AI Accountability: Approaches, Affecting Factors, and Challenges

Ashwin Kumar Raja
Birla Institute of Technology & Science Pilani, Hyderabad Campus, India
Email: f20170599@hyderabad.bits-pilani.ac.in

Jianlong Zhou*
Data Science Institute, University of Technology Sydney, Australia
Email: jianlong.zhou@uts.edu.au
* Corresponding author

***Abstract*—Accountability is one of core ethical principles for AI systems. It is often cited as a must-have property especially in high-stake applications. This survey paper investigates AI accountability from different key perspectives for its effective operationalisation. Major approaches for achieving accountability are reviewed and typical factors affecting accountability are identified. It is argued that accountability needs to be considered at each phase of the AI life cycle. Accountability in three typical sectors and other application areas which have already addressed accountability are exemplified to motivate approaches for achieving AI accountability. Challenges and obstacles are finally identified for future work directions.**

Index Terms: Accountability, Artificial Intelligence, AI Ethics

ARTIFICIAL Intelligence (AI) and Machine Learning (ML) algorithms are increasingly used in various domains such as banking, insurance, medical care, criminal justice, predictive policing, and hiring by making decisions with legal and ethical impacts. The new Worldwide Artificial Intelligence Spending Guide from International Data Corporation (IDC) forecasts global spending on AI systems will jump from $85.3 billion in 2021 to more than $204 billion in 2025. Compared with humans, AI-informed decision-making can lead to faster and better decision outcomes. However, because of the complexity and black-box nature of AI models, it is often hard for users and even their designers to understand how the data is processed by AI for specific decision-making. Therefore, the deployment of AI algorithms especially in high stake domains usually requires testing and verification for reasonability and reliability as well as other concerns such as fairness by domain experts not only for safety but also for legal reasons. These concerns have led to AI ethical principles which have been intensively discussed during the last several years. Despite the large number and variety of ethical guidelines for AI, accountability forms one of the core AI ethical principles which are fairness, privacy and security, transparency, and accountability as well as others. Therefore, accountability has been increasingly investigated in recent years as shown in Figure 1. It is often cited as a must-have ethical principle, but the reasons for this, as well as what it entails, are rarely explained [1].

This paper aims to deliver a focused review of AI accountability in different key aspects. The concept of accountability is first introduced followed by approaches for achieving accountability. Different factors affecting AI accountability are then reviewed. Accountability is also investigated from the perspectives of different sectors (e.g. government sector, private sector, and academia) and different levels (e.g. data, algorithm, and developer) respectively. Accountability in other areas such as cloud and network technology, medicine and clinical trials, as well as law and policy-making, is also reviewed.

We searched key databases such as IEEE xplorer, ACM Digital Library, Springer as well as Google with key words of "AI accountability" and "AI ethics" to collect literature for analysis.

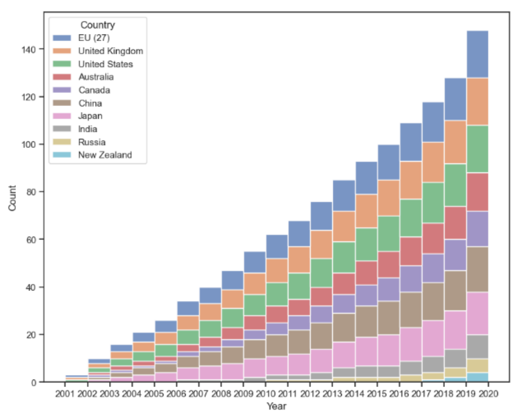

Figure 1. Cumulative publications in accountability in AI from 2000-2020. Source of Data: Microsoft Academic Graph

What is accountability?

Accountability delegates responsibility to appropriate agencies for actions, decisions, products, and policies. It helps keep social order and run society. Accountability is relational, as it depends on a multitude of factors. Accountability's definition generally includes three main elements: responsibility, answerability, and sanctionability. It includes taking responsibility for one's acts and decisions. It deals with answerability—the capacity to reveal the reasons for decisions. Also, sanctionability and liability to the actions [1].

Why do we need accountability?

The biggest question on people's minds is: "Who is accountable for the AI systems?" Almost any government or organisational ethics guidelines include accountability. It helps build trust among the stakeholders in AI systems, and resolve conflicts. It paves the way to realising the complete potential of AI systems. AI has spread its wings in many domains, and accountability has been considered as a should-have rather than a mere desire. An AI system consisting of safe components may still be risky, and vice versa. Diverse teams unify to build an AI system. Thus, to make it trustworthy, accountability serves as a fulcrum. Also, explicability incorporates accountability as its major component. It is thus important to discuss accountability and liability for AI.

Where does accountability fall in the AI life cycle?

Integrating ethical testing into the development life cycle of AI is the first step toward responsible AI development. Software release cycles already include defined phases such as unit testing, load testing, and user testing. There must be bias, transparency, predictability testing, and many new and undiscovered testing in the future. An AI life cycle consists of multiple stages as shown in Figure 2. There are three parties that hold their percentage of shares in the accountability chart, namely the data, the developer and the model. It is essential to understand the decision making process and steps followed to make accountable AI systems.

Accountability should be treated as a requirement during the whole AI life cycle. For example, in the design phase, it is important to lay out all the underlying assumptions and clear deliverables. There must be no bias or ethical concerns associated with them. Accountability plays a vital role in the development phase. We have elaborated on them in the following sections on data and model accountability. It involves validating and testing the system from end to end. The deployment phase involves ensuring regulatory compliance and evaluating the user experience. It must be fair and reliable. There must be clear

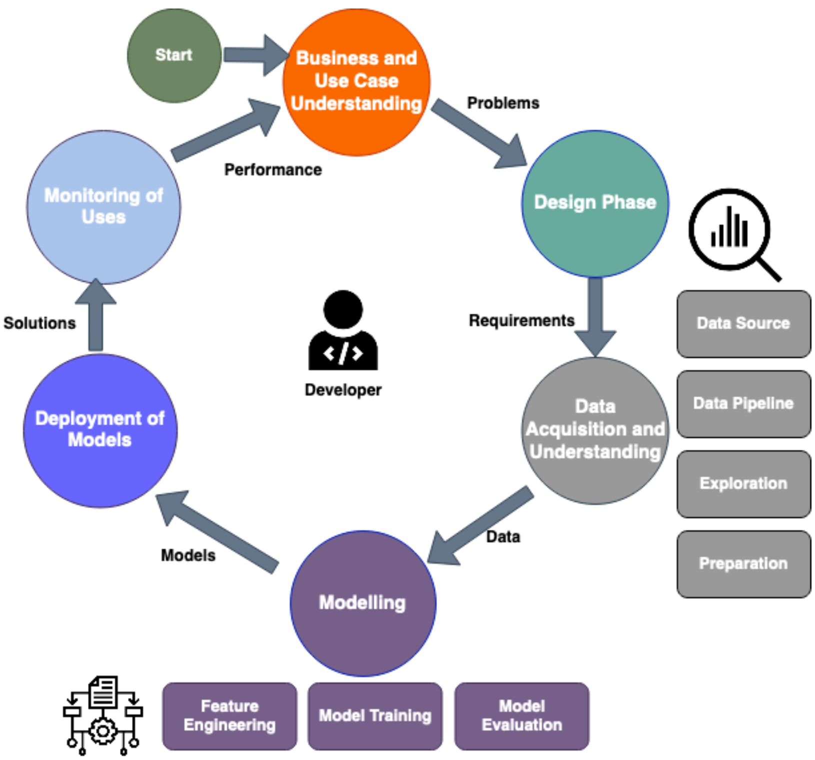

Figure 2. Illustration of the AI life cycle.

fallback procedures. In the monitoring phase, there is impact assessment and conflict resolution. Accountability is thus observed in various forms at various stages.

Approaches for Achieving Accountability

There has been substantial research trying to imbibe accountability in AI systems. From government policies to redesign organizational structures and guidelines concerning the model and the system itself and finally, from the point of view of data, accountability, in turn, influences the assessment of such a system and effective decision making. One can enforce them in multiple ways. Singh et al. [2] propose decision provenance as a critical element of this accountability 'jigsaw' with many possibilities for dealing with concerns of responsibility, technical compliance, and increased user agency in systems.

Algorithmic Impact Assessments provide a framework for agencies to comprehend the automated decision systems they acquire and increase transparency into the work of automated decision systems for the public to hold them accountable. Gebru et al. [3] propose to accompany every dataset with a datasheet that details the dataset's motivation, composition, gathering procedure, recommended usage.

Raji et al. [4] present an auditing methodology that helps close the accountability gap in developing and deploying large-scale artificial intelligence systems by incorporating a comprehensive auditing procedure. Incorporating disclosure and scope into the AI system's development and deployment process fosters proactive engagement between both parties while also ensuring that subjects are made aware of the system's decisions and can challenge them. Unlike transparency, which has several unsolved issues, disclosure and scope are essentially policy-based and can be incorporated into any AI systems reasonably straightforward. Once correctly created, these standards make transparent algorithms usable, but they also educate all stakeholders on the benefits and drawbacks of AI systems. Smith et al. [5], in their literature survey, talk about the accountability of an opaque AI system when used for clinical decision making. It addresses accountability issues and the allocation of responsibility and legal liability as applied to the clinician and the technologist. They do so by a four-step process: formulate the review question and eligibility, identify all of the literature that meet the criteria, extract and synthesise pertinent data, and finally derive the results organised by themes.

Anne et al. [6] talk about how accountability can be put forward in pursuit through the developer. First, check the model for bias in various methods, preferably with various training datasets. Second, transparency has a limited role to play. Provenance and lineage play a significant role. The third is to track its impacts on social minorities and parties and check the vulnerable accordingly. Fourth, they must bear the consequences and sanction them accordingly. Better governance is a fifth mechanism for increasing accountability. Finally, consider the government's role in encouraging algorithmic accountability, Miguel et al. [7] introduce the Global-view Accountability Framework (GAF) to put accountability into practice. The GAF examines auditability and redress of contradicting information resulting from a context with two or more AI systems, which can have a detrimental influence in the case of autonomous vehicles. Researchers will have to help people navigate contentious ideals, and they will not always have the freedom to pursue the kinds of accountability and justice that they prefer. Such difficulties should be welcomed. To do so, trust between public bodies and researchers will need to be built. The accountability framework has a Goldilocks effect. A lack of one feature might lead to a deficit. An overabundance of one aspect might lead to an overload. If we want to develop a successful accountability system, we'll need to balance the scales of each part (for example, making sure the accountability is complete but not overly detailed) in order to produce workable accountability. Algorithms must be built in an empathetic and socially acceptable way which helps areas of accountability to prosper.

Factors affecting accountability

This section itemises typical widely investigated factors that affect AI accountability. These factors are: explainability, fairness, transparency, empathy, and uncertainty.

Explainability

One must be open about the trade-offs between model explainability and performance as it helps to explain the decisions and address the accountability of an issue better [8]. Explaining models is much more complex than it needs to be, and different types of explanations may not have the same trade-offs. Looking at it from a broader perspective and looking at it from different angles help trade-offs. Explainability is not a binary, on-or-off quality because it may be measured in degrees. As a result, the most transparent white-box model and the most opaque black-box model have different levels of explainability. For high-stake applications or applications that impact humans, more explainable models should be used. There are also design approaches for AI explainability where there is a proliferated set of rules for different case scenarios that finally help achieve accountability. The degree to which the model is explainable is important. It helps to quantify the answerability aspect of accountability. It helps to understand the degree to which the model is responsible. It enables accountability by providing key insights. XAI (eXplainable Artifical Intelligence) techniques can help improve explainability. This in turn improves accountability.

Fairness and transparency

There have been multiple instances where there could be unwanted bias in AI systems, resulting in fairness issues. These unwanted biases could arise due to data and the algorithm. Different approaches have been developed to mitigate various biases and obtain fairer algorithms.

High transparency is another form of soft accountability. By this, Fox [9] argues that there is an overlap of them in the case of institutional answerability. They differ in the case of access to information and compensation, which are opaque transparency and hard accountability, respectively. While academics in machine learning and AI are attempting to make their algorithms more understandable, they are not focusing on useable, practical, and effective transparency that benefits humans. Human users can be involved in the AI lifecycle to help to close the gap by ensuring that AI is built from the ground up to be understandable. One of the main factors that affect transparency is the secrecy involved in decision making and the building of software. If there is opacity, users will be deprived of transparency, leading to deprivation of accountability for the same.

Empathy

Empathy is a social characteristic and does not have a fixed definition but can be defined using a reaction to emotional states affectionately and cognitively [10].

Considering everyone as a unique individual, not everyone thinks or behaves in the same way. While empathy and compassion are not innate qualities, one can learn them. Non-active listening, discomfort with emotion, or simply being more focused on the business task at hand and less on the individuals who will execute the assignment are all factors that prevent people from being as empathic as they could be. Trust, empathy, and mutual understanding are the foundations of well-ordered and cohesive societies.

Therefore, the administrative law of procedural due process needs to be extended beyond its current concentration on eliminating errors and cutting costs because of that human trait of empathy. Research shows that people's conduct, especially loyalty, is influenced by emotional intelligence and empathy rather than information. Therefore, more empathy is poured into businesses operations. A culture of high accountability and empathy is also important for an organisation, in which management and staff devote themselves to the organization's requirements and one another. Research shows that those who score higher on empathy also have more socially responsible attitudes and a greater understanding of how their management decisions affect society. This could lead to better socially just decision-making and more social justice activities in their management.

As AI becomes increasingly common, these connections must be maintained. Human makes decisions based on various factors, including values, ethics, morality, empathy, and a sense of right and wrong – all of which AI lacks. There would be four layers to an effective process: One begins with their values. What exactly do we want AI to accomplish for us? Then one goes on to ethical concepts. What is the best way for AI to accomplish its goals? Then one can make policy recommendations and finally look at the technical controls required to put that policy into action.

AI researchers have been focusing on artificial empathy. Similarly to how societal ideals, moral codes, and social behaviour norms help humans live better in society, empahty can be integrated into AI systems to benefit AI accountability. For example, humans are now teaching the AI algorithms to be more sympathetic to people who are upset about misplaced luggage, damaged products, or their cable service being out after repeated attempts to fix it.

Uncertainty

Uncertainty is a quantitative factor that is an inevitable property for AI solutions. It helps to get the accurate prediction results with a significant probability and help discriminate between low and high confidence level predictions. There are different categories of uncertainty. Informational uncertainty, environmental uncertainty and intentional uncertainty are significant in decision-making processes and understanding individual intentions. Informational uncertainty talks about the completeness of information, its quality, and its conception. Environmental uncertainty talks about the environment where it was taken. Intentional uncertainty deals with heuristics, prospects, preferences and intentions while making decisions. Epistemic uncertainty, uncertainty related to the model's parameters, and Aleatoric uncertainty, uncertainty related to noise and randomness, are the two essential kinds of uncertainties besides semantic uncertainty.

Different approaches have been investigated to measure uncertainty. For example, Tensorflow helps to develop standardized benchmarks for the uncertainty quantification. Semantic uncertainty is also used by researchers to evaluate completeness in data and labels. It can be used as an additional feature to falsify misjudgements. The Monte Carlo Dropout method in deep ensemble models has been used to quantify uncertainty. Others use similar Monte Carlo sampling and Bayesian Belief networks including sensor noise and data for improved predictions and better results.

Among these previous work, most of the research tends to use Bayesian methods and their variants to quantify uncertainty, e.g. the use of dropout method to incorporate uncertainty quantification in reinforcement learning. Multi-label classification is a beneficial but challenging task, and uncertainty quantification poses an even more significant challenge. However, there have been successful efforts using Bayesian learning semantic uncertainty. Abdar et al. [11] have compiled a very comprehensive review of uncertainty measures, all the way from classical models to reinforcement models. It gives a good picture of the Aleatoric and Epistemic uncertainties and their evaluations, especially in ensemble techniques and complicated models, where the opacity improves with an increase in the number and architecture of the models. Bhatt et al. [12] discuss how the uncertainty of the model's predictions can help improve transparency to stakeholders about the model.

Though it is highly important, there is a little research on the effects of uncertainty on accountability.

Levels of Accountability

As shown in the AI life cycle in the previous section, the data, developer, and model are three key components in the AI life cycle. This section investigates accountability from three levels of data, developer, and model.

Data

Data inevitably plays a significant role in the whole AI life cycle. Much research tries to improve data accountability and data privacy. How data is handled from conception and used is essential to assess data accountability. Thus, tools like data provenance and data lineage are very effective in understanding the whole life cycle of a data entity. Primarily in cases where big data is processed in sizeable readable data sets, we must ensure completeness and correctness in data [13]. Also, another critical criterion is to be able to handle sensitive and private data. Redacted data sets, i.e. access to only a certain level and amount of data and not the complete data based on the level of access, come very handily as it makes data available but not visible and has to go through security measures to be completely accessible. Furthermore, adversarial learning is a unique idea where adding bad examples as the input helps the model distinguish from adversarial examples and misinformation in data. While there is quite some evidence for it to work, there is not much to understand why. Thus to improve explainability and accountability in this scenario, it is of utmost importance to incorporate the above principles like data lineage, data provenance, redacted data sets and datasheets. Uncertainty can also be incorporated into data to understand the behaviour of them for accountability.

Model

Model behaviours are changed with the changes of parameters and hyperparameters. Though the data also plays a pivotal role, the model differs significantly from the former. In a big data context, accountability for models can only exist in a culture where models are more than just data processing operations. Because the implications of models reach beyond data security, so does the approach to algorithmic accountability [14]. There are multiple ways to incorporate model accountability in AI. For example, in decision trees and random forests, entropy and information gain help to understand the model. Bayesian methods help calculate and quantify uncertainty, which helps us understand and make decisions. Finally, it comes to the experts to acknowledge the uncertainty in models and make decisions. Neural networks, the black boxes, have complicated structures, making them challenging to analyze. However, the uncertainty measure can serve as a bridge to quantify such complications, especially in ensemble models and neural networks where it is tough to explain the model. They can significantly improve the decision-making process, thereby improving accountability. One must appropriately handle data bias and others that lead to additional uncertainties – models' impact on whether and how people have access to social goods and rights.

Developer

Since AI systems and their opacity pose the accountability problem, the developer accountability is the ultimate idea for AI accountability. Dishing out the responsibility on the model and data makes little sense when the humans are the ones affected the most in the whole AI-based decision making process. Thus, empathy is an essential factor in building software that will ensure accountability. The critical difference between humans and AI systems is that they think differently and emotionally. AI Systems cannot effectively make decisions as they cannot understand the consequences and the developers building them can understand. At the same time, there are also attempts for making the AI System do so to build the bridge of accountability and trust initially. Markovic et al. [15] use a suite of tools to improve accountability. They have provenance detectors and traces throughout the whole cycle that help improve accountability. It is argued that developers should adjust system specifications to ultimately address biases in algorithmic decisions. While regulators should pay special attention to not only what AI systems should learn but also to what they should avoid learning. In the case of driverless cars, for instance, learning algorithms should include restrictions to learning dangerous learning behaviors. Developers must go through different trainings and certifications to understand how they can be accountable and also understand the sensitivity of the issues they are working on. Some of them include the Health Insurance Portability and Accountability Act (HIPAA) certification, Certified Information Privacy Professional(CIPP) certification etc. It is also argued that developers must ensure they are open about the decision-making process and sign up to take responsibility when things go through for AI accountability. Understanding the risks and consequences and taking appropriate measures to understand the level of uncertainty in the data and model also lies in the developer's interest. All these are example significant factors of developers for AI accountability.

Accountability in Different Sectors

Accountability is an important component in different sectors. This section focuses on the investigation of accountability in three major sectors of government, private sector, and academia. Such investigation not only helps to understand the status of accountability in these sectors, but also motivates approaches to implement and achieve AI accountability in these sectors.

Government Sector

The World Health Organization (WHO) guidelines have a strong mention of accountability and emphasize that it is the responsibility of stakeholders to ensure that they can perform those tasks and that AI is used under appropriate conditions and by appropriately trained people. Reliance on AI technologies entails responsibility, accountability, liability, and compensation for any undue damage.

There are different ethics guidelines from government institutions and governments like the EU, US Department of Defense, US Government Accountability Office (GAO) etc. GAO created an AI accountability framework to assist managers in ensuring AI accountability in government programs and operations. The governance, data, performance, and monitoring concepts are arranged around four complementing principles in this framework. India's Niti Aayog talks about "many hand's problem", i.e. due to the complexity and many actors involved it becomes difficult to find the one responsible.

Existing legal systems assign responsibility for actions and consequences based on the assumption that a human agent is involved. The Australian government has well-defined AI ethics standards that urge AI systems that impact majorly on individual's rights to be subject to external assessment, which includes giving fast, accurate, and comprehensive data to independent oversight authorities. The Chinese government anticipates that a closer examination will reveal numerous loopholes and exceptions that will allow the government to circumvent privacy and protection while raising accountability concerns. As a result, a requirement is to address them contextually.

Private sector

Big companies such as Google, Microsoft, IBM, among others, have published ethical principles they follow and pursue, and put special emphases on accountability in their product development. Different industries such as manufacturing, healthcare and E-commerce that leverage AI technologies also imbibe ethical practices in the software, and there have been consequences of not doing so. Results from a recent survey of physicians show that nearly half (49.3%) believe responsibility should lie with doctors, 31.2% with patients who consented to its use, and only 19.4% with the company which created the AI-based tool [16]. The consequences of algorithms are incomprehensible by managers, and administrative decisions are difficult to make for developers. Thus they have to work together closely to avoid unintended outcomes. Most of the research and development teams are well aware of the principles. Though profit and business seeking the companies are, there is an absolute requirement for imbibing ethical principles being transparent and accountable for one's actions. Lawsuits against unethical practices are not new and have been dealt with considerably.

Academia

Most of the research journals, publications include Code of Ethics, Open Source organizations, and academic institutions have more freedom and are independent of selfish motivations, which makes it under-shadowed in the question of accountability. On conducting a survey on public's opinion on AI, Henriksen et al. [17] suggest AI researchers to perform the octagon measurement of AI ethics. It will help them understand where they stand with the public against the ethical principles to empathize with them [17]. Academia have the freedom to take calculative risks and are willing to share and be transparent about the whole process and work. Unlike the government or private institutions, they are entirely transparent and fair with their approaches and act as an exemplary example in multiple scenarios. It is also fair not to compare the government and private sectors as they are very competitive and cannot be transparent in all cases, but that is the most significant setback.

Accountability in Other Areas

Accountability is an age old question. Some of other areas have already addressed them. Cloud technology has logging to help improve answerability. Clinical trials have well established policies to address the concerns and take responsibility. Law gives a clear understanding on the liability aspect of accountability. This section investigates how these areas have addressed accountability by identifying their main concerns and the gaps, and addressing them. AI systems can derive motivation from these areas for the AI accountability.

Cloud and network technology

Since most companies are going serverless and Cloud is the place to store a massive amount of data, we must discuss accountability in areas such as the Cloud. There are multiple cloud security protocols and agreements to maintain the integrity of the Cloud. Distributed recognition and accountability improve the accountability of cloud systems and help understand their vulnerabilities and weaknesses. It tracks users as they move across the cloud nodes and make sure their identity is in sync with the root node. It also checks for intrusions and back-door attacks. Data provenance and logging in the distributed system enables data accountability. Accountability in distributed systems depending on the granularity is discussed in detail by Tan et al. [18]. As the Open Systems Interconnection (OSI) model has seven layers, namely the physical, data link, network, transport, session, presentation and application layer. The problem can be dealt at the level it occurs.

Medicine and clinical trials

The human body and its complications have always posed a challenge. Diseases and problems have been innumerable. Medicine and clinical trials also need accountability as they are life-threatening. They always involve the risks and side effects of them and not only the one who develops, but the consumer is also accountable as they sign up for them understanding the side effects. The Belmont Report attempts to summarize the basic ethical principles identified by the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research in the course of its deliberations. The Health Insurance Portability and Accountability Act of 1996 (HIPAA) is a federal law in USA that requires the creation of national standards to protect sensitive patient health information from being disclosed without the patient's consent or knowledge. The US Department of Health and Human Services (HHS) issued the HIPAA Privacy Rule to implement the requirements of HIPAA. The HIPAA Security Rule protects a subset of information covered by the Privacy Rule.

Members participating in research, accessing the sensitive patient data, must comply with the rules and regulations. Thus rules and regulations plays a pivotal role in the accountability in medicine and clinical trials.

Law and policy making

In law and public policy design, there must be accountable stakeholders as it affects people's daily lives. Public authorities make expenditures in the name of public. Accountability of public authorities means that they are democratic and transparent at that rate.The Oxford Handbook of Law and Politics considers rule making in a democracy, focusing on technocracy and public accountability [19]. Unless done so, there will be nothing but chaos and misunderstanding. To have a well-functioning society, one must take responsibility for their actions. There have been undue consequences of overthrowing government leaders when unclear accountability and pointing fingers across the table. Especially in democratic government's where leaders do not be accountable for their decisions are replaced or sacked by higher authorities.

Discussion

Accountability is becoming one of indispensable ethical principles that AI needs to follow as AI is widely used in different domains especially high-stake domains for prediction, automation, planning, targeting, and personalisation as well as others. This paper investigated AI accountability from different perspectives of concepts, approaches for achieving AI accountability, factors affecting AI accountability, accountability from three key components in the AI life cycle, AI accountability in three typical sectors, and accountability in other areas which may motivate approaches for achieving AI accountability.

It was found that the current work on AI accountability mostly focuses on the clarification of accountability in AI systems and the identification of factors that affect AI accountability. Despite the importance of these investigations, there are still different challenges especially on operationalising AI accountability in practices. This paper articulates the following challenges as examples of making AI accountable:

- **A systematic framework for achieving AI accountability**. It was shown that the consideration of accountability in each phase of AI life cycle is highly necessary for achieving AI accountability. However, it still lacks a systematic framework to achieve AI accountability by considering all related phases in the AI life cycle.
- **Metrics**. From the operable perspective [20], there is a lack of widely acceptable metrics that are used to evaluate AI accountability. The metrics could include qualitative metrics and quantitative metrics that can be evaluated easily by developers and justified by AI ethics authorities.
- **Standardisation**. The standardisation can help to simplify the complexity and provide a common baseline in order to make AI accountability operable. The standards are not only on which components of accountability should be validated, but also on how those components should be validated, and what are the criteria that AI solutions "pass" the validation.
- **Factors affecting AI accountability**. Despite the investigation of affecting factors such as explainability and uncertainty, it is still a challenge how to systematically mitigate negative factors and make full use of positive factors for achieving AI accountability.

There are also obstacles that affect the operationalisation of AI accountability. In addition to obstacles that are common to all AI ethical principles such as communication between AI research field and other disciplines, complexity, subjectivity, risk and government policies/rules [20], AI accountability faces specific obstacles in its operationalisation. For example, different domains and sectors may have different requirements on the accountability, which needs special considerations for each domain and sector. Such considerations propose obstacles for the generalisation of operationalisation of AI accountability. Fortunately, we are seeing the increasing collaboration among different disciplines which would be the good start point for effective AI accountability. Big companies in different sectors are also setting up their own guidelines for AI ethics. It is hoped that this paper can motivate researchers and stakeholders to propose effective approaches to solve challenges and eliminate obstacles in operationalising AI accountability.

Conclusion

This paper reviewed concepts of accountability and investigated major approaches for achieving accountability in AI systems. Typical factors affecting accountability were also identified. We then discussed accountability from three key components of data, model and developer in the AI life cycle. Furthermore, Accountability in different sectors of government, private sector and academy was reviewed to motivate approaches to achieve AI accountability in these sectors. Since accountability is an age old question and some of other areas have already addressed them, this paper investigated how these areas have addressed accountability by identifying their main concerns and approaches addressing them, which could motivate approaches for achieving AI accountability. This paper articulated key aspects to consider in making AI accountability operable and initiated discussions on challenges and obstacles needed to mitigate in making AI accountability operable.

REFERENCES

1. M. Loi and M. Spielkamp, “Towards accountability in the use of artificial intelligence for public administrations,” *CoRR*, vol. abs/2105.01434, 2021. [Online]. Available: https://arxiv.org/abs/2105.01434
2. J. Singh, J. Cobbe, and C. Norval, “Decision Provenance: Harnessing data flow for accountable systems,” *IEEE Access*, vol. 7, pp. 6562–6574, 2019.
3. T. Gebru, J. Morgenstern, B. Vecchione, J. W. Vaughan, H. Wallach, H. Daumé III, and K. Crawford, “Datasheets for Datasets,” *arXiv:1803.09010 [cs]*, Mar. 2020, arXiv: 1803.09010. [Online]. Available: http://arxiv.org/abs/1803.09010
4. I. D. Raji, A. Smart, R. N. White, M. Mitchell, T. Gebru, B. Hutchinson, J. Smith-Loud, D. Theron, and P. Barnes, “Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing,” *arXiv:2001.00973 [cs]*, Jan. 2020, arXiv: 2001.00973. [Online]. Available: http://arxiv.org/abs/2001.00973
5. H. Smith, “Clinical AI: opacity, accountability, responsibility and liability,” *AI & SOCIETY*, vol. 36, no. 2, pp. 535–545, Jun. 2021. [Online]. Available: https://doi.org/10.1007/s00146-020-01019-6
6. A. Henriksen, S. Enni, and A. Bechmann, “Situated accountability: Ethical principles, certification standards, and explanation methods in applied ai,” New York, NY, USA, p. 574–585, 2021. [Online]. Available: https://doi.org/10.1145/3461702.3462564
7. B. S. Miguel, A. Naseer, and H. Inakoshi, “Putting Accountability of AI Systems into Practice,” in *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*. Yokohama, Japan: International Joint Conferences on Artificial Intelligence Organization, Jul. 2020, pp. 5276–5278. [Online]. Available: https://www.ijcai.org/proceedings/2020/768
8. D. Minh, H. X. Wang, Y. F. Li, and T. N. Nguyen, “Explainable artificial intelligence: A comprehensive review,” *Artif. Intell. Rev.*, vol. 55, no. 5, p. 3503–3568, jun 2022. [Online]. Available: https://doi.org/10.1007/s10462-021-10088-y
9. J. Fox, “The uncertain relationship between transparency and accountability,” *Development in Practice*, vol. 17, no. 4-5, pp. 663–671, Aug. 2007. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/09614520701469955
10. A. Paiva, I. Leite, H. Boukricha, and I. Wachsmuth, “Empathy in Virtual Agents and Robots: A Survey,” *ACM Transactions on Interactive Intelligent Systems*, vol. 7, no. 3, pp. 11:1–11:40, Sep. 2017. [Online]. Available: https://doi.org/10.1145/2912150
11. M. Abdar, F. Pourpanah, S. Hussain, D. Rezazadegan, L. Liu, M. Ghavamzadeh, P. W. Fieguth, X. Cao, A. Khosravi, U. R. Acharya, V. Makarenkov, and S. Nahavandi, “A review of uncertainty quantification in deep learning: Techniques, applications and challenges,” *CoRR*, vol. abs/2011.06225, 2020. [Online]. Available: https://arxiv.org/abs/2011.06225
12. U. Bhatt, J. Antorán, Y. Zhang, Q. V. Liao, P. Sattigeri, R. Fogliato, G. G. Melançon, R. Krishnan, J. Stanley, O. Tickoo, L. Nachman, R. Chunara, M. Srikumar, A. Weller, and A. Xiang, “Uncertainty as a Form of Transparency: Measuring, Communicating, and Using Uncertainty,” *arXiv:2011.07586 [cs]*, May 2021, arXiv: 2011.07586. [Online]. Available: http://arxiv.org/abs/2011.07586
13. R. Thiago, R. Souza, L. Azevedo, E. Soares, R. Santos, W. Santos, M. De Bayser, M. Cardoso, M. Moreno, and R. Cerqueira, “Managing Data Lineage of O & G Machine Learning Models: The Sweet Spot for Shale Use Case,” *arXiv:2003.04915 [cs]*, Mar. 2020, arXiv: 2003.04915. [Online]. Available: http://arxiv.org/abs/2003.04915
14. A. Vedder and L. Naudts, “Accountability for the Use of Algorithms in a Big Data Environment,” Social Science Research Network, Rochester, NY, SSRN Scholarly

Paper ID 3407876, mar 2017. [Online]. Available: https://papers.ssrn.com/abstract=3407876

15. M. Markovic, I. Naja, P. Edwards, and W. Pang, "The accountability fabric: A suite of semantic tools for managing ai system accountability and audit," *CEUR Workshop Proceedings*, Jul. 2021, supported by the award made by the UKRI Digital Economy programme to the RAInS project (ref: EP/R033846/1).; The 20th International Semantic Web Conference, ISWC 2021 ; Conference date: 24-10-2021 Through 28-10-2021. [Online]. Available: https://iswc2021.semanticweb.org/
16. S. Oh, J. H. Kim, S.-W. Choi, H. J. Lee, J. Hong, and S. H. Kwon, "Physician confidence in artificial intelligence: An online mobile survey," *J Med Internet Res*, vol. 21, no. 3, p. e12422, Mar 2019. [Online]. Available: http://www.jmir.org/2019/3/e12422/
17. Y. Ikkatai, T. Hartwig, N. Takanashi, and H. M. Yokoyama, "Octagon measurement: Public attitudes toward ai ethics," *International Journal of Human–Computer Interaction*, vol. 0, no. 0, pp. 1–18, 2022. [Online]. Available: https://doi.org/10.1080/10447318.2021.2009669
18. Y. S. Tan, R. K. Ko, and G. Holmes, "Security and Data Accountability in Distributed Systems: A Provenance Survey," in *2013 IEEE 10th International Conference on High Performance Computing and Communications 2013 IEEE International Conference on Embedded and Ubiquitous Computing*, Nov. 2013, pp. 1571–1578.
19. S. Rose-Ackerman, *Law and Regulation*. Oxford University Press, 2008.
20. J. Zhou and F. Chen, "Ai ethics: From principles to practice," *AI & Society*, 2022.

Ashwin Kumar Raja received his M.Sc. degree in Mathematics and B.E. degree in Computer Science from BITS Pilani, India in 2022. He is currently with Amazon, Chennai, India. He is working to build innovative products in the lines of demand forecasting and price optimization using Causal Machine Learning. His research interests include Causal Machine Learning, AI ethics, explainability and Uncertainty in AI. Contact him at his personal email at ashwinkumarraja@gmail.com or work email at ashwinkw@amazon.com.

Jianlong Zhou is an Associate Professor in Computer Science at University of Technology Sydney, Sydney, Australia. He holds a PhD degree in Computer Science from The University of Sydney, Australia. His research interests include AI for social good, AI fairness, AI explainability, smart agriculture, visual analytics, behaviour analytics, human-computer interaction, and related applications. He is a Senior Member of IEEE and a Senior Member of ACM. Contact him at jianlong.zhou@uts.edu.au.