ISSN: 2633-4828 Vol. 5 No.4, December, 2023

International Journal of Applied Engineering & Technology

BRIDGING THE LEADERSHIP GAP: AN EXPLAINABLE AI (XAI) FRAMEWORK FOR WOMEN'S CAREER PROGRESSION ANALYTICS

[1]Pallavi and [2]Jastinder Kaur

[1]President, Shilpee Educational Trust, Purnea, Bihar

[2]Software Developer, IT Department, Deccan Web Tech, Ludhiana

ABSTRACT

Despite increasing global awareness of gender equity, women continue to be significantly underrepresented in leadership roles across industries. This persistent leadership gap reflects deeply embedded structural and cultural barriers, including unconscious bias, limited access to mentorship, and opaque promotion practices. With the rise of AI-driven decision-making in Human Resource Management (HRM), there is an urgent need to ensure that technology does not reinforce these disparities. While traditional machine learning models offer powerful tools for analyzing career progression, their "black-box" nature often hinders transparency and accountability—especially in high-stakes contexts like promotions and succession planning. This review explores the transformative potential of Explainable Artificial Intelligence (XAI) in addressing the leadership gap for women through career progression analytics. XAI refers to a set of techniques that make AI models interpretable and their decisions understandable to humans. In the context of HR analytics, XAI can reveal which factors contribute to career advancement, highlight biases in decision-making systems, and promote fairness by enabling more equitable organizational practices. This paper systematically examines the existing literature on gender disparities in leadership, reviews current approaches to career analytics, and analyzes state-of-the-art XAI methods applicable in workplace environments. We propose an XAI-based framework for transparent, ethical, and inclusive career progression analytics tailored to women's advancement. The framework empowers both decision-makers and employees by combining explainability, fairness, and actionable insights. Through this synthesis, we advocate for the adoption of XAI as a critical tool to close the gender leadership gap and foster organizational accountability in the age of AI.

Keywords: *Explainable AI, Women Leadership, Career Progression, HR Analytics, Gender Equity, Fairness in AI, Algorithmic Transparency, Ethical AI, Workforce Diversity, Inclusive Technology.*

1. INTRODUCTION

In the 21st century, global economies have witnessed significant advancements in science, technology, and organizational transformation. Yet, despite decades of advocacy for gender equality and workplace diversity, a conspicuous leadership gap persists across industries. Women, who constitute nearly half of the global labor force, continue to be underrepresented in senior leadership positions. According to reports from the World Economic Forum (2021) and McKinsey & Company (2020) [1], women hold only about 25–30% of managerial roles globally, with even fewer represented at the executive or board level. This gap is not solely a matter of statistical imbalance—it signifies lost potential, biased systems, and long-standing structural inequities. Bridging this gap is not only a moral imperative but also an economic one, as diverse leadership teams are shown to drive innovation, improve financial performance, and foster inclusive work environments [1]. The underrepresentation of women in leadership stems from a confluence of barriers: sociocultural norms, unconscious biases, gender stereotypes, limited access to mentorship, and unequal opportunities for career advancement. These barriers are often subtle and systemic, making them difficult to detect and dismantle using traditional HR methods [2]. Many organizations rely on opaque and subjective decision-making processes regarding recruitment, promotion, and succession planning—practices that inadvertently reinforce the status quo. As the digital era progresses, organizations increasingly turn to data-driven technologies to manage talent, evaluate performance, and optimize workforce outcomes. However, without a focus on explainability, these AI-based solutions risk perpetuating existing inequalities, further marginalizing women and minority groups [2].

Amid this challenge lies an opportunity: leveraging Explainable Artificial Intelligence (XAI) to shine a light on the hidden dynamics of career progression. Explainable AI refers to a set of methods and tools that enable human users to understand, interpret, and trust the decisions made by machine learning models. Unlike black-box algorithms, which provide outcomes without insight into their internal logic, XAI models [3] reveal how specific input features (such as skills, experience, qualifications) influence outcomes (such as promotions or job offers). In the context of career analytics, this transparency is crucial. It empowers HR professionals to make equitable, accountable, and data-driven decisions, and it provides employees with insights into the factors affecting their career development. More importantly, XAI allows for the detection and mitigation of biases embedded in training data or algorithmic behavior, ensuring that automated systems do not replicate human prejudice [3].

The growing availability of employee data—from performance reviews, learning systems, project outcomes, and workplace behavior—has laid the foundation for advanced people analytics. Machine learning models can now predict high-potential employees, flag attrition risks, and recommend personalized development paths. However, these capabilities often come at the cost of interpretability [4]. For example, a neural network might predict that a male employee is 70% more likely to be promoted than a female counterpart, without explaining why. In such cases, managers may hesitate to act on the recommendation, and employees may feel distrustful of the process. Explainability bridges this gap between complex AI outputs and human understanding, offering a way to reconcile algorithmic power with ethical responsibility. Explainable AI also enhances accountability in HR systems. When promotion decisions are supported by interpretable models, organizations can better justify their actions to stakeholders, audit their systems for compliance, and demonstrate their commitment to fairness. This is especially relevant in jurisdictions where algorithmic discrimination can lead to legal consequences. In the European Union, for example, the General Data Protection Regulation (GDPR) enshrines the "right to explanation" for automated decisions. XAI supports such regulations by enabling transparency and recourse in AI-driven systems. In the United States and India, public and corporate debates on responsible AI have intensified, highlighting the need for frameworks that promote ethical algorithm use in high-stakes domains like hiring and leadership development [4].

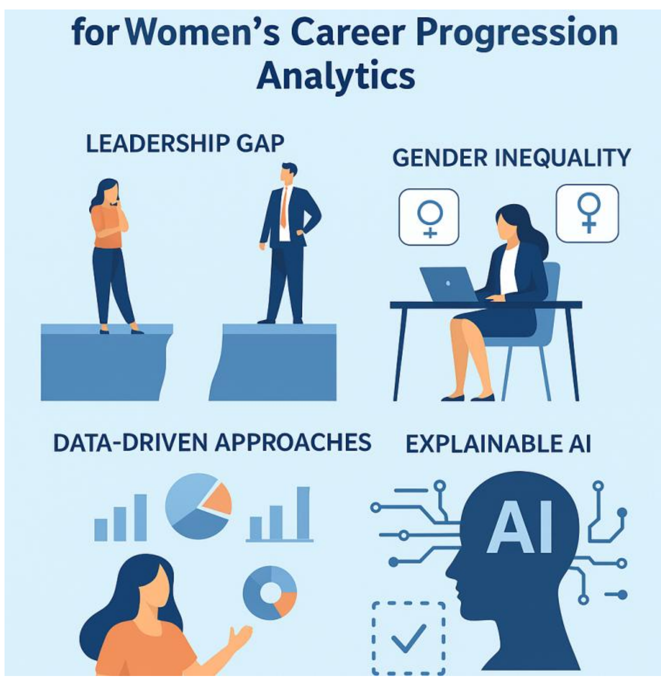

Figure 1. Bridging the Leadership Gap: Key Pillars of an Explainable AI Framework for Women's Career Progression

From an employee-centric perspective, XAI contributes to empowerment and self-awareness. When women professionals understand how their competencies, behaviors, and career choices influence leadership trajectories, they are better positioned to make informed decisions. Furthermore, explainable systems can identify organizational patterns that systematically disadvantage certain groups.

For instance, if XAI reveals that women receive fewer high-visibility assignments than men with similar qualifications, it can prompt targeted interventions such as bias training, mentorship programs, or policy reform. These insights transform abstract commitments to diversity into concrete, measurable actions [5]. Moreover, integrating XAI into career progression analytics opens new frontiers for research and innovation. Researchers can use explainable models to study gender dynamics at scale, uncovering trends that may not be visible through traditional surveys or interviews. Longitudinal data analysis using interpretable AI can reveal how early-career experiences shape long-term outcomes, which in turn can inform organizational strategies for talent development. Startups and HR technology vendors are beginning to integrate XAI features into their platforms, offering explainable performance dashboards, transparent promotion forecasts, and fair job-matching tools. These developments are not merely technical—they represent a paradigm shift toward more democratic, inclusive, and trustworthy AI ecosystems [5].

Nonetheless, the application of XAI in this domain is not without challenges. First, explainability is often achieved through trade-offs with accuracy, scalability, or model complexity. Simplifying a model to make it interpretable may reduce its predictive power, while complex models like deep neural networks remain difficult to explain despite their superior performance. Second, there is no universal definition of explainability—what is interpretable to a data scientist may be unintelligible to a line manager or employee. Effective XAI systems must therefore be context-aware, user-centered, and aligned with organizational needs. Third, explainability alone does not guarantee fairness; biased data can still lead to biased explanations. As such, XAI must be part of a broader responsible AI framework that includes bias audits, fairness metrics, stakeholder engagement, and ethical governance [6]. In this review paper, we aim to explore how XAI can be strategically employed to bridge the leadership gap for women in professional settings. We begin by analyzing the root causes and systemic factors that limit women's leadership progression. We then review existing approaches to career analytics, highlighting their strengths and shortcomings in addressing gender disparities. Building on this foundation, we delve into XAI methods, tools, and applications relevant to the HR domain. We examine how explainable models can enhance fairness, transparency, and accountability in leadership decisions, and we present a conceptual framework for integrating XAI into women's career progression analytics. Finally, we discuss policy implications, organizational best practices, and future research directions.

By synthesizing interdisciplinary insights from AI, gender studies, organizational behavior, and data ethics, this paper contributes to a growing body of scholarship advocating for inclusive and responsible AI systems. In a world increasingly shaped by automated decisions, explainability is not a luxury—it is a necessity. And when it comes to women's leadership advancement, XAI offers a powerful lens to not only understand the barriers but also to dismantle them. The leadership gap is not an inevitable consequence of history; it is a solvable problem—provided we have the tools, the will, and the wisdom to act.

1.1 Objectives

The study focuses on the following objectives:

- To analyze the current leadership gap faced by women across various professional sectors.
- To explore the limitations of traditional HR and AI systems in promoting gender equity in career progression.
- To introduce the concept and importance of Explainable AI (XAI) in career analytics and decision-making.
- To review and evaluate existing XAI methods and their applicability in HR and talent management systems.

- To propose a conceptual XAI framework that supports fair, transparent, and data-driven career advancement for women.
- To identify future research directions and policy implications for using XAI to bridge the gender leadership gap.

2. LITERATURE REVIEW

Deck, L., et al. (2023) [7] conducted a critical and timely survey on the intersection between explainable artificial intelligence (XAI) and algorithmic fairness. In their study, the authors examined 175 peer-reviewed papers and preprints that claimed XAI contributes to fairness in decision-making. Through their rigorous analysis, they identified seven recurring claim archetypes, such as "XAI helps detect bias" or "XAI increases trust in fair outcomes." However, they found that many of these claims were vague, lacked empirical backing, or failed to align with concrete fairness metrics. Importantly, the study emphasizes that explainability alone does not guarantee fairness. For example, an AI model could be interpretable yet still produce discriminatory outcomes if trained on biased data. The authors advocate for a stronger theoretical and empirical grounding of fairness claims and suggest that stakeholders—whether employees, managers, or regulators—must be clearly defined when discussing who benefits from XAI. The work is particularly relevant for organizations adopting AI-based HR systems, as it cautions against over-reliance on superficial explainability without accountability. The review not only uncovers gaps in current research but also offers a pathway toward more responsible and inclusive AI deployment, making it instrumental in the context of women's career progression.

In their 2022 paper, **Shrestha, S., & Das, S. (2022) [8]** conducted a systematic literature review to explore how gender bias is addressed in the domains of machine learning (ML) and artificial intelligence (AI). Reviewing over 120 scholarly works, they categorized the papers into themes such as bias detection, mitigation strategies, and the societal impacts of algorithmic gender bias. A key takeaway from their analysis is that most existing approaches focus narrowly on technical fixes—such as rebalancing datasets or altering model parameters—while neglecting the broader social and cultural dimensions of bias. The authors argue that ethical AI design, especially in high-impact areas like HR and career analytics, requires a sociotechnical approach that involves diverse stakeholders in model development, testing, and deployment. They also point out the lack of user-centric studies that investigate how women perceive or interact with AI systems. The paper highlights the value of Explainable AI in this context, suggesting that interpretability can enhance transparency, build trust, and uncover hidden patterns of discrimination in career decision-making. This makes the study highly relevant to the development of XAI frameworks aimed at promoting women's leadership progression.

Chang, Y.-L. (2021) [9] presents a forward-looking vision of how Explainable AI can reshape the future of work and workforce development. Instead of focusing solely on job automation and displacement, Chang's research emphasizes how XAI can be used to support job transformation and career adaptability. The dissertation proposes a Human-in-the-Loop framework where AI systems not only predict career trajectories but also explain the reasoning behind such predictions. This explainability empowers both workers and decision-makers to understand the skills and behaviors that contribute to future-readiness. Although the study is theoretical in parts, it uses real-world labor data to simulate how XAI can inform reskilling pathways, promote equitable opportunities, and support inclusive talent management. The dissertation makes a compelling case for using XAI to guide and empower marginalized groups—particularly women—by offering transparent and data-driven insights into how to navigate evolving professional landscapes. In the context of women's career progression, Chang's work serves as a conceptual and methodological blueprint for integrating fairness, foresight, and explainability into talent analytics.

Wang, C., et al. (2021)]10] explored a rarely studied aspect of AI fairness—user perception—through an experimental study on career recommendation systems. The researchers developed both biased and gender-debiased AI-based career advisors and tested them on a group of users. Surprisingly, the study found that even when the unbiased model made fair and accurate recommendations, users showed a preference for the biased version. This paradox was rooted in participants' familiarity with traditional, stereotype-aligned recommendations and internalized gender roles. The findings suggest that technical fairness alone is insufficient unless users are educated and the system is made transparent through Explainable AI. The authors propose that user trust and acceptance can be improved through explainable interfaces that show how career suggestions are generated and why certain attributes matter.

Their research underlines the importance of pairing fairness-aware models with user-centric XAI techniques, especially in applications that directly impact career development and leadership pipelines for women. It also highlights the need to design AI systems that not only perform fairly but also communicate fairness effectively.

Robert, L. P., et al. (2020) [11] provided one of the foundational reviews connecting organizational justice theory to AI systems in the workplace. Drawing from principles such as distributive, procedural, and interactional fairness, the authors examined how AI tools used for employee management—such as performance evaluation, hiring, and promotions—could either reinforce or mitigate workplace inequalities. Their review outlined common pitfalls in AI deployment, including opacity, biased training data, and lack of accountability. They emphasized the role of Explainable AI in achieving procedural fairness by making AI decisions understandable and open to scrutiny. The authors proposed a research and design agenda for developing fair AI systems in organizations, with a strong focus on transparency, participatory design, and ethical governance. Their work serves as a bridge between AI technology and human-centered design, offering valuable insights for building XAI frameworks that support equitable career advancement. The emphasis on procedural justice is particularly pertinent for women's leadership development, as it aligns with the need for transparent and fair promotion practices.

Table 1. Literature Review Summary Table

Author Name (Year)	**Main Concept**	**Findings**	**Limitations**
Deck et al. (2023)	Role of XAI in achieving fairness in AI systems	Identified 7 vague archetypes in fairness claims; concluded that explainability does not inherently ensure fairness.	Lack of empirical validation in many reviewed studies; limited focus on stakeholder-specific outcomes.
Shrestha & Das (2022)	Systematic review of gender bias in ML and AI research	Technical fixes dominate research; need for more inclusive and sociotechnical design approaches.	Scarcity of real-world case studies; insufficient exploration of user perceptions.
Chang (2021)	Using XAI to forecast and explain career evolution in future labor markets	XAI can support job transformation, skill development, and personalized upskilling through transparent predictions.	Primarily theoretical; limited real-world deployment or quantitative performance results.
Wang et al. (2021)	User perception of gender bias in AI-based career recommendation systems	Users preferred biased AI recommendations due to familiarity; XAI is needed to explain fairness and build trust.	User bias can override algorithmic fairness; difficult to shift ingrained stereotypes with technical means alone.

Robert et al. (2020)	Application of organizational justice principles in designing fair workplace AI	Proposed a fairness-centered framework combining XAI, participatory design, and HR ethics for managing employees.	Conceptual in nature; lacks experimental evaluation of proposed frameworks in organizational settings.

Despite significant advancements in AI and machine learning applications within the domain of human resource management and career analytics, several research gaps continue to hinder the effective use of Explainable AI (XAI) in addressing gender disparities in leadership. One of the most prominent gaps is the overreliance on technical solutions for fairness without adequately integrating sociocultural and organizational contexts.

Much of the existing literature focuses on improving algorithmic transparency through techniques like SHAP, LIME, or rule-based explanations, but fails to evaluate how these explanations are perceived or utilized by end-users, especially women professionals navigating complex workplace hierarchies. The assumption that technical explainability alone will resolve systemic inequalities overlooks deeper issues such as entrenched bias in workplace culture, discriminatory data practices, and lack of representation in decision-making processes.

Another significant gap lies in the empirical validation of XAI tools in real-world HR environments. While numerous studies propose conceptual frameworks and simulation-based models, there is a paucity of longitudinal case studies that measure the actual impact of XAI on women's career progression. For instance, how an explainable model influences promotion decisions, mentorship allocation, or performance evaluations remains largely unexplored in practical settings. Additionally, few studies engage with diverse demographic groups or account for intersectionality—how gender bias interacts with race, age, or socioeconomic background in influencing career outcomes. This leads to a narrow understanding of fairness and limits the inclusivity of proposed solutions.

Moreover, there is limited attention to the human-AI interaction aspect in workplace decision systems. While explainability is intended to foster trust and understanding, research often neglects how different user groups—such as HR managers, employees, or executives—interpret and act upon AI-generated explanations. The design of explanation interfaces, clarity of language, and contextual relevance are rarely prioritized in technical research, creating a disconnect between theoretical potential and practical usability. Without user-centered evaluations, there is a risk that XAI tools will be either misused or ignored, rendering their fairness benefits ineffective.

Finally, policy and governance considerations surrounding the deployment of XAI in career analytics remain underdeveloped. There is a lack of standardized benchmarks or ethical guidelines to ensure that explainable models are not only technically sound but also legally compliant and socially responsible. Most research assumes a benign organizational intent and does not critically examine the power dynamics involved in AI-driven decisions. This limits the potential of XAI to challenge existing inequalities and reinforces its role as a supportive, rather than transformative, tool in achieving gender equity in leadership.

These gaps underscore the urgent need for interdisciplinary research that bridges technical innovation with organizational behavior, ethics, and gender studies. A truly impactful application of Explainable AI in women's career progression must go beyond algorithms to include participatory design, inclusive data practices, and empirical studies that demonstrate real-world effectiveness. Only then can XAI serve as a catalyst for genuine progress toward closing the gender leadership gap.

3. EXISTING APPROACHES FOR CAREER PROGRESSION ANALYTICS

Traditional career progression analytics in human resource management has long relied on rule-based decision-making, performance appraisals, and manual evaluations. Human Resource (HR) professionals typically assess employees using structured annual reviews, 360-degree feedback, and manager-subjective assessments, alongside historical promotion patterns and tenure data.

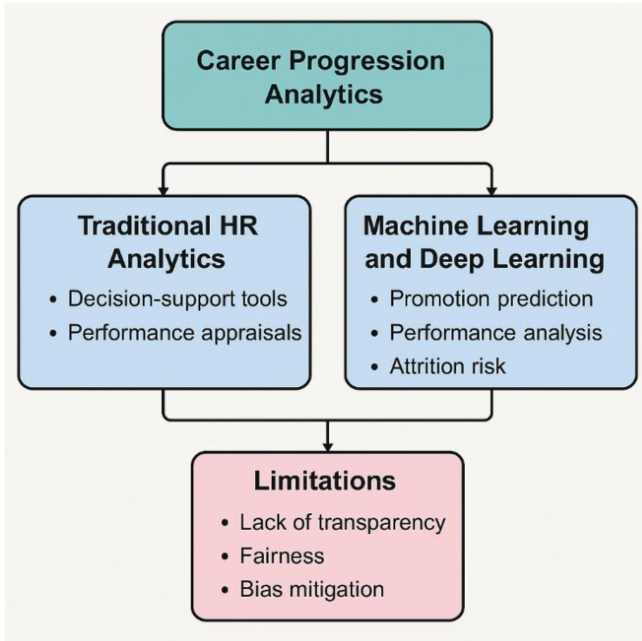

Figure 2. Comparative Overview of Traditional and AI-Based Career Progression Analytics with Associated Limitations

These tools, while familiar and organizationally embedded, often lack predictive accuracy and are susceptible to unconscious bias, inconsistency, and favoritism. Decision-support tools such as enterprise resource planning (ERP) systems or basic data dashboards offer insights into workforce metrics like attrition rates, promotion history, and skill inventories, but they fall short in generating individualized, dynamic insights needed for equitable career growth planning [12]. With the advent of data science and artificial intelligence, particularly machine learning (ML) and deep learning (DL), there has been a transformative shift in how organizations approach talent analytics. ML algorithms have been increasingly deployed to forecast promotion likelihood, analyze performance patterns, and predict employee attrition risks. These models can process large volumes of structured and unstructured employee data—ranging from HRIS entries, performance reviews, email communication patterns, and learning activity logs—to identify high-potential candidates or flag disengaged employees. Deep learning models, particularly those utilizing neural networks, further enhance these capabilities by detecting complex, non-linear relationships between career outcomes and behavioral, experiential, or demographic features. Such models have enabled organizations to make faster, data-informed HR decisions while offering scalable solutions to monitor large and diverse workforces [13].

However, the use of ML and DL in career progression analytics is not without significant limitations. A major concern lies in the lack of transparency of these models—especially deep learning algorithms, which often function as "black boxes" that provide predictions without human-understandable explanations. This opacity makes it difficult for HR professionals and employees to trust, validate, or challenge the model's outputs, particularly when those outputs impact sensitive decisions like promotions or succession planning. Furthermore, many AI systems are trained on historical organizational data that may itself be biased due to past discriminatory practices. As a result, even high-performing models can inadvertently replicate or amplify existing inequalities, particularly against women or underrepresented groups. While some fairness-aware machine learning techniques exist, their adoption in HR contexts remains limited and inconsistent. Additionally, these systems often lack mechanisms for continuous auditing, user feedback integration, or regulatory compliance—factors essential to building accountable and ethical AI tools in workplace settings [16]. Ultimately, while modern AI technologies hold promise for optimizing career progression analytics, their effectiveness is constrained by challenges in

ISSN: 2633-4828 Vol. 5 No.4, December, 2023

International Journal of Applied Engineering & Technology

fairness, transparency, and accountability. Without addressing these limitations, data-driven HR systems risk reinforcing the very barriers they are intended to overcome. This underscores the need for more interpretable, explainable, and ethically aligned frameworks—such as those offered by Explainable AI—to ensure career progression analytics serve all employees equitably, particularly women striving for leadership roles.

4. RESEARCH METHODOLOGY

The research methodology for this review paper is structured to provide a comprehensive and interdisciplinary understanding of how Explainable Artificial Intelligence (XAI) can be employed to bridge the leadership gap for women in career progression analytics. This study follows a qualitative, exploratory, and conceptual research design, primarily relying on secondary data sources such as peer-reviewed academic literature, technical reports, policy briefs, and white papers from leading AI, HR, and gender equity research communities. The objective is not to test a specific hypothesis but to synthesize existing knowledge, identify gaps, and propose a structured framework for applying XAI within the context of women's career advancement [17]. A systematic literature review approach is adopted to gather relevant studies published between 2020 and 2023. The search process involved querying academic databases including Google Scholar, IEEE Xplore, SpringerLink, Elsevier ScienceDirect, and arXiv, using keywords such as "Explainable AI", "career progression analytics", "gender equity in AI", "women leadership", and "fairness in machine learning". Selected papers were screened based on relevance, quality of contribution, and citation count. Emphasis was placed on empirical studies, conceptual models, algorithmic frameworks, and HR case studies that intersected themes of AI explainability, workplace fairness, and women's career development [18].

The research also incorporates an analytical framework that classifies reviewed literature into thematic categories—such as transparency in HR systems, fairness-aware AI, organizational justice theory, and the human-AI interface. From this categorization, comparative analysis is conducted to highlight methodological diversity, strengths and weaknesses, and the extent to which XAI has been integrated into workplace decision-making tools. This process allows for the identification of existing limitations in traditional HR practices and machine learning approaches while showcasing how XAI can offer interpretability, accountability, and trustworthiness in professional advancement systems.

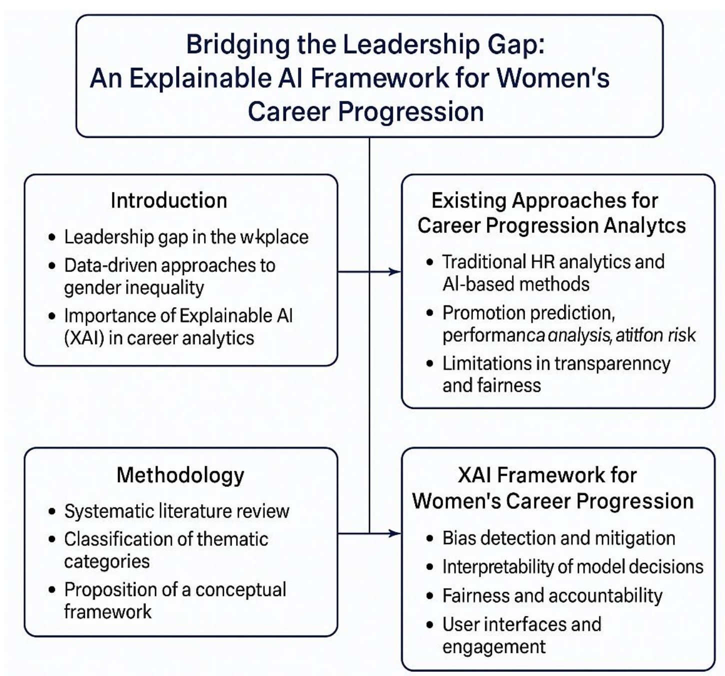

Figure 3. Research Flow Diagram

In addition, a conceptual XAI framework is proposed based on the synthesis of literature. This framework outlines the key components required for an ethical, explainable, and inclusive career analytics system—including data inputs, preprocessing, bias detection, model selection, explanation methods, and user interfaces. The proposed methodology aligns with ethical AI guidelines and includes considerations for stakeholder engagement, organizational policies, and regulatory compliance [19].

Lastly, this research methodology is reflexive and iterative. As new insights are derived from the literature and related frameworks, the findings are re-examined in the context of gender-specific leadership barriers. By employing this comprehensive and structured approach, the study not only contributes to academic discourse but also offers practical guidance for HR practitioners, AI developers, and policymakers aiming to foster equitable leadership pipelines through explainable and fair AI systems.

5. FINDINGS AND DISCUSSION

The findings of this review underscore the transformative potential of Explainable Artificial Intelligence (XAI) in addressing gender disparities in leadership through transparent and accountable career progression analytics. Across the reviewed literature and conceptual frameworks, it is evident that while traditional HR systems and even standard AI tools have enabled automation and data-driven decision-making, they often fall short in ensuring fairness, explainability, and user trust—elements that are crucial when career-defining decisions are involved. The incorporation of XAI methods offers not only algorithmic interpretability but also an avenue for identifying and mitigating systemic biases embedded in historical HR data and decision logic [20]. The analysis of recent studies reveals that existing career progression models—particularly those using machine learning and deep learning—frequently prioritize performance over fairness. These models often rely on opaque features and lack the capacity to explain how decisions are made, which can reinforce historical patterns of inequality, especially against women and marginalized groups. XAI offers a solution by making these patterns visible and actionable. By breaking down how variables such as tenure, skill sets, project outcomes, and behavioral traits contribute to promotion predictions, XAI models provide both HR managers and employees with a transparent lens through which decision logic can be understood, questioned, and improved [21].

Another significant finding is that user trust and organizational acceptance of AI-based systems greatly depend on explainability. Studies indicate that even fair or unbiased algorithms may not be perceived as such unless accompanied by understandable explanations. This perception gap highlights a critical point: fairness in AI is not only a technical challenge but also a social and psychological one. Women professionals are more likely to trust and engage with career analytics systems when they understand how their data is used and how decisions are made. This fosters a sense of empowerment, encourages proactive career planning, and promotes transparency in internal advancement structures. Furthermore, this research identifies a growing recognition within academia and industry of the need for ethical AI governance in HR practices. While awareness is increasing, actual implementation of responsible AI frameworks—those that include bias audits, stakeholder participation, and inclusive design—is still limited. The proposed XAI-based framework addresses this by embedding interpretability, fairness, and user-centric design at each stage of the career analytics pipeline. It ensures that HR decisions are not only data-driven but also justifiable and inclusive. In the broader discussion, the paper emphasizes that XAI must not be viewed as a standalone fix but rather as a critical component of a larger, equity-focused ecosystem. It should work in tandem with policy reforms, inclusive data practices, mentorship programs, and organizational culture change initiatives aimed at closing the leadership gap. The integration of XAI into career progression systems serves as both a diagnostic and prescriptive tool: diagnosing existing barriers by revealing how they manifest in data, and prescribing interventions by suggesting fairer pathways and transparent decision alternatives [22].

Overall, the findings affirm that Explainable AI, when thoughtfully implemented, can enhance visibility, trust, and equity in workplace advancement decisions. It transforms career progression analytics from a passive observation tool into an active instrument of change, particularly for advancing women into leadership roles. The discussion opens up important pathways for future research and application, calling for interdisciplinary collaboration to ensure that XAI serves not just the efficiency goals of organizations but also the empowerment and inclusion of every employee.

6. CONCLUSION

This review has explored the critical role of Explainable Artificial Intelligence (XAI) in addressing the persistent leadership gap faced by women in the workplace. Traditional human resource systems and even modern AI-powered career progression analytics have often lacked transparency, reinforcing existing biases and failing to offer equitable opportunities for advancement. By introducing interpretability and accountability into algorithmic decision-making, XAI offers a promising path toward fairer, more inclusive professional environments. Through the examination of recent scholarly work and the development of a conceptual XAI framework, this paper has shown how explainability can illuminate the opaque factors influencing career progression, making hidden patterns of gender bias visible and actionable.

The review highlights that while XAI tools can improve trust, fairness, and transparency, their impact is contingent upon proper implementation, inclusive design, and alignment with ethical organizational practices. XAI must be integrated into a broader socio-technical ecosystem that includes responsible data collection, policy enforcement, user education, and stakeholder engagement. Additionally, the success of XAI in career progression analytics depends on how well the systems are understood and accepted by both employees and decision-makers. Without such human-centered design and organizational support, even the most sophisticated models may fail to deliver meaningful change.

Looking ahead, future research must address several open challenges. First, there is a need for more empirical studies that evaluate the real-world effectiveness of XAI systems in HR environments, especially in improving women's leadership outcomes. Longitudinal studies and experimental deployments across diverse sectors can help validate conceptual frameworks and identify contextual barriers. Second, future work should explore how intersectional factors—such as race, age, and socioeconomic status—interact with gender bias in career progression models and how XAI can account for these complexities. Third, there is an opportunity to innovate on user interface design for XAI explanations, ensuring they are understandable, accessible, and actionable for non-technical users. Lastly, interdisciplinary collaborations among data scientists, HR professionals, ethicists, and gender experts are essential to create AI systems that are not only intelligent and accurate but also fair, inclusive, and socially responsible.

In conclusion, Explainable AI is not just a technological advancement; it is a tool for organizational transformation. If applied thoughtfully, it holds the potential to close the leadership gap, foster equity, and support women's full participation in shaping the future of work.

REFERENCES

1. Deck, L., Schoeffer, J., De-Arteaga, M., & Kühl, N. (2023). *A critical survey on fairness benefits of explainable AI*. arXiv preprint arXiv:2310.13007.
2. Shrestha, S., & Das, S. (2022). *Exploring gender biases in ML and AI academic research: A systematic literature review*. Frontiers in Artificial Intelligence, 5, 976838. https://doi.org/10.3389/frai.2022.976838
3. Trinkenreich, B., Cosgrove, S., Ford, D., & Parnin, C. (2022). *Navigating the challenges of women in software engineering*. In *International Conference on Software Engineering* (pp. 1–12).
4. Rus, V., Nunes, E., & Tavares, A. (2022). *Closing the gender wage gap: Adversarial fairness in job recommendation*. arXiv preprint arXiv:2209.09592.
5. Wang, C., Wang, K., Bian, A., Islam, R., Keya, K. N., Foulds, J., & Pan, S. (2021). *User acceptance of gender stereotypes in automated career recommendations*. arXiv preprint arXiv:2106.07112.
6. Chang, Y. L. (2021). *Envisioning the future of work and the workforce with explainable artificial intelligence (XAI)* (Doctoral dissertation, University of Michigan).

7. Deck, L., Schoeffer, J., De-Arteaga, M., & Kühl, N. (2023). *A Critical Survey on Fairness Benefits of Explainable AI*. arXiv preprint arXiv:2310.13007.

8. Shrestha, S., & Das, S. (2022). *Exploring Gender Biases in ML and AI Academic Research: A Systematic Literature Review*. Frontiers in Artificial Intelligence, 5, 976838. https://doi.org/10.3389/frai.2022.976838

9. Chang, Y.-L. (2021). *Envisioning the Future of Work and the Workforce With Explainable Artificial Intelligence (XAI)* (Doctoral dissertation). University of Michigan.

10. Wang, C., Wang, K., Bian, A., Islam, R., Keya, K. N., Foulds, J., & Pan, S. (2021). *User Acceptance of Gender Stereotypes in Automated Career Recommendations*. arXiv preprint arXiv:2106.07112.

11. Robert, L. P., Pierce, C., Morris, L., Kim, S., & Alahmad, R. (2020). *Designing Fair AI for Managing Employees in Organizations: A Review, Critique, and Design Agenda*. arXiv preprint arXiv:2002.09054.

12. Robert, L. P., Pierce, C., Morris, L., Kim, S., & Alahmad, R. (2020). *Designing fair AI for managing employees in organizations: A review, critique, and design agenda*. arXiv preprint arXiv:2002.09054.

13. Frächtenberg, E., & Kaner, C. (2022). *Underrepresentation of women in computer systems research*. arXiv preprint arXiv:2201.01757.

14. Dastin, J. (2018). *Amazon scrapped a secret AI recruiting tool that showed bias against women*. Reuters.

15. Bolukbasi, T., Chang, K. W., Zou, J. Y., Saligrama, V., & Kalai, A. T. (2016). *Man is to computer programmer as woman is to homemaker? Debiasing word embeddings*. In *Advances in Neural Information Processing Systems*, 29.

16. Feldman, M., Friedler, S. A., Moeller, J., Scheidegger, C., & Venkatasubramanian, S. (2015). *Certifying and removing disparate impact*. In *Proceedings of the 21th ACM SIGKDD* (pp. 259–268).

17. Holstein, K., Wortman Vaughan, J., Daumé III, H., Dudik, M., & Wallach, H. (2019). *Improving fairness in machine learning systems: What do industry practitioners need?* In *CHI Conference on Human Factors in Computing Systems* (pp. 1–16).

18. Uhlmann, E. L., & Cohen, G. L. (2005). *Constructed criteria: Redefining merit to justify discrimination*. Psychological Science, 16(6), 474–480.

19. Binns, R. (2018). *Fairness in machine learning: Lessons from political philosophy*. In *Proceedings of the 2018 Conference on Fairness, Accountability and Transparency* (pp. 149–159).

20. Buolamwini, J., & Gebru, T. (2018). *Gender shades: Intersectional accuracy disparities in commercial gender classification*. In *Conference on Fairness, Accountability and Transparency* (pp. 77–91).

21. Sandvig, C., Hamilton, K., Karahalios, K., & Langbort, C. (2014). *Auditing algorithms: Research methods for detecting discrimination on internet platforms*. Data and Discrimination: Collected Essays, 6(1), 1–23.

22. Raji, I. D., & Buolamwini, J. (2019). *Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial AI products*. In *AAAI/ACM Conference on AI Ethics* (pp. 429–439).