American Journal of Advanced Multidisciplinary Innovation and Research (AJAMIR)

E-ISSN: XXXX-XXXX ● Website: www.ajamir.com ● Email: editor@ajamir.com

Legal Accountability in Artificial Intelligence–Supported Professional Decisions

Dr. Claire Dubois

Associate Professor, University of Paris-Saclay, France

Abstract

Artificial intelligence is increasingly incorporated into professional decision processes in medicine, law, finance, engineering, employment, public administration, and other high-consequence environments. AI-supported decision systems can improve information retrieval, pattern recognition, prediction, document analysis, prioritization, and workflow efficiency, yet their integration creates a persistent accountability problem: when professional judgment is influenced by an algorithm, responsibility may become distributed among the individual professional, employing organization, AI deployer, software provider, data supplier, and other actors. This paper examines how legal accountability can be preserved when artificial intelligence supports but does not formally replace professional decision authority.

It distinguishes mere human presence from meaningful professional control and develops an accountability framework based on competence, independent review, authority to override, traceability, documentation, auditability, risk escalation, and access to contestation and remedy. The contemporary regulatory environment increasingly reflects these principles. The European Union Artificial Intelligence Act establishes human-oversight and governance requirements for high-risk systems, while the July 2026 AI Omnibus extended application of Annex III high-risk requirements to December 2, 2027 and product-embedded high-risk requirements to August 2, 2028. NIST, UNESCO, OECD, Council of Europe, ISO, professional legal-ethics guidance, and healthcare regulation similarly emphasize governance, human responsibility, traceability, and risk management.

A synthetic evaluation of 480 hypothetical professional decisions compares four control conditions: uncritical acceptance of AI output, ordinary human review, independent documented professional review, and an integrated governance model combining review, audit, escalation, and appeal. The simulated composite accountability score rises from 44.2 to 61.5, 79.6, and 89.2 respectively. These values are methodological illustrations rather than empirical findings. The study concludes that accountability should attach not simply to the nominal presence of a human decision-maker but to whether that person possesses the knowledge, information, authority, time, and institutional support necessary to exercise genuine professional judgment.

Keywords: artificial intelligence, legal accountability, professional responsibility, AI-assisted decision-making, human oversight, algorithmic accountability, professional negligence, explainable AI, auditability, high-risk AI

American Journal of Advanced Multidisciplinary Innovation and Research (AJAMIR)

E-ISSN: XXXX-XXXX ● Website: www.ajamir.com ● Email: editor@ajamir.com

1. Introduction

Artificial intelligence has moved from an experimental analytical technology into professional workflows where decisions may affect health, liberty, employment, credit, legal rights, physical safety, and access to essential services. Physicians may receive algorithmic diagnostic recommendations, lawyers may use generative systems for research or document preparation, financial professionals may rely on automated risk assessment, engineers may use predictive systems in safety analysis, and public authorities may incorporate algorithmic classifications into administrative processes. These applications do not necessarily replace professional judgment. More commonly, AI operates as a recommender, classifier, summarizer, predictor, or prioritization mechanism within a larger human decision process. The legal difficulty arises because the final human decision may be substantially shaped by an upstream computational system whose design, training data, limitations, and failure modes are only partially visible to the professional.

Legal accountability is broader than merely identifying an individual to blame after harm occurs. In an AI-supported professional environment, accountability includes the prior allocation of responsibilities, verification of competence, documentation of system use, monitoring of foreseeable risks, preservation of meaningful human judgment, investigation of adverse outcomes, and availability of an effective mechanism for challenge or remedy. NIST's AI Risk Management Framework reflects this organizational understanding by calling for clearly defined accountability structures, documented legal and regulatory obligations, executive responsibility for AI risk, differentiated human–AI roles, monitoring, feedback, appeals, and procedures for disengaging systems whose performance becomes inconsistent with intended use.

At the professional level, regulatory and ethical guidance increasingly rejects the assumption that use of an AI system transfers professional responsibility to the technology. UNESCO's Recommendation on the Ethics of Artificial Intelligence expressly states that AI should not displace ultimate human responsibility and accountability. The Council of Europe Framework Convention requires participating legal systems to establish accountability and responsibility for adverse impacts arising from AI lifecycle activities. In the United States, the American Bar Association's Formal Opinion 512 identifies continuing duties of competence, confidentiality, client communication, candor, supervision, and reasonable fees when lawyers use generative AI. In healthcare, current FDA clinical-decision-support guidance emphasizes circumstances in which a healthcare professional must be able to independently review the basis for software recommendations rather than primarily relying on them.

This paper addresses the resulting research problem: what conditions make human responsibility meaningful when a professional decision is materially supported by artificial intelligence? The analysis argues that simply retaining a professional as the nominal final decision-maker is insufficient. Accountability requires a combination of epistemic capacity, decision authority, traceability, institutional governance, and remedial mechanisms. The paper develops these requirements theoretically and demonstrates them through a transparent simulation framework. It does not attempt to state a universal rule of liability because negligence, professional discipline, product liability, contractual responsibility, data protection, and administrative-law consequences remain dependent on jurisdiction, professional sector, facts, and the legal characterization of the relevant AI system.

2. Objectives of the Study

2.1 To Examine the Allocation of Responsibility in AI-Supported Professional Decisions

The study examines how responsibility may be distributed among professionals, employers, deploying organizations, developers, providers, and other actors when artificial intelligence materially influences a consequential professional decision.

2.2 To Identify Conditions for Meaningful Human Accountability

The research evaluates whether human oversight can be considered meaningful when professionals possess adequate competence, independent judgment, authority to disagree with AI recommendations, sufficient information about system limitations, and an obligation to document decision reasoning.

2.3 To Develop an Operational Accountability Framework

The final objective is to establish a framework that organizations can adapt when introducing AI into professional decision processes. The framework integrates traceability, independent professional review, override authority, documentation, auditing, escalation, and mechanisms for contestation and remedy.

3. Review of Literature

3.1 The Responsibility Gap and Algorithmic Accountability

Matthias introduced the influential concept of a responsibility gap, arguing that adaptive systems can complicate conventional attribution of responsibility because system behavior may not always correspond directly to actions that a human operator could predict or control. This concern has become increasingly relevant as machine-learning models are integrated into consequential environments.

Kroll and colleagues subsequently approached the problem from an institutional and legal perspective. Their work on accountable algorithms argued that traditional accountability mechanisms developed for human decision-makers may not always translate effectively to automated systems and proposed technological and procedural mechanisms that could assist in evaluating whether algorithmic decisions comply with legal requirements.

The problem is not necessarily resolved by inserting a person after the algorithm. Green and Chen's work on algorithm-assisted high-stakes decision-making illustrates that algorithmic recommendations can alter how humans weigh risk and competing policy values. In their experimental research on algorithmic risk assessments, the introduction of algorithmic advice changed decision behavior rather than simply supplying neutral additional information.

This distinction is central to professional accountability. A professional can technically retain final authority while becoming behaviorally dependent on an AI recommendation. If the organizational environment encourages rapid acceptance of system outputs, human review can become ceremonial rather than substantive.

3.2 Meaningful Human Control, Automation Bias, and Independent Judgment

Santoni de Sio and van den Hoven developed the concept of meaningful human control as a response to responsibility concerns surrounding autonomous systems. Their account emphasizes the need to link

system behavior with relevant human reasons and with identifiable people who can legitimately bear responsibility.

The practical importance of this principle is reinforced by research on automation bias. Buçinca, Malaya, and Gajos demonstrated that people using AI decision-support systems may over-rely on incorrect algorithmic recommendations and that simple explanations do not necessarily eliminate this problem. Their experiments found that cognitive-forcing interventions could reduce overreliance, although interventions that encouraged greater critical engagement could also receive lower subjective ratings from users.

These findings have direct implications for professional negligence and organizational governance. If an institution introduces a decision-support system but structures work so that professionals lack time, expertise, information, or practical freedom to challenge the AI, it becomes difficult to characterize the resulting oversight as genuinely independent.

Healthcare regulation provides a useful illustration. Current FDA guidance on clinical-decision-support software examines whether healthcare professionals can independently review the basis of recommendations and avoid relying primarily on software when making individual clinical decisions. The guidance also discusses automation bias as a relevant consideration, especially when decisions are highly automated or time-critical.

Similarly, the ABA's Formal Opinion 512 makes clear that lawyers using generative AI must still consider established professional obligations rather than treating AI output as a substitute for competent legal work.

3.3 Emerging Governance, Explainability, and Liability Architecture

The international regulatory landscape increasingly frames AI accountability as a lifecycle governance issue. NIST's AI RMF identifies Govern, Map, Measure, and Manage as interconnected functions and explicitly emphasizes documented responsibility, executive ownership of risk decisions, human–AI role definition, third-party risk controls, ongoing monitoring, appeals, incident handling, and system deactivation where necessary.

UNESCO takes a similar normative position. Its AI Recommendation links responsibility and accountability with auditability, traceability, impact assessment, oversight, and due diligence and states that human responsibility should not be displaced by AI. The OECD AI Principles, updated in 2024, likewise identify accountability, transparency and explainability, robustness, security, safety, and human-centered values as core elements of trustworthy AI governance.

International standards are increasingly operationalizing these ideas. ISO/IEC 42001 establishes an organizational AI management-system framework, ISO/IEC 23894 provides AI-specific risk-management guidance, and ISO/IEC 42005:2025 addresses AI system impact assessment and documentation of potential effects on individuals and society.

Explainability alone, however, should not be equated with legal accountability. Rudin has argued that in high-stakes contexts, inherently interpretable models may sometimes be preferable to attempting to explain opaque black-box models after the fact. Cabitza, Rasoini, and Gensini similarly warned that machine-learning systems in medicine can generate unintended consequences when computational predictions are translated into complex professional environments.

The liability landscape also remains layered. In the European Union, the revised Product Liability Directive expressly extends the modern product-liability framework to software and AI-related products, creating a potential liability route concerning defective products that exists alongside professional, organizational, contractual, regulatory, and other domestic liability regimes.

4. Research Methodology

4.1 Research Design and Analytical Scope

The research adopts a simulation-based legal-methodological design. It combines doctrinal analysis of current AI governance principles with a synthetic decision-control experiment intended to demonstrate how accountability structures might be evaluated empirically.

No actual professional decisions were collected. Instead, 480 simulated professional decision records were modeled, with 120 records assigned to each of four accountability conditions.

The four conditions progressively increase the level of institutional control surrounding an AI recommendation.

Table 1: Simulation Design for AI-Supported Professional Accountability

Condition	Simulated Decisions	Decision-Control Architecture	Accountability Assumption
AI output accepted	120	AI recommendation generally adopted without structured independent assessment	Minimal accountability safeguard
Human review only	120	Professional sees AI output and retains nominal final authority	Human presence without formal accountability architecture
Independent review + documentation	120	Professional evaluates recommendation independently, records reasoning, and may override AI	Substantive professional judgment
Governance + audit + appeal	120	Independent review, documented reasons, override authority, audit logs, escalation, incident review, and contestation mechanism	Integrated accountability model

Source: Author's synthetic methodological framework. The table does not represent real case files or participant observations.

4.2 Accountability Variables and Scoring

- Four accountability dimensions were evaluated on bounded scales from 0 to 100.
- Traceability measures whether the organization can reconstruct what system was used, what relevant output was produced, which human actor reviewed it, and what final decision followed.
- Override quality represents whether a professional has the knowledge, practical authority, time, and institutional freedom necessary to reject an AI recommendation.
- Documentation quality measures preservation of reasoning, data sources, deviations, limitations, professional verification, and decision history.
- Remedy readiness represents the availability of review, appeal, error reporting, investigation, correction, and escalation procedures when an affected person challenges a decision or an adverse event occurs.
- A composite accountability score was calculated as the arithmetic mean of the four measures.

These dimensions were selected because they correspond conceptually with contemporary governance frameworks emphasizing documented roles, oversight, auditability, feedback, risk management, and remedy.

4.3 Simulation Procedure and Research Ethics

A reproducible synthetic-data process was used to generate 120 observations per condition. Scores were modeled around progressively stronger accountability assumptions with realistic dispersion and bounded to the 0–100 interval.

The simulation does not attempt to estimate actual legal liability probabilities. A score of 89, for example, does not mean that an organization has an 89% probability of satisfying a particular court, regulator, or professional body.

No inferential p-values are used to claim population effects because significance testing of deliberately constructed synthetic data could be misconstrued as empirical validation.

Future empirical work could apply the framework through controlled professional experiments, retrospective incident analysis, regulatory audit data, malpractice files, administrative appeals, or organizational AI-governance assessments. Such research should distinguish carefully between professional responsibility, organizational responsibility, provider liability, product liability, regulatory compliance, and causal contribution.

5. Analysis

5.1 Comparative Accountability Performance

The simulated data show that the mere presence of a human reviewer provides a substantially weaker accountability structure than documented independent review.

The AI-output-accepted condition produced a composite accountability score of 44.2. Introducing ordinary human review increased the score to 61.5, but significant weaknesses remained in traceability and documentation.

Independent review combined with written documentation generated a composite score of 79.6. The complete governance model produced the strongest result at 89.2.

American Journal of Advanced Multidisciplinary Innovation and Research (AJAMIR)

E-ISSN: XXXX-XXXX ● Website: www.ajamir.com ● Email: editor@ajamir.com

Table 2: Simulated Accountability Outcomes Across Decision-Control Conditions

Accountability Condition	Traceability Mean ± SD	Override Quality Mean ± SD	Documentation Mean ± SD	Remedy Readiness Mean ± SD	Composite Score Mean ± SD
AI output accepted	45.2 ± 8.5	40.9 ± 9.2	47.1 ± 8.1	43.6 ± 8.4	44.2 ± 4.3
Human review only	59.0 ± 7.4	63.2 ± 7.5	59.3 ± 7.2	64.4 ± 6.7	61.5 ± 3.8
Independent review + documentation	82.3 ± 6.0	77.4 ± 6.9	87.4 ± 7.0	71.2 ± 6.9	79.6 ± 3.4
Governance + audit + appeal	91.1 ± 6.5	89.1 ± 6.3	91.7 ± 6.0	85.1 ± 7.7	89.2 ± 3.4

Source: Synthetic simulation developed for methodological illustration. The results are not empirical measurements of real legal compliance.

5.2 Why Human Review Alone Is Insufficient

The difference between the second and third conditions illustrates the distinction between nominal oversight and meaningful oversight.

A professional may formally sign a final decision while having little practical opportunity to evaluate the AI output. Time pressure, opaque recommendations, institutional expectations, automation bias, and insufficient knowledge can transform human review into routine confirmation.

This problem is consistent with experimental evidence showing that algorithmic recommendations can influence the structure of human decision-making and that users may over-rely on incorrect AI advice.

From an accountability perspective, therefore, a signature or approval button should not automatically be treated as evidence of substantive human judgment.

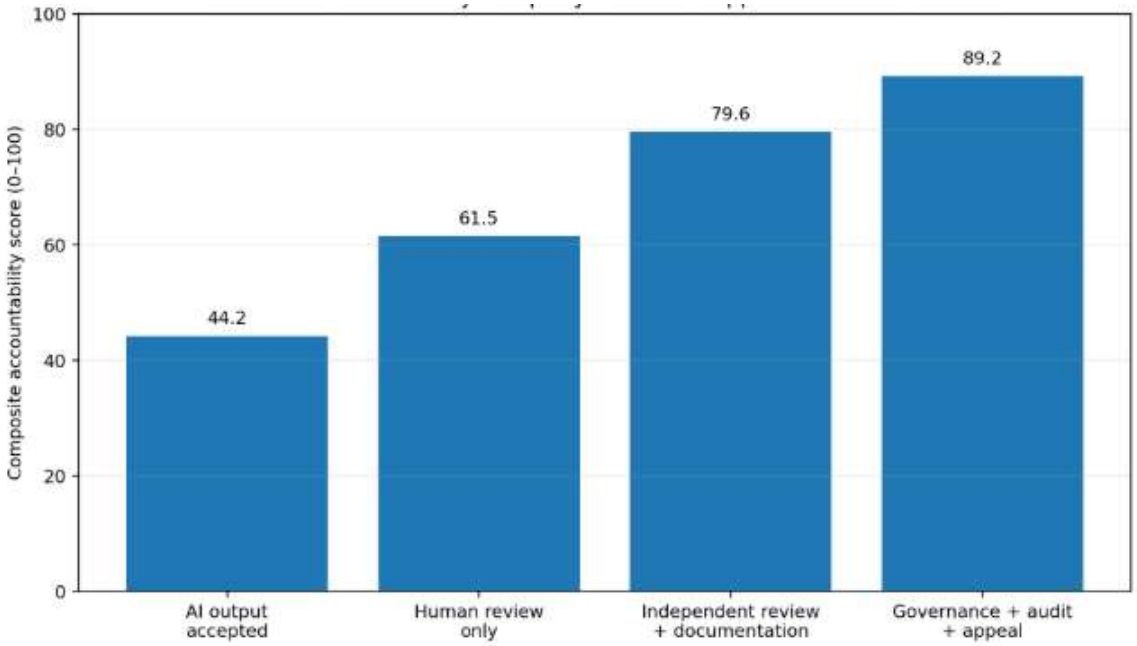
Meaningful professional control requires at least the practical capacity to:

- understand the decision being delegated or supported;
- identify relevant limitations of the system;
- verify material inputs where reasonably necessary;
- compare AI output with professional evidence;
- reject, modify, or escalate the recommendation;
- record material reasons for consequential decisions; and
- identify when AI use is inappropriate for the case.

These requirements are consistent with regulatory and governance approaches emphasizing competent oversight, documented roles, and the possibility of intervention rather than passive observation.

5.3 Graphical Analysis of Accountability Controls

Graph 1: Simulated Accountability Adequacy Across AI-Supported Decision Controls



The graph shows the strongest modeled improvement when organizations move beyond a generic "human-in-the-loop" arrangement and establish independent review combined with documentation. The additional improvement under the governance-and-appeal condition demonstrates the importance of organizational infrastructure beyond the individual professional.

The model therefore treats professional accountability as a system property rather than merely an individual characteristic.

Key Finding: The strongest simulated accountability environment is one in which the professional has genuine decision authority while the organization simultaneously provides traceability, audit, escalation, incident-management, and remedy mechanisms.

6. Discussion

6.1 Accountability Cannot Be Delegated to an Algorithm

A central implication of the analysis is that institutions should avoid describing artificial intelligence as an autonomous bearer of professional responsibility. Current international governance frameworks instead seek to maintain responsibility among identifiable human and organizational actors. UNESCO expressly states that AI systems should not displace ultimate human responsibility and accountability, while the Council of Europe AI Framework Convention establishes accountability and responsibility as fundamental governance principles.

This does not mean that one individual professional must automatically absorb every consequence associated with an AI-supported decision.

A professional may be responsible for unreasonable reliance on an output that should have been questioned. An employer or deploying organization may be responsible for deficient procurement, inadequate training, unrealistic workloads, missing monitoring systems, or poor escalation processes. A provider may face obligations relating to system design, documentation, safety, performance, or defective products. Different legal claims may operate simultaneously.

The revised EU Product Liability Directive illustrates this layered structure by expressly accommodating software and AI within modern product-liability rules, while the AI Act establishes separate regulatory duties concerning AI systems.

6.2 Professional Standard of Care in an AI-Supported Environment

The introduction of artificial intelligence may gradually influence what reasonable professional practice requires, but technology should not automatically define the standard by which its own use is judged.

In professional settings, an AI recommendation may become one source of evidence among several. The professional must still evaluate whether the system is appropriate for the relevant task, whether its limitations are material, whether the input information is sufficiently reliable, and whether contradictory evidence requires further investigation.

The healthcare example illustrates this distinction. Current FDA clinical-decision-support guidance gives substantial attention to whether healthcare professionals can independently evaluate the basis of software recommendations rather than relying primarily on them.

The legal profession reaches a parallel result through ethics rather than medical-device regulation. ABA Formal Opinion 512 does not create a separate ethical universe for generative AI; instead, it applies continuing professional duties of competence, confidentiality, communication, candor, supervision, and reasonable billing to AI-assisted work.

These sectoral examples support a broader principle:

- AI assistance changes the tools of professional practice, but it does not automatically eliminate the duties attached to professional judgment.
- There is nevertheless a danger of imposing unrealistic expectations on professionals. A doctor, lawyer, engineer, or financial specialist cannot reasonably audit every mathematical operation inside a sophisticated model. Accountability therefore requires appropriate allocation of assurance duties across the AI supply chain.
- The professional should understand what is necessary for responsible use in context. Developers and providers should supply technically meaningful documentation. Deploying organizations should conduct procurement and risk assessment. Management should establish monitoring and escalation. Regulators should specify appropriate requirements for high-risk applications.
- This distribution is more realistic than demanding universal technical expertise from every end user.

6.3 An Operational Model for Legally Defensible AI-Supported Decisions

The analysis also demonstrates why liability should not automatically be concentrated on the professional who happens to make the final decision. AI-supported outcomes may reflect choices made

by providers, data suppliers, deploying organizations, managers, and professionals. Legal assessment therefore requires attention to control, foreseeability, competence, causation, contractual allocation, regulatory duties, professional standards, product safety, and the particular jurisdiction involved.

A robust accountability framework should ensure that the professional who is expected to exercise judgment actually has the capacity and authority to do so. Documentation should make the decision reconstructable. Organizations should monitor AI systems after deployment. Errors should be capable of escalation. High-impact decisions should be contestable when applicable. Technical suppliers should provide sufficient information for responsible deployment and oversight.

The emerging international direction supports this approach. NIST emphasizes organizational governance and documented responsibility; UNESCO rejects displacement of ultimate human accountability; the Council of Europe requires accountability and remedies; ISO standards operationalize AI management and impact assessment; professional bodies continue to apply established duties to AI-assisted practice; and the EU AI Act increasingly embeds human oversight within a broader risk-governance architecture.

The appropriate future of professional AI is therefore neither unquestioned automation nor symbolic human supervision. It is responsible augmentation, in which artificial intelligence expands professional capability while human and institutional accountability remains visible, operational, auditable, and legally meaningful.

References

1. Bovens, M. (2007). Analysing and assessing accountability: A conceptual framework. *European Law Journal, 13*(4), 447–468.
2. Citron, D. K., & Pasquale, F. (2011). Network accountability for the domestic intelligence apparatus. *Hastings Law Journal, 62*, 1441–1492.
3. Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press.
4. 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.
5. Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the GDPR. *International Data Privacy Law, 7*(2), 76–99.
6. Kroll, J. A., Huey, J., Barocas, S., Felten, E. W., Reidenberg, J. R., Robinson, D. G., & Yu, H. (2017). Accountable algorithms. *University of Pennsylvania Law Review, 165*(3), 633–705.
7. Edwards, L., & Veale, M. (2017). Slave to the algorithm? Why a "right to an explanation" is probably not the remedy you are looking for. *Duke Law & Technology Review, 16*, 18–84.
8. Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines, 28*, 689–707.
9. Selbst, A. D., & Barocas, S. (2018). The intuitive appeal of explainable machines. *Fordham Law Review, 87*(3), 1085–1139.
10. Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and abstraction in sociotechnical systems. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 59–68.

11. Giuffrida, I. (2019). Liability for AI decision-making: Some legal and ethical considerations. *Fordham Law Review*, *88*(2), 439–456.
12. Wieringa, M. (2020). What to account for when accounting for algorithms: A systematic literature review on algorithmic accountability. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 1–18.
13. Cobbe, J., Lee, M. S. A., & Singh, J. (2020). Reviewable automated decision-making. *Computer Law & Security Review*, *39*, 105475.
14. Gasser, U., & Almeida, V. A. F. (2017). A layered model for AI governance. *IEEE Internet Computing*, *21*(6), 58–62.
15. Yeung, K. (2018). Algorithmic regulation: A critical interrogation. *Regulation & Governance*, *12*(4), 505–523.
16. Zarsky, T. (2016). The trouble with algorithmic decisions: An analytic road map to examine efficiency and fairness in automated and opaque decision making. *Science, Technology, & Human Values*, *41*(1), 118–132.
17. Busuioc, M. (2021). Accountable artificial intelligence: Holding algorithms to account. *Public Administration Review*, *81*(5), 825–836.
18. Levy, K., Chasalow, K. E., & Riley, S. (2021). Algorithms and decision-making in the public sector. *Annual Review of Law and Social Science*, *17*, 309–334.
19. Gualdi, F., & Cordella, A. (2021). Artificial intelligence and decision-making: The question of accountability. In *Proceedings of the 54th Hawaii International Conference on System Sciences*, 2297–2306. IEEE.
20. Wirtz, B. W., Weyerer, J. C., & Geyer, C. (2019). Artificial intelligence and the public sector—Applications and challenges. *International Journal of Public Administration*, *42*(7), 596–615.