

# Trust: The missing link in scaling AI

Deloitte Trustworthy AI™ Services: **Human Trust with AI**

Organizations are moving fast to invest in Generative AI (GenAI) but for many, the impact isn't keeping pace with the ambition. According to Deloitte's State of *Generative AI in the Enterprise Q4* report, only 11% of organizations that have rolled out GenAI tools to employees say those tools are used routinely. Nearly half of those with low usage report that their most advanced initiatives are falling short of ROI expectations.<sup>1</sup>

## What's getting in the way?

At its core, the challenge is trust—or, more accurately, the lack of it. When AI is introduced into the workplace, trust declines by an average of 39%. But when trust is present, the outcomes shift. Employees are 2.6 times more likely to use AI tools and use them consistently. On average, they save 2.2 hours per person, per week.<sup>2</sup>

The message is clear: Building technically trustworthy AI is essential. But to unlock its full value, organizations must also design systems and experiences that people genuinely trust.

## So, how do you define trust?

At Deloitte, we've spent years studying how trust operates inside organizations. We've made a substantial investment in research and built one of the world's most comprehensive trust datasets. Our TrustID® methodology is grounded in that work, drawing on annual trust scores from more than 250,000 customers and 60,000 employees to provide a real-time view of how trust is built, lost, and rebuilt across different industries, roles, and technologies.

From that research, we've identified the **Four Factors of Trust™**—the fundamental drivers influencing whether people engage or opt out. It doesn't matter whether you're showing up as a customer or an employee, here or abroad, in the C-suite or on the front line. These four factors are simply human:

- **Humanity:** Demonstrates empathy and kindness, and treats everyone fairly
- **Capability:** Creates quality products, services, and/or experiences
- **Transparency:** Openly shares information, motives, and choices in clear, plain language
- **Reliability:** Consistently delivers on promises and expectations

Let's bring this to life. Imagine you are part of a sales team, and your company rolls out a new virtual AI agent to support prospect engagement. It works—accurately handling common questions, surfacing relevant materials, and routing leads based on stated preferences. But trust doesn't follow automatically, even when the tool performs as promised.

You're curious about what this means for your role. No one has quite explained how the tool will help you grow, make your work more impactful, or free you up to focus on what really matters. That's a miss on **humanity**. You're not sure how your personal data is being used—or whether the company is tracking your interactions with the tool. That's a gap in **transparency**. The tool might schedule meetings for you, but it won't let you reschedule or add context when plans change. That's a **capability** issue—it solves part of the problem, but not the whole workflow. And when it comes to **reliability**, the challenge is knowing how and when to use it. Can it take notes? Write emails? Prioritize tasks? Some days it's helpful; other days it feels like more work to guide it than to do it yourself.



### Human Trust with AI in practice: A case from within

The best way to test any trust methodology is to use it ourselves. So, when Deloitte launched an internal GenAI chatbot for our professionals, we didn't just track usage—we applied our trust framework to understand what was working and pinpoint what was driving trust and where it was falling short.

The analysis revealed two trust gaps among professionals who hadn't used the tool or only used it occasionally—transparency and reliability scored notably lower than other trust factors. In many cases, professionals had questions about how their data would be used and whether the tool's outputs were safe and appropriate for client work. Regarding reliability, the issue wasn't the tool's performance—it was a lack of awareness of how many ways it could be applied.

To address the transparency gap, we developed a series of communications that broke down the chatbot's design in approachable, human language. We explained how the model was trained, what types of data it draws from, and—critically—what guardrails are in place to protect personal and client information. We also hosted live "Ask Me Anything" sessions with the AI team, where employees could raise concerns and get answers directly from those who built the tool. These conversations helped demystify the technology and replaced uncertainty with understanding.

To build reliability, we turned to the voices people trust most: their peers. We curated real world success stories from early adopters and showcased them across internal forums and team meetings, demonstrating how the chatbot was delivering consistent, high quality results across a variety of use cases. These examples gave users a clear picture of what "good" looked like and what to expect.

We also ran a series of interactive "prompt a thons" designed to spark creativity, reinforce various use cases, and let users experience and experiment with the tool's capabilities firsthand. These sessions generated excitement and surfaced new ways to use the chatbot across the business, turning skeptics into champions and observers into explorers. The results were compelling. We saw a 65% increase in visits per user, a 13% lift in repeat users, and significant improvements in trust scores—56% in transparency and 51% in reliability.

Now, we're taking that approach further by embedding the Four Factors of Trust into the design process. That means adding trust based questions in user testing, real time trust signal tracking, and AI experiences built with the human response in mind.

Because when people trust the technology, they don't just try it. They return to it, rely on it, and help it scale.

### Practical steps bring trust to life

When organizations work with Deloitte to build trust in AI, they gain access to engineers, human centered designers, and change leaders. Just as importantly, they tap into the behavioral science that reveals how trust is formed, measured, and sustained at scale.

Because trust is measurable and predictive, we're not just identifying where adoption is stalling. We can quantify how much trust influences performance and even model the likely outcomes of targeted interventions. That allows us to prioritize what matters most, move faster with more confidence, and focus resources where they'll have the most significant impact. Here's what that might look like at each stage:

**Design:** Teams begin by assessing current levels of trust in the organization. These insights help identify gaps and guide the prioritization of features and requirements that are most likely to drive confidence and adoption.

**Build and test:** Trust is continuously reinforced through user experience evaluations beyond binary metrics during build and testing. Instead of simple "thumbs up or down" feedback, we gather insights grounded in the Four Factors, allowing for more meaningful iteration and refinement.

**Deploy and scale:** Once the solution is launched and scaled, trust is monitored across different user populations in real time. This ongoing pulse helps inform targeted change management efforts and ensures feedback loops stay active, so trust isn't just protected; it continues to grow as the system evolves.

The impact of putting human trust at the center of AI development is clear. Organizations gain access to more rigorous, data backed methods for measuring performance, grounded in science, benchmarking, and behavioral insight. They see stronger adoption, often with double digit increases in tool usage. And they walk away with a defined, repeatable strategy for embedding and sustaining trust across teams, roles, and use cases.

### 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.

### The key to AI's enterprise potential

AI isn't new anymore, but it can feel new every day. From robotic process engineering to machine learning to GenAI and now agentic AI, new capabilities continue to unfold, which means the role of AI within an organization is always in motion. Putting that fast evolving technology to practical and profitable uses takes human and organizational resilience, and trust is the foundation.

Awareness of AI risks will always be part of responsible design and use, but a foundation of trust can free an organization to focus just as much on upside and value. If AI were all about risk, no one would use it. Deloitte's approach to Human Trust with AI helps organizations realize greater value from their AI investments by treating people and technology as an integrated system



**Ashley Reichheld**  
**Principal**  
Deloitte Consulting LLP  
[areichheld@deloitte.com](mailto:areichheld@deloitte.com)



**Ryan Youra**  
**Senior Manager**  
Deloitte Consulting LLP  
[ryoura@deloitte.com](mailto:ryoura@deloitte.com)

<sup>1</sup> Jim Rowan et al., [State of Generative AI in the Enterprise –Quarter four report](#), Deloitte, January 2025.

<sup>2</sup> Deloitte, TrustID® Index Benchmark Report, 2021–2024.

<sup>3</sup> Ashley Reichheld et al., ["If you want your team to use Gen AI, focus on trust,"](#) Harvard Business Review, January 24, 2025.