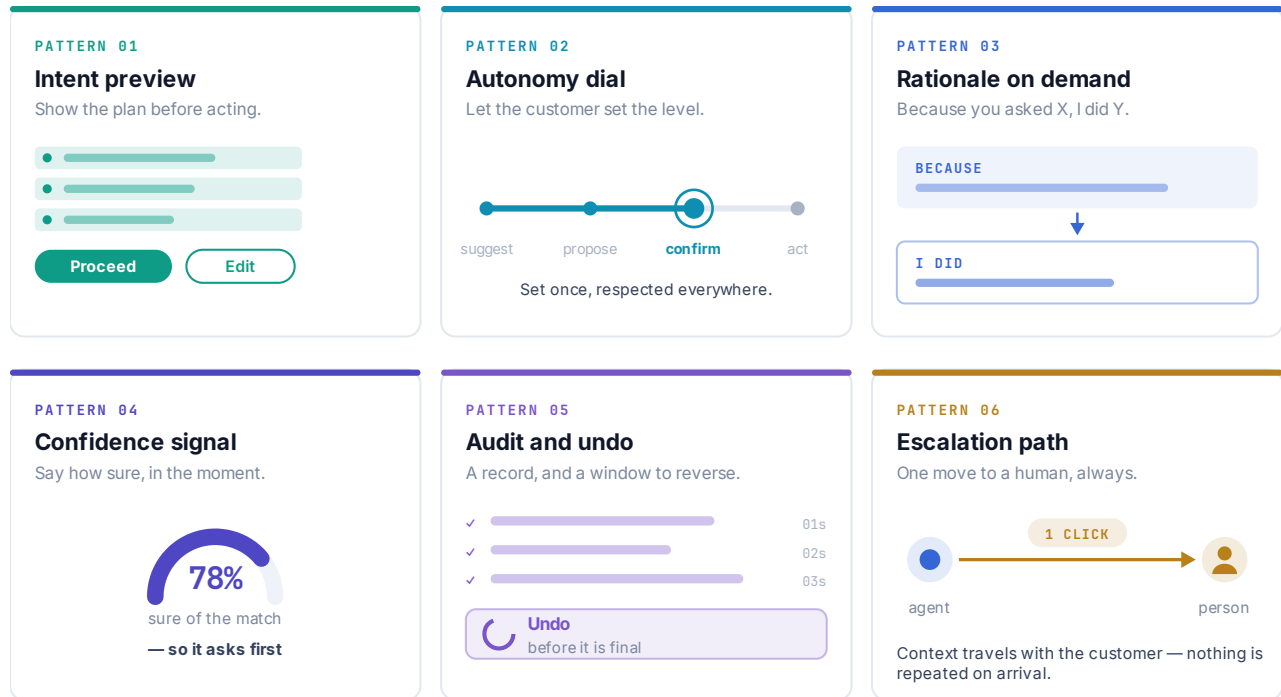
One owner, not a committee

Authority tiers are a customer-experience decision with legal and risk input, not the reverse. If nobody can name the owner, the tiers will drift back to being decided per team.

Six patterns that buy you autonomy

Trust is not a tone of voice. It is a set of shipped mechanisms that let a customer see what is being done on their behalf, judge whether it was right, and change it if it was not. These six are the working set. Each has a metric, and each is testable before launch.

FIGURE 13 • THE PATTERN SET



Six mechanisms, drawn as what they do. Synthesized from published practice — Google's People + AI Guidebook, Microsoft's human-AI interaction guidelines, IBM's Carbon AI disclosure components, and the agentic UX pattern work of Victor Yocco and Notch. None of them is exotic. All of them are usually missing.

How to use them

Attach patterns to tiers rather than sprinkling them everywhere. Tier one needs disclosure and an undo window. Tier two adds intent preview and confidence. Tier three is an escalation design problem, not an interface one.

Two of the six are commonly done badly.

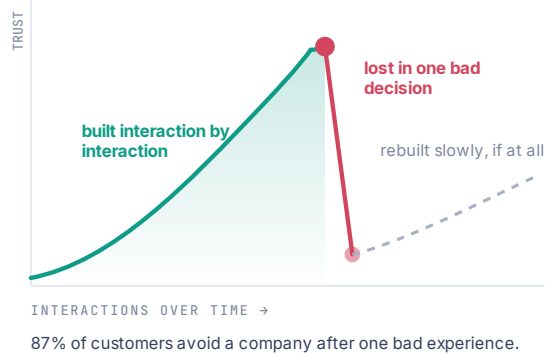
Rationale gets built as a technical trace rather than a sentence a customer can read. And **confidence** gets shown as a number with no consequence — a confidence signal only earns its place if a low score visibly changes what the agent does next.

What they are worth

- A clear escalation path makes people **45% more likely** to use an agent at all.
- Explaining the agent's logic adds a further **44%**.
- **95%** expect an explanation for an AI-made decision, and four in five CX leaders now expect transparency to be a requirement rather than a differentiator.

These are the cheapest points of leverage in the entire program. They cost interface work, not model work.

FIGURE 14 • TRUST COMPOUNDS SLOWLY, COLLAPSES AT ONCE



The asymmetry is the whole argument for over-investing in the failure path. Accenture puts the consumer behavior behind this curve bluntly: 87% say they will avoid a company after a single bad experience.

FIGURE 15 • DISCLOSURE, AS IT NOW HAS TO LOOK



Article 50 of the EU AI Act requires notice at first interaction, in a clear and distinguishable manner. Verify current dates against the Regulation before relying on them.

Treat trust as a balance, not a mood

The Conversation Design Institute calls it *conversational capital*: trust accumulates through consistent good interactions, buys you the right to higher autonomy, and can be destroyed faster than it was built. It is a useful frame because it makes trust a tracked balance with a trajectory, rather than a sentiment measured after the fact.

The practical version: before widening an agent's authority, ask what the trust line has done since the last widening. If it is flat, you have not earned the next level.

Disclosure is now a legal surface

Telling customers they are dealing with an agent stopped being a design preference. In the EU it is an obligation with a date. California has required bot disclosure in commercial contexts since 2019, Utah requires it on request, and Colorado has its own regime.

Design for the strictest of these and the rest follows. Practically that means: state it at the first interaction, in the conversation itself, in plain language — and pair it with what the agent can and cannot do, which is the part customers actually want.

TEST IT LIKE THIS

The screenshot test

Take one screenshot of the first thing a customer sees. If disclosure, scope and the escape route are not all visible in it, none of them exists as far as the customer is concerned.

AND LIKE THIS

The complaint test

Pick a past complaint. Could you now reconstruct what the agent did, why, and on whose authority — from the record alone? If not, the trust surface has no floor under it.

WHAT NOT TO DO

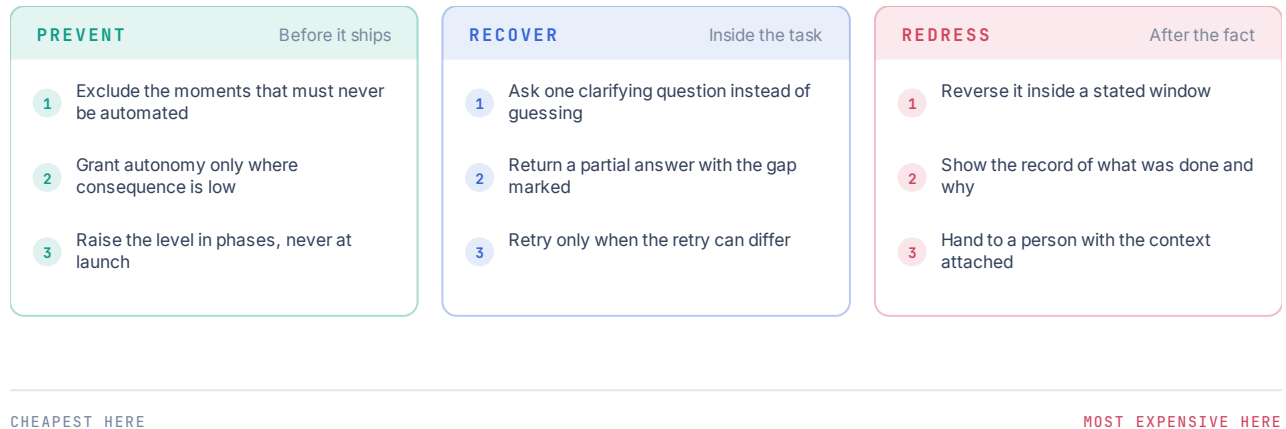
The persona hedge

Giving the agent a human first name and a photograph, then disclosing its nature in the footer. It reads as concealment when discovered, and in the EU it does not satisfy the obligation.

Design the failure path first

Customers form their view of an agent at the moment it fails. That moment is the most under-designed surface in the category — and the one place where an hour of design work is worth a quarter of model work.

FIGURE 16 • THREE LAYERS, USUALLY COLLAPSED INTO ONE



Prevention, in-task recovery and after-the-fact redress are different disciplines with different owners. Most programs build only the middle column, then discover the first and third the hard way.

Prevention is the cheapest column, and the one that gets skipped

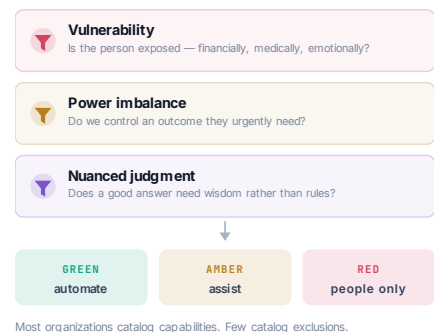
Nearly every organization maintains a list of what agents can do. Almost none maintains a list of what they must never do. That second list is the higher-value document, and it takes a workshop, not a platform.

The filter at right is deliberately blunt. It asks three questions about the moment rather than about the technology, and it produces three zones you can staff and audit. A customer in financial distress, a bereavement, a safety complaint, a dispute where you control the outcome they urgently need — these do not go to an agent because it would be technically difficult. They do not go to an agent because being right on average is not good enough there.

THE COUNTERINTUITIVE FINDING

More approval prompts make oversight worse. A year of Microsoft's agentic red teaming found human-in-the-loop bypass the most exploited failure mode, largely through consent fatigue. Approvals are a scarce resource — spend them where consequence is real.

FIGURE 17 • THE EXCLUSIONS FILTER



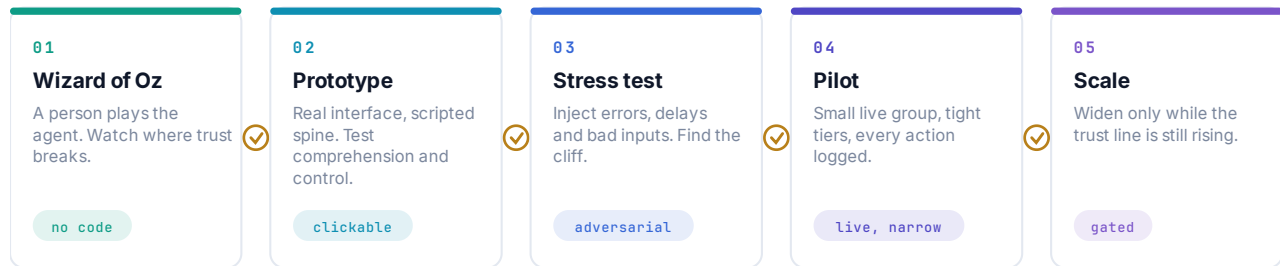
Three questions, three zones. After the Conversation Design Institute's exclusions framework.

Prove it in a room before you prove it on customers

The reason so much agentic work dies in pilot is that the pilot was the first honest test. There are four cheaper tests before it, and running them is what separates an innovation practice from a procurement exercise.

FIGURE 18 • THE EVIDENCE LADDER

EACH STEP IS GATED — EVIDENCE OPENS THE NEXT ONE



Each step is gated: evidence opens the next one. The first three need a room, a script and a willingness to be wrong early — no platform, no integration, no committed roadmap.

The two steps almost nobody runs

Wizard of Oz. A person plays the agent behind the interface. It is a forty-year-old research method and it remains the fastest way to find out how customers feel about delegation before you have built anything. Watch for the moment someone stops trusting the system, and ask what they thought had just happened.

Stress testing. Introduce errors, delays, bad data and deliberate manipulation. The goal is not a pass mark; it is to locate the cliff — the point at which a customer stops believing the system and starts trying to escape it. Prototypes tested only with cooperative users tell you nothing about production, where users are adversarial, distracted or distressed.

Three questions to ask in every session

- **Control:** “At what point did you feel you had lost track of what was happening?”
- **Transparency:** “Do you understand why it chose that?”
- **Comfort:** “Would you let this run overnight without you?”

The third question is the useful one. It converts a vague attitude into a delegation decision, and the reasons people give for saying no are your design backlog.

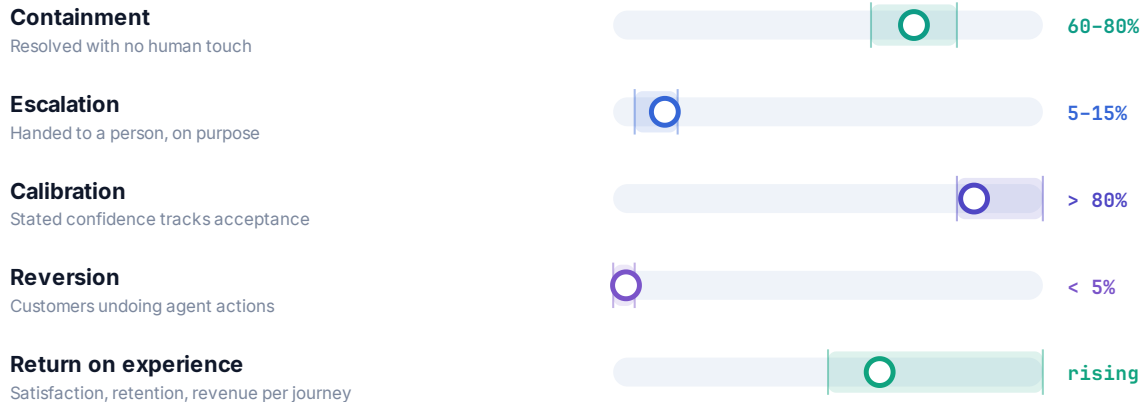
WHY THIS IS AN INNOVATION FUNCTION

Maturity, not adoption, is what correlates with results: 87% of teams at mature deployment report improved metrics, against 62% overall. Maturity comes from redesigning the work and testing it — which is a practice you build, not a product you buy.

Five numbers, read together

Most agentic programs are reported on one number: how much got handled without a person. On its own it is the most gameable metric in customer experience. Read against four others, it becomes an instrument panel.

FIGURE 19 • THE INSTRUMENT PANEL



SHADED BAND = HEALTHY RANGE

CIRCLE = AN ILLUSTRATIVE READING

The bands matter more than the values. Containment outside its band is usually a sign of something else: too low and the agent is not trusted with anything, too high and customers cannot get out. Read on its own it rewards the wrong behavior — a team will hit the number by trapping people in a loop they cannot leave.

What each one is actually telling you

- **Containment** measures reach. Salesforce's telemetry across a year of production agents puts roughly seven in ten conversations resolved without escalation — a fair benchmark for tier-one work, not a target for everything.
- **Escalation** measures honesty. A healthy agent hands over deliberately and often enough to notice.
- **Calibration** measures self-knowledge: does stated confidence actually predict whether customers accept the answer? Poor calibration is what produces confident wrong answers.
- **Reversion** measures consent. People undoing agent actions is the clearest signal that autonomy was granted where it was not wanted.
- **Return on experience** measures the point of the exercise — satisfaction, retention and revenue per journey, argued alongside ROI rather than instead of it.

Two things to add to the board pack

Cost per resolution, with a horizon. Gartner expects generative AI cost per resolution to exceed three dollars by 2030, surpassing many offshore human agent costs. A business case resting on indefinite unit cost decline is resting on nothing. Full automation is unlikely to be the cheapest answer; the right mix is.

Time to first measurable value. Seven in ten service organizations adopting agents observe value within sixty days of deployment. If your program cannot see anything in that window, the problem is almost certainly scope rather than technology.

ONE MEASURE TO STOP USING

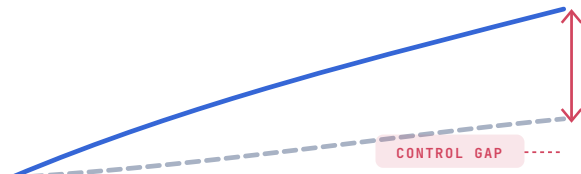
Deflection. It names the goal as getting rid of the customer. Every organization that manages to it eventually gets what it measured.

Control is the throttle, not the brake

The instinct is to treat governance as the thing that slows agentic deployment down. The data says the opposite: the organizations with the strongest control layer run the most agents, with fewer incidents and better margins. Control is what makes speed survivable.

FIGURE 20 • ACCOUNTABILITY IS OUTRUNNING CONTROL

Agents deployed, decisions delegated



Ability to see, explain and stop them

- 77%** say adoption already outpaces governance
- 2 in 3** are accountable for systems they do not control
- 11%** call themselves ready for next year's scale

IBM INSTITUTE FOR BUSINESS VALUE • 2,000 CIOs AND CTOs

Deployment and delegation are rising faster than the ability to see, explain and stop what agents do. IBM Institute for Business Value, 2026.

FIGURE 21 • THE FOUR LAYERS THAT EARN AUTONOMY

Policy in code

Limits enforced outside the model

Provenance

Every action, input and reason logged

Interception

Hard stops the prompt cannot argue with

Standing review

A register, an owner, a kill switch

Teams that build this deploy 16× more agents with 25% fewer incidents. Control is the throttle, not the brake.

Each layer removes a reason to say no. Together they are what lets you raise a tier without arguing about it.

What to stand up, in plain terms

- 01 An agent register.** Every agent in production, its tier, its owner, its tools, its last review. If you cannot produce this in an afternoon, you do not know what is running.
- 02 Interception in the infrastructure.** Thresholds, spend limits and forbidden actions enforced by code the model cannot reach or reason with. Track them across a session, not per call.
- 03 A replayable record.** Intent, reasoning, tool calls, arguments, guardrail evaluations and the interface state the customer actually saw. This is what makes a complaint answerable.
- 04 A drift review.** Watch for an agent stating an intent that violates policy even when the backend blocked it — that is prompt instability. Watch step counts and loops too.
- 05 One stop button, and a named person who may press it.**

The internal argument is easier than it looks. Organizations that embed control report 25% fewer incidents, deploy 16× more agents and run 18% higher operating margins than those governing manually — useful in any meeting where governance is being described as friction.

Two in three technology leaders report being held accountable for AI systems they do not fully control.

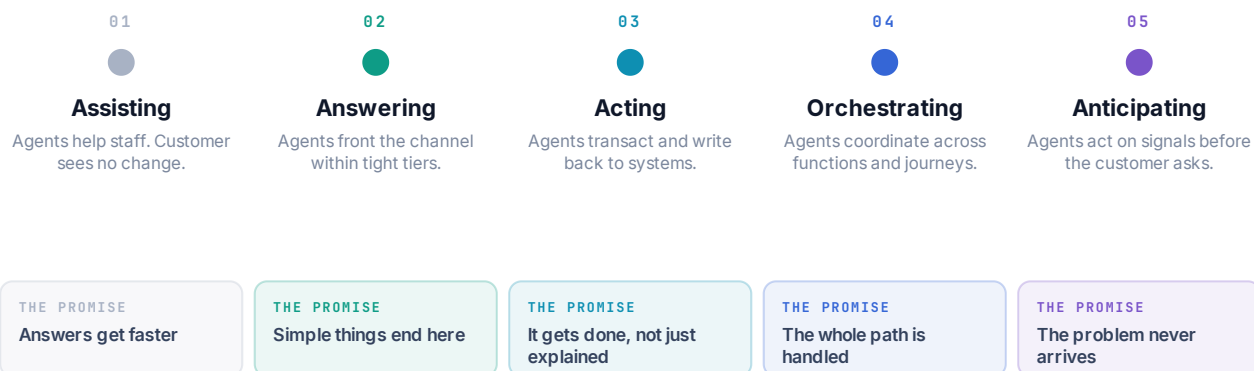
IBM INSTITUTE FOR BUSINESS VALUE, 2026

The experience consequence is specific: without a replayable record, you cannot tell a customer why something happened. Every trust pattern in section five depends on this layer.

Five stages, five different promises

Maturity models usually describe what the technology can do. This one describes what you are entitled to promise a customer at each stage — because that is the claim your brand makes, and the one you have to be able to keep.

FIGURE 22 • THE AGENTIC EXPERIENCE MATURITY MODEL



Stages are not a queue to be rushed — each one is a different promise, and each promise raises the standard of proof. Moving from *answering* to *acting* requires write access and an authority model. Moving to *anticipating* requires consent, because acting before you are asked is either delightful or intrusive with very little in between.

How to place yourself honestly

Take one journey, not the organization. Most enterprises are at stage two in service, stage one in sales, and nowhere in billing — and the average is meaningless. The useful question is which single journey is furthest along, and what it needed that the others do not yet have.

Then look at the gap between where you operate and what your marketing already implies. That gap is where trust gets spent.

The two stalls to expect

Between answering and acting. This is the write-access wall, and it is a data and integration problem wearing an experience costume. Data readiness is the most-cited blocker among service operations teams for exactly this reason.

Between acting and orchestrating. Here the obstacle is organizational: agents cannot coordinate across functions that will not share a definition of the customer, a queue or a P&L. No platform resolves that.

STAGE TWO

Can it end the conversation?

If every path still ends in a handoff, the agent is assisting, not answering — whatever containment says.

STAGE THREE

Can it change a record?

Write access, an authority tier, a replayable log. Without all three, acting is a demo, not a capability.

STAGE FOUR

Does it cross a function?

One journey, one agent network, one definition of the customer. An organizational test, not a technical one.

STAGE FIVE

Did the customer consent?

Acting before you are asked needs permission on the record. Without it, anticipation reads as surveillance.

Six moves, in this order

None of these requires a platform decision, and none of them is a technology project. They are the sequence that turns agentic ambition into something you can defend to a customer, a regulator and a board.

FIGURE 23 • THE SEQUENCE

01 Map the moments

Pick the journeys where a wrong answer is cheap and a right one compounds.

02 Write the exclusions

Name what will never be automated, and publish it internally.

03 Set the tiers

One authority model for the whole estate, not one per team.

04 Build the trust surface

Disclosure, rationale, undo and escalation as shipped features.

05 Instrument before scaling

Stand up the five measures first. Widen only while they hold.

06 Govern from the center

One register, one owner, one stop button.

Order matters more than pace. Teams that start at move four build a trust surface for an authority model they have not agreed. Teams that skip move two eventually automate a moment they should not have touched.

What good looks like at the end of the sequence

- A named list of journeys where agents act, and a named list where they never will.
- One authority model, enforced outside the model, understood by every team.
- Disclosure, rationale, undo and escalation shipped as features with owners.
- Five measures on one page, reviewed on a standing cadence, with bands rather than targets.
- A register that tells you what is running, under whose authority, and when it was last reviewed.

And the honest caveat

This field is young enough that the pattern vocabulary is still moving. One of the more credible practitioner groups published that there were no settled patterns for agentic experience, then published a sixteen-pattern library four months later. Treat any framework in this space — including this one — as a working instrument to be revised against your own evidence.

What is not moving is the underlying requirement. Customers will delegate to a system that discloses what it is, explains what it did, and lets them undo it. Everything else is implementation.

The enterprises that win this will not automate faster. They will decide better, act more accountably, and be able to show their work.

WHERE THIS GOES NEXT

Design it. Prove it. Stay accountable for it.

f/cx.ai is a customer experience design and innovation agency for the AI era. We design, prototype and prove the journeys, service models and AI agents enterprise organizations run on.

IF YOU ARE AT STAGE ONE OR TWO

Start with the exclusions and the tiers

The highest-value work is deciding what agents may do, before deciding what they will do. It is a room, a facilitator and a set of published decisions.

IF YOU ARE STALLED IN PILOT

Run the evidence ladder in reverse

Most stalled pilots skipped a step. Find which one, run it properly, and the argument for scaling either makes itself or ends cheaply.

IF YOU ARE ALREADY ACTING

Instrument, then widen

Stand up the five measures and the register. Then raise one tier on one journey and watch what the trust line does.

Customers will delegate to a system that discloses what it is, explains what it did, and lets them undo it.
Everything else is implementation.

THE CONVERSATION WE HAVE MOST OFTEN

Where agentic systems belong in your customer path — which moments, at which authority, and what has to be true before you widen. If that is the question in front of your leadership team, it is the one we are set up for.

CONTACT

Skip Vanderburg

Founder & Chief Experience Officer

innovate@fractionalcx.ai

fractionalcx.ai

Where every external number came from

Every figure in this paper is original. External data points are attributed on the page where they appear and listed here in full. Three notes on care: the MIT NANDA figures circulated in 2025 in a paraphrased form that overstates them; one Salesforce trust dataset is a 2024 baseline, flagged as such; and vendor-published performance data is self-interested, so it is paired with independent sources wherever a claim rests on it. Regulatory dates should be verified against the primary legal text before being relied on, as application timing has been amended.

ADOPTION, SCALING AND THE PILOT GAP

McKinsey 2025	<i>Seizing the agentic AI advantage.</i> Fewer than 10% of deployed use cases pass pilot; up to 80% of common incidents resolvable autonomously under process reinvention; 60–90% resolution-time reduction.
McKinsey 2026	<i>The State of AI: global survey.</i> 40% of large organizations report scaling AI agents, up from 27%.
McKinsey 2026	<i>Rewiring customer experience for the agentic era.</i> 41% of customer-facing AI deployments fully scaled, 3.5× the rate of other domains; named program outcomes (+24% FCR, +6% CSAT, €40M margin).
MIT NANDA 2025	<i>The GenAI Divide: State of AI in Business.</i> Task-specific enterprise tools: 60% evaluated, 20% piloted, 5% in production. Widely paraphrased; the headline finding concerns zero measurable return, not pilot failure.
Gartner 2025	<i>Press release: agentic AI project cancellations.</i> Over 40% of agentic AI projects expected to be canceled by end of 2027; “agent washing”; ~130 genuine agentic vendors.
Intercom 2026	<i>Customer Service Transformation Report.</i> 10% report mature deployment; 87% of mature teams report improved metrics against 62% overall.

SERVICE PERFORMANCE AND ECONOMICS

Salesforce 2026	<i>AI service agents and customer satisfaction.</i> AI agent adoption in service rose from 39% to 66%; CSAT the most-improved KPI; 70% see value within 60 days; data readiness cited as a major blocker by 72%.
Salesforce 2026	<i>Agentic Enterprise Index.</i> Seven in ten conversations resolved without escalation; 32% escalation rate steady across 170× conversation growth; +10 NPS on a named concierge deployment.
Salesforce 2025	<i>State of Service, 7th edition.</i> 30% of cases AI-handled, expected to reach half by 2027.
Gartner 2026	<i>GenAI cost per resolution forecast.</i> Cost per resolution to exceed \$3 by 2030, surpassing many offshore human agent costs; regulatory change to raise assisted interactions 30% by 2028.
Gartner 2026	<i>Customer service leadership survey.</i> 91% of service leaders under pressure to implement AI; service model redesign underway in ~80% of organizations.
Gartner 2025	<i>Agentic AI resolution forecast.</i> 80% of common service issues resolved autonomously by 2029, with a 30% operational cost reduction.

CUSTOMER TRUST AND SENTIMENT

Zendesk 2026	<i>CX Trends.</i> 95% expect an explanation for AI-made decisions; 79% want plain-language reasoning; 74% frustrated by repeating themselves; 81% expect continuity.
Intercom 2026	<i>AI Sentiment Report.</i> 54% would trust an agent with simple issues only; easy escalation and upfront disclosure are the top two comfort factors.
Qualtrics 2025	<i>AI-powered customer service failure rates.</i> Nearly one in five see no benefit from AI in service — about four times the rate for AI generally; 53% fear data misuse; 39% trust responsible data use.
Salesforce 2024	<i>State of the AI Connected Customer.</i> ~75% want to know they are dealing with an agent; 45% more likely to use one given a clear escalation path; 44% given explained logic. 2024 baseline.
Accenture Song 2025	<i>Customer Service on the Brink.</i> 87% likely to avoid a company after one bad experience; 35% do not trust AI in service.
PwC 2025	<i>AI Agent Survey.</i> 79% report agent adoption; executive trust for autonomous customer interactions at 22%; customer service the leading use case at 57%.
Forrester 2025	<i>2026 B2C Marketing, CX and Digital Predictions.</i> A third of companies expected to harm experiences with frustrating AI self-service.

DESIGN FRAMEWORKS AND PATTERNS

John Maeda 2026	<i>Design in Tech Report — From UX to AX.</i> The move from the gulf of execution to the gulf of evaluation; the agentic loop of action, outcome, feedback, correction.
Knight Institute, Columbia 2025	<i>Levels of Autonomy for AI Agents.</i> Five levels named by the human's role: operator, collaborator, consultant, approver, observer.
Anthropic 2026	<i>Measuring AI agent autonomy in practice.</i> Autonomy scored against risk; autonomy as an emergent property of a deployment rather than a model.
Conversation Design Institute 2026	<i>Agentic Experience Design.</i> Conversational capital; three authority tiers; the exclusions framework of vulnerability, power imbalance and nuanced judgment.
Google PAIR 2024–2026	<i>People + AI Guidebook.</i> Automate more when risk is low; let users supervise automation; automate in phases; give control back when automation fails.
Microsoft 2019–2026	<i>Guidelines for Human-AI Interaction (HAX Toolkit).</i> Eighteen guidelines including scope services when in doubt, support efficient correction, and make clear why the system did what it did.
IBM 2026	<i>Carbon for AI.</i> The AI label as a disclosure primitive with in-context explainability and revert behavior.
Victor Yocco 2026	<i>Practical UX patterns for agentic AI (Smashing Magazine).</i> Intent preview, autonomy dial, explainable rationale, confidence signal, action audit and undo, escalation pathway — with proposed target bands.
Notch 2026	<i>What is agentic experience; AX pattern library.</i> Human-in-the-loop checkpoint, autonomous action log, graceful agent failure; the field's own admission that patterns are unsettled.
EY Studio+ 2025	<i>How agentic AI enables a new approach to UX design.</i> Intent-driven rather than navigation-driven design; graceful intervention; trust protocols.
Nielsen Norman Group 2024–2026	<i>Generative UI; AI agents as users.</i> Outcome-oriented design within designer-defined guardrails; agents as a second class of user.

GOVERNANCE, FAILURE AND REGULATION

IBM Institute for Business Value 2026	<i>The AI control gap.</i> 54 agent incidents a year requiring human correction, 17% high severity; 77% say adoption outpaces governance; two in three accountable for systems they do not control; control-mature organizations report 25% fewer incidents and 16× more agents.
Microsoft AI Red Team 2026	<i>Taxonomy of failure modes in agentic AI systems, v2.</i> Human-in-the-loop bypass the most consistently exploited failure mode, largely via consent fatigue; goal hijacking; session context contamination.
European Union 2024	<i>Regulation (EU) 2024/1689, Article 50.</i> Transparency obligations: notice at first interaction, clear and distinguishable, plus machine-readable marking of synthetic output. Confirm dates against the Regulation.
State of California 2019	<i>B.O.T. Act, Bus. & Prof. Code §§ 17940–17942.</i> Bot disclosure required in commercial and electoral contexts.
Salesforce 2024	<i>Responsible Agentic AI Guidelines.</i> Accuracy, safety, honesty, empowerment, sustainability; explicit AI-generated content disclosure.
Lee & See 2004	<i>Trust in Automation: Designing for Appropriate Reliance (Human Factors).</i> The foundation for trust calibration and appropriate reliance.

Prepared by f/cx.ai — a customer experience design and innovation agency for the AI era. Sample data inside figures is illustrative and is not any named client's results. Frameworks presented here are working instruments and should be revised against your own evidence.