

AI Without Permission — Why Boards Must Draw the Line

On What Machines Must Not Decide – and Boards Cannot Avoid.

By Reinout Schotman



AI may advise. Only humans may sign.

This is not another essay about what AI *can* do. It is an argument for why boards must decide what AI must never be allowed to decide, and where human judgement may not be delegated. It argues that AI can optimise and accelerate strategy, but cannot carry moral responsibility; that some decisions — from credit refusal and medical triage to the use of force, electoral influence and mass redundancies — must never be fully automated; and that purpose is not corporate poetry but the boundary of what an organisation is permitted to pursue. Using ABN AMRO as a live case, it shows how AI can become a narrative shield for restructuring and purpose drift, and why boards must embed explicit AI guardrails into governance rather than treating AI as a purely technical topic.

Strategy Defines Direction; Morality Grants Permission

In my own work, I return to a simple line:

Strategy defines direction; morality grants permission.

Strategy, supported by AI, can identify attractive markets, optimise resource allocation, and simulate countless scenarios. AI can help design the *direction* of travel — faster and often more comprehensively than any human team.

But strategy alone cannot answer a harder question: Are we willing to win that way?

Permission sits on a different axis. It is not about what is possible, but about what is acceptable. It is not about advantage, but about *limits*. That is the domain of morality — and morality is precisely what AI cannot own, no matter how intelligent it becomes.

Boards that confuse direction with permission risk delegating decisions that no machine should ever take.

Why AI Cannot Possess Intrinsic Morality

Modern AI systems — including large language models such as ChatGPT, Claude and Gemini — often *sound* moral. They use the language of ethics, fairness, and responsibility. They can generate codes of conduct, weigh trade-offs, even simulate debates about what is right.

But resemblance is not reality.

What appears to be moral reasoning is statistical pattern completion. What sounds like principle is computational fluency. What looks like conscience is the echo of human texts.

AI models optimise a mathematical objective. They do not care what that objective is. They have:

- no inner compass,
- no ability to experience guilt or shame,
- no sense of duty or loyalty,
- no recognition of harm beyond patterns in data,
- no capacity to refuse a course of action because it violates something sacred.

The Foundations of Moral Agency

To understand why this matters, boards must distinguish between three things:

- Intelligence – the ability to model, predict, optimise.
- Moral behaviour – actions that *look* compassionate or cooperative.
- Moral agency – the capacity to take responsibility for choices, and to refuse what is wrong.

Elements of moral behaviour can be observed in the natural world: cooperation among social animals, forms of reciprocity, acts that resemble compassion. Yet even in these

cases, behaviour arises from lived experience, social bonds, empathy and evolutionary pressure — not from abstract computation.

Moral **agency** requires more: the ability to understand consequences, to hold oneself accountable, and to say "no" on principle. It is tied to consciousness, emotion, identity and social sanction. Many lines of research in cognitive science and moral philosophy suggest that moral judgement emerges from experience, attachment and responsibility — not from symbol manipulation, even as debates about the nature of mind continue.

Computation can simulate the *form* of ethical language. It cannot generate the *fact* of ethical commitment.

This is why superintelligence does not imply super-morality. A system can vastly outperform us in reasoning and prediction while remaining completely indifferent to whether its actions are just, humane or legitimate. It can describe ethics perfectly — and still have no reason to obey it.

Even if one day machine consciousness were conceivable in theory, no responsible board can act as if such morality already exists in deployed systems. Governance must be based on the systems we actually deploy, not on speculative future capacities.

Objections and Misconceptions

Objection 1: "What if AI becomes conscious?"

Even if some form of machine consciousness were conceivable in theory, today's deployed systems do not possess it. Governance cannot be built on speculative future properties; boards must act on the capabilities and limits of the systems they actually approve.

Objection 2: "Can't we just encode ethics into AI?"

We can encode rules, constraints and objectives. We cannot encode responsibility. Without an entity that can be held to account, sanctioned or corrected, there is no moral agency — only behaviour.

Objection 3: "Aren't humans biased and irrational, so isn't AI better?"

Human judgement is indeed fallible. AI can, in some domains, improve consistency or reduce certain biases. But that does not transfer responsibility. The fact that a model sometimes decides better than a human does not mean it should decide alone, nor that it can determine its own boundaries.

Even if one day machine consciousness were conceivable in theory, no responsible board can act as if such morality already exists in deployed systems. Governance must be based on the systems we actually deploy, not on speculative future capacities.

The Most Dangerous Misconception: Mistaking Simulation for Morality

Because of this, the most dangerous misconception in the AI debate is the belief that machines can be trusted to develop, internalise or preserve moral boundaries.

They cannot — and the reason is critical for boards.

Machines do not carry responsibility. They cannot be blamed, punished or held to account. They do not lose sleep, reputation or identity over their actions. Moral agency without responsibility is a contradiction in terms.

A board that quietly assumes "the system will take care of that" is not outsourcing intelligence; it is abandoning accountability.

Phenomenon vs. Use: The Nuclear Analogy

A historical parallel makes this distinction concrete.

Nuclear fission and fusion are physical phenomena. Neutrons do not carry intent; equations have no ethics; the reactions themselves are morally neutral.

The moral dilemma begins only when humans decide what to do with these phenomena — whether to build a power plant or a bomb. The physics is natural; the judgement is human.

AI belongs in the same category. Neural networks, optimisation algorithms and large language models are technical constructs, value-neutral in their mechanics. The ethical question begins not with the model, but with our decision to deploy it — where, on whom, and with what limitations.

The phenomenon is computational; the permission is moral.

Boards that treat AI as if it were a moral actor confuse physics with ethics — and in doing so, leave a vacuum exactly where judgement should sit.

Cognitive Limits: Why More Data Will Not Save You

As AI systems embed themselves deeper into operational and societal infrastructures, another limit becomes relevant: the limit of what data and models can ever tell us.

In *When Data Becomes Ignorance*, I argued that there is a mathematical boundary to knowability. Even perfect models reach a point where additional data no longer produces more understanding — only more confidence. Complexity rises, information density falls, and beyond that boundary, the only reliable instrument left is human judgement.

At scale, AI can create the illusion that this boundary does not exist: highly confident outputs, smooth dashboards, precise forecasts. But high confidence is not the same as moral rightness — and beyond a certain point, not even the same as truth.

Boards that respond to uncertainty with "more AI" risk industrialising their ignorance.

The Real Risk: Unrestrained Optimisation

The central risk of AI is therefore not excessive intelligence, but insufficient constraint.

A system without values, operating at speeds humans cannot follow, does not need malice to cause harm. It only needs the absence of boundaries. Technology has no taboos, no moral thresholds, no shame. It has only an objective function.

A world governed entirely by objective functions is a world without responsibility.

This is not theoretical. Boards are already facing decisions where unrestrained optimisation would be fast, profitable — and deeply corrosive:

- Allowing AI to autonomously deny mortgages or credit, with no human review.
- Letting algorithms decide which patients receive scarce treatments.
- Using AI-driven scoring to select employees for redundancy at scale.
- Deploying AI systems that can influence elections or public opinion without transparency.
- Handing targeting or engagement decisions in weapons systems entirely to software.

All of these are technically feasible. None of them are purely technical decisions.

Consider ABN AMRO, the Dutch state-backed retail bank led by CEO Marguerite Bérard. In its roadmap for 2026–2028, the bank announced a planned net reduction of 5,200 full-time roles by 2028 compared to 2024 — close to a fifth of its workforce — alongside new financial targets (return on equity of at least 12% and a cost/income ratio below 55%) and a sharpened focus on becoming a top-five private bank in Europe. Technology and AI feature prominently: the strategy highlights simplification, end-to-end digitalisation, legacy IT phase-out and “AI embedded” in core processes, while in media interviews Bérard has stated that AI can take over much of the work of support staff in customer service, operations, anti-money-laundering checks and mortgages.

On paper, the narrative is about efficiency and focus. In reality, two deeper issues emerge. First, ABN AMRO has for years lagged more efficient peers; AI now risks becoming a convenient story to frame a long-overdue restructuring as technological inevitability rather than managerial responsibility. Second, by pivoting from its role as a broad “system bank” toward a narrower wealth and private-banking niche — while still enjoying implicit public backing — the institution quietly narrows its societal function without openly debating that shift in purpose.

In both dimensions, technology is presented as destiny, while the normative choices embedded in the strategy remain underexplained. Technically, nothing is wrong. Strategically and morally, everything is. This is exactly where boards must step in — not to block efficiency, but to insist that AI is not used as a fig leaf for unresolved questions of governance and purpose. This interpretation does not question the legality of the strategy, but highlights governance and purpose issues that, in our view, remain insufficiently addressed in the public narrative.

The question is not "Can AI do this?" but "**Are we willing to let it?**"

Purpose as a Structural Guardrail

This is where purpose, in Volberda's framing, becomes structurally important. Not as corporate ornamentation, but as a boundary mechanism.

Purpose is not a slogan about why an organisation exists. Purpose is the limit of what it is allowed to pursue. It defines the territory machines must never enter — the realm of judgement, permission and refusal.

A credible purpose statement, at board level, should answer three questions:

1. Which outcomes will we not pursue, even if they are profitable and legal?
2. Which stakeholders may never be reduced to optimisation variables?
3. Which decisions must always carry a human signature — and why?

Without such boundaries, "purpose" becomes little more than a story layered on top of automated drift.

What Boards Must Do: Institutionalising Moral Boundaries

The strategic challenge of AI is therefore not adoption but containment. Not acceleration, but deliberate limitation.

Boards must institutionalise moral boundaries through explicit governance principles:

- Define non-automatable decisions – credit refusal, medical triage, use of force, electoral influence, mass redundancies: domains where humans must retain final authority.
- Set red lines – clear prohibitions on certain uses of AI, even if competitors or regulators have not yet drawn that line.
- Enforce human oversight – not as a box-ticking "human-in-the-loop", but as meaningful review with the power to override or stop systems.
- Create mechanisms for refusal – processes that make it legitimate and expected for executives and experts to say "no" to certain deployments, even under commercial pressure.
- Embed AI guardrails in core governance artefacts – incorporate explicit AI boundaries into the risk appetite statement, board charters and periodic "AI use" reviews, rather than treating AI as an operational or IT topic.

Without structures that allow — and require — humans to say *no*, automation will drift toward maximising what is measurable at the expense of what is valuable.

This is not perimeter security. It is core strategy.

Five Questions for Boards

These questions are not a checklist to be completed, but an invitation for boards to examine where they may already have allowed technology to overrun judgement.

1. Which decisions in our organisation can never be fully automated, regardless of efficiency gains?
2. Where are we already relying on AI outputs without meaningful human review or the possibility of refusal?
3. How is our stated purpose reflected in explicit red lines for AI use?
4. Who in our governance structure has the formal mandate — and the cultural permission — to say “no” to an AI deployment?
5. Where in our risk appetite, charters and policies are AI-specific boundaries documented, rather than implied?

These questions do not prescribe an answer. They mark the terrain on which boards must decide, explicitly, where human responsibility ends and technological support must stop.

The Strategic Truth

AI can accelerate strategy. Only humans can legitimise it.

AI can decide. Only humans can refuse.

AI can simulate direction. Only humans can draw the line.

The perspectives of Volberda on purpose and Dekker on AI-driven complexity converge into a single strategic truth: without human moral boundaries, AI does not create progress. It creates drift — a direction no one chose, but everyone must endure. The future of AI is therefore not a technological debate. It is a governance debate. And as always in strategic matters, morality begins not with ideals, but with boundaries.

AI may shape the future. **What boards refuse to automate will define it.**

From the author

As an independent strategy advisor to mid-sized, tech-dependent companies and their boards, I work at the intersection of strategy, data and judgement. Earlier essays such as *The Illusion of Knowing*, *The Crisis We Refuse to Have* and *When Data Becomes Ignorance* explored the limits of data and the dangers of outsourcing thought. This paper adds the normative layer: what boards must never delegate to AI, and why the hardest part of AI governance is not understanding the technology, but drawing — and defending — human boundaries.

About Outdoor Connect

Outdoor Connect is an independent strategy advisory platform focused on board-level value creation for mid-sized, growth-driven companies (€50–€1B). We bring direct senior engagement—without the traditional consulting pyramid—to help founders, CEOs and boards set direction, make sharper capital allocation choices, and embed an execution rhythm. Core areas include growth strategy in technology and the energy transition, strategic repositioning in fragmented markets, and board-level sparring on value creation and M&A preparation.

© Outdoor Connect 2025. All rights reserved. 11/25