Algorithms and Models for BIAS Reduction in AI Implementation for Dynamic Organizational Environments

Liliana VUȘEAN (IOSIF)[1]
Nicoleta Valentina FLOREA[2]

Abstract

This paper examines the role of Artificial Intelligence in facilitating the management of organisational change, with a particular focus on the use of data analytics and algorithms to identify possible sources of biases in automated decision-making processes. The main objective is to identify techniques and models that can minimise these biases and to assess the impact of their reduction on organisational performance and culture.

In order to extend the applied research to the implementation of AI solutions in a branch of a leading company in the FMCG industry, present in multiple market segments both in Romania and in other European countries, it is essential to identify in advance the potential sources of biases that could negatively influence the activity and communication within the organisation.

Thus, this research investigates algorithms and methods to reduce biases, proposing appropriate solutions to minimise them in automation and change processes by integrating AI in organisational activity.

Keywords: *Artificial Intelligence (AI), change management, organisational communication, biases, algorithms and models.*

JEL classification: C8, M12, O33

DOI: 10.24818/RMCI.2025.4.712

1. Introduction

Artificial Intelligence (AI) plays a key role in transforming the modern organisational landscape, providing innovative solutions for managing change and adapting business structures to dynamic market demands. In the digital age, an organisation's ability to integrate advanced technology with effective change processes is crucial for maintaining competitiveness and improving performance. Today, organizations are adopting advanced digital solutions (Șișu et al., 2024), as artificial intelligence (AI) has become the key disruptive force across all industries) and is also shaping national economic outcomes (Ioan-Franc & Gâf-Deac, 2024).

[1] Liliana Vușean (Iosif), PhD student, Valahia University of Targoviste, E-mail: l.iosif@yahoo.com
[2] Nicoleta Valentina Florea, Valahia University of Targoviste, E-mail: floreanicol @yahoo.com

The traditional change management model, based on long-term methodologies (Parnell, 2013), implies a defined beginning and end for projects, with the direct involvement of the organisation's management. However, these approaches do not always keep pace with the accelerated pace of technological innovation. In this context, AI introduces a new dimension to the coordination and implementation of organisational change. With AI, organisations can process large volumes of data to identify trends, automate processes and improve internal and external communication. These factors accelerate change adoption and increase the level of adaptability of employees and organisational structures.

By applying innovative techniques and models, management processes can become more efficient. Advanced decision-making strategies and methods help to increase operational efficiency and facilitate organisational change (McKinsey, 2011).

2. Key Concepts – Definitions

The Bias in Artificial Intelligence (AI) refers to the systematic tendency of a model to generate erroneous results due to implicit biases or preferences, which may stem from the data used, algorithms or internal processes (Mitchell, 1997, p. 54). Correcting these biases is essential, as bias can lead to inequitable results, inaccurate decisions and, in some cases, discrimination in the use of AI systems.

Integrating AI into change management also brings challenges such as those related to ethics, data privacy and the impact on internal and external communication. As organisations adopt automated systems to increase the efficiency and accuracy of decisions, concerns about inherent biases that may affect these decisions become prominent. Biases can have substantial negative consequences, affecting not only individual outcomes, but also the overall culture and performance of the organisation.

Reducing biases in artificial intelligence models and organisational decision-making processes not only improves the performance of the models but can also have a positive impact on organisational culture. The present research investigates the role of AI in facilitating effective communication and smooth adaptation in changing organisations, addressing the identification of biases as well as the selection of specific algorithms and models for their mitigation in AI implementation processes.

These issues can be analysed from multiple perspectives: ethical, operational or managerial. Improving performance by minimising biases is reflected by increasing the accuracy of datasets and decreasing the risk of making inappropriate or incorrect decisions. An organisation that aligns its models and processes with principles of fairness and equity fosters an inclusive work environment, attracts talent and improves employee retention, thereby increasing motivation and achieving higher performance results (Negnevitsky, 2024).

2.1 Conceptual Model and Research Hypotheses

This research focuses on identifying the causes generating biases in the processes of implementing artificial intelligence (AI) in an organisation undergoing change. The research also proposes methods to minimise these biases, using specific techniques such as pre-processing, fair learning algorithms and post-processing, to ensure a digital transition without systemic biases.

The sources of the occurrence of biases in AI implementation processes are related to the data used during algorithm testing, which can lead to faulty AI implementations. These include the use of algorithms based on implicit assumptions that are not universally applicable or the application of algorithms in inappropriate contexts. In general, there is a tendency to consider statistically based managerial decisions to be superior; however, when statistical data are flawed and automatically processed, they can become sources of unfairness and bias (Stevenson & Doleac, 2022).

To identify the sources of biases, we reviewed the literature detailing how these biases arise and affect decision-making systems. Biases in automated decisions can have multiple sources (Barocas, Hardt & Narayanan, 2019): data biases (reflecting unfair or biased practices), algorithm biases (favouring certain patterns without prior data checking), and biases induced by those who implement and monitor the system (users who choose parameter settings introduce certain assumptions).

A general review of the different sources of biases - stemming from data, algorithms or human interventions (Mehrabi et al., 2021) - emphasises how the lack of data diversity and assumptions related to the choice of algorithms contribute to uncertainty and the occurrence of biases.

The first hypothesis of the research suggests that the implementation of algorithms as early as the data preprocessing phase contributes significantly to reducing biases in the AI models to be implemented in the organisation. AI algorithms need to be trained on diverse and representative datasets to avoid perpetuating existing biases, this is particularly important in changing organisations to adapt to new contexts and needs (Sheremetov & Diachenko, 2025).

The second hypothesis states that machine learning models incorporating AI decision auditing and monitoring techniques contribute to process transparency and bias reduction.

The third hypothesis argues that continuous training and diversity of teams responsible for AI development and implementation reduce the risk of introducing biases in implementation processes (Raji & Buolamwini, 2019) and in the use of new systems.

Hypothesis four addresses organisational culture that promotes AI ethics and social responsibility, thereby generating change processes that lead to bias reduction.

The proposed conceptual model directs the detailed analysis of the interdependence between the variables associated with each hypothesis and the reduction of biases in AI implementation processes.

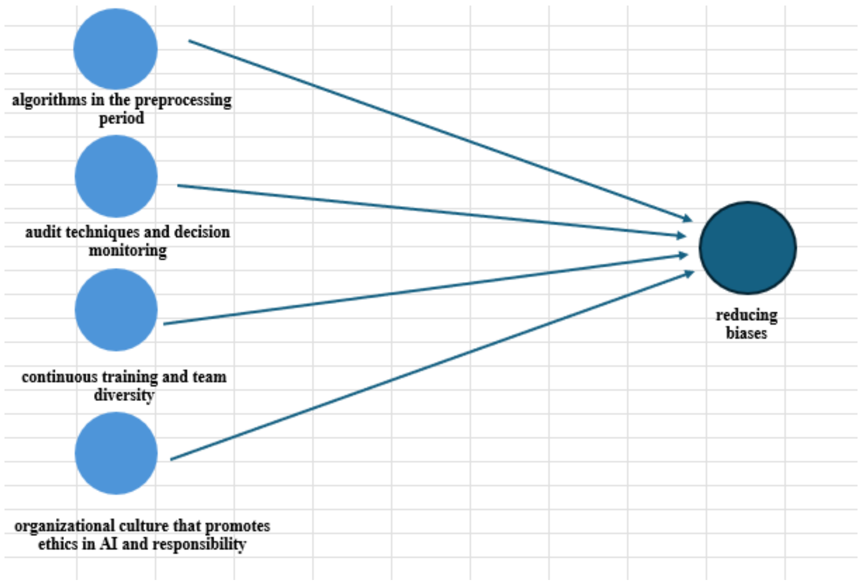

Figure 1. Conceptual Map
Source: authors work

3. Methodology, Instruments and Research Design

In order to collect the relevant data needed for this analysis and to test the hypotheses developed, we used survey techniques. To this end, we developed and applied a questionnaire consisting of 15 statements distributed along five main dimensions: application of algorithms in the data pre-processing phase, techniques for auditing and monitoring the decisions made, continuous training and diversity of the teams involved, organisational culture promoting ethics and accountability, and bias reduction. This questionnaire was implemented in a branch of an organisation where AI solutions have already been implemented.

The questionnaire was administered to a sample of 203 respondents, selected from a total of 650 employees at the head office of a well-known food manufacturing and distribution company. The participants in our study were employees working in various departments of the company, including logistics, warehouse, communication, marketing, marketing, human resources and commercial/distribution.

A first step in the research involved analysing demographic data on respondents' characteristics, such as age, gender, education level, residence, and position held in the company. These characteristics are detailed in the table below.

Demographic characteristics of respondents

Table 1

Demographic characteristics		n = 203	%
Age (years)	16-24	36	17.7
	25-34	47	23.2
	34-44	53	26.1
	45-54	40	19.7
	54-64	27	13.3
	Total	**203**	**100**
Gender	Men	80	39.4
	Women	123	60.6
	Total	**203**	**100**
Studies	secondary	113	55.7
	university	67	33
	postgraduate	19	9.4
	doctoral	4	2
	Total	**203**	**100**
Residence	urban	114	56.2
	rural	89	43.8
	Total	**203**	**100**
Position level	managerial	28	13.8
	execution	175	86.2
	Total	**203**	**100**

Source: authors work

4. Demographic Structure of Participants

Analysing the demographic structure of the participants revealed significant aspects, showing that almost half of the participants, 49.3%, are in the 25-44 age range. This indicates a concentration of the labour force in the phase of maximum productivity, considerably influencing this segment's outlook on challenges and opportunities. The 45-54 age group (19.7%) and the 54-64 age group (13.3%) also bring a diversity of experiences, from the enthusiasm of young professionals to the wisdom of experienced employees, enriching the discussion on the adoption of new technologies.

The gender analysis showed an imbalance, with 60.6% of respondents being female, emphasising not only a significant female presence in the workplace, but also possible differences in perceptions of artificial intelligence, influenced by varied communication styles.

In terms of education level, the majority of respondents (55.7%) are high school graduates, suggesting a prevalence of executive roles. The urban-rural distribution (56.2% urban vs. 43.8% rural) indicates differences in access to advanced technologies.

An analysis of the roles occupied in the organisation reveals that the clear majority of employees (86.2%) are in executive positions. This suggests that perceptions of the integration of artificial intelligence into communication processes are mainly influenced by those who implement the technologies, rather than those who strategise.

This demographic diversity contributes significantly to the external validity of the study, strengthening the ability to generalise the results to different organisations and socio-economic contexts. The representation of participants by department is summarised in Table 2 and the graph shown below (Figure 2).

The representation of participants by department

Table 2

Department	Respondents
Production&Quality	71
Logistics&Warehouse	45
Technical	22
Commercial/Distribution	20
Accounts	14
Procurement	14
Marketing & PR	11
HR	6
Total	**203**

Source: authors work

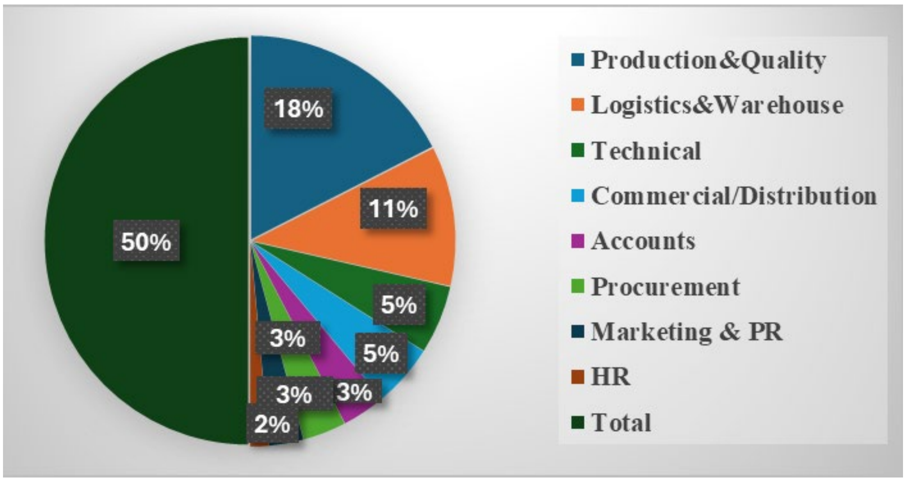

Figure 2. Graphical representation of respondents/departments
Source: authors work

A significant percentage, over 50 per cent, of the respondents work in the production, quality, technical and commercial/distribution departments. Therefore, it can be deduced that these demographic groups considerably influence the results of the applied questionnaire.

When designing the questionnaire, we have integrated questions that cover a broad spectrum of perspectives, thus minimising the risk of favouring one particular opinion. Analysing responses according to demographic variables such as age, gender or location may also reveal potential systematic biases.

Incorporating the consideration of cognitive biases, according to Ariely (2008), involved monitoring those common biases, such as confirmation bias and authority bias, in both the wording and interpretation of questions. To this end, we applied methods such as factor analysis and regression techniques, which are capable of identifying patterns indicative of biases.

Factor analysis, a statistical technique that allows the discovery of relationships between observed variables, has been used to detect real-life patterns in the collected responses, which could signal the presence of cognitive biases (Fabrigar et al., 1999).

Cognitive biases can have a significant impact on the way questions are constructed and interpreted, which can lead to erroneous conclusions (Korteling & Toet, 2020). For example, confirmation bias occurs when there is a tendency to privilege information that confirms pre-existing beliefs to the neglect of contradictory data, and questions are framed to suggest a specific preferred response.

In the case of authorities who influence participants and where the questions appear to support expert opinions, authority bias arises, which can affect respondents' perceptions.

The fact that the questionnaire was administered to a diverse group, with different roles, education levels, age and residence, allowed the identification of questions that could stimulate bias.

The meticulous design of the questions, with clear and precise language and avoidance of technical or difficult to understand terms, as well as neutral questions, allowed for the collection of objective data and prevented response bias (Bradburn, Sudman & Wansink, 2004).

Bias identification methods using exploratory data analysis (EDA) are essential as they facilitate a deep understanding of the structure and characteristics of the data and allow early detection and correction of possible biases and errors. Data visualisation, as a key element of exploratory data analysis, helps to identify and understand patterns, relationships and anomalies, providing clarity on the distribution of variables (Bruce et al., 2020).

Averaging can reveal significant differences between groups, suggesting the presence of bias. The arithmetic mean, defined as the sum of the values divided by the total number of observations, was used to establish the central tendency of the dataset (Groves, 2009). Comparing mean values between distinct groups (e.g. gender, age) helped us to identify possible biases.

Although the demographic variable "gender" includes both females and males, there is an underrepresentation of males (only 39.4%), which may indicate selection bias. Comparison of the centrality metrics (mean, median) for different subsets of the data can reveal significant differences; using descriptive statistics, we can detect possible skewness in the distributions of the variables.

Graphs and visual representations help to effectively communicate the results of analyses, facilitating informed decisions, generating new hypotheses and research questions, and providing directions for more advanced statistical analyses.

5. Research Results

In terms of participation and involvement in the pre-operation period of the AI solution in the company, the following results were obtained (Figure 3).

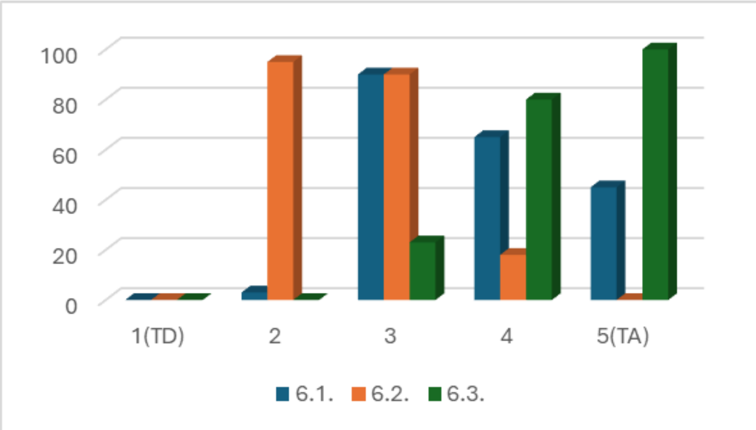

Figure 3. Participation and involvement in the pre-operation of the AI solution within the company
Source: authors work

The majority of respondents felt moderately involved in the data pre-processing period, with a predominant score of 3. This suggests that although a significant number of staff actively participated, there is untapped potential to increase staff engagement at this critical stage.

Related to the period of auditing and monitoring decision making by the implemented AI solution a significant proportion of respondents (100 people for each of scores 2 and 3) felt that the algorithms had a positive-moderate influence on the AI solution. However, only 53 respondents felt a significantly positive influence (score 4 and 5), suggesting opportunities to improve the impact of algorithms at the organisational level.

The majority of respondents (180 out of 203) felt that the pre-processing period was the right time to find the right algorithms for effective AI implementation (scores 4 and 5). This indicates a high appreciation for following a clear and concise, strategically defined and organised plan in establishing the algorithms and underlines the importance of adequate preparation at this stage.

Following the application of the questionnaire question on auditing and monitoring the decision making of AI solutions in the organisation (Figure 4), the responses of the 203 respondents were collected and analysed and the following conclusions were drawn.

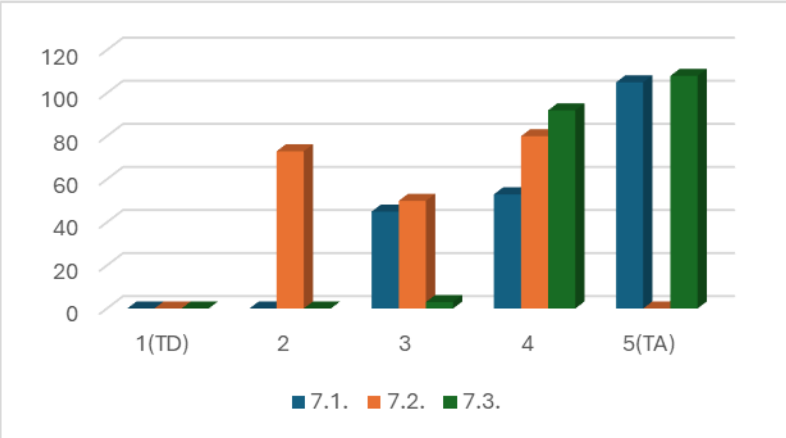

Figure 4. Auditing and monitoring how AI decision making is made

Source: authors work

An overwhelming majority of respondents believe that ongoing auditing of AI solutions post-implementation is essential and this emphasises the importance of continuous monitoring to ensure the reliability and efficiency of AI operation, as well as to identify and correct any emerging issues.

Responses varied considerably, with a trend towards scepticism about the infallibility of AI decisions, with a significant number of respondents (123 out of 203) rating the statement with scores of 2 and 3, signalling that there are concerns about the accuracy of AI decisions. However, 80 respondents have a more positive view (score 4) and this suggests the need to improve the verification and validation processes of AI decisions to increase their confidence and reliability.

Also, the overwhelming majority of respondents (200 out of 203) believe that the use of AI contributes significantly to the transparency of processes, with scores of 4 and 5, suggesting that in the general perception, AI is seen as a tool that can facilitate clarity and openness of organisational decisions, as long as they are properly implemented with audit and communication mechanisms.

The results indicate a widespread appreciation for the need for continuous audit, showing that organisations recognise its importance in maintaining the integrity of AI solutions. However, there is a moderate concern about the infallibility of AI decisions, suggesting an area where investment can be made in developing confidence by optimising existing models and improving their interpretability.

In terms of transparency, positive perceptions indicate that AI is associated with clear benefits in terms of openness and visibility of decision-making processes, potentially improving organisational and collaborative processes. In conclusion, by balancing audit, accuracy and transparency, organisations can better harness the benefits of AI in their decision-making processes.

From the perspective of the involvement of the team of managers in the implementation of the AI project, the results of the questionnaire allow a detailed assessment of the perception and interaction of team members in the development and maintenance process (Figure 5).

The majority of respondents (162 out of 203) rated their direct involvement with scores of 4 and 5, suggesting an active participation in the implementation team, this indicating a high level of personal involvement, reflecting a good commitment to the project. However, there is still a smaller proportion of respondents who felt less involved (score 3), where more communication and clear distribution of roles may be needed.

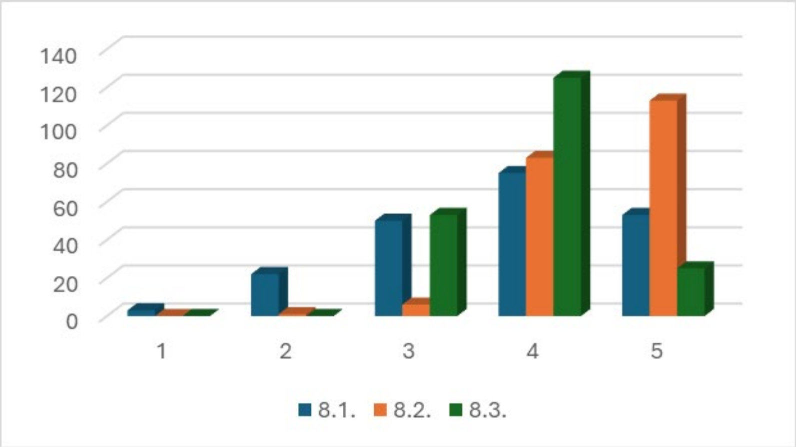

Figure 5. Team responsible for implementing the AI project
Source: authors work

Nearly all respondents actively participate in professional development activities, indicating a strong commitment to continuous learning and improvement in the use of AI systems.

This favorable attitude towards continuous training supports operational efficiency and the ability to respond quickly to errors or technological challenges.

The majority of respondents felt that the implementation team was competent enough to ensure the effective operation of the AI solution, with 162 out of 203 rating the team positively (scores of 4 and 5.) However, 41 respondents gave a score of 3, suggesting room for improvement in integrating technical expertise or collaboration within the team.

In conclusion, the results of question 8 of the questionnaire indicate a high level of involvement and commitment of the employees surveyed in the AI project implementation process. Of note, investment in continuous training and professional development should continue, as these are perceived positively and are essential to maintain and improve the performance of AI solutions.

While the overall perception of the quality of the implementation team is positive, there is room for developing team skills and improving collaboration, which could further increase the efficiency and success of AI implementations. Organising constant feedback and various internal initiatives on this segment can help in addressing these challenges.

In terms of the use and benefits of AI project implementation, the results of the questionnaire provided valuable insights into the use, acceptance and impact of AI solutions within the company (Figure 6).

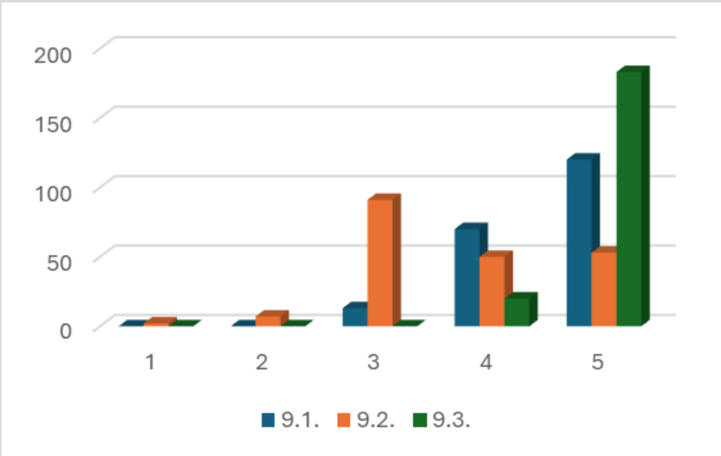

Figure 6. The use and benefits of AI project implementation
Source: authors work

An overwhelming majority of participants (190 out of 203) believe that they use AI ethically and responsibly, reflecting a strong commitment to ethical practices in the use of technology. This indicates awareness of and respect for the standards and norms required to operate AI in the organisation.

Whilst there is a general trend of acceptance of AI decisions, with 103 respondents giving scores of 4 and 5, there is a significant proportion (91 respondents) who rated the statement with a 3. This suggests that around half of the respondents may have reservations or requirements for further justification for AI decisions, highlighting opportunities for improving the communication and explainability of AI solutions.

The overwhelming majority of respondents (183 out of 203) believe that the implementation of AI has significantly facilitated and supported change processes within the company, highlighting the positive impact of AI on organisational innovation and adaptability. This result suggests that AI not only improved efficiency, but also provided a supportive framework for beneficial change.

The results indicate a favorable opinion towards the ethical use and positive impact of AI within the company. Whilst users are largely satisfied with the ethics of implementation and the beneficial effects of AI on change initiatives, there is still some reluctance in unconditional acceptance of AI decisions. Solutions in this regard could be: to invest in AI training, to reassure users that AI decisions are well informed and about the need to improve the interpretability of AI-generated results to increase user confidence and align technology decisions with organisational values.

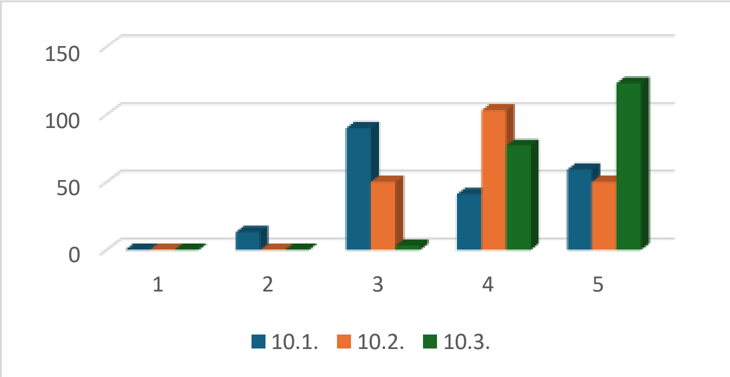

Figure 7. Reducing biases (preconceptions) when implementing the AI solution
Source: authors work

A significant proportion of respondents (90 out of 203) reported a modest assessment (score 3) of the organisation of audit and recalibration actions of the AI system to reduce bias, indicating a perception of the need to improve efforts or communication on these initiatives. However, 100 respondents have a more positive perception (scores 4 and 5), emphasizing a recognition of the efforts made. Reinforcement and visibility of adjustment programs can increase confidence in these processes.

The majority of respondents (153 out of 203) believe that AI decisions contribute positively to organisational culture and change, reflecting a perceptible impact of AI on organisational adaptability. However, 50 respondents have a more neutral view (score 3), suggesting that further presentation of the benefits and integration of the AI vision into organisational strategy may be needed.

An overwhelming majority (200 out of 203) agree that reducing preconceptions in AI systems contributes significantly to increased confidence in the data provided and the effective use of AI solutions, with a large number of respondents giving maximum scores (5). This emphasizes the importance of continued efforts to identify and mitigate biases to maintain and increase trust and engagement in the use of AI.

The results indicate a general recognition of the need for and benefits of mitigating biases in AI solutions and a significant positive impact on organisational culture and change. However, the varied perceptions of the audit and recalibration efforts suggest that there are opportunities to improve the communication and implementation of this initiative in a visible way.

Reinforcing these practices through discussion and training sessions and high transparency of results can help align employee perceptions with the bias reduction strategy and could amplify the positive impact AI is already having on the organisation.

In terms of monitoring cognitive biases, from our analysis we found that the majority of respondents feel a significant commitment to the ethical use of AI and participation in ongoing training. This high commitment could reflect a positive employee bias towards notions of ethics and responsibility, reinforced by organisational policies and values.

Despite the commitment to ethics, there is some reluctance to accept AI decisions (50% being neutral or reluctant) which could indicate a confirmation bias, where employees rely on their own knowledge and judgements and are more skeptical of automated solutions that are not sufficiently clearly explained to them.

The implementation of AI is perceived as having a positive impact on organisational processes and facilitating change. This optimism might be influenced by a bias where respondents tend to assess the impact of new technologies more favorably due to initial enthusiasm and high expectations.

Also, the observed reluctance towards AI decisions (in 10.2 and 9.2, where many responses were 3 or lower) may suggest a confirmation bias, where employees prefer information and decisions that confirm their initial views and may be skeptical of AI solutions.

In questions 9.3 and 10.3, the positive perception of the impact of reducing AI biases and facilitating change indicates an optimism bias, where respondents expect the positive dimensions of AI to predominate due to the potential benefits.

The results for question 10.1 indicate a modest assessment of audit and recalibration efforts, possibly reflecting a status quo bias - where employees are influenced by familiarity with existing systems and processes and have a reservation against constant change or new interventions, even in the form of recalibration.

The hypothesis that the implementation of algorithms as early as in the pre-processing period contributes significantly to reducing biases in the AI models to be implemented within the organisation (Table 3) is not fully supported by the available data.

Although there is a positive perception of pre-processing as an appropriate time to choose algorithms (question 6.3), the positive impact of these algorithms on the subsequent use of the AI solution is not clearly evident (low scores in question 6.2). Many respondents had medium to high involvement in the preprocessing period (with a mean score of 3.75), indicating significant participation in the process of identifying specific algorithms.

The overwhelming majority of respondents felt that preprocessing was the right time to identify the correct algorithms, indicating a positive perception of the timing and synchronization of the timing of the introduction of these algorithms, with a mean score of 4.38.

Participation and involvement in the pre-operation of the AI solution in the company

Table 3

	1	2	3	4	5	**average/vari**
6.1. I was involved in the data pre-processing period in finding specific algorithms	0	3	90	65	45	**3.75**
6.2. the implemented algorithms have positively influenced the subsequent use of the AI solution	0	95	90	18	0	**2.62**
6.3. the preprocessing period was the right time to find the right algorithms	0	0	23	80	100	**4.38**

Source: authors work

Respondents are, however, skeptical that the algorithms implemented in preprocessing had a positive influence on subsequent use, with a mean score of 2.62. This suggests that the anticipated benefits of these algorithms were not clearly manifested or were limited.

As for the validity of the second hypothesis - which refers to machine learning models that integrate AI decision auditing and monitoring techniques specifying that they contribute to process transparency and thus reduce bias - this is partially supported by the data collected (Table 4).

Auditing and monitoring AI decision-making

Table 4

	1	2	3	4	5	**average/variable**
7.1. permanent audit of the AI solution post implementation is absolutely necessary	0	0	45	53	105	4.3
7.2.decisions and solutions proposed through AI are always correct	0	73	50	80	0	2.32
7.3. the use of AI in decision making ensures transparency of processes	0	0	3	92	108	4.52

Source: authors work

With a mean-ranked score of 4.30, the responses indicate a strong consensus that post-implementation auditing is crucial for the proper functioning of AI solutions suggesting a clear recognition of the role of auditing in maintaining the integrity of AI systems.

With an average score of 2.32, there is significant scepticism that AI decisions are always correct, this highlights concerns about the inherent biases and current limitations of AI systems. The majority of respondents agree that the use of AI contributes significantly to the transparency of decision-making processes (with an average score of 4.52), which can make it easier to identify and manage biases.

More specifically, common agreement on the need for auditing suggests that it is seen as an essential practice for transparency and bias reduction and the high perception of AI transparency suggests that models that integrate auditing and monitoring can contribute significantly to exposing and managing biases. However, responses indicate scepticism about the error-free nature of AI decisions, signalling that the integration of auditing alone is not sufficient to ensure absolute accuracy and complete bias reduction.

Implement measures such as: Developing and deploying advanced algorithms that include active self-correcting mechanisms and real-time identification of biases, increasing the level of transparency in technological processes that could include detailed reporting and user accessibility to how algorithms work to build trust in AI, training employees and end-users on the limitations and capabilities of AI, and on how auditing is used to ensure fair decision-making processes,establishing continuous feedback mechanisms for detecting and correcting deviations in AI decisions would improve the effectiveness of AI models and strengthen the hypothesis that integrating auditing and monitoring contributes to transparency and bias reduction.

The hypothesis that continuous training and diversity of teams contribute to reducing the risk of introducing biases is supported by the results of the analysis (Table 5) as a strong commitment to continuous training confirms that teams are proactive in managing errors and how they perceive and resolve biases, and the recognition of teams' expertise and its impact on the effective functioning of AI solutions underlines that professional diversity and expertise are seen as key elements in reducing biases.

Team responsible for implementing the AI project

Table 5

	1	2	3	4	5	**average/variable**
8.1.I was directly involved in the implementation team	3	22	50	75	53	**3.54**
8.2.I still attend trainings and workshops to be permanently connected to solutions for system generated errors	0	1	6	83	113	**4.51**
8.3.the implementation team consisted of specialists so that the implemented AI solution works very well	0	0	41	120	42	**4**

Source: authors work

With an average score of 3.54, a significant number of respondents indicate a moderate to high level of involvement in the implementation team, suggesting that most people took an active role in the process.

There is a strong commitment to ongoing training (mean score of 4.51), with the majority of respondents actively participating in trainings and workshops to stay informed about bug fixes. This indicates a recognition of the importance of ongoing training and a positive perception of the competence of the implementation team and the positive impact of diversity on the functioning of the AI solution (mean score of 4.00).

The fourth hypothesis is strongly supported by the collected data because: the ethical use of AI is well embedded in the organisational culture, which can lead to better bias management and implementation of responsible AI practices; and AI implementation is perceived as an important catalyst for change, suggesting that AI not only supports existing processes but also stimulates organisational improvements.

Use and advantages of implementing the AI project

Table 6

	1	2	3	4	5	**average/variable**
9.1.use AI ethically and responsibly within the company	0	0	13	70	120	**4.53**
9.2. easily accept a solution or decision generated by the implemented AI system	2	7	91	50	53	**3.65**
9.3. consider that the implementation of the AI solution has facilitated and supported change processes within the company	0	0	0	20	183	**4.9**

Source: authors work

With an extremely high mean score of 4.53, many respondents believe that the use of AI in their company is ethical and responsible with a strong commitment to ethical principles, which is essential for reducing bias. Also, with a mean score of 4.90, it indicates overwhelming agreement that AI solutions have facilitated positive change in the company (AI is seen as a catalyst for organisational change). However, the responses indicate moderate to high openness (mean score 3.65) in terms of acceptance of AI solutions. This result suggests that there is a level of trust in AI decisions, but also some caution.

The hypothesis that reducing biases in AI solution implementation is critical to improving the functioning of and trust in these systems is strongly supported by the evaluation results (Table 7) because: there is strong recognition of the importance of reducing biases in increasing trust and effectiveness of AI systems and AI is perceived as a positive factor in influencing organisational culture and sustaining change, suggesting that the effects of AI are generally well received.

Reduction of biases (preconceptions) in AI solution implementation

Table 7

	1	2	3	4	5	average/variable
10.1.within the company we have observed that audit and recalibration sessions of the AI system have been organised in order to reduce biases (preconceptions)	0	13	90	41	59	**3.59**
10.2.decisions and solutions given by the implemented AI system positively influence the organisational culture and support organisational change	0	0	50	103	50	**4**
10.3.I believe that reducing biases in the AI solution helps to increase confidence in the data provided by the system and its efficient use	0	0	3	77	123	**4.59**

Source: authors work

With a mean-ranked score of 4.59, there is a broad consensus that reducing biases increases trust in AI systems and makes them more effective hence the importance of combating biases to increase trust and acceptance of AI solutions.

There is a general agreement that AI decisions positively influence organisational culture and support change (mean score of 4.00) which demonstrates that AI is a positive factor in transforming organisational culture.

At the same time, with a mean score of 3.59, there is a moderate perception that auditing and recalibration to reduce biases are practised in the company.

6. Conclusions and Recommendations

The results of this questionnaire indicate varying staff involvement in the pre-processing stage, with an overall moderate to positive judgement on the influence of the implemented algorithms and the timing of selection. However, there is room to increase the level of involvement and maximise the positive effects through more effective communication strategies and the provision of adequate resources for all stages of AI development.

It also concludes that facilitating a collaborative environment and continuous information sharing can increase employee engagement and positive perceptions of the impact of AI in the organisation.

Among the solutions emerging from the research we identified the need to ensure that the data is clean, complete and diverse before being tested on the new models and it is imperative to develop and utilise algorithms that are sensitive to potential biases and allow for automatic correction whenever necessary.

The framework within which the AI solution is implemented must be transparent and fair so that both the decision-makers in the organisation and the users of the system can understand the AI decisions and check them when necessary.

Implementing initiatives that enable users to understand the decisions and recommendations made by AI systems always ensures better informed feedback. Evaluation should be done based on identifying robust and reliable methods that identify and correct biases throughout the AI lifecycle.

The organisation's internal culture needs to emphasise the ethical use of AI and also take into account the risks of bias. We identified that , within the analysed organisation various training initiatives, especially for key departments related to the use of AI focus on educating on the risks and responsibilities related to automated decisions.

By integrating these solutions, organisations can reap the benefits of automation and AI, while simultaneously reducing the risk of decision bias and ensuring a fair and efficient integration of these technologies into current processes. This research thus proposes a holistic and proactive approach to bias management in the age of digitisation and can be extrapolated to groups of firms or organisations in different business domains.

Although we have analysed the distribution of respondents by department and a range of their demographics, it would be useful in future research to consider whether or not respondents represent a significant percentage of the total employees in the departments to which they belong.

Applying test datasets to evaluate the performance of bias reduction algorithms is a recommendation that should be implemented in future projects.

Also, the use of correction algorithms when appropriate, the organisation of AI implementation processes in a transparent and accountable way, continuous monitoring by constantly evaluating the performance of algorithms to correct biases are essential elements that should be implemented.

Adjusting data to ensure a balanced representation, increasing transparency about AI decisions and explaining the benefits of bias reduction can help reduce confirmation bias. At the same time, promoting open communication and holding feedback sessions can be useful to address cognitive biases and adjust strategies according to the most common perceptions and maintaining an ongoing dialogue about the impact of AI and recalibration measures can improve adaptability and reduce reluctance to change, decreasing status quo bias.

References

1. Ariely, D. (2008). *Predictably irrational: The hidden forces that shape our decisions*. New York: Harper Collins.
2. Barocas, S., Hardt, M., & Narayanan, A. (2019). *Fairness and machine learning: Limitations and opportunities*. FairML Book, pp. 45-60.
3. Bradburn, N.M., Sudman, S., & Wansink, B. (2004). *Asking questions: The definitive guide to questionnaire design*. Hoboken, NJ: Wiley.

4. Bruce, P., Bruce, A., & Gedeck, P. (2020). *Practical statistics for data scientists: 50+ essential concepts using R and Python*. 2nd ed. Sebastopol, CA: O'Reilly Media, pp. 50-51.
5. Fabrigar, L.R., Wegener, D.T., MacCallum, R.C., & Strahan, E.J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), pp. 272-299.
6. Groves, R.M., Fowler, F.J., Couper, M.P., Lepkowski, J.M., Singer, E., & Tourangeau, R. (2009). *Survey methodology*. 2nd ed. Hoboken, NJ: Wiley, pp. 9-12.
7. Ioan-Franc, V., & Gâf-Deac, I.I. (2024). Participation of artificial intelligence in economic growth in Romania, *Amfiteatru Economic*, 26(67), pp. 944-956.
8. Korteling, J.E., & Toet, A. (2020). Cognitive biases. In: G.F. Koob, M. Le Moal & R.F. Thompson, eds. *Encyclopedia of behavioral neuroscience*. 2nd ed. Amsterdam: Elsevier.
9. McKinsey, J.O. (2011). *Management techniques*. 1st ed. New York: McGraw-Hill Education, p. 394.
10. Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys*, pp. 8-12.
11. Mitchell, T.M. (1997). *Machine learning*. New York: McGraw-Hill, p. 54.
12. Negnevitsky, M. (2024). *Artificial intelligence: A guide to intelligent systems*. 4th ed. Harlow: Pearson Education Limited, pp. 8-12.
13. Parnell, J.A. (2013). *Strategic management: Theory and practice*. 4th ed. Thousand Oaks, CA: SAGE Publications, Inc., p. 664.
14. Raji, I.D., & Buolamwini, J. (2019). Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial AI products. In: *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*. Available at: https://dl.acm.org/doi/10.1145/3306618.3314244 (Accessed: 2 April 2025).
15. Sheremetov, D., & Diachenko, A. (2025). Biases in artificial intelligence: How to detect and reduce bias in AI models. *Onix Systems Blog*, 24 February. Available at: https://onix-systems.com/blog/ai-bias-detection-and-mitigation (Accessed: 3 April 2025).
16. Stevenson, M.T., & Doleac, J. (2022). Algorithmic risk assessment in the hands of humans. *SSRN*. Available at: https://ssrn.com/ (Accessed: 1 April 2025).
17. Șișu, J.A. et al. (2024). The Role of Digital Solutions in Enhancing Organizational Efficiency., *Review of International Comparative Management*/Revista de Management Comparat International, 25(4).