ARTIFICIAL INTELLIGENCE AND HUMAN DECISION-MAKING: A MULTIDISCIPLINARY ANALYSIS OF OPPORTUNITIES, RISKS AND FUTURE GOVERNANCE

Vlad Daniel VALERIUS, Associate Professor Ph.D.,
Bogdan Vodă University of Cluj-Napoca

ABSTRACT

This paper provides an integrated multidisciplinary analysis of the impact of artificial intelligence (AI) on human decision-making. The argument is supported by examples from the Romanian regional ecosystem (e.g., the investment in the Artificial Intelligence Research Institute - AIRi in Cluj-Napoca and the Smart City Strategy in Oradea), illustrating the disparities in adoption and institutional capacity. The research adopts the methodology of systematic literature review (according to the PRISMA guideline) and analyzes the transition from classical decision theories (based on bounded rationality and human cognitive errors) to hybrid human-machine decision-making models. The paper assesses the major opportunities of algorithmic augmentation (accuracy, speed, predictive capacity, personalization) in parallel with the associated systemic, ethical and legal risks (algorithmic bias, opacity of "black box" models, dilution of legal liability, data opacity and cognitive dependence). From a multidisciplinary perspective – legal, economic, managerial, psychological, sociological and sports sciences – it is demonstrated that AI should not replace the human decision-maker, but complement him in a rigorous governance framework. As an original contribution, the author proposes the Integrated Governance Model of AI-Assisted Decision (AIGF), structured on six interconnected pillars, and formulates concrete public policy recommendations for governments, universities, the private sector and the judiciary.

KEYWORDS: *Artificial Intelligence, Decision Making, Human Decision, Governance, Ethics, Algorithmic Decision, Trust, Public Administration, Law, Leadership, Behavior.*

J.E.L. Classifications: L86, O33, D81, D 83, D91, H83, K24, Z13

1. INTRODUCTION

1.1. Artificial Intelligence as a New Decision-Making Environment

The contemporary decision-maker no longer acts in isolation, relying exclusively on intuition or manual data analysis. The evolution of IT brings to mind the words of the columnist "It is not the

times that are beneath man, but the poor man beneath the times". Machine learning algorithms, large-scale linguistic models and Big Data analysis systems have become a component part of any relevant institutional choice.

Two regional indicators confirm the pace of this transformation.

In Cluj-Napoca, the Technical University is building, through an investment of over 105 million lei supported by its own revenues, the Artificial Intelligence Research Institute (AIRi) — the first research center built from scratch in post-December Romania, integrated into the Romanian Artificial Intelligence Hub (HRIA) and scheduled for inauguration in spring 2026.

In Oradea, the City Hall launched in September 2025 the consultations for the city's Smart City Strategy, financed by a European grant of approximately 104,770 euros through the Digital Europe and PoCIDIF programs, with the declared objective of integrating "the advantages of the technological leap, including in the AI area" into local public services — completing the already functional digital.oradea.ro platform, developed to simplify administrative procedures.

The difference in scale between the two projects — a national research institute and an incipient municipal strategy — illustrates exactly the tension that we will discuss in section 7.2: AI adoption is not progressing uniformly, but on uneven territorial layers.

The OECD, UNESCO, and World Economic Forum reports, along with the analyses published in the Stanford AI Index, converge towards the same finding: modern institutional governance can no longer be thought of independently of the analytical capabilities of artificial intelligence.

1.2. Research problem

The specialized literature remains fragmented by field. Lawyers analyze the impact of AI on law, doctors — the applicability of algorithms in diagnosis, and management specialists — the optimization of economic decisions. Rarely do these perspectives converge in a single model, capable of reconciling the legal, economic, psychological, and socio-organizational dimensions of AI-assisted decision-making. The gap is not just theoretical — it is felt practically, when a local institution must simultaneously decide on the legal compliance, implementation costs and social acceptance of an automated system.

1.3. Objectives

In the paper, we dare to propose four specific objectives:

1. identifying and evaluating the analytical opportunities offered by AI;
2. mapping the systemic and ethical risks associated with algorithmic implementation;

3. comparing classical decision theories with emerging hybrid decision-making models;

4. proposing a governance framework adapted to public and private institutions.

1.4. The didactic principle "a good question provokes a good answer" led us to pay attention to research questions (Research Question RQ).

RQ1: How does AI modify the fundamental mechanisms of the human decision-making process?

RQ2: What are the strategic opportunities emerging from the integration of AI systems?

RQ3: What systemic, ethical, and legal risks affect the quality of technology-assisted decision-making?

RQ4: What governance models can balance innovation with human control?

RQ5: What are the main research gaps in the current multidisciplinary literature?

2. METHODOLOGY

The research adopts a systematic literature review, structured according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines.

2.1. Database selection

The query was conducted in Scopus, Web of Science, IEEE Xplore, SpringerLink, ScienceDirect and Google Scholar.

2.2. Inclusion criteria

Works published between 2018 and 2026, of peer-reviewed type or monographs, indexed in Scopus/WoS, written in English were retained.

2.3. Exclusion criteria

Conference abstracts, duplicate articles, editorial materials, grey literature and strictly technical studies, without a human or institutional dimension, were eliminated.

3. HUMAN DECISION BEFORE ARTIFICIAL INTELLIGENCE

3.1. Classical Decision Theories

Herbert Simon, Daniel Kahneman, Amos Tversky, James March, and Henry Mintzberg have demonstrated, each from their own perspective, that the classical decision-making process depends on the limited capacity of the human brain to process information.

3.2. Perfect Rationality versus Bounded Rationality

Classical economic models assumed a perfectly rational decision-maker — Homo Economicus. Simon dismantled this fiction through the concept of bounded rationality: the human decision-maker cannot access all relevant information and cannot calculate all possible consequences, so he chooses the "satisfactory" solution, not the optimal one.

In accordance with recent analyses of risk governance in organizations (Marian, 2026a), the structural limits of information processing make institutional decision-making vulnerable precisely under conditions of high uncertainty — precisely the situations in which AI systems are introduced today with the highest expectations.

3.3. Systematic Cognitive Errors

Building on the foundation laid by Kahneman and Tversky, the psychological literature has mapped a series of recurring judgmental distortions:

• confirmation bias — the selective search for information that validates a preexisting belief;
• availability bias — the overvaluation of recent or easily recalled events;
• anchoring — excessive reliance on the first information received;
• overconfidence — overestimating the accuracy of one's own predictions;
• loss aversion — the emotional disproportion between the fear of losing and the desire to win.

4. ARTIFICIAL INTELLIGENCE AS A DECISION SUPPORT SYSTEM

4.1. The Evolution of AI

We could recall the history, the evolution, the path even if they are known, for the clarity of the argument: expert systems based on rigid rules from the 1980s, followed by machine learning algorithms capable of recognizing patterns in data, then deep neural networks and, currently, generative artificial intelligence and large-scale linguistic models — capable of synthesizing large volumes of text and interacting in natural language with the decision-maker.

4.2. Decision-making augmented by AI

The integration of AI does not eliminate the human component from the decision, but rather amplifies it. Four parameters change visibly: processing speed decreases from days to milliseconds; accuracy increases by reducing calculation errors; analytical capacity extends to

volumes of data impossible to process manually; and forecasting gains in accuracy through predictive models.

4.3. Application areas

The effect of AI is felt in most institutional sectors, as summarized in: Table 1.

Table 1. Application areas of AI in decision-making support

Area Decision-making role of AI Application examples

Law / Justice Case law analysis, judicial predictability Recidivism risk assessment, preliminary drafting of documents

Management / Business Resource optimization, financial analysis Capital allocation, credit scoring, recruitment

Medicine / Health Precision diagnosis Personalized therapeutic decisions, imaging analysis

Public administration Service automation, fraud detection Social protection, tax optimization and audit

Finance / Banking Algorithmic trading, scoring Portfolio management, money laundering detection

Sport Injury prevention, tactics Physical load monitoring, talent identification

5. OPPORTUNITIES CREATED BY AI

5.1. Better decisions

Partial neutralization of human cognitive errors — fatigue, impulsivity, emotional state of the moment — allows decisions to be more firmly anchored in empirical evidence.

5.2. Faster decisions

Real-time analytics gives decision-makers in management, healthcare, or administration the ability to react almost instantly to crises.

5.3. Predictive capacity

Predictive modeling moves management from reactive to proactive. Some authors emphasize, in a note that directly applies to the AI context, that the exploitation of an algorithmic prediction requires a rigorous risk governance framework — otherwise, prediction remains a technical exercise without real institutional consequence (Marian (2026a).

5.4. Personalization

AI allows for large-scale personalization: genetically calibrated treatments in medicine, learning paths adapted to the pace of each student, public services individualized to the citizen's profile.

5.5. Organizational performance

At the micro and macroeconomic level, AI streamlines internal processes and supports product innovation.

6. RISKS OF AI-BASED DECISION-MAKING

6.1. Algorithmic Bias

Algorithms learn from historical data that already carries structural discrimination. When training data is flawed, AI amplifies — not corrects — racial, gender, or socioeconomic discrimination in hiring, lending, or bail decisions.

6.2. Lack of Transparency

Many deep learning models function as a “black box”: reconstructing the logic that led to a conclusion is often impossible, which affects the right to explanation and undermines institutional trust.

6.3. Accountability

When an AI system generates an erroneous decision, the question of responsibility arises: the programmer, the developer company, the user institution, or the human decision-maker who validated the recommendation?

6.4. Data Protection

The massive data collection required to train algorithms often conflicts with regulations such as GDPR, raising risks of unauthorized surveillance.

6.5. Overreliance (automation bias)

Decision-makers tend to uncritically accept algorithm recommendations. In the long term, this phenomenon leads to cognitive disqualification (deskilling) and the erosion of critical thinking.

6.6. Ethical challenges

Human dignity, fairness, and individual autonomy can be undermined when major decisions regarding a person’s life or freedom are delegated to automated systems.

7. MULTIDISCIPLINARY PERSPECTIVES

7.1. The legal perspective

The main legal challenge remains the establishment of the civil and criminal liability regime. The European Regulation on Artificial Intelligence (EU AI Act) takes a risk-level approach, with strict requirements for systems intended for justice, public order or credit assessment. In an analysis dedicated to budgetary sustainability and the institutional foundations of public financial law, it is shown that the function of regulation is not to block innovation, but to draw clear limits to it — a principle equally valid for AI governance (Marian, 2026b).

7.2. The economic perspective

Economists are tracking the impact of AI on productivity and labor market dynamics. The economic benefits are real, but so is the risk of labor displacement and widening gaps between advanced and developing economies. The contrast already described in section 1.1 — a research institute of 105 million lei in Cluj-Napoca, versus a municipal strategy of around 105,000 euros in Oradea — is not a simple budget difference, but an indicator of the institutional capacity gap between a consolidated university hub and a medium-sized urban center, a gap that the National Strategy for Artificial Intelligence 2024-2027 explicitly recognizes as a risk of regional fragmentation of technological adoption. The asymmetry of scale and speed observed between university centers and local administrations (as illustrated by the contrast between the AIRi Cluj project and the Oradea Smart City Strategy) does not represent a simple budgetary difference, but signals the structural risk of regional fragmentation of AI governance.

While large technology hubs develop their own analytical capabilities and advanced decision support models, mid-sized administrations remain dependent on external solutions or gradual adoption processes. Without an integrated national governance framework (such as the AIGF model proposed in this paper), these gaps in institutional capacity can deepen the rift between innovative urban centers and resource-constrained administrations, equally affecting the equity of public services and the uniform application of automated decision regulations.

7.3. The managerial perspective

AI is changing the role of the organizational leader: he no longer just distributes tasks, but orchestrates human-machine collaboration, integrating organizational learning with strategic change management.

7.4. The psychological perspective

Psychology investigates trust in technology, cognitive load, and acceptance of hybrid systems. People oscillate between irrational fear of the machine and blind faith in the algorithm—both attitudes distort the final decision.

7.5. The sociological perspective

The expansion of AI influences social cohesion. Disparities in access to technology generate new forms of digital inequality, and unchecked automation can erode the legitimacy of public institutions.

7.6. Sports Sciences

In sports science, AI is reconfiguring tactical modeling and talent identification. The final decision, in moments of extreme pressure, however, remains dependent on psychological factors and team cohesion. based on the Belbin model applied to group dynamics, it is shown that organizational performance results from the balance between human resilience and analytical tools—not from the substitution of one for the other (Marian, 2022, 2023).

8. TOWARDS A GOVERNANCE FRAMEWORK IN PERSPECTIVE

8.1. Existing governance models

OECD guidelines, UNESCO recommendations, the EU AI Act, the NIST AI Risk Management framework and ISO standards provide valuable principles, but often remain general or regionally fragmented.

8.2. Proposed Integrated Governance Framework

As an original contribution of the paper, we propose the AI Governance Framework (AIGF), structured on six interconnected dimensions, applicable incrementally (each new step is added on top of the old one, improving it, without spoiling what was done before):

AI-Assisted Decision Governance Framework (AIGF)

No. crt.	Governance Dimension	Description and Key Components	Objective / Impact in the Decision-Making Process
1	**Legal compliance**	Strict and continuous alignment with national, European (e.g. EU AI Act) and international regulations in force.	Preventing legal non-compliance and positioning the system at the appropriate level of legal risk.
2	**Ethical principles**	Respect for the principles of equity, non-discrimination, inclusion and protection of fundamental human rights.	Mitigating the takeover/amplification of historical discrimination and protecting human dignity.
3	**Transparency and Explainability (XAI)**	Using explainable algorithms (Explainable AI) and providing the right to explanation.	Eliminating the "black box" effect and strengthening institutional trust.
4	**Human supervision**	Mandatory application of the Human-in-the-Loop (HITL) principle in all administrative, medical or judicial decisions with major impact.	Maintaining human control and critical thinking over automated decisions.
5	**Institutional responsibility**	Clear and unambiguous delimitation of the chain of legal responsibility	Eliminating uncertainty regarding civil or criminal

No. crt.	Governance Dimension	Description and Key Components	Objective / Impact in the Decision-Making Process
		(developer, supplier, user institution, human decision-maker).	liability in the event of algorithmic errors.
6	**Continuous learning and auditing**	Post-implementation monitoring of algorithm performance, periodic independent auditing and model retraining.	Preventing data degradation (data drift) and adapting the system to environmental dynamics.

8.3. Public policy recommendations

For governments and administrations — establishing national bodies to audit public sector algorithms. For universities — interdisciplinary programs in ethics and technological governance; such a program, built on common modules in law, management and organizational psychology, could be piloted at the level of a single faculty before regional expansion, from Cluj-Napoca to smaller university centers such as Oradea. For companies — internal ethics committees, active before the market launch of any automated decision. For justice — guaranteeing the right to human review of any automated administrative or judicial decision. It is worth pointing out that the information about the AIRi Institute (UTCN, Cluj-Napoca) comes from economic press sources (Transilvania Business, October 2025) and can be additionally verified on the UTCN website. The information about the Oradea Smart City Strategy comes from Transilvania Business (September 2025) and from the press releases of the Oradea City Hall.

9. FUTURE RESEARCH AREAS

There are areas that invite, in the future, systematic and in-depth research:

• AI and leadership — there is a lack of empirical studies on how AI modifies moral authority and leadership style.

• AI in public order — there are no comparative models on the efficiency and ethical impact of police decisions based on predictive algorithms.

• AI and sports integrity — there is a lack of a dedicated legal framework for predictive analytics used in preventing match-fixing.

• AI and citizen trust — we still know little about the psychological mechanisms by which citizens accept or reject automated administrative decisions.

CONCLUSIONS

Artificial intelligence does not replace the human decision-maker; it irreversibly transforms the nature and dynamics of his decision. The major benefits — precision, speed, predictive capacity — are only realized through careful calibration of input data, coupled with the maintenance of human critical thinking.

The risks—algorithmic bias, opacity, uncertainty of accountability, overreliance—remain real threats to human rights and institutional stability. Adopting a six-tier governance framework, such as the one proposed here, is the condition that can transform the transition to hybrid human-machine decision-making into an engine of institutional progress, rather than an additional source of fragility.

REFERENCES

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

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1, 389–399.

Kahneman, D. (2011). Thinking, Fast and Slow. New York: Farrar, Straus and Giroux.

Marian, A. S. (2026a). Risk Management and Governance in Organizations — Legal and Economic Foundations and Practical Applications. Cluj-Napoca: Galaxia Gutenberg Publishing House. ISBN 978-630-337-242-6.

Marian, A. S. (2026b). Public Financial Law — Legal Regime of Public Finance: From Institutional Foundations to Budget Sustainability. Cluj-Napoca: Galaxia Gutenberg Publishing House. ISBN 978-630-337-201-3.

Marian, A. S. (2022). How Can We Build Effective Teams? Belbin Model and Tests. In: Nové trends vzdělání polístico IV. Holešov: Vyšší poliční škola a Střední poliční škola Ministerstva vnitra v Holešově. ISBN 978-80-7616-177-1.

Marian, A. (2023). The contemporary world means a world of teams. Do we want efficient teams? How do we make them? Cluj University Journal, Interdisciplinary: Social Sciences and Humanities, 1(1).

Mintzberg, H. (1979). The Structuring of Organizations. Englewood Cliffs, NJ: Prentice-Hall.

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

Simon, H. A. (1997). Administrative Behavior (4th ed.). New York: Free Press.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185(4157), 1124–1131.