

Purpose determines performance

Deloitte Trustworthy AI™ Solutions: **AI Strategy**

Among the disciplines that help bring trust to artificial intelligence (AI), some are operational, such as workforce training or regulatory compliance. Some are more foundational. At the beginning of it all lies strategy: The precursor that sets the tone for technology before the technology begins and remains vital as long as it operates.

This is especially true for businesses and organizations as enterprise AI bears little resemblance to the consumer-facing, open-source AI tools that may assist people with academic research or travel plans. Large organizations look to AI to transform the way they work in ways that are only partly about technology—and to help operate those transformed organizations by making hugely consequential decisions.

The uses, potential benefits and potential risks of this commitment vary from company to company and industry to industry. A health care network operates in a trust and stakeholder environment that is different from what a financial services company faces, while a less regulated organization such as a retailer may be in a situation that looks nothing like either of those.

For that reason, beginning with the tech is shortsighted—and embracing the tech for its own sake, or because it's exciting, can be a recipe for regret. A *strategy* is the foundation that connects such an organization's ambitions with the steps it needs to take and the outcomes it hopes to achieve—not only in technology but across the operation.


Toys no more

A small business whose employees query open-source AI tools with simple prompts may not need a comprehensive strategy for doing so. But a large, complex organization that sets out to design and operate AI platforms is not merely using a new technology, it is fundamentally redesigning the way it does business. There's a difference between punching up a paragraph and transforming a clinical trial process.

However, many organizations in the latter category still have fragmented approaches to AI that lack centralized governance, goals for return on investment, or even an understanding of the metrics that can tell them if those goals are within reach. In contrast, a handful of organizations in sectors such as finance,

utilities and pharmaceuticals are accustomed to more regulation and regimentation in their work—which makes them more aware of the stakes but also makes it more complicated to design and operate AI platforms that can meet their standards.

Because the resultant need for trust is so high with AI and the consequences of failure potentially so large, its use at enterprise scale represents a fundamental shift in risk tolerance. Some organizations recognize this and work to drive AI strategy from the top down. But some aggregate what feels like a strategy from many discrete use cases across the organization. But “nothing has gone wrong yet” is a naïve understanding, especially in a circumstance where something that does go wrong may go unnoticed.



Building toward trust

Why is that? Unlike traditional software, AI doesn't crash or return error codes. It *always* offers an answer, the result of a probabilistic process that isn't the same twice and remains comparatively opaque to its human operators. When it fails—when the offered answer is not right or correct—it fails silently, and there's little to deter a human from putting the output into consequential use.

That means that at the strategy level, trust in AI is an exercise in risk mitigation. Organizations that are used to the more deterministic, transparent functions of non-AI systems have to change not only the tools they use, but the ways they work. (Including their protocols and procedures aligned to risk mitigation.) It's one thing to develop and run a proof of concept that passes tests on the lab bench. It's quite another to scale that system up and let it make meaningful decisions on the direction of the enterprise. And that's where trust in AI needs to be.

A strategy that builds trust into AI design and use can proceed in stages: from the aforementioned proofs of concept to internal-facing solutions whose risks may be easier to contain. From there, a next step might be customer-facing applications that expose the organization to risk based on identifiable processes. At the highest level of trust, agentic systems can bring a significant degree of automation and autonomy to external-facing activities where the risks are large and failure is difficult to contain.

With the right strategy, AI can even contribute to its own trustworthiness. One possible element of a proof of concept is an agentic element designed for the purpose of integrating trust—a “guardian agent” that watches over the others.

How Deloitte can help

By combining broad AI competencies with first-hand industry knowledge, Deloitte can help accelerate trustworthy AI™ transformation and value realization by accounting for the whole picture. That includes change management, financial and operational understanding alongside the technology itself. The result is a scalable and responsible roadmap and governance structure for each organization's AI journey that accounts equally for AI capabilities, business goals, and standards for ethics, fairness and security.

In practice this can include the following steps, among others:

- AI transformation strategy for trust
- Assessment of business objectives and how to identify and move key performance indicators
- Mapping AI capabilities to business goals
- Understanding stakeholders and their expectations
- Tech, AI and data-value potential
- Scalable roadmaps
- Build vs. buy decisions and vendor selection
- Value realization workstreams to manage return on investment
- Definition and architecture of AI and data technical capabilities



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 Trustworthy AI™ offering—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.

Remembering the why

AI can exceed human performance in speed and other parameters, but it's a common fallacy to say it "thinks." AI operates, and it needs to be surrounded by controls that come only from a deliberate strategy. That is why, in establishing trustworthiness and reliability, a sound strategy is the AI pillar that happens before any AI does.

There is more to strategy than safeguards. It is also the link between an organization's most advanced capabilities and its most fundamental aims. An AI project can meet all its technical specifications and still leave a management team or board wondering how, or whether, it had any impact on business performance or share price. But when the purpose comes first and the tool is made to meet it, the result can be worth the investment over the long term.

Contacts



Mike Segala

Principal

Deloitte Consulting LLP
msegala@deloitte.com



Baris Sarer

Principal, TMT Strategy and Analytics

Deloitte Consulting LLP
bsarer@deloitte.com

As used in this document, "Deloitte" means Deloitte Consulting LLP, a subsidiary of Deloitte LLP. Please see www.deloitte.com/us/about for a detailed description of our legal structure. Certain services may not be available to attest clients under the rules and regulations of public accounting.

This publication contains general information only and Deloitte is not, by means of this publication, rendering accounting, business, financial, investment, legal, tax, or other professional advice or services. This publication is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action that may affect your business. Before making any decision or taking any action that may affect your business, you should consult a qualified professional advisor. Deloitte shall not be responsible for any loss sustained by any person who relies on this publication.

Copyright © 2025 Deloitte Development LLC. All rights reserved.