



Responsible AI Is Becoming a Growth Strategy

Harvard Business Review (2026)

Context

As AI moves into products, services and everyday decisions, the same capabilities that personalise and automate also surveil, manipulate and discriminate. The authors argue that trust, long assumed and rarely measured, has become a scarce currency that customers, regulators and investors now price in. Corporate Digital Responsibility (CDR), the discipline of governing how firms build and deploy data and AI, is therefore shifting from a compliance obligation into a source of competitive advantage for the firms that invest early.

Key Insights

Why trust is now a commercial issue

- 3 waves of AI are arriving at once: generative (content and decisions at scale), agentic (systems acting autonomously), physical (service robots)
- Each widens the gap between what tech can do and what firms can account for, which is where trust risk concentrates
- One industry index reports 93% of IT leaders use or plan AI while only 23% of consumers trust firms to handle AI and their data responsibly, a gap the authors read as commercial, not just reputational (Thales Digital Trust Index 2026)

The CDR Calculus, a decision tool for AI deployment

- Temptations are the sharpest risk, as systems trained on behavioural data can identify and exploit individual psychological vulnerabilities at a scale no human sales team could match
- The tool's value is forcing leaders to place a specific system deliberately, rather than discovering their position after a regulator calls or a video goes viral

The Corporate Digital Responsibility (CDR) Calculus

Every AI initiative creates trade-offs between business performance and stakeholder well-being. The CDR Calculus helps leaders distinguish between practices that build lasting trust and those that create hidden risks.

Stakeholder impact		Business impact	
		NEGATIVE	POSITIVE
POSITIVE	Sacrifices Good for stakeholders, costly for business	Failures Bad for stakeholders and business	The Sweet Spot Good for stakeholders and business
	Transparent pricing Robust privacy controls Sustainable data centers Strict data minimization	Buggy features Poor user experience Inefficient operations Reputational damage	Trust-enhancing features Ethical personalization Value-based loyalty programs Green AI initiatives
NEGATIVE			
			Temptations Profitable, but harmful to stakeholders Dark patterns Predatory pricing Privacy-invasive tracking Unsustainable data centers

The three-stage CDR Playbook

- **Stage 1: Know what you have.** A living registry of every model, its decision scope and named owner, a risk tier per system, and a short set of published non-negotiables.
→ **Diagnostic [Openness & Delegation]:** Can we name every AI system currently making decisions that affect our customers, and do we know who is accountable for each?
- **Stage 2: Build oversight into the machine.** Govern at the design stage, not after failure, using red-teaming and bounded autonomy so agents hold only the permissions a task needs.
→ **Diagnostic [Openness & Delegation]:** If our most consequential AI system made a harmful decision tomorrow, could we reconstruct exactly what it did, why it did it, and who is accountable?
- **Stage 3: Turn CDR into competitive signal.** Auditable public transparency, compensation aligned to CDR objectives, and a hand in shaping emerging standards rather than waiting for them.
→ **Diagnostic [Openness & Opinions]:** Could a well-informed regulator, customer, or institutional investor review our public CDR disclosures and conclude that we govern AI more responsibly than our competitors?