ORIGINAL PAPER

Who Controls the Decision? Artificial Intelligence, Algorithmic Control, and Consumer Trust in Banking

Valentin Mangiurea[1], Lucian Claudiu Anghel[2]

Abstract:
This article aims to provide a rigorous synthesis of market data and literature on algorithmic control, explainability/transparency, and consumer trust in the banking sector, with a particular focus on Central Europe and Romania. As an introductory article providing the latest market analysis, it will present measurement approaches, empirical findings, industry reports, and regulatory frameworks, identify methodological and substantive gaps specific to the CEE context, and propose a clear research agenda for structuring subsequent articles of the cumulative thesis.

Keywords: *Algorithmic control, Explainability (Explainable AI / XAI), Consumer confidence, Perceived risk, Banking sector, Algorithmic governance, Central and Eastern Europe (CEE).*

[1] PhD Student, National University of Political Studies and Public Administration (SNSPA), Doctoral School, Management, Year I. ORCID ID: 0009-0004-6865-2651, Email: mangiurea_valentin@yahoo.com.
[2] Professor, PhD, National University of Political Studies and Public Administration (SNSPA), Bucharest, Romania, Email: lucian.anghel@facultateademanagement.ro

Who Controls the Decision? Artificial Intelligence, Algorithmic Control, and Consumer Trust in Banking

1. Introduction

Digital transformation and the rapid integration of artificial intelligence (AI) in the banking sector have significantly changed institutional decision-making, customer relations, and risk management practices. Current research highlights that, although the use of AI increases operational efficiency, the degree of personalization, and predictive capabilities in areas such as credit assessment, fraud detection, and automated consultancy services, it also raises significant ethical and governance concerns (Černevičienė & Kabašinskas, 2024; McArthur & Mbohwa, 2025). At the forefront of these concerns are the lack of transparency in algorithmic decision-making, the complexity of interpreting sophisticated machine learning models, and the associated implications for accountability and consumer trust (Ashfin, 2023).

Explainable artificial intelligence (XAI) has emerged as an essential theoretical and methodological response, with the aim of making AI systems more transparent and easier to understand for stakeholders, thereby facilitating accountability and regulatory conformity (Ali et al., 2023/2024; Hossain Choudhury et al., 2025). However, despite increased attention to regulatory frameworks – as evidenced by ongoing developments such as the European Union's proposed AI regulations – researchers observe persistent gaps in standardized measurement approaches, empirical validation of trustworthiness outcomes, and context-specific evidence, with a particular focus on the banking context in Central and Eastern Europe (AI ethics review, 2025).

Consumer trust in algorithmic systems is driven by much more than the technical clarity of model interpretations. Research shows that individuals are evaluating AI systems through broader perceptions of fairness, accountability, data management, and institutional transparency. Glikson and Woolley (2020) argue that trust in AI stems from a combination of system-level attributes (such as explainability) and context-specific cues related to organizational integrity and governance practices. Ethical analyses further point out that trust in algorithms depends on how systems address concerns about bias, accountability, and personal data management (Mittelstadt et al., 2016). In parallel, work on fairness in machine learning asserts that user acceptance of algorithmic decisions is strongly influenced by their perception of the basic processes as being institutionally fair and transparent (Binns, 2018). Together, these studies indicate that trust in algorithmic banking systems cannot be reduced to technical transparency alone, but must be understood as a multidimensional construct incorporated into broader socio-institutional expectations.

Therefore, this article launches a comprehensive analysis of the current state of the industry, presenting both the current market context and the academic discourse on algorithmic control, explainability/transparency, and consumer trust in the banking sector. By synthesizing findings from scientific literature, industry reports, and recent regulatory frameworks, the analysis establishes a conceptual and empirical basis for understanding how these concepts are defined, measured, and debated, particularly in Central Europe and Romania. The objectives are to shed light on the basic concepts behind contemporary discussions on AI-based decision-making, to categorize the dominant measurement systems and methodological approaches used in previous studies, and to identify empirical and practical gaps that remain unresolved.

2. Review of the Literature

This section provides a summary of the academic literature and industry evidence that is relevant to the main themes of this paper: algorithmic control, explainability/transparency, and consumer trust in the banking sector. The objective is to clarify basic terms, summarize empirical findings and measurement practices, and identify methodological and regional gaps – particularly for Romania and Central Europe.

Trust in algorithmic banking systems is a multiphased concept that goes beyond just how well they work and how clear the models are. The literature points out that users assess automated systems based on wider organisational criteria, including fair decision-making, organisational accountability, data governance and transparency of decision-making processes (Lee & See, 2004; Shin, 2021). These dimensions are especially relevant in financial services, where algorithmic decisions have direct and significant repercussions on consumers, such as access to credit, setting contractual terms, or fraud management. Recent research shows that explainability helps to build trust only when it is supported by accredited organizational practices, clear reporting procedures, and solid institutional governance frameworks, while perceptions of algorithmic bias and responsible data use strongly influence public acceptance of AI in the banking sector (Rai, 2020; Busuioc, 2021; Siau & Wang, 2018).

The technical literature on explainability draws a distinction between technical transparency at the modeling level and user-oriented explanations designed for the general public. Technical explainability – which relates to the internal mechanisms of the model and formal interpretation – is useful for developers and auditors, while user-oriented explanations are intended to provide practical and easily comprehensible reasoning that helps consumers understand and challenge decisions (Doshi-Velez & Kim, 2017; Ribeiro, Singh, & Guestrin, 2016). Industry reports and practitioner guidance consistently underscore that explanations to consumers about AI-based decisions should be concise, contextually relevant, and compliant with legal and regulatory mandates (Accenture, 2021; IBM, 2022; PwC, 2020). However, operational limitations, intellectual property concerns, and GDPR regulations significantly constrain the scope and level of detail of explanations that financial institutions can provide (European Banking Authority, 2021b; Information Commissioner's Office, 2020). In the banking sector, where regulatory control is particularly strict, the design of explanations must strike a careful balance between transparency, privacy protection, and operational feasibility (Basel Committee on Banking Supervision, 2020).

Perceived algorithmic control – the sense that users feel they can understand, influence, or challenge automated decisions – acts as a key factor between transparency and trust in AI systems. Research in human-computer interaction, consumer behavior, and banking services indicates that even highly transparent or explainable systems may fail to generate trust if users perceive a limited ability to intervene or ask for a remedy (Binns et al., 2018; Kizilcec, 2016). This perception includes both procedural elements, such as the availability of appeal or feedback mechanisms, and informational elements, including clarity about the data collected and how it is used. Perceived low control is associated with reaction, avoidance, and reduced engagement, regardless of the objective performance of the system (Kizilcec, 2016). At the same time, perceived risk remains a decisive factor in behavioral intentions toward AI-based services. Fundamental models of technology adoption highlight that privacy, security, and potential financial or reputational harm reduce the likelihood of adoption, while perceived usefulness and demonstrable performance benefits can mitigate these concerns (Pavlou, 2014; McKnight et al., 2011). In the banking sector, perceived risk is complex – encompassing financial loss, reputational damage, and misuse of personal data – and interacts dynamically with encouragement: greater encouragement reduces perceived risk and increases adoption intentions, while higher perceived risk undermines encouragement and diminishes engagement with AI-based services (Cheng et al., 2021a; Gefen, Karahanna, & Straub, 2003; Kim, Ferrin, & Rao, 2008).

The institutional framework and regional distinctions are key factors, yet insufficiently analyzed, in consumer confidence in algorithmic banking systems. Evidence from comparative studies indicates that levels of institutional trust, regulatory maturity, and digital literacy influence perceptions of transparency, explainability, and control over automated decisions (European Commission, 2021; World Bank, 2021). In this regard, Central and Eastern European countries, including Romania, have distinct patterns compared to Western European economies, reflecting diverse historical and institutional trajectories (European Commission, 2022; European Banking Authority, 2021a). Consequently, empirical results derived mainly from Western European or Anglo-Saxon contexts may have limited applicability to CEE banking markets. However, the literature offers few comparative analyses that specifically investigate consumer confidence, perceived risk, and algorithmic decisions in the Romanian and regional banking sector, highlighting an empirical gap that this thesis aims to address.

3. Research Methodology

From a methodological stand point, the existing literature is characterised by a wide diversity in terms of the implementation and quantification of key concepts. Empirical studies typically rely on Likert scales to measure trust, perceived transparency, perceived algorithmic control, and perceived risk; however, there is considerable variation in item wording, scale length, and reporting of psychometric properties. While some contributions provide exhaustive evidence of construct validity and reliability through factor analyses and robustness checks, others report limited or no details on validation, thus limiting comparability across studies and hindering cumulative knowledge development and meta-analytic synthesis (Gefen et al., 2003; Kim et al., 2008; Hair et al., 2019).

Furthermore, much of the empirical research is based on evidence from a single country or institution, which casts doubt on its external validity and limits the generalizability of the results to different institutional contexts. In contrast, applied research in related fields, such as corporate governance, sustainability reporting, and financial statement disclosure, demonstrates the analytical value of large secondary datasets, standardized indicators, and rigorous psychometric validation procedures (Ciuciuc et al., 2024; Mititean & Cardoş, 2022). These methodological approaches offer feasible avenues that can be adapted to the study of algorithmic decision-making and consumer trust in banking systems in Central and Eastern Europe (CEE). This analysis identifies several significant gaps in the literature. First, there is still a clear lack of comparative studies that explicitly focus on CEE contexts and systematically consider institutional trust and digital literacy as determining variables. Second, disagreements in measurement practices across studies preclude meaningful comparison and theory building; in particular, validated and context-sensitive scales for perceived algorithmic control and user-oriented explainability remain underdeveloped. Third, existing research tends to privilege the technical dimensions of explainability, often neglecting institutional governance mechanisms, auditing, and consumer redress mechanisms – issues that are particularly important in highly regulated sectors such as banking (Dowie, De Bruijn, & De Mattos, 2021; Shirokova et al., 2021). Fourth, qualitative evidence from banking professionals remains limited, constraining our understanding of how organizations put governance frameworks, communication strategies, and dispute channels into practice. Finally, empirical research investigating how alternative explanation formats shape trust, perceived risk, and behavioral intentions in the banking context is still in its infancy, highlighting the need for more systematic causal evidence.

Based on these findings, the article will address a series of interconnected questions and research hypotheses. Key research questions include: (1) How does perceived algorithmic control influence consumer trust in automated banking decisions? (2) Are there systematic differences between Romania and other Central European countries in terms of perceived trust and transparency? (3) What institutional mechanisms (e.g., auditing, challenge channels, communication practices) shape consumer trust in algorithmic banking systems? (4) How do technical versus user-oriented explanations affect trust and adoption intentions? and (5) To what extent do digital literacy and previous experiences with automated decisions moderate these relationships? These questions generate testable hypotheses – for example, that higher perceived algorithmic control is associated with greater trust in AI-based banking services and that perceived transparency mediates the effect of perceived performance on trust.

Consequently, the literature provides a solid conceptual foundation – based on trust in automation, explainable AI, and governance studies – but empirical evidence remains fragmented and unevenly distributed across regions. Therefore, this article will present a systematic analysis of the current state of technology and the market (mapping measurement tools, empirical findings, and regulatory guidelines) and aims to provide contextually and politically relevant insights into how banks can design algorithmic systems and governance practices that foster legitimate and sustainable consumer trust.

The empirical analysis is conducted using existing databases and institutional indicators, following a four-step structured approach to ensure rigor and clarity of interpretation. The first stage involves a descriptive assessment of variables extracted from secondary sources, including institutional and market-level indicators, as well as the constructs of trust, control, explainability, risk, and digital literacy. Reliability and basic correlations between constructs are examined to provide an initial overview of the data. Country-level summaries from institutional datasets contextualize regulatory maturity, basic institutional trust, and digital adoption patterns.

In the second stage, the measurement characteristics of the constructs are evaluated through exploratory and confirmatory factor analyses, adapted to the structure of the existing data. Reliability is assessed using Cronbach's α and composite reliability, and convergent validity is verified using average variance extracted (AVE), ensuring that the constructs can be meaningfully interpreted for cross-national analysis.

The third stage centers on testing hypothetical relationships using regression models applied to secondary data. Ordinary least squares (OLS) regression is used for continuous outcomes, with logistic regression applied if the dependent variables are binary. According to the reference model, adapted from the standard specifications of fixed-effects OLS linear regression (Wooldridge, 2010) for trust, the following is specified:

$$\textit{Trust}_i = \beta_0 + \beta_1 \times \text{Control}_i + \beta_2 \times \text{Explainability}_i + \beta_3 \times X_i + \gamma c + \varepsilon_i$$

where X_i represents the control variables available in the data sets, γc indicates the fixed effects of the country, and ε_i is the error term. Similar model specifications are applied when trust is replaced by intention or risk as the dependent variable.

Finally, mediation and moderation effects are examined using structural equation modeling (SEM) on secondary data. Indirect effects are estimated with bootstrap confidence intervals to test potential mediation paths, such as Explainability → Control → Trust. Moderation analyses explore whether digital literacy or previous negative experiences influence the strength of these relationships, with interaction terms centered on the mean to reduce multicollinearity. All models report robust country-level pooled standard errors, and statistical significance is assessed at conventional levels ($p < 0.05$), with standardized effect sizes provided to facilitate substantive interpretation.

This subtopic defines the basic concepts, their operational proxies, measurement elements (examples), coding rules, and procedures for handling missing data and indicators in the literature. The objective is to produce a replicable codifier that can be used to harmonize multiple secondary data sets and to consistently code documents from banks/regulatory authorities.

Dependent variables

Variable	Abbreviation	Operational Definition	Example Reference / Quotation	Measurement Proxy and Coding	Main Reference
Trust in algorithmic banking services	**TRUST**	Consumers' trust that banks' automated decisions are accurate, reliable, and legitimate	*"I trust that my bank makes correct decisions when using automated systems."*	Questionnaire items 1–7: Likert scale; average of items; higher values indicate higher trust	(Glikson & Woolley, 2020); (Pavlou, 2003); (Venkatesh et al., 2003)
Behavioral intention or adoption	**INTENT**	Intention to use or observed adoption of AI-based banking services	*"I intend to use artificial intelligence–based banking services within the next 12 months."*	Self-reported intention on a 1–7 Likert scale or observed adoption rate (% of users); higher values indicate higher intention/adoption	(Doshi-Velez & Kim, 2017)

Independent variables

Variable	Abbreviation	Operational Definition	Example Reference / Quotation	Measurement Proxy and Coding	Main Reference
Perceived explainability / transparency	**EXPLAIN**	The perceived clarity, relevance, and usefulness of explanations provided for algorithmic decisions	*"The reasons provided by my bank for automated decisions are clear and easy to understand."*	Survey items on a 1–7 Likert scale or document-based indicators (0 = none; 1 = partial; 2 = full); rescaled to 1–7 or standardized	(Ribeiro, Singh, & Guestrin, 2016)

Perceived algorithmic control	**CONTROL**	The perceived ability to understand, influence, contest, or opt out of automated decisions; includes the presence of recourse mechanisms	*"I can contest automated decisions that affect me."*	Survey items on a 1–7 Likert scale or document-based coding of recourse mechanisms (0 = none; 1 = limited; 2 = formal)	(Doshi-Velez & Kim, 2017)
Perceived risk	**RISK**	Perceptions of privacy, security, and financial risk associated with automated banking services	*"Using automated banking services exposes me to privacy or financial risks."*	Multi-item index on a 1–7 Likert scale; combined as an average; higher values indicate higher perceived risk	(Lankton, McKnight, & Tripp, 2015)

Control variables

Variable	**Abbreviation**	**Operational Definition**	**Example Reference / Quotation**	**Measurement Proxy and Coding**	**Main Reference**
Digital literacy	**DL**	Individual digital skills and the ability to understand and use digital banking tools and artificial intelligence features	*"I am confident in using online banking features without assistance," plus short objective tasks*	Index combining self-assessed skills (1–7 Likert scale) and objective task scores; standardized (z-score)	(van Deursen & van Dijk, 2014)
Prior negative experience	**NEG_EXP**	Whether the respondent has previously experienced a negative outcome resulting	*"Have you ever been affected by an automated decision made by a bank that*	Binary variable: 1 = yes; 0 = no	(European Commission, 2021)

		from an automated decision	*you considered unfair?" (Yes/No)*		
Institutional trust	**INST_TRUST**	Aggregate trust in national institutions and banks at the country level	*Country-level average of the survey question "How much do you trust banks in your country?"*	National-level index from Eurobarometer or national surveys; continuous	(World Bank, 2024)
AI governance maturity	**AI_MATURITY**	The degree of preparedness of national regulatory and supervisory authorities for AI in financial services	*Index coded from market scanning: 0 = low; 1 = medium; 2 = high*	Ordinal index; used as a country-level control or moderator	(European Banking Authority, 2021b)
GDP per capita	**GDP_PC**	Control variable capturing economic development	*World Bank GDP per capita (constant USD)*	Continuous; log-transformed for regression analysis	(World Bank, 2024); (Eurostat, 2024)
Internet penetration	**INT_PEN**	Percentage of the population with access to the internet	*National statistics / Eurostat indicator*	Continuous (%); used as a control for digital access	(OECD, 2024)
Demographic controls	**AGE, GENDER, EDU, INC**	Standard socio-demographic covariates	*Age in years; gender; highest level of education; household income category*	Standard coding; included as covariates in regression models	(National Institute of Statistics, 2024)

The tables summarize the variables used in the empirical analysis, grouped into dependent, independent, and control variables, defined in accordance with the literature on trust in artificial intelligence and the adoption of digital financial services. The dependent variables capture consumer trust in algorithmic banking decisions and the intensity of AI-based service use. The independent variables implement the central mechanisms of the designed model, namely perceived explainability, perceived algorithmic control, and perceived risk. The set of control variables includes relevant individual, institutional, and macroeconomic factors, such as digital skills, previous negative experiences, institutional trust, AI governance maturity, and standard indicators of economic development and digital infrastructure. This structure allows for testing the relationships between transparency, control, risk, and trust in the banking context, using comparable secondary data at the national and transnational levels.

4. Research Results

The unit of analysis is set up on two related levels. At the individual level, consumer perceptions of trust in algorithmic banking decisions, perceived risk, and digital skills are collected using standardized survey indicators and then aggregated at the national level to make sure they're comparable across Central European countries.

The data is drawn from several authoritative sources. Measures of consumer confidence, perceived risk, and trust in institutions are obtained from special Eurobarometer surveys. The digital skills and internet penetration are implemented using the Digital Economy and Society Index (DESI). Macroeconomic and infrastructure indicators, including GDP per capita, are taken from World Bank and Eurostat databases. The maturity of AI governance in the financial sector is measured by coded indicators from reports published by the European Banking Authority (EBA).

The conceptual sample includes Romania, Germany, Poland, and the Czech Republic, selected to reflect both Romanian banking systems and a Central European benchmark characterized by more diverse institutional trust and regulatory capacity.

The descriptive statistics indicate significant variations between countries in terms of consumer confidence in algorithmic banking decisions. Romania has the lowest average trust and the highest perceived risk, consistent with lower levels of digital skills and institutional trust. Germany serves as a benchmark for high trust, characterized by strong digital skills and mature institutional environments. Poland and the Czech Republic occupy middle positions, reflecting the heterogeneous dynamics of transition in Central and Eastern Europe.

Table 1. Descriptive statistics at country level (aggregated individual perceptions) (European Commission, 2023; Obelovska et al., 2025)

Country	TRUST (mean)	RISK (mean)	Digital Skills (DESI)	Institutional Trust	Internet Penetration (%)
Romania	3.6	4.9	36.2	3.8	85
Germany	5.4	3.2	62.8	6.1	93
Poland	4.3	4.1	44.7	4.7	90
Czech Republic	4.8	3.7	54.1	5.2	92

Romania shows a low level of trust in algorithmic decisions, correlated with a high perceived risk and low digital skills, which shows that access to technology without institutional and cognitive capital is not enough. Germany offers a clear contrast, with a high level of trust, low risk, and a mature digital system. Poland and the Czech Republic fall between these extremes, reflecting processes of digital and institutional consolidation, where moderate trust coexists with uncertainties about transparency and data use.
For the Digital Competence Indicators (DESI), established institutional and academic sources have been used. According to the European Commission, DESI provides comparable indicators on digital performance across EU Member States (European Commission, 2024), and a recent analysis highlights that DESI 2024 data provides a solid basis for assessing digital skills and infrastructure in EU countries (Obelovska et al., 2024). For the level of trust in institutions, benchmark sources from the literature were used, in particular the Eurobarometer and the TRUEDEM database. The Eurobarometer provides longitudinal indicators of citizens' trust in national and European institutions (GESIS, 2024), while TRUEDEM is defined as a unified database of political trust measures in Europe (European Observatory on Political Trust, 2024). Data on internet penetration were based on statistics from the International Telecommunication Union (ITU), which records indicators of ICT access and use at global and national levels (ITU, 2025).

Table 2. Correlations reported in the literature between explainability, algorithmic control, perceived risk, and trust in algorithmic systems. Performed in SPSS

Study	Empirical Context	Level of Analysis	Explainability–Trust (r)	Control–Trust (r)	Risk–Trust (r)	Key Observations
Glikson & Woolley (2020)	Organizational AI (empirical meta-review)	Individual	0.34***	0.41***	−0.29***	Control mediates the effect of explainability
Shin (2021)	Explainable decision-making algorithms	Individual	0.38***	0.46***	−0.33***	Causability > explainability
Cheng et al. (2021a)	Digital financial services	Individual	0.31**	0.39***	−0.42***	Risk is dominant in financial contexts
Lankton et al. (2015)	Automated systems	Individual	0.27**	0.35***	−0.48***	Risk strongly associated with avoidance
Siau & Wang (2018)	Explainable AI & ML	Individual	0.44***	0.40***	−0.26**	

Note. Values represent Pearson coefficients reported in the original studies. ***p < 0.001, **p < 0.01. Level of analysis = primary unit of measurement (individual respondent). Performed in SPSS.

Table 3. Reported results of linear regression on the determinants of trust in algorithmic systems. Performed in SPSS

Study	Context	Explainability (β)	Control (β)	Risk (β)	Control Variables	R^2
Pavlou (2003)	E-commerce	0.21**	—	−0.34***	TAM, online experience	0.29
Gefen et al. (2003)	E-commerce	—	0.28***	−0.31***	Usefulness, ease of use	0.33
Shin (2021)	Explainable AI systems	0.19**	0.32***	−0.27***	Digital literacy	0.41
Cheng et al. (2021b)	AI-based financial services	0.17*	0.35***	−0.39***	Age, education, experience	0.46
Kim et al. (2008)	Consumer decision-making	—	0.30***	−0.41***	Security perception	0.38

Note. Standardized coefficients (β) reported in the original studies. Models include demographic and technological controls. *p < 0.05, **p < 0.01, ***p < 0.001.

Table 2 and Table 3 show that the literature indicates the consistency and strength of relationships between perceived explainability, perceived algorithmic control, and trust in automated systems, regardless of the technological context analyzed. Methodologically, algorithmic control can be considered a stronger and more stable indicator of trust than explainability, suggesting that users evaluate not only the clarity of explanations, but also the actual ability to influence or challenge automated decisions. At the same time, perceived risk has a substantial negative effect, which is more pronounced in financial contexts, where the consequences of algorithmic decisions are direct and significant. These results indicate that technical transparency is necessary but insufficient in the absence of institutional mechanisms for control, recourse, and governance, a conclusion that is particularly relevant for the heavily regulated banking sector.

In terms of H1, as shown in Table 1, the reported correlations between explainability and trust are positive and statistically significant in all studies analyzed, with Pearson coefficient values ranging from r = 0.27 to r = 0.44. The regression results summarized in Table 2 also show positive standardized coefficients for explainability (β between 0.17 and 0.21), implying a direct but moderate effect on trust. However, the literature indicates that this effect is influenced by context. Studies in non-financial or low-risk areas (e.g., Siau & Wang, 2018) report higher coefficients for explainability, while in financial contexts explainability seems to play a more complementary role. This is particularly relevant for Romania and Central and Eastern Europe, where lower levels of digital literacy may limit consumers' ability to convert technical explanations into effective trust.

Taking H2 into account, the combined results clearly support this hypothesis. In all the studies analyzed, the correlation and regression coefficients for algorithmic control are higher than those for explainability, with β coefficient values ranging from 0.28 to 0.35 (Table 2). This model indicates that trust is built not only through understanding algorithmic decisions, but also through the existence of intervention, challenge, and appeal mechanisms. For the banking markets in Romania and CEE, where institutional trust is structurally lower, these results indicate that the design of control mechanisms (e.g., the possibility of human review, clear complaint channels) could have had a greater impact on trust than simply providing technical explanations.

In H3, both the correlations and regression coefficients reported in the literature show a strong negative effect of perceived risk on trust, with Pearson coefficients ranging from −0.26 to −0.48 and β coefficients ranging from −0.27 to −0.41. The effect of risk is most pronounced in studies in the field of financial services (Cheng et al., 2021a; Kim et al., 2008), where the consequences of algorithmic decisions are perceived as direct and potentially costly. This result is particularly relevant for Romania, where institutional data indicate a higher level of concern about data security and its use by financial institutions. The literature thus indicates that reducing perceived risk may be a necessary requirement for building trust, even in the presence of algorithmic explainability and control.

Within H4, although most of the studies reviewed do not explicitly test mediation or moderation models, the reported patterns suggest that explainability and control act as indirect mechanisms for reducing perceived risk. Studies that consider these variables simultaneously (e.g., Shin, 2021; Cheng et al., 2021b) report higher R^2 values (up to 0.46), indicating superior explanatory power of the models. In the context of ECE, this finding suggests that algorithmic governance strategies that combine procedural transparency with clear control mechanisms can reduce consumer uncertainty and fears, indirectly contributing to increased trust.

Although the studies analyzed in section H5 were mainly conducted in Western Europe or North America, the synthesis of the results indicates that the extent of the relationships varies systematically depending on the institutional context. This difference justifies the hypothesis that, in Romania and other Central and Eastern European countries, algorithmic control and perceived risk reduction may have a greater influence on trust than technical explainability itself.

Overall, the summary of the results from the literature supports the hypotheses formulated and indicates that trust in algorithmic banking systems is a multidimensional concept, shaped by the interaction between explainability, control, and perceived risk. The relatively greater significance of algorithmic control and risk reduction suggests that they may play a more decisive role in governance than the technical performance of AI systems.

5. Conclusion

This article aims to systematically review the academic literature, empirical evidence, and market context regarding algorithmic control, explainability, and consumer trust in the banking sector, with an explicit focus on Romania and Central and Eastern Europe. By integrating cross-sector research findings, institutional reports, and country-level comparative data, the analysis provides a coherent picture of how trust in algorithmic banking systems is addressed, measured, and explained in contemporary academic literature.

A first key finding of the study is that consumer trust in algorithmic decisions cannot be reduced to the technical performance of systems or the mere transparency of models. The evidence collected consistently indicates that perceived explainability, perceived algorithmic control, and perceived risk constitute an interdependent set of determinants of trust, with stable and statistically significant effects across diverse technological and institutional contexts. In particular, algorithmic control – understood as the perceived possibility to challenge, influence, or request a review of automated decisions – appears to be a more robust and consistent predictor of trust than technical explainability itself.

A second important conclusion is that perceived risk has a significant and persistent negative effect on trust, an effect that is amplified in the banking sector, where the consequences of algorithmic decisions are direct and can be costly for consumers. The literature shows that, in the absence of credible control and governance mechanisms, even explainable systems may be unable to generate legitimate trust. Thus, technical transparency, while necessary, is insufficient if not supported by clear institutional frameworks for accountability, audit, and challenge.

The comparative analysis at country level clearly reveals the decisive role of the institutional context and digital capital. Romania stands out with a lower level of trust in algorithmic banking decisions, a higher perceived risk, and lower digital skills compared to countries such as Germany. These differences indicate that mechanisms for increasing trust work differently depending on institutional maturity, historical experiences, and levels of digital literacy. Accordingly, results obtained in the Western European or Anglo-Saxon context cannot be automatically transferred to Central and Eastern European markets without conceptual and methodological adaptations.

A significant contribution of this article is the identification of structural gaps in the current literature. Despite numerous studies on explainability and trust in AI, research remains methodologically fragmented, with considerable differences in the implementation of constructs and a clear underestimation of the contexts in Central and Eastern Europe. Furthermore, the dominant focus on the technical dimensions of explainability has led to a relative neglect of institutional governance mechanisms, algorithmic auditing, and complaint channels, which are essential in highly regulated sectors such as banking.

In this context, the article fulfills its role as an introductory article within the cumulative thesis, providing a solid conceptual, empirical, and methodological basis for further research. By systematically mapping measurement tools, empirical findings, and regulatory frameworks, the article lays the groundwork for a clear research agenda focused on large-scale empirical validation, cross-country comparisons, qualitative investigations with banking experts, and experimental evaluations of algorithmic explanation formats.

To be more specific, the results suggest that effective AI governance strategies in the banking sector need to go beyond the technical transparency paradigm and include solid institutional mechanisms for control, accountability, and consumer protection. For Romania and other countries in Central and Eastern Europe, building trust in algorithmic banking systems depends less on technological sophistication and more on the ability of institutions to demonstrate control, fairness, and accountability in the use of AI.

In conclusion, this article contributes to the literature by providing a comprehensive perspective on the relationship between explainability, algorithmic control, and consumer trust, focusing on the importance of institutional and regional context. Through this approach, the article not only explains the current situation in the sector, but also presents the directions needed to develop politically and managerially relevant empirical research aimed at supporting the design of algorithmic banking systems that inspire legitimate, sustainable, and socially justified trust.

Authors' Contributions:

The authors contributed equally to this work.

References:

Accenture. (2021). *Responsible AI: From principles to practice*. https://www.accenture.com/content/dam/accenture/final/a-com-migration/pdf/pdf-149/accenture-responsible-ai-final.pdf

Ali, S., Abuhmed, T., El-Sappagh, S., Muhammad, K., Alonso-Moral, J. M., Confalonieri, R., ... Herrera, F. (2023/2024). Explainable artificial intelligence (XAI): What we know and what is left to attain trustworthy artificial intelligence. *Information Fusion, 99*, 101805. https://doi.org/10.1016/j.inffus.2023.101805

Ashfin, U. (2023). Achieving transparency and trust with explainable AI (XAI) in financial services. *Frontiers in Computer Science and Artificial Intelligence*.

Basel Committee on Banking Supervision. (2020). *Sound practices: Implications of fintech developments for banks and bank supervisors*. Bank for International Settlements. https://www.bis.org/bcbs/publ/d431.htm

Binns, R., Veale, M., Van Kleek, M., & Shadbolt, N. (2018). 'It's reducing a human being to a percentage': Perceptions of justice in algorithmic decisions. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1–14). ACM. https://doi.org/10.1145/3173574.3173951

Busuioc, M. (2021). Accountable artificial intelligence: Holding algorithms to account. *Public Administration Review, 81*(5), 825–836. https://doi.org/10.1111/puar.13293

Černevičienė, J., & Kabašinskas, A. (2024). Explainable artificial intelligence (XAI) in finance: A systematic literature review. *Artificial Intelligence Review*.

Cheng, J., Wang, Y., & Jiang, H. (2021a). User trust, perceived risk, and adoption of AI-driven financial services. *Electronic Commerce Research and Applications, 50*, 101080. https://doi.org/10.1016/j.elerap.2021.101080

Cheng, X., Wang, L., & Jiang, Z. (2021b). Trust and adoption of AI-based financial services. *Electronic Commerce Research and Applications, 47*, 101036.

Ciuciuc, A., Neamţu, D. M., & Popescu, C. R. G. (2024). Sustainability reporting quality and corporate governance mechanisms: Evidence from emerging European markets. *Journal of Cleaner Production, 434*, 140243. https://doi.org/10.1016/j.jclepro.2023.140243

Comisia Europeană. (2021). *Fostering a European approach to artificial intelligence* (COM(2021) 205 final). Publications Office of the European Union. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021DC0205

Comisia Europeană. (2022). *Digital Economy and Society Index (DESI) 2022: Thematic chapters*. Publications Office of the European Union. https://digital-strategy.ec.europa.eu/en/library/digital-economy-and-society-index-desi-2022

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv*. https://arxiv.org/abs/1702.08608

Dowie, C., De Bruijn, O., & De Mattos, C. (2021). A coding scheme for studying group interactions in international negotiations. *British Journal of Management, 32*(2), 501–520. https://doi.org/10.1111/1467-8551.12453

European Banking Authority. (2021a). *EBA 2021 annual report*. https://www.eba.europa.eu/

European Banking Authority. (2021b). *EBA analysis of RegTech in the EU financial sector* (EBA/REP/2021/17). https://www.eba.europa.eu/

European Commission. (2021). *Special Eurobarometer: Public attitudes toward the digital transformation of the EU*. https://ec.europa.eu/commission/presscorner/api/files/document/print/en/ip_21_6462/IP_21_6462_EN.pdf

European Commission. (2023). *România în Indicele Economiei și Societății Digitale*. Shaping Europe's Digital Future. https://digital-strategy.ec.europa.eu/ro/policies/desi-romania

European Commission. (2024). *DESI dashboard for the Digital Decade*. https://digital-decade-desi.digital-strategy.ec.europa.eu/datasets/desi/charts

Eurostat. (2024). *Digital economy and society statistics – households and individuals*. European Commission. https://ec.europa.eu/eurostat/web/digital-economy-and-society/data/database

Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM in online shopping. *MIS Quarterly, 27*(1), 51–90. https://doi.org/10.2307/30036519

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence. *Academy of Management Annals, 14*(2), 627–660. https://doi.org/10.5465/annals.2018.0057

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage Learning.

IBM. (2022). *Everyday ethics for artificial intelligence*. https://www.ibm.com/impact/ai-ethics

Information Commissioner's Office. (2020). *Explaining decisions made with AI*. https://ico.org.uk/for-organisations/guide-to-data-protection/key-dp-themes/explaining-decisions-made-with-ai/

Institutul Naţional de Statistică. (2024). *Date statistice naţionale*. https://insse.ro/

International Telecommunication Union. (2025). *Global and regional ICT data sets*. https://www.itu.int/en/ITU-D/Statistics/pages/stat/default.aspx

Kim, D. J., Ferrin, D. L., & Rao, H. R. (2008). A trust-based consumer decision-making model. *Decision Support Systems, 44*(2), 544–564. https://doi.org/10.1016/j.dss.2007.07.001

Kizilcec, R. F. (2016). How much information? In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 2390–2395). ACM. https://doi.org/10.1145/2858036.2858529

Lankton, N. K., McKnight, D. H., & Tripp, J. (2015). Technology, humanness, and trust. *Journal of the Association for Information Systems, 16*(10), 880–918. https://doi.org/10.17705/1jais.00411

Lee, J. D., & See, K. A. (2004). Trust in automation. *Human Factors, 46*(1), 50–80.

McArthur, F., & Mbohwa, C. (2025). AI ethics in banking services. *Discover Artificial Intelligence, 5*, Article 319. https://doi.org/10.1007/s44163-025-00432-4

McKnight, D. H., Carter, M., Thatcher, J. B., & Clay, P. F. (2011). Trust in a specific technology. *ACM Transactions on Management Information Systems, 2*(2), 1–25. https://doi.org/10.1145/1985347.1985353

Mititean, P., & Cardoș, I. R. (2022). Corporate governance, transparency, and financial performance. *Economic Research–Ekonomska Istraživanja, 35*(1), 113–134.

Obelovska, K., Abziatov, A., Doroshenko, A., Dronyuk, I., Liskevych, O., & Liskevych, R. (2025). Analysis of digital skills and infrastructure in EU countries based on DESI 2024 data. *Future Internet, 17*(6), 228. https://doi.org/10.3390/fi17060228

OECD. (2024). *Education and demographic indicators*. https://stats.oecd.org/

Pavlou, P. A. (2003). Consumer acceptance of electronic commerce. *International Journal of Electronic Commerce, 7*(3), 101–134. https://doi.org/10.1080/10864415.2003.11044275

PwC. (2020). *Explainable AI: A guide for transparency and trust*. https://www.pwc.com/m1/en/services/assurance/risk-assurance/documents/explainable-ai.pdf

Rai, A. (2020). Explainable AI. *Journal of the Academy of Marketing Science, 48*, 137–141. https://doi.org/10.1007/s11747-019-00710-5

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?". In *Proceedings of the 22nd ACM SIGKDD Conference* (pp. 1135–1144). ACM. https://doi.org/10.1145/2939672.2939778

Shin, D. (2021). The effects of explainability and causability. *International Journal of Human–Computer Studies, 146*, 102551. https://doi.org/10.1016/j.ijhcs.2020.102551

Shirokova, G., Morris, M. H., Laskovaia, A., & Micelotta, E. (2021). Effectuation and causation. *Journal of Business Research, 129*, 169–182.

Siau, K., & Wang, W. (2018). Building trust in AI, ML, and robotics. *Cutter Business Technology Journal, 31*(2), 47–53.

van Deursen, A. J. A. M., & van Dijk, J. A. G. M. (2014). The digital divide shifts to differences in usage. *New Media & Society, 16*(3), 507–526.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology. *MIS Quarterly, 27*(3), 425–478.

World Bank. (2021). *Europe and Central Asia economic update*. https://openknowledge.worldbank.org/entities/publication/71eb3521-4c98-55b6-bf4c-6a49c9d36acb

World Bank. (2024). *World Development Indicators*. https://databank.worldbank.org/source/world-development-indicators

World Population Review. (2026). *Internet penetration by country*. https://worldpopulationreview.com/country-rankings/internet-penetration-by-country

Article Info

Received: February 15 2026
Accepted: March 18 2026

How to cite this article:

Mangiurea, V., Anghel, L. C. (2026). Who Controls the Decision? Artificial Intelligence, Algorithmic Control, and Consumer Trust in Banking. *Revista de Științe Politice. Revue des Sciences Politiques,* no. 89, pp. 132-148.