Minnesota Journal of Business Law and Entrepreneurship
Volume 2026, No. 2
ISSN: 1540-3270

Mitigating Biases in Managerial Decision Making with Assistance from Large Language Model Tools: The Case of Artificial Intelligence Auditing

[1]Kali Charan Sabat, [2]Som Sekhar Bhattacharyya, [3]Pranjali Kulkarni

[1]Professor, Department of Management Studies,

Dayananda Sagar College of Engineering,
Bangalore, Karnataka, India
Email: kalicharan330@gmail.com
[2]Professor of Strategy & Entrepreneurship,
Indian Institute of Management (IIM),
Nagpur, Maharashtra, India
Email: som@iimnagpur.ac.in
[3]Visvesvaraya National Institute of Technology Nagpur (VNIT Nagpur)
Nagpur, Maharashtra, India
Email: pranjaliyk1236@gmail.com

Abstract

Managers have historically relied on data for decision-making. The purpose of this study is to explore the potential enhancement of managerial decision-making capabilities through the usage of Large Language Models (LLMs). The study aims to investigate how LLM-based tools could provide cues to managers, influencing their decision-making process and contributing to better organizational decision-making. The research employed a theoretical approach to examine the role of LLMs in managerial decision-making. It assessed how LLM tools offered decision choices to managers and explored the convergence divergence between LLM suggestions and human intelligence. The study emphasized the phenomenon of artificial intelligence avoidance and underscored the need for artificial intelligence auditing at both the planning and operational phases to ensure effective decision-making. The findings highlighted that LLM-based tools could significantly impact managerial decision-making. When LLM suggestions aligned with managers' human intelligence, decision certainty could be enhanced. However, if there was a divergence, managers might experience uncertainty due to artificial intelligence avoidance. The study emphasized the importance of auditing LLM tools, particularly in the context of business organizations, to address biases and ensure reliable decision support for managers. This study contributes to the literature by providing insights into artificial intelligence auditing in the context of business organizations. It underscores the originality of exploring the convergence and divergence between LLM suggestions and human intelligence, shedding light on the nuanced dynamics of managerial decision-making in the presence of artificial intelligence tools.

Keywords: Artificial intelligence, Auditing, Business, Bias, Decision making, Regulation, entrepreneurship, LLM

1. Introduction

In the dynamic landscape of artificial intelligence (AI), the advent of large language models (LLMs) has ushered in a new era of natural language processing, sentiment analysis, and content generation (Hadi *et al.,* 2023). However, this technological progress has not been without its challenges, as concerns were raised by both practitioners and managers regarding biases embedded in the outputs of these powerful models. This study aims to delve into the intricate realm of LLM biases, acknowledging the ethical dilemmas and societal implications associated with biased AI tools (Hadi *et al,* 2023). At the forefront of this investigation is the proposal of a comprehensive auditing framework designed to systematically identify and mitigate biases within LLMs (McIntosh *et al.*, 2024). The urgency of this research stems from the realization that biased outputs can significantly influence decision-making processes, shape public discourse, and perpetuate existing inequalities (Tokayev, 2023). By combining quantitative and qualitative analyses, the auditing framework not only scrutinizes biases in model outputs but also investigates their origins within the training data, offering a nuanced understanding of the biases at play (Prabhu and Birhane, 2020). In undertaking this research, the objective was to contribute valuable insights to the discourse on ethical AI development, striving to foster responsible AI integration while upholding transparency and accountability in the development process (Schiff *et al.*, 2020).

This study sets the stage for a detailed exploration of LLM biases, the innovative auditing methodology employed, and the subsequent findings, with the overarching goal of guiding the responsible deployment of AI systems in diverse applications (Mökander *et al.*, 2023). In recent years, the pervasive impact of AI technologies on our daily lives has underscored the critical importance of addressing biases within these systems. The biases often originate from the training datasets, reflecting and perpetuating societal prejudices (Leavy *et al.,* 2020). This study recognizes the multifaceted nature of biases within LLMs, acknowledging their potential to exacerbate existing disparities and hinder the pursuit of fairness and inclusivity in AI applications (Schwartz *et al.*, 2022). Consequently, the proposed auditing framework not only aims to identify and rectify biases but also provides an avenue for fostering greater awareness and understanding of the ethical dimensions of AI development.

The research acknowledges the limitations of traditional auditing methods in addressing the nuances of language models, emphasizing the need for a tailored approach (Hadi *et al.*, 2023). By advocating for the integration of this framework, the study seeks to catalyse a paradigm shift towards more responsible AI practices, where developers, policymakers, and society at large engage in an ongoing dialogue about the ethical implications of these powerful technologies (Bankins and Formosa, 2023). The ensuing analysis and findings are poised to contribute not only to academia but also to industry practices, ensuring a more equitable and ethically sound future for AI integration (Dwivedi *et al.*, 2021). Ultimately, this research strives to bridge the gap between technological advancements and ethical considerations, ushering in an era where AI systems align with societal values and uphold principles of fairness and inclusivity (Walsh *et al.*, 2019).

This article is organized as follows. First, it examines the background for the need of bias mitigation in the managerial decisions while using AI tools and systems for decision making. Next, it explores bias detection and mitigation strategies, including domain-specific auditing techniques and methods to ensure fairness in recommendations and pricing. The subsequent sections focus on the ethical aspects of LLM auditing such as transparency, explainability, and user-centric mechanisms, highlighting the importance of interpretable algorithms and user feedback systems in building trust. This is followed by a discussion on proposed reforms on ethical governance and privacy protection, addressing regulatory compliance, ethical guidelines, and the safeguarding of user data. Further, the article investigates the future implications of LLM audits in the form of human-in-the-loop systems, periodic audits, and collaborative approaches. for business and entrepreneurship. Finally in the conclusions section, it discusses capacity building and external validation, including training initiatives, third-party audits, and open collaboration to strengthen accountability and promote best practices. Through this structured approach, the research work aims to provide a comprehensive foundation for developing fair, transparent, and ethically responsible LLM-driven decision-making systems.

2. Background

Large language models (LLMs) represented a ground-breaking technological leap (Zhou *et al.,* 2023). Applications of LLMs presented immense potential for creating superior competitive positions for business firms operating in the digital age. Further, one could argue that LLMs were advanced artificial intelligence (AI) tools (Kar *et al,* 2023). LLMs possessed a substantive capacity to comprehend and generate text that closely mimicked human language (Ruan *et al.,* 2023). Business managers and entrepreneurs could thus apply LLMs across a wide spectrum of business operations. LLMs if deployed by managers properly would benefit the firm. Thus, it was essential to recognize the applicability and utility of LLMs (Meyer *et al.,* 2023).

One could argue that from the perspective of resource-based view (Voelpel *et al.*, 2004), LLMs could be classified as strategic assets capable of reshaping the way businesses operated and competed in the marketplace. Usage of LLMs could augment business and entrepreneurial decision-making capabilities. LLMs could help managers both sense and seize new market opportunities. Deployment of LLMs in organizations could also drive operational efficiencies at firm customer interface (Kar *et al.,* 2023). This was especially by transforming customer interactions and crafting highly personalized marketing campaigns. This could be possible by both automating complex customer-based data analysis and content creation for customers. In a world where data was abundant and effective communication was of paramount importance, application of LLMs could create unique advantage (Kasneci *et al.*, 2023). This could be achieved by creating compelling narratives and making data-informed decisions.

The primary target for incorporating the LLMs was in the business-to-customer (B2C) context (De Caigny *et al.*, 2021). This was because LLMs application involved direct communication with a large number of customers. Large number of customers brought in a large number of iterative interactions for a typical B2C firm (De Caigny *et al.*, 2021). The firm customer interaction volume was substantive, owing to the varied context of scenarios and the frequency of interactions. Thus, LLMs were apt tools for such high-volume firm customer interaction. LLMs could be deployed at the customer interface like based upon an app or a website to facilitate effective communication between customers and a firm.

One must note that firms interact with customers at three stages namely pre-purchase, purchase and post purchase (Park *et al.*, 2015). At the pre-purchase stage, customers explored the product and service offerings of a firm from digital channels like mobile apps or websites. At the purchase stage customers interact with a firm to make the payment transaction. At the post-purchase stage, customers received after sales services regarding the products and services the firm had offered to customers. This again could be provided through digital channels. LLMs could be deployed by firms to cater to all the three stages of firm customer interaction. The most common use cases for the firms have been tabulated in table-1.

Table-1- Use cases of deployment of LLMs during firm-customer interaction

Pre-purchase	**Purchase**	**Post-purchase**
Market Research: LLMs can analyse vast amounts of online content, customer reviews, and social media data to provide valuable insights about market trends, customer sentiment, and competitors. This information can inform product development and marketing strategies. ***Content Generation:*** LLMs can assist in creating engaging and informative pre-purchase content such as blog posts, product descriptions, and marketing materials. They can help ensure consistent messaging and SEO optimization. ***Chatbots:*** LLM-powered chatbots can provide instant responses to customer inquiries, assisting potential buyers in the early stages of their research. They can answer common questions, offer product recommendations, and guide users to relevant resources.	***Personalized Recommendations:*** LLMs can analyse user behaviour and preferences to provide personalized product recommendations. This can increase the likelihood of successful conversions. ***Chatbots and Virtual Assistants:*** LLM-driven chatbots can facilitate the purchase process by answering questions about product specifications, pricing, and availability. They can even assist with completing transactions. ***Content Customization:*** LLMs can dynamically adjust website content and product listings based on user profiles, enhancing the user experience and potentially increasing sales.	***Customer Support:*** LLM-powered chatbots and virtual assistants can provide post-purchase support, helping customers with order tracking, returns, and warranty information. ***Feedback Analysis:*** LLMs can analyse customer feedback and reviews to identify areas for product improvement and customer satisfaction enhancement. ***Content Marketing:*** LLMs can generate follow-up content like thank-you emails, product usage guides, and post-purchase offers, maintaining engagement with customers after their initial purchase. ***Cross-selling and Upselling:*** LLMs can analyse customer data to suggest complementary products or upgrades, increasing revenue from existing customers.
Source: Authors' own conceptualization		

One could argue based upon the observation in table-1 that LLMs could play a versatile role during the pre-purchase, purchase, and post-purchase phases of a customer's interaction with a firm. LLMs could contribute to improved customer engagement, personalized experiences, more efficient operations, and data-driven decision-making. Leveraging LLMs effectively in these areas could lead to enhanced customer satisfaction, increased sales, and overall business growth.

3. LLM auditing: Ethical or Unethical

Managers had always relied on data for decision making. Given the abundance and ubiquitous presence of data in the era of big data, appropriate data analysis has become a very potent and strategic organizational capability (Sivarajah *et al.*, 2017). The coming up of various artificial intelligence (AI) tools further helped managers in getting assistance in their decision-making process. Usage of large language model (LLM) tools could provide managers substantive scope to augment their individual decision-making capabilities which would in turn facilitate better organizational decision making (Mokander *et al.*, 2023). LLM based tools could provide cues and stimuli to managers to undertake a particular decision from a possible decision choice set (Guler *et al.*, 2025). If the choice offered by LLM to managers converges with the managers own human intelligence then the managers become more certain of taking a decision (Mokander *et al.*, 2023). However, if the choice offered by LLM to managers diverged from the managers own human intelligence then the managers might become uncertain of taking a decision (Guler *et al.*, 2023). This would be because of the phenomenon of AI avoidance. It has often been pointed out that LLMs provided biased perspectives. For managers to undertake effective decision making the deployed LLM tool must be audited both at the planning as well as at the operations phase (Mokander *et al.*, 2023). Thus, AI auditing particularly for managers getting inputs for organizational decision making need to be carried out.

Table 2: Decision choice sets

	Choice option-1	Choice option- 2	Choice option- 3	Choice option- 4
Decision choice offered by LLM to a manager	No	No	Yes	Yes
Managers own individual intelligence-based choice	No	Yes	No	Yes
Resultant decision	Not selecting the option	Uncertainty regarding choice selection	Uncertainty regarding choice selection	Definitely selecting the option

Source: Authors' own conceptualization

Table 2 conceptualized the interaction between LLM-generated recommendations and a manager's individual judgment as a set of four decision scenarios, highlighting how alignment or misalignment between the two influenced decision outcomes. When both the LLM and the manager rejected an option (Choice 1), the outcome was clear and led to non-selection, indicating strong consensus and low cognitive conflict. In contrast, when there was disagreement, either the manager supported the option while the LLM did not (Choice 2), or the LLM recommended the option while the manager did not (Choice 3), the result was uncertainty, reflecting cognitive dissonance and the need for further evaluation, validation, or additional information. These scenarios emphasized the tension between human intuition and machine-generated insights. Finally, when both the LLM and the manager aligned in favor of the option (Choice 4), the decision became definitive, suggesting that convergence between artificial and human intelligence enhanced confidence and reduced ambiguity.

Table 3: Factors to be considered for LLM auditing

S. No	Factors	Remarks
1	Diverse and Representative Data	In e-commerce, use diverse and representative datasets during training to minimize biases related to product recommendations, user reviews, and purchasing behaviours. Ensure that data sources encompass a wide range of product categories, customer demographics, and geographical regions.
2	Bias Audits	Conduct thorough bias audits specific to the e-commerce domain. Analyse training data and model outputs to identify biases in product recommendations,

		pricing, and customer reviews. Utilize specialized tools and methodologies designed for bias detection in e-commerce.
3	Bias Mitigation Techniques	Implement bias mitigation techniques tailored to e-commerce applications. For instance, focus on de-biasing product recommendations to avoid promoting certain products or excluding others based on biased data.
4	Transparency and explanation ability	Make the decision-making processes of recommendation algorithms transparent and explainable to users. Explain why certain products are recommended and provide information about data sources.
5	User Feedback Mechanisms	Develop feedback mechanisms that allow e-commerce platform users to report biased or inappropriate product recommendations, reviews, or pricing. Actively address and correct issues based on user input.
6	Ethical Guidelines	Establish clear ethical guidelines for e-commerce operations. Ensure that recommendations, pricing strategies, and marketing practices adhere to ethical principles. Develop governance structures to ensure compliance with these guidelines.
7	Human-in-the-Loop Review	Incorporate human reviewers to evaluate and refine product recommendations, pricing, and reviews. Particularly in e-commerce, human oversight is critical for assessing and mitigating biases in real-time.
8	Regular Model Audits	Regularly audit and evaluate the performance of recommendation algorithms and pricing models for biases and ethical concerns. Continuously fine-tune these models to enhance fairness and accuracy.
9	Data Labelling Guidelines	Provide clear guidelines for data labelling, especially when human annotators are involved in labelling product data. Educate annotators about potential biases and ethical considerations specific to e-commerce contexts.
10	Privacy Protection	Implement robust data privacy protections to safeguard users' personal information and shopping history. Ensure compliance with e-commerce data protection regulations and privacy policies.
11	Context Sensitivity	Train e-commerce recommendation algorithms to consider the context of a user's browsing and purchasing history. Adapt recommendations based on user preferences and real-time behaviour.
12	Bias Impact Assessments	Conduct impact assessments to understand how biased product recommendations and pricing may affect different customer segments and communities within the e-commerce ecosystem.
13	Stakeholder Engagement	Engage with diverse stakeholders, including customers and sellers, to gather input and feedback on the impact and performance of product recommendations and pricing strategies. Address concerns and adapt to feedback.
14	Continuous Education and Training	Provide ongoing education and training to developers, e-commerce platform users, and reviewers to promote awareness of biases and ethical considerations specific to the e-commerce industry.
15	Third-Party Audits	Consider third-party audits and assessments of your e-commerce recommendation algorithms and pricing models to provide an independent evaluation of bias and ethics in the e-commerce context.
16	Open Collaboration	Collaborate with the e-commerce and AI communities to share research and best practices for mitigating bias and promoting ethical use of recommendation algorithms and pricing strategies.
Source: Authors' own conceptualization		

The factors presented in Table 3 were collectively examined to understand how effective auditing of LLMs in the e-commerce domain could be operationalized. The findings indicated that ensuring diverse and representative data was foundational, as it helped reduce biases in product recommendations, customer reviews, and purchasing behaviour. It was observed that datasets encompassing varied product categories, demographics, and geographic regions contributed to more balanced model outputs. The study further showed that bias audits played a critical role in identifying systemic issues within both training data and generated outputs. These audits enabled the detection of unfair patterns in pricing, recommendations, and reviews. In conjunction with this, the implementation of bias mitigation techniques was found to be essential in correcting such imbalances, particularly in preventing the over-promotion or exclusion of certain products.

Moreover, transparency and explainability were emphasized as key factors in building user trust. It was noted that when users were provided with clear explanations for recommendations and insights into data sources, their confidence in the system improved. Similarly, user feedback mechanisms were found to be valuable in capturing real-time issues, allowing platforms to address biased or inappropriate outputs effectively. The role of ethical guidelines was identified as a guiding framework that ensured recommendation systems and pricing strategies adhered to acceptable standards. Supporting this, human-in-the-loop review processes (Mosqueira-Rey *et al.,* 2023) were found to enhance oversight, enabling continuous monitoring and correction of model behaviour, especially in complex or sensitive scenarios. In addition, regular model audits were highlighted as necessary for maintaining long-term fairness and performance. These audits facilitated ongoing evaluation and fine-tuning of models. The importance of data labelling guidelines (Fredriksson *et al.,* 2020) was also evident, as clear instructions and awareness among annotators helped minimize the introduction of bias during dataset preparation.

Privacy protection (Yu *et al.,* 2024) emerged as a critical consideration, ensuring that users' personal and transactional data were safeguarded in compliance with relevant regulations. Alongside this, context sensitivity was found to improve recommendation relevance by adapting outputs based on user behaviour and preferences. The study also revealed that conducting bias impact assessments provided deeper insights into how different customer segments were affected, enabling more equitable system design. Stakeholder engagement further strengthened this process by incorporating feedback from customers, sellers, and other participants in the e-commerce ecosystem. Finally, continuous education and training were recognized as vital for maintaining awareness of ethical and bias-related challenges among developers and users (Hanna *et al.,* 2025). The inclusion of third-party audits added an additional layer of accountability by offering independent evaluations (Raji *et al.,* 2023). Furthermore, open collaboration with the broader AI and e-commerce communities was found to support the sharing of best practices and foster innovation in ethical AI deployment (Zhang *et al.,* 2026).

4. Proposed Reforms

To manage the biases in the business decisions we propose Coleman bathtub model (Coleman, 1986, 1987, 1990) to conduct LLM audits. The Coleman bathtub model (CBM) is a sociological framework developed by James Coleman to explain how individual actions within a social system contribute to broader social outcomes, focusing on the interplay between social structures and individual behaviors (Coleman, 1992). It is especially valuable in understanding how macro-level conditions shape micro-level actions, which in turn lead to outcomes that impact or reinforce the original social structure (Sultan *et al.,* 2024). Figure-1 provides an in-depth look at each component of the CBM, along with a breakdown of its core principles and applications in LLM audits.

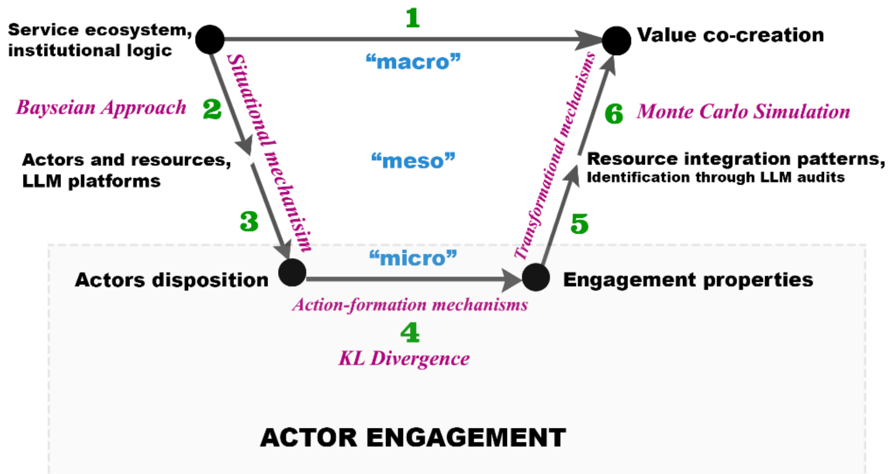

Figure 1: The Coltman Bathtub Model: LLM value co-creation through mathematical treaties for auditing (Source: Adapted from Storbacka *et al.,* 2016)

The model is called a "bathtub" because it resembles a flow that begins at a higher level (the macro level), goes down to an individual level (the micro level), and then returns to another collective level (macro-outcome) (Storbacka *et al.,* 2016). Each part of this flow represents a critical link between organizational structures, individual managers' decisions, and resulting business outcomes. The model has four key components:1) Macro-to-Micro Link (Top of the Bathtub)-organizational structures influence individual managers; 2) Micro-Level Decision-Making (Inside the Bathtub)- Individual Managers act based on LLM influences; 3) Micro-to-Macro Link (Bottom of the Bathtub)- Individual managers' actions aggregate into collective outcomes; 4) Feedback Loop- Outcomes at the macro level may reinforce or change initial organizational structures.

4.1 Detailed Breakdown of each component Coleman bathtub model

4.1.1. Macro-to-micro-Link: organizational structures influence individual managers

This first step represents how large-scale organizational factors such as norms, cultural values, laws, or institutional policies influence the way individual managers think and behave. Coleman (1987) suggested that individuals don't operate in a vacuum; rather, their decisions and actions are influenced by the environment they are part of. This can include everything from organizational policies to social expectations and market conditions. For example, in an organizational context, if a company has a strong innovation culture, managers may feel encouraged to experiment and take risks. This overarching structure influences individual decisions by setting the "rules" or expected behaviors within the organization.

4.1.2. Micro-Level Decision-Making: Individuals Make Choices

Once organizational structures have influenced individual motivations, managers act based on their own beliefs, incentives, and rationality. CBM emphasizes rational choice theory (Raub and Voss, 2017) here, suggesting that individuals make decisions that they believe will maximize their benefits or achieve their goals within the context provided by the macro structures. This step is highly variable, as managers interpret and react to business situations differently based on personal experiences, attitudes, and motivations. For example, managers might choose to contribute to new product development based on their perceived benefits, such as career growth or recognition. Even if the organization encourages innovation, managers' responses will vary based on their individual goals.

4.1.3. Micro-to-Macro Link: Collective Actions Lead to Social Outcomes

Individual actions don't happen in isolation. When many people within an organizational structure act in certain ways, their actions collectively lead to observable organizational outcomes. This aggregation of individual actions can create trends, societal shifts, or large-scale consequences that can then be observed and measured (Storbacka *et al.,* 2016). For example, if many employees in an organization are motivated by an innovation culture to develop new ideas, the cumulative effect might be a surge in product innovation. This collective outcome can improve the organization's market position, enhance its brand, and create a culture that attracts more talent.

4.1.4. Feedback Loop: Social Outcomes Reinforce or Transform Social Structures

The model includes a feedback loop where the outcomes generated can reinforce or change the original organizational structures (Storbacka *et al.,* 2016). If the collective outcome aligns well with organizational goals, it may reinforce the existing structures. However, if the outcomes are misaligned or lead to negative consequences, the macro conditions may be adjusted (Coleman, 1990). For example, if an organization's innovation culture leads to successful outcomes, it may reinforce the organization's focus on creativity and invest further in resources to support innovation. Conversely, if managers' actions lead to issues such as excessive risk-taking or reduced efficiency, the company may adjust its policies to create a more balanced environment (Coleman, 1992).

4.2 Application areas of the Coleman Bathtub Model

The CBM is particularly useful for analysing the dynamics within organizational systems, as it provides a structured approach to assess how collective outcomes emerge from individual actions within a framework of constraints and incentives (Coleman, 1992). Some common applications of CBM include:

- **Organizational Behaviour**: Understanding how organizational policies influence employee actions and lead to specific outcomes, such as productivity or innovation.
- **Public Policy Analysis**: Examining how governmental policies shape individual behaviours and result in social outcomes, like health, education, or economic activity.
- **Sociology and Social Change**: Analysing how cultural norms or social institutions shape individual behaviours and lead to societal changes over time.
- **LLM and AI Auditing**: Exploring how the structure (training data, model design) influences model behaviour (responses) and leads to macro-level impacts (user experience and societal impact).

In general, the CBM provides a way to understand complex social systems by dissecting the relationship between structures, individual behaviours, and collective outcomes (Coleman, 1992). By following this structured approach, organizations, policymakers, and sociologists can identify the forces at play within any organizational

Minnesota Journal of Business Law and Entrepreneurship
Volume 2026, No. 2
ISSN: 1540-3270

or social system, allowing them to predict outcomes, design interventions, and understand how individual actions contribute to broader social patterns (Coleman, 1990).

4.3 Coleman Bathtub Model and LLM Auditing

The CBM framework can be applied to LLM auditing by breaking down the complex relationships between system structures (e.g., training data, design choices), model behaviours, and the resultant impacts on end-users and society. In this context, it helps audit teams analyse how individual model actions (like generating responses) lead to broader outcomes and how systemic factors influence these actions. The model can help guide LLM audits to ensure transparency, fairness, and alignment with ethical standards. Figure 2 presents a flow chart of how the CBM could be applied for LLM auditing:

Figure 2: Application of CBM in LLM auditing

The CBM provides a structured approach for auditing LLMs by examining the influences on model behaviour (macro-level conditions), analysing model responses (micro-level actions), and evaluating the larger societal and user impacts (macro-level outcomes). By focusing on these interconnected layers, the model helps LLM auditors identify specific areas for improvement, ensure ethical behaviour, and refine LLMs to align more closely with intended purposes and values.

4.4. Mathematical treaties for auditing and reducing bias in LLMs:

There were multiple methods that could have been applied for auditing LLMs (Hadi *et al.,* 2023; McIntosh *et al.*, 2024). These methods were normally applied at the development stage of the LLM so as to reduce the biases from LLM outcomes (Peng *et al.,* 2024). The authors in this research focused on three major mathematical approaches to reduce the bias: The Bayesian Approach, Kullback-Leibler (KL) Divergence, and Monte Carlo Simulations. In the following section each of the approaches were discussed in details.

4.4.1. Bayesian Approach:

A Bayesian approach was used to audit and reduce bias in LLMs by modelling the uncertainty and potential sources of bias in the data and the model (González-Alday *et al.,* 2023).

Scenario presented to Managers: *You made the decision to offer consumers in all income ranges unbiased, customized product recommendations. In order to do this, you should utilize the following procedures to evaluate and reduce any potential bias in product recommendations depending on customers' income levels:*

1) Data Collection and Pre-processing:

Collect historical user interaction data, including user profiles, product views, purchases, and income information. Annotate the data with user income information that can be obtained from user profiles or inferred based on behaviour.

2***) Apply Bayesian Modelling for Bias Assessment:***

a) **Bayesian Data Model:**
Define a Bayesian model that describes the likelihood of a user engaging with a product given their income level, product category, and other relevant features:
P (Engagement | Income, Product Category, Features)

b) **Incorporate Prior Information:**
Include prior information that represents prior beliefs about potential bias in product recommendations based on income levels:
P (Parameters) ~ Prior

c) **Parameter Estimation:**
Use Bayesian inference techniques to estimate the model parameters, including the conditional probabilities of user interactions with products for different income brackets:
P (Parameters | Data, Income) ∝ P (Engagement | Income, Product Category, Features) * P (Parameters)

d) **Bias Quantification:**
Compute posterior distributions of model parameters that describe user behaviour for each income level, representing the uncertainty in the model:
Parameter ~ Distribution (a, b)

3) Bias Assessment:

a) **Hypothesis Testing:**
Formulate hypotheses related to income-based bias in product recommendations, such as: *H_0: There is no significant difference in user engagement between income brackets.*
H_1: There is a significant difference in user engagement between income brackets.
Use the posterior distributions and test statistics to perform Bayesian hypothesis testing and assess statistical significance.

b) **Credible Intervals:**
Compute credible intervals for model parameters to quantify the uncertainty in user behaviour for each income level. If the credible intervals do not overlap, it suggests significant differences in engagement.

4) Bias Mitigation:

a) **Prior Adjustment:**
Adjust the prior information used in the Bayesian model to reflect a more balanced representation of income-related product preferences, aiming to reduce bias:
P (Parameters) ~ Updated_Prior

b) **Fine-tuning:**
Fine-tune your recommendation algorithm using fairness-specific objectives to reduce bias in product recommendations across income levels.

c) **Regularization:**
Apply regularization techniques within the Bayesian model to encourage fairness and reduce disparities in user engagement with products between income brackets.

5) Evaluation and Monitoring:

Continuously monitor user interactions and product recommendations, updating the Bayesian model as new data becomes available.

Managers could explicitly model the uncertainty and integrate previous views about potential causes of bias in the suggestions by using a Bayesian method (Debnath *et al.,* 2023). This method allowed the managers to evaluate income bias in the product suggestions statistically and adjust it as needed.
The authors have depicted this in figure 3.

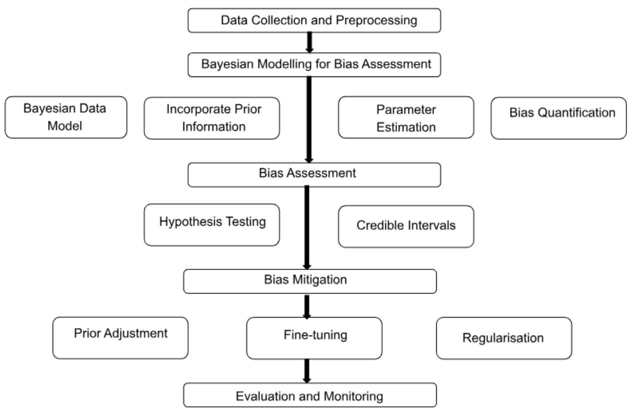

Figure 3: Bayesian approach for reducing LLM biases ***(Authors' own conceptualisation)***

4.4.2. Kullback-Leibler Divergence:

Kullback-Leibler (KL) divergence was a mathematical concept used to measure the difference between two probability distributions (Tang *et al.,* 2025). In the context of auditing and reducing bias in LLMs, KL divergence was employed to assess the disparity between the model's behaviour for different demographic groups (Mylrea and Robinson, 2023).

Scenario presented to Managers: *To spot possible shifts in consumer preferences, you were interested in comparing the similarity between two distributions of the buying behaviour of your customers. You had to use the following procedures to evaluate and reduce any potential bias in consumer preferences in order to accomplish this:*

1) Data Collection and Pre-processing:

Collect historical data on customer purchases, including product categories and quantities purchased.

2) Compute the Reference Distribution:

a) Select a Reference Period:

Choose the previous quarter as the reference period to represent the baseline behaviour.

b) Calculate the Reference Distribution:

Calculate the probability distribution of product category purchases during this reference period. This can be represented as:

P (Product Category | Reference Period)

c) Compute the Current Distribution:

Calculate the probability distribution of product category purchases for the current quarter. This can be represented as:

P (Product Category | Current Period)

d) Calculate KL Divergence:

Compute the KL divergence between the reference and current distributions using the following equation:

$$KL (P \| Q) = \sum (P(x) * \log(P(x) / Q(x)))$$

In this context:

- *P(x) represents the probability of a product category purchase in the reference distribution.*
- *Q(x) represents the probability of a product category purchase in the current distribution.*
- *The summation (∑) is performed over all product categories.*

e) Interpretation:

Interpret the KL divergence results as follows:

- A low KL divergence (close to 0) indicates that customer purchase behaviour in the current period is similar to the baseline, suggesting stable preferences.
- A high KL divergence suggests that there have been significant changes in customer preferences between the reference and current periods.

f) Threshold or Significance Level:

Set a threshold or significance level of 0.05 to determine the significance of KL divergence

g) Alert or Action:

If the computed KL divergence exceeds the threshold, generate an alert or take action to investigate the underlying causes of the change in customer preferences.

h) Monitoring and Feedback Loop:

Continuously monitor customer purchase behaviour and regularly recalculate the KL divergence to track changes over time. Utilize this feedback loop to modify product suggestions, inventory control, and marketing tactics to conform to changing consumer tastes.

Through the application of KL divergence, the discrepancy between the LLM's suggestions for two consumer purchasing behavior distributions and those for other demographic groups may be objectively measured. This method would assist in locating and addressing any potential bias in the model's content recommendations (Zhang et al., 2023). The authors have depicted this in figure 4.

ISSN: 1540-3270

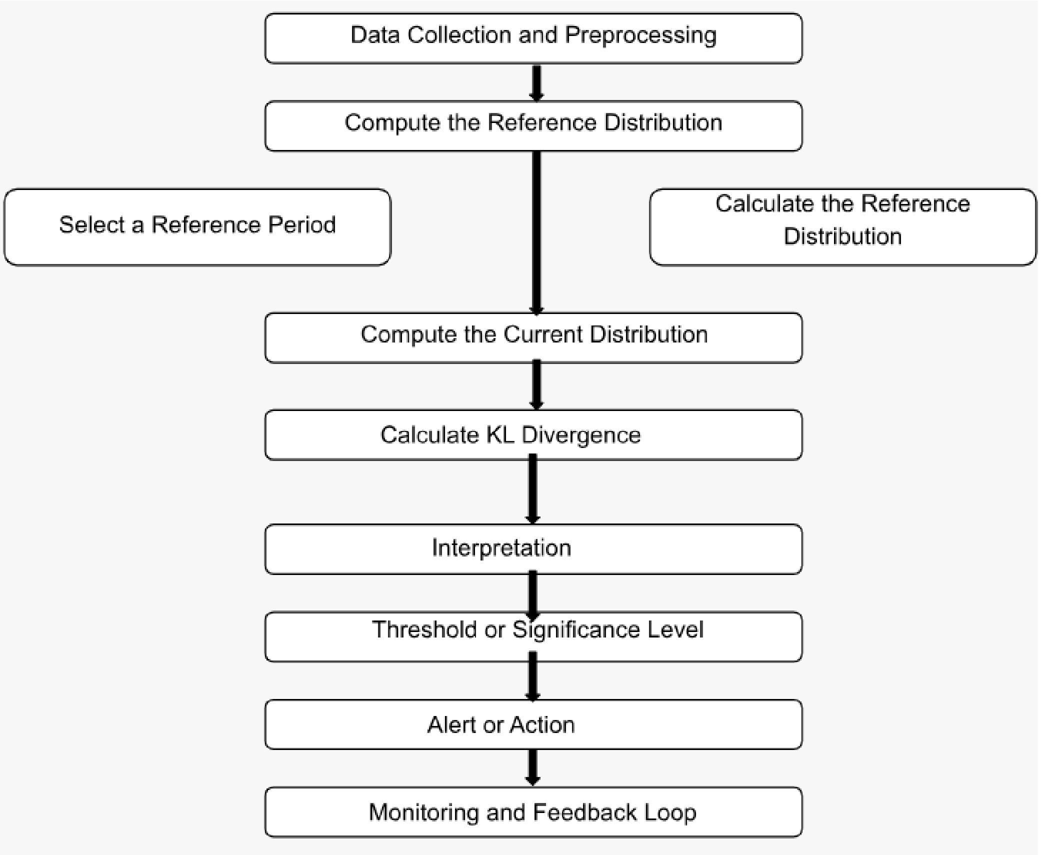

Figure 4: KL Divergence for reducing LLM biases (Authors' own conceptualization)

4.4.3. Monte Carlo Simulations:

Monte Carlo simulations could be used to assess and mitigate bias in LLMs by generating synthetic data and analysing how the model behaves with different demographic groups (Bazzoli, 2023).

Scenario presented to Managers: *You made the decision to offer consumers a portfolio of new products. For the same you had to assess the risk and potential return from the product portfolio. Use Monte Carlo simulations to estimate the portfolio's future value and the range of possible outcomes using the following steps:*

1) Data Collection and Pre-processing:

Gather historical data for different products, including daily or monthly price returns and any relevant financial information.

Calculate the expected returns (μ) and standard deviations (σ) for each stock in the portfolio based on historical data.

2***) Apply Monte Carlo Simulation:***

a) **Define the Portfolio**:

Specify the composition of the portfolio, including the allocation of funds to each new product.

b) Generate Random Returns:

Use Monte Carlo simulation to generate random returns for each new product. These returns should be sampled from probability distributions fitted to historical data.

For product price simulation, you can use the Geometric Brownian Motion Model:

$dP\_i (t) = \mu\_i * P\_i(t) * dt + \sigma\_i * P\_i(t) * dW\_i(t)$

Apply this model to generate random returns for the new products in each simulation iteration.

c) Simulate Portfolio Value:

Calculate the portfolio value for each simulated scenario by applying the generated returns to the portfolio composition:

$V (t) = \Sigma(W\_i * P\_i(t))$

d) Repeat Simulations:

Repeat the Monte Carlo simulations a large number of times (e.g., 10,000 iterations) to generate a distribution of possible future portfolio values.

ISSN: 1540-3270

e) Calculate Portfolio Metrics:
Calculate various portfolio metrics for each simulated scenario, including the final portfolio value, annualized return, standard deviation (portfolio risk), and other risk measures.

f) Generate Probability Distributions:
Analyse the distributions of these portfolio metrics.

g) **Risk Assessment**:
Assess the probability of various outcomes, such as the likelihood of portfolio losses or underperformance, based on the simulated distributions.

h) Portfolio Optimization:
Use linear programming techniques to select an optimal portfolio allocation that balances risk and return based on simulation results.

i) Decision-Making:
Make informed investment decisions based on the analysis of the Monte Carlo simulations. Adjust the portfolio allocation or investment strategy as needed to align with risk tolerance and return goals.

j) Monitoring and Scenario Testing:
Continuously monitor the portfolio's performance and update the Monte Carlo simulations as new data becomes available or investment objectives change. Conduct scenario testing by modifying input parameters such as expected returns and standard deviations to assess how the portfolio may perform under different economic conditions.

By using Monte Carlo simulations, one can create controlled experiments that simulate different scenarios and assess how your LLM behaves with synthetic data (Chen *et al.,* 2024). This approach allows detecting and quantifying bias, evaluating the effectiveness of bias mitigation strategies, and improving the fairness of recommendations provided by the model (Pierson *et al.,* 2022). The authors depicted this in figure 5.

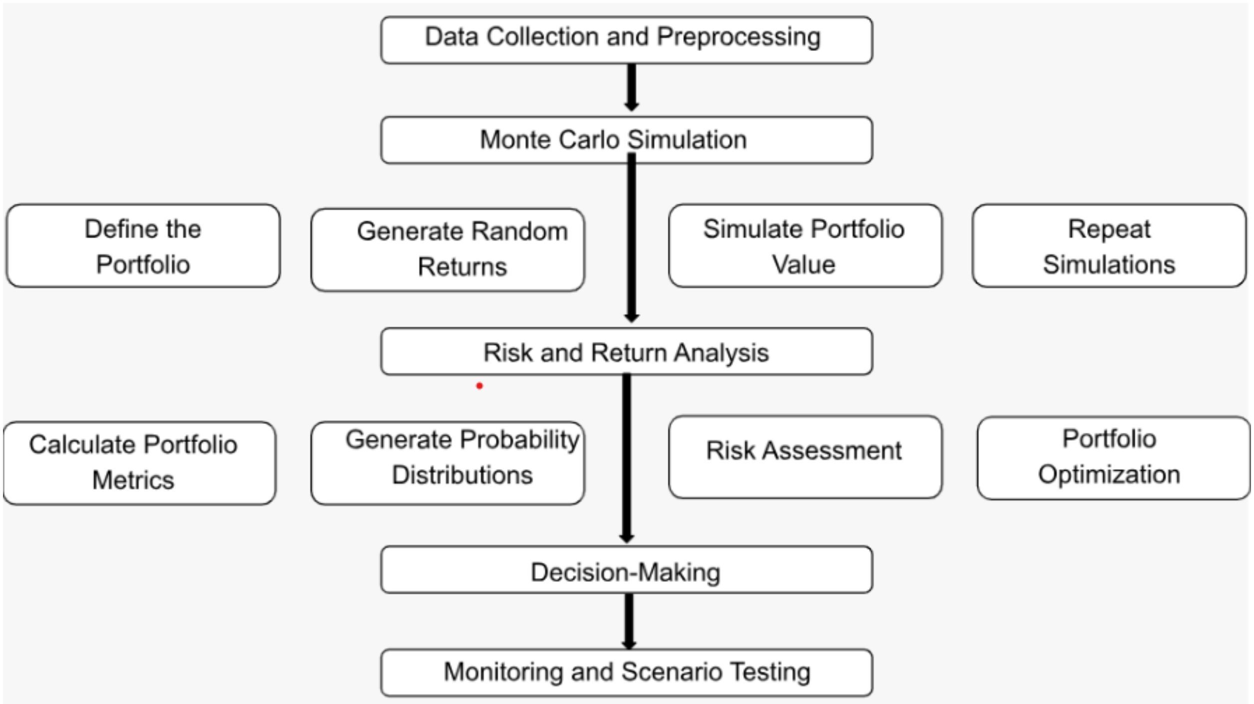

Figure 5: Monte Carlo Simulation for reducing LLM biases (*Authors' own conceptualization*)

5. Future Implications for Businesses and Entrepreneurship

The findings of this research carried significant implications for businesses and entrepreneurship, particularly as organizations increasingly integrated Large Language Models (LLMs) into e-commerce ecosystems. It was evident that firms adopting structured auditing frameworks would be better positioned to build trustworthy and responsible Artificial Intelligence (AI) - driven platforms, which in turn could enhance customer loyalty and brand credibility. Businesses that proactively addressed bias, transparency, and ethical concerns were likely to gain a competitive advantage in markets where consumers were becoming more aware of algorithmic fairness and accountability. From an entrepreneurial perspective, the research suggested the emergence of new venture opportunities centred around AI auditing, bias detection tools, and ethical compliance solutions. Startups could leverage these insights to develop specialized services such as automated bias monitoring systems, explainability tools, and third-party auditing platforms tailored for e-commerce. This indicated a growing ecosystem where ethical AI was not only a compliance requirement but also a value-creating domain.

The study also implied that organizations would need to invest in interdisciplinary capabilities, combining technical expertise with ethical governance and regulatory awareness. This shift could reshape hiring practices, encouraging the inclusion of roles such as AI ethicists, audit specialists, and compliance officers. For entrepreneurs, this opened pathways to build consultancies and training platforms focused on equipping businesses with these competencies. In addition, the emphasis on user feedback mechanisms and stakeholder engagement pointed toward a more participatory model of innovation. Businesses that incorporated customer and seller feedback into algorithmic improvements were likely to achieve more inclusive and adaptive systems. This participatory approach could redefine customer relationships, transforming users from passive consumers into active contributors to platform governance.

The importance of privacy protection and regulatory compliance highlighted in the research also suggested that future businesses would operate in increasingly regulated environments. Entrepreneurs would need to design solutions that aligned with evolving data protection laws and ethical standards across different regions. This could drive the development of scalable compliance technologies and region-specific customization strategies. Furthermore, the research indicated that continuous auditing and model refinement would become standard operational practices rather than one-time interventions. This ongoing process could influence business models by introducing recurring costs but also creating opportunities for subscription-based auditing services and AI maintenance solutions. Finally, the emphasis on open collaboration and third-party audits suggested a shift toward more transparent and cooperative innovation ecosystems. Businesses that engaged in knowledge sharing and adopted industry best practices were likely to accelerate innovation while minimizing risks. For entrepreneurs, this created opportunities to build platforms that facilitated collaboration, benchmarking, and shared learning across organizations.

6. Conclusion

Reducing bias in Large Language Models (LLMs) was paramount for ensuring their utility in managerial decision-making processes. As organizations increased their reliance on these models to analyse vast datasets and generate insights, addressing bias had become crucial for fair and informed decision-making. By implementing strategies such as using diverse and representative data, conducting bias audits, and employing mitigation techniques, businesses have tried enhancing the ethical foundations of LLMs. Reducing bias in LLMs was not only in alignment with principles of fairness and inclusivity but also had safeguarded against unintentional reinforcement of societal prejudices. Ethical considerations were especially pertinent in managerial decisions that impact diverse stakeholders. Transparent decision-making processes, user feedback mechanisms, and continuous model audits had contributed to a culture of accountability and responsiveness.

The utility of LLMs in managerial decisions lied in their ability to process and interpret complex information, automate routine tasks, and generate valuable insights. In areas such as risk assessment, market analysis, and customer engagement, unbiased LLMs were capable of offering a more accurate and holistic view, empowering managers to make well-informed choices. Moreover, LLMs also contributed in efficiency gains, freeing up managerial time for strategic thinking and innovation. In conclusion, mitigating bias in LLMs enhanced their utility for managerial decisions by fostering fairness, transparency, and reliability. Organizations that prioritized bias reduction in LLM deployment not only adhere to ethical standards but also were able to harness the full potential of these models to drive informed and equitable managerial strategies.

References

[1] Bankins, S., & Formosa, P. (2023). The ethical implications of artificial intelligence (AI) for meaningful work. *Journal of Business Ethics*, *185*(4), 725-740.

[2] Bazzoli, A. (2023). *Analyzing workplace accident underreporting: a systematic review, a Monte Carlo simulation, and a real data application*. Washington State University.

[3] Birhane, A., & Prabhu, V. U. (2021, January). Large image datasets: A pyrrhic win for computer vision?. In *2021 IEEE Winter Conference on Applications of Computer Vision (WACV)* (pp. 1536-1546). IEEE.

[4] Chen, Q., Liu, J., & Yang, B. (2024). Kinetic Monte Carlo Simulations Coupling the Isomerization and Diffusion of Xylenes in HZSM-5. *The Journal of Physical Chemistry C*, *128*(18), 7529-7535.

[5] Coleman, J. S. (1986). Social theory, social research, and a theory of action. *American journal of Sociology*, *91*(6), 1309-1335.

[6] Coleman, J. S. (1987). Microfoundations and macrosocial behavior. *The micro-macro link*, *153*, 173.
[7] Coleman, J. S. (1990). *Foundations of social theory*. Harvard university press.
[8] Coleman, J. S. (1992). The vision of foundations of social theory. *Analyse & Kritik*, *14*(2), 117-128.
[9] Debnath, B., Shakur, M. S., Bari, A. M., & Karmaker, C. L. (2023). A Bayesian Best–Worst approach for assessing the critical success factors in sustainable lean manufacturing. *Decision Analytics Journal*, *6*, 100157.
[10] De Caigny, A., Coussement, K., Verbeke, W., Idbenjra, K., & Phan, M. (2021). Uplift modeling and its implications for B2B customer churn prediction: A segmentation-based modeling approach. *Industrial Marketing Management*, *99*, 28-39.
[11] Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., ... & Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International journal of information management*, *57*, 101994.
[12] Fredriksson, T., Mattos, D. I., Bosch, J., & Olsson, H. H. (2020, November). Data labeling: An empirical investigation into industrial challenges and mitigation strategies. In *International conference on product-focused software process improvement* (pp. 202-216). Cham: Springer International Publishing.
[13] González-Alday, R., García-Cuesta, E., Kulikowski, C. A., & Maojo, V. (2023). A scoping review on the progress, applicability, and future of explainable artificial intelligence in medicine. *Applied Sciences*, *13*(19), 10778.
[14] Guler, N., Cahalane, M., Kirshner, S., & Vidgen, R. (2025). The role of roles: When are LLMs behavioural in information systems decision-making. *Australasian Journal of Information Systems*, *29*.
[15] Hadi, M. U., Qureshi, R., Shah, A., Irfan, M., Zafar, A., Shaikh, M. B., ... & Mirjalili, S. (2023). Large language models: a comprehensive survey of its applications, challenges, limitations, and future prospects. *Authorea preprints*, *1*(3), 1-26.
[16] Hanna, M. G., Pantanowitz, L., Jackson, B., Palmer, O., Visweswaran, S., Pantanowitz, J., ... & Rashidi, H. H. (2025). Ethical and bias considerations in artificial intelligence/machine learning. *Modern Pathology*, *38*(3), 100686.
[17] Kanbach, D. K., Heiduk, L., Blueher, G., Schreiter, M., & Lahmann, A. (2024). The GenAI is out of the bottle: generative artificial intelligence from a business model innovation perspective. *Review of Managerial Science*, *18*(4), 1189-1220.
[18] Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences*, *103*, 102274.
[19] Kar, A. K., Varsha, P. S., & Rajan, S. (2023). Unravelling the impact of generative artificial intelligence (GAI) in industrial applications: A review of scientific and grey literature. *Global Journal of Flexible Systems Management*, *24*(4), 659-689.
[20] Leavy, S. (2020). Uncovering gender bias in media coverage of politicians with machine learning. *arXiv preprint arXiv:2005.07734*.
[21] McIntosh, T. R., Liu, T., Susnjak, T., Watters, P., & Halgamuge, M. N. (2024). A reasoning and value alignment test to assess advanced gpt reasoning. *ACM Transactions on Interactive Intelligent Systems*, *14*(3), 1-37.
[22] Meyer, J. G., Urbanowicz, R. J., Martin, P. C., O'Connor, K., Li, R., Peng, P. C., ... & Moore, J. H. (2023). ChatGPT and large language models in academia: opportunities and challenges. *BioData mining*, *16*(1), 20.
[23] Mökander, J., Sheth, M., Watson, D. S., & Floridi, L. (2023). The switch, the ladder, and the matrix: Models for classifying AI systems. *Minds and Machines*, *33*(1), 221-248.
[24] Mosqueira-Rey, E., Hernández-Pereira, E., Alonso-Ríos, D., Bobes-Bascarán, J., & Fernández-Leal, Á. (2023). Human-in-the-loop machine learning: a state of the art. *Artificial Intelligence Review*, *56*(4), 3005-3054.
[25] Mylrea, M., & Robinson, N. (2023). Artificial Intelligence (AI) trust framework and maturity model: applying an entropy lens to improve security, privacy, and ethical AI. *Entropy*, *25*(10), 1429.
[26] Osterwalder, A., Pigneur, Y., Oliveira, M. A. Y., & Ferreira, J. J. P. (2011). Business Model Generation: A handbook for visionaries, game changers and challengers. African journal of business management, 5(7), 22-30.
[27] Osterwalder, A., Pigneur, Y., Smith, A., & Etiemble, F. (2020). The invincible company: how to constantly reinvent your organization with inspiration from the world's best business models . Haboken , New Jersey; John Wiley & Sons.
[28] Park, I., Cho, J., & Rao, H. R. (2015). The dynamics of pre-and post-purchase service and consumer

evaluation of online retailers: a comparative analysis of dissonance and disconfirmation models. *Decision Sciences*, *46*(6), 1109-1140.
[29] Peng, J. L., Cheng, S., Diau, E., Shih, Y. Y., Chen, P. H., Lin, Y. T., & Chen, Y. N. (2024). A survey of useful llm evaluation. *arXiv preprint arXiv:2406.00936*.
[30] Pierson, D., Lohse, K. A., Wieder, W. R., Patton, N. R., Facer, J., de Graaff, M. A., ... & Will, R. (2022). Optimizing process-based models to predict current and future soil organic carbon stocks at high-resolution. *Scientific Reports*, *12*(1), 10824.
[31] Pigneur, Y., & Osterwalder, A. (2010). Business model generation: A handbook for visionaries, game changers, and challengers. Haboken , New Jersey; John Wiley & Sons.
[32] Raji, I. D., Chock, S. C., & Buolamwini, J. (2023). Change from the outside: towards credible third-party audits of AI systems. *Missing links in AI governance*, *5*.
[33] Raub, W., & Voss, T. (2017). Micro-macro models in sociology: Antecedents of Coleman's diagram. *Social dilemmas, institutions, and the evolution of cooperation*, 11-36.
[34] Ruan, J., Chen, Y., Zhang, B., Xu, Z., Bao, T., Du, G., ... & Zhao, R. (2023). TPTU: large language model-based AI agents for task planning and tool usage. *arXiv preprint arXiv:2308.03427*.
[35] Schiff, D., Biddle, J., Borenstein, J., & Laas, K. (2020, February). What's next for ai ethics, policy, and governance? a global overview. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (pp. 153-158).
[36] Schwartz, R. (2022). Towards a standard for identifying and managing bias in artificial intelligence.
[37] Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of business research*, *70*, 263-286.
[38] Storbacka, K., Brodie, R. J., Böhmann, T., Maglio, P. P., & Nenonen, S. (2016). Actor engagement as a microfoundation for value co-creation. *Journal of business research*, *69*(8), 3008-3017.
[39] Sultan, M., Hussain, G., Ismail, W. K. W., & Rashid, M. A. (2024). From entrepreneurial leadership to new product development performance: A study of the Coleman bathtub model. *European Journal of Innovation Management*, *27*(8), 2623-2645.
[40] Tang, Y., Ma, Y., & Li, B. (2025). A KL-divergence-based test for elliptical distribution. *Journal of Nonparametric Statistics*, 1-27.
[41] Tokayev, K. J. (2023). Ethical implications of large language models a multidimensional exploration of societal, economic, and technical concerns. *International Journal of Social Analytics*, *8*(9), 17-33.
[42] Yu, S., Carroll, F., & Bentley, B. L. (2024). Insights into privacy protection research in AI. *Ieee Access*, *12*, 41704-41726.
[43] Voelpel, S. C., Leibold, M., & Tekie, E. B. (2004). The wheel of business model reinvention: how to reshape your business model to leapfrog competitors. *Journal of change management*, *4*(3), 259-276.
[44] Walsh, K. R., Mahesh, S., & Trumbach, C. C. (2021). Autonomy in AI systems. *The Journal of Technology Studies*, *47*(1), 38-47.
[45] Zhang, Y., Pan, J., Li, L. K., Liu, W., Chen, Z., Liu, X., & Wang, J. (2023). On the properties of kullback-leibler divergence between multivariate gaussian distributions. *Advances in neural information processing systems*, *36*, 58152-58165.
[46] Zhang, X., Li, K., Wu, Y., Liang, S., & Yu, M. (2026). Transforming E-Commerce with AI: Navigating Innovation, Personalization, and Ethical Challenges. *Journal of Theoretical and Applied Electronic Commerce Research*, *21*(1), 29.
[47] Zhou, X., Wang, Q., Wang, X., Tang, H., & Liu, X. (2023). Large language model soft ideologization via AI-self-consciousness. *arXiv preprint arXiv:2309.16167*.