INTERNATIONAL JOURNAL FOR LEGAL RESEARCH AND ANALYSIS

Open Access, Refereed Journal Multi Disciplinary
Peer Reviewed

www.ijlra.com

 ISSN: 2582-6433

DISCLAIMER

 ISSN: 2582-6433

EDITORIALTEAM

EDITORS

Dr. Samrat Datta

Dr. Samrat Datta Seedling School of Law and Governance, Jaipur National University, Jaipur.Dr. Samrat Datta is currently associated with Seedling School of Law and Governance, Jaipur National University, Jaipur. Dr. Datta has completed his graduation i.e., B.A.LL.B. from Law College Dehradun, Hemvati Nandan Bahuguna Garhwal University, Srinagar, Uttarakhand. He is an alumnus of KIIT University, Bhubaneswar where he pursued his post-graduation (LL.M.) in Criminal Law and subsequently completed his Ph.D. in Police Law and Information Technology from the Pacific Academy of Higher Education and Research University, Udaipur in 2020. His area of interest and research is Criminal and Police Law. Dr. Datta has a teaching experience of 7 years in various law schools across North India and has held administrative positions like Academic Coordinator, Centre Superintendent for Examinations, Deputy Controller of Examinations, Member of the Proctorial Board

Dr. Namita Jain

Head & Associate Professor

School of Law, JECRC University, Jaipur Ph.D. (Commercial Law) LL.M., UGC -NET Post Graduation Diploma in Taxation law and Practice, Bachelor of Commerce.

Teaching Experience: 12 years, AWARDS AND RECOGNITION of Dr. Namita Jain are - ICF Global Excellence Award 2020 in the category of educationalist by I Can Foundation, India.India Women Empowerment Award in the category of "Emerging Excellence in Academics by Prime Time &Utkrisht Bharat Foundation, New Delhi.(2020). Conferred in FL Book of Top 21 Record Holders in the category of education by Fashion Lifestyle Magazine, New Delhi. (2020).Certificate of Appreciation for organizing and managing the Professional Development Training Program on IPR in Collaboration with Trade Innovations Services, Jaipur on March 14th, 2019

 ISSN: 2582-6433

Mrs.S.Kalpana

Assistant professor of Law

Mrs.S.Kalpana, presently Assistant professor of Law, VelTech Rangarajan Dr.Sagunthala R & D Institute of Science and Technology, Avadi.Formerly Assistant professor of Law,Vels University in the year 2019 to 2020, Worked as Guest Faculty, Chennai Dr.Ambedkar Law College, Pudupakkam. Published one book. Published 8Articles in various reputed Law Journals. Conducted 1Moot court competition and participated in nearly 80 National and International seminars and webinars conducted on various subjects of Law. Did ML in Criminal Law and Criminal Justice Administration.10 paper presentations in various National and International seminars. Attended more than 10 FDP programs. Ph.D. in Law pursuing.

Avinash Kumar

Avinash Kumar has completed his Ph.D. in International Investment Law from the Dept. of Law & Governance, Central University of South Bihar. His research work is on "International Investment Agreement and State's right to regulate Foreign Investment." He qualified UGC-NET and has been selected for the prestigious ICSSR Doctoral Fellowship.He is an alumnus of the Faculty of Law, University of Delhi. Formerly he has been elected as Students Union President of Law Centre-1, University of Delhi.Moreover, he completed his LL.M. from the University of Delhi (2014-16), dissertation on "Cross-border Merger & Acquisition"; LL.B. from the University of Delhi (2011-14), and B.A. (Hons.) from Maharaja Agrasen College, University of Delhi. He has also obtained P.G. Diploma in IPR from the Indian Society of International Law, New Delhi.He has qualified UGC – NET examination and has been awarded ICSSR – Doctoral Fellowship. He has published six-plus articles and presented 9 plus papers in national and international seminars/conferences. He participated in several workshops on research methodology and teaching and learning.

 ISSN: 2582-6433

ABOUT US

INTERNATIONAL JOURNAL FOR LEGAL RESEARCH & ANLAYSIS ISSN

2582-6433 is an Online Journal is Monthly, Peer Review, Academic Journal, Published online, that seeks to provide an interactive platform for the publication of Short Articles, Long Articles, Book Review, Case Comments, Research Papers, Essay in the field of Law & Multidisciplinary issue. Our aim is to upgrade the level of interaction and discourse about contemporary issues of law. We are eager to become a highly cited academic publication, through quality contributions from students, academics, professionals from the industry, the bar and the bench. INTERNATIONAL JOURNAL FOR LEGAL RESEARCH & ANALYSIS ISSN 2582-6433 welcomes contributions from all legal branches, as long as the work is original, unpublished and is in consonance with the submission guidelines.

 ISSN: 2582-6433

"UNVEILING BIAS: EXPLORING GENDER AND RACIAL DISCRIMINATION IN ARTIFICIAL INTELLIGENCE AND ITS IMPACT ON CONVICTIONS AND WORKFORCE EQUITY"

AUTHORED BY - JAISVICA KAUR BAGGA & NIVEDHA K.V

Abstract

It is indeed a scary thing to imagine that your next job opportunity might be shaped by an unseen hand - computer algorithm. In the realm of Artificial Intelligence, the invisible biases ingrained in algorithms wield significant influence, particularly in matters of gender and racial discrimination. Brace yourself for a journey into the heart of this digital labyrinth, where we unveil the biases dictating not just employment prospects but also impacting legal convictions and workforce equity. In an era driven by technology breakthroughs, the widespread use of Artificial Intelligence (AI) has prompted worries about the presence of bias, particularly in terms of gender and race discrimination. As AI algorithms progressively affect decision-making processes, a rigorous assessment of the biases inherent in these systems is required to build a fair and inclusive society.

The discernible bias intrinsic to artificial intelligence within gender-centric domains, particularly in the realm of occupational settings and employment opportunities, has become a focal point of this paper's scrutiny. A plethora of empirical studies elucidate that AI algorithms are prone to perpetuating gender stereotypes, thereby engendering discriminatory outcomes within the ambit of recruitment processes. Legal perspectives underscore the imperativeness of incorporating ethical AI practices, as underscored by instances such as the Amazon hiring tool, which exhibited a proclivity toward favoring male candidates. Illustrative case scenarios, elucidating instances of gender-based discrimination experienced by female job seekers attributable to biased algorithms, serve to accentuate the exigency of addressing this multifaceted issue.

This paper directs its attention towards delineating the accurate assessment of the correlation between AI bias and the maltreatment of women within workplace contexts. The primary objective is to underscore the imperative for heightened ethical standards in AI to mitigate the

 ISSN: 2582-6433

aforementioned repercussions. Furthermore, the paper seeks to elucidate strategies for minimising this disparity, thereby contributing to the advancement of equitable practices in artificial intelligence.

Keywords - Algorithm, Artificial intelligence, Bias, Employment, Discrimination, Workplace

A woman's journey in the world of AI: An introduction

The ever-evolving landscape of technology demands constant adaptation. This will be especially true for women, who are always seated in areas of conflict and vulnerability. As the quote by Anu Madgavkar suggests, "Women are going to be among the groups who are going to have to learn how to adapt and use many of these technologies". This underscores the significance of comprehending the perpetually shifting technological landscape and acknowledging the unique challenges confronted by women in the tech sphere. In light of the above statement and considering the exponential growth of the global artificial intelligence (AI) market, virtually all sectors are poised to experience significant disruption due to the implementation of AI systems. This disruption will manifest in the digitization and automation of work processes, leading to transformative shifts within labor markets. Ultimately it can be expected that the impact will be multifaceted, encompassing the number and profiles of available jobs across industries, as well as the evolving skill sets required of the workforce. This paper adopts a gender-based analytical framework to examine the potential for a vicious cycle of digital gender inequality to emerge[1]. This cycle is hypothesized to be driven by the interplay between pre-existing gender stereotypes and occupational segregation within the workforce, alongside the digitalization and automation spurred by AI implementation. The objective is to offer insights into the gender-specific ramifications of AI technologies. This knowledge is critical for mitigating the potential for social stratification and marginalization.

Dr. Muneera Bano underscores the risk of perpetuating sexism through AI systems, as they can mirror and perpetuate biased patterns encoded in their data as she put forths by stating that "If there is sexism embedded within the data, they will pick up that pattern and exhibit the same sexist behavior in their output. And unfortunately, the workforce in AI is male dominated."

It does not come as a surprise that there are concerns regarding the potential for AI to perpetuate

[1] Mirzazadeh I, "Artificial Intelligence (AI) and Violation of Human Rights" [2023] SSRN Electronic Journal

 ISSN: 2582-6433

sexism. Datasets used to train AI may contain inherent gender biases, which the AI can then learn and replicate. Furthermore, the male-dominated nature of the AI workforce raises concerns that these biases might go unnoticed. From a legal standpoint, discriminatory outcomes based on gender could violate existing anti-discrimination laws for example a scenario where the translation of the English term "nurse" consistently utilizes a feminine noun, while the term "doctor" is translated using a masculine noun. This exemplifies the potential for AI to perpetuate pre-existing societal biases regarding gender roles within professions. Such biases could have far-reaching consequences, including the exacerbation of the gender gap in certain fields and the potential for negative downstream impacts, such as the endangerment of women's lives if, for example, a medical AI system exhibits gender bias in its recommendations. The birthplace of a lot of these biases are simply the stereotypes and normalized norms that exist in society at large.

The role of gender stereotypes in shaping AI culture

It is common knowledge that Stereotypes, both negative and positive, exert significant influence on individuals' consciousness, opinions, judgments, attitudes, and behaviors. Gender-based stereotypes persist in society and contribute to workplace gender inequality and likewise unconscious biases shape one's feelings and thoughts toward certain topics or individuals, often leading to negative associations and decision-making. Implicit biases, influenced by experiences, environments, and societal norms, are universal and can affect everyone[2]. These biases can be mitigated through de-biasing techniques, such as changing the way information is presented or providing tools to overcome unconscious predispositions. In the realm of artificial intelligence (AI), biases are increasingly problematic, particularly when algorithms trained with biased data result in algorithmic discrimination. Underrepresented datasets can lead to biased samples, reinforcing inequality, and perpetuating unfair treatment, especially in critical areas like healthcare and employment. The prevalence of gender biases in AI algorithms impacts various aspects of society, from search engine results to hiring decisions. Despite efforts to reduce biases, algorithmic discrimination persists, contributing to gender disparities in the tech industry and beyond. The deep rootedness of this must be analyzed and one way or another it always goes back to the societal perception of women and races. Representation of gender in society is deeply entrenched in cultural norms and expectations. Men are often portrayed as powerful and dominant, while women are expected to conform to

[2] Rifat Ara Shams, Didar Zowghi and Muneera Bano, "AI and the Quest for Diversity and Inclusion: A Systematic Literature Review" [2023] AI and Ethics

 ISSN: 2582-6433

traditional femininity, perpetuating gender stereotypes and reinforcing unequal power dynamics. Challenging these stereotypes and promoting diverse representations is crucial for fostering inclusivity and dismantling gender-based discrimination. One such manifestation of the society's male centric gaze and an aggregate of patriarchal ideas resulted in the fantasization of Sophia, a humanoid robot.

The development of Sophia, a humanoid robot created by Hanson Robotics, raises questions about the feminization of AI and its implications. While Sophia is designed to resemble a female figure, her capabilities and functionalities primarily revolve around human-like interactions and expressions rather than gender-specific traits. However, the case study of Sophia also sheds light on the potential risks associated with hypersexualization and gender stereotypes in AI. Sophia's appearance, modeled after Audrey Hepburn, incorporates feminine features such as expressive eyes, a slender smile, and delicate facial features. While this aesthetic choice may aim to make Sophia more relatable and appealing to humans, it also reflects societal norms and expectations regarding female beauty standards.

Furthermore, Sophia's interactions, including banter and wisecracks, are programmed responses aimed at mimicking human conversation and emotions. However, the extent to which these interactions reinforce gender stereotypes or contribute to the objectification of women depends on the context and intention behind Sophia's design and programming. The hypersexualization of Sophia, or any AI with a female persona, can perpetuate harmful stereotypes and undermine efforts towards gender equality. By portraying female AI as subservient, passive, or primarily focused on pleasing others, there is a risk of reinforcing traditional gender roles and expectations. Moreover, the lack of agency or autonomy in Sophia's decision-making process raises ethical concerns about the implications of feminized AI. If Sophia's responses are primarily designed to cater to male fantasies or expectations, it could contribute to the objectification and dehumanization of women, even in non-human forms. In conclusion, while Sophia represents an impressive advancement in robotics and AI technology, her case study highlights the need for careful consideration of the implications of feminization in AI development. By promoting diversity, inclusivity, and ethical standards in AI design and programming, we can mitigate the risks of perpetuating harmful stereotypes and biases in AI systems.

 ISSN: 2582-6433

A roadmap to eliminating gender bias in AI

Over the years AI technology has become a powerful driver of innovation, impacting everything from entertainment to healthcare. However, recent events have shed light on potential legal concerns surrounding bias in AI. The concern lies in the possibility that AI systems, particularly those relying on machine learning, can inherent biases from the data they are trained on. This raises the possibility of unfair or discriminatory outcomes, depending on the application of the AI[3]. Fortunately, researchers have developed techniques to address this issue. These techniques focus on different stages of the AI development process. Some methods involve pre -processing the data used to train the AI, while others involve modifying the algorithms during training itself. Finally, there are techniques that can be applied after the training data has been processed. By implementing these bias mitigation techniques throughout the development process, organizations can help ensure their AI systems are fair and comply with evolving legal standards. Machine learning models can be trained in ways that encourage fairness and minimize prejudice. Techniques used during training, known as in-processing models, provide unique chances to accomplish this. One technique uses a concept known as regularization. This prevents the model from becoming unduly reliant on specific features in the training data, such as a person's race, which may be irrelevant to the main task at hand. For example, in a university admissions situation, the model should prioritize characteristics such as GPA and test scores over the applicant's race. Regularization helps to prevent the model from overfitting the training data and prioritizing irrelevant elements. Adversarial training is another strategy used with in-processing models[4]. This strategy pits two models against one another during training[5]. One model focuses on the primary goal, such as deciding on university admissions. The second model attempts to forecast a sensitive attribute, such as race, based on the first model's decisions. If the second model accurately predicts race, the first model is penalized. This inhibits the primary model from using race or other sensitive information to make decisions, resulting in more equitable outcomes.

To effectively analyze the trajectory of adverse bias within Artificial Intelligence (AI) systems,

[3] World Health Organization. (2021). Ethical use of artificial intelligence: principles, guidelines, frameworks and human rights standards. In *WHO Consultation Towards the Development of guidance on ethics and governance of artificial intelligence for health: Meeting report Geneva, Switzerland, 2–4 October 2019* (pp. 8–11). World Health Organization. http://www.jstor.org/stable/resrep35680.8

[4] Hershock, P. D. (2020). *Humane Artificial Intelligence: Inequality, Social Cohesion and the Post Pandemic Acceleration of Intelligent Technology*. East-West Center. http://www.jstor.org/stable/resrep25513

[5]Agrawal, A., Gans, J. S., & Goldfarb, A. (2019). Artificial Intelligence: The Ambiguous Labor Market Impact of Automating Prediction. *The Journal of Economic Perspectives*, *33*(2), 31–50. https://www.jstor.org/stable/26621238

 ISSN: 2582-6433

a fundamental understanding of certain core AI concepts is necessary. One such concept critical to this examination is propensity data modeling. Propensity modeling encompasses a collection of techniques employed to construct models capable of predicting the likelihood of an individual undertaking a specific action. These models achieve their predictive power by analyzing historical behavioral patterns within a designated target audience. In simpler terms, imagine you want to advertise a new fitness program in this scenario, propensity modeling can help us identify people who have shown interest in fitness activities in the past, making them more likely to be interested in a program. However, the seemingly innocent nature of this example transforms significantly when applied to broader contexts, particularly those with the potential to significantly impact individuals' lives. For instance, a financial institution might utilize propensity modeling to make loan approval decisions. If the model is trained on historical data exhibiting a gender disparity where men have a higher application and approval rate for loans, a concerning outcome may emerge. The model, conditioned on this biased data, could predict that future male applicants are more likely to be successful borrowers. This scenario has the potential to perpetuate discriminatory practices and lead to the unjustified rejection of loan applications submitted by qualified women. This example underscores the critical need to acknowledge the potential for historical behavioral data to perpetuate existing societal biases within AI systems. Mitigating such biases necessitates a comprehensive understanding of the underlying mechanisms at play and the implementation of robust safeguards within the development and deployment of AI technologies. In a case study conducted on amazon's recruitment model it was seen that an attempt to leverage machine learning for recruitment has raised concerns about perpetuating gender bias. The model, trained on historical applicant data, mirrored the existing imbalance in the tech industry, favoring resumes with characteristics historically associated with male candidates. This bias stemmed from the inherent lack of gender neutrality in the training data, reflecting the real-world dominance of men in tech roles. While Amazon claims to have addressed this by modifying the algorithms, the possibility of new biases emerging based on keywords or other resume features remains. This case highlights the challenges of achieving gender based fairness in AI-powered hiring systems. Machine learning expert Assistant Professor Nihar Shah emphasizes the ongoing need for development in this area, showcasing the complexities of eliminating bias entirely from algorithmic decision-making. The jarring gender imbalance inadvertently manifested in Amazon's initial attempt at an AI-powered recruitment system. The model, trained on existing data, learned to penalize resumes containing the word "women's," highlighting the potential for bias within such algorithms. Recognizing this issue, Amazon's

 ISSN: 2582-6433

researchers implemented a more refined approach. They created a collection of 500 specialized models, each focusing on specific job functions and locations[6]. These models were trained to analyze a vast array of parameters (over 50,000) extracted from applicant resumes. Notably, the algorithms were programmed to de-emphasize the significance of generic skills commonly listed by many applicants, such as programming languages or frequently used platforms. This shift towards targeted parameters and the downplaying of generic skill sets demonstrates a positive step by Amazon toward achieving fairness and mitigating potential gender bias in its recruitment process[7]. The case emphasizes the crucial need for constant vigilance in identifying and addressing potential biases within AI systems. This commitment is essential for ensuring the responsible development and deployment of AI technology in our society. Despite efforts to achieve gender parity, leading U.S. tech companies continue to struggle with a significant gender gap, particularly in technical roles like software development where men remain vastly overrepresented. The ethical models of a few leading tech companies are analyzed below to gain a more nuanced and informed look into this specific matter.

Ethical AI models: A look into industry pioneers

As most of us are aware, Microsoft and Google are at the forefront of AI innovation, and their initiatives reflect a commitment to ethical considerations in AI development. Let's delve deeper into specific examples from both companies and explore how an Ethical Impact Assessment (EIA) might apply to these cases.

Ece Kamar is a senior researcher in Microsoft's research lab in Redmond, Washington. Her research focuses on AI tools that assist engineers in identifying blind spots in training data, such as the underrepresentation of darker-skinned women, which can lead to AI systems with unacceptable mistake rates on gender classification tasks[8]. A more sophisticated difficulty, she added, is figuring out how and when to intervene and reduce AI systems that reflect and magnify societal biases, not because datasets are incomplete or algorithms are insufficient, but

[6] Agrawal, A., Gans, J. S., & Goldfarb, A. (2019). Artificial Intelligence: The Ambiguous Labor Market Impact of Automating Prediction. *The Journal of Economic Perspectives*, *33*(2), 31–50. https://www.jstor.org/stable/26621238

[7] Herrbach, O., & Mignonac, K. (2012). Perceived Gender Discrimination and Women's Subjective Career Success: The Moderating Role of Career Anchors. *Relations Industrielles/Industrial Relations*, *67*(1), 25–50. http://www.jstor.org/stable/41634302

[8] Farris, C., Jaycox, L. H., Schell, T. L., Street, A. E., Kilpatrick, D. G., & Tanielian, T. (2015). Sexual Harassment and Gender Discrimination Findings: Active Component. In T. L. Schell, A. R. Morral, & K. L. Gore (Eds.), *Sexual Assault and Sexual Harassment in the U.S. Military: Volume 2. Estimates for Department of Defense Service Members from the 2014 RAND Military Workplace Study* (pp. 31–54). RAND Corporation.

 ISSN: 2582-6433

because human cultures are prejudiced. For example, if you search the internet for the term "CEO," you will most certainly find information about top leadership positions in companies and organizations all over the world, as well as a few photographs, most of which are of men. This is not surprising given that women account for less than 5% of Fortune 500 CEOs. At Microsoft, this has prompted an ongoing investigation by engineers at the search engine Bing, in collaboration with experts such as Wallach, into how to best surface results that reflect the active discussion in boardrooms, academia, and on social media about the lack of female CEOs - and efforts to change that.

Similarly over the course of this year, Google has made an effort to promote justice and decrease prejudice in machine learning. The most recent update in this effort addresses gender prejudice by offering feminine and masculine translations of select gender-neutral words on the Google Translate website. Google Translate learns from hundreds of millions of previously translated instances from the web. Historically, it has offered only one translation for a query, despite the fact that the translation could be feminine or masculine. So, when the model produced one translation, it unintentionally repeated existing gender biases[9]. For example, it would be masculine for words like "strong" or "doctor," yet feminine for others, such as "nurse" or "beautiful."

For every single term, such as "surgeon," you will now receive a translation in both the feminine and masculine forms when translating from English into French, Italian, Portuguese, or Spanish. When converting words and sentences from Turkish to English, you'll also receive translations in both languages. For example, if you input "o bir doktor" into Turkish, the gender-specific translations that appear are "she is a doctor" and "he is a doctor."

As mentioned above companies are required to compose an ethical framework to minimize the adverse effects of AI bias, however companies can only go so far on its own, there comes a time where the government will need to utilize its resources and intervene on this matter, as the matter of gender disparity comes under the purview of governance. The imperative for global governance to deliberate and intervene on the issue of AI bias in gender stems from the profound impact that biased artificial intelligence systems can have on societies worldwide. A world government intervention is essential to establish standardized ethical frameworks and

[9] Maas, V. S., & Torres-González, R. (2011). Subjective Performance Evaluation and Gender Discrimination. *Journal of Business Ethics*, *101*(4), 667–681

 ISSN: 2582-6433

guidelines that ensure the development and deployment of AI systems free from gender bias. By fostering international collaboration through meetings and conventions governments can collectively address the challenges associated with AI bias, implement corrective measures, and safeguard against the amplification of gender disparities. This proactive approach not only advances fairness and equity but also reinforces the commitment to harnessing the potential of artificial intelligence for the betterment of global societies.

The Intersection of AI and Governance

On that note the annual meeting of the Commission on the Status of Women (CSW) is a significant event that brings together activists, advocates, experts, and government representatives from across the globe. Taking place in New York, CSW serves as a pivotal platform for stakeholders to assess progress on gender equality, reaffirm commitments, and strategize for the future. CSW67, scheduled for March 6-17, 2023, has identified "Innovation and Technological Change, and Education in the Digital Age for Achieving Gender Equality and the Empowerment of all Women and Girls" as its priority theme. Despite substantial strides in integrating gender equality into international discourse and action, women and girls continue to face entrenched disparities and discrimination worldwide. This session of CSW aims to address these challenges by focusing on the intersection of innovation, technology, and gender equality.

The theme underscores the imperative of ensuring that rapid technological advancements benefit women and girls equitably, rather than exacerbating existing inequalities. Emphasizing a human-centric approach to digitalization, CSW67 seeks to embed feminist principles of inclusion and intersectionality into technological innovation. Central to this effort is the need to close the gender gap in digital access, foster inclusive innovation ecosystems, and eliminate technology-facilitated gender-based violence. Highlighting the economic ramifications of excluding women from the digital sphere, CSW67 underscores the importance of bringing women closer to technology. Failure to do so not only stifles creativity and innovation but also incurs significant economic costs, as evidenced by the substantial GDP losses attributed to women's digital exclusion. The official documentation of CSW67 includes the Report of the Secretary-General, which provides a comprehensive analysis of harnessing technology in the digital age to advance gender equality. The report emphasizes the transformative potential of the digital age for women's empowerment while underscoring the imperative of prioritizing The annual meeting of the Commission on the Status of Women (CSW) is a significant event

 ISSN: 2582-6433

that brings together activists, advocates, experts, and government representatives from across the globe. Taking place in New York, CSW serves as a pivotal platform for stakeholders to assess progress on gender equality, reaffirm commitments, and strategize for the future. CSW67, scheduled for March 6-17, 2023, has identified "Innovation and Technological Change, and Education in the Digital Age for Achieving Gender Equality and the Empowerment of all Women and Girls" as its priority theme. Despite substantial strides in integrating gender equality into international discourse and action, women and girls continue to face entrenched disparities and discrimination worldwide. This session of CSW aims to address these challenges by focusing on the intersection of innovation, technology, and gender equality.

The theme underscores the imperative of ensuring that rapid technological advancements benefit women and girls equitably, rather than exacerbating existing inequalities. Emphasizing a human-centric approach to digitalization, CSW67 seeks to embed feminist principles of inclusion and intersectionality into technological innovation. Central to this effort is the need to close the gender gap in digital access, foster inclusive innovation ecosystems, and eliminate technology-facilitated gender-based violence. Highlighting the economic ramifications of excluding women from the digital sphere, CSW67 underscores the importance of bringing women closer to technology. Failure to do so not only stifles creativity and innovation but also incurs significant economic costs, as evidenced by the substantial GDP losses attributed to women's digital exclusion. The official documentation of CSW67 includes the Report of the Secretary-General, which provides a comprehensive analysis of harnessing technology in the digital age to advance gender equality. The report emphasizes the transformative potential of the digital age for women's empowerment while underscoring the imperative of prioritizing.

Algorithms of oppression

In Europe, a group of experts advised the European Commission that AI systems should treat everyone fairly and avoid discrimination. They also said the data used to train these systems should be inclusive and represent a variety of people. The European Commission is also working on laws for AI that would balance innovation with protecting people's rights. These laws would make sure AI developers and users are held accountable for their creations. In the UK, a government advisory group suggested that existing principles for ethical conduct in public service should also apply to AI. They recognized that AI can have unforeseen consequences, so they proposed a flexible approach to using these principles with AI.A famous

 ISSN: 2582-6433

study on the 'Algorithms of Oppression' (2018) asserts that "artificial intelligence will become a major human rights issue in the twenty-first century". In this way, national and international institutions have already begun to develop policies and procedures for identifying and mitigating biases. In 2019, the European Commission's independent High-Level Expert Group on Artificial Intelligence published a study titled 'Ethics Guidelines for Trustworthy AI'. The document supports "equality, non-discrimination, and solidarity" as a fundamental right, emphasizing the importance of ensuring that systems do not produce unfairly biased outcomes, such as employing inclusive data that represents different populations. The European Commission also intends to establish a legal framework for AI that is consistent with the EU Charter of Fundamental Rights and strives to define the duties of users and providers (2022).This charter seeks to strike a balance between the creativity and benefits of new AI technologies and the fundamental rights of EU people, while also aligning them with EU ideals. It offers a 'risk-based' approach, with the goal of prohibiting artificial intelligence practices that are contrary to EU rules and principles. Although the majority of the case studies in this paper focus on certain characteristics of AI bias, the legal framework seeks to prevent genuine biases in decision-making that may damage its inhabitants. A similar perspective was expressed in a review of Artificial Intelligence and Public Standards by the Committee on Standards in Public Life, an independent organization that advises the UK government. Their paper (2022) concludes that the current 'Seven Principles of Public Life' (selflessness, integrity, objectivity, accountability, openness, honesty, and leadership) should be used as a guideline for integrating AI technologies into public life. Understanding that AI may have far-reaching and unanticipated consequences, the review presents a rough overview of how the Seven Principles can be translated into practice for the use of AI[10]. Overall, it is obvious that current policy recommendations for the regulation of AI focus on general concepts and guidelines, reflecting the ongoing and developing spectrum of difficulties that may require attention.

Furthermore It should be underlined that AI bias is not a hypothetical or theoretical concept. In industry and government, algorithms that provide healthcare, assist with employment, lead police, and establish creditworthiness have shown uneven treatment and impact on identically situated individuals, locations, and objects. As highlighted throughout the study, examples of human language models with a bias against women and persons with disabilities have appeared in recent years; the accuracy of some well-known speech recognition and facial recognition

[10] BOBBITT-ZEHER, D. (2011). GENDER DISCRIMINATION AT WORK: Connecting Gender Stereotypes, Institutional Policies, and Gender Composition of Workplace. *Gender and Society*, *25*(6), 764–786

 ISSN: 2582-6433

algorithms is skewed against African Americans. Even when unintentional, algorithmic bias can result from using unrepresentative training data or biased historical data, or from failing to correct statistical biases. Algorithmic prejudice can occasionally lead to unlawful discrimination in addition to being immoral. However, there is no reason why this new technology must take a bad turn. In order to use algorithms for efficiency and good change, the government will be crucial in minimizing algorithmic bias as both a consumer and a regulator. Governments have significant market power and influence over numerous significant algorithmic use cases as consumers of algorithms. Governments can play this role by establishing guidelines, norms, and policies that lessen algorithmic prejudice. Prior to implementation, the public sector should win over the people's trust with algorithmic use cases and make the use of AI in choices that have a major impact on individuals transparent. All UK agencies are required by law to document any possible or predicted algorithmic discrimination before it is used. This obligation cannot be delegated. Impact assessments should be used to examine the usage of algorithms, just like with any other policy decision. Analyzing real-world performance across demographic groupings is important because data interpretations and algorithms can have unanticipated and unwanted consequences.[11] The public sector should apply a consistent national approach while taking local demands into account. The method of detecting AI bias will involve formulating AI requirements, investigating how to construct compliant algorithms, and then enforcing those criteria. Governments can utilize verification, audit, and certification to ensure compliance, as well as risk assessments, audits, and continuing testing. Requirements will differ depending on industry and setting; for example, in contexts with considerable uncertainty and variance, some differential influence may be tolerated, but in others it must be avoided. Guidelines must mix principles and exact thresholds. Algorithms require a fundamental fairness test to avoid severe disparities in impact. Guiding principles can help to prevent prejudice in critical instances that disparate impact studies may ignore. To design and implement regulations, governments must also increase their technological skills.

Beyond the binary - Conclusion

Gender prejudice has been demonstrated at many levels across society. Women encounter difficulties beginning with early education and continuing through various stages of employment; for example, STEM-related jobs have significant impediments at the executive level. As a result, it is critical to evaluate developing technologies that may perpetuate societal

[11] Ridgeway, C. L. (1997). Interaction and the Conservation of Gender Inequality: Considering Employment. *American Sociological Review*, *62*(2), 218–235

 ISSN: 2582-6433

bias. This study provides a better understanding of gender bias in AI and shows evidence from the diverse literature that gender bias exists in AI systems. As a result, the first step in addressing gender bias in AI is to investigate the underlying causes of this bias that function as a catalyst in AI systems. This study not only argues for the concept of justice in gender bias in society as a whole, but it also paves the way for other researchers and users to gain a better understanding of the notion of gender bias in AI, its contributing factors, and potential mitigation approaches as reported in the current multidisciplinary literature. Fair decisions are seen to have a substantial impact on an individual's contentment and loyalty to a specific product, company, or relationship. The purpose of this research is to examine the elements that contribute to gender bias in AI, as well as the ways that could be used to mitigate gender bias. We expect this research to lay the groundwork for new paradigms that, as our literature survey indicates, are still very much required.

After investigating the role of bias in AI, the deep rootedness of its existence, and means to minimize it through the use of ethical machine learning models supplemented by the helping hands of good governance, it is clear that the primary goal of avoiding gender bias in AI is not only to bring women into equal footing, but also to enable and encourage future participation of women in STEM. In the symphony of technological progress, the inclusion and empowerment of women are not just ethical imperatives; they are the keystones to harmony and innovation. As we explore the limitless possibilities of artificial intelligence, it is critical that women are active participants in defining the algorithms that define our digital future. The collaborative brilliance of various minds drives remarkable advancements while also ensuring that technology's ethical compass remains true. Empowering women in technological settings, particularly in the field of artificial intelligence, is more than just a matter of equality; it is a deliberate decision to unleash the whole range of human potential and creativity. The future of technology must be laced with threads of inclusivity, with women serving not only as beneficiaries but also as architects, guiding us toward a future in which invention recognizes no gender boundaries. Overall, it must be noted that a limitation to this paper is that current efforts to regulate AI focus on broad ideas and guidelines, because the technology is still developing and new issues might arise in the future.