

Artificial Intelligence and the Principle of Decision-Making Transparency: From Opaque Algorithms to Explainable Justice in European and Romanian Law

PhD Student **Gabriel Florinel ION**¹

Abstract

This research analyzes the legal challenges generated by the increasing use of artificial intelligence systems in public and private decision-making processes, with particular emphasis on the principle of decision-making transparency. The central hypothesis is that algorithmic opacity may undermine fundamental rights, effective judicial protection, and democratic accountability, thereby requiring a reconceptualization of traditional transparency requirements in the context of automated decision-making. The methodology combines doctrinal analysis, comparative examination of European and Romanian legislation, and the study of relevant case law and regulatory approaches. Particular attention is given to the GDPR, the EU Artificial Intelligence Act, the Romanian administrative-law framework, and emerging requirements concerning explainability, human oversight, and algorithmic auditing. The analysis identifies persistent tensions between technological complexity, trade-secret protection, and the need to provide individuals and courts with meaningful explanations of decisions affecting legal rights and interests. The study argues that explainability should operate as a substantive legal safeguard rather than merely a technical feature of AI systems and proposes legislative and institutional mechanisms capable of ensuring effective oversight, contestability, and accountability. These findings have broader implications for administrative law and for the development of a coherent rights-based framework governing artificial intelligence.

Keywords: artificial intelligence, decisional transparency, opaque algorithms, explainable justice, European law, GDPR, AI Act.

JEL Classification: K24, K40

DOI: <https://doi.org/10.62768/ADJURIS/2026/7/15>

Please cite this article as:

Ion, Gabriel Florinel. "Artificial Intelligence and the Principle of Decision Making Transparency: From Opaque Algorithms to Explainable Justice in European and Romanian Law." In *Comparative and International Law in an Era of Global Transformation: Sovereignty, Justice, Cooperation, and Emerging Technologies*, edited by Liliana Tymchenko and Paulina Guli Hoti, 258–277. ADJURIS – International Academic Publisher, 2026.

<https://doi.org/10.62768/ADJURIS/2026/7/15>

¹ Gabriel Florinel Ion - Doctoral School of Socio-Human Sciences, "Dunarea de Jos" University of Galati, Romania. ORCID: <https://orcid.org/0009-0008-2491-3273>. E-mail: gabrielgaby14598@gmail.com.

1. Introduction

Artificial intelligence has become ubiquitous in contemporary decision-making systems, from the algorithms used by banking institutions to assess creditworthiness to the automated systems that manage the granting of social benefits, the distribution of public resources or the evaluation of professional performance. This technological transformation redefines the way decisions with legal effect are adopted and creates substantive challenges for the legal order, directly affecting the fundamental principles of transparency, accountability and democratic control over decision-making.

The proliferation of artificial intelligence systems in the public and private sectors has generated the emergence of a new type of administrative and civil decision, algorithmic decision, which does not fit into traditional legal categories. If in classical law the decision was the expression of the will of a natural person or legal entity, in the case of AI systems the decision becomes the result of an automated, complex and often inaccessible process to direct human explanation. This change raises questions about the nature of the legal act, the actual holder of the Will and the ways in which constitutional principles of transparency, motivation and access to information can be guaranteed in an opaque technological framework.

Decision-making transparency is a pillar of the rule of law, enshrined in fundamental international instruments. The European Convention on human rights and the Charter of Fundamental Rights of the European Union provide for the right of individuals to understand decisions affecting them and to benefit from appropriate procedural safeguards. In Romanian law, the principle of administrative transparency is enshrined in the 1991 constitution, revised in 2003. Article 31 guarantees the right to information and Article 52 ensures the right of persons to apply to the authorities by petitions and to challenge administrative acts. These rules presuppose the effective possibility of knowing the reasons for the decision, a condition that is difficult to fulfill when the basis of the decision is opaque algorithms.

Machine learning algorithms, especially advanced models of deep neural networks, operate on the basis of complex mathematical structures, the internal processes of which remain unintelligible even for their developers. This technical opacity, referred to in the literature as the black box phenomenon, calls into question the compatibility between AI systems and the obligation of authorities or private entities to provide reasoned and comprehensible decisions. In the absence of an understanding of algorithmic reasoning, the affected individuals cannot effectively exercise the right to challenge the decision and the courts cannot carry out an effective check on its legality and proportionality.

This research aims to analyze the compatibility of artificial intelligence

systems with the principle of decision-making transparency in European and Romanian law, assessing to what extent the requirements of the rule of law can be maintained in a technological context marked by automation. The adopted methodology combines doctrinal and normative analysis with a comparative study of relevant regulations and jurisprudence, with the ultimate objective of identifying legal solutions that allow reconciling technological innovation with the fundamental requirements of human rights protection. The study aims both to assess the legal risks associated with algorithmic decisions and to formulate regulatory proposals capable of ensuring a balance between technological efficiency and respect for democratic principles.

2. Theoretical and Legal Foundations of Decision Transparency

Decision-making transparency is the obligation of public authorities and, in certain situations provided for by law, also of private entities, to make public the reasons, criteria and procedures on the basis of which decisions with an impact on individual rights and freedoms are adopted. This principle is of central importance in the architecture of the rule of law, since it allows the exercise of democratic control over public power and ensures the effective protection of individuals against arbitrariness. Decision-making transparency is manifested in three essential dimensions. The first is the accessibility of information, which implies that the data relevant to understanding the decision must be made available to the data subject in a clear and complete form. The second dimension is the intelligibility of the decision-making process, which implies the possibility of understanding the reasoning or criteria used in making the decision. The third is the possibility of appeal, which presupposes the existence of effective procedural mechanisms by which the decision can be subject to administrative or judicial review.

In private law, the principle of transparency acquires specific valences, adapted to the nature of contractual legal relations and professional obligations. In modern contractual relationships, especially in digital media, transparency becomes an essential component of information obligations and standards of good faith. Private entities providing services of general interest, such as utilities, financial institutions or digital platforms, have increased transparency obligations due to their dominant position or systemic impact on users.

In the context of the use of algorithms, these obligations include providing clear information on the automated criteria used to establish credit scores, assess risk, moderate content or personalize commercial offers. Thus, transparency extends beyond the traditional communication of contract terms, including the need to explain how digital systems work that can significantly influence the legal and economic position of users.

The European Court of human rights has developed a consistent case law on the obligation of authorities to motivate decisions impacting on the rights of

individuals, considering that reasoning is an essential guarantee against arbitrariness and an indispensable element of democratic decision-making. In many cases, the court has emphasised that the persons concerned must be able to understand the reasons behind the decision in order to assess the appropriateness of challenging it and to benefit from effective judicial review. This requirement applies to both administrative and jurisdictional decisions, enshrining a high standard of procedural transparency. In European Union law, the principle of good administration, expressly enshrined in Article 41 of the Charter of Fundamental Rights, includes the right of every person to receive a clear and understandable statement of reasons for administrative decisions concerning him or her. This right provides a sound regulatory framework for examining the compatibility of algorithmic decisions with European requirements on transparency and accountability.

Transparency is not an autonomous goal, but an indispensable legal instrument for the realization of other fundamental principles of the rule of law, such as legality, proportionality and accountability. Firstly, transparency makes it possible to verify the legality of the decision-making, since an opaque decision cannot be effectively controlled in terms of compliance with the law. Second, proportionality can only be assessed if the criteria used by the authority and how they have been applied to the concrete situation are known. Thirdly, legal responsibility depends on the possibility of identifying the factors that caused the decision and understanding the internal mechanisms of the decision-making process. In the context of algorithmic decisions, transparency is of critical importance: without access to the relevant information on how the intelligent system works, the affected persons cannot effectively exercise their procedural rights and the authorities cannot demonstrate compliance with constitutional and European standards. Thus, transparency becomes the essential condition for the legitimate integration of artificial intelligence into public and private decision-making processes.

Article 6 of the European Convention on Human Rights enshrines the right to a fair trial, which implicitly includes the obligation for courts to provide sufficient and understandable reasons for judgments given. The European Court of human rights has consistently pointed out that the right of reasons is not only a procedural formality, but a substantial guarantee that allows the parties to understand the court's reasoning and assess the appropriateness of the appeal. Proper reasoning ensures the transparency of the act of justice, fosters public confidence in the functioning of the judicial system and allows the exercise of democratic control over judicial authority.

The case law of the Strasbourg court has gradually extended this requirement also to administrative decisions affecting civil rights, especially in situations where administrative decisions have direct consequences on the legal situation of the person. The court held that, when an administrative authority adopts decisions affecting rights recognised by the legal order, the person concerned

must be able to understand the reasons for the measure taken. This reasoning is based on the idea that the protection of civil rights cannot be effective if the administration adopts non-transparent or opaque decisions that escape judicial review. Therefore, the principles enshrined in Article 6 are also reflected in the obligation of the administrative authorities to ensure clear, comprehensible and accessible reasoning.

The Constitution of Romania enshrines, in Article 124, the principles governing the administration of justice, including judicial independence. The obligation to provide reasons for judgments is developed by procedural legislation and is reinforced by the guarantees of Article 6 of the European Convention on Human Rights. This requirement also supports effective judicial review of administrative action, since the legality and proportionality of measures affecting rights or legitimate interests cannot be assessed without adequate reasons.

Romanian administrative law also requires administrative decisions to be adequately reasoned in the circumstances provided by applicable legislation and developed through case law. Reasoning must enable the person concerned and the reviewing court to identify the factual and legal basis of the measure. This requirement is essential for protection against arbitrariness and for effective judicial review, strengthening the ability of the individual to assess whether a decision complies with legality and proportionality.

3. Automated Decisions and the Challenge of Opaque Algorithms

Automated decisions are characterized by a number of obvious operational advantages, such as high processing speed, the ability to simultaneously analyze very large amounts of data and scalability that allows the same criteria to be applied to large populations. These systems can identify subtle statistical patterns and correlations that are inaccessible to traditional analysis, generating predictions or classifications with a high level of accuracy in certain areas, such as financial risk analysis or fraud detection. The appearance of objectivity results from the fact that algorithms apply uniform mathematical rules and are perceived as devoid of subjective preferences or influences, which creates the impression of impartiality superior to human.

However, this appearance can be misleading. Artificial intelligence systems are deeply dependent on the data that is provided to them at the training stage, and real data inevitably reflects the social structures from which they come, including existing prejudices, inequalities and discrimination. In the absence of adequate corrective measures, algorithms can reproduce these biases, amplifying them on a large scale. The result is the emergence of forms of systemic discrimination, which are generated not by a direct intention, but by the automated way of functioning of the system.

This phenomenon becomes particularly problematic in sensitive areas, where automated decisions can profoundly affect the lives of individuals. In

criminal justice, risk assessment systems may disadvantage certain ethnic or socioeconomic groups. In lending, predictive models can exclude individuals based on statistical correlations that reflect historical inequalities. In recruitment, algorithms can privilege profiles traditionally associated with certain categories of employees, reducing the chances of candidates who do not align with this pattern. In all these situations, automated decisions can generate discriminatory effects that are difficult to challenge, especially in the absence of transparency on how the system works.

Algorithmic opacity has diverse and interrelated causes. First, the technical complexity of machine learning algorithms, especially deep models, makes internal computing processes difficult to interpret even for specialists². Decisions are generated based on mathematical weights that are continuously adjusted according to training data, and the logical route cannot be reconstructed intuitively. Second, developers frequently invoke trade secrecy to prevent disclosure of the source code or the system's internal architecture, making it difficult to externally analyze its operation. Thirdly, the lack of clear legal standards on the degree of transparency required in automated decisions contributes to the absence of a uniform reporting and control framework. All this leads to an obvious paradox: the more efficient and sophisticated the algorithm becomes, the opaquer and more difficult to verify the decision logic becomes.

From a legal perspective, the tension between the protection of intellectual property and the demand of transparency is one of the central issues of the use of artificial intelligence in decision-making processes. On the one hand, developers have a vested interest in protecting innovation, as algorithms represent substantial investments in research and development. On the other hand, public authorities and private entities with functions of general interest must provide adequate explanations to those affected by automated decisions, in order to ensure respect for fundamental rights and the principles of the rule of law. Finding a balance between these interests is essential and involves the development of legal mechanisms that allow access to information relevant to the control of the legality of the decision, without compromising the developers' trade secrets entirely.

Algorithmic opacity generates major risks for the protection of fundamental rights. First, the inability to understand decision logic can hide situations in which the algorithm produces arbitrary or discriminatory effects. Without transparency, the damages remain unidentified and the affected persons cannot know the real reasons for the treatment applied. Secondly, the lack of access to relevant information makes it difficult or even impossible to pursue redress.

Algorithmic discrimination can take direct or indirect forms, and under-

² Sylvia Lu, "Algorithmic Opacity, Private Accountability, and Corporate Social Disclosure in the Age of Artificial Intelligence," *Vanderbilt Journal of Entertainment and Technology Law* 23, no. 1 (2020): 99–159, <https://scholarship.law.vanderbilt.edu/jetlaw/vol23/iss1/3>.

standing these mechanisms is essential for assessing the impact of decision-making technologies on fundamental rights. Direct discrimination occurs in situations where the algorithm explicitly uses protected characteristics, such as gender, ethnicity, religion, or disability, to determine the decision outcome. This occurs either by including sensitive variables in the predictive model or by introducing rules that lead to differential treatments on prohibited criteria. In such cases, the violation of the right to equality is obvious and legal responsibility can be established relatively quickly, since the discriminatory decision is explicitly based on a protected category.

Indirect discrimination is much more difficult to identify, but just as problematic from a legal point of view. This is manifested when the algorithm uses proxy variables, that is, information that is not directly protected, but closely correlates with sensitive features. For example, the zip code may function as a proxy for socioeconomic status or ethnic affiliation, and employment history may reflect gender inequalities. Such correlations are commonly found in real data sets and, if not adequately managed, can reproduce and amplify existing structural biases in society. Indirect discrimination produced by algorithms is particularly dangerous because they act invisibly, generating unequal treatment without these sensitive criteria being formally recognized in the decision-making process.

The identification and Combating of these forms of discrimination depend essentially on access to information on the functioning of the algorithm, the types of data used and the internal logic of the decision-making process. Without transparency, it is impossible to verify whether proxy variables have generated disproportionate effects on a protected group and the affected persons cannot demonstrate the existence of indirect discrimination. Therefore, the legal framework must ensure the possibility of auditing algorithms, access to understandable explanations and effective mechanisms of appeal. Only in this way can it be guaranteed that technology does not become a tool for perpetuating inequalities, but contributes to a fair and responsible administration of decisions that affect the rights of individuals.

4. Applicable European and Romanian Regulatory Framework

Article 22 of the GDPR³ recognises individuals with the right not to be subject to a decision based solely on automated processing, including profiling, where this decision produces legal effects on them or significantly affects them. This provision represents the first explicit legal recognition at European level of the risks arising from algorithmic decisions, marking an essential step in shaping

³ Regulation (EU) 2016/679 of the European Parliament and of the Council of April 27, 2016, on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation), art. 22(1), *Official Journal of the European Union* L 119 (May 4, 2016): 1–88, <https://eur-lex.europa.eu/eli/reg/2016/679/oj>.

effective protection of individuals from autonomous decision-making technologies.

Article 22 also imposed the obligation to provide adequate safeguards in situations where automated processing is permitted by exception, such as significant human intervention, the possibility for the data subject to express his or her point of view and the right to challenge the decision. These safeguards reflect the need to maintain effective human control over algorithms, ensuring that the final decision is not the product solely of opaque mathematical logic, but the result of a process in which standards of fairness, transparency and proportionality are met.

Recital 71 of the GDPR⁴ states that data subjects should be provided with significant information on the logic used in automated processing processes, in particular in the case of decisions that produce legal or similar effects to an important extent. This wording has been interpreted by part of the doctrine as equivalent to the establishment of a right to explanation, in the sense of the obligation of the controller to disclose the decision-making reasoning of the algorithm and to clarify how the data of the data subject was evaluated. In doctrine, however, this interpretation is strongly disputed. Although the wording in the recital appears to grant a substantial right to transparency, the normative text of Article 22 and related provisions does not explicitly enshrine such a right. Consequently, many authors emphasise the rather indicative nature of the recital, in the absence of an overriding rule requiring operators to provide detailed explanations on the operation of algorithms. More recent developments have substantially clarified this debate, particularly following the recognition of a genuine right to explanation under Article 15(1)(h) GDPR and the introduction of a distinct right to explanation in Article 86 of the AI Act.⁵

AI Act, adopted in 2024⁶, it represents the first truly comprehensive global regulatory framework dedicated to the regulation of artificial intelligence systems. The regulation introduces a methodical approach based on the classification of AI systems according to the level of risk they pose to Fundamental Rights, Public Security and legal order. In this regulatory architecture, systems considered high-risk are subject to a substantially stricter legal regime, which includes firm transparency standards, detailed data governance requirements, mandatory human surveillance and continuous monitoring mechanisms.

Among the central obligations imposed on these systems, transparency

⁴ Andrew D. Selbst and Julia Powles, “Meaningful Information and the Right to Explanation,” *International Data Privacy Law* 7, no. 4 (2017): 233–242, <https://doi.org/10.1093/idpl/ix022>.

⁵ Lars Arnesen, “Comparing the Right to Information about Automated Decision-Making to the New Right to Explanation,” *International Data Privacy Law* 16, no. 2 (2026), <https://doi.org/10.1093/idpl/ipag008>.

⁶ Regulation (EU) 2024/1689 of the European Parliament and of the Council of June 13, 2024, laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), *Official Journal of the European Union* L 2024/1689 (July 12, 2024), <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>.

occupies an essential place. The regulation provides that suppliers are obliged to provide complete technical documentation to allow supervisory authorities to assess how the system works and identify possible risks. This documentation shall include details of the algorithmic architecture, data sets used, testing and validation procedures, and risk mitigation measures. In addition, AI Act establishes the obligation to report incidents and to ensure a rigorous traceability of the decisions generated by the system.

Article 13 of the AI Act⁷ it explicitly enshrines the principle of operational transparency for high-risk AI systems. The legal text establishes that these systems must be designed and developed in a way that allows end-users to understand, to a reasonable extent, how results are generated. The purpose of this provision is not to fully disclose the technical structure of the algorithms, but to give users the opportunity to correctly interpret the predictions, classifications or recommendations generated by the system. Through this approach, the regulation seeks to prevent the automatic and uncritical use of algorithmic results, promoting the active involvement of the human factor in decision-making processes that can produce significant legal effects or consequences.

Overall, the transparency obligations imposed by the AI Act are an important step towards making technology operators accountable and ensuring that AI systems are compatible with the rule of law principles. They create a framework in which technological innovation can be harnessed without compromising the protection of fundamental rights and without allowing the emergence of areas of opacity incompatible with contemporary legal requirements.

The current Romanian Administrative Code was adopted by Government Emergency Ordinance No. 57/2019⁸, which lays down general principles governing public administration. The obligation to state reasons for administrative acts is developed through applicable legislation and case law.

This requirement is not only a procedural formality, but is an indispensable tool for the control of legality and for the Prevention of administrative arbitrariness. By means of reasoning, the necessary link is made between the exercise of Public Power and the person's right to understand the reasons underlying a decision affecting his legal situation.

In the context of the emergence of algorithmic decisions, the application of art. 10, however, raises obvious difficulties. A machine learning algorithm does not operate through explicitly formulable reasoning, but through complex statistical processes that are difficult to translate into intelligible legal language. Thus, the obligation to state reasons, although present in the legal text, risks becoming inoperable when the public authority uses automated systems whose internal mechanisms cannot be explained in traditional terms. This situation reveals

⁷ Regulation (EU) 2024/1689, art. 13(1).

⁸ Government Emergency Ordinance No. 57 of July 3, 2019, regarding the Administrative Code, *Monitorul Oficial* no. 555 (July 5, 2019), <https://legislatie.just.ro/Public/DetaliiDocumentAfis/249654>.

the need to reinterpret and adapt the norm to the specifics of modern technology, in order to avoid a vacuum of legal protection.

Law No. 190/2018⁹, which establishes national measures for implementing certain provisions of the GDPR, does not provide further clarification on the application of the rules on decisions based solely on automated processing. Although the law supplements the European regulation at national level, it does not introduce detailed rules on the obligations of transparency, auditability or motivation of algorithmic decisions taken by public authorities or private entities covered by administrative law. Thus, the lack of specific provisions generates an obvious gap between the European requirements and the internal normative framework.

This legislative loophole is particularly problematic in the context of the accelerated digitization of Public Administration, where artificial intelligence systems are beginning to be used in sensitive areas such as the allocation of social benefits, the management of administrative applications or risk assessment. In the absence of clear rules on responsibility, motivation and access to Algorithmic Information, data subjects may find themselves unable to effectively challenge automated decisions affecting their rights. Therefore, the urgent adaptation of national legislation becomes essential to ensure compatibility between the use of technology and the fundamental principles of the rule of law.

5. Explicable Justice: Between Normative Ideal and Technological Reality

The concept of explainable artificial intelligence, known in the literature as explainable AI or XAI, represents an essential technological direction in the evolution of modern algorithmic systems¹⁰. The central goal of this field is the development of artificial intelligence models whose decision-making processes can be tracked and understood by users, be they experts or people without technical training. Unlike black box algorithms, XAI aims to make the internal reasoning of the system visible, by generating coherent explanations of the reasons that determined the final result.

From a legal perspective, explainability is not only a technological advantage, but an indispensable condition for the compatibility of automated decisions with the fundamental principles of the rule of law. An algorithmic system affecting legitimate rights and interests cannot remain opaque, since the lack of

⁹ Law No. 190 of July 18, 2018, on measures for implementing Regulation (EU) 2016/679, *Monitorul Oficial* no. 651 (July 26, 2018), <https://legislatie.just.ro/Public/DetaliiDocumentAfis/203151>.

¹⁰ David Gunning, "DARPA's Explainable Artificial Intelligence (XAI) Program," in *Proceedings of the 24th International Conference on Intelligent User Interfaces (IUI '19)* (New York: ACM, 2019), ii, <https://doi.org/10.1145/3301275.3308446>.

transparency makes it impossible to check legality, does not allow the identification of possible discrimination and does not give the data subject the real opportunity to bring an effective appeal.

In order to have legal relevance, explainability must meet certain essential criteria. First of all, the explanation provided must be understandable to the addressee, which implies the adaptation of the technical language to a level accessible to an ordinary person. Secondly, the explanation must be sufficiently detailed to allow verification of the legality of the decision, including the identification of the criteria used by the algorithm and how they influenced the outcome. Third, the explainability must provide the person affected with sufficient information to bring a reasoned appeal capable of triggering an effective judicial or administrative review.

There are three broad categories of technological approaches to achieving algorithmic explainability, each with specific advantages and limitations. The first category is interpretable models, designed from the beginning to be completely transparent.

These models use simple decision structures, such as decision trees or linear regressions, which allow accurate tracking of the logic that leads to the final result. Their advantage lies in the fact that they provide native and understandable explainability. However, this transparency is often achieved at the expense of performance, especially in complex contexts that require processing large volumes of data or identifying subtle patterns. For this reason, interpretable models are rarely used in advanced artificial intelligence applications.

The second category is that of post hoc explanations, which seek to provide an intelligible justification for the decisions of an opaque model after they have been generated. Methods such as relevance maps, influence vectors, or sensitivity analysis allow one to roughly reconstruct the model's internal reasoning without fully revealing its operation. While these techniques can improve transparency, they usually produce approximate explanations, raising questions about their legal value and the degree to which they can serve as a basis for challenging automated decisions. In the absence of complete transparency, the persistence of the risk of misinterpretation remains significant.

The third category is hybrid models, which combine the robustness and performance of opaque algorithms with dedicated explanatory capabilities. These systems use a complex model for making the actual decision, but resort to a separate interpretable model to generate the necessary explanations. This approach allows maintaining the high level of accuracy, specific to neural networks or other advanced machine learning techniques, while providing consistent and verifiable justifications in parallel. From a legal perspective, hybrid models are the most promising solution, as they offer a balance between technological performance and legal requirements for transparency and accountability¹¹.

¹¹ Zachary C. Lipton, "The Mythos of Model Interpretability," *Communications of the ACM* 61,

Technical explainability does not always overlap with legal explainability, and this difference produces important consequences in assessing the legality of automated decisions. A technical explanation may correctly describe the internal functioning of an algorithm, highlighting statistical weights, activation functions or influence variables, but this information may be irrelevant or inaccessible from the perspective of an ordinary person or even a practitioner of law. Thus, an adequate technical explanation is not necessarily a sufficient legal explanation to satisfy the requirements as to the reasons for the act, its intelligibility and the possibility of effective appeal.

The difficulty is amplified in situations where explanations provided by algorithmic systems can be manipulated or oversimplified. Some systems can produce explanations built to apparently satisfy the motivation requirement without actually reflecting the internal logic of the algorithm. Other situations raise privacy concerns, as providing a detailed explanation may involve the disclosure of protected information such as trade secrets, sensitive data or proprietary parameters. For this reason, explainability is constantly in a delicate balance between transparency and the protection of economic and technological interests.

Courts face significant difficulties in assessing algorithmic explanations because of their deeply technical nature. Judges are called upon to assess the validity and accuracy of decision-making mechanisms that often go beyond traditional legal competences, which requires the use of independent experts in computer science and data science. This technical dependence can affect the process balance, as parties with superior technical and financial resources can produce more elaborate expertise, while vulnerable persons may have obvious difficulties in combating the conclusions presented by powerful entities such as banks, corporations or digitized public authorities.

These challenges naturally raise questions about respect for the principle of equality of arms and effective access to justice. If a person is not able to understand the rationale for an algorithmic decision, nor to challenge it in a reasoned way, the right to an effective remedy becomes purely formal. In this context, legal explainability must be conceived not only as a translation of technical terms into an accessible language, but as a mechanism that allows the actual exercise of procedural rights. Therefore, the development of a regulatory framework that imposes clear standards on the content, form and degree of detail of algorithmic explanations becomes indispensable for ensuring effective protection of those affected.

6. Algorithmic Probation and Access to Justice

Probation in disputes involving artificial intelligence systems raises significant difficulties that go beyond the traditional evidentiary framework. First,

access to relevant technical information is inherently limited, since the architecture of algorithms, training data and operating parameters are often protected by the legal regime of trade secrets or intellectual property rights¹².

This protection, justified from the perspective of developers, can, however, become a major obstacle to the exercise of the rights of Defense for people affected by automated decisions. In the absence of legal mechanisms allowing controlled access to this information, the interested party cannot obtain the necessary elements to challenge the legality or correctness of the functioning of the algorithmic system.

Second, even when technical information becomes accessible, its interpretation requires advanced knowledge in Computer Science, Statistics and machine learning theory. This complexity makes seemingly relevant evidence unintelligible to the parties, lawyers and sometimes even the court. In such situations, the real problem lies not only in access to information, but also in the effective ability to use it in the judicial process. The appeal to experts becomes inevitable, which can extend the duration of the proceedings and considerably increase the costs of litigation.

The burden of proof can turn into a disproportionate burden for the person challenging an algorithmic decision. If the plaintiff has to prove the existence of damage, the identification of its technical cause and the causal link with the operation of the system, the burden of proof becomes practically impossible to fulfill. This situation directly affects the principle of equality of procedural weapons, since entities using AI systems, such as financial institutions or public authorities, have a monopoly on relevant information and have superior technical resources to support their own arguments.

Algorithmic auditing is increasingly emerging as an indispensable tool for verifying the compliance of artificial intelligence systems with applicable legal requirements, in particular in terms of transparency, non-discrimination and accountability¹³. The audit is intended to assess not only the accuracy and technical performance of the algorithm, but also its impact on fundamental rights and compatibility with relevant legal principles. Depending on the entity that carries it out, the audit can be internal, external or institutional.

Internal audit is carried out by the developer of the system or the organization that uses it. It has the advantage of full access to the algorithm architecture, training data and technical parameters. However, internal audit is vulnerable to conflicts of interest as the assessed entity has a direct interest in maintaining a

¹² Joshua A. Kroll, "The Fallacy of Inscrutability," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376, no. 2133 (2018): 20180084, <https://doi.org/10.1098/rsta.2018.0084>.

¹³ Samuele Bertaina et al., "Fundamental Rights and Artificial Intelligence Impact Assessment: A New Quantitative Methodology in the Upcoming Era of AI Act," *Computer Law & Security Review* 56 (2025): 106101, <https://doi.org/10.1016/j.clsr.2024.106101>.

favourable image of its own system. For this reason, although useful as a preliminary verification mechanism, internal audit cannot substitute for independent control.

External audit is carried out by independent entities specializing in the analysis and evaluation of algorithmic systems. It is considered the most credible form of audit, as it guarantees a high level of impartiality and objectivity. External audit is essential in areas with a high impact on fundamental rights, such as lending, recruitment, criminal justice and digitised public services. In these situations, independent verification of algorithmic models is a condition for ensuring effective control over automated decisions.

Institutional audit can be carried out by specialised public authorities such as data protection authorities, administrative control institutions or newly established entities with a supervisory role for artificial intelligence. This form of audit is particularly important because it allows the exercise of public control over AI systems used by state institutions or in areas of general interest. In the absence of such public scrutiny, the risk of abuse or non-transparent administrative decisions increases significantly.

Judicial technical expertise must, in turn, be deeply adapted to meet the specificity of artificial intelligence algorithms. Traditional expertise is not sufficient in disputes involving machine learning models, neural networks or probabilistic scoring systems. It is necessary to develop a new type of expertise, based on standardised methodologies and clear criteria for assessing performance, bias and transparency¹⁴. Judicial experts must also have interdisciplinary skills, integrating knowledge from law, Computer Science, Statistics and data ethics.

The case *Loomis v. Wisconsin*, settled in the United States, represents one of the first times that a court was called to rule on the use of algorithms in criminal proceedings and on their compatibility with the constitutional rights of the defendant¹⁵. In this case, the court considered the legality of using the risk scores generated by the Compass system in determining the penalty.

The Wisconsin Supreme Court allowed the use of the algorithmic system, considering that it can be a complementary tool in the process of individualizing punishment. However, the court stressed the need for additional safeguards, in particular the prohibition on using the risk score as the sole basis for determining the penalty, as well as the obligation to inform the defendant of the limitations of the instrument. The *Loomis* case opened the global debate over the right to transparency in criminal justice and the need to subject algorithmic systems to rigorous judicial review.

In Europe, the SyRI case in the Netherlands has had a remarkable impact on how artificial intelligence systems can be used in the administration of public

¹⁴ Alexander Boer, Léon de Beer, and Frank van Praat, "Algorithm Assurance: Auditing Applications of Artificial Intelligence," in *Advanced Digital Auditing*, ed. Egon Berghout et al. (Cham: Springer, 2023), 149–183, https://doi.org/10.1007/978-3-031-11089-4_7.

¹⁵ *State v. Loomis*, 2016 WI 68, 371 Wis. 2d 235, 881 N.W.2d 749 (Wis. 2016).

services¹⁶. The SyRI system was created to automatically detect fraud in the field of social assistance by correlating a large volume of personal data from various public databases. The Hague tribunal decided to ban this system, considering that the interference in privacy was disproportionate to the intended purpose. The court held that the system operated in an opaque manner, did not provide adequate safeguards against the risk of discrimination and did not allow data subjects to understand how suspicions of fraud were generated. With this decision, the SyRI case reinforced the importance of the proportionality test as a fundamental legal instrument in assessing the use of AI systems in the public sector.

7. Regulatory Perspectives and Best Practices

France has adopted one of the most advanced European approaches to algorithmic transparency through the law for a Digital Republic of 2016¹⁷. This law explicitly enshrines the right of the person to receive an explanation of administrative decisions generated or assisted by IT systems. The French regulation obliges public administrations to disclose the essential criteria used within the algorithms, including the relative importance of each parameter, the general logic of processing and the data envisaged. This legal obligation is not just a technical tool, but a fundamental guarantee for the protection of citizens' rights, since it allows to verify the legality of administrative acts and challenge them effectively. The French model subsequently influenced other European states, being considered a reference point in the development of regulations on automated decisions in the public sector.

Canada has adopted a structured approach through the automated decision-making directive, an innovative regulatory tool that classifies algorithmic systems according to the level of impact on the rights and interests of individuals¹⁸. The directive imposes differentiated requirements for transparency, audit and risk assessment, depending on the risk category assigned to the system. In the case of high-impact systems, administrations are required to carry out impact assessments on algorithms, document decision-making logic and ensure clear mechanisms for challenge. This gradual approach allows regulatory adaptation to technological complexity and guarantees protection commensurate with the risks generated.

Romania adopted the National Strategy on Artificial Intelligence 2024–

¹⁶ The Hague District Court, *SyRI*, ECLI:NL:RBDHA:2020:1878, February 5, 2020, <https://uitspraken.rechtspraak.nl/details?id=ECLI:NL:RBDHA:2020:1878>.

¹⁷ France, *Code des relations entre le public et l'administration*, art. L311-3-1, created by Law No. 2016-1321 of October 7, 2016, art. 4, https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000033205535/.

¹⁸ Government of Canada, Treasury Board of Canada Secretariat, *Directive on Automated Decision-Making*, effective April 1, 2019, <https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592>.

2027, but still faces the need to develop detailed legal and institutional mechanisms for managing the challenges generated by artificial intelligence systems in the public and private sectors¹⁹. In the absence of a clear and coherent regulatory framework, the risks associated with automated decisions can directly affect fundamental rights, administrative efficiency and public trust in authorities. Therefore, an essential first step is the adoption of a specific law on the use of artificial intelligence systems in public administration. This law should expressly establish transparency obligations, define minimum criteria for explicability and impose clear control, audit and reporting mechanisms. The rules should also provide for the conditions under which algorithmic decisions can produce legal effects, as well as procedures for challenging them.

A second fundamental element of an effective national strategy is the establishment of a specialised authority to oversee the use of artificial intelligence. This entity could operate either within the National Supervisory Authority for Personal Data Processing or as an independent institution, following the model of some European states that have created dedicated agencies. Such an authority would have the role of monitoring the implementation of algorithmic systems in the public and private sector, carrying out technical and legal audits, issuing best practice guidelines and responding to complaints from affected persons.

It could also have powers to authorise high-risk systems, ensuring that they are only implemented under the conditions of rigorous prior assessment.

The implementation of such an institutional structure would allow Romania to effectively manage the risks associated with automated decisions, ensuring an adequate balance between technological innovation and the protection of fundamental rights. In addition, the existence of a specialized authority would facilitate the alignment of the national regulatory framework with emerging European and international standards in the field of artificial intelligence regulation.

Public institutions must systematically develop their internal capacities for the evaluation, monitoring and supervision of the artificial intelligence systems they use. The implementation of these technologies cannot be left solely to external providers, as decisions generated by algorithms produce direct legal effects on citizens and imply a clear institutional liability. In this regard, it is necessary to train administrative staff in the field of digital technologies, through continuous training programs covering issues such as the principles of operation of algorithms, associated risks and transparency requirements. Institutions should also develop dedicated internal procedures for the pre-evaluation of AI systems, monitoring their operation and managing any algorithmic incidents. The implementation of periodic audit mechanisms, carried out either internally or by independent entities, is an essential condition for maintaining effective control

¹⁹ Government of Romania, Government Decision No. 832 of July 11, 2024, approving the National Strategy on Artificial Intelligence 2024–2027, *Monitorul Oficial* no. 730 (July 25, 2024), <https://legislatie.just.ro/Public/DetaliiDocumentAfis/285931>.

over these technologies and for compliance with the principles of legality and proportionality.

Academia plays a central role in this transformation, as training specialists able to understand both the legal and technological dimensions of artificial intelligence is essential for the coherent functioning of the legal system. Universities must integrate into their legal study programs courses dedicated to the interaction between law and technology, in which students acquire basic notions regarding algorithms, data, explainability and relevant international regulations. At the same time, it is necessary to create specialized master's programs in technology law, which train experts capable of managing the legal challenges generated by digitization. The development of interdisciplinary research, at the intersection of law, Informatics, ethics and data science, can contribute to the theoretical foundation of appropriate legislative solutions and to the strengthening of the national legal culture in the field of artificial intelligence.

8. Conclusions

The analysis highlights the fact that the integration of artificial intelligence systems in decision-making processes causes profound transformations in the application and guarantee of the principle of decision-making transparency. Algorithmic opacity is not limited to a technical difficulty, but becomes a genuine legal and Democratic problem. It affects the ability of individuals to understand the reasons for a decision, the ability of courts to control its legality and the legitimacy of the decision-making process as a whole. Therefore, its management requires a holistic approach, combining legal regulation, technical tools and institutional supervision.

At the same time, recent developments in the field of explainable artificial intelligences demonstrate that algorithmic transparency is not a technological utopia, but an achievable goal. Explicability models, either natively interpretable, post hoc or hybrid, provide concrete opportunities for reconciling algorithmic performance with legal requirements for reasoning of acts and verifiability of decisions. In order for these opportunities to become effectively functional, it is necessary to develop clear legal standards, adapted both to the needs of justice and to technological limitations. These standards need to be strengthened through effective oversight mechanisms capable of monitoring the deployment and use of AI systems in critical sectors.

The challenges posed by the use of artificial intelligence cannot be solved exclusively by traditional legal instruments. The technological complexity of algorithms requires close collaboration between lawyers, computer scientists, ethics experts and data science specialists. Only through an interdisciplinary approach can viable solutions be developed, capable of responding to emerging risks and responsibly harnessing the benefits of technology.

A strategic element of the regulation is the extension of the principle of

privacy by design to a genuine transparency by design. This implies the integration of transparency requirements from the design stage of algorithmic systems, so that explicability is not a later addition, but a structural feature²⁰. A proactive approach of this type is much more effective than subsequent attempts to address transparency deficiencies and significantly reduces the risks of discrimination or arbitrariness.

Future research directions should focus on developing robust metrics for measuring algorithmic transparency and evaluating the effectiveness of various explicability models. At the same time, it is necessary to deepen the analysis on the societal impact of the use of AI systems, including their effects on public trust, social cohesion and access to justice. The development of standardised tools for assessing societal impacts would contribute to responsible and well-founded regulation.

From a regulatory perspective, an essential priority is the effective implementation of the AI Act and the adaptation of national legislation to the specific challenges identified. Romania has the opportunity to position itself as a regional leader in developing a balanced legal framework for AI, capable of protecting fundamental rights without discouraging technological innovation. Such an approach can turn technology into a factor of social progress, while strengthening the values of the rule of law.

Bibliography

I. Books and Articles

1. Arnesen, Lars. "Comparing the Right to Information about Automated Decision-Making to the New Right to Explanation." *International Data Privacy Law* 16, no. 2 (2026). <https://doi.org/10.1093/idpl/ipag008>.
2. Bertaina, Samuele, Ilaria Biganzoli, Rachele Desiante, Dario Fontanella, Nicole Inverardi, Ilaria Giuseppina Penco, and Andrea Claudio Cosentini. "Fundamental Rights and Artificial Intelligence Impact Assessment: A New Quantitative Methodology in the Upcoming Era of AI Act." *Computer Law & Security Review* 56 (2025): 106101. <https://doi.org/10.1016/j.clsr.2024.106101>.
3. Binns, Reuben. "Algorithmic Accountability and Public Reason." *Philosophy & Technology* 31, no. 4 (2018): 543–556. <https://doi.org/10.1007/s13347-017-0263-5>.
4. Boer, Alexander, Léon de Beer, and Frank van Praat. "Algorithm Assurance: Auditing Applications of Artificial Intelligence." In *Advanced Digital Auditing*, edited by Egon Berghout, Rob Fijneman, Lindy Hendriks, Matthijs de Boer, and Bas Jan Butijn, 149–183. Cham: Springer, 2023. https://doi.org/10.1007/978-3-031-11089-4_7.

²⁰ Eveline H. van Beem, "Explainable throughout the AI Lifecycle: The Evolution of the Explainability Provisions in the EU AI Act," *International Review of Law, Computers & Technology* (2026): 1–35, <https://doi.org/10.1080/13600869.2026.2668320>.

5. Gunning, David. "DARPA's Explainable Artificial Intelligence (XAI) Program." In *Proceedings of the 24th International Conference on Intelligent User Interfaces (IUI '19)*, ii. New York: ACM, 2019. <https://doi.org/10.1145/3301275.3308446>.
6. Kroll, Joshua A. "The Fallacy of Inscrutability." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376, no. 2133 (2018): 20180084. <https://doi.org/10.1098/rsta.2018.0084>.
7. Lipton, Zachary C. "The Mythos of Model Interpretability." *Communications of the ACM* 61, no. 10 (2018): 36–43. <https://doi.org/10.1145/3233231>.
8. Lu, Sylvia. "Algorithmic Opacity, Private Accountability, and Corporate Social Disclosure in the Age of Artificial Intelligence." *Vanderbilt Journal of Entertainment and Technology Law* 23, no. 1 (2020): 99–159. <https://scholarship.law.vanderbilt.edu/jetlaw/vol23/iss1/3>.
9. Selbst, Andrew D., and Julia Powles. "Meaningful Information and the Right to Explanation." *International Data Privacy Law* 7, no. 4 (2017): 233–242. <https://doi.org/10.1093/idpl/ixp022>.
10. van Beem, Eveline H. "Explainable throughout the AI Lifecycle: The Evolution of the Explainability Provisions in the EU AI Act." *International Review of Law, Computers & Technology* (2026): 1–35. <https://doi.org/10.1080/13600869.2026.2668320>.

II. Legislation and Official Sources

1. European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council of April 27, 2016, on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). *Official Journal of the European Union* L 119 (May 4, 2016): 1–88. <https://eur-lex.europa.eu/eli/reg/2016/679/oj>.
2. European Union. Regulation (EU) 2024/1689 of the European Parliament and of the Council of June 13, 2024, laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the European Union* L 2024/1689 (July 12, 2024). <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>.
3. France. *Code des relations entre le public et l'administration*, art. L311-3-1, created by Law No. 2016-1321 of October 7, 2016, art. 4. https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000033205535/.
4. Government of Canada, Treasury Board of Canada Secretariat. *Directive on Automated Decision-Making*. Effective April 1, 2019. <https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592>.
5. Government of Romania. Government Decision No. 832 of July 11, 2024, approving the National Strategy on Artificial Intelligence 2024–2027. *Monitorul Oficial* no. 730 (July 25, 2024). <https://legislatie.just.ro/Public/DetaliiDocumentAfis/285931>.
6. Government of Romania. Government Emergency Ordinance No. 57 of July 3, 2019, regarding the Administrative Code. *Monitorul Oficial* no. 555 (July 5, 2019). <https://legislatie.just.ro/Public/DetaliiDocumentAfis/249654>.
7. Romania. Law No. 190 of July 18, 2018, on measures for implementing

Regulation (EU) 2016/679. *Monitorul Oficial* no. 651 (July 26, 2018). <https://legislatie.just.ro/Public/DetaliiDocumentAfis/203151>.

III. Cases and Decisions

1. Čiapas v. Lithuania, no. 4902/02. European Court of Human Rights, November 16, 2006. <https://hudoc.echr.coe.int/eng?i=001-78053>.
2. The Hague District Court. *SyRI*, ECLI:NL:RBDHA:2020:1878, February 5, 2020. <https://uitspraken.rechtspraak.nl/details?id=ECLI:NL:RBDHA:2020:1878>.