Deloitte.

Together makes progress

Trust by design: Building AI customers will want to use

Deloitte Trustworthy AI™ Services: **Human trust in AI**

Why the most powerful AI experiences aren't just built to work—they're built on trust

AI has a trust problem

Artificial intelligence isn't just a technology challenge it's a human one.

Deloitte's TrustID® index shows that customer trust drops by an average of 46% when AI enters an experience. That drop represents more than a perception problem; it's a strategic and financial risk. When trust falters, customers hesitate to share data, resist automation, and disengage from digital channels stalling innovation before it starts. When trust is high, they do the opposite: they experiment, advocate and spend. In fact, customers with high trust in a brand are 1.9x more likely to engage with its AI tools and 2.8x more likely to buy new products or services launched by that brand.

Leading organizations are treating AI implementation as more than a technical challenge they're designing AI experiences that earn trust from the start.

The sections that follow explore why trust in AI isn't experienced equally, how to design for it, and how to measure and refine it over time because the future of AI will be defined not by how intelligent it becomes, but by how much people trust it.

AI isn't experienced equally

Trust in AI isn't evenly distributed. It varies by **who people are** and **how they think** their experiences, expectations and comfort with technology all shape how willing they are to let AI in.

Who you are: Trust in AI begins with personal experience and that experience isn't the same for everyone. Our research shows that older generations are far more skeptical: Trust drops by 76% among baby boomers when brands introduce AI, compared with just 22% among Gen Z and 24% among millennials.

Education and income also shape trust. People with higher education levels or household incomes above $100,000 report smaller declines in trust when encountering AI, likely reflecting greater exposure and familiarity. By contrast, lower income or less tech exposed groups show deeper skepticism, particularly around data use and control.

While personal background shapes how people perceive AI, the environment they encounter it in the industry or service setting can be just as influential. Sectors where reliability and accountability are already built into the customer experience such as life sciences,

health care and financial services — retain significantly more confidence when introducing AI, with trust declines around 40%, compared with drops exceeding 55% in lower trust categories like technology, media and telecom.

How you think: Beyond identity and industry, trust depends on mindset — the beliefs and emotions people bring to each interaction. Deloitte's Mindsets Matter research finds that people cluster into distinct trust profiles when it comes to AI:[1]

- **Optimists** see AI as a leap forward. They expect personalization, speed and creativity, and they'll reward capability and reliability.

- **Maybes** are willing to experiment — but only once they see a clear benefit. Their trust grows when the value is obvious and the path forward is understandable.

- **Unawares** need an invitation and clearer cues. They're not opposed, but they don't engage until the experience feels seamless, safe and intuitive.

- **Avoidants** ask the hard questions before they engage: How am I in control? How is my data used? Where's the human in the loop? They need transparency and assurance before they'll commit.

These mindsets reveal that trust isn't static — it flexes with emotion, context and experience. The same person can move from skeptic to believer as AI earns confidence over time.

Consider customers on opposite ends of that spectrum. Gen Z optimists might eagerly try a new AI shopping assistant — expecting personalization, quick results, and a touch of fun. They'll reward capability and reliability first: Does it get them what they want, fast? Gen X pragmatists or avoidants, on the other hand, will engage only once they understand how recommendations are made and that their data won't be misused. For them, transparency and control aren't optional; they're the entry ticket.

For organizations, the takeaway is clear: Trust must be built through experience. Designing AI that adapts to both who people are and how they think is what turns hesitation into adoption, and adoption into advocacy.

How we measure trust

At Deloitte, we measure trust the way markets track performance — quantitatively, consistently and over time. TrustID is an always-on index that continuously captures perceptions of trust across industries, reaching more than 240,000 customers and employees each year.

Respondents answer a consistent set of questions assessing how well a brand demonstrates the Four Factors of Trust™ — whether people feel seen and valued (humanity), informed and in control of their own data (transparency), confident in quality and competence (capability), and assured that the organization keeps its promises (reliability). The average across these factors forms a composite TrustID score, which has proven to be highly predictive of human behavior.

The result is a data-rich, predictive measure that helps leaders pinpoint where trust is building or breaking down — and link those shifts to tangible business outcomes like adoption, advocacy and willingness to pay. By tracking trust alongside financial and operational metrics, organizations can not only understand what's happening in the moment but also anticipate where future growth — or resistance will occur.

And the financial impact isn't linear — it's compounding. Analysis of TrustID data shows that each one-point increase in trust at a lower level (for instance, moving from 30 to 31) corresponds to roughly a 3% rise in expected stock returns. The same one-point increase at a higher level (from 60 to 61) yields nearly double the effect — about 6%. In other words, the more trust an organization builds, the greater the momentum it creates — like a flywheel that accelerates with every turn.

Design for trust

Once organizations understand who they're designing for and how those people think, the next step is to design experiences that earn and sustain their confidence. Deloitte's Four Factors of Trust — **humanity**, **transparency**, **capability** and **reliability** — provide a roadmap for doing exactly that.

1. Humanity: [The AI tool] directly supports my specific needs and enhances my experience.

2. Transparency: The outputs of [the AI tool] are explainable and use straightforward and plain language to share key information, its purpose, data usage and usage policies.

3. Capability: [The AI tool] produces accurate, unbiased materials/outputs aligned with its stated purpose.

4. Reliability: [The AI tool] consistently and dependably delivers upon its purpose.

1. Humanity: Design AI for me.

In the airline industry, many carriers are experimenting with AI enabled virtual service agents to help travelers rebook flights, manage upgrades or track baggage — tasks that are urgent but routine enough to be resolved through efficient, automated channels. Yet when emotions run high — say, after repeated cancellations or a missed connection — the same system could recognize frustration and respond with empathy: "I understand this has been stressful. Would you like me to connect you with someone who can help?"

That distinction matters. Humanity in AI isn't about replacing people — it's about recognizing people. Research shows that hybrid models, where customers can escalate to a human, perform as well as the more costly human only service model across most satisfaction measures. Blending AI efficiency with human reassurance helps customers feel seen, understood and in control.[2]

The trust impact is measurable. For example, when service issues are resolved quickly and to the customer's satisfaction, individual trust scores more than double compared to baseline, with the strongest gains in **humanity** and **transparency**. Conversely, when problems linger or require multiple interactions, trust can drop sharply — by as much as 57%.

> **The lesson:**
> Designing for humanity isn't about making AI act like a person—it's about ensuring people still feel like one.

2. Transparency: Make the invisible visible — without overwhelming

Transparency in AI design is a balancing act. Too much detail and customers tune out; too little, and skepticism takes hold. The goal is to give its B2B customers enough clarity to build confidence, revealing how AI works without burying them in complexity.

Autodesk software offers a glimpse of what that balance looks like. The company introduced AI transparency cards plain language summaries that outline what data the model uses, how it's trained, and where humans remain in control. Much like a nutrition label, each card gives customers the essential facts at a glance: what's inside, how it works, and what to watch for.

The results were measurable. In Deloitte's pilot research, Autodesk's transparency scores increased by 26% and overall trust rose by 7% after customers viewed the cards. More important, by turning transparency into a design feature — not a disclaimer — Autodesk made AI feel more understandable, not more complicated.

> **The lesson:**
> Striking the right level of transparency means revealing enough to build confidence—without flooding people with information they can't use.

3. Capability: Make it work and make it work for the customer

Capability is table stakes for trust. Customers assume AI will work; what they're watching is whether it works for them. If the system fails, delivers inconsistent results or simply doesn't meet expectations, confidence evaporates.

That means designing with the user's needs, context and goals in mind and being open about how well (or not well) the AI is performing. Leading organizations are already doing this:

- Some e commerce platforms now include "why you're seeing this" explanations or confidence scores showing how well recommendations fit.

- Many financial institutions publish model accuracy ranges and retraining frequencies so clients can see when updates occur.

- A few health care innovators now display model performance data to help patients and clinicians understand when to trust the algorithm and when to rely on human judgment

Our research on trust finds that perceived capability the sense that an AI system is accurate, consistent and credible is one of the strongest predictors of adoption.

> **The lesson:**
> It's not enough for AI to work; it has to show its work—and show it's working for the customer.

4. Reliability: Keep your promises even when things change

Reliability is the quiet promise behind every AI interaction: It will do what it says, every time. Customers judge it in simple ways does the outcome stay consistent across channels and releases, and does the brand resolve issues quickly when something goes wrong?

Imagine a retailer's AI service that tracks orders, processes return and manages price adjustments. When the data is clear, it confidently delivers a consistent answer whether you ask in the app, on the website or through chat and meets that promise. When signals conflict or confidence dips, it fails safe: widening the delivery window, explaining the cause ("carrier scan delayed due to weather"), and offering a quick handoff to a human instead of guessing. It also makes stability visible, showing a "last checked" time stamp and a change log whenever rules or models update ("new carrier data source added"). Customers can see what changed before they feel it building trust through transparency, not mystery, and demonstrating reliability as a result.

Designing for reliability pays off. Customers who perceive an AI system as reliable are 1.8 times more likely to opt for a company's assistant and twice as likely to engage with its digital tools, such as chatbots. Consistency and clarity don't just improve satisfaction they increase adoption and long term loyalty.

> **The lesson:**
> Reliability isn't just uptime—it's consistency you can feel. Make stability visible, fail safe when unsure, and close the loop fast when things break.

Measuring and maintaining trust over time

Trust isn't static   it evolves with every interaction. Designing for trust is only the beginning; maintaining it requires ongoing measurement and adaptation.

Traditional metrics track output   such as deadlines met, budgets held and systems deployed   but they often miss the human signals that predict lasting loyalty. To understand the true return on AI, companies should connect operational, experiential and financial data into a single, continuous feedback loop:

- Operational metrics capture how the system performs by evaluating efficiency, speed and consistency of processes and technology. However, operational success can sometimes create experiential friction. For example, many contact centers measure average handle time (AHT) to understand the cost of a call. While companies aim to reduce AHT to manage costs, incentivizing agents to shorten interactions can make customers feel rushed or unheard. Leaders need to assess how operational gains impact the customer experience   and balance both accordingly.

- Experiential metrics capture how people feel as a result of operations and marketing. They measure confidence, understanding and emotional connection   how well a brand demonstrates humanity, transparency, capability and reliability in every interaction. Experience is shaped through operations   whether it's a conversation with a service representative (or a chatbot), a website interaction, or direct use of a product.

- Financial metrics capture how trust pays off. They measure the tangible outcomes that follow   growth in adoption, loyalty, conversion and share of wallet. When connected to experiential and operational signals, they complete the trust performance equation.

When these signals are measured together, leaders can see how operations drive experience   and how experience drives performance.

For example, imagine a financial services company rolling out an AI powered advisory assistant. Operationally, the tool reduces response time by 30% and increases portfolio recommendations per adviser by 20%. On the surface, it looks like success. But experiential data reveals a deeper issue: Trust scores in transparency and humanity fall after launch, as clients question how recommendations are generated and whether human advisers remain involved. Within six months, attrition rises among high value accounts   revealing the financial consequence of eroding trust.

Imagine instead that the company pulses the "Four Factors" as it evaluates returns. This allows leaders to detect and address trust gaps early. By adding clearer model explanations and human validation checkpoints, trust rebounds   followed by renewed growth in assets under management.

When these signals are measured together, leaders can see how operations drive experience   and how experience drives performance. Deloitte's TrustID research shows that shifts in trust are predictive, not reactive   able to forecast changes in customer purchase intent four months in advance. In other words, trust doesn't just explain performance; it anticipates it.

The trust advantage

As AI becomes universal, the next competitive edge won't come to which company deploys it fastest but to which one customers trust. Trust leaders aren't just better liked they're more resilient. TrustID data shows that high trust brands are 2.6x less likely to experience major trust declines when introducing AI than their lower trust peers.

Trust by design isn't ethical window dressing   it's a growth strategy. It guides where to invest, how to design, and how to measure success. In an era where algorithms can replicate almost anything, trust is what remains uniquely human   and uniquely valuable.

The Deloitte difference

At Deloitte, we believe trust is essential to scaling AI with confidence. It must operate on two levels: the system must be designed to perform reliably, and people must feel secure in using it.

That's why we've built an integrated platform for Trustworthy AI™ one that brings together machine level governance and human centered design. It's engineered to help organizations develop AI systems that are secure, transparent, explainable, and aligned with intended outcomes.

Backed by Deloitte's AI Institute, supported by global research, and informed by deep experience across industries, this platform helps organizations embed trust into AI development from day one transforming it from a reactive concern into a proactive capability for responsible growth and long term value.

Ashley Reichheld
Principal
Deloitte Consulting LLP
areichheld@deloitte.com

Dorsey McGlone
Managing Director
Deloitte Consulting LLP
dmcglone@deloitte.com

Courtney Sherman
Principal
Deloitte Consulting LLP
csherman@deloitte.com

Ryan Youra
Senior Manager
Deloitte Consulting LLP
ryoura@deloitte.com

1. Deloitte Digital, "Mindsets matter, Four keys to creating GenAI experiences consumers trust," March 27, 2025, accessed December 12, 2025. https://www.deloittedigital.com/us/en/insights/research/mindsets-matter-generative-ai-experiences-digital-product.html CHECK FORMATTING 1 & 2

2. Authors, **Can chatbot customer service match human service agents on customer satisfaction? An investigation in the role of trust**, Journal of Retailing & Consumer Services, Vol. 78, 2024, 103701, https://doi.org/10.1016/j.jretconser.2023.103701. this link is wonked; it went to article/paper "ARvolution: Decoding consumer moivation and value dimensions in augmented reality