European Economic Letters
ISSN 2323-5233
Vol 14, Issue 4 (2024)
http://eelet.org.uk

AI Ethics in Practice: A Literature Review on AI Professional's perception and attitude towards Ethical and Governance principles of AI.

Ms. Nidhi Kshirsagar[1]

Assistant Professor (Institute of Business Management and Research),

IPS Academy, Indore (MP)

Dr. Ekta Rokade[2]

Assistant Professor, School of Economics,

Devi Ahilya Vishwavidyalaya, Indore (MP)

Abstract :

As the need for AI Ethics is rising, we have analysed prior studies and available literature for understanding the perception and attitude of AI professionals towards Ethical and governance principles of AI in current scenario, those principles that are prioritized over others, the challenges and the barriers faced by the practitioners and the ways and means by which the practical implementation of these principles can be done. Key principles include transparency, accountability, privacy, fairness and human agency. Literature reveals that there is often challenges and barriers arising from organization's side as they are not keen because of time and cost limitations. Also the principles are vague and abstract that makes it all the more difficult for implementation in practical way.

Keywords : AI Ethics, AI ethics principles, AI ethics practitioners, AI ethics frameworks, AI ethics challenges.

Introduction :

Today, we are in the Age of Artificial Intelligence, surrounded by applications that have implications on our daily lives. It is indeed central to our daily course of actions. In businesses it boosts performances, reduces and cuts down costs, offers personalized services to the clients and paves way for digitalization. In spite of its many uses and applications there are often unsolicited instances that forms the headlines of news. The instances of violation of Data privacy, human rights and Undue surveillance, instances of bias and discrimination by the AI enhancing the social, economic, cultural and gender inequalities have got the spotlight for all the wrong reasons.

AI systems, mimic human intelligence. A majority of the data scientists and AI practitioners feel that training data on its own is neutral and cannot be the cause ethical harm, however these get generated by the AI systems which more often learn and get trained on real time with real world happenings. We humans have common sense for judgement, can explain our underlying assumptions, and become accountable for our decisions but these basic characteristics of decision making are unavailable in AI systems and thus in the real world gives immense scope for ethical risks to emerge. In light of these there has been rise in issuance of ethical guidelines, frameworks, preferred ethical practices and code of conduct by several standardizations' bodies, government bodies, private and public academic institutes and business organizations since 2016, and is still a widely discussed topic. Standardization bodies like IEEE have issued guidelines for developers and AI practitioners and even is in process for forming the standards for integrating organizational ethics in projects and processes

European Economic Letters
ISSN 2323-5233
Vol 14, Issue 4 (2024)
http://eelet.org.uk

involving artificial intelligence(IEEEP7999™, P7999.1™, P7999.2™)[i1] also IEEE has been actively promoting the certification program IEEE CertifAIEd™[2] Assessor Training which would enable the authorized AI assessor to assess the ethics of AIS so that the AI system can gain more trust and acceptability amongst the stakeholders and businesses.

AI systems thus need to be under ethical and regulatory framework and needs to pass some checks and balances to be able to be trusted to be allowed to mimic humans and their capabilities. In AI systems trust is assessed through presence of factors such as transparency, fairness, respect for privacy and accountability to name a few. These however are understood in the way the AI systems present a safe behaviour, provide reasonable responses and give predictable outcomes for the given problems.

Image Source : IEEE SA (https://blended-learning.ieee.org/LMSiLMS/Learner/Home#/Program?pgmId=10509)

It indeed calls for a major role of the AI developers and practitioners. In the wake of the same several research articles have been written in the past emphasizing the need for developers themselves to posses the awareness of ethical concepts, theories and behaviour as only then the resultant AI systems would depict the desired ethical qualities. Thus, the onus of Trustworthy and Responsible AI thus falls on the shoulders of AI developers and Practitioners. In a way it important that AI developers become more empowered, self-responsible, and be ethically and morally upright and there are several frameworks that AI developers can refer the GEDAI Framework(Vanhée, L., & Borit, M. (2022))[3].

Objective:

The purpose of this review paper is to understand the perception and attitude of the AI developers and practitioners towards ethical and governance principles in current scenario, those that are prioritized over others, the challenges and the barriers faced and the ways and means by which practical implementation of these principles can be done as explored in the existing literature and prior studies.

[1] https://standards.ieee.org/ieee/7999/11740/?mkt_tok=MjExLUZZTC05NTUAAAGXi6h75qftV1ElldZd5DoWhFVZnP0K1l6dY5qNdTxaV8cPs3eAf6WCmQFasHwNIM0ZYwZkRcRrseLJftlqNaaxuem7M3vZmXEpI63Y3Mm-NJk

[2] https://engagestandards.ieee.org/ieeecertifaied.html?_gl=1*1kd3stx*_ga*MTQ0Mjg4MjQ5Ni4xNzM1NjI2NjUy*_ga_XDL2ME6570*MTczNTYyNjY1Mi4xLjEuMTczNTYyNzMxMy41MC4wLjA.

[3] Vanhée, L., & Borit, M. (2022). Ethical by designer-how to grow ethical designers of artificial intelligence. *Journal of Artificial Intelligence Research*, *73*, 619-631.

Review of Literature:

Peters, D., Vold, K., Robinson, D., & Calvo, R. A. (2020)[4]. The authors and this paper has presented two complementary frameworks the spheres of Technology experience and the responsible design process in order to operationalize ethics design and Technology these frameworks are designed and developed to translate abstract ethical principles of safety, inclusivity, autonomy, justice and alignment with human well-being into actionable practices. In the responsible design process, the framework integrates ethical decision making at all stages and emphasizes the impact evaluation and the inclusive stakeholder environment where as the sphere for technology experience framework analysis ethical impact across six spheres like adoption interface task behaviour life and society and other broader societal context. The authors have made use of case study in digital mental health technologies specifically the online therapy for depression and anxiety and recommends that five ethical principles of autonomy beneficence non-maleficent justice and explicability be adhere, the finding suggest that there is ethical trade-off between privacy and transparency and also lack of universal standards for integrating ethics into design. The authors recommend developers to construct multi-disciplinary teams for addressing ethical dimensions and corporate evaluations of real-world impact post deployment. For organisations it is recommended that they establish systematic processes for Ethical decision making also develop systems for trainings to support professionals in applying ethical frameworks in last researchers or advised to conduct for the studies for creation of ethical frameworks and further explore ethical trade off solutions in complex socio technical systems.

Ryan, M., & Stahl, B. C. (2020)5. This paper provides and elaborate review for existing ethical guidelines which are relevant for AI development. The paper identifies 11 important principles such as transparency fairness accountability justice non-maleficence as core for practical and normative implications while AI development. The study reviews that compliant to all of these principles result and bringing trust for the Ness and better adoption of AI systems. In all a structured literature review has been done for 91 publicly available AI ethics guidelines with the help of thematic analysis data analysis has been done. The study reviews that there are several times conflicting interest like choosing between principals of privacy and transparency also there is the height and need for implementation mechanism and audits to check for fairness of the AR algorithms. Author suggests that ethical guidelines should be integrated into the organisation framework for better implementation. Further it is suggested that concrete standardised guidelines have to be tailor for developers and users' conflicts have to be addressed between principals by establishing priority mechanism and also encourage interdisciplinary collaborations for refining and applying the guidelines

Cihon, P., Kleinaltenkamp, M. J., Schuett, J., & Baum, S. D. (2021)[6]. In this paper the authors have explored the role of AI certification for meeting the ethical challenges raised by AI developers and the external stakeholders the certifications in general a considered as a government school that ensure compliance to standards for transparency and accountability. The authors have reviewed existing air certificate programs giving important information about the limitations and suggest ways and means for effective certificate that aligns with societal and ethical objectives. The authors have reviewed and analysed existing as well as proposed AI certificate programs issued by IEEE maltas AI and European commission's notification for analysis thematic analysis of these certification frameworks along with their effectiveness have been done.

[4] Peters, D., Vold, K., Robinson, D., & Calvo, R. A. (2020). Responsible AI—two frameworks for ethical design practice. *IEEE Transactions on Technology and Society*, *1*(1), 34-47.

[5] Ryan, M., & Stahl, B. C. (2020). Artificial intelligence ethics guidelines for developers and users: clarifying their content and normative implications. *Journal of Information, Communication and Ethics in Society*, *19*(1), 61-86.

[6] Cihon, P., Kleinaltenkamp, M. J., Schuett, J., & Baum, S. D. (2021). AI certification: Advancing ethical practice by reducing information asymmetries. *IEEE Transactions on Technology and Society*, *2*(4), 200-209.

Henriksen, A., Enni, S., & Bechmann, A. (2021, July)[7].In this paper the authors have tried to examine how the accountability in ai is understood and practiced by the developers in several Scandinavian AI companies the paper focuses on three mechanisms that is ethical principles the certification standards and the explanation methods which reviews the limitations and the concerns of the developers the study has used ethnographic method for data analysis and through this method has tried to identify the gap between the policy level account abilities with their practical implications and implementations. The paper reviews that developers find it difficult and field misaligned with the guidelines and the real-world needs of accountability in ai. The ethnographic case study was conducted between 2018 to 2020 with more than 20 semi structured interviews being conducted among the AI developers the managers and the data scientist in the study the independent variables ethical principles certification standards and the explanation method used by the developers the dependent variables however included to the account ability measures by the developers and the perceived utility of this mechanism. The important findings of the paper include that developers find it difficult and challenging for adopting the ethical guidelines as they are two abstract and missilened with the practical AI operations also please policies sometimes are even prove to be harmful for business operations in the context of certification standards there are lack of guidance on implementing the several ISO standards for as specific applications creating frustration among the developers are there is a lot of gap between the principles and their practical application within the given current frameworks

Zhang, B., Anderljung, M., Kahn, L., Dreksler, N., Horowitz, M. C., & Dafoe, A. (2021)[8]. In this paper investigation has been made by the authors regarding the ethical and governance related challenges in artificial intelligence by the machine learning researchers the study has been conducted based on the responses from 524 researchers who have made their publications in leading AI conferences. The objectives of the paper are to understand the AI/ML researchers' perspectives on the ethical challenges and governance priorities and also how much they trust their Institute or organisation for Ethical AI development. This paper also explores their attitude towards AI safety and the likelihood of involving stakeholders and policy makers for effective AI governance strategies. The important independent variables included in the paper are the type of organisation they work with, the AI governance and safety frameworks they use and the challenges they say while implementing. The dependent variables are researchers' trust levels and their attitude towards prioritizing AI safety. For data analysis, a regression model has been used to find associations between the demographic characteristics and the responses. Findings of the study suggest that these practitioners prioritise data privacy safety of autonomous systems and prevention of mass surveillance practices also they give these variables higher priority as compared to the general public. The AI practitioners have high trust in international organisations and their policies such as UN and EU they also have moderate crust in the organisations and have low trust in the military and Chinese based tech companies. Majority of the respondents conveyed that AI safety was of top most priority to them.

Hedayati-Mehdiabadi, A. (2022)[9]. The ongoing impact of emerging technologies on shaping human lives underscores the need to prioritize the ethical development of the future computing workforce. To enhance ethical

[7] Henriksen, A., Enni, S., & Bechmann, A. (2021, July). Situated accountability: Ethical principles, certification standards, and explanation methods in applied AI. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 574-585)

[8] Zhang, B., Anderljung, M., Kahn, L., Dreksler, N., Horowitz, M. C., & Dafoe, A. (2021). Ethics and governance of artificial intelligence: Evidence from a survey of machine learning researchers. *Journal of Artificial Intelligence Research*, *71*, 591-666.

[9] Hedayati-Mehdiabadi, A. (2022). How do computer science students make decisions in ethical situations? Implications for teaching computing ethics based on a grounded theory study. *ACM Transactions on Computing Education (TOCE)*, *22*(3), 1-24.

decision-making among computer science students, it is crucial to understand how they navigate ethical dilemmas. This article advances research and practice in computer ethics education by identifying key factors influencing the ethical decision-making of computer science students and offering actionable insights for improvement. Employing a constructivist grounded theory approach, the study analysed data from students' discussion posts on three ethical scenarios in computer science and follow-up interviews. The analysis revealed that connecting scenarios to real-life stories, reflecting on responsibilities associated with technical expertise, demonstrating concern for users or affected parties, and recognizing fallacies fostered better ethical decision-making. Conversely, susceptibility to fallacies and excessive empathy for developers hindered the process. The study proposes a model outlining the factors influencing ethical decision-making in computer science students and provides recommendations for researchers and educators in computer ethics.

Vanhée, L., & Borit, M. (2022)[10]. This paper has introduced the concept of GEDAI that is growing ethical designers of artificial intelligence. This paper emphasis importance of virtues and cultivating the ethical behaviours among the AI developers through integrative and interdisciplinary educational strategies the authors have proposed that current ethical training is in sufficient and what you based ethics approach be integrated along with model psychology pedagogy and technical Education to embed ethical behaviour as a transferable skill amongst the AR developers. The objective of this paper is to breach the gap between the abstract ethical principles related to Ai whereas is practical application this paper also emphasizes for string a habitual ethical behaviour by the practitioners which would support the responsible AI development. Authors have propagated a framework development through 4 component models of moral behaviour emphasizing sensitivity judgment motivation and action. The authors suggest that the ethical behaviour is framed as a practical skill which is adaptable across the professional contacts and would also encourage self-responsibility and moral reasoning in Ai designs however there are short and challenges with respect to current AI ethics curriculum and difficulty and assessing the ethical growth quantity. The paper provides recommendations in the line of developing ethics-based AI framework for promotion of inter disciplinary collaborations between educators in ethics AI and establishment of clear assessment tools for tracking ethical skill development

Cox, A. (2023)[11]. In the paper the author have identified eight scenarios where the AI professionals need to understand the issues related to AI ethics. It further discussed the ethical issues that are faced by information technology professionals. Professional ethics guides at the time of any ethical doubts and dilemmas. AI has posed several ethical issues concerning around data and its misuse. The authors suggest that the IT professionals should not depend on their vendors for ethical products rather they themselves should apply their ethical principles independently, and they themselves should know clearly what ethical issues can arise of the AI systems. There are chief principles identified for IT and AI professionals like Confidentiality and privacy, Professional development, Integrity, Conflict of interest and personal gain, Free and equal access to information, Censorship, Responsibilities to profession, consent and transparency. The AI systems are much impacted by the Professional ethics of AI professionals and also by the big corporates which are driven by commercial logics, also AI is at nascient stage so it is easier to mold and use/reject in accordance to the corporate motives. Hence If AI has to be used for social beneficiance, it is important it is positively driven by the two ends.

[10] Vanhée, L., & Borit, M. (2022). Ethical by designer-how to grow ethical designers of artificial intelligence. *Journal of Artificial Intelligence Research*, *73*, 619-631.

[11] Cox, A. (2023). How artificial intelligence might change academic library work: Applying the competencies literature and the theory of the professions. *Journal of the Association for Information Science and Technology*, *74*(3), 367-380.

Khan, A. A., Akbar, M. A., Fahmideh, M., Liang, P., Waseem, M., Ahmad, A., ... & Abrahamsson, P. (2023)[12]. In the study about 99 ai professionals and lawmakers from 20 different nations are surveyed to explore their ideas and perception for AI ethics principle and challenges in the study principles just transparency accountability privacy are considered as most important AI ethics principle however the study reviews that there is dearth ethical knowledge lack of legal governors frameworks and infective and insufficient monitoring bodies as major challenges. The study tries to bridge a gap between ethics theories and practices in Ai systems for data analysis thematic mapping of responses has been done and statistical analysis is done to identify correlation and differences in perception. Findings reveal that transparency accountability and privacy are the most important principles as transparency ensures openness, understand ability of decision making for the stakeholders however lack of ethical knowledge and absence of legal frameworks or significant barriers also the study reviews that law makers emphasize fairness and regulatory measures more than the AI practitioners whereas ai professionals give more weightage to transparency and accountability the paper also gives recommendations for organisations to develop structure training programs in order to meet the ethical knowledge gap also to establish monitoring bodies and boards to ensure the others of ethical principles in all levels for policy makers the researchers have suggested to create more comprehensive legal framework that collaborate with industry and alliance with ethics and practical realities

Li, B., Qi, P., Liu, B., Di, S., Liu, J., Pei, J., ... & Zhou, B. (2023)[13]. In this research paper the key dimensions of trust for the Ness for AI have been identified such as robustness fairness explainable at transparency privacy preservations and accountability the authors have tried to give a comprehensive view and practical guide for building trust for the AI to these principles and dimensions the authors for this propose implementation of systematic life cycle framework covering several stages of data acquisition deployment and governance. This paper presence practical guides for creating trustworthy AI through a manage life cycle process. There are certain challenges that have been identified such as finding a trade of between including One dimension with others impacting the Ai system also have widely expressed that there is a lack of standardisation and need of multi discipline in collaborations for bringing trust in the Ai systems

Rismani, S., & Moon, A. (2023, August)[14]. This paper explorers the emerging importance of responsible AI practitioners and also discusses the skills qualifications and responsibilities required for filling such positions the authors in the paper have analysed job postings and also conducted several expert interviews in order to create the ontology for the roles of technical researchers and data scientist to policy analyst and executive managers. The authors have highlighted then there is a need for structured educational support along with the organisation support to build the capacity for responsible AI the objective of the paper is to provide the competency framework that inform education HR policies and organisation. The authors have identified 79 job postings and reviewed them with the perspective of identifying rolls related to Ai ethics responsible AI and similar areas, further 14 interviews were conducted with the practitioners who are involved and responsible AI profiles in the diverse industries for data analysis thematic coding of job description and job interviews were done in order to identify key responsibilities required skills qualifications and interpersonal skills. Important findings suggest that these practitioners face barriers of integrating ethics into their AI work flows due to organisational barriers and resistance also there are lack of standardization in job description and skills expectations. Authors have proposed a framework which includes specific skills and qualification for each roll and focus on aspect such as adaptability collaboration and passion for Ethical AI. The recommendations for this paper are for organisations for educators and for the practitioners for shaping responsible AI practices.

[12] Khan, A. A., Akbar, M. A., Fahmideh, M., Liang, P., Waseem, M., Ahmad, A., ... & Abrahamsson, P. (2023). AI ethics: an empirical study on the views of practitioners and lawmakers. *IEEE Transactions on Computational Social Systems*, *10*(6), 2971-2984.

[13] Li, B., Qi, P., Liu, B., Di, S., Liu, J., Pei, J., ... & Zhou, B. (2023). Trustworthy AI: From principles to practices. *ACM Computing Surveys*, *55*(9), 1-46

[14] Rismani, S., & Moon, A. (2023, August). What does it mean to be a responsible AI practitioner: An ontology of roles and skills. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 584-595).

Tribelhorn, B., & Nuxoll, A. (2023)[15].This paper focuses to teach ethical principles related to ai and computing to both stem and non-stem students so that it can encourage enter disciplinary approach for addressing global issues using AI also this would improve the students analytical and communication skills to projects and discussions. In this paper the authors focus on including the AI concepts ethics and real-world applications into the course design for the students both of non-stem and stem also the use of simplified AI algorithms that do not require prior programming knowledge should be induced to the non-stem students. The key findings for this paper include that there is a positive reception to air applications and ethical discussions among the students which also means that the students are interested to understand the ethical implications and ethics for AI. One of the important findings of this particular paper is that the students struggled more with the conceptual problem solving than understanding the AI applications this paper suggest that there is an important need for inducing ethics-based curriculum to the AI learning students.

Baldassarre, M. T., Gigante, D., Kalinowski, M., Ragone, A., & Tibidò, S. (2024, June)[16]. In this paper the authors have tried to investigate the challenges the AI practitioners face while implementing trustworthy AI, through the software development life cycle or sdlc. The paper identifies the gaps between the current practices and the tools available for creating trust for the AI, it also elaborates on the business conference and limited focus on the TAI principles which include aspect such as human agency and oversight technical robustness and safety, privacy and data governance, transparency, diversity, not discrimination, fairness, society and environment well being and accountability these are required to ensure that air systems respect the fundamental rights privacy become transparent and avoid biosis take hold of participation also bringing society. In the paper the authors have conducted service and interviews among the practitioners and highlights the need for better tools guidelines and mechanism for bringing trust for the Ness to the AI systems with the analysis the following key findings are explode that is that trust for the air principles are mostly used during the design and development phase whereas little attention has been paid during the deployment and post deployment monitoring stage. Majority of practitioners believe that improving data set quality and using algorithms would increase explain ability whereas using automated tools would be least beneficial for addressing trust and AI issues. Major business constraints like time cost and performance leads to deviation from following TAI issues. Professionals strongly believe that privacy and transparency are most emphasize principles as a part of regulated requirement and fairness is under priorities does the paper recommends that throughout sdlc post-employment and monitoring phases needs emphasis for the tools to be created and business-related constraint should be mitigated to bridge the gap between principals and actions

Hodent, C., Blumberg, F., & Deterding, S. (2024)[17]. In this paper the authors at talked about developing the guidelines to safeguard the game developers with emphasis on creation of comprehensive code of ethics for the industry generally it has been seen that the game developers face a lot of harassment and jobs in security they also create the content which leads to gaming addiction and the deceptive monetisation practices among the players thus it is important that a unified global code of ethics be created to address the issues comprehensively

[15] Tribelhorn, B., & Nuxoll, A. (2023). A Course Model for Ethics Education in Computer Science. *Journal of Computing Sciences in Colleges*, *39*(1), 48-55.

[16] Baldassarre, M. T., Gigante, D., Kalinowski, M., Ragone, A., & Tibidò, S. (2024, June). Trustworthy AI in practice: an analysis of practitioners' needs and challenges. In *Proceedings of the 28th International Conference on Evaluation and Assessment in Software Engineering* (pp. 293-302).

[17] Hodent, C., Blumberg, F., & Deterding, S. (2024). Ethical Games: Toward Evidence-Based Guidance for Safeguarding Players and Developers. *ACM Games: Research and Practice*, *2*(2), 1-11.

Lancaster, C. M., Schulenberg, K., Flathmann, C., McNeese, N. J., & Freeman, G. (2024)[18]. In this paper the authors have explored professionals perception of accountability in mitigating the AI bases, in the paper the authors explain that despite the digital and technical knowledge professionals often this engage morally and transfer the countability to the organisational structures or any other external factors the paper I went to 5 individual and organisational barriers to ethical AI development and offers guidelines for improving the collective accountability the paper has several objectives such as to understand the AI professionals perception of a virus and their accountability for bias mitigation also to explain explore how the organisation and external factors shape these perception and limit actions to provide high level guidelines to enhance the under which will an organisation accountability for the responsible air development in this paper there are independent and dependent variables identified independent variables include professionals technical knowledge and their understanding of AI bias in the system also the organisation policy is culture and power dynamics plain important role as an independent variable also the client and external pressures for important independent variables in the study the dependent variables include professional sense of accountability adoption of AI mitigation practices usability of tools to identify and mitigate Pisces and also the organisation commitment to responsible air development in this paper is level and from different organisation types. The data is collected through snowball sampling method reflecting US based participation data analysis has been done through thematic coding inspired by ground theory methods supported by social technical system and moral disengagement frameworks analysis is also done of individuals and organisational factors which mediate the account ability the paper has following he findings there are individual level barriers like where the professionals often you buyers as inevitable or outside their control leading to model dyse engagement the knowledge does not always translate to actions due to perceived limitations in their roles and responsibilities and also the objectivity is often over size side lining the ethical concentrations there are short and organizational barriers also like the power dynamics and articles structures which inhibit the junior employees from raising and voiceing their concerns regarding the ethical risk and also the organisation give importance to the clients demands and the financials over any of the Other ethical air practices reinforcing the model disegagement organisational values and public private contradictions influence the actions at the organisational level in the paper certain external pressures have also be in identified like the client requirement and resource conference limiting the proactive bio's mitigation efforts at the professionals and organisational side the professionals feel restricted by competing priorities between ethical responsibilities and client satisfaction the paper in the end give the following recommendations like we need to empower professionals across their roles and responsibilities by giving them clear account ability frameworks and the ethics training also the paper emphasizes fostering a culture of responsible and collaborative organisations that would initiate to mitigate hierarchical barriers the paper emphasizes that organisation values need to be aligned with the actionable practices that would priority is the ethical AI development in the end the paper encourages the external stakeholders to support the value responsible AI practices conducted by the professionals and AI organisations

Lassiter, T. B., & Fleischmann, K. R. (2024)[19] This paper reports findings from interviews with 19 AI auditing stakeholders to understand how AI auditing professionals seek to create calibrated trust in AI tools and AI audits. The key objective for this paper are to analyse how the AI audits influence trust in ai systems and audits also this paper explores professionals prospective on trust and trust worthiness in context to ai audits also this paper identifies the barriers and the best practices that can be included for creating impact full orders this paper for the exams the current state of AI auditing standards and how they affect on trust building the important research questions include how the AI auditing professionals view AI ordered role in trust building and this paper the important variables are audit frameworks methodology expertise and accreditation of the auditors and transparency and independence and the audit practices the dependent variables are the increased users trust effectiveness of audits and enhance public and organisational person of the audit outcomes the study has been

[18] Lancaster, C. M., Schulenberg, K., Flathmann, C., McNeese, N. J., & Freeman, G. (2024). “It’s Everybody’s Role to Speak Up... But Not Everyone Will”: Understanding AI Professionals’ Perceptions of Accountability for AI Bias Mitigation. *ACM Journal on Responsible Computing*, *1*(1), 1-30.

[19] Lassiter, T. B., & Fleischmann, K. R. (2024). " Something Fast and Cheap" or" A Core Element of Building Trust"?-AI Auditing Professionals' Perspectives on Trust in AI. *Proceedings of the ACM on Human-Computer Interaction*, *8*(CSCW2), 1-22

conducted with the use of thematic exploration of audit methods trust perception and challenges for data analysis thematic coding using Atlas.ti applying Braun and Clarks reflective thematic analysis method is used also authors have used comparative analysis of existing literature on AI audits with respect to trust building s include the following the AI audit industry is fragmented and lacks universal standards or definition leading to differentiated practices on auditors independence processes and good communication for order results to be authentic transparency are important elements for building the trust. Client needs audits of AI systems to ensure that ethical concerns and risks are effectively managed however and important finding that there is a need of uniform regulations when it comes to creating audit and enhanced public trust.

Lu, Q., Zhu, L., Xu, X., Whittle, J., Zowghi, D., & Jacquet, A. (2024)[20]. In this paper the authors have introduced responsible AI pattern catalogue which is a comprehensive framework providing guidelines to operationalise responsible AI throughout the AI life cycle. It provides patterns for governance processes and design. The authors have tried to bridge the gap between high level ethical principles and their practical implementation thus it addresses the gap in the current practices and offers practical solutions for Ethical risks in governance engineering and operational context so that AI systems can be made trustworthy reliable and are able to engage stakeholders. The data collection method use is multi vocal literature review which combines academic literature and grey literature for identifying suitable solutions the literature has been gathered from 25 academic items and 69 grey items focusing around the standards guidelines and case studies. The authors have recommended the implementation of multi-level governance frameworks which should specially be designed for industry and organisations use of standardised reporting and role base accountability for Ethical compliance. They also suggest developing ethical training programs among the stakeholder and continuously monitoring and adapting patterns to address emerging risk in ai systems

Pant, A., Hoda, R., Spiegler, S. V., Tantithamthavorn, C., & Turhan, B. (2024)[21] In this paper the authors have investigator the AI practitioners awareness about AI ethics and what challenges the generally face when trying to other to ethical principles in the development phase in this paper a survey of 100 AI practitioners has been done trying to identify the key barriers and giving suggestion about how ethical practices can be enhance at development phase. The independent variables in the study include formal education workplace policies and experience levels and the dependent variable at the awareness levels about ai ethics challenges as seen while deployment and development and adherence to the ethical principles. The paper uses mix method approach using quantitative descriptive statistics and so show technical ground theory for qualitative data analysis. The important findings of the study are around 41% of the AI practitioners reported they they were reasonable familiar with the AI ethics according to them privacy protection and security for recognised as the most important ethical principle also formal education aided ethical preparation but wasn't highly impactful. The AI practitioners face challenges in balancing the ethical principles, har time constraints, had resource limitations also they lacked high quality unbiased training data and faced challenges like human bias lack of foresight and also defining the universal ethical standards they can include. Their organisation also had no consensus regarding the ethical principles across the regions and human biosis within the organisation frequently transfer into the AI systems both intentionally and unintentionally. Finally the authors recommended the need for formal education and training for AI ethics practices developing standardised global guidelines and principles for ai ethics and the need for promoting interdisciplinary collaborations for addressing the ethical challenges effectively.

[20] Lu, Q., Zhu, L., Xu, X., Whittle, J., Zowghi, D., & Jacquet, A. (2024). Responsible AI pattern catalogue: A collection of best practices for AI governance and engineering. *ACM Computing Surveys*, *56*(7), 1-35.

[21] Pant, A., Hoda, R., Spiegler, S. V., Tantithamthavorn, C., & Turhan, B. (2024). Ethics in the age of AI: an analysis of ai practitioners' awareness and challenges. *ACM Transactions on Software Engineering and Methodology*, *33*(3), 1-35.

Popova, K., Figueras, C., Höök, K., & Lampinen, A. (2024)[22]. The authors in this paper have investigator how the technology practitioner attribute responsibility in their work it has been seen that they identify ethical stances at 3 levels the I stance, We stance and They stance for acting on the ethical issues the authors have conducted interviews with 23 practitioners in Scandinavia. The resort reviews that developers feel optional responsibility and a burden when they feel that they need to act on the ethical issue whereas when collectively taking a responsibility in teams generally it is a way responsibility and they often try to distance the responsibility to others through they stand does there is always a tension between vulnerability and distance and offering insights into the emotional dimension of ethical decision making. Through the research following findings reveal that the I stance or the individual responsibility comes with high emotional stakes and pressures of shame and guilt whereas the We stance observes share responsibility within the teams and the organisations which reduces the burden however it often lacks the clarity on concrete actionable steps under the they stance responsibility is outsource to the external entities such as clients and regulators there by creating an emotional distance to the risks of ethical oversight. The paper brings are recommendation that there needs to be a systematic mechanism for distributing the ethical responsibility within the teams and the organisations it also suggest that there has to be balance between the individual and the collective responsibilities in order to mitigate the emotional pressure and borders with in the individual practicioners. Author suggest that there is height and need for bringing ethical awareness among the technology development and deployment teams and also they need to be provided with training and support in order to navigate between the ethical complexities.

Reinmund, T., Salvini, P., Kunze, L., Jirotka, M., & Winfield, A. F. (2024).[23] This paper explorers variable autonomy in robotics which allows the robots to dynamically adjust the levels of autonomy needed based on the social and environmental context in this paper 42 research articles have been identified for developing a unified definition for variable autonomy so as to establish design guidelines for a responsible robotic applications in the paper ethical considerations with respect to autonomy levels human robot interaction and robotic capabilities have been assessed. The key findings of this paper includes coming up with a unified definition of variable autonomy the four dimensions of variable autonomy like who is the initiator of the change what aspects of autonomy shall be adjusted and when the adjustment occurs and why the changes in the tournament levels are triggered several design guidelines like account ability and transparency stakeholder engagement adaptability are important. Several challenges have also be an identified such as misalignment of theoretical frameworks and the practical applications and need for global standards and defining and evaluating autonomy levels

Discussions & Findings :

The collective findings of the above mention studies reveal landscape which illustrates ethical barriers and challenges relating to AI with their dimensions ranging between technical, organisational and societal in nature. The above researches explorer that the AI professionals have a consensus on the importance relating to AI ethics principles like transparency, accountability, fairness, robustness and justice etc. but putting them for practical application, is a difficult task. The challenges come from guidelines which are ambiguous in nature, limitations of resources and also organisations resistance to change post major barriers. Organisations also think that time and cost implications for incorporating AI ethics make it difficult for them and professionals to incorporate ethics in practical applications. Most of the researchers believe that there is a greater need for interdisciplinary collaborations, ethics trainings and stakeholder engagement as important tools for in depth alignment of ethical

[22] Popova, K., Figueras, C., Höök, K., & Lampinen, A. (2024). Who Should Act? Distancing and Vulnerability in Technology Practitioners' Accounts of Ethical Responsibility. *Proceedings of the ACM on Human-Computer Interaction*, *8*(CSCW1), 1-27.

[23] Reinmund, T., Salvini, P., Kunze, L., Jirotka, M., & Winfield, A. F. (2024). Variable Autonomy through Responsible Robotics: Design Guidelines and Research Agenda. *ACM Transactions on Human-Robot Interaction*, *13*(1), 1-36.

AI principles, further these should be intertwined with the organisational goals to foster trust and long-term society benefits.

Key findings of the studies reveal that there is gap in operationalising the ethical principles which are abstract and offer limited guidance with their practical applications. Also the current tools and frameworks available needs to be more adaptive and context specific. Stakeholders have important role as their collaboration with developers auditors policy makers and end users would reduce the risks of leaving out marginalising groups based on economics social and gender related criteria. The literature provides context to the challenges that have been related to resource constraints conflicting ethical priority and organisations lack of interest and commitment leading to ethics washing. Findings reveal that ethical AI practices fosters stakeholder or trust mitigates the legal risk and enhances long term organisational resilience.

Recommendations :

Based on the studies we recommend that the organisations should established dedicated ethics committees and boards that would integrate ethics reviews into their existing work flows also they need to invest in ethics training programs to enhance air professionals awareness and their capabilities in ethical AI. For policy makers we recommend that enforceable standards needs to be developed and incentives be given out to the practitioners an organisations favouring ethical AIR development also promotion of multi disciplinary collaborations for refinement and validation of ethical frameworks needs to be done. Developers are recommended to make a use of context sensitive AI that can detect ethical challenges. They need to prioritise transparency accountability privacy and human centre City within the design to build trust and inclusivity in the systems. For Academia suggested that for the empirical studies be conducted to validate the effectiveness of ethical guidelines principles and frameworks across divorce contacts. Academia also needs to collaborate with industry for developing scalable tools for operationalization of AI Ethics. Also efforts be made for bringing awareness amongst the public regarding ethical and societal implications of AI ethics.

Conclusion :

The reviewed articles explore various dimensions of AI ethics, focusing on operationalizing ethical principles, addressing challenges, and integrating ethics into AI development and business practices. Key themes include the gap between high-level principles and actionable implementation, the role of stakeholders in fostering responsible AI, the financial and organizational costs of ethical compliance, and frameworks for ensuring fairness, accountability, transparency, and sustainability. The studies emphasize the need for systemic approaches, collaboration, and practical tools to address challenges like bias, resource constraints, and regulatory pressures.

Reference :

Peters, D., Vold, K., Robinson, D., & Calvo, R. A. (2020). Responsible AI—two frameworks for ethical design practice. *IEEE Transactions on Technology and Society*, *1*(1), 34-47.

Ryan, M., & Stahl, B. C. (2020). Artificial intelligence ethics guidelines for developers and users: clarifying their content and normative implications. *Journal of Information, Communication and Ethics in Society*, *19*(1), 61-86.

Cihon, P., Kleinaltenkamp, M. J., Schuett, J., & Baum, S. D. (2021). AI certification: Advancing ethical practice by reducing information asymmetries. *IEEE Transactions on Technology and Society*, *2*(4), 200-209.

Henriksen, A., Enni, S., & Bechmann, A. (2021, July). Situated accountability: Ethical principles, certification standards, and explanation methods in applied AI. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 574-585)

Zhang, B., Anderljung, M., Kahn, L., Dreksler, N., Horowitz, M. C., & Dafoe, A. (2021). Ethics and governance of artificial intelligence: Evidence from a survey of machine learning researchers. *Journal of Artificial Intelligence Research*, *71*, 591-666

Hedayati-Mehdiabadi, A. (2022). How do computer science students make decisions in ethical situations? Implications for teaching computing ethics based on a grounded theory study. *ACM Transactions on Computing Education (TOCE)*, *22*(3), 1-24.

Vanhée, L., & Borit, M. (2022). Ethical by designer-how to grow ethical designers of artificial intelligence. *Journal of Artificial Intelligence Research*, *73*, 619-631.

Cox, A. (2023). How artificial intelligence might change academic library work: Applying the competencies literature and the theory of the professions. *Journal of the Association for Information Science and Technology*, *74*(3), 367-380.

Khan, A. A., Akbar, M. A., Fahmideh, M., Liang, P., Waseem, M., Ahmad, A., ... & Abrahamsson, P. (2023). AI ethics: an empirical study on the views of practitioners and lawmakers. *IEEE Transactions on Computational Social Systems*, *10*(6), 2971-2984.

Li, B., Qi, P., Liu, B., Di, S., Liu, J., Pei, J., ... & Zhou, B. (2023). Trustworthy AI: From principles to practices. *ACM Computing Surveys*, *55*(9), 1-46

Rismani, S., & Moon, A. (2023, August). What does it mean to be a responsible AI practitioner: An ontology of roles and skills. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 584-595).

Tribelhorn, B., & Nuxoll, A. (2023). A Course Model for Ethics Education in Computer Science. *Journal of Computing Sciences in Colleges*, *39*(1), 48-55.

Baldassarre, M. T., Gigante, D., Kalinowski, M., Ragone, A., & Tibidò, S. (2024, June). Trustworthy AI in practice: an analysis of practitioners' needs and challenges. In *Proceedings of the 28th International Conference on Evaluation and Assessment in Software Engineering* (pp. 293-302).

Hodent, C., Blumberg, F., & Deterding, S. (2024). Ethical Games: Toward Evidence-Based Guidance for Safeguarding Players and Developers. *ACM Games: Research and Practice*, *2*(2), 1-11.

Lancaster, C. M., Schulenberg, K., Flathmann, C., McNeese, N. J., & Freeman, G. (2024). “It’s Everybody’s Role to Speak Up... But Not Everyone Will”: Understanding AI Professionals’ Perceptions of Accountability for AI Bias Mitigation. *ACM Journal on Responsible Computing*, *1*(1), 1-30.

Lassiter, T. B., & Fleischmann, K. R. (2024). " Something Fast and Cheap" or" A Core Element of Building Trust"?-AI Auditing Professionals' Perspectives on Trust in AI. *Proceedings of the ACM on Human-Computer Interaction*, *8*(CSCW2), 1-22

Lu, Q., Zhu, L., Xu, X., Whittle, J., Zowghi, D., & Jacquet, A. (2024). Responsible AI pattern catalogue: A collection of best practices for AI governance and engineering. *ACM Computing Surveys*, *56*(7), 1-35.

Pant, A., Hoda, R., Spiegler, S. V., Tantithamthavorn, C., & Turhan, B. (2024). Ethics in the age of AI: an analysis of ai practitioners’ awareness and challenges. *ACM Transactions on Software Engineering and Methodology*, *33*(3), 1-35.

Popova, K., Figueras, C., Höök, K., & Lampinen, A. (2024). Who Should Act? Distancing and Vulnerability in Technology Practitioners' Accounts of Ethical Responsibility. *Proceedings of the ACM on Human-Computer Interaction*, *8*(CSCW1), 1-27.

Reinmund, T., Salvini, P., Kunze, L., Jirotka, M., & Winfield, A. F. (2024). Variable Autonomy through Responsible Robotics: Design Guidelines and Research Agenda. *ACM Transactions on Human-Robot Interaction*, *13*(1), 1-36.

[i]https://standards.ieee.org/ieee/7999/11740/?mkt_tok=MjExLUZZTC05NTUAAAGXi6h75qftV1ElldZd5DoWhFVZnP0K1l6dY5qNdTxaV8cPs3eAf6WCmQFasHwNIM0ZYwZkRcRrseLJftlqNaaxuem7M3vZmXEpI63Y3Mm-NJk