Article

Ethical Considerations in AI Applications in Smart Grid

Kadhim Hayawi , Sakib Shahriar * , A. R. Al-Ali

Posted Date: 23 December 2024

doi: 10.20944/preprints202412.1883.v1

Keywords: Ethical AI; smart grids; power systems; ethics

Article

Ethical Considerations in AI Applications in Smart Grid

Kadhim Hayawi [1], Sakib Shahriar [2,*] and A. R. Al-Ali [3]

[1] College of Interdisciplinary Studies, Zayed University, Abu Dhabi, United Arab Emirates
[2] School of Computer Science, University of Guelph, Guelph, ON, Canada
[3] Department of Computer Science and Engineering, American University of Sharjah, Sharjah, United Arab Emirates
* Correspondence: shahrias@uoguelph.ca

Abstract: The rise of artificial intelligence (AI)-driven applications due to the advancement of large language models and generative AI has sparked growing discussions about the ethics and governance of such algorithms. As the global adoption of smart grid and digital power infrastructure accelerates, the role of AI becomes increasingly significant in optimizing operations like forecasting, maintenance, and energy distribution. Despite several works in the literature relating to ethical AI in domains like medicine and agriculture, a framework and recommendation are missing from the power grid lens. In this study, we propose a framework of ethical principles, including transparency, accountability, fairness, and privacy, tailored to the smart grid context. Through this framework, we identify key methods and strategies, present case studies, and discuss challenges associated with ethical AI implementation in smart grids. Our findings highlight the importance of balancing technological innovation with moral considerations, ensuring equitable and transparent energy management. We also discuss significant challenges and outline future research directions for ethical AI in power grid systems.

Keywords: ethical AI; smart grids; power systems; ethics

1. Introduction

In May 2023, Denmark witnessed one of the most considerable cyberattacks in its history; twenty-two energy companies were targeted in a coordinated assault [1]. Hackers exploited vulnerabilities in system firewalls, which allowed them to breach critical infrastructure. The attacks forced several companies to disconnect from the primary grid and operate independently. The incident raised serious concerns about the security of smart grids, as it demonstrated the potential for cyberattacks to disrupt essential services. With the rise of AI-driven systems in smart grids ([1–3]), ethical considerations have become more critical than ever. Imagine a scenario where energy grids rely on AI decision support systems like predictive maintenance algorithms [4], demand response optimization [5], or autonomous grid management [6]. In such cases, attacks like adversarial manipulation [7] or model poisoning [8] could severely disrupt operations, leading to widespread blackouts, loss of control over energy distribution, and compromised consumer data. These risks highlight the urgent need to build AI systems for smart grids with ethics and security in mind, ensuring transparency, fairness, and resilience against cyber threats. Ethical AI can safeguard the future of our energy infrastructure.

AI revolutionized the power sector by enriching the efficiency, reliability, and overall performance of power systems [9]. The emerging predictive maintenance, load forecasting, and energy management have helped AI minimize operational costs and total hours of downtime while optimizing decision-making in the smart grid [10]. From a broad perspective, AI-driven technologies

[1] https://therecord.media/danish-energy-companies-hacked-firewall-bug

are now being used in renewable energy systems huge function changes are driving and helping transition the power sector to cleaner, more sustainable forms of energy [11,12]. On the other hand, such rapid integration also raises serious ethical issues at the front of probably lost human jobs, cybersecurity risks, and control over automated systems. There is a need to ensure that the development and deployment of AI are done in ways that give primacy to transparency and accountability parameters in order not to undermine these challenges into advantages within the power sector. Another important aspect will be to balance technological advancement requirements with ethics in order to shape a future with AI that is efficient and fair concerning energy management.

1.1. Importance of Ethical Considerations in AI

While the benefits brought forth by AI in smart grid systems are great, their deployment also raises a raft of ethical issues that need to be discussed if they are ever to be used responsibly and equitably. The major ethical issues are transparency and accountability, fairness, and privacy-because they touch on the trustworthiness and acceptance of AI applications in the power sector and their long-term viability.

Transparency in the operation of AI systems lets stakeholders understand and, therefore, trust the decisions of such systems. It implies explanation and interpretation of the AI model and its decisions, and it means access to these models and decision processes. Next is accountability, which involves defining and enforcing the responsibilities of developers, operators, and regulators of AI for the ethical use of AI. It requires fairness in AI algorithms to avoid bias, hence discriminatory practices that can further lead to unequal treatment of individuals or groups. The vast use of AI applications involves the collection and analysis of data; therefore, several privacy concerns are raised. Thus, sensitive information requires robust protection and measures for legal and regulatory compliance. Addressing these ethical considerations is not only a matter of compliance but also needed to increase public trust and acceptance of AI. Ethical AI practices can thus enhance the legitimacy and social license of AI applications in smart grid systems.

1.2. Research Purpose

This study explores the ethical considerations associated with AI applications in smart grid systems. The primary objectives are to highlight the importance of ethical considerations in AI applications for power systems, introduce fundamental ethical principles such as transparency, accountability, fairness, and privacy, and discuss specific ethical challenges faced in the power sector. Furthermore, the research proposes strategies and best practices for ensuring ethical AI deployment in smart grids.

2. Related Works

The rapid proliferation of AI has sparked significant discourse on its ethical implications. While these advancements have introduced groundbreaking capabilities, such as automating complex tasks, enhancing decision-making, and enabling more efficient workflows, they pose substantial ethical risks, including bias, discrimination, and privacy violations. Researchers have emphasized the need for transparency and explainability in AI systems to foster trust and reduce opacity, particularly in decision-making processes where ethical missteps can lead to societal harm ([13–15]). The challenges extend to managing the inherent biases embedded in AI training data, which can perpetuate stereotypes and exclusion, especially in underrepresented languages and cultures ([16], [14]). Researchers have proposed tools such as bias detection toolkits and ethical management frameworks to help developers, managers, and stakeholders systematically address these risks, with some focusing on enabling organizations to integrate ethical considerations into their operational and strategic decision-making ([13,17]). Moreover, the ethical concerns tied to integrating AI with big data, such as ensuring data privacy, informed consent, and security, underscore the need for robust ethical frameworks that guide the responsible development and deployment of AI technologies [15]. These frameworks call for collaborative efforts between technologists, managers, and policymakers to harmonize innovation with ethical accountability across diverse applications ([15,17,18]). While these general discussions on ethical AI and LLMs provide valuable insights, there is a critical need for domain-specific considerations tailored to power systems and smart grids. The unique operational contexts and the handling of sensitive consumer data introduce specific risks and scenarios that require targeted ethical frameworks and solutions.

Domain-specific AI applications introduce unique ethical considerations that differ significantly across contexts, demanding tailored approaches to address risks and challenges effectively. In healthcare, AI systems must navigate critical issues such as patient safety, algorithmic transparency, and the preservation of trust in the patient-physician relationship, highlighting the importance of an "Ethics by Design" approach during development ([19–21]). Similarly, the integration of AI in agriculture raises ethical concerns about farmers' privacy, animal welfare, and accountability for outcomes, requiring frameworks that prioritize fairness, sustainability, and reliability to mitigate unintended harm [22]. In education, while AI has the potential to personalize learning and support diverse student needs, it also risks diminishing the social aspects of education and perpetuating inequalities if not carefully managed ([23,24]). These domain-specific challenges show the need for ethical norms, standards, and proactive governance strategies that are context-sensitive. Particularly in emerging applications like LLMs, the lack of established conventions or guidelines poses ethical uncertainties, necessitating immediate efforts to develop standards that align with the requirements of each domain [25]. While there is extensive research on general ethical considerations in AI and growing attention to domain-specific ethics in areas like healthcare, agriculture, and education, the ethical implications of AI in the smart grid sector remain underexplored. Our research thus seeks to address the ethical considerations of AI from the context of unique risks and operational challenges of smart grids.

3. Ethical Principles in AI

3.1. Definition of Ethics

Ethics in AI encompasses the moral compass and guiding principles that shape how we create, implement, and utilize AI algorithms. The importance of ethics in AI lies in its ability to ensure that these technologies are designed and implemented in ways that benefit individuals and society, avoiding harm and promoting trust [26]. Ethical considerations are the bedrock in confronting the potential downfall of AI, such as bias, discrimination, and privacy breaches. Following ethical principles will help developers and other stakeholders to reduce these risks, gain public trust, and make AI applications more acceptable to the general public.

3.2. Core Principles

In the context of smart grid system applications, we consider four primary ethical principles. Smart grids involve extensive data collection and real-time decision-making [27], which require transparency to ensure stakeholders understand and trust the outcomes of AI systems. Accountability is essential for clarifying responsibilities across the AI lifecycle when errors or unintended consequences arise in critical operations like energy distribution or maintenance. Since AI systems in smart grids rely on diverse datasets, fairness becomes pivotal in mitigating biases and ensuring equitable outcomes. Finally, integrating AI introduces privacy risks due to the sensitivity of consumer energy data, necessitating robust data protection measures. Since these ethical principles address interconnected challenges, there may be overlaps or trade-offs between them. For instance, ensuring transparency in AI data or model processes could potentially compromise privacy by exposing sensitive information. Figure 1 summarizes the four principles, along with their examples.

Transparency: Transparent AI relates to making the processes, data, and algorithms used in AI systems understandable and accessible to stakeholders [28]. It ensures that the decisions made by AI systems can be explained and justified. This principle is needed for building trust by allowing users to understand the rationale behind decisions and see the "how" and "why" behind them. Transparency also facilitates the detection and correction of errors or biases within AI systems.

Accountability: AI accountability refers to the obligation of AI developers, operators, and users to take responsibility for the actions and outcomes of AI systems. It involves establishing clear roles and responsibilities for those involved in the AI lifecycle and implementing mechanisms to ensure compliance with ethical standards and legal requirements [29]. Embedding accountability measures would enable recourse for addressing harm or unintended consequences of AI applications.

Fairness: Fairness in AI deals with building AI systems that do not perpetuate or exacerbate biases and inequalities [30]. It is necessary to avoid discrimination based on race, gender, social status, or other unfair elements. To achieve fairness, we must keep working to find, measure, and reduce biases in AI systems. This process requires ongoing efforts to identify, measure, and mitigate biases in AI systems to promote equitable outcomes.

Privacy: Data privacy involves protecting personally identifiable information from unauthorized access, use, or disclosure. In this regard, AI systems should comply with data protection regulations and ethical guidelines to protect individual privacy. Privacy-preserving techniques, such as data anonymization and encryption, are often applied to minimize risks while enabling the beneficial use or utility of data in AI applications [31].

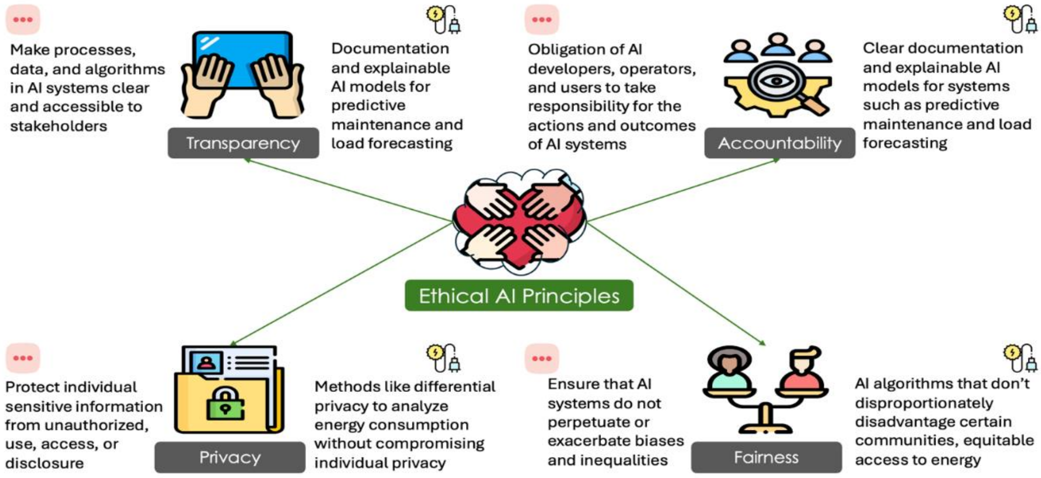
Make processes, data, and algorithms in AI systems clear and accessible to stakeholders
Transparency
Documentation and explainable AI models for predictive maintenance and load forecasting
Obligation of AI developers, operators, and users to take responsibility for the actions and outcomes of AI systems
Accountability
Clear documentation and explainable AI models for systems such as predictive maintenance and load forecasting
Ethical AI Principles
Protect individual sensitive information from unauthorized, use, access, or disclosure
Privacy
Methods like differential privacy to analyze energy consumption without compromising individual privacy
Ensure that AI systems do not perpetuate or exacerbate biases and inequalities
Fairness
AI algorithms that don't disproportionately disadvantage certain communities, equitable access to energy

Figure 1. Primary Ethical AI Principles in Smart Grids.

4. AI Transparency in Smart Grid

As society around the globe integrates into the smart grid by directly providing or consuming electrical power, they have a right to know how AI decision-making can impact them. Factors such as time of use (TOU) expenses, renewable energy delivery information, and relevant analytics should be easily accessible to the public. The following are some of the primary reasons why AI models and decision-making processes should be transparent:

1. **Trust and Acceptance:** A more transparent technology fosters trust in regulators, operators, and the public. The more transparent and comprehensible the decision made by AI systems is, the easier it is for stakeholders to accept and support the technology [32].
2. **Error Detection and Correction**: Although AI systems are generally accurate after extensive training, they may inadvertently contain errors or biases on occasion. Transparency in this can make the detection and rectification of incorrect detection much easier.
3. **Compliance with Regulations:** Generally speaking, algorithmic transparency is governed by regulations. Allowing AI systems to be transparent assists an organization in meeting the required legal and ethical requirements, which may be breached and result in fines as well as reputational damage [31].

4.1. Methods to Enhance Transparency

Several methods can be employed to enhance transparency in AI models and decision-making processes:

4.1.1. Explainable AI (XAI)

Explainable AI (XAI) is a set of techniques and methods developed to make the output of AI models understandable to humans [33]. In power systems, XAI is essential for improving the transparency and trustworthiness of AI applications like predictive maintenance, load forecasting, and fault detection. XAI can help operators, engineers, and regulators by providing clear explanations of how AI systems make their decisions. For instance, it would be feasible to trace back a decision about load delivery pricing to a given region based on factors like TOU and explain why a suburb is allocated more renewable energy. This understanding aids in the reliable and safe operation of smart grid systems.

Techniques like decision trees and rule-based models inherently offer interpretability, which makes it easier to understand the decision pathways and logic. Additionally, post-hoc explanation methods such as Shapley Additive Explanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME) can be applied to more complex black-box models [34]. These methods show how different features contribute to the model's final predictions. Understanding AI decision-making can help find potential errors, meet regulatory requirements, and develop trust in the AI systems used in smart grids.

4.1.2. Documentation

Documentation guarantees the transparency and reliability of AI models for use in smart grids. This relates to providing detailed documentation on the proposed AI system, its design, and the data sources and algorithms feeding the decision mechanisms [31]. Extensive documentation will be produced by referring to model specifications, type, and any result from the processing of their training data, performance measures, and known limitations.

In smart grids, accessible and updated documentation allows all stakeholders involved to understand and inspect the AI systems in current use. It aids in effective communication and troubleshooting and maintains the application of AI in concert with organization objectives and ethical policies. For example, in a load forecasting application, documentation would detail the data inputs (like weather conditions and economic indicators), model architecture (the type of neural network or regression model used), feature selection process (which variables were chosen and why), and the rationale behind specific forecasting decisions (how and why some predictions are made).

4.1.3. Audits

Audits consist of the independent review and assessment of AI systems to ensure that operations occur as instructed and in full compliance with ethical and regulatory standards [35]. Smart Grid systems enhance transparency, equity, and compliance in the application of AI through regular internal and external audits. Such audits can discover potential biases, errors, or avenues for improvement that make sure AI models remain aligned with organizational objectives and ethical policies. Auditing of an AI-based fault detection system would review, for example, the model for its accuracy, if it was nondiscriminatory in its decisions, and if it followed all the regulations. Such routine auditing helps maintain integrity, reliability, and ethical usage of AI systems engaged in the power section. They review AI systems on a routine basis, enabling them to refresh and enhance AI models with updated data and technology. An AI system becomes increasingly superior in performance and more trustworthy with such little steps taken over time.

4.2. Transparency Issues Case Studies

In smart grids, the implementation of AI technologies can sometimes face transparency challenges, which can hinder their effectiveness and acceptance. Several case studies illustrate these issues and the solutions that have been employed to address them.

4.2.1. Case Study 1: Predictive Maintenance in Smart Grids:

When an AI system deployed for predictive maintenance on the power grid did not provide enough insight into its decision-making process, operators suspiciously shied away from using it. Without any transparency in their decisions, operators could not understand the reasoning behind suggested maintenance actions, leading to low man-machine trust and, consequently, underutilization of capabilities. This has been overcome by making recommendations provided by the system more interpretable through explainable AI with techniques like decision trees and feature importance scoring. Further, detailed documentation was provided, and regular training sessions were conducted with the operators. In this way, their confidence in the acceptance of the AI system increases, improving utilization and effectiveness for maintaining the power grid.

4.2.2. Case Study 2: Load Forecasting and Energy Management:

A utility company using AI for load forecasting faced regulatory scrutiny due to a lack of transparency in its forecasting models. Regulators demanded a clearer understanding of how the models made predictions and the data they used. To address this, the company adopted post-hoc explanation methods like SHAP to clarify the model predictions. Additionally, they created comprehensive documentation detailing the model's data sources, algorithms, and performance metrics, which was then shared with regulators. The documentation satisfied regulatory requirements and maintained compliance while also improving the overall transparency and credibility of the load forecasting system.

4.2.3. Case Study 3: Fault Detection and Isolation:

An AI-based fault detection system in a power network produced numerous false positives, which led to unnecessary maintenance actions and operational disruptions. The lack of transparency in the AI model made it difficult to identify the root causes of these false positives. An independent audit was done to check how transparent and effective the AI system was. The audit uncovered biases in the training data, which were corrected by retraining the model with more representative data. The system was also improved to make it easier for operators to understand how it works using explainability features. These improvements reduced the incidence of false positives and also boosted the reliability of the fault detection system.

5. AI Accountability in Smart Grid Systems

Despite stringent transparency measures and proper implementation, any systems is vulnerable to failure. However, the consumers and operators of smart grids have the right to know the parties responsible for mismanagement and failure to ensure adequate investigation and subsequent improvements to the system. In this context, accountability refers to the obligation developed and held by developers, operators, and regulators to ensure that AI systems are designed, implemented, and used in accordance with ethical and responsible practices. This is a problem of accountability among proponents for the actions and decisions made by AI systems, particularly those affecting individual or group life and asylum. Establishing clear roles and responsibilities is necessary to address any AI-related difficulties and harm.

5.1. Stakeholder Responsibilities

Effective accountability in AI applications necessitates that all stakeholders diligently fulfill their responsibilities, including AI developers, operators, and regulatory bodies.

AI Developers: Developers bear a notable responsibility in the lifecycle of AI systems as they are tasked with designing systems that are ethical, fair, and transparent. Firstly, data used to train models should be representative of the populations and conditions under which the models will operate. To this end, developers must engage in meticulous data curation, selecting diverse and comprehensive datasets that minimize the risk of bias [36]. Once the data is secured, developers are to rigorously test and validate AI systems through extensive trials and simulations. This involves stress-testing AI models under various scenarios to uncover and mitigate potential risks.

Additionally, developers should provide thorough documentation detailing every aspect of the AI system, including design choices, data sources, algorithmic logic, performance metrics, and known limitations [37].

AI Operators: The responsibilities of AI operators are equally significant. They are tasked with the deployment and continuous monitoring of AI systems to ensure their correct functioning in real-world conditions. This involves overseeing the initial implementation and also engaging in regular maintenance and updates. Operators must establish robust monitoring frameworks that track the AI system's outputs and performance metrics, and detect anomalies or deviations promptly [38]. In addition, operators have the responsibility to ensure that end-users are adequately trained to interact with AI systems and comprehend their outputs. This training should cover how to interpret AI-generated insights, understand system limitations, and take appropriate actions based on AI recommendations.

Regulators: Regulators should establish standards and guidelines that govern the ethical deployment of AI systems. Establishing comprehensive regulatory frameworks that outline the legal and ethical requirements for AI applications [39] is fundamental for accountability. In addition to establishing frameworks, it is the duty of regulators to monitor that operators of AI systems comply with these standards by conducting regular inspections. In cases where entities fail to meet these standards, regulators are tasked with taking corrective action, which may include imposing penalties, mandating system modifications, or halting the deployment of non-compliant AI systems. Additionally, regulators must safeguard public interests by ensuring that AI systems do not cause harm and that there are mechanisms for redress. This involves setting up channels for reporting and addressing grievances related to AI applications, thus ensuring that individuals and communities affected by AI decisions can seek justice and remediation [40].

5.2. Methods to Ensure Accountability

Ensuring accountability in AI applications requires the implementation of robust methods and mechanisms across various dimensions, including legal frameworks, ethical guidelines, and industry standards.

Legal Frameworks: Governments and international organizations have established extensive laws for AI applications. Legal frameworks establish clear expectations for AI developers, operators, and organizations, facilitating enforcement processes [41]. Regulations may require regular audits and strict data protection to ensure openness in AI decision-making. Legal regimes sometimes include liability provisions to hold parties liable for damages caused by AI systems. Accountability allows individuals and entities affected by AI malfunctions or unexpected consequences to seek justice and get compensation.

Ethical Guidelines: Professional organizations and bodies contribute to AI accountability by developing codes of conduct that explain ethical principles and best practices for AI development and use [42]. These guidelines provide moral advice to creators and operators, guaranteeing that their decisions meet ethical standards. To reinforce compliance, many organizations establish ethical review boards responsible for evaluating AI projects before deployment, ensuring alignment with ethical principles. Because of this, ethical review boards prevent AI systems that may present ethical risks from being released and make sure AI applications consider a strong base of ethics in both their design and implementation.

Industry Standards: Setting up industry standards has also played a significant role in enforcing accountability in AI. These standards set the best ways to develop, deploy, and maintain AI; they ensure that AI systems are robust, reliable, and ethical [43]. Common elements of standards include model validation, data quality, security, and an overall framework on how to create trustworthy AI systems. In addition, a number of industries are establishing special certification programs for AI systems. This can serve to further enhance accountability because stakeholders will see that an AI system has undergone deep levels of evaluation to attain such certification and that the performance and ethical standards required thereof are significant. Certification provides, therefore, a motivation for organizations to aim for excellence in their AI practices.

5.3. Examples of Accountability Measures

Several examples illustrate how accountability measures are applied in different contexts.

Predictive Maintenance Systems: With respect to periodic audits and monitoring, the AI systems for predictive maintenance on smart grids should be checked up routinely and periodically to avoid sudden collapse or malfunction. These audits are very important in recognizing any potential problem or malfunction that may occur in the systems. Continuous oversight maintains the reliability and effectiveness of the AI models, hence protecting the integrity of the smart grid and those issues arising that could lead to big operational problems related to maintenance.

Load Forecasting Model: Accountability in load forecasting models is maintained by transparency reports. It gives an elaborate description of the model prediction, the used data, and the algorithms that have produced forecasts. By giving a clear insight into how decisions are made, transparency reports develop confidence among stakeholders, that is, regulatory compliances that could be trusted by utility companies, regulators, and consumers.

Smart Grid Management: Ethical governance is fundamental for accountability in smart grid management systems. These systems are governed by ethical guidelines that mandate regular review and oversight. Ethical governance ensures that AI systems used in smart grid management are efficient, fair, and ethical in their operation. Regular ethical reviews assess the impact of AI decisions on various stakeholders and oversee that the systems operate in a manner that upholds ethical standards and promotes public trust.

6. AI Bias and Fairness in Smart Grid Systems

Imagine a scenario where an AI algorithm trained on historical data from affluent neighborhoods disproportionately allocated power resources to these areas during peak demand and neglected lower-income regions. This bias created blackouts in marginalized communities and led to skepticism regarding AI use in smart grids. Bias in AI relates to systematic errors that result in unfair outcomes like favoritism towards certain groups or discrimination against others. These biases can stem from various sources, including biased training data, algorithmic design flaws, and unintentional human biases embedded in the development process [44]. Fairness in AI algorithms aims to prevent such biases, ensuring that AI systems treat all individuals and groups equitably and without any form of discrimination [45]. In the context of smart grids, fairness is critical as AI algorithms are increasingly used to make decisions that affect resource distribution, access to services, and overall system efficiency.

6.1. Identifying and Mitigating Biases in AI Applications for Smart Grids

To address biases in AI applications for smart grid systems, it is essential first to identify their sources and manifestations. Biases can be detected through various methods, such as examining the training data for representativeness, analyzing algorithmic outputs for patterns of discrimination, and soliciting feedback from affected stakeholders [46]. Once biases are identified, several strategies can be employed to mitigate them:

1. **Data Preprocessing:** This ensures that the training data is balanced and representative of all relevant groups. It involves techniques such as data augmentation, re-sampling, and the removal of biased data points [47].
2. **Algorithmic Adjustments:** Modifying algorithms can lead to the correction of identified biases. This includes adjusting the weighting of certain features, incorporating fairness constraints into the optimization process, and using bias mitigation algorithms [48].
3. **Post-processing Corrections:** This involves applying adjustments to the outputs of AI models to ensure fairer outcomes. Techniques such as re-ranking or recalibrating predictions to align with fairness criteria [49] can improve AI outputs.

6.2. Strategies for Ensuring Equitable AI Outcomes

This calls for addressing the issue from multiple perspectives in order to ensure that AI applications in smart grid systems result in equitable outcomes. Some of the important strategies include:

1. **Stakeholder Involvement:** Various sets of stakeholders must be involved in the design and deployment of AI systems so that all viewpoints and requirements of varied groups can be considered in order to provide fair and inclusive AI solutions [50].
2. **Auditing Fairness**: Regular fairness audits of the AI systems contribute to studying its impact on different groups and locating probable biases. These audits must be independent in order to give credibility and objectivity to the issue at hand [37].
3. **Transparency and Accountability:** It is important to have in place processes and mechanisms that ensure transparency and hold developers and operators responsible for fairness in AI systems. This includes clear documentation, explainability of decisions, as well as avenues for redress in case of perceived unfair outcomes.[51].
4. **Regulatory Compliance:** Conformity with legal and ethical standards that call for fairness in AI applications, among them anti-discrimination laws, data protection regulations, and sectorial regulations is very essential [52].

6.3. Case Studies on Fairness Issues

6.3.1. Case Study 1: Smart Grids-Load Distribution:

An AI system that was being used for the purpose of distributing the load in a smart grid was found to distribute undue share of power resources to rich neighborhoods, yielding frequent shortages in underserved communities. The problem lay in biases in the training data that had an overrepresentation of high-income areas. This skewed resource allocation decisions made by the AI model. Such biases were identified through a full-scale fairness audit. Later, they re-trained the AI model on a balanced dataset representative of all neighborhoods. Besides that, they added some fairness constraints in the algorithm to provide equal power to each and every community. Thus, they were able to rectify this disparity and worked toward equal access to energy resources for all communities.

6.3.2. Case Study 2: Predictive Maintenance and Worker Safety:

A predictive maintenance AI system was installed in a power plant and showed bias, favoring the shutdown of older pieces of equipment, which increased the frequency and amount of work in certain sections. This created grave areas of safety concern where the biased AI system would impact specific groups of people disproportionately. Subsequently, an independent review has shown this bias was enshrined in the algorithm. A revised algorithm was developed to consider those same factors non-discriminatorily, namely the age and condition of the equipment. The developers filtered biases and incomplete predictions of the need for maintenance. Additionally, training was given to operators to enhance their knowledge about the revised system, ensuring unbiased and timely maintenance. This has resulted in improving operational safety and distributing workloads more evenly across the workforce.

6.3.3. Case Study 3: Renewable Energy Allocation:

An AI model, built to perform the functions of renewable energy credit allocation, happened to favor large-scale commercial organizations rather than small businesses and residential users of such energies. This propensity resulted in an imbalance in the benefit distribution of renewable energy. Thus, smaller-scale users should benefit equally from such allocations. As a result, this required to be addressed through stakeholder engagements with representatives from small enterprises, residential user groups, and other affected parties. These discussions provided significant feedback that could be used to make modifications to the AI model while taking into account the demands and capacities of smaller users. The redesigned methodology distributes renewable energy credits in a more equitable manner while ensuring access to renewable energy supplies. Thus, it assisted many more users in their effort to be sustainable.

7. Privacy Concerns in Smart Grids

Integrating AI into smart grid systems brings about privacy implications, particularly concerning the collection, usage, and sharing of data. Security and privacy are inherently interconnected, as vulnerabilities in AI systems can expose sensitive data to malicious cyber attacks. Cyber threats, such as unauthorized access, data breaches, and adversarial attacks on AI models, compromise the integrity of power grid operations and amplify privacy risks for consumers [53]. Thus, a dual focus on robust security mechanisms and privacy-preserving measures is needed to ensure that personal information remains protected while maintaining the reliability of power systems.

AI-driven smart grids rely heavily on vast amounts of data to optimize operations, predict maintenance needs, manage loads, and integrate renewable energy sources. This data often includes sensitive information about consumers' energy usage patterns, personal habits, and even real-time locations [54]. The collection of such granular data raises concerns about the potential for unauthorized access, misuse, and surveillance. Additionally, the usage of this data for training AI models and making operational decisions necessitates robust safeguards to ensure that personal information is protected and that data privacy is maintained. The sharing of data among various stakeholders, including third-party vendors, further complicates the privacy landscape, which requires stringent controls and clear guidelines to prevent breaches and unauthorized dissemination of sensitive information [31].

7.1. Best Practices

To address privacy concerns in AI-driven power systems, adopting best practices for data protection and privacy preservation is crucial. Key practices include:

1. **Data Minimization:** Collect only the data that is strictly necessary for the AI system to function effectively. This reduces the risk of sensitive information being exposed or misused [55].
2. **Anonymization and Pseudonymization:** Transform personal data into anonymous or pseudonymous forms, ensuring that individuals cannot be easily identified from the data sets used in AI models [56].
3. **Encryption:** Implement robust encryption techniques for data at rest and in transit to protect it from unauthorized access and breaches [57].
4. **Access Controls:** Establish strict access controls to limit who can view or manipulate sensitive data. This includes role-based access and regular audits to ensure compliance [58].
5. **Transparency:** Maintain transparency with consumers about what data is being collected, how it is used, and with whom it is shared. Providing clear privacy policies and obtaining informed consent are essential components of this practice [59].
6. **Data Governance:** Develop and enforce comprehensive data governance policies that outline procedures for data handling, storage, and sharing [60]. This ensures consistency and compliance with privacy standards across the organization.

7.2. Legal and Regulatory Frameworks Governing Data Privacy

Legal and regulatory frameworks play a pivotal role in governing data privacy in AI-driven smart grid systems. These frameworks set the standards for how personal data should be handled and protected by making sure that privacy rights are upheld. Notable regulatory instruments include:

General Data Protection Regulation (GDPR): Applicable in the European Union, GDPR imposes strict requirements on data processing activities, including obtaining explicit consent, ensuring data portability, and implementing the right to be forgotten [61].

1. **California Consumer Privacy Act (CCPA):** In the United States, CCPA grants consumers rights over their personal data, including the right to access, delete, and opt out of the sale of their data [62].

2. **Smart Grid Privacy Policies:** Various jurisdictions have specific policies aimed at protecting consumer privacy within smart grid systems [63]. These policies mandate secure data handling practices and consumer rights protection.
3. **Industry Standards:** Organizations must also comply with industry-specific standards and guidelines, such as those from the National Institute of Standards and Technology (NIST) [64] or the International Organization for Standardization (ISO) [65], which provide frameworks for data protection and privacy.

7.3. Privacy-Preserving Methods

Several privacy-preserving techniques can be employed in AI applications within smart grids to ensure data privacy while enabling the benefits of AI:

1. **Differential Privacy:** This technique introduces noise into data sets to prevent the identification of individuals, ensuring that AI models can learn from the data without compromising privacy [66].
2. **Federated Learning:** Instead of centralizing data, federated learning allows AI models to be trained across decentralized devices or servers holding local data samples [67]. This approach keeps personal data on local devices, reducing privacy risks.
3. **Homomorphic Encryption:** This advanced encryption method allows computations to be performed on encrypted data without decrypting it first [68]. This ensures that sensitive data remains protected even during processing.
4. **Secure Multi-party Computation (SMPC):** SMPC enables multiple parties to jointly compute a function over their inputs while keeping those inputs private [69]. This is particularly useful for collaborative AI projects involving multiple stakeholders.
5. **Data Masking:** Involves obfuscating specific data within a data set to protect sensitive information [70]. This technique is useful for sharing data with third parties while preserving privacy.

8. Ethical Challenges in AI Integration

The role of AI in the smart grid systems is a balancing act that is both about the encouragement of innovation and following ethical considerations. The development and deployment of AI systems must be done responsibly, through thorough testing and validation, including adherence to ethical guidelines [71]. Indeed, it would be this kind of balance that would help prevent misuse and negative, unintended consequences that undermine the reliability and trustworthiness of power systems.

8.1. Ethical Dilemmas and Societal Implications

Ethical dilemmas arise about the deployment and decision-making of AI in instances when the goals of efficiency and optimization conflict with ethical values. [15]. For example, even though an AI system can be trained to distribute energy efficiently according to programmed aims, it may end up settling on economic efficiency and discrimination between areas, which translates to inequitable access to power. Such AI-driven decisions might result in existing biases or new forms of discrimination unless checked. These dilemmas outline the need for a robust ethical framework in AI development and implementation to ensure that decisions are made transparently and inclusively, with AI benefits shared equitably.

AI-powered electricity systems will have profound impacts on society, starting from early energy access to job markets. While AI has the potential to revolutionize the sector through increased efficiency and resilience, it also introduces risks linked to job displacement due to automation and may further exacerbate various social inequalities [72]. The use of AI in smart grid systems will bring large social changes, and thus great care has to be taken to see the greater implications of its impacts. This means acquiring an understanding of how AI decisions affect so many kinds of communities, especially vulnerable populations, and mitigating those impacts by considering plans and policies that truly include them.

The involvement of several types of stakeholders makes sure that a wide array of perspectives and insights is considered; this is a way of ensuring that AI solutions are more inclusive and equitable. Independent bodies should perform ethical reviews of the suitability of AI applications to ensure ethical alignment prior to their deployment [73]. Therefore, there is a need for continuous monitoring of AI systems, whereby any issue arising thereof is identified and rectified during maintenance and processing while observing ethical principles during the entire life cycle of the AI application.

8.2. Recommendations for AI Applications in Smart Grid Systems

According to energy reports, the global smart grid market was estimated to be almost 50 billion USD in 2022 and the smart electricity meters are forecasted to increase by 45% between 2021 and 2027[2]. No doubt with the growth in digital infrastructure and smart grid adoption, AI systems will be deployed increasingly for various applications and thus it is necessary to prioritize ethical considerations. Future developments in AI should focus on making AI models more transparent and easier to understand to build trust among stakeholders. Investing in XAI techniques and ensuring detailed documentation and clear communication about AI processes can help achieve this goal. In addition, a culture of continuous learning and adaptation within organizations will help keep up with evolving AI technologies and new ethical challenges. Training programs and workshops for developers, operators, and policymakers can enhance their understanding of ethical AI practices and encourage the adoption of strong transparency measures [74].

Smart grid systems should integrate better mechanisms for oversight and governance to enhance accountability. This can include setting up independent ethical review boards that evaluate AI projects before and after deployment to ensure they meet ethical standards. Regular audits and evaluations should be established to monitor AI systems' performance and adherence to ethical guidelines [35]. Also, adopting industry-wide standards and certifications for ethical AI can provide a framework for consistent and accountable AI practices across the sector [75]. To develop these standards for the latest advancements, there must be collaborations between the industry, regulatory bodies, and academia.

A proactive and comprehensive approach is needed to obtain fairness and privacy in AI applications. Organizations should implement data governance policies that focus on data minimization, anonymization, and secure data handling practices to protect user privacy [76]. Further, the participation of diverse groups in the design and implementation of AI systems can ensure that the perspectives and needs of different communities are considered, which would lead to more inclusive and fair AI solutions. Last but not least, privacy-by-design principles should be integrated into every aspect of AI development in the smart grid [77]. This approach ensures that privacy is embedded at the core of the AI system rather than being an afterthought.

9. Conclusion

As smart grid adoption continues to grow along with the digital infrastructure, AI applications in smart grid operations will also increase to improve efficiency, predictive capabilities, and optimized resource management. However, for global safety and alignment, it is necessary to ensure that ethical standards are first defined and upheld. Throughout this discourse, we proposed a framework grounded in transparency, accountability, fairness, and privacy tailored to the unique challenges of power grid systems. Through practical recommendations and real-world case studies, we illustrated how these principles can guide the ethical deployment of AI in the power sector. This framework is designed to support key stakeholders like power operators, policymakers, and technology developers in navigating the complex ethical dilemmas associated with AI-driven smart grids. Future work should focus on refining and operationalizing this framework through empirical testing and developing domain-specific metrics to evaluate ethical AI systems in smart grids.

[2] https://www.statista.com/topics/9349/smart-grids/#topicOverview

Additionally, addressing emerging challenges, such as the moral implications of integrating generative AI technologies into smart grids, will be crucial.

Contributions: Conceptualization, Kadhim Hayawi and Sakib Shahriar; Methodology, Kadhim Hayawi, Sakib Shahriar and A. R. Al-Ali; Supervision, Kadhim Hayawi and A. R. Al-Ali; Writing – original draft, Sakib Shahriar; Writing – review & editing, Kadhim Hayawi.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Z. Shi et al., “Artificial intelligence techniques for stability analysis and control in smart grids: Methodologies, applications, challenges and future directions,” *Appl. Energy*, vol. 278, p. 115733, Nov. 2020, doi: 10.1016/j.apenergy.2020.115733.
2. T. Mazhar et al., “Analysis of Challenges and Solutions of IoT in Smart Grids Using AI and Machine Learning Techniques: A Review,” *Electronics*, vol. 12, no. 1, Art. no. 1, Jan. 2023, doi: 10.3390/electronics12010242.
3. J. Ramesh, S. Shahriar, A. R. Al-Ali, A. Osman, and M. F. Shaaban, “Machine Learning Approach for Smart Distribution Transformers Load Monitoring and Management System,” *Energies*, vol. 15, no. 21, Art. no. 21, Jan. 2022, doi: 10.3390/en15217981.
4. M. A. Mahmoud, N. R. Md Nasir, M. Gurunathan, P. Raj, and S. A. Mostafa, “The Current State of the Art in Research on Predictive Maintenance in Smart Grid Distribution Network: Fault’s Types, Causes, and Prediction Methods—A Systematic Review,” *Energies*, vol. 14, no. 16, Art. no. 16, Jan. 2021, doi: 10.3390/en14165078.
5. U. Assad et al., “Smart Grid, Demand Response and Optimization: A Critical Review of Computational Methods,” *Energies*, vol. 15, no. 6, Art. no. 6, Jan. 2022, doi: 10.3390/en15062003.
6. M. Selim, R. Zhou, W. Feng, and P. Quinsey, “Estimating Energy Forecasting Uncertainty for Reliable AI Autonomous Smart Grid Design,” *Energies*, vol. 14, no. 1, Art. no. 1, Jan. 2021, doi: 10.3390/en14010247.
7. A. Chakraborty, M. Alam, V. Dey, A. Chattopadhyay, and D. Mukhopadhyay, “A survey on adversarial attacks and defences,” *CAAI Trans. Intell. Technol.*, vol. 6, no. 1, pp. 25–45, 2021, doi: 10.1049/cit2.12028.
8. T. T. Nguyen et al., “Manipulating Recommender Systems: A Survey of Poisoning Attacks and Countermeasures,” *ACM Comput Surv*, Jul. 2024, doi: 10.1145/3677328.
9. V. Franki, D. Majnarić, and A. Višković, “A Comprehensive Review of Artificial Intelligence (AI) Companies in the Power Sector,” *Energies*, vol. 16, no. 3, Art. no. 3, Jan. 2023, doi: 10.3390/en16031077.
10. Y. S. Afridi, K. Ahmad, and L. Hassan, “Artificial intelligence based prognostic maintenance of renewable energy systems: A review of techniques, challenges, and future research directions,” *Int. J. Energy Res.*, vol. 46, no. 15, pp. 21619–21642, 2022, doi: https://doi.org/10.1002/er.7100.
11. S. Shahriar, A. R. Al-Ali, A. H. Osman, S. Dhou, and M. Nijim, “Machine Learning Approaches for EV Charging Behavior: A Review,” *IEEE Access*, vol. 8, pp. 168980–168993, 2020, doi: 10.1109/ACCESS.2020.3023388.
12. K. Hayawi, S. Shahriar, and H. Hacid, “Climate Data Imputation and Quality Improvement Using Satellite Data,” *J. Data Sci. Intell. Syst.*, Jun. 2024, doi: 10.47852/bonviewJDSIS42022857.
13. M. Bahrami, R. Sonoda, and R. Srinivasan, “LLM Diagnostic Toolkit: Evaluating LLMs for Ethical Issues,” in *2024 International Joint Conference on Neural Networks (IJCNN)*, Jun. 2024, pp. 1–8. doi: 10.1109/IJCNN60899.2024.10650995.
14. B. K. Konidena, J. N. A. Malaiyappan, and A. Tadimarri, “Ethical Considerations in the Development and Deployment of AI Systems,” *Eur. J. Technol.*, vol. 8, no. 2, Art. no. 2, Mar. 2024, doi: 10.47672/ejt.1890.
15. A. Nassar and M. Kamal, “Ethical Dilemmas in AI-Powered Decision-Making: A Deep Dive into Big Data-Driven Ethical Considerations,” *Int. J. Responsible Artif. Intell.*, vol. 11, no. 8, Art. no. 8, Aug. 2021.

16. C. B. Head, P. Jasper, M. McConnachie, L. Raftree, and G. Higdon, "Large language model applications for evaluation: Opportunities and ethical implications," *New Dir. Eval.*, vol. 2023, no. 178–179, pp. 33–46, 2023, doi: 10.1002/ev.20556.

17. A. B. Brendel, M. Mirbabaie, T.-B. Lembcke, and L. Hofeditz, "Ethical Management of Artificial Intelligence," *Sustainability*, vol. 13, no. 4, Art. no. 4, Jan. 2021, doi: 10.3390/su13041974.

18. N. M. Safdar, J. D. Banja, and C. C. Meltzer, "Ethical considerations in artificial intelligence," *Eur. J. Radiol.*, vol. 122, p. 108768, Jan. 2020, doi: 10.1016/j.ejrad.2019.108768.

19. A. Čartolovni, A. Tomičić, and E. Lazić Mosler, "Ethical, legal, and social considerations of AI-based medical decision-support tools: A scoping review," *Int. J. Med. Inf.*, vol. 161, p. 104738, May 2022, doi: 10.1016/j.ijmedinf.2022.104738.

20. C. Wang, S. Liu, H. Yang, J. Guo, Y. Wu, and J. Liu, "Ethical Considerations of Using ChatGPT in Health Care," *J. Med. Internet Res.*, vol. 25, no. 1, p. e48009, Aug. 2023, doi: 10.2196/48009.

21. T. Dave, S. A. Athaluri, and S. Singh, "ChatGPT in medicine: an overview of its applications, advantages, limitations, future prospects, and ethical considerations," *Front. Artif. Intell.*, vol. 6, May 2023, doi: 10.3389/frai.2023.1169595.

22. R. Dara, S. M. Hazrati Fard, and J. Kaur, "Recommendations for ethical and responsible use of artificial intelligence in digital agriculture," *Front. Artif. Intell.*, vol. 5, Jul. 2022, doi: 10.3389/frai.2022.884192.

23. M. J. Reiss, "The Use of Al in Education: Practicalities and Ethical Considerations," *Lond. Rev. Educ.*, vol. 19, no. 1, 2021, Accessed: Dec. 15, 2024. [Online]. Available: https://eric.ed.gov/?id=EJ1297682

24. D. Schiff, "Education for AI, not AI for Education: The Role of Education and Ethics in National AI Policy Strategies," *Int. J. Artif. Intell. Educ.*, vol. 32, no. 3, pp. 527–563, Sep. 2022, doi: 10.1007/s40593-021-00270-2.

25. R. Watkins, "Guidance for researchers and peer-reviewers on the ethical use of Large Language Models (LLMs) in scientific research workflows," *AI Ethics*, vol. 4, no. 4, pp. 969–974, Nov. 2024, doi: 10.1007/s43681-023-00294-5.

26. L. Floridi et al., "AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations," *Minds Mach.*, vol. 28, no. 4, pp. 689–707, Dec. 2018, doi: 10.1007/s11023-018-9482-5.

27. A. Zainab, A. Ghrayeb, D. Syed, H. Abu-Rub, S. S. Refaat, and O. Bouhali, "Big Data Management in Smart Grids: Technologies and Challenges," *IEEE Access*, vol. 9, pp. 73046–73059, 2021, doi: 10.1109/ACCESS.2021.3080433.

28. U. Ehsan, Q. V. Liao, M. Muller, M. O. Riedl, and J. D. Weisz, "Expanding Explainability: Towards Social Transparency in AI systems," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, in CHI '21. New York, NY, USA: Association for Computing Machinery, May 2021, pp. 1–19. doi: 10.1145/3411764.3445188.

29. C. Novelli, M. Taddeo, and L. Floridi, "Accountability in artificial intelligence: what it is and how it works," *AI Soc.*, Feb. 2023, doi: 10.1007/s00146-023-01635-y.

30. L. Devillers, F. Fogelman-Soulié, and R. Baeza-Yates, "AI & Human Values," in *Reflections on Artificial Intelligence for Humanity*, B. Braunschweig and M. Ghallab, Eds., Cham: Springer International Publishing, 2021, pp. 76–89. doi: 10.1007/978-3-030-69128-8_6.

31. S. Shahriar, S. Allana, S. M. Hazratifard, and R. Dara, "A survey of privacy risks and mitigation strategies in the artificial intelligence life cycle," *IEEE Access Pract. Innov. Open Solut.*, vol. 11, pp. 61829–61854, 2023, doi: 10.1109/ACCESS.2023.3287195.

32. H. Felzmann, E. F. Villaronga, C. Lutz, and A. Tamò-Larrieux, "Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns," *Big Data Soc.*, vol. 6, no. 1, p. 2053951719860542, Jan. 2019, doi: 10.1177/2053951719860542.

33. A. Barredo Arrieta et al., "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," *Inf. Fusion*, vol. 58, pp. 82–115, Jun. 2020, doi: 10.1016/j.inffus.2019.12.012.

34. A. Gramegna and P. Giudici, "SHAP and LIME: an evaluation of discriminative power in credit risk," *Front. Artif. Intell.*, vol. 4, p. 752558, 2021.

35. G. Falco et al., "Governing AI safety through independent audits," *Nat. Mach. Intell.*, vol. 3, no. 7, pp. 566–571, Jul. 2021, doi: 10.1038/s42256-021-00370-7.

36. A. S. Albahri et al., "A systematic review of trustworthy and explainable artificial intelligence in healthcare: Assessment of quality, bias risk, and data fusion," *Inf. Fusion*, vol. 96, pp. 156–191, Aug. 2023, doi: 10.1016/j.inffus.2023.03.008.

37. I. D. Raji et al., "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*, in FAT* '20. New York, NY, USA: Association for Computing Machinery, Jan. 2020, pp. 33–44. doi: 10.1145/3351095.3372873.

38. F. Bayram, B. S. Ahmed, and A. Kassler, "From concept drift to model degradation: An overview on performance-aware drift detectors," *Knowl.-Based Syst.*, vol. 245, p. 108632, Jun. 2022, doi: 10.1016/j.knosys.2022.108632.

39. A. Walz and K. Firth-Butterfield, "Implementing Ethics into Artificial Intelligence: A Contribution, from a Legal Perspective, to the Development of an AI Governance Regime," *Duke Law Technol. Rev.*, vol. 18, p. 176, 2020 2019.

40. A. Lior, "Insuring AI: The Role of Insurance in Artificial Intelligence Regulation," *Harv. J. Law Technol. Harv. JOLT*, vol. 35, p. 467, 2022 2021.

41. N. A. Smuha, "From a 'race to AI' to a 'race to AI regulation': regulatory competition for artificial intelligence," *Law Innov. Technol.*, vol. 13, no. 1, pp. 57–84, Jan. 2021, doi: 10.1080/17579961.2021.1898300.

42. P. Boddington, *Towards a Code of Ethics for Artificial Intelligence*. in Artificial Intelligence: Foundations, Theory, and Algorithms. Cham: Springer International Publishing, 2017. doi: 10.1007/978-3-319-60648-4.

43. P. Cihon, "Standards for AI governance: international standards to enable global coordination in AI research & development," 2019.

44. R. Schwartz et al., *Towards a standard for identifying and managing bias in artificial intelligence*, vol. 3. US Department of Commerce, National Institute of Standards and Technology, 2022.

45. E. Ferrara, "Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies," *Sci*, vol. 6, no. 1, Art. no. 1, Mar. 2024, doi: 10.3390/sci6010003.

46. N. Shahbazi, Y. Lin, A. Asudeh, and H. V. Jagadish, "Representation Bias in Data: A Survey on Identification and Resolution Techniques," *ACM Comput Surv*, vol. 55, no. 13s, p. 293:1-293:39, Jul. 2023, doi: 10.1145/3588433.

47. Y. Li and N. Vasconcelos, "REPAIR: Removing Representation Bias by Dataset Resampling," presented at the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 9572–9581. Accessed: Jul. 03, 2024. [Online]. Available: https://openaccess.thecvf.com/content_CVPR_2019/html/Li_REPAIR_Removing_Representation_Bias_by_Dataset_Resampling_CVPR_2019_paper.html

48. Z. Wang et al., "Towards Fairness in Visual Recognition: Effective Strategies for Bias Mitigation," presented at the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 8919–8928. Accessed: Jul. 03, 2024. [Online]. Available: https://openaccess.thecvf.com/content_CVPR_2020/html/Wang_Towards_Fairness_in_Visual_Recognition_Effective_Strategies_for_Bias_Mitigation_CVPR_2020_paper.html

49. A. Ashokan and C. Haas, "Fairness metrics and bias mitigation strategies for rating predictions," *Inf. Process. Manag.*, vol. 58, no. 5, p. 102646, Sep. 2021, doi: 10.1016/j.ipm.2021.102646.

50. M. Madaio, L. Egede, H. Subramonyam, J. Wortman Vaughan, and H. Wallach, "Assessing the Fairness of AI Systems: AI Practitioners' Processes, Challenges, and Needs for Support," *Proc ACM Hum-Comput Interact*, vol. 6, no. CSCW1, p. 52:1-52:26, Apr. 2022, doi: 10.1145/3512899.

51. J. A. Kroll, "Outlining Traceability: A Principle for Operationalizing Accountability in Computing Systems," in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, in FAccT '21. New York, NY, USA: Association for Computing Machinery, Mar. 2021, pp. 758–771. doi: 10.1145/3442188.3445937.

52. D. Kumar and N. Suthar, "Ethical and legal challenges of AI in marketing: an exploration of solutions," *J. Inf. Commun. Ethics Soc.*, vol. 22, no. 1, pp. 124–144, Jan. 2024, doi: 10.1108/JICES-05-2023-0068.

53. Z. Zhang et al., “Vulnerability of Machine Learning Approaches Applied in IoT-Based Smart Grid: A Review,” *IEEE Internet Things J.*, vol. 11, no. 11, pp. 18951–18975, Jun. 2024, doi: 10.1109/JIOT.2024.3349381.
54. T. Jakobi, S. Patil, D. Randall, G. Stevens, and V. Wulf, “It Is About What They Could Do with the Data: A User Perspective on Privacy in Smart Metering,” *ACM Trans Comput-Hum Interact*, vol. 26, no. 1, p. 2:1-2:44, Jan. 2019, doi: 10.1145/3281444.
55. R. Mukta, H. Paik, Q. Lu, and S. S. Kanhere, “A survey of data minimisation techniques in blockchain-based healthcare,” *Comput. Netw.*, vol. 205, p. 108766, Mar. 2022, doi: 10.1016/j.comnet.2022.108766.
56. Z. Zuo, M. Watson, D. Budgen, R. Hall, C. Kennelly, and N. A. Moubayed, “Data Anonymization for Pervasive Health Care: Systematic Literature Mapping Study,” *JMIR Med. Inform.*, vol. 9, no. 10, p. e29871, Oct. 2021, doi: 10.2196/29871.
57. S. Shukla, J. P. George, K. Tiwari, and J. V. Kureethara, “Data Security,” in *Data Ethics and Challenges*, S. Shukla, J. P. George, K. Tiwari, and J. V. Kureethara, Eds., Singapore: Springer, 2022, pp. 41–59. doi: 10.1007/978-981-19-0752-4_3.
58. J. G. Cederquist, R. Corin, M. A. C. Dekker, S. Etalle, J. I. den Hartog, and G. Lenzini, “Audit-based compliance control,” *Int. J. Inf. Secur.*, vol. 6, no. 2, pp. 133–151, Mar. 2007, doi: 10.1007/s10207-007-0017-y.
59. A. Rossi and G. Lenzini, “Transparency by design in data-informed research: A collection of information design patterns,” *Comput. Law Secur. Rev.*, vol. 37, p. 105402, Jul. 2020, doi: 10.1016/j.clsr.2020.105402.
60. J. Kuzio, M. Ahmadi, K.-C. Kim, M. R. Migaud, Y.-F. Wang, and J. Bullock, “Building better global data governance,” *Data Policy*, vol. 4, p. e25, Jan. 2022, doi: 10.1017/dap.2022.17.
61. “Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).” 2016. [Online]. Available: https://eur-lex.europa.eu/eli/reg/2016/679/oj
62. “California consumer privacy act of 2018.” 2018. [Online]. Available: https://oag.ca.gov/privacy/ccpa
63. M. A. Brown, S. Zhou, and M. Ahmadi, “Smart grid governance: An international review of evolving policy issues and innovations,” *WIREs Energy Environ.*, vol. 7, no. 5, p. e290, 2018, doi: 10.1002/wene.290.
64. National Institute of Standards and Technology, “National institute of standards and technology.” [Online]. Available: https://www.nist.gov/
65. International Organization for Standardization, “International organization for standardization.” [Online]. Available: https://www.iso.org/
66. C. Dwork, “Calibrating noise to sensitivity in private data analysis,” in *Proceedings of the third conference on theory of cryptography (TCC)*, Berlin, Heidelberg: Springer-Verlag, 2006, pp. 265–284. [Online]. Available: https://doi.org/10.1007/11681878_14
67. H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, “Communication-efficient learning of deep networks from decentralized data,” *Proc. 20th Int. Conf. Artif. Intell. Stat. AISTATS*, pp. 1273–1282, 2017.
68. C. Gentry, “Fully homomorphic encryption using ideal lattices,” phd, Stanford University, 2009. [Online]. Available: https://crypto.stanford.edu/craig/craig-thesis.pdf
69. A. C.-C. Yao, “Protocols for secure computations,” *Proc. 23rd Annu. Symp. Found. Comput. Sci. FOCS*, pp. 160–164, 1982.
70. S. Wohlgemuth, I. Echizen, and N. Sonehara, “Privacy protection by context-sensitive data masking,” *Proc. 2007 ACM Workshop Priv. Electron. Soc. WPES*, pp. 105–108, 2007.
71. N. Díaz-Rodríguez, J. Del Ser, M. Coeckelbergh, M. López de Prado, E. Herrera-Viedma, and F. Herrera, “Connecting the dots in trustworthy Artificial Intelligence: From AI principles, ethics, and key requirements to responsible AI systems and regulation,” *Inf. Fusion*, vol. 99, p. 101896, Nov. 2023, doi: 10.1016/j.inffus.2023.101896.
72. N. R. Mannuru et al., “Artificial intelligence in developing countries: The impact of generative artificial intelligence (AI) technologies for development,” *Inf. Dev.*, p. 02666669231200628, Sep. 2023, doi: 10.1177/02666669231200628.

73. C. Burr and D. Leslie, “Ethical assurance: a practical approach to the responsible design, development, and deployment of data-driven technologies,” *AI Ethics*, vol. 3, no. 1, pp. 73–98, Feb. 2023, doi: 10.1007/s43681-022-00178-0.

74. M. Fengchun, H. Wayne, R. Huang, H. Zhang, and UNESCO, *AI and education: A guidance for policymakers*. UNESCO Publishing, 2021.

75. B. Shneiderman, “Bridging the Gap Between Ethics and Practice: Guidelines for Reliable, Safe, and Trustworthy Human-centered AI Systems,” *ACM Trans Interact Intell Syst*, vol. 10, no. 4, p. 26:1-26:31, Oct. 2020, doi: 10.1145/3419764.

76. E. Eryurek, U. Gilad, V. Lakshmanan, A. Kibunguchy-Grant, and J. Ashdown, *Data Governance: The Definitive Guide*. O’Reilly Media, Inc., 2021.

77. S. Yanisky-Ravid and S. K. Hallisey, “Equality and Privacy by Design: A New Model of Artificial Intelligence Data Transparency via Auditing, Certification, and Safe Harbor Regimes,” *Fordham Urban Law J.*, vol. 46, p. 428, 2019.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.