AI applications in SG for reliability, security, and stability.

Theodore Kindong[1]

[1] Linköping University, SE-581 83 Linköping, Sweden

Abstract

The new paradigm in clean, sustainable, dependable, and efficient energy generation, and delivery, has led to the transformation and innovation of traditional grid to smart grid. The transformation and innovation use advanced technologies such as AI and IoT to monitor and control power generation, transmission, and distribution processes in a smart grid (SG). AI applications in SG have emerged as an innovation that guarantees effective, flexible, reliable, sustainable, decentralized, secure, and cost-effective distribution and management of energy in SG. This study is a research proposal for AI Applications in SG for Reliability, Security, and Stability. It begins by introducing the SG and its related challenges, followed by AI applications in SG and its implementation challenges. The study identifies research problems in AI applications in SG to be AI interpretability and formulates three research questions that can help address the problem identified. Also, this paper presents the results of the literature review conducted to provide a sufficient grounding for this study and discusses the following concepts of AI application in SG. Predictive Analytics in SG, AI-enabled Demand Response in SG, AI-enabled Control and Coordination in SG, AI-enabled security, stability, and reliability analysis in SG, and Implementation challenges of AI applications in SG. The study proceeds to discuss the proposed theoretical approach and the chosen research methodology and then concludes with the expected study contribution to research and practice.

Keywords

Smart grid, artificial intelligence, predictive analytics, SG stability, and AI interpretability

1. Introduction

The term "smart grid" refers to the transformation of a traditional electric power grid, which was originally regulated using electromechanical methods, into a network that is controlled using information and communication technologies (ICTs). The Smart Grid (SG), as outlined in the US Department of Energy's Smart Grid System Report, includes information management, control technologies, digitally based sensors, information and communication technologies (ICTs), and field devices[1]. Its deployment has arisen as a viable way to improve energy efficiency and tackle the problems presented by increasing energy consumption and environmental issues. Its rise has been due to the integration of Internet of Things (IoT) technology and other emerging technologies into the traditional grid. Which has facilitated the implementation of sophisticated monitoring and control systems to enhance energy management efficiency. SG system employs IoT-enabled sensors and devices to gather real-time data on energy usage, generation, and environmental conditions [2]. These sensors offer a plethora of data that may be utilized to study energy patterns, detect inefficiencies, and make well-informed decisions to enhance energy efficiency.

Due to the expansion of the SG, which includes more interconnection, greater integration of renewable energy, widespread use of direct current power transmission systems, and the liberalization of electricity markets[1-3]. The smart grid facilitates the gathering of vast quantities of complex and diverse data on the operations of the electric power grid [1]. This is

BIR-WS 2024: BIR 2024 Workshops and Doctoral Consortium, 23rd International Conference on Perspectives in Business Informatics Research (BIR 2024), September 11-13, 2024, Prague, Czech Rep.
theodore.kindong@liu.se (T. Kindong)
0009-0001-6502-8325 (T. Kindong)

achieved by combining modern metering infrastructure, control technologies, and communication technologies[1, 2]. This rapid expansion requires effective stability analysis and control to ensure a stable, reliable, and secure operation of the smart grid. Thus, the stability characteristics of the smart grid have become significantly more difficult compared to the past. Necessitating a transition from conventional stability analysis and control methodologies to more sophisticated stability analysis and control due to the limitations of traditional stability analysis and controls in terms of speed, efficacy, and cost [3].

More so, the vast amount and multidimensional data collected pose analysis challenges for conventional modeling, optimization, and control technologies that are restricted in their ability to handle large amounts of data, leading to a growing recognition of the need for artificial intelligence (AI) techniques in the smart grid [1]. Hence, various artificial intelligence algorithms are being applied in SG for stability analysis and control, such as load forecasting, power grid stability evaluation, fault detection, and security issues. Furthermore, research asserts that the utilization of AI approaches can augment and enhance the dependability and robustness of smart grid systems[1]. This has continued to be the case with the rise in machine learning (ML) applications and the growing use of data-driven approaches to improve resilience, stability, reliability, and security in SG. This shows significant prospects for utilizing ML techniques in SG to anticipate power outages (POP), forecast energy demand, perform security, analyze stability, and control. Hence, artificial intelligence (AI) applications in SG offer robust and encouraging tools for analyzing stability and controlling smart grids, which continue to gain increasing attention in academics and practice.

Despite the overwhelming interest in the application of AI in SG, they have yet to significantly address how AI applications in SG can be transparent and human-understandable. Humans are increasingly interested in understanding and interpreting AI decisions within SG. However, existing literature on AI applications in SG has not given sufficient attention to the necessity to address AI algorithms' inherent opacity and their functioning as "black boxes"[4]. This raises the challenge of transparency and AI interpretability, which is the capacity to provide explanations to humans about the functions of AI and clarity about AI decisions. Hence, this study seeks to explore AI interpretability within SG can be addressed. To meet users of AI applications continuous desire to comprehend what the AI model has acquired from the data, understand and interpret AI decisions, and determine bias when AI makes false predictions as well as return the autonomy to make the ultimate decision. Furthermore, the absence of interpretability in current literature poses a significant risk, particularly in application scenarios that demand high levels of security, such as the smart grid. Therefore, the objective of this study is to investigate how interpretable models can be applied in SG. Interpretable models offer the ability to identify and pinpoint the underlying source of abnormalities, providing valuable insights into the reasons for their failure[3]. The remainder of this study covers the background and research problem, previous research, knowledge gaps, the study's significance, and research questions. The second section of this work presents a literature review and chosen theoretical lens. The chosen research methodology follows this and finally concludes with the expected study's contribution to research and practice.

1.1. Background and Research Problem

Electricity has been a paramount and extensively utilized kind of energy since the 19th century[3]. The electric power grid, responsible for transmitting electricity from power plants to end-users, has been hailed as the most significant engineering accomplishment of the 20th century[5] and has become an essential component of contemporary civilization. Nevertheless, the electricity grid continues to evolve. The growing integration of renewable energy sources, such as wind and solar energy, into power networks is being driven by worries over fossil fuel depletion, climate change, global warming, and the decreasing prices associated with these sources[3]. The evolution of the electricity grid has continued with the integration of emerging disruptive technologies such as IoT and AI tackled primary obstacles in the energy system such as environmental awareness, inadequate clean and sustainable energy management, insufficient optimization of energy distribution and transmission, high costs associated with power transfer, and increased customer awareness of energy charges [6].

Similarly, SG relies on communication and IT for its operation [7], which involves a network of sensors in transmission-distribution infrastructure, remote monitoring, SCADA, lumbar systems, and household appliances. Meters, sensors, and synchrophasors provide significant data [5, 8]. This vast amount of data requires the best mechanisms for handling, investigating, and evaluating it to ensure SG stability, security, reliability, and resilience. Hence, AI techniques are being adopted to produce intelligent electricity that is stable and secure. AI techniques enable control hubs and energy supply structures to remotely monitor power-driven campaigns to advance and analyze the power system in real time, reducing principal times [9].

Also, integrating renewable energy sources and technologies such as IoT raises the problem of SG's stability, reliability, security, control, and resilience. Hence, to ensure widespread access to the grid while prioritizing high levels of stability, control, security, and resilience, AI applications in SG have emerged as a potential solution. AI applications in SG have been used to enable real-time grid analysis and predictive analysis using historical and current data to provide an accurate representation of the current grid state to producers and consumers.

Additionally, it can predict future grid status. Therefore, it is feasible to detect power grid transmission loss, locate the overheated line, identify missing power connections, make decisions, and implement self-healing measures[6]. AI techniques promise to help prevent power system failures, including minor outages and major blackouts, through power outage prediction and weather forecasting. These types of events provide relief to both customers and the nation, as they prevent significant economic and business losses.

Despite the promising future of AI applications in SG for communicating, retaining information, and making decisions through analysis. There has been growing interest in addressing AI interpretability in SG. As SG combines human and technological interactions, there is a need for humans within SG to be able to interpret AI decisions and their operations for a more accurate analysis, intuitive, and collaborative system. AI interpretability in SG provides an opportunity to effectively integrate the behaviors of producers, consumers, and users to provide reliable, cheap, and reliable power sources.

1.2. Previous Research

In recent years, there has been increasing interest in both academia and industry on AI applications in SG. This is due to a global increase in the popularity of artificial intelligence (AI) and its transformative nature in other domains. Their inventive and disruptive nature has led to their widespread recognition and practical applications in several industries [1, 10]. Previous research has examined the implementation of artificial intelligence in the SG and substantial changes resulting from integrating and utilizing artificial intelligence (AI) in SG [1, 5, 6, 8, 11, 12].

More so, past studies have explored how AI techniques such as machine learning can be used in power outage prediction and energy demand forecast [9, 11, 13-16]. Kehkashan et al. [16], proposed a systematic approach for selecting optimal machine learning models to improve the performance of the SG. Their study examines the efficacy of machine learning algorithms by utilizing performance evaluation measures as the fundamental criteria for selecting the most optimal model [16]. Similarly, other studies have sought to address security concerns in SG using AI techniques, such as the work of Abdullah et al. [17]. Their work examined how the integration of AI with Blockchain distributed ledger technology (BDLT) can be applied in SG in the realm of renewable energy and associated power automation [17]. Also, Liu et al. [18] address security in SG by exploring a new method for detecting a specific type of attack, known as Bad Data Injection

(BDI), in the Smart Grid. Their method combines information from the network traffic flow and the power system's physical laws to create a unified model called Abnormal Traffic-indexed State Estimation [18].

Furthermore, prior research has mostly addressed distinct issues, such as the analysis and control of stability in an efficient manner [3], as well as the utilization of Deep Learning (DL) for predicting power consumption during peak hours[19, 20]. Similarly, Barth et al. [14] investigate how distributed reinforcement learning can enhance decision-making autonomy and collaboration in SG.

However, despite the overwhelming interest in AI applications in SG, there needs to be more literature on how AI interpretability is addressed to improve trust and transparency. The absence of technical expertise and the inability to comprehend AI judgments [21] have not been resolved, resulting in trust issues around the use of AI in SG. Current research has been scarce on how artificial intelligence's predictive models in SG can be made transparent and easily comprehensible to people. Hence, this study aims to build on the fact that human intervention can improve AI applications when AI systems fail or exhibit bias [22, 23]. To explore how AI interpretability can be integrated into SG when solving issues related to stability control, dependability, security, and transmission cost systems in SG. It will also investigate how AI interpretability in SG can increase the accuracy of predictive models and support human operators in obtaining and exploiting data throughout the grid. The study will build on the growing interest in AI applications in SG to explore the integration of AI interpretability for all participants' more effective, intuitive, and collaborative use. This approach aims to achieve global optimization and distribute the resulting expenses across all participants. AI applications in the smart grid offer several advantages, including global optimization, cost-sharing among players, cost-effectiveness, environmental friendliness, stability analysis, and developing a dependable SG [1, 3].

1.3. Knowledge Gaps

Current studies on AI applications in SG have often neglected the aspect of interpretability, and the research on this topic is still in its early stages, with only a limited number of references available. The black-box aspect of AI poses a challenge to its implementation in smart grids, as most AI agents cannot achieve perfect accuracy and lack an understanding of the specific reasons behind AI failures. Given the importance of SG stability, security, and management for safety, operators must comprehend the decisions made by AI agents and identify the source of any abnormalities, which has not been the case in the current literature. AI must earn the trust of operators before it can be extensively utilized for smart grid stability analysis and control.

Also, studies around the AI applications in SG have predominantly addressed the technical aspect of AI, principally focusing on improving AI performances and ignoring AI algorithms' potential biases and failures. The research in this area is still in its early stages, with only a few references available. The black-box aspect of AI poses a challenge to its implementation in smart grids as most AI agents are unable to achieve perfect accuracy, and we need an understanding of the specific reasons behind AI failures. Given the importance of power grid stability and management for safety, operators must comprehend the decisions made by AI agents and identify the source of any abnormalities. Only until AI gains the trust of operators can it be extensively utilized for smart grid stability analysis and control.

1.4. Importance and Significance of the Research

With the rising adoption of AI techniques in SG to guarantee stability, reliability, resilience, and security and the growing reliance on these techniques, it is crucial to understand the trade-off between Accuracy and Interpretability for successful implementation in SG. Hence, interpretability has been a significant focus of AI research. However, this feature has received limited attention in the field of SG research. Typically, the effectiveness and comprehensibility of AI methods are two variables that must be balanced against each other [24] , making this study very important. The current approaches of AI applications in SG have achieved exceptional performance at the expense of extensive abstraction, hence the necessity to achieve a somewhat effective equilibrium between the accuracy and interpretability of AI.

Furthermore, SG is a critical infrastructure requiring effective security, stability, reliability, and resilience. Thus, the application of AI in SG must adopt an effective method for creating AI algorithms that are easier to understand by incorporating a user interface that presents the AI algorithm's results and reveals part of the underlying reasoning behind the AI decision process [21]. The need to make AI techniques in SG offers explanations for important AI decisions that, when partially accurate and, in some cases, can give misleading reassurance, [21] reinforces the importance of this study. Also, it is important to develop new training methods for creating high-quality AI algorithms in SG that are easily understandable for security assessment. A study that integrates AI interpretability in SG by providing a coherent explanation for the AI algorithms' output and identifying the factors contributing to system instability and security issues is crucial.

1.5. Research Questions

The objective of this study is to examine the current body of literature on the use of artificial intelligence (AI) in smart grids (SG), assess specific instances of AI implementation, and establish the boundaries for ensuring the interpretability of AI in SG. Existing research on AI in SG primarily emphasizes AI's technical capabilities while neglecting the crucial aspect of human comprehension and interpretation of AI decisions. In contrast, this study takes a distinct approach by addressing AI's technical abilities and the human perspective.

Therefore, our research aims to utilize the latest advancements in artificial intelligence (AI) and its applications, as well as big data and data generated by smart grids (SG), to investigate how AI can be applied transparently to smart grids. The objective is to augment human comprehension of AI determinations and promote their involvement and self-governance to optimize energy distribution and administration in smart grids. This will be achieved by answering the following research questions:

RQ1: What is the present state of AI research in SG?

RQ2: How can we balance AI's effectiveness in SG and its ease of understanding?

RQ3: How does AI interpretability in SG affect its reliability, security, and stability?

To provide sufficient grounding for the study and answer the above research questions, the next sections of the paper present a preliminary literature review.

2. Literature review

A comprehensive analysis of the existing literature was undertaken to provide additional clarity regarding the framework and themes of this study. A comprehensive literature evaluation is essential for any study. So, it is for this PhD thesis as it gives ideas and concepts to support the chosen approach to the issue, helps determine the appropriate methodology, reveals areas where information is lacking, and demonstrates the distinctive contribution of the thesis[25]. The literature evaluation aimed to facilitate the formulation of theories, address research gaps, and identify areas that require more investigation [26].

2.1. Search Strategy

The comprehensive literature requires a search strategy to identify relevant literature. Thus, a keyword search was conducted in the relevant databases LiU UniSearch, Scopus, IEEE Xplore, Google Scholar, ACM, and Web of Science to collect pertinent publications related to the subject of interest. This search followed the systematic procedure of doing a literature review outlined by [26] . A comprehensive literature search was conducted, which included a systematic search of relevant titles in reverse order to ensure thorough coverage. This involved searching for related literature among the references cited by studies identified in the initial keyword search.

The author initially identified the key themes of "Smart Grids," artificial intelligence," “reliability,” “predictive analytics,” "smart grid security," and "stability of smart grid" in Scopus, a comprehensive database of peer-reviewed literature. Four concepts were discovered while investigating AI-enabled Reliability, Security, and Stability in Smart Grid Systems: Demand response, artificial intelligence techniques, intrusion detection, and security. Subsequently, we identified specific terms or acronyms associated with each concept. For instance, "power outage prediction" and "deep learning" serve as terms or acronyms for the concept of "machine learning". Ultimately, we employed the OR operator to combine terms or concepts associated with the same concept, whereas the AND operator was utilized to merge unrelated concepts. Due to the query phrase employed in our search on Scopus, we obtained many search results. The same was done for other databases, and inclusion and exclusion criteria were used to select quality papers for analysis. The selected articles were synthesized into the themes below.

2.2. Predictive Analytics in SG

Predictive analytics are crucial in the administration of SG systems. Forecasting models for unregulated factors (such as the generation of renewable energy sources and building energy usage) are necessary for the optimal management of SG, as these models enable informed decision-making and facilitate fault identification and diagnostics[13]. SG consists of smart meters and sensors that capture real-time data on energy use, voltage, current flows, and other factors across the grid. The collected data is processed, aggregated, and stored in a central repository called the data management system[2]. This system integrates data from several sources to provide a complete perspective of grid performance. It uses information, two-way communication technologies, and computational intelligence to ensure stability, security, reliability, resilience, sustainability, and efficiency[7, 12, 27]. Thus, predictive analytics have emerged in the literature, especially for real-time data collection and analysis in SG, which is used for efficient control and coordination through data-decision making. Also, with the growing demand for environmentally friendly local electricity production and delivery in an urban area, predictive analytics is used as an optimization technique to assess historical data and forecast future energy demand patterns[13, 16, 20, 28].

More so, predictive analytics makes modern AI models essential for risk prediction and decision-making [23]and has gained interest in the area of AI applications in SG. It uses data mining, predictive modelling, and machine learning to analyze historical and real-time data. The technique is the “modern oracle of our networked digital age” [23]. SG uses AI to improve predictive analytics for fault diagnosis, prediction, decision-making, and optimization, such as in the work of [11, 13, 15, 20, 28].

Also, Ahmad et al. [13] comprehensive study compares tree-based ensemble machine learning models (random forest – RF and extra trees – ET), decision trees (DT), and support vector regression (SVR) to forecast solar thermal collector system useable hourly energy[13]. Their approach involved training and testing machine-learning models with experimental data. Similarly, AI-enabled predictive analysis helps fight climate change, lower energy transmission costs, and predict energy demand and grid stability[16]. Thus, energy producers and consumers accept predictive analytics and seek ways to improve it [28]. SG power systems benefit greatly from AI techniques, according to Bose [11]. He gives a brief but comprehensive explanation of expert systems (ES), fuzzy logic, and artificial neural networks.

Furthermore, Zhongtuo et al. [3] present a clear summary of SG's use of predictive analytics for grid stability analysis and control. AI in SG analyzes smart grid security, stability, fault diagnostics, and stability control. Abdullah et al.'s study focuses on One of the real-time analyses of the physical layer of the smart grid, leveraging predictive analytics for intelligent information processing (IIP) to SG and its management and bidirectional data channel applications, ensuring secure communication through an effective control mechanism. Predictive analytics in SG improves supervisory control and data acquisition (SCADA) systems' monitoring, control, and coordination[29]. It also allows for reconfiguring the power system, advanced metering infrastructure (AMI), protection, distribution automation (DA), and embedded intelligence that prioritizes self-healing, optimization, and recovery from anomalies, which are also improved [30].

2.3. AI-enabled Demand Response in SG

Demand response (DR) is a key concept in SG, encouraging users to discharge non-essential electricity during peak hours to balance peak-hour electricity supply. AI applications in SG have also been used to optimize DR by integrating different AI techniques in SG to estimate consumer electricity consumption and automate DR. DR in SG has been addressed in the literature, such as the work of Qunzhi et al. [31]. They used semantic Web approaches to create an integrated Smart Grid information model and present semantic information case studies for dynamic DR. Their work demonstrates that the semantic model simplifies information integration and knowledge representation for subsequent Smart Grid applications[31].

Similarly, José R. and Zoltán [32], in their work, reviewed the use of reinforcement learning (RL) for demand response applications in the SG to control diverse energy systems such as electric vehicles, heating, ventilation, and air conditioning (HVAC) systems, smart appliances, or batteries[32]. More so, Ma et al.[20] , in their work, explored the method of distribution optimization on the multi-agent system to determine the ideal network weights for different stakeholders, resulting in an optimal dynamic pricing strategy that provides smart grid economic efficiency and security[20]. Their proposed method improves the optimization of the RL algorithm, stakeholder privacy, and decision-making autonomy in SG. Furthermore, Boopathy et al. [19] , in their study, cover deep learning (DL) applications for intelligent smart grid demand response. Presented DL fundamentals in SG demand response and examined cutting-edge DL applications in SG, including electric load forecasting, state estimation, energy theft detection, energy sharing, and trading[19]. Finally, we discuss existing research problems, critical issues and potential paths in DL for smart grids and demand response. Deep learning (DL) models can discover patterns from the massive SG network data and estimate electricity demand and peak hours. Several studies have examined DL principles for DR in SG.

2.4. AI-enabled Control and Coordination in Smart grids

Control and coordination are becoming vital as smart grids transmit large amounts of information. Hence, existing literature has examined how SG uses cutting-edge power electronics, computer systems, information technology, emerging technologies, and cyber technology. To coordinate, produce, distribute, and use electricity sustainably, environmentally friendly, and reliable[3]. Qunzhi et al.[31] argue that Smart Grids' ability to collect huge information enables new software applications and tools to revolutionize macro and micro power consumption management to satisfy rising electricity demand[31]. Also, AI techniques have emerged as more agile tools for controlling and coordinating generation transmission and distribution of electricity in SG from and to the power grid, as well as real-time pricing by third-party service providers to control home energy use[31].

Zhongtuo et al. [3] argue that transmissive information in SG requires sophisticated information processing to add value and meaning in real time to ensure effective stability analysis, control, and coordination. The expansion of the smart grid, which includes more interconnection, greater integration of renewable energy, widespread use of direct current power transmission systems, and the liberalization of electricity markets, has made grid stability more complex and communication more transmissive[3]. AI-enabled Intelligent information processing allows smart grids to integrate computer processing logic, vast database repositories, and communications network connectivity, expanding the concept of control and coordination before "facilitating or enabling certain tasks"[3, 11, 14]. Conversely, growing artificial intelligence (AI) approaches offer robust and promising tools for smart grid stability analysis and control and are gaining interest.

2.5. AI-enabled security, stability, and reliability analysis in SG

SG integrates renewable energy sources, decentralized power generation units, energy storage, and plug-in hybrid electric vehicles (PHEV), raising security, stability and reliability concerns. Hence, current literature has explored AI-enabled security approaches and stability and reliability analysis[3, 9, 15, 17, 18, 33]. Rahman et al. [33] proposed a new approach for customer reliability in SG by forecasting distribution power systems using a fault tree technique with customer-weighted component failure frequencies and downtimes[33]. Their method goes a notch higher than traditional electric grid customer reliability forecast that uses system average (SA) component failure frequency and downtime weighted by component amount[33]. To include weight component, failure frequency and downtime predictions with customer disturbance data. Similarly, other studies have investigated how different AI systems can detect abnormalities and predict assaults to protect stakeholders' privacy by prohibiting the flow of personally identifiable information[9, 15, 17, 18].

More so, Zhongtuo et al.[3] provide a thorough and lucid overview of the current progress of AI applications in SG with a comprehensive introduction to AI, encompassing its definitions, historical background, and cutting-edge approaches. Their study thoroughly examines how it might be applied to evaluate security, measure stability, diagnose faults, and control stability in smart grids[3]. Also, Abdullah et al.[17] review the latest integrated artificial intelligence and blockchain-enabled smart grid and power distribution automation scheduling, management, optimization, privacy, and security[17]. Their research focuses on real-time smart grid physical layer analysis and automation using AI and blockchain.

2.6. Implementation challenges of AI applications in SG

Despite the growing interest in AI applications in SG, AI capabilities and the ability to address many more difficult problems than conventional mechanism-based approaches. Also, AI applications in SG have been extensively researched and yielded outstanding results. The practical implementation of AI applications in SG confronts obstacles such as high data requirements, learning from imbalanced data, interpretation, transfer learning, robustness to communication quality, and robustness to attack or adversarial examples[3].

More so, Smart grids face the ongoing issue of managing massive amounts of data with significant fluctuation due to their reliance on power systems for communication networks[1]. This also raises another significant barrier to AI applications in SG, especially for power outage prediction, which is data quality, since it depends on how the ML algorithms are trained and tested[16].

Finally, the issues of AI interpretability in SG hugely affect the implementation of AI applications in SG, as it limits the integration of domain expertise. AI algorithms can be complex and challenging to comprehend[1, 3, 16], and the absence of interpretability in AI models presents a difficulty to leveraged domain expertise gained over the years, of power system operation. The expertise reduces AI's data dependence and boosts performance. The skill could be used to integrate data-driven approaches with symbolic AI or knowledge engineering.

2.7. Proposed Theoretical Approach

SG has emerged as an innovative technology that addresses the growing energy demand, and it has continued to evolve with the application of AI in SG. AI applications in SG are a recent technology that requires methods that should be used immediately in their design and implementation to ensure their adoption. Involving an inter-organizational community with many different types of people creates and uses a vision of the SG innovation that is key to its early and later spread [34]. This vision is very important for understanding, validating, setting up and accelerating economic tasks and changes in SG. Several institutional forces affect how an organizing idea grows and what effect it has[34]. Also, considering the early stage of AI applications in SG, there are many doubts about its benefits, usability in diverse contexts, usage patterns, and future[35]. Hence, organizations need to understand a new idea before adopting it. This study adopts organizational vision theory as its theoretical lens to investigate how organizations understand AI applications in SG. According to what Creswell and Creswell [36] call a "theoretical lens," this study uses organizing vision theory to shed light on AI applications in SG. We use this idea on purpose because organizations and individuals try to understand innovations before adopting them, and by using this knowledge, people become adopters or non-adopters[37]. The comprehension process offers much room for interpretation regarding how IT artefacts work collaboratively with humans in a team. The team members are analyzed regarding their motives and behaviours [38].

To better understand how humans benefit from the computing capabilities of AI and how AI uses information and large amounts of data generated from the SG. The organizational vision, a “focal community idea for applying information technology in organizations”, [34] influence potential adopters' decision-making and comprehension. The organizing vision is shaped by public discourse between suppliers, consultants, journalists, university researchers, early adopters, practitioners, and executive groups. The community calls each organizational vision by its name and incorporates metaphors, scenarios, stories, difficulties, and dilemmas [34]. The organizational vision will enhance AI interpretability, allowing users to interpret, legitimize, and mobilize new technologies. Interpretation clarifies the AI's existence, decisions, and function to reduce uncertainty. Legitimization relates AI applications in SG to benefits and commercial operations[34] such as stability, reliability, and security.

3. Methodology

The methodology encompasses planning and procedures for data collecting, analysis, and interpretation, from basic assumptions to specific methods. It also presents the philosophical tradition, methodological approach, and method of data collection and analysis.

3.1. Philosophical Tradition (Paradigm)

The philosophical paradigm is the researcher's philosophical assumptions, that informed the strategies, data collecting, analysis, and interpretation methodologies that should guide this selection. To construct an effective plan, researchers must consider their philosophical worldview assumptions, the inquiry strategy associated with this worldview, and the research procedures or processes that translate the approach into practical application[39]. Information systems research is predominantly categorized into positivism, interpretivism, and critical realism. This research focuses on comprehending the epistemological assumptions, which are the principles for creating and assessing reliable knowledge about a particular phenomenon [40], hence adopting the interpretive paradigm.

The choice for interpretive paradigm is because it is primarily linked to qualitative research as it asserts that reality is socially constructed (ontology) and can only be comprehended by interpreting the underlying significance that people attribute to it (epistemology) [36]. This study seeks to understand events by accessing the subjective interpretations that individuals attribute to them. Unlike the previous "descriptive" studies on AI applications in SG, this study does not accept the notion of an "objective" or "factual" explanation of events and situations. Instead, it aims to achieve a relativistic yet commonly understood comprehension of phenomena[40]. The study's objective is to heavily depend on the participants' viewpoints regarding the topic being examined.

3.2. Methodological Approach

This study adopts ethnomethodologically informed ethnography (EM) as a research method, as it seeks the integration of ethnography into the systems development process. This is driven by the belief that the social context in which systems are put greatly impacts their success[41]. AI applications in SG are utilized in social situations, regardless of their technological features, investigating their design and use is best suited in their natural setting. Hence, ethnography, which focuses on observing interactions in natural situations, can provide a social perspective on system design[42]. The choice of EM is because it might inform system design and the challenges of integrating research findings with information system designers' needs. It is suitable for addressing challenges in transdisciplinary collaboration between designers and EM researchers. This study explores how developing information systems under a specific framework can address the valid complaints of thoughtless technical implementation[42].

More so, examining circumstances in their authentic contexts and comprehending or interpreting occurrences related to the significance individuals attribute to them[43], offers a better approach to addressing the issues of AI interpretability in SG. Also, Christin's [4] research expands upon the work of prior ethnographers such as Seaver, presenting an alternative epistemological perspective that diverges from the conventional "black box" framework[4]. This design involves a comprehensive description, analysis, and interpretation of social expressions between people and groups, which typically refers to a program, event, activity, process, or one or more individuals[36]. This research design is selected because the study focuses on a current subject that necessitates a specific time frame and organized research activities for the researcher to gather comprehensive information using different data-collecting methods during the specified duration.

3.3. Methods of Data collection

The method of data collection employs "scavenging" methods to collect relevant information from various sources (such as informal conversations, official announcements, reviews from AI apps in SG, and industry conference areas). To illuminate the intricate relationship between computer systems' social, cultural, and technological components in our daily lives.

More so, the study will observe and interview specialists to acquire actual data on algorithm creation and use in SG. It will include participatory and non-participatory observations supported by semi-structured qualitative interviews with administrators, data scientists, and community users.

3.4. Method of Data Analysis

To get clarity and valuable information from the collected data, the study will use thematic analysis. A technique for finding, examining, and interpreting meaning patterns or "themes" in qualitative data[44]. Hence, the study will record and evaluate qualitative inquiry transcripts using coding methods comparable to Cech [45]. To identify similarities between workplace cultures and algorithmic design, interviews and field notes will be transcribed and categorized using grounded theory and thematic analysis[46]. To show how social practices influence algorithm use and user impact.

4. Expected Contribution

This study uses research techniques to solve a practical problem. Thus, it will benefit practitioners working on AI applications in SG. Initially, AI app developers for SG will gain, by learning how to incorporate domain expertise to design and deploy interpretable AI in SG. The findings will also assist systems or software developers, IT security specialists in creating secure, robust, and interpretable AI models, and company leaders and managers preparing to use AI. This PhD thesis research could help make AI technology more inclusive by incorporating domain expertise, semantic information, and interpretability into its material design.

Also, energy producers and consumers will utilize this knowledge to construct stable, reliable, and secure SG using interpretable AI algorithms for real-time detection of anomalies and intrusions, power outage prediction, and demand response. According to Gregor [47], the finding's theoretical contribution to practitioners is design and action. Similarly, the study's contribution to knowledge according to Gregor's categories [47] will be explanatory because it explains interpretable AI in SG. AI applications in SG are gaining tremendous interest, hence the impact of AI's "black box" characteristics on smart grid operations and how smart grid operators can tackle this issue. It will help academics create human-centered AI solutions in SG. This study will demonstrate how interpretable AI solutions in SG might benefit a growing research area.

References

[1] Olufemi A., O. and N. Haoran, Artificial Intelligence Techniques in Smart Grid: A Survey Smart Cities, 2021.

[2] Sunkari, P., et al., Analysis and Functioning of Smart Grid for Enhancing Energy Efficiency Using Optimization Techniques with IoT, in International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA). 2023, IEEE.

[3] Zhongtuo, S., et al., Artificial intelligence techniques for stability analysis and control in smart grids: Methodologies, applications, challenges and future directions. Applied Energy, Elsevier 2020.

[4] Christin, A., The ethnographer and the algorithm: beyond the black box. Theory and Society: Springer Nature 2020.

[5] Arnold, G.W., Challenges and Opportunities in Smart Grid: A Position Article. Proceedings of the IEEE, 2011.

[6] Mhamud, H.S., et al., Towards electric digital twin grid: Technology and framework review. Elsevier, Energy and AI, 2023. **11**.

[7] Faheem, M., et al., Smart grid communication and information technologies in the perspective of Industry 4.0: Opportunities and challenges. Computer Science Review, 2018.

[8] Bhuiyan, E.A., et al., Towards next generation virtual power plant: Technology review and frameworks. Renewable and Sustainable Energy Reviews, 2021.

[9] Beniwal, R.K., et al., A critical analysis of methodologies for detection and classification of power quality events in smart grid. IEEE Access, 2021.

[10] Wang, J., et al., The evolution of the Internet of Things (IoT) over the past 20 years. Computers & Industrial Engineering, 2021. **155**.

[11] Bose, B.K., Artificial Intelligence Techniques in Smart Grid and Renewable Energy Systems—Some Example Applications. IEEE XPLORE, Proceedings of the IEEE, 2017. **105**(11): p. 2262 - 2273.

[12] Dhara, S., A.K. Shrivastav, and P.K. Sadhu, Smart grid modernization: Opportunities and challenges. Electric Grid Mod, 2022.

[13] Ahmad, M.W., J. Reynolds, and Y. Rezgui, Predictive modelling for solar thermal energy systems: A comparison of support vector regression, random forest, extra trees and

[14] Barth, D., B. Cohen-Boulakia, and W. Ehounou, Distributed Reinforcement Learning for the Management of a Smart Grid Interconnecting Independent Prosumers. Energies, 2022. **15**.
[15] Cao, H., D. Zhang, and S. Yi, Real-Time Machine Learning-based fault Detection, Classification, and locating in large scale solar Energy-Based Systems: Digital twin simulation. Solar Energy, 2023.
[16] Kehkashan, F., et al., Machine learning for power outage prediction during hurricanes: An extensive review. The International Journal of Intelligent Real-Time Automation, 2024.
[17] Abdullah, A.K., et al., Artificial intelligence and blockchain technology for secure smart grid and power distribution Automation: A State-of-the-Art Review. Sustainable Energy Technologies and Assessments, 2023.
[18] Liu, T., et al., Abnormal traffic-indexed state estimation: A cyber–physical fusion approach for Smart Grid attack detection. Future Generation Computer Systems, 2015.
[19] Boopathy, P., et al., Deep learning for intelligent demand response and smart grids: A comprehensive survey. Elsevier, 2024. **15740137**.
[20] Ma, H., et al., Optimal demand response based dynamic pricing strategy via Multi-Agent Federated Twin Delayed Deep Deterministic policy gradient algorithm. Engineering Applications of Artificial Intelligence, Elsevier, 2024.
[21] Burrell, J., How the machine 'thinks': Understanding opacity in machine learning algorithms. sage Journals: Big Data & Society, 2016. **Vol. 3**(No 1).
[22] Chandre, P.R., et al., Explainable AI for Intrusion Prevention: A Review of Techniques and Applications, in ICT with Intelligent Applications. 2023, Springer Link.
[23] Lazaro, C. and M. Rizzi, Predictive analytics and governance: a new sociotechnical imaginary for uncertain futures. International Journal of Law in Context, 2023.
[24] Shuang, W., et al., Improved Deep Belief Network and Model Interpretation Method for Power System Transient Stability Assessment. Journal of Modern Power Systems and Clean Energy, 2020. **Vol. 8**(No. 1).
[25] Hart, C., Doing a Literature Review: Releasing the Social Science Research Imagination. Sage London, 1998.
[26] Levy, Y. and T.J. Ellis, A Systems Approach to Conduct an Effective Literature Review in Support of Information Systems Research. Informing Science: The International Journal of an Emerging Transdiscipline, 2006.
[27] Falvo, M.C., et al., Technologies for smart grids: A brief review, in 12th International Conference on Environment and Electrical Engineering. 2013, IEEE: Wroclaw, Poland
[28] Feng, L., et al., Complex-valued artificial hummingbird algorithm for global optimization and short-term wind speed prediction. Expert Systems with Applications, 2023.
[29] Rizzetti, T.A., et al., Cyber security and communications network on SCADA systems in the context of Smart Grids, in 50th International Universities Power Engineering Conference (UPEC). 2015, IEEE: Stoke on Trent, UK.
[30] El-Hawary, M., The Smart Grid—State-of-the-art and future trends, in Eighteenth International Middle East Power Systems Conference (MEPCON). 2016, IEEE: Cairo, Egypt.
[31] Qunzhi, Z., et al., Semantic Information Modeling for Emerging Applications in Smart Grid, in 2012 Ninth International Conference on Information Technology - New Generations. 2012, IEEE: Las Vegas, NV, USA.
[32] José R., V.-C. and N. Zoltán, Reinforcement learning for demand response: A review of algorithms and modeling techniques. Applied Energy, Elsevier, 2019.
[33] Rahman, F.A., et al., Application of fault tree analysis for customer reliability assessment of a distribution power system. Reliability Engineering & System Safety, 2013. **111**: p. 76-85.
[34] Swanson, E.B. and N.C. Ramiller, The Organizing Vision in Information Systems Innovation. Organization Science 1997. **8**(5).

[35] Tona, O. and S.A. Carlsson. THE ORGANIZING VISION OF MOBILE BUSINESS INTELLIGENCE. in European Conference on Information Systems. Association for Information Systems.
[36] Creswell, J.W. and J.D. Creswell, Research Design: Qualitative, Quantitative, and Mixed Methods Approaches FIFTH EDITION. . 2017: SAGE Publications, Inc.
[37] Swanson, E.B. and N.C. Ramiller, Innovating Mindfully with Information Technology. MIS Quarterly, 2004. **Vol. 28**(No. 4): p. pp. 553-583.
[38] Walsham, G., Doing interpretive research. European Journal of Information Systems, 2006. **15**.
[39] Creswell, J.W. and J.D. Creswell, Research Design: Qualitative, Quantitative and Mixed Methods Approaches 3rd Edition. SAGE. 2008, Sage London.
[40] Orlikowski, W.J. and D. Robey, Information Technology and the Structuring of Organizations. Information Systems Research, 1991. **Vol. 2**(No. 2).
[41] Crabtree, A., et al., Ethnomethodologically informed ethnography and information system design. JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE., 2000. **Vol. 51**(No. 7): p. pp. 666 – 682.
[42] Srinivasan, R., Ethnomethodological Architectures: Information Systems Driven by Cultural and Community Visions. JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY, 2007. **Vol. 58**(No. 5): p. pp. 723–733.
[43] Crang, M. and I. Cook, Doing Ethnographies. 2007, London, Thousand Oaks, New Delhi: SAGE Publications.
[44] Clarke, V. and V. Braun, Thematic Analysis. The Journal of Positive Psychology, 12(3), 297–298. . 2016.
[45] Cech, F. Tackling Algorithmic Transparency in Communal Energy Accounting through Participatory Design. in Proceedings of the 10th International Conference on Communities & Technologies - Wicked Problems in the Age of Tech. 2021. Vienna, Austria: ACM.
[46] Passi, S. and S.J. Jackson. Trust in Data Science: Collaboration, Translation, and Accountability in Corporate Data Science Projects. in Proceedings of the ACM on Human-Computer Interaction. 2018. New York, NY: ACM.
[47] Gregor, S., The Nature of Theory in Information Systems. MIS Quarterly, 2006. **30**(3): p. 611-642.