



© Banstanks/AdobeStock

LIGHTS - Technology

Algorithmic governance and the diffusion of responsibility: What becomes of accountability in systems without subjects?

ESCP Impact Paper No.2026-44-ENG

Benoît HEILBRUNN
ESCP Business School

ESCP RESEARCH INSTITUTE OF MANAGEMENT (ERIM)

[Algorithmic governance and the diffusion of responsibility: What becomes of accountability in systems without subjects?

Benoît Heilbrunn*

ESCP Business School

Abstract

Artificial intelligence and algorithmic governance are commonly presented as challenges of transparency, explainability, and ethical oversight. This paper argues that the deeper transformation concerns responsibility itself. Rather than eliminating responsibility, algorithmic governance reconfigures the conditions under which it can be exercised. Drawing on Günther Anders' concept of Promethean gap and Hannah Arendt's analysis of the banality of evil, the paper shows how contemporary organizations increasingly separate action from comprehension and judgment from decision-making. As organizational decisions emerge from complex socio-technical systems, responsibility becomes progressively detached from identifiable actors and redistributed across infrastructures, procedures, and multiple forms of agency. The paper's main contribution is to distinguish accountability from auditability. While organizations have become increasingly capable of tracing decisions, documenting processes, and demonstrating procedural compliance, they have become less able to identify subjects genuinely capable of answering for the consequences of those decisions. The paper concludes by discussing the implications of this transformation for organizations and business schools, arguing that preserving responsibility requires protecting spaces of human judgment rather than relying solely on technical mechanisms of transparency and oversight.

Keywords : responsibility, algorithmic governance, promethean gap, distributed agency

*Professor, ESCP Business School

Algorithmic governance and the diffusion of responsibility: What becomes of accountability in systems without subjects?

Introduction: from accountable actors to auditable systems

When an applicant is rejected by an AI-assisted recruitment system, who is responsible? The recruiter may explain that the recommendation came from the software. The software provider may point to the statistical model. Engineers may argue that they merely designed the architecture, while managers insist that they only followed organizational procedures. The decision can often be reconstructed, documented and audited, yet no individual can fully explain or assume responsibility for it. The process is traceable, but responsibility remains elusive. This situation is no longer exceptional. Across contemporary organizations, algorithmic systems increasingly mediate decisions concerning recruitment, credit allocation, performance evaluation, insurance, logistics, customer management, and resource allocation. These systems do not simply automate existing decisions. They redefine the conditions under which decisions are produced by selecting relevant variables, structuring information, generating predictions, and delimiting the range of acceptable actions. Human actors continue to participate in these processes, yet their agency becomes embedded within technical infrastructures whose complexity frequently exceeds individual understanding.

This transformation challenges one of the central assumptions of modern organizations: that responsibility can ultimately be attributed to identifiable decision-makers. Contemporary debates on artificial intelligence have largely focused on transparency, explainability, fairness, and bias. Considerable attention has been devoted to developing mechanisms capable of rendering algorithmic systems more explainable and auditable. Yet a more fundamental question remains insufficiently explored: what becomes of responsibility when decisions emerge from socio-technical systems whose operation no single actor fully controls or comprehends? This paper argues that algorithmic governance does not eliminate responsibility; it fundamentally transforms it. More precisely, it transforms accountability into auditability. Organizations increasingly develop sophisticated mechanisms capable of tracing decisions, documenting procedures, and verifying compliance, while simultaneously weakening the possibility of identifying subjects genuinely capable of answering for those decisions. Responsibility becomes simultaneously omnipresent in organizational discourse and increasingly difficult to locate in organizational practice. The paper first positions its argument within current debates on AI governance before mobilizing Anders and Arendt to explain why algorithmic systems transform accountability into auditability. It concludes by discussing the implications for organizations and management education.

From explainability to accountability: beyond the AI governance debate

The rapid diffusion of artificial intelligence has generated an extensive literature on algorithmic governance and Responsible AI. Much of this work has focused on improving the transparency, fairness, explainability, and accountability of algorithmic systems (Pasquale, 2015; Burrell, 2016; Mittelstadt et al., 2016; Floridi & Cowls, 2019). Scholars have examined algorithmic opacity, bias, explainable AI, human oversight, and regulatory frameworks such as the European AI Act, all of which seek to ensure that automated

decision-making remains trustworthy and compatible with democratic and organizational values.

These contributions have considerably advanced our understanding of the technical and ethical challenges posed by artificial intelligence. Yet they largely share a common assumption: that responsibility remains fundamentally intact and that the central task is to reconnect algorithmic decisions to accountable actors through better governance mechanisms. Explainability, transparency, documentation, and human oversight are therefore presented as ways of restoring accountability within increasingly automated decision processes.

This paper adopts a different perspective. Rather than asking how algorithms can be made more accountable, it asks how algorithmic governance transforms the very conditions under which accountability becomes possible. The question is therefore not simply whether an algorithmic decision can be explained, but whether contemporary organizations still produce identifiable subjects capable of exercising judgment and assuming responsibility for the consequences of that decision.

This distinction requires differentiating three concepts that are often conflated. Explainability concerns the intelligibility of algorithmic outputs. Auditability concerns the capacity to reconstruct decision processes and verify procedural compliance. Accountability, however, refers to something fundamentally different: the existence of actors capable of justifying their decisions, exercising reflective judgment, and answering for their consequences.

Contemporary organizations increasingly strengthen the first two dimensions while inadvertently weakening the third. They become remarkably effective at documenting data flows, recording model versions, tracing decision pathways, and demonstrating regulatory compliance. At the same time, responsibility itself becomes progressively dispersed across software developers, data scientists, managers, compliance officers, and algorithmic infrastructures. Decisions become easier to reconstruct but increasingly difficult to attribute.

The argument developed here therefore shifts the discussion from the governance of artificial intelligence to the anthropology of responsibility. Rather than treating responsibility as a property that can simply be added to technological systems through better regulation, it examines how algorithmic governance reconfigures the relationship between action, judgment, and responsibility itself. To explore this transformation, the paper mobilizes Günther Anders' concept of the *Promethean gap* and Hannah Arendt's analysis of the *banality of evil*. Together, these perspectives suggest that the central challenge posed by algorithmic organizations is not merely opacity but the gradual displacement of responsibility from accountable subjects to auditable systems.

The Promethean gap and the limits of comprehension

Consider a credit officer in a large bank using an AI-assisted credit scoring system. The officer receives a recommendation generated from hundreds of variables processed through a machine-learning model. Although the recommendation influences the final decision, neither the officer nor the branch manager can fully explain how the score was produced. Data scientists understand the architecture of the model but not every individual prediction. Senior executives approve the use of the system without mastering its technical functioning. Each actor contributes to the decision, yet not a single one possesses a complete understanding of the process that produced it. This situation

illustrates what Günther Anders described as the *Promethean gap*. In *Die Antiquiertheit des Menschen* (1956/2016), Anders argues that modern technological societies are characterized by a growing discrepancy between what human beings are capable of producing and what they are capable of imagining, understanding, or experiencing. Human action increasingly generates consequences that exceed human comprehension. Technology extends our capacity to act while simultaneously reducing our capacity to understand what our actions ultimately produce. Artificial intelligence represents perhaps the most advanced contemporary manifestation of this Promethean condition.

Although Anders developed this argument in response to industrial technologies and nuclear weapons, his diagnosis appears remarkably relevant to contemporary algorithmic systems. Artificial intelligence dramatically expands organizations' capacity to process information, generate predictions, and coordinate decisions at unprecedented speed and scale. Yet this increase in computational power comes with a corresponding decline in individual intelligibility. Decisions increasingly emerge from interactions between data infrastructures, machine-learning models, organizational routines, and multiple human actors whose combined effects no participant fully understands.

The Promethean gap therefore no longer concerns only the disproportion between production and imagination. It also characterizes organizational decision-making itself. Engineers design models whose behaviour evolves through continuous learning. Managers rely on outputs whose internal logic they cannot reconstruct. Executives authorize systems whose downstream effects extend far beyond their initial intentions. The organization becomes capable of producing decisions that no single individual could have produced — or fully understood — alone.

This transformation has profound implications for responsibility. Classical theories of accountability presuppose that actors understand the decisions for which they are expected to answer. Responsibility requires more than causal participation; it presupposes a meaningful relationship between action, comprehension, and consequence. The Promethean gap progressively erodes this relationship. Individuals continue to make decisions, yet increasingly within socio-technical systems whose complexity exceeds their cognitive horizon.

Responsibility therefore does not disappear because humans are replaced by machines. Rather, it becomes destabilized because action is progressively detached from understanding. The more organizations rely on algorithmic infrastructures, the greater their capacity to act, while their members become progressively less capable of comprehending the totality of what these actions produce. Responsibility becomes structurally fragile precisely because comprehension no longer goes with action. This first transformation prepares the second one. If actors can no longer fully understand the systems within which they operate, the role of judgment itself begins to change. It is this second transformation that Hannah Arendt's analysis helps illuminate. If Anders explains why we understand less, Arendt explains why we judge less.

The banality of action and the erosion of judgment

Imagine a manager responsible for evaluating employee performance. Rather than relying primarily on direct observation or discussion with team members, she consults a dashboard aggregating productivity indicators, behavioural metrics, attendance records, customer feedback, and algorithmic performance scores. The system recommends identifying one employee as underperforming. The manager follows the recommendation. When challenged, she does not justify the decision through her own

judgment but by referring to the indicators generated by the system. The decision appears objective precisely because it has become procedural.

This situation illustrates a transformation that Hannah Arendt helps us understand. In *Eichmann in Jerusalem* (1963/2006), Arendt famously argued that modern bureaucratic systems do not necessarily produce evil through malicious intentions but through the gradual disappearance of reflective judgment. What she called the *banality of evil* did not describe ordinary crimes; it described ordinary forms of obedience in which individuals cease to think critically about the meaning and consequences of their own actions. The problem was not ignorance or incompetence but what Arendt called *thoughtlessness*: the inability — or refusal — to exercise judgment independently of institutional procedures.

Algorithmic governance does not reproduce the historical circumstances analysed by Arendt, but it reinforces a similar organizational dynamic. Contemporary organizations increasingly replace deliberation with procedural execution. Decisions become embedded within standardized workflows supported by predictive models, automated recommendations, and performance dashboards. Rather than asking whether a decision is justified, actors are encouraged to verify whether the prescribed procedure has been correctly followed.

The authority traditionally attached to managerial judgment is progressively transferred to technical systems. Algorithmic recommendations derive their legitimacy not from deliberation but from their apparent objectivity, statistical sophistication, and computational neutrality. The question shifts from "What should we do?" to "What does the system recommend?" Judgment is not prohibited; it simply becomes increasingly unnecessary.

This evolution transforms the role of managers. Rather than acting as authors of decisions, they become interpreters, validators, or implementers of algorithmically generated outputs. Their expertise increasingly consists in managing procedures rather than exercising independent judgment. Responsibility consequently changes its meaning. Instead of asking whether a decision is right, organizations increasingly ask whether the established procedure has been correctly followed.

Arendt's analysis therefore illuminates a second dimension of algorithmic governance. Anders showed that technological systems exceed our capacity for comprehension. Arendt helps explain what happens once this gap becomes institutionalized. As understanding declines, organizations compensate by strengthening procedures. Reflective judgment is progressively replaced by procedural compliance, and responsibility becomes increasingly detached from individual acts of deliberation.

This does not imply that managers become morally indifferent or passive. On the contrary, they may remain highly committed, competent, and conscientious. Yet their commitment is increasingly directed toward ensuring that organizational processes function correctly rather than questioning whether those processes themselves deserve to be followed. Responsibility thus shifts from judging actions to administering procedures.

The result is a profound transformation of organizational accountability. Decisions continue to be made by human beings, but they are increasingly experienced as the implementation of system outputs rather than as acts of personal judgment. The actor remains visible, while the subject of responsibility gradually disappears.

The transformation of accountability

The transformations described by Anders and Arendt converge towards a more fundamental shift in contemporary organizations. The central issue is no longer simply that algorithmic systems are opaque or that human judgment is weakened. Rather, the very meaning of organizational responsibility is being redefined.

These two notions are often treated as interchangeable. They are not. Accountability presupposes identifiable actors capable of explaining, justifying, and assuming responsibility for their decisions. It implies a relationship between action, judgment, and consequence. An accountable manager is not simply someone who participated in a decision but someone who can answer for it, justify it, and accept its consequences. Auditability refers to something different. It concerns the possibility of reconstructing processes, verifying procedural compliance, documenting decisions, and tracing the sequence of operations that produced a particular outcome. Auditability asks whether procedures have been correctly followed; accountability asks who should answer for the decision itself.

Algorithmic governance increasingly gives precedence to the second logic rather than the first. Organizations invest heavily in documentation, logging systems, traceability, compliance procedures, and explainability mechanisms. Every stage of decision-making can often be reconstructed with remarkable precision. Data inputs are archived. Model versions are recorded. Human interventions are timestamped. Validation procedures are documented. From a technical perspective, decisions become more traceable than ever before.

Yet this expansion of traceability does not necessarily strengthen responsibility. On the contrary, it often contributes to its diffusion.

Consider again an AI-assisted recruitment process. The data scientist designed the model. HR specialists selected the variables. External software providers supplied the technological infrastructure. Recruiters interpreted the recommendations. Managers approved the hiring decision. Compliance officers ensured that regulatory procedures were respected. Each participant can legitimately explain their own contribution. Yet none can fully account for the decision as a whole. Responsibility becomes fragmented across a chain of actors, each of whom remains accountable only for a small segment of the process while the overall decision escapes personal appropriation.

The paradox is therefore striking. The more organizations improve their capacity to audit decisions, the more difficult it becomes to identify actors capable of assuming responsibility for those decisions. Traceability expands while accountability shrinks.

Paradoxically, however, this discursive expansion coexists with a practical weakening of responsibility. As decision-making becomes distributed across increasingly complex socio-technical infrastructures, responsibility itself becomes progressively detached from identifiable subjects. It circulates through procedures, protocols, governance mechanisms, and technological systems without ever stabilizing in actors capable of genuinely answering for the consequences of action.

Responsibility therefore undergoes a profound ontological transformation. It ceases to function primarily as an attribute of persons and increasingly becomes a property of systems. Organizations no longer rely principally on responsible individuals but on responsible architectures, responsible processes, responsible governance frameworks, and

responsible technologies. The language of responsibility remains ubiquitous, yet the subject traditionally capable of assuming it becomes increasingly elusive.

This is the central paradox of algorithmic governance. Modern organizations have become extraordinarily effective at demonstrating that decisions comply with established procedures. They have become considerably less effective at identifying who ultimately stands behind those decisions. Auditability progressively substitutes for accountability.

This distinction has important implications. If the problem were merely one of algorithmic opacity, better explanations or more transparent models might suffice. But if the deeper transformation concerns the displacement of responsibility from subjects to systems, then technical improvements alone cannot resolve the problem. What is at stake is not simply how organizations explain decisions, but whether they continue to preserve identifiable spaces in which human judgment remains possible and responsibility can still be genuinely assumed.

This transformation should not be understood as the disappearance of responsibility. Organizations continue to produce an abundance of responsibility discourse. Corporate reports emphasize ethical AI, responsible innovation, sustainability, governance, and stakeholder engagement. Business schools multiply courses devoted to responsible leadership and ethical decision-making. Regulatory frameworks increasingly require organizations to document responsible AI practices. Responsibility has become one of the dominant moral vocabularies of contemporary organizations.

Implications for organizations: preserving responsibility in algorithmic systems

If the central challenge of algorithmic governance lies not merely in opacity but in the diffusion of responsibility, then improving algorithmic transparency alone cannot suffice. Organizations increasingly invest in explainable AI, ethical guidelines, documentation procedures, and governance frameworks. These initiatives are valuable, yet they often remain focused on the technical properties of algorithms rather than on the organizational conditions that make responsibility possible.

The first implication is that organizations should distinguish more clearly between auditability and accountability. A decision that can be reconstructed is not necessarily a decision for which someone can genuinely answer. Compliance mechanisms, documentation, and traceability should therefore not be confused with responsibility itself. Organizations need identifiable actors capable of exercising judgment and assuming responsibility for the consequences of decisions, even when those decisions are supported by algorithmic systems.

Second, organizations should preserve spaces for human deliberation. Artificial intelligence should assist managerial judgment rather than replace it. Human oversight should not become a symbolic validation step performed after the algorithm has already determined the outcome. Instead, organizations should create institutional arrangements in which managers remain able — and expected — to question algorithmic recommendations, challenge underlying assumptions, and justify departures from automated outputs.

Third, responsibility should be understood as a collective organizational capability rather than merely an individual virtue. Distributed decision-making does not eliminate responsibility; it requires organizations to clarify how responsibility is shared across

technical, managerial, and strategic functions. Engineers, data scientists, operational managers, compliance officers, and executives should not simply contribute sequentially to algorithmic decisions; they should also participate in mechanisms that enable shared reflection on their consequences.

Finally, organizations should recognize that not every dimension of managerial work can or should be optimized through algorithmic procedures. Many managerial decisions involve uncertainty, contextual interpretation, competing values, and ethical dilemmas that cannot be fully reduced to predictive models. Preserving responsibility therefore means preserving institutional spaces where judgment remains possible precisely because procedures do not determine every outcome.

The challenge posed by algorithmic governance is therefore not to slow technological innovation but to prevent organizations from confusing procedural correctness with responsible action. Auditability strengthens organizational control; accountability preserves organizational legitimacy. Future governance systems will require both, but they should never allow the former to substitute for the latter. Organizations should not merely ask whether algorithms comply with procedures but whether they preserve identifiable subjects capable of judgment.

Business schools and the education of responsible judgment

Business schools occupy a particularly strategic position in this transformation. Business schools increasingly teach students how to manage algorithms. They should also teach them when to resist them. They do not merely observe the diffusion of algorithmic governance; they actively contribute to shaping the future managers, consultants, and executives who will design, deploy, and govern these systems. As such, they stand at the intersection of two potentially contradictory imperatives. On the one hand, they seek to equip students with the analytical and technological capabilities required by increasingly data-driven organizations. On the other, they continue to affirm responsibility, ethical leadership, and critical thinking as defining characteristics of managerial education.

This dual ambition reveals a deeper tension. Contemporary business schools increasingly operate within the same socio-technical logic that characterizes the organizations they serve. Their own performance is measured through rankings, accreditation standards, employability statistics, citation indexes, publication metrics, student satisfaction surveys, and salary outcomes. These indicators do not merely evaluate educational performance; they shape institutional priorities, pedagogical choices, and definitions of success. Like contemporary corporations, business schools are themselves governed through increasingly sophisticated systems of measurement and audit.

This institutional environment inevitably influences the way responsibility is taught. Responsibility tends to be presented as a managerial competency that can be acquired through dedicated courses, ethical frameworks, sustainability modules, or responsible leadership programmes. Such initiatives undoubtedly remain valuable. Yet they also risk reducing responsibility to another measurable competency within an already highly proceduralized educational environment.

If responsibility is increasingly transformed into a systemic rather than individual property, management education faces a more fundamental challenge. The objective cannot simply be to teach future managers how to use artificial intelligence responsibly. It must also prepare them to recognize situations in which algorithmic systems weaken the very conditions that make responsibility possible.

This implies moving beyond technical AI literacy. Future managers need to understand not only how algorithms function, but also how algorithmic infrastructures reshape organizational authority, redistribute agency, and transform judgment. They should be trained to recognize that explainability is not equivalent to accountability, that procedural compliance is not synonymous with responsible action, and that technical optimization cannot substitute for ethical deliberation.

Business schools should therefore cultivate forms of education that preserve reflective judgment alongside analytical competence. Students should be encouraged to question algorithmic recommendations rather than merely interpret them, to deliberate collectively about situations characterized by uncertainty and conflicting values, and to accept responsibility even when decisions emerge from highly distributed socio-technical systems. The capacity to challenge an algorithm may become as important as the capacity to build one.

Ultimately, the purpose of management education is not simply to produce managers capable of operating increasingly sophisticated technologies. It is to educate subjects capable of judgment. Algorithmic governance makes this task more difficult, but also more necessary. If organizations risk confusing auditability with accountability, business schools should become places where this distinction remains visible and where responsibility continues to be understood not as a procedural requirement but as a fundamentally human practice.

Conclusion: what becomes of responsibility in algorithmic systems

Artificial intelligence has often been presented as a technological challenge requiring better regulation, greater transparency, and stronger governance mechanisms. These objectives remain essential. Yet they address only part of the problem.

The deeper transformation concerns the nature of responsibility itself. Algorithmic governance does not simply automate decisions; it reorganizes the relationship between action, judgment, and accountability. As organizations increasingly rely on socio-technical systems that distribute agency across infrastructures, procedures, and multiple actors, responsibility gradually loses its traditional anchor in identifiable subjects.

This paper has argued that the central transformation is the progressive substitution of accountability by auditability. Contemporary organizations have become remarkably effective at documenting decisions, tracing processes, and demonstrating procedural compliance. At the same time, they have become less capable of identifying who ultimately answers for the consequences of those decisions. Responsibility has not disappeared. It has become systemic.

This distinction carries important practical consequences. The challenge facing organizations is not merely to design more transparent algorithms but to preserve institutional spaces where human judgment remains possible and where responsibility can still be assumed rather than merely documented. Similarly, the challenge facing business schools is not only to teach Responsible AI or ethical leadership, but to educate managers capable of resisting the quiet displacement of judgment by procedure.

The risk posed by algorithmic governance is therefore neither technological nor moral in the conventional sense. It is institutional. Modern organizations increasingly know how to trace decisions without knowing who should answer for them. They excel at auditing processes while gradually losing the subjects upon whom responsibility ultimately

depends. Algorithmic governance does not eliminate responsibility. It changes its location. The danger is therefore not that nobody is responsible, but that responsibility is invoked everywhere and assumed nowhere. Contemporary organizations increasingly know how to audit decisions without knowing who should answer for them

References

Anders, G. (2016). "The obsolescence of human beings: On the soul in the age of the second industrial revolution". Polity Press.

Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, 20(3), 973–989.

Arendt, H. (2006). "Eichmann in Jerusalem: A report on the banality of evil". Penguin Books.

Burrell, J. (2016). How the machine "thinks": Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1), 1–12.

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1).

Kitchin, R. (2017). Thinking critically about and researching algorithms. *Information, Communication & Society*, 20(1), 14–29.

Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 267, 1–38.

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 1–21.

Pasquale, F. (2015). "The black box society: The secret algorithms that control money and information". Harvard University Press.