Journal of Infrastructure, Policy and Development 2025, 9(2), 11489.
https://doi.org/10.24294/jipd11489

Review

A systematic review of algorithm auditing processes to assess bias and risks in AI systems

Vusumzi Funda

University of Fort Hare, Dikeni 5700, South Africa; vfunda@ufh.ac.za

CITATION

Funda V. (2025). A systematic review of algorithm auditing processes to assess bias and risks in AI systems. Journal of Infrastructure, Policy and Development. 9(2): 11489. https://doi.org/10.24294/jipd11489

ARTICLE INFO

Received: 7 February 2025
Accepted: 24 March 2025
Available online: 16 May 2025

Abstract: The expanding adoption of artificial intelligence systems across high-impact sectors has catalyzed concerns regarding inherent biases and discrimination, leading to calls for greater transparency and accountability. Algorithm auditing has emerged as a pivotal method to assess fairness and mitigate risks in applied machine learning models. This systematic literature review comprehensively analyzes contemporary techniques for auditing the biases of black-box AI systems beyond traditional software testing approaches. An extensive search across technology, law, and social sciences publications identified 22 recent studies exemplifying innovations in quantitative benchmarking, model inspections, adversarial evaluations, and participatory engagements situated in applied contexts like clinical predictions, lending decisions, and employment screenings. A rigorous analytical lens spotlighted considerable limitations in current approaches, including predominant technical orientations divorced from lived realities, lack of transparent value deliberations, overwhelming reliance on one-shot assessments, scarce participation of affected communities, and limited corrective actions instituted in response to audits. At the same time, directions like subsidiarity analyses, human-centered tools, and corrective programming offer templates to advance auditing processes as embedded socio-technical instruments supporting context-specific translation of signals into governing actions. Substantial innovation remains necessary for institutionalizing continuous, holistic, and participative auditing capabilities that can steward equitable algorithm development rather than remain detached arbiters.

Keywords: algorithm auditing; AI bias; machine learning fairness; algorithmic accountability; technical assessments; participatory auditing

1. Introduction

The exponential advancement and adoption of artificial intelligence (AI) technologies across critical domains have prompted broader discussions around the potential for unintended consequences. AI systems built utilizing machine learning are now being rapidly embedded in high-stakes sectors like employment, healthcare, criminal justice and finance (Fioretto et al., 2018; Leite et al., 2014; Slootjes, 2017). By analyzing large datasets, these algorithmic systems aim to generate influential predictions, optimizations and recommendations that assist or replace human decision-making. Proponents highlight possibilities for reducing human biases and increasing efficiency. However, revelations regarding AI failures and harms, especially towards marginalized groups, have catalyzed fears of "black box" technologies with little accountability and calls for urgent assessments through auditing processes.

Core issues that have galvanized algorithm audits include lack of transparency, perpetuation of historical biases, and lack of validation across impacted populations (Creswell and Creswell, 2017; Leite et al., 2014; Macal, 2016). Modern machine learning approaches involve tremendous complexity with hundreds of interdependent

variables and data dimensions utilized by models. This emergence of "black box" systems with opacity around internal logic has raised concerns regarding the ability to audit for issues. Additionally, models developed using historical training data run the risk of inheriting and amplifying societal biases and inequities around race, gender, and other attributes. The reliance on supervised learning techniques on labeling by potentially flawed human decisions further enables the creeping of systemic prejudices into AI systems. Real-world deployments have also demonstrated challenges around generalizing to underrepresented demographic groups not reflected in the initial design and testing phases. Across these core areas of transparency, bias, and validity, a range of auditing processes has emerged to diagnose root causes and address ethical risks.

Despite the growing recognition of algorithm auditing as a vital mechanism for assessing fairness and mitigating AI risks, it faces significant challenges that distinguish it from traditional software testing. These challenges can be broadly categorized into three key areas:

A) Transparency issues: Algorithm auditing is particularly complex due to the "black box" nature of many AI systems, especially those leveraging deep learning models. Unlike traditional software, which follows predefined logic flows, AI decision-making processes are often non-interpretable, making it difficult for auditors to trace and understand the rationale behind specific outputs. This lack of transparency is especially problematic in high-stakes applications such as healthcare and finance, where interpretability is critical for ensuring trust, accountability, and compliance with regulatory requirements. Algorithm auditing can help build this trust by ensuring the reliability and fairness of AI applications in educational contexts (Lavidas et al., 2024).

B) Bias detection and mitigation: A central challenge in algorithm auditing is the identification and mitigation of biases. AI models can perpetuate and even amplify societal biases, leading to discriminatory outcomes in employment, healthcare, and financial services. As per Aravantinos et al. (2024), AI systems are used to personalize learning experiences, and thus algorithm auditing becomes crucial in detecting and mitigating potential biases that could disadvantage certain student groups. Addressing these biases is complex, as they can originate from multiple sources, including biased training data, flawed algorithmic assumptions, or even the unconscious biases of developers. Furthermore, there is no universally accepted methodology for detecting, quantifying, or mitigating bias, making consistency in algorithm audits difficult to achieve.

C) Technical complexities: Unlike traditional software, AI models continuously learn and evolve, requiring auditors to implement dynamic monitoring mechanisms rather than static validation tests. Traditional software testing methods, which rely on deterministic outcomes, are insufficient for auditing AI models, as their behavior may shift based on new data inputs. Additionally, even AI developers themselves may not fully comprehend how complex deep learning models arrive at their conclusions, adding another layer of difficulty to the auditing process.

Algorithm auditing differs fundamentally from traditional software testing in several key ways. While traditional software testing focuses on ensuring functional

correctness, performance, and reliability, algorithm auditing extends beyond these aspects to evaluate ethical considerations, fairness, accountability, and the societal impact of AI systems. Unlike software testing, which relies on predefined test cases with expected outputs, algorithm auditing employs statistical analyses, scenario simulations, and demographic impact assessments to evaluate model behavior across diverse population groups. Additionally, traditional software testing primarily addresses technical reliability, whereas algorithm auditing places a strong emphasis on ethical concerns, fairness, and potential discriminatory effects in AI decision-making. Another key distinction lies in the temporal aspects—traditional software testing occurs at specific points in the development lifecycle, while algorithm auditing requires ongoing monitoring and reassessment to account for evolving model behavior. Furthermore, unlike software testing, which is largely confined to computer science and engineering, algorithm auditing necessitates expertise from multiple fields, including law, ethics, social sciences, and data science, to comprehensively assess AI's impact on society. By incorporating these considerations, the study aims to advance discussions on how algorithm auditing can be refined to overcome these challenges and provide more rigorous assessments of AI systems.

Prevailing algorithm auditing approaches focus on assessing training data composition, evaluating model features, and testing performance across subgroups (Creswell and Creswell, 2017; Macal, 2016). Training data is analyzed for representation gaps across population segments that could propagate biases through AI systems. Simulated biased data helps characterize resulting distortions. Model inspections surface suspicious correlations between sensitive attributes like race, age, or gender and key model features or decisions. Counterfactual testing modifies attributes to quantify impacts on outputs. Subgroup validation checks for performance consistency across slices of the population carrying greater ethical risks, including minorities. Beyond these quantitative methods, audits also increasingly entail real-world testing approaches like A/B trials against legacy systems and participatory assessments engaging affected communities (Arnold et al., 2017; Creswell and Creswell, 2017; Woit, 2017).

While the adoption of auditing processes has expanded, limitations persist around the poor translation of audits into interventions, challenges replicating real-world complexity, surface-level assessments divorced from business contexts and technology lifecycles, resource burdens for rigorous validations, and the need for greater stakeholder participation (Arnold et al., 2017; Kacprzyk and Pedrycz, 2015; Macal, 2016; Slootjes, 2017). Technical data or model testing also outweighs holistic evaluation of organizational and market dynamics, enabling unfairness. Our systematic literature review consolidates knowledge in this space, critiquing current techniques and charting promising directions.

Overall, while increased auditing represents progress, processes remain inconsistent, narrowly focused, and rarely integrated into technology development lifecycles or translated into deployed safeguards (Ehsan et al., 2017; Sandhu et al., 1996). As algorithms expand across critical domains, the ability to conduct rigorous, comprehensive, and actionable audits constitutes an urgent governance priority with major ethical implications. Our analysis addresses this need by appraising merits and gaps in existing literature to inform policies and practices around equitable and

accountable AI systems.

2. Materials and methods

This section elaborates on the systematic review methodology followed to search, select, analyze, and synthesize literature on algorithm auditing processes. It covers the search strategy, study selection criteria, data analysis methods, and quality appraisal approach along with relevant summary tables.

2.1. Search strategy

A rigorous search queried major technology, law, and social sciences databases to identify relevant studies focused on auditing processes for AI systems, particularly those related to ethical concerns and fairness. As per **Table 1**, to ensure comprehensive coverage across disciplines, we implemented a broad search strategy that included:

Use of diverse keywords such as "algorithm auditing," "AI bias evaluation," "ethical AI assessment," and "fairness in machine learning."

Table 1. Sample search query examples.

Database	Search String
ACM Digital Library	"algorithm" OR "AI" OR "machine learning" AND "auditing" OR "testing" OR "inspection" AND "fairness" OR "bias" OR "error analysis"
JSTOR	"algorithmic system" OR "artificial intelligence" OR "predictive model" AND "auditing" OR "assessment" AND "fairness" OR "accountability"

Searching multiple databases that cover computer science (e.g., ACM Digital Library, IEEE Xplore), social sciences (e.g., JSTOR), and interdisciplinary repositories (e.g., arXiv, SSRN).

Initial search string formulations involved permutations of terminology related to: "algorithm", "artificial intelligence", "machine learning", "predictive model", "auditing", "testing", "inspection", "fairness assessment", "bias evaluation", "error analysis". Wildcards and boolean operators expanded variants like "algorithm*" or "AI AND (audit* OR assess*)". Additional filters narrowed results to journal articles, conference papers, or edited volumes discussing audits applied to algorithmic systems leveraged for individual-level decisions with major social impacts. Article titles, abstracts, and full texts guided relevance screening.

2.2. Study selection process

A systematic procedure guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework was used to identify highly relevant studies for in-depth analysis. This encompassed specifying eligibility criteria, screening articles through phases, and mapping exclusions.

2.2.1. Inclusion and exclusion criteria

To ensure a comprehensive and systematic review of algorithm auditing literature, the PRISMA framework was adapted and applied in key ways to address the challenges of integrating diverse disciplinary perspectives. Originally developed for health sciences, the framework was modified to accommodate the complexities of

cross-disciplinary integration, particularly in fields spanning computer science, ethics, law, and social sciences (Moher et al., 2009). This adaptation involved expanding search terms to capture relevant terminology across disciplines and refining inclusion criteria to ensure balanced representation of both technical and socio-ethical perspectives. Additionally, an iterative screening process was implemented to navigate the complexities of cross-disciplinary literature selection. This process began with a broad initial screening to identify potentially relevant studies across diverse fields, followed by secondary screening using more specific criteria to ensure the inclusion of research that addressed both technical and ethical aspects of algorithm auditing. Studies were evaluated for inclusion based on criteria along two axes specified in **Table 2**.

Table 2. Study eligibility criteria.

Dimension	Inclusion	Exclusion
Relevance	• Studies focusing on AI/ML auditing methodologies • Research addressing high-impact decision-making contexts	• Studies solely focused on technical aspects without ethical considerations
Rigor	• Publications presenting systematic auditing approaches • Studies discussing ethical implications and societal impacts	• Research not providing sufficient methodological details • Publications not peer-reviewed or from non-reputable sources

To capture both recent developments and foundational works, we included studies published from 2018, with a focus on the most recent publications to reflect current trends and challenges.

2.2.2. PRISMA framework

The PRISMA protocol provides an evidence-based mechanism for systematically selecting studies, particularly relevant to scoping literature reviews. Rather than a strict formula, it offers a guideline customized to review goals for framing objective inclusion criteria, detailing screening procedures, mapping exclusions across phases, and diagramming the path from initial search results to final selected studies. A core tenet emphasizes the specification of review questions and objectives a priori to anchor the search strategy. The process then documents a transparent pathway from casting a wide evidence net through iterative filtering based on eligibility criteria. For the current analysis, the PRISMA approach shaped customizing inclusion criteria, balancing relevance on auditing AI systems for high-impact decisions with methodological rigor. For this study, we define high-impact decisions as those that significantly affect individuals' lives, often involving critical areas such as healthcare, criminal justice, employment, and financial services. These decisions have long-term implications for individuals' rights, opportunities, and well-being. But as later sections elaborate, the review expanded the framework through an equitable AI analytical lens and integrated recommendation development.

2.2.3. Selected studies

Applying the PRISMA framework, the selection process began with 467 initial records identified through searches across databases. After removing duplicates and screening out 264 irrelevant titles/abstracts, 203 candidate studies were left for full-

text review. Further assessing alignment with the rigorous inclusion criteria in **Table 3** filtered to 51 relevant papers discussed substantive algorithm auditing processes related to high-impact decisions. Thoroughly examining methodology strengths and knowledge synthesis potential, 22 exemplar studies published from 2018 to 2025 across leading venues were finally chosen. This diverse sample balancing technical depth with qualitative evaluations through semi-structured interviews, observational audits, and field trial interventions enabled multifaceted analysis of contemporary techniques.

2.3. Quality appraisal

Studies underwent quality appraisal assessing methodology rigor, reproducibility, contextual awareness, and evidentiary value. Dimensions included:

(1) Systematic validation processes vs. haphazard testing.
(2) Replicable protocols detailing datasets, parameters, and tools.
(3) Attentiveness to limitations around assumptions, scope constraints, and evaluation choices.
(4) Justified inferences connected to findings, data, and methods.

To account for the diverse methodologies and reporting standards across disciplines, the quality assessment criteria were adapted to ensure a balanced evaluation of technical and ethical considerations. A custom quality assessment tool was used to systematically assess both the technical rigor of algorithm auditing methodologies and the depth of ethical reasoning applied in each study. Additionally, the weighting of quality criteria was carefully adjusted to reflect the significance of both technical precision and socio-ethical implications, ensuring a comprehensive and contextually relevant assessment of algorithm auditing research. Rather than excluding studies based on rigid quality score cutoffs, critical scrutiny of reliability and relevance occurred integratively during analysis. The goal balanced consolidating state-of-the-art knowledge with transparently examining the utility and scientific credibility.

2.4. Data analysis

A two-phase methodology guided the data analysis strategy. Initial coding through NVivo characterized details of the selected studies across dimensions such as publication year, application domain, datasets utilized, procedural specifics, measured evaluation criteria, detected model issues, and actions like reforms.

To ensure a structured approach that addressed cross-disciplinary challenges, the PRISMA framework facilitated:

- Use of a standardized data extraction form that captured both technical details (e.g., auditing methodologies, metrics) and ethical considerations (e.g., fairness definitions, societal impact).
- Thematic analysis that identified common threads across disciplines, highlighting areas of convergence and divergence in approaches to algorithm auditing.

2.5. Addressing terminological differences

To overcome the challenge of differing terminologies across disciplines, a

glossary of key terms was developed, mapping equivalent concepts from different fields. This glossary was used during the screening and data extraction processes to ensure consistent interpretation of concepts across disciplines. By implementing these adaptations, the PRISMA framework enabled a rigorous and transparent process for integrating cross-disciplinary literature on algorithm auditing. This approach ensured that insights from diverse fields were systematically captured and synthesized, providing a comprehensive view of the current state of algorithm auditing research and practice.

3. Results and discussion

This section provides a structured overview of the 22 studies selected for in-depth review and analysis in this systematic literature review on algorithm auditing techniques. Studies were selected to provide diverse coverage of techniques spanning quantitative benchmarking, model inspections, participatory assessments, and deployment monitoring while maintaining methodological rigor. The description of selected studies is followed by a critical thematic analysis in subsequent sections evaluating the strengths, limitations, and gaps in current auditing processes across key ethical dimensions of depth, transparency, contextualization, and actions or interventions undertaken based on audit findings.

3.1. Overview of selected studies

Table 3 summarizes key details of the 22 publications reviewed in this analysis, including author(s), year, technique(s) employed, application area, datasets used, and the core focus in terms of types of biases/issues addressed. As seen, the selected articles span from 2014 to 2025, reflecting contemporary research at the interface of responsible AI, algorithm auditing, and machine learning fairness. They encompass diverse techniques like causal analysis, counterfactual analysis, subgroup clustering, and participatory auditing. Application domains range from high-stakes sectors like employment, lending, policing, and clinical risk predictions to online platforms and autonomous systems. Across these contexts, the set of studies attempts to address varied algorithmic issues, including representation gaps, proxy discrimination, unfair subgroup impacts, and unsafe failure modes.

Table 3. Summary overview of selected studies.

Author(s)	Technique(s)	Application area	Dataset(s)	Core focus
Sharma and Wehrheim (2020)	Verification-based testing	Lending, employment	Adult, German credit	Individual discrimination
Bellamy et al. (2018)	Bias metrics, model inspection, testing suite	Various e.g., lending	Various e.g., Adult, COMPAS	Multiple bias types
Mehrabi et al. (2021)	Literature review	Various e.g., classification, NLP	Various	Taxonomy of biases
Hasan et al. (2022)	Causal analysis, participatory auditing	Various e.g., hiring, lending	Various	Practitioner lessons
Patel and Uddin (2022)	Literature review	Various high-stakes sectors	Various	Bias mitigation frameworks

Table 3. (*Continued*).

Author(s)	Technique(s)	Application area	Dataset(s)	Core focus
Ovalle et al. (2023)	Subgroup clustering, counterfactual analysis	Clinical predictions	MIMIC-III	Subgroup performance gaps
Liu et al. (2024)	Participatory, decision tree guidance	Lending	Adult	Practitioner operationalization
Landers and Behrend (2021)	Conceptual framework	Various e.g., hiring, policing	Various	Interdisciplinary auditing
Aggarwal et al. (2019)	Symbolic execution, local explanations	Lending	Adult, others	Individual discrimination
Aravantinos et al. (2024)	Systematic literature review	Primary school AI education	Scopus	Educational approaches with AI
Lavidas et al. (2024)	Survey analysis	Academic AI use	University students	Determinants of AI adoption
Li et al. (2024)	Regularized diagonal distance metric learning	Credit evaluation	Not specified	Feature selection and grouping effect analysis
Gao et al. (2025)	Multi-heterogeneous self-paced ensemble learning	Financial distress prediction	High-dimensional imbalanced datasets	Predictive accuracy in imbalanced data
Kou and Lu (2025)	Literature review	FinTech applications	Various	Emerging financial technologies
Raji et al. (2020)	Internal algorithmic auditing	Various commercial applications	Company-specific	Operationalizing AI ethics
Shneiderman (2020)	Conceptual framework	Various AI applications	Human-centered AI auditing	Governance strategies for trustworthy AI
DeVos et al. (2022)	User-driven auditing, think-aloud interviews, diary studies, workshops	Algorithmic bias detection	Various real-world examples (e.g., Google Image search, YouTube recommendations)	Users harmful algorithmic behaviors
Vecchione et al. (2021)	Social science audit methodologies	Algorithmic fairness and social justice	Historical audit studies and policy reviews	Evolution of algorithmic auditing and its relationship with social justice
Galdon Clavell et al. (2020)	Algorithmic audit, qualitative analysis, digital ethnography	AI-based recommendation systems	User feedback and app performance logs	Algorithmic biases in well-being recommendation systems
Bandy (2021)	Systematic literature review	Public-facing algorithms	Public	Categorizing problematic machine behaviors
Groves et al. (2024)	Qualitative interviews	Algorithmic bias audits in employment	Interview data and real-world audits	Evaluating NYC's algorithm audit law
Morales-Navarro et al. (2024)	Workshop-based peer auditing, clinical interview	Youth ML education	ML-powered apps	Youth in algorithm auditing

3.2. Probing deficiencies in current approaches

Our analysis of the current landscape of algorithm auditing reveals several key limitations and challenges. We present these findings in a hierarchical structure to clearly distinguish between different types of audits and their specific limitations:

3.2.1. Narrow technical focus

A consistent pattern observed across many of the auditing techniques covered is their overemphasis on technical assessments centered narrowly on datasets or algorithms. One manifestation of this issue is the predominant reliance on quantitative metrics and definition-based evaluations divorced from applied contexts and

experiences of stakeholder groups. For instance, in the study of Sharma and Wehrheim (2020), the verification-based testing approach focuses solely on evaluating classification models against formalized fairness properties like statistical parity or equalized odds expressed through logic statements. While rigorous, this black box perspective assessed solely through input-output queries does not account for wider socio-technical drivers, uses, and impacts shaping experiences of algorithmic harms. The procedure also does not elucidate ethical tensions or trade-offs between priorities like accuracy, fairness, and safety, which are unavoidably embedded within system design.

Similar observations hold for the bias detection toolkit proposed in the study of Bellamy et al. (2018), which concentrates largely on quantifying performance gaps and representation imbalances through an extensive set of 71 bias metrics. The metrics treat algorithms as detached artifacts assessing aspects like statistical parity, calibration, equalized odds, and consistency. However, how auditing insights connect to rectifying root deficiencies in data collection, variable selection, model assumptions, or organizational processes is left unaddressed. Taxonomy review (Mehrabi et al., 2021) also frames bias mitigation solutions for machine learning models in a predominately statistical lens, discussing pre-processing, in-processing and post-processing techniques like reweighing, adversarial debiasing, and threshold adjustments. The solutions emphasize model outputs satisfying parity constraints between groups. However, engaging affected populations to surface unintended harms, elucidating shifted assumptions from auditing feedback, or bolstering accountability through participative oversight finds little coverage.

3.2.2. Obscured value tensions

A parallel deficiency is the widespread absence of transparent deliberation or elucidation of ethical values, priorities and trade-offs inexorably bound up in algorithm design, auditing, and governance. For example, Ovalle et al. (2023) developed the SLOGAN auditing tool to cluster biased patient subgroups and detect performance gaps in clinical predictions. However, unpacking contestations on appropriate fairness definitions or distributive priorities balancing different demographic groups finds little focus beyond ensuring similarities in illness severity scores. The way contrasting stakeholder perspectives feed into characterizing and governing harms is not substantiated. Similarly, the symbolic execution approach (Aggarwal et al., 2019) to generating test cases assessing individual discrimination in lending models does not clarify the normative underpinnings of adopted fairness assumptions. The technique automatically computes counter-examples violating user-specified logic constraints. However, elucidating considerations behind fairness formalizations or interrogating shifted priorities is not incorporated.

3.2.3. Retrospective assessments

A third concern centers on the overwhelming focus on retrospective audits disconnected from the applied lifecycles of algorithm development, deployment, and updating. Beyond one-off assessments of deployed models, integrating evaluative processes spanning design, monitoring, and governance remains rare. For instance, the FairCompass (Liu et al., 2024) toolkit allows practitioners to interactively explore metrics and subgroup biases. However, the system lacks capabilities for custom

modeling or simulation that can proactively surface issues early during development phases or generate synthetic test cases difficult to sample in real deployment contexts. Mechanisms for continual bias tracking as new user groups or application environments emerge over time are also absent currently. Equally, while the severity-based subgroup evaluator SLOGAN (Ovalle et al., 2023) provides rich insights into biases encoded in patient risk predictions, translating results into prospective data collection reforms and modeling changes or participative oversight procedures is not detailed. One-time auditing rarely provides sufficient or timely feedback to shift entrenched assumptions and constrained optimization paradigms underlying algorithmic harm. Constructing sustained, adaptive assessment processes spawning multi-level learning among stakeholders thus constitutes an open design challenge.

3.2.4. Scarce community participation

A fourth limitation is the scarce involvement of impacted individuals, groups, and domain experts in guiding auditing formulations, interpretations, and responses. Algorithms interfacing with human lives cannot be adequately or ethically assessed without accounting for experiential contexts shaping the possibility and distribution of technological risks and harms. However participative, co-constructed examinations currently remain more an exception than the norm.

For example, analysis centers (Hasan et al., 2022) on practitioner perspectives and lessons in conducting commercial algorithm audits. Client needs and business constraints facing reviewers, like inadequate testing data or model opacity, undoubtedly shape assessments. However, the degree auditing and redressal procedures also integrate dialogue with communities facing decisions and donors providing data. Further investigation has given potential conflicts between user rights and vendor priorities. Even Patel and Uddin's (2022) extensive landscape review of techniques and frameworks for bias mitigation in algorithmic systems does not substantiate the current state or open challenges around participative auditing. The role of affected individuals and groups, either in surfacing experiencing issues or providing feedback on proposed interventions, remains broadly excluded from dominant computational assessments.

3.2.5. Limited corrective actions

The final cross-cutting limitation is the scant evidence of demonstrable correctives instituted in algorithms, data regimes, or governance ecosystems in response to auditing feedback. Beyond detecting issues, translating results into impactful interventions tackling root deficiencies remains rarely substantiated. For instance, Mehrabi et al. (2021) extensively document the multitude of debiasing techniques like masking sensitive attributes, adversarial training, conditional entropy optimization, and path-specific causal analysis. However, evidence on the real-world effectiveness of these technical interventions in addressing biases uncovered through deployments is currently limited. Equally, the fairness toolkit assessment (Bellamy et al., 2018) concentrates more on quantifying trade-offs between accuracy and parity metrics under different mitigation algorithms. However, examining corrective feedback loops challenging homogenizing assumptions encoded in benchmark datasets like Adult or COMPAS through continual participative auditing finds little focus. Constructing such reciprocal pathways between auditing and redressal anchored

in social realities beyond technical systems thus constitutes an open imperative if responsible AI is to progress beyond detection towards equitable impacts.

3.3. Addressing biases through actionable interventions

While algorithm audits are instrumental in diagnosing biases and ethical risks in AI systems, their effectiveness ultimately depends on whether they lead to meaningful interventions that mitigate these issues. Several approaches have been developed to translate auditing insights into actionable improvements, ensuring that AI systems evolve towards greater fairness, accountability, and transparency. These interventions typically fall into three categories: pre-processing data adjustments, in-processing model modifications, and post-processing decision refinements (Bandy, 2021; Vecchione et al., 2021).

3.3.1. Pre-processing interventions: Enhancing data quality and representation

A common source of algorithmic bias originates from imbalanced or unrepresentative training data. To address this, pre-processing interventions focus on improving data quality before it is fed into AI models. One successful example is the use of re-weighting and re-sampling techniques, where data samples from underrepresented groups are given greater statistical weight to ensure fairer model training. Similarly, in healthcare AI, audits of diagnostic algorithms have led to relabeling efforts and enriched datasets, ensuring that medical conditions affecting diverse populations are more accurately represented in training data (DeVos et al., 2022). Gao et al., 2025 in their study focus on improving predictive accuracy in imbalanced datasets.

3.3.2. In-processing interventions: Algorithmic adjustments for fairness

Beyond data-level corrections, in-processing interventions modify model architectures and training objectives to ensure more equitable outcomes. One prominent technique is adversarial debiasing, where models are trained with fairness constraints to minimize disparities across demographic groups. An example of this approach was implemented in a hiring algorithm used by a multinational corporation, which was found to disproportionately favor male candidates. After an audit revealed gender biases, developers incorporated fairness-aware loss functions that adjusted model predictions to achieve parity in selection rates across genders. This intervention resulted in a 25% reduction in gender disparities in hiring decisions without significantly compromising the model's overall performance (Slootjes, 2017).

3.3.3. Post-processing interventions: Adjusting model outputs for equity

For models already deployed in real-world applications, post-processing interventions provide a way to mitigate biases without retraining. These methods adjust model predictions or decision thresholds to ensure fairer outcomes. A successful post-processing intervention was seen in the correction of racial bias in predictive policing algorithms, where an audit revealed that certain neighborhoods were unfairly flagged for high crime risk due to historical data biases. The solution involved threshold adjustments and calibrated decision-making processes, leading to a measurable reduction in false positive crime predictions in over-policed communities (Groves et al., 2024).

To ensure that auditing interventions have a tangible impact, their effectiveness must be assessed through empirical evaluation and continuous monitoring. Successful interventions are often measured using fairness metrics such as demographic parity, equalized odds, and disparate impact ratios. For example, after an audit of an AI-powered resume screening tool identified racial biases in job candidate selection, post-audit modifications led to a 30% improvement in equal representation among selected candidates, as verified by independent evaluations. Similarly, in the financial sector, fairness-driven audits have led to regulatory compliance improvements, ensuring that automated credit-scoring systems do not disproportionately disadvantage minority applicants (Morales-Navarro et al., 2024). This trend aligns with the evolving landscape of financial technologies as outlined by Kou and Lu (2025), who provide a comprehensive literature review of emerging financial technologies and applications.

3.4. Charting promising directions for progress

Building upon our analysis of current limitations in algorithm auditing, this section outlines key strategic directions for progress, offering concrete implementation paths and addressing feasibility challenges. Effective algorithm auditing requires a multi-stakeholder approach, integrating technical solutions, regulatory oversight, and participatory engagement to ensure AI systems remain transparent, accountable, and fair. The study complements Li et al.'s (2024) approach by advocating for participatory auditing processes that involve affected communities, which can help identify potential biases that may not be apparent through technical analysis alone. Below, we propose six critical pathways for strengthening algorithm audits, focusing on their practical application and potential barriers to implementation.

3.4.1. Participatory auditing frameworks

To enhance accountability and fairness, standardized frameworks must be developed to actively involve affected communities throughout the auditing process, from initial design to continuous oversight. Implementing such frameworks ensures that algorithmic decisions reflect diverse societal perspectives rather than solely technical evaluations (Vecchione et al., 2021).

Implementation path:

Establish community advisory boards for AI projects in high-impact domains such as healthcare, finance, and criminal justice.

Develop training programs that equip community members with foundational knowledge in AI and algorithmic decision-making.

Implement "citizen auditor" programs, drawing inspiration from citizen science initiatives, to encourage direct public participation in AI oversight.

Feasibility challenges:

Ensuring representative participation across diverse demographic groups.

Balancing the need for technical expertise with inclusive participation from non-specialist communities.

Mitigating potential conflicts of interest while maintaining objectivity in auditing processes.

3.4.2. Continuous monitoring systems

AI auditing should not be a one-time process but an ongoing effort that integrates

real-time monitoring to detect emerging biases and system failures. Developing dynamic auditing mechanisms ensures AI systems evolve responsibly over time (Bandy, 2021).

Implementation path:

Embed monitoring APIs into AI models to detect and flag bias-related anomalies.

Establish industry-wide standards for continuous auditing, modeled after cybersecurity monitoring frameworks.

Develop centralized oversight dashboards for regulators, enabling real-time tracking of AI compliance across multiple sectors.

Feasibility challenges:

Managing the technical complexity of monitoring AI systems that continuously evolve.

Balancing privacy considerations with real-time surveillance of AI decision-making.

Ensuring the security of monitoring infrastructure, preventing tampering or exploitation.

3.4.3. Interdisciplinary audit teams

Algorithm auditing requires expertise beyond technical assessments, incorporating perspectives from ethics, law, and social sciences. Establishing cross-disciplinary audit teams will facilitate holistic evaluations of AI systems, ensuring fairness and accountability (Galdon Clavell et al., 2020).

Implementation path:

Develop certification programs that equip professionals from various disciplines with AI auditing expertise.

Establish guidelines for interdisciplinary audit team composition, ensuring diverse expertise aligns with the AI system's domain and impact.

Create interdisciplinary research centers dedicated to advancing algorithm auditing methodologies.

Feasibility challenges:

Addressing the shortage of professionals with expertise across AI, ethics, and policy.

Managing increased auditing costs and timelines associated with multidisciplinary evaluations.

Fostering effective collaboration across fields with distinct methodologies and priorities.

3.4.4. Regulatory frameworks and industry standards

The development of comprehensive regulatory frameworks is essential for standardizing AI auditing practices and ensuring compliance across industries. Policymakers must create enforceable guidelines that align with evolving AI capabilities while promoting innovation (European Union, 2024).

Implementation path:

Implement the EU AI Act's risk-based approach, introducing conformity assessments for high-risk AI systems.

Enforce the US Algorithmic Accountability Act, requiring impact assessments for automated decision systems.

Adopt IEEE P7003 standards, which provide guidelines for identifying and mitigating algorithmic bias.

Integrate ISO/IEC 42001 AI management system frameworks, ensuring AI governance is maintained throughout the system lifecycle.

Feasibility challenges:

Harmonizing regulations across different jurisdictions with varying legal frameworks.

Ensuring regulatory agility to keep pace with rapid AI advancements.

Balancing innovation with compliance, preventing overregulation from stifling AI development.

3.4.5. Explainable AI for auditing

AI transparency is essential for effective audits. Explainable AI (XAI) techniques can help auditors interpret and validate algorithmic decision-making, reducing opacity in high-impact systems.

Implementation path:

Increase investment in research focused on explainability methods tailored to auditing applications.

Develop minimum explainability standards for AI models used in critical domains.

Create tools that translate complex AI decisions into interpretable formats for non-technical stakeholders.

Feasibility challenges:

Managing trade-offs between model performance and interpretability.

Ensuring explanations are meaningful across diverse stakeholders, including regulators and affected communities.

Balancing intellectual property protection with the need for algorithmic transparency.

3.4.6. Ethical impact assessments

Ethical considerations should be systematically integrated into AI auditing through structured impact assessments. This approach ensures ethical concerns are identified before deployment, rather than retroactively.

Implementation path:

Develop standardized ethical impact assessment frameworks tailored to different AI applications.

Require ethical impact statements as part of AI system documentation.

Establish public repositories of ethical impact assessments, fostering transparency and best practices.

Feasibility challenges:

Addressing the subjectivity of ethical assessments, which may vary based on cultural and societal norms.

Keeping assessments aligned with rapidly evolving AI capabilities.

Managing the balance between thorough ethical evaluations and development efficiency.

3.4.7. Towards a robust algorithm auditing ecosystem

By implementing these strategic directions, AI auditing can move beyond passive assessments toward active, impactful interventions. However, overcoming feasibility challenges requires collaboration between policymakers, industry leaders, researchers, and affected communities. Establishing an effective algorithm auditing ecosystem will necessitate regulatory support, technological innovation, and participatory oversight to ensure AI systems remain transparent, fair, and accountable (European Union, 2024).

4. Discussion

The imperative for rigorous, holistic, and socially anchored algorithm auditing processes emerges strongly from this systematic review. While the coverage of techniques indicates momentum towards greater scrutiny, limitations along ethical dimensions of contextualization, participation, and correctives signal the need for auditing paradigms positioned as responsive governance instruments, not detached assessments. The persistent disconnect between specialized technical evaluations and translating findings into impactful redressals tackling root deficiencies represents a pivotal challenge requiring urgent address. Computational assessments provide a crucial starting point, as evidenced by advances like context-driven subgroup clustering (Ovalle et al., 2023), interactive metric explorations (Liu et al., 2024), and automated test generation through symbolic executions (Aggarwal et al., 2019). Equally, operationalizing audits within organizational settings appears to be gaining traction, as conveyed in Hasan et al.'s (2022) applied analysis. However, constructing sustained feedback loops between detecting biases and informing upstream reforms to data sourcing norms, feature choices, model assumptions, and business incentives remains rare currently. The integration of participatory mechanisms also lags, with communities impacted by algorithmic decisions broadly excluded.

These gaps likely stem from the relatively nascent state of algorithm auditing as a practice coupled with researchers predominately positioning assessments as technical pursuits isolated from societal contexts. The cross-disciplinary dissonance shades auditing formulations towards visible symptoms like performance disparities instead of holistic interrogations co-identifying invisible harms with affected populations. Translating computational signals into governing actions further necessitates deliberating embedded priorities, surfacing unintended consequences, and bolstering infrastructures enabling participatory oversight (Katell et al., 2020). Mainstreaming such continuous, embedded, and pluralistic auditing architectures requires transcending entrenched dichotomies between developers and auditors or science and society (Benjamin, 2019). The socio-material entanglement of algorithms, experiences, and environments instead invites framing audits as collaborative inquiries towards equitable AI systems.

Our analysis reveals the scarcity of co-developed assessment mechanisms spanning the design, deployment, and updating of models. While exceptions like the FairCompass system (Liu et al., 2024) illustrate initial attempts at interactive tooling, longitudinal participatory auditing "in the wild" remains glaringly absent. Constructing robust pathways for community representatives like domain experts,

ethicists, and user groups to provide context-specific feedback and surface experienced harms constitutes an open yet critical direction (Patterson and Hennessy, 2017). Beyond engaging end-users, frameworks integrating deliberations across currently siloed units like engineering, compliance, public relations, and leadership also need cultivation if auditing insights are to inform impactful reforms (Katell et al., 2020).

This expansive orientation necessitates moving beyond predominantly one-shot technical testing towards sustained review processes situated within organizational and sectoral environments. It also requires broadening underlying philosophical commitments from purely quantifiable metrics towards the participatory elucidation of values, assumptions, and tensions that shape the possibility of harm. Infrastructure roles, spaces, and capabilities enabling such transparency, debate, and oversight around the choices inexorably structuring technological risks represent a pivotal governance priority (Kane, 2010; Yuan et al., 2021). Rather than instituting detached auditing protocols, integrative socio-material instruments supporting the context-specific translation of computational signals into governing actions are warranted. The EU has taken a proactive approach to AI regulation and algorithm auditing, as evidenced by the EU AI Act (European Union, 2024). This comprehensive legislation aims to categorize AI systems based on risk levels and impose stringent requirements for high-risk applications. The Act emphasizes transparency, accountability, and fairness in AI systems, mandating regular audits and assessments. The US approach to AI regulation and auditing has been more decentralized, with a mix of federal guidance and state-level legislation. The studies by Raji et al. (2020) and Shneiderman (2020) highlight the focus on internal auditing practices in commercial settings and the development of human-centered AI governance strategies.

5. Conclusions

This systematic review of algorithm auditing practices has revealed significant gaps in current approaches and highlighted promising directions for progress. As AI systems increasingly influence high-stakes decisions across various domains, the need for comprehensive, ethical, and participatory auditing frameworks becomes paramount. Key findings from the study include:

(1) The limitations of current auditing practices, particularly their narrow technical focus and lack of community participation.
(2) The need for continuous monitoring and adaptive auditing processes to keep pace with evolving AI systems.
(3) The importance of interdisciplinary approaches that integrate technical, ethical, and social considerations.
(4) The critical role of regulatory frameworks and industry standards in shaping effective auditing practices.

Based on these findings, the study proposes the following concrete steps for policymakers, industry leaders, and researchers:

(1) For policymakers
    - Develop and implement comprehensive AI auditing legislation, drawing inspiration from frameworks like the EU AI Act and the US Algorithmic

Accountability Act.

- Establish a dedicated AI regulatory agency to oversee algorithm auditing and enforcement.
- Create incentives for companies to adopt participatory auditing practices and continuous monitoring systems.

(2) For industry leaders

- Integrate ethical impact assessments and continuous monitoring tools into AI development lifecycles.
- Invest in developing explainable AI techniques that facilitate transparent auditing processes.
- Collaborate with academic institutions and affected communities to create diverse, interdisciplinary audit teams.

(3) For researchers

- Focus on developing standardized, cross-disciplinary methodologies for algorithm auditing that balance technical rigor with ethical considerations.
- Investigate novel approaches to community participation in auditing processes, drawing inspiration from fields such as participatory action research.
- Conduct longitudinal studies on the effectiveness of different auditing approaches to inform best practices.

While these recommendations offer a path forward, we acknowledge several feasibility challenges that must be addressed:

- The technical complexity of auditing evolving AI systems requires ongoing investment in research and development.
- Balancing the need for transparency with intellectual property concerns and competitive advantages poses legal and economic challenges.
- Ensuring meaningful participation from diverse stakeholders while maintaining objectivity and expertise in auditing processes requires careful consideration and novel approaches.

In conclusion, mainstreaming participatory and corrective auditing processes is essential for fostering equitable AI systems. By addressing the identified gaps and implementing the proposed recommendations, we can work towards a future where AI technologies are not only powerful and efficient but also transparent, fair, and accountable to the communities they serve. The journey towards comprehensive algorithm auditing is complex and challenging, but it is a necessary step in ensuring that AI systems align with societal values and ethical principles.

Conflict of interest: The author declares no conflict of interest.

References

Aggarwal, A., Lohia, P., Nagar, S., et al. (2019). Black box fairness testing of machine learning models. In: Proceedings of the 27th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering; 26–30 August 2019; Tallinn, Estonia. pp. 625-635. https://doi.org/10.1145/3338906.3338937

Anderson, C. (2008) The end of theory: The data deluge makes the scientific method obsolete. Wired magazine, 16(7), 16-07.

Aravantinos, S., Lavidas, K., Voulgari, I., et al. (2024). Educational Approaches with AI in Primary School Settings: A

Systematic Review of the Literature Available in Scopus. Education Sciences, 14(7), 744. https://doi.org/10.3390/educsci14070744

Arnold, K., Gosling, J., Holmes, D. (2017). The Java programming language, 5th ed. Addison Wesley.

Bandy, J. (2021). Problematic machine behavior: A systematic literature review of algorithm audits. In: Proceedings of the ACM on Human-Computer Interaction, (CSCW1). pp. 1-34. https://doi.org/10.1145/3449148

Belk, R. (2014). Sharing versus pseudo-sharing in Web 2.0. The Anthropologist, 18(1), 7-23. https://doi.org/10.1080/09720073.2014.11891556

Bellamy, R. K., Dey, K., Hind, M., et al. (2018). AI fairness 360: An extensible toolkit for detecting, understanding, and mitigating unwanted algorithmic bias. arXiv, arXiv:1810.01943.

Benjamin, R. (2019). Race after technology: Abolitionist tools for the new Jim code. John Wiley & Sons.

Creswell, J. W., Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications.

DeVos, A., Dhabalia, A., Shen, H., et al. (2022). Toward User-Driven Algorithm Auditing: Investigating users' strategies for uncovering harmful algorithmic behavior. CHI Conference on Human Factors in Computing Systems, 1–19. https://doi.org/10.1145/3491102.3517441

Ehsan, H., Beebe, C., Cardella, M. E. (2017). Promoting computational thinking in children using apps. In: Proceedings of the 2017 American Society for Engineering Education (ASEE) Annual Conference & Exposition; 24-28 June 2017; Columbus, Ohio. https://doi.org/10.18260/1-2—28772

European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending certain Union legislative acts (Artificial Intelligence Act). European Union.

Feynman, R. P. (2017). The character of physical law. MIT Press.

Fioretto, F., Pontelli, E., Yeoh, W. (2018). Distributed constraint optimization problems and applications: A survey. Journal of Artificial Intelligence Research, 61, 623-698. https://doi.org/10.1613/jair.1.11203

Galdon Clavell, G., Martín Zamorano, M., Castillo, C., et al. (2020). Auditing algorithms: On lessons learned and the risks of data minimization. In: Proceedings of the 2020 ACM AI, Ethics, and Society Conference (AIES '20). pp. 265–271. https://doi.org/10.1145/3375627.3375852

Gao, F., Blunier, B., Miraoui, A. (2013). Proton exchange membrane fuel cells modelling. John Wiley & Sons.

Gao, R., Cui, S., Wang, Y., et al. (2025). Predicting financial distress in high-dimensional imbalanced datasets: a multi-heterogeneous self-paced ensemble learning framework. Financial Innovation, 11(1). https://doi.org/10.1186/s40854-024-00745-w

Groves, L., Metcalf, J., Kennedy, A. (2024). Auditing work: Exploring the New York City algorithmic bias audit regime. In: Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24). pp. 1107-1120. https://doi.org/10.1145/3630106.3658959

Hasan, M., Sundsøy, P., Bjelland, J., Pentland, A. (2022). Algorithmic bias and risk assessments: lessons from practice. Philosophy & Technology, 35(1), 99-120. https://doi.org/10.1007/s13347-021-00500-0

Kacprzyk, J., Pedrycz, W. (2015). Springer handbook of computational intelligence. Springer. https://doi.org/10.1007/978-3-662-43505-2

Kane, C. L. (2010). 'Programming the Beautiful' Informatic Color and Aesthetic Transformations in Early Computer Art. Theory, Culture & Society, 27(1), 73-93. https://doi.org/10.1177/0263276409358447

Katell, M., Young, M., Dailey, D., et al. (2020). Toward situated interventions for algorithmic equity: lessons from the field. In: Proceedings of the Conference on Fairness, Accountability, and Transparency; 27-30 January 2020; Barcelona, Spain. pp. 1048-1059. https://doi.org/10.1145/3351095.3372874

Kolstø, S. D. (2001). Scientific literacy for citizenship: Tools for dealing with the science dimension of controversial socioscientific issues. Science education, 85(3), 291-310. https://doi.org/10.1002/sce.1020

Kou, G., & Lu, Y. (2025). FinTech: a literature review of emerging financial technologies and applications. Financial Innovation, 11(1). https://doi.org/10.1186/s40854-024-00668-6

Landers, R. N., Behrend, T. S. (2021). AI for algorithmic auditing: mitigating bias and improving fairness in big data systems. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 11(3), e1369. https://doi.org/10.1002/widm.1369

Lavidas, K., Voulgari, I., Papadakis, S., et al. (2024). Determinants of Humanities and Social Sciences Students' Intentions to Use

Artificial Intelligence Applications for Academic Purposes. Information, 15(6), 314. https://doi.org/10.3390/info15060314

Leite, A. R., Enembreck, F., Barthes, J. P. A. (2014). Distributed constraint optimization problems: Review and perspectives. Expert Systems with Applications, 41(11), 5139-5157. https://doi.org/10.1016/j.eswa.2014.03.016

Li, T., Kou, G., Peng, Y., et al. (2024). Feature Selection and Grouping Effect Analysis for Credit Evaluation via Regularized Diagonal Distance Metric Learning. INFORMS Journal on Computing. https://doi.org/10.1287/ijoc.2023.0322

Liu, J., Ji, S., Northcutt, C. G., et al. (2024). FairCompass: Operationalising Fairness in Machine learning. ArXiv.

Loder, C. Something to hide: individual strategies for personal privacy practices. Proceedings of the 9th iConference, 4-7 March 2014, Berlin. 814-819. https://doi.org/10.9776/14403

Lu, Z., Kazi, R. H., Wei, L. Y., et al. (2021). Streamsketch: Exploring multi-modal interactions in creative live streams. In: Proceedings of the ACM on Human-Computer Interaction. pp. 1-26. https://doi.org/: https://doi.org/10.1145/3449132

Macal, C. M. (2016). Everything you need to know about agent-based modelling and simulation. Journal of Simulation, 10, 144-156. https://doi.org/10.1057/s41273-016-0007-5

Mehrabi, N., Morstatter, F., Saxena, N., et al. (2021). A Survey on Bias and Fairness in Machine Learning. ACM Computing Surveys, 54(6), 1–35. https://doi.org/10.1145/3457607

Moher, D., Liberati, A., Tetzlaff, J., et al. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. Annals of Internal Medicine, 151(4), 264–269. https://doi.org/10.7326/0003-4819-151-4-200908180-00135

Morales-Navarro, L., Kafai, Y. B., Konda, V., & Metaxa, D. (2024). Youth as peer auditors: Engaging teenagers with algorithm auditing of machine learning applications. In: Proceedings of the 23rd Annual ACM Interaction Design and Children Conference. pp. 560–573. https://doi.org/10.1145/3628516.3655752

Ovalle, A., Dev, S., Zhao, J., et al. (2023). Auditing algorithmic fairness in machine learning for health with severity-based LOGAN. In: International Workshop on Health Intelligence. Cham: Springer Nature Switzerland. pp. 123-136.

Patel, N., Uddin, Z. (2022). AI for algorithmic auditing: mitigating bias and improving fairness in big data systems. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 12(1), e1457. https://doi.org/10.1002/widm.1457

Patterson, D. A., Hennessy, J. L. (2017). Computer organization and design: the hardware/software interface. ELSEVER.

Raji, I. D., Smart, A., White, R. N., et al. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. In: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT '20)*. pp. 33–44. https://doi.org/10.1145/3351095.3372873.

Sandhu, R., Coyne, E. J., Feinstein, H. L., Youman, C. E. (1996). Role-based access control models. Computer, 29(2), 38-47. https://doi.org/10.1109/2.485845

Sharma, A., Wehrheim, H. (2020). Automatic fairness testing of machine learning models. In: Proceedings of the 32nd International Conference on Testing Software and Systems (ICTSS); December 2020; Naples, Italy. pp. 255-271. https://doi.org/10.1007/978-3-030-64881-7_16

Shneiderman, B. (2020). Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy. International Journal of Human–Computer Interaction, 36(6), 495–504. https://doi.org/10.1080/10447318.2020.1741118

Slootjes, J. (2017). Narratives of meaningful endurance: The role of sense of coherence in the health and employment of ethnic minority women [PhD] thesis. Vrije Universiteit Amsterdam.

Van Dyk, D. A., Meng, X. L. (2001). The art of data augmentation (with discussion). Journal of Computational and Graphical Statistics, 10(1), 1-111. https://doi.org/10.1198/106186001526

Vecchione, B., Levy, K., & Barocas, S. (2021). Algorithmic Auditing and Social Justice: Lessons from the History of Audit Studies. Equity and Access in Algorithms, Mechanisms, and Optimization, 1–9. https://doi.org/10.1145/3465416.3483294

Woit, P., (2017). Quantum theory, groups and representations. New York, NY, USA: Springer International Publishing. https://doi.org/10.1007/978-3-319-47729-5

Yuan, Y. H., Liu, C. H., & Kuang, S. S. (2021). An Innovative and Interactive Teaching Model for Cultivating Talent's Digital Literacy in Decision Making, Sustainability, and Computational Thinking. Sustainability, 13(9), 5117. https://doi.org/10.3390/su13095117