Markets, Organizations & Public Value
GOVERNANCE PERSPECTIVE | COMMISSIONED SYNTHESIS

Corporate Governance for AI-Enabled Decision Systems

Mark Collins - Corresponding author

Abstract

This governance perspective examines corporate governance for AI-enabled decision systems. The organizing question is how boards and executives can oversee material algorithmic decisions without reducing oversight to generic principles. Ten related scholarly sources are synthesized through a decision-centered framework spanning problem definition, mechanism, measurement, evaluation, implementation, and governance. The review does not invent experiments, pooled estimates, or unreported quantitative results. It instead evaluates the strength and transferability of the available evidence, with particular attention to assuming technical validation alone satisfies fiduciary and stakeholder obligations. The resulting framework links technical or empirical performance to explicit use conditions and identifies tests that should precede wider adoption in firms using AI in finance, employment, marketing, and operations.

Keywords: corporate governance, artificial intelligence, boards, model risk, accountability

Synthesis lens	Principal question
Mechanism	the chain from strategic purpose and data sourcing through model ownership, deployment, monitoring, and remediation
Evaluation	decision materiality, model risk, third-party dependence, human override, stakeholder effects, and control effectiveness
Primary risk	assuming technical validation alone satisfies fiduciary and stakeholder obligations

1. Introduction

Research on corporate governance for AI-enabled decision systems sits at the boundary between a promising component and a consequential decision. The core question is how boards and executives can oversee material algorithmic decisions without reducing oversight to generic principles. Answering it requires more than assembling favorable results. It requires a chain of evidence that makes the problem boundary, mechanism, measurement, comparator, uncertainty, and accountable decision visible to the reader.

This article presents a structured narrative review of ten related publications. Sources were selected for topical relevance, traceable publication metadata, and complementary coverage of technical, empirical, and governance questions. No meta-analytic pooling is attempted because the included studies differ in designs, populations, system boundaries, and outcomes. The purpose is decision-oriented synthesis: to identify what each source can support, where transfer remains uncertain, and which evidence would justify the next stage of use.

The central risk is assuming technical validation alone satisfies fiduciary and stakeholder obligations. A top-line metric or broad policy label can appear decisive while concealing the conditions that produced it. Accordingly, the synthesis separates observation from interpretation and implementation from validation. This separation is especially important when the same system creates benefits for one user or institution while transferring cost, risk, or administrative burden to another.

2. Evidence Base and Problem Boundary

A defensible review begins by defining the unit of analysis. For the present topic, the relevant boundary includes inputs, institutional or technical context, the mechanism producing an output, the person or system acting on that output, and the downstream outcome. Evidence falls outside the boundary when it measures only an intermediate proxy yet is used to claim an end-to-end benefit.

Chowdhury et al. (2022) addresses "Unlocking the value of artificial intelligence in human resource management through AI capability framework." In this synthesis, that work is used as a source on problem definition within corporate governance for AI-enabled decision systems. The connection is deliberately bounded: the cited study informs the evidence map, but it does not by itself establish readiness for firms using AI in finance, employment, marketing, and operations. That distinction keeps the review close to the published record and prevents an adjacent result from being converted into a broader causal claim.

Novelli et al. (2023) addresses "Accountability in artificial intelligence: what it is and how it works." In this synthesis, that work is used as a source on conceptual boundary within corporate governance for AI-enabled decision systems. The connection is deliberately bounded: the cited study informs the evidence map, but it does not by itself establish readiness for firms using AI in finance, employment, marketing, and operations. That distinction keeps the review close to the published record and prevents an adjacent result from being converted into a broader causal claim.

Hilb (2020) addresses "Toward artificial governance? The role of artificial intelligence in shaping the future of corporate governance." In this synthesis, that work is used as a source on measurement within corporate governance for AI-enabled decision systems. The connection is deliberately bounded: the cited study informs the evidence map, but it does not by itself establish readiness for firms using AI in finance, employment, marketing, and operations. That distinction keeps the review close to the published record and prevents an adjacent result from being converted into a broader causal claim.

Together, these works provide complementary entry points, but their conclusions should not be flattened into a single ranking. Differences in data generation, setting, temporal horizon, and outcome definition may be substantively meaningful. A review should therefore preserve those differences and state where analogy, rather than direct replication, connects one domain to another.

3. Mechanism and System Design

The operative mechanism is the chain from strategic purpose and data sourcing through model ownership, deployment, monitoring, and remediation. This formulation discourages component-only reasoning. A model, platform, policy, assay, or infrastructure service acquires practical meaning only when its interfaces and use conditions are specified. The evidence chain should describe what information is available, what transformation occurs, which assumptions make that transformation credible, and who is authorized to act.

Camilleri (2023) addresses "Artificial intelligence governance: Ethical considerations and implications for social responsibility." In this synthesis, that work is used as a source on system mechanism within corporate governance for AI-enabled decision systems. The connection is deliberately bounded: the cited study informs the evidence map, but it does not by itself establish readiness for firms using AI in finance, employment, marketing, and operations. That distinction keeps the review close to the published record and prevents an adjacent result from being converted into a broader causal claim.

Rodgers et al. (2022) addresses "An artificial intelligence algorithmic approach to ethical decision-making in human resource management processes." In this synthesis, that work is used as a source on representation choice within corporate governance for AI-enabled decision systems. The connection is deliberately bounded: the cited study informs the evidence map, but it does not by itself establish readiness for firms using AI in finance, employment, marketing, and operations. That distinction keeps the review close to the published record and prevents an adjacent result from being converted into a broader causal claim.

Mariani and Dwivedi (2024) addresses "Generative artificial intelligence in innovation management: A preview of future research developments." In this synthesis, that work is used as a source on failure analysis within corporate governance for AI-enabled decision systems. The connection is deliberately bounded: the cited study informs the evidence map, but it does not by itself establish readiness for firms using AI in finance, employment, marketing, and operations. That distinction keeps the review close to the published record and prevents an adjacent result from being converted into a broader causal claim.

These studies also show why modular claims require interface tests. A component may perform well while the complete pathway fails because inputs drift, users interpret outputs differently, recovery semantics are ambiguous, or institutional incentives change. Design documentation should therefore include not only the intended pathway but also foreseeable deviations, degraded modes, and conditions under which the system should stop or defer.

4. Evaluation and Uncertainty

Evaluation should center on decision materiality, model risk, third-party dependence, human override, stakeholder effects, and control effectiveness. Average performance remains useful, but it should be accompanied by distributions, error categories, relevant subgroups or operating regimes, and uncertainty at the point where a decision changes. Comparators must isolate the value of the proposed addition rather than compare a mature intervention with an intentionally weak baseline.

Fenwick and Vermeulen (2018) addresses "Technology and Corporate Governance: Blockchain, Crypto, and Artificial Intelligence." In this synthesis, that work is used as a source on comparative evaluation within corporate governance for AI-enabled decision systems. The connection is deliberately bounded: the cited study informs the evidence map, but it does not by itself establish readiness for firms using AI in finance, employment, marketing, and operations. That distinction keeps the review close to the published record and prevents an adjacent result from being converted into a broader causal claim.

Elliott et al. (2021) addresses "Towards an Equitable Digital Society: Artificial Intelligence (AI) and Corporate Digital Responsibility (CDR)." In this synthesis, that work is used as a source on uncertainty within corporate governance for AI-enabled decision systems. The connection is deliberately bounded: the cited study informs the evidence map, but it does not by itself establish readiness for firms using AI in finance, employment, marketing, and operations. That distinction keeps the review close to the published record and prevents an adjacent result from being converted into a broader causal claim.

External validation should introduce meaningful shifts in setting, time, population, workload, market structure, or environmental conditions. A nominally independent sample can still be too similar to test transfer. Conversely, a failed transfer is scientifically informative when the changed conditions are measured and the mechanism of failure is examined rather than hidden in an aggregate score.

5. Translation and Governance

Translation to firms using AI in finance, employment, marketing, and operations depends on more than technical readiness. Decision rights, documentation, maintenance, monitoring, appeal, and recovery are part of the intervention. The governance requirement is clear accountability, board information rights, independent challenge, audit evidence, and incident escalation. Each control should be connected to a named failure mode and should identify an owner, evidence source, review interval, and escalation path.

Dunleavy and Margetts (2023) addresses "Data science, artificial intelligence and the third wave of digital era governance." In this synthesis, that work is used as a source on implementation within corporate governance for AI-enabled decision systems. The connection is deliberately bounded: the cited study informs the evidence map, but it does not by itself establish readiness for firms using AI in finance, employment, marketing, and operations. That distinction keeps the review close to the published record and prevents an adjacent result from being converted into a broader causal claim.

Katzenbach and Ulbricht (2019) addresses "Algorithmic governance." In this synthesis, that work is used as a source on governance within corporate governance for AI-enabled decision systems. The connection is deliberately bounded: the cited study informs the evidence map, but it does not by itself establish readiness for firms using AI in finance, employment, marketing, and operations. That distinction keeps the review close to the published record and prevents an adjacent result from being converted into a broader causal claim.

A credible implementation plan also defines sunset and revalidation conditions. New data, software updates, institutional restructuring, population change, or shifts in incentives can invalidate previously reasonable assumptions. Monitoring should therefore test the validity of the evidence chain, not merely whether a service remains online or a headline metric remains stable.

6. Research Agenda

Future studies should preregister or otherwise predefine the intended decision, principal outcomes, exclusions, and analysis plan when feasible. They should report missingness, sensitivity analyses, negative or null findings, and the cost of errors to differently situated stakeholders. Reproducible artifacts should include sufficient metadata to reconstruct data provenance, model or analytical versions, parameter choices, and the sequence from evidence to conclusion.

Cross-disciplinary work needs a shared object of explanation rather than a collection of adjacent vocabularies. For corporate governance for AI-enabled decision systems, that shared object is the complete decision pathway. Technical, clinical, social, environmental, or economic expertise should each change the design or interpretation of the study. When evidence remains incomplete, the useful output is a bounded hypothesis, a transparent uncertainty statement, or a request for additional measurement.

7. Conclusion

The literature supports a disciplined approach to corporate governance for AI-enabled decision systems: define the decision, preserve system boundaries, test consequential failure modes, connect uncertainty to action, and assign accountable control. The strongest conclusion is not that one method is universally superior. It is that credible use in firms using AI in finance, employment, marketing, and operations requires an auditable link from source evidence to design choice, evaluation result, and governed decision.

References

Camilleri, M. A. (2023). Artificial intelligence governance: Ethical considerations and implications for social responsibility. Expert Systems, 41(7). https://doi.org/10.1111/exsy.13406

Chowdhury, S., Dey, P., Joel-Edgar, S., Bhattacharya, S., Rodriguez-Espindola, O., Abadie, A., & Truong, L. (2022). Unlocking the value of artificial intelligence in human resource management through AI capability framework. Human Resource Management Review, 33(1), 100899. https://doi.org/10.1016/j.hrmr.2022.100899

Dunleavy, P., & Margetts, H. (2023). Data science, artificial intelligence and the third wave of digital era governance. Public Policy and Administration, 40(2), 185-214. https://doi.org/10.1177/09520767231198737

Elliott, K., Price, R., Shaw, P., Spiliotopoulos, T., Ng, M., Coopamootoo, K., & Moorsel, A. V. (2021). Towards an Equitable Digital Society: Artificial Intelligence (AI) and Corporate Digital Responsibility (CDR). Society, 58(3), 179-188. https://doi.org/10.1007/s12115-021-00594-8

Fenwick, M., & Vermeulen, E. P. M. (2018). Technology and Corporate Governance: Blockchain, Crypto, and Artificial Intelligence. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3263222

Hilb, M. (2020). Toward artificial governance? The role of artificial intelligence in shaping the future of corporate governance. Journal of Management & Governance, 24(4), 851-870. https://doi.org/10.1007/s10997-020-09519-9

Katzenbach, C., & Ulbricht, L. (2019). Algorithmic governance. Internet Policy Review, 8(4). https://doi.org/10.14763/2019.4.1424

Mariani, M., & Dwivedi, Y. K. (2024). Generative artificial intelligence in innovation management: A preview of future research developments. Journal of Business Research, 175, 114542. https://doi.org/10.1016/j.jbusres.2024.114542

Novelli, C., Taddeo, M., & Floridi, L. (2023). Accountability in artificial intelligence: what it is and how it works. AI & Society, 39(4), 1871-1882. https://doi.org/10.1007/s00146-023-01635-y

Rodgers, W., Murray, J. M., Stefanidis, A., Degbey, W. Y., & Tarba, S. Y. (2022). An artificial intelligence algorithmic approach to ethical decision-making in human resource management processes. Human Resource Management Review, 33(1), 100925. https://doi.org/10.1016/j.hrmr.2022.100925