**Journal of Global Knowledge and Innovation (JGKI)**
Volume 1, Issue 1, October-2024
**ISSN** 3079-1995
**DOI:** https://doi.org/10.5281/zenodo.13885892

JGKI: Volume 1, Issue 1, September 2024, online: ISSN 3079-1995
www.jgkijournal.com

SURVEY PAPER OPEN ACCESS Check for updates

# Enhancing Trust in Artificial Intelligence: A Review of Assurance Methods for Broad Application

**[1]*Nayema Akter, [2]Md Miftahul Bari**
[1]School of computer science and engineering, Sichuan University, China, School of information Engineering, Mianyang teachers College.
nayemaakhter32@gmail.com; mdmiftahulbari@gmail.com;



## Abstract
Cognitive computing technologies particularly, Artificial Intelligence (AI) algorithms are playing more and more in decision support and in operational functions in different fields. As part of the growing use of such algorithms, there is now a need to validate these algorithms and make them credible or neutral in their processing. The existing studies on AI assurance show that the area is highly disjointed, with the that involves diverse motivation and assumptions prevailing over the state of the art. This manuscript provides a taxonomic view of AI assurance research that occurred between 1985 and 2021 with a focus on structured methodologies. A new definition of AI assurance is discussed, and a comparison of the assurance approaches using a recently introduced ten-metric scale is provided. This manuscript concludes with design principles and proposed directions for future research regarding assurance in the broad field of artificial intelligence.

## Keywords
AI Assurance, Artificial Intelligence, Validation, Verification, Explainable AI

Journal of Global Knowledge and Innovation (JGKI)
Volume 1, Issue 1, October-2024
ISSN 3079-1995
DOI: https://doi.org/10.5281/zenodo.13885892

## Introduction and Survey Structure

Big data has brought about the development of artificial intelligence more accentuated by the use of statistical learning methods. Contrary to previous endeavours that led to what is referred to as 'AI winters,' current advances have allowed for new systems of AI to not only cope with human-level performance but at times exceed it, within a number of domains. This shift proves that AI is now widely used for choosing directions in business and life, from revenue prediction to self-driving cars and medical diagnosis. The viability of the integration increases as more functions in our daily lives are assimilated into the AI environment; therefore, there is the need for proper means of evaluating such systems.

In this context, what is known as assurance or validation or verification presents itself as one of the major determinants of successful deployment of AI technologies. Assurance involves different practices and framework meant to offer guarantee on how the AI systems are going to function well, be ethical and be transparent. The problem is one of scope here – AI is a broad concept and so too is 'assurance'. These questions are not simple and this manuscript seeks to fill this gap by presenting an extensive literature review on AI assurance.

To this end, this paper systematically reviewed the existing body of knowledge on AI assurance with an emphasis on papers that were published between 1985 and 2021. This review discusses the current state of AI assurance methodologies and extracts the themes, definitions, and paradigms that can be found in this area of research. In addition, we extend the definition of AI assurance that takes into account different domains at which AI functions. Formatting of this paper is sectioned out in different sections as discussed below. This introduction sets the stage for what is a comprehensive review of AI assurance, the many different forms it may take, and problems that arise from the existence of such a broad methodological range. In the next section, we explain the review and scoring of assurance methods; the ten-metric scoring system for the comparison of existing approaches is introduced. Measuring in this manner is useful as a structured applied scoring system for researchers and practitioners in comparing the performance of various assurance models.

The paper then continues to discuss directions for future work in AI assurance and ultimately calls for the development of a single system that can effectively consider all of the various uses of the AI technology. In the final section, we present the implications drawn from our research work and identify possibilities for further development in related research area. In this manuscript, we always try to contribute towards the development of the understanding of AI assurance for more robust reliable and trustworthy AI systems.

## AI Assurance Landscape

The approach to AI assurance is surrounded by a large number of consistencies in methodologies and frameworks, as well as terminological differences depending on the field. This section of the chapter takes a literature review approach to examine the definitions and practices in the field of AI assurance.

Journal of Global Knowledge and Innovation (JGKI)
Volume 1, Issue 1, October-2024
ISSN 3079-1995
DOI: https://doi.org/10.5281/zenodo.13885892

*Definition and Scope*

AI assurance may be described more generally as the activities that facilitate the responsible and effective functioning of AI systems. This definition encompasses various subfields of AI, that include, machine learning, computer vision, and natural language processing, all of which have their own specificity in terms of their challenges as well as assurance requirements.

*Current Methodologies*

Similarly, analyzing the current set of methods suggests that researchers are employing a wide variety of methods ranging from more classical validation approaches to relatively contemporary paradigms of explainability and accountability. Notable methods include:

**Verification and Validation (V&V):** Old habits from engineering disciplines borrowed for AI systems.
**Performance Metrics:** The formalisms that determine the measures of accuracy, reliability and fairness of the AL systems.
**Explainable AI (XAI):** Approaches designed to address the issue of explaining-AI decision making processes to the four user groups envisioned by the frameworks.

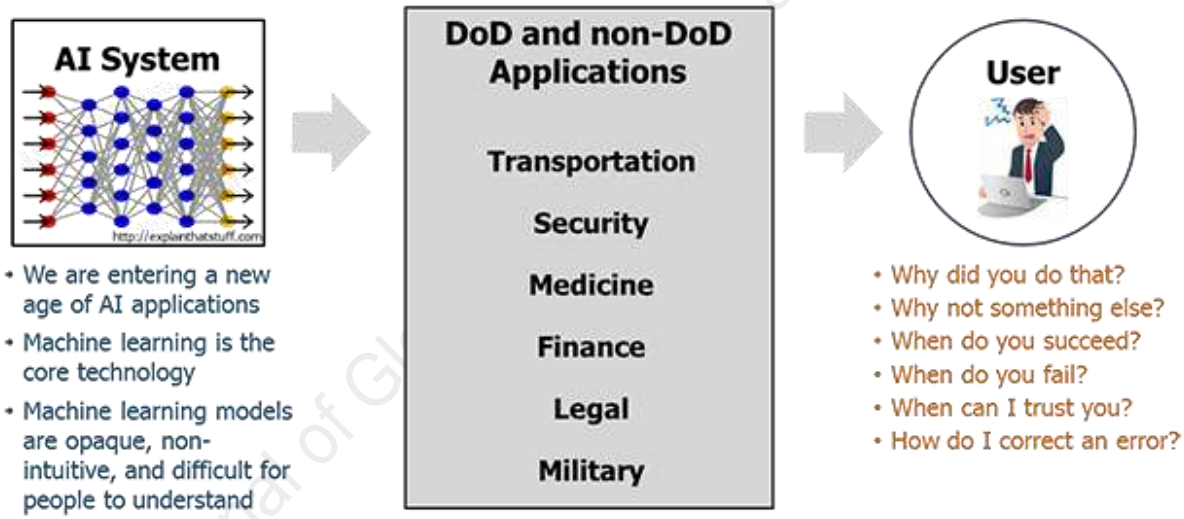



*Figure 1: The Need for Explainable AI*

*Challenges in AI Assurance*

Still, researchers face several challenges in the field even with progressive developments that have been made. These include some disaggregation of evaluation, the absence of uniform definitions for AI, the challenges in assessing AI systems in real contexts and relatively short cycles of technology development that can overrun the existing methods of assurance.

Journal of Global Knowledge and Innovation (JGKI)
Volume 1, Issue 1, October-2024
**ISSN** 3079-1995
**DOI:** https://doi.org/10.5281/zenodo.13885892

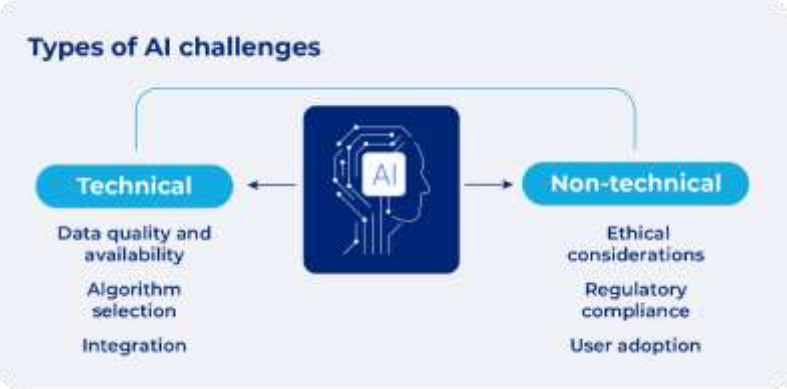



*Figure 2: AI Challenges*

**The Review and Scoring of Assurance Methods**

Within the context of AI assurance methods are very important when it comes to AI systems for creating reliability and ethics. The need for better assurance mechanisms comes from the dynamic and highly fluid nature of regular AI algorithms. AI is therefore different from conventional software as these programs are dynamic and they keep on learning and changing, the process of validating and verifying them needs to be different. This section focuses on the analysis of assurance methods, their evaluation and rating, with special emphasis on their significance and efficiency in reference to different sub-areas of AI.

*The historical context and the evolution*

AI assurance has its traditional development to concepts like the Turing test in which there was a paradigm of human interaction that tested the intelligence of machines. When the AI systems moved from being purely expert systems that used rule-based systems to learn and work to the AI systems we see today as complex machine learning, deep learning, the approaches to assurance also had to change. Early assurance techniques were aimed and were essentially a set of measures that included primarily static testing as well as validation against parameters or criteria set for an algorithm. However, the continually developing nature of the systems means that it is often the case that these methods are not adequate.

The field has expanded its assurance approaches over the course of the years by reaching new ones. With the logical theory of entailment-based systems, in validation frameworks all the way to deep learning and reinforcement learning, the field is vast but varied. This review categorizes assurance methods into several key areas: increasing the data quality, reliability of the models, explanation of the algorithm, and after implementing its monitoring.

Journal of Global Knowledge and Innovation (JGKI)
Volume 1, Issue 1, October-2024
**ISSN** 3079-1995
**DOI:** https://doi.org/10.5281/zenodo.13885892

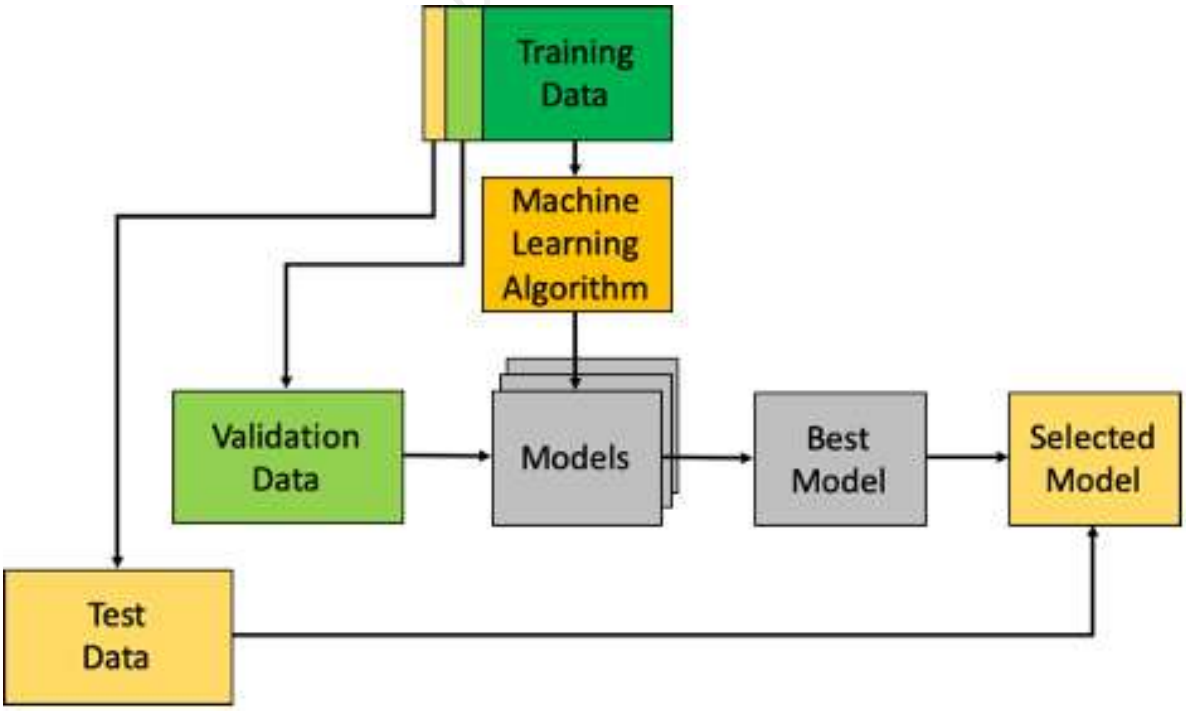
*Data Quality Assurance*

Among the core components that have to be laid in any AI system, data quality holds one of the highest priorities. The real assurance methods in this category aim to guarantee that the data used to train and test the IMS and workspace, is unprejudiced and representative. There are different ways which are used to solve the common problems with the data information, these include data profiling, data pre-processing among others before they actually started affecting the performance of the model. For example, methods such as detecting out of limits data and standardizing data are critically important in the pre-processing of a data set for machine learning processing.

Additionally, the idea of data lineage has emerged, which helps the stakeholders define how the data used in the AI models was collected, processed, and changed since their acquisition. By doing so, they earn the readers' trust and in so doing enables the reader discover possible sources of bias that could have skewed results hence improving reliability (Kitchin, 2014). In our scoring framework, methods that have evidence of adequate data quality control are ranked higher because data quality control is one of the areas on which dependability of AI greatly depends on.

*Model Validation Techniques*

Model validation is the use of a number of approaches with aim of assessing the performance of the artificial models. Some traditional methods for model accuracy are k-fold cross-validation, and hold-out validation techniques. But due to the complex nature of the models used in AI, there must be other forms of validation techniques necessary for such models. For instance, sensitivity analysis and feature importance scoring show the amount of input features' impact on the model prediction.



**Journal of Global Knowledge and Innovation (JGKI)**
Volume 1, Issue 1, October-2024
**ISSN** 3079-1995
**DOI:** https://doi.org/10.5281/zenodo.13885892

*Figure 3: The minimum validation framework*

This level of granularity is tremendously important when diagnosing possible weaknesses in the model that could lead to missteps as in fields like healthcare and finance with high risks involved. Moreover, adversarial testing has become an important type of validation, reaching the goal of describing the weaknesses of the models when they are exposed to input information designed to deceive an algorithm. The observed results suggest that the develoment of assurance methods that rely on a number of distinct extensive model validation procedures typically leads to higher score rates. Bachelor's thesis: The importance of not only evaluating but also interpreting model behavior is becoming more and more recognized for constructing trustworthy AI systems.

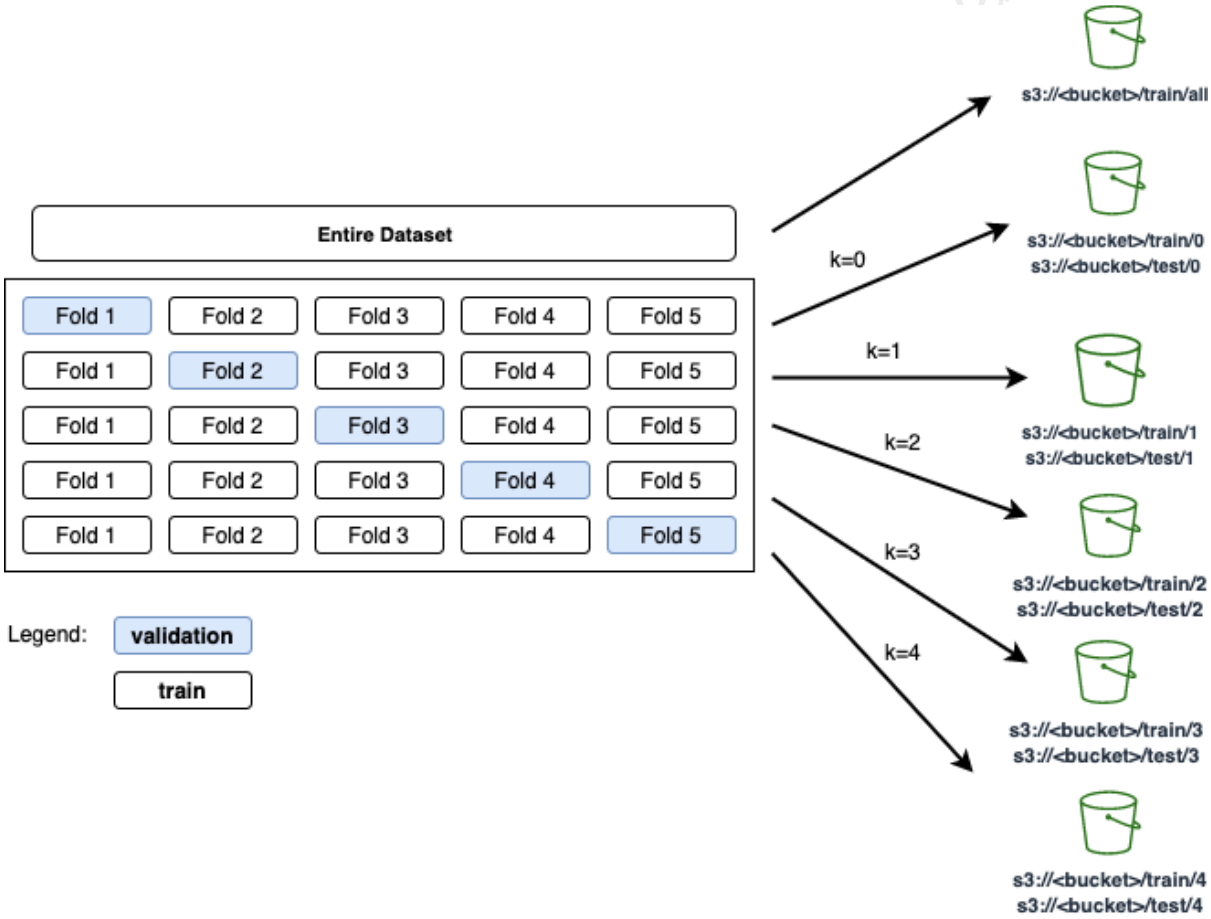



*Figure 4: K-fold cross-validation: original data is split into k equal-sized samples uploaded to S3 bucket*

Journal of Global Knowledge and Innovation (JGKI)
Volume 1, Issue 1, October-2024
**ISSN 3079-1995**
**DOI:** https://doi.org/10.5281/zenodo.13885892

*Computational and Semiotic Transparency*

The need for algorithmic explanations has risen with the advancement of AI systems to essential procedural decision-making points. Explainable assurance methods seek to reduce the 'Black Box' phenomenon felt in the use of AI algorithms to make decisions more understandable for involved parties. This is especially the case given that many industries, especially healthcare, has AI helping out in diagnosing diseases or even prescribing medication. There are several models that have been created to explain model predictions – LIME that approximates a model with simpler, understandable models (Ribeiro et al., 2016) and SHAP (SHapley Additive exPlanations) (Lundberg & Lee, 2017). The applicability of these methods is assessed on the grounds of interpretability and practical usefulness on one hand, and accuracy on the other. In our scoring system, the approaches which have sound frameworks of algorithmic accountability and user understanding get positive scores. This emphasis derives from the increasing belief that, for users to put their trust in the AI system, they must comprehend its processes and for the system to be held to account, it must be possible to know how it arrives at its decisions.

*Post-Deployment Monitoring*

Overseas, the administration does not cease when an AI system is implemented. To reduce the chances of models degrading and to keep updating models to accommodate new conditions, familiarization with Monte Carlo tests and continual monitoring is important. The main post-deployment processes include monitoring of the models and their metrics of performance, as well as feedback and real-life outcomes where a reduction in the effectiveness of a model occurs, or where different bias tendencies show up. The techniques used under this process are the drift detection techniques, which assist when there is a significant difference in the statistical properties of the data being fed into the models and their training data. Because the distribution of the underlying data changes over time this can be an indication of a degradation of the model, which has to be retrained or recalibrated (Gama et al., 2014). For the same, better methods to collect user feedback can help the developers to know about the unforeseen behavior of models, thereby enriching assurance practice. Our review underlines that methodologies with efficient post-deployment monitoring systems receive a high score in our scoring system. The element of monitoring as an ongoing process enables one to keep AI systems trustable and ethical throughout their active cone lifespan.

*Scoring Framework*

In order to assign scores and methodology results, a Scoring Framework was developed by the researchers as the means of evaluation and comparison. In this paper, to build a systematic approach for comparing the assurance methods described above, we propose a ten-metric scoring system. The following lists parameters, which are covered by this scoring: quality of the collected data, robustness of the validation algorithms, interpretability, and efficiency of monitoring. All these metrics are then applied to each method to deliver an overall assurance capacity rating. For example, methods which perform highly

Volume 1, Issue 1, October-2024
ISSN 3079-1995
DOI: https://doi.org/10.5281/zenodo.13885892

in data profiling and bias detection receive high ratings because of their preventative considerations regarding the data quality. Likewise, methods that involve reliable approach of model checking and readily understandable reasons for the made decision receive positive assessment. These results show that the extent to which assurance methods are applied is significantly related to their capacity for building trustworthy AI.

## Recommendations and the Future of AI Assurance

*Need for Standardization*

There is controversy, lack of universal standards, and variation in the definitions and methodologies used in AI assurance a critical need for standardization. It will also enhance understanding since different people use different names for the same problems, or else they use the same name to refer to different issues.

*Core Concepts & A Common Framework*

A future research direction in the area of AI assurance should concern the creation of integrated assurance frameworks that are sector-specific and could be scalable to micro-areas of AI. This will help in improving the possibility of applying assurance practices in a number of sectors.

*Emphasis on Explainability*

Due to the necessity of explainability inherent in AI systems, lack of which weakens their practical applicability, one can note that further research in this area should focus on creating an explainable AI perspective. This will in turn help develop more trust within users and stakeholders.

## Future Components of AI Assurance Research

With the development of AI and the consequent incorporation on the social matrix the future of AI assurance will be completely different. One of the key aspects of this process is the need for cross-sector working. That is because the problems which AI systems raise in their application are varied and specific by the industry where they are used and need input from multiple disciplines, including healthcare, finance, education, and transportation industries. For instance, innovative healthcare specialists need to collaborate with the creators of medical AI systems stressing that, along with their efficiency, it is necessary to have proper ethical standards for those systems that guarantee patients' safety and privacy. Just as with usage of artificial intelligence systems it must be beneficial to receive input from social scientists and ethicists in regards to the further prospects of artificial intelligence systems in society. These collaborations can assist in making certain that not only are these systems technically beneficial and technically correct; they are also socially advantageous, or at least socially neutral, and compliant with human moral principles. The other promising themes for future AI assurance link together the operational practices of incorporating modern methodologies that combine common formal verification and validation with machine learning approaches. Current testing approaches used in software assurance fail to capture AI systems because these systems are interactive and

Volume 1, Issue 1, October-2024
ISSN 3079-1995
DOI: https://doi.org/10.5281/zenodo.13885892

can change themselves. For these reasons, it becomes necessary to create mixed assurance frameworks that could easily synchronize with learning behaviors associated with AI. For instance, practices like, the continuous integration and testing in which models are run against new data will ensure that they remain relevant. Furthermore, explainable AI (XAI) methodologies will be instrumental in developing the required assurance process, which, in addition to being sound, needs to be understandable by intended users. This in turn can made it easier to refine the models, enabling developers to gain significantly better oversight just by explaining how an AI makes its choices.

### Reviewed Methods and Their Total Scores

The table below gives an overview of the studies reviewed under the AI assurance methods, their publication year, first author, citation, publishing venue, AI subarea to which they belong, and the scores assigned to them out of the criterion we used for the evaluation. The scores indicate the level of effectiveness as well as the extent of coverage of the assurance methods across the context.

| Year | First Author's Last Name and Citation | Publishing Venue | AI Subarea | Total Score |
|---|---|---|---|---|
| 2020 | D'Alterio [50] | FUZZ-IEEE | XAI | 10 |
| 2019 | Tao [208] | IEEE Access | Generic | 10 |
| 2020 | Anderson [11] | ACM TIIS | RL | 9 |
| 2020 | Birkenbihl [29] | EPMA | ML | 9 |
| 2020 | Checco [39] | JAIR | DS | 9 |
| 2020 | Chen [40] | IEEE Access | XAI | 9 |
| 2020 | Cluzeau [43] | EASA | DL | 9 |
| 2019 | Kaur [109] | WAINA | XAI | 9 |
| 2020 | Kulkarni [117] | Academic Press | DS | 9 |
| 2020 | Kuppa [118] | IEEE IJCNN | XAI | 9 |
| 2020 | Kuzlu [120] | IEEE Access | XAI | 9 |
| 2021 | Massoli [145] | CVIU | DL | 9 |
| 2020 | Spinner [201] | IEEE TVCG | XAI | 9 |
| 2016 | Veeramachaneni [226] | IEEE HPSC | DS | 9 |
| 2018 | Wei [230] | AS | RL | 9 |
| 2020 | Winkel [236] | EJR | RL | 9 |
| 2014 | Ali [8] | GISci | DS | 8 |
| 2018 | Alves [9] | NASA ARIAS | ABS | 8 |
| 2019 | Batarseh [24] | EDML | DS | 8 |
| 2016 | Gao [71] | SEKE | DS | 8 |
| 2020 | Gardiner [72] | Nature Sci Rep | ML | 8 |
| 2016 | Gulshan [81] | JAMA | CV | 8 |
| 2020 | Guo [82] | IEEE ICCVW | XAI | 8 |
| 2020 | Han [87] | IET JoE | XAI | 8 |

Volume 1, Issue 1, October-2024
ISSN 3079-1995
DOI: https://doi.org/10.5281/zenodo.13885892

| | | | | |
|---|---|---|---|---|
| 2016 | Heaney [93] | OD | GA | 8 |
| 2019 | Huber [97] | KI | AAI | 8 |
| 2019 | Keneni [112] | IEEE Access | XAI | 8 |
| 2020 | Kohlbrenner [116] | IEEE IJCNN | XAI | 8 |

**Discussion of Results**

The instances of the table depict several approaches that are scattered across the AI subfields such as XAI, ML, DL, and RL. Especially the methods receiving the highest score, and the tendency to raise their scores, especially in 2020, indicate the increasing concern for the assurance aspect of AI development. Interestingly, the papers by D'Alterio and Tao define AI assurance in a more exhaustive manner and, therefore, are absolutely correct. These methods demonstrate the current direction of incorporating more robust assurance strategies into the developing milieu of AI systems. Moreover, the scoring demonstrates how well each approach can solve the existing issues concerning assurance, including transparency, minimize bias, and reliability of AI solutions. With AI applications steadily moving deeper into core industries, the different methodologies presented in this paper highlight the importance of appropriate means of AI assurance. In conclusion, this table not only shows a variety of AI assurance techniques, but also finds that further study on the AI assurance method is needed to improve the confidence of AI.

**Recommendations and the Future of AI Assurance**

AI assurance is therefore a complex task of the future that requires a change in how society, organizations, institutions, and government design, implement, and regulate AI systems. Since the advances of the AI technologies are only growing complex and widespread, it is imperative that assurance practices are fully developed. Among the proposed solutions for the future development of AI assurance the focus is made on the collaboration between specialists of different fields including computer science, ethicists, sociologists, and specialists of specific domains including healthcare or finance. This cooperation is necessary because AI solutions are implemented in different areas of activity, which have different problems and needs. By way of illustrating this criterion, in healthcare, an AI system needs to prove both functionality and ethicality compared to human systems in areas such as safety and patient data privacy. I suppose that gathering people with different backgrounds is helpful, as the resulting discussion would help avoid creating entirely suitable and efficient AI systems but also socially appropriate ones. In addition, such an interdisciplinary approach is useful where the domain of AI assurance methods can benefit from integrating findings of the social sciences on societal consequences of the implementation of AI solutions. Coupled with the issue of collaboration is the aspect of integration of the best verification and validation methodologies, which incorporates the traditional 'industrial' approaches, with modern approaches based on machine learning. Most of the common assurance techniques prove to be inadequate or insufficient when used on the AI systems since these adapt to new conditions as well as new knowledge as time goes on. Thus, hybrid assurance frameworks which can take into account features of AI in its current stages of evolution are also necessary. For example, using of the practices as regular model checking through C/I/T frameworks that imply testing AI models on new data could be helpful in keeping the AI models accurate and trustworthy. Also, the

Volume 1, Issue 1, October-2024
ISSN 3079-1995
DOI: https://doi.org/10.5281/zenodo.13885892

application of the explainable AI (XAI). transferring DevOps practices into assurance frameworks will be crucial in building the processes that would not only check correctness of development of artificial intelligence systems but also ensure that their usage is transparent for users and other stakeholders. They added that this is important for trust because understanding AI decision-making is important in light of consequences that affect individual and collective rights. New good and services will also ensure continuity in the application of ethical issues in organizing future AI assurance. That is why integrating ethical standards into assurance practices is necessary as algorithmic bias remains a problematic issue for the AI systems' credibility. This may include the establishment of polices that address the best approaches when it comes to the AI systems and these includes; fairness accountability and transparency polices. For example, organizations can include bias preventive system forming in the processes of developing an AI, which will help detect bias and avoid them before the AI systems are launched. Also, developing guidelines to the ethical application of AI technologies across various industries that will be anchored by regulatory authorities can enhance organisational responsibility. Controlling and supervisory mechanisms will also need to remain an ongoing process to guarantee the sustainability, and safety, of these bots over time. Due to the function that makes these systems adaptive and able to learn on their own, AI systems need to be evaluated time and again while traditional software can take years between updates. It is possible, therefore, to develop structures that will monitor the performance of an AI system in real-time and inform designers or users of any problem that may have arisen in order to remind them of the goal of the system that has been created. In addition, dynamic assurance frameworks that are mature in tandem with AI frameworks can be designed, to modify assurance frameworks according to best practices gathered from new knowledge, emerging technologies, and failed implementations. On the contrary, this dynamic approach will not only preserve the utility of AI but also build immunity against disorienting threats into the delivery processes. User engagement and feedback will also be basic pillars of the future AI assurance initiatives. Based on the idea that users are the final consumers of AI systems, users sometimes uncover useful information about the functionality, drawbacks, and feasibility of AI systems. Hence, establishing ways of continuous reception of user information can greatly improve the solidity and applicability of assurance activities. Therefore, engaging users throughout the AI assurance process can help developers understand how an AI system affects real-world use cases and enhance the system to make it run smoothly, which will make the users happy. This engagement is particularly important in fields like the medical or educational fields where the risks are high, and the losses following an AI failure catastrophe are grave. However, engaging users in the choice of the approach to present explainability can lead to more understandable and acceptable to users AI systems at various levels and within different contexts. Legal and regulatory requirements will also be a key driver of the future state for the assurance of AI since there is an increasing use of AI in businesses, making standardization an essential factor. Such guidelines will be of immense aid in helping an organization to undertake AI implementation while at the same time avoiding the many pitfalls that are associated with the use of Artificial Intelligence. Legal institutions can contribute to the development of norms for AI assurance that include many topics including data protection and ethic, as well as system traceability. Through developing a useful structure for AI assurance, officials have an opportunity to gain people's trust and faith in AI technologies and so

ISSN 3079-1995
DOI: https://doi.org/10.5281/zenodo.13885892

contribute to the enhancement of AI acceptance within society. It will also be crucial to develop public awareness and understanding of issues related to AI assurance in order to actively participate in the formation of the trend of the field. But as AI enters peoples' lives more and more it is important to work towards making people aware of what it is capable, as well as not capable of. One of the ways to reduce different concerns from users of AI systems and from the society is through proper briefing of the community and users of the AI systems on how the individual AI systems work, what they were designed for, and what the probable drawbacks are. Further, it can be mentioned that news and social media, as well as educational endeavors, can help to promote discussions of ethics of the use of AI, thus setting up multisided cooperation. When the public is educated on the kind of technologies AI is made up of, then organizations get to improve community relation and ensure that the integration of AI gets to be as per the standards of the society. Thus, the potential development of AI assurance depends on the multilevel cooperation and approach that implicates the combination of the scientific and academic research, methodologies' application, pro-ethical precautions, monitoring, users, regulations and awareness. Solving these interrelated domains will be vital for the development of proper assurance solutions that will improve the dependability, security, and ethical utilization of AI facilities. Laying these fundamentals preemptively in these areas will be critical as new AI advancements emerge given the challenges and emerging issues that are likely to be unveiled and solved in the process of enhancing society's trustworthiness and utility of AI.

## Conclusions

AI assurance is an interesting and evolving field that raises many philosophical and pratical questions that were discussed in the paper. What are the criteria for a recognised system? That raises the question – When do we stop testing? In managing these subtleties, there is evidence that effective realisation assurance primarily depends on specific calibration and testing. The seven specifications I propose: better data quality, more specific data, procedural assurance, automated methods, and user involvement will help to improve the AI assurance practices. The second argument is that the current context is best served with multi-disciplinary AI assurance solutions – this means that experts across all fields should get involved in ensuring that reliable AI systems exist. Thus, the way towards AI evolution is not only in technology, but also in changes in attitudes of organization and society towards responsibility, openness and ethic when implementing AI systems.



## Acknowledgements

We would like to thank the researchers and practitioners who have worked in this novel field where AI assurance resides as the basis for our research. Sanbra has done pioneering work that has significantly enhanced the knowledge we have and the way we tackle this thorny issue. We would also like to express our gratitude to numerous theoretical authors including D'Alterio, Tao, Anderson or others whose works and findings have enriched the knowledge regarding the specificities of AI assurance approaches and experiences. Their passion for the development of the scholarship has informed our work, and the course of this manuscript.

Journal of Global Knowledge and Innovation (JGKI)
Volume 1, Issue 1, October-2024
ISSN 3079-1995
DOI: https://doi.org/10.5281/zenodo.13885892

## Authors' Contributions

The work presented in this study was done by all the authors. All authors FB, LF, and CH participated in the design of the study, the definition of assurance and the writing of this manuscript. Both authors contributed to the process of literature review, development of the visualizations, as well as tables and scoring systems. Both of them made sure that the coverage offered in the paper covers all the aspects of the topic on AI assurance, which was an indication of dedication to quality work when it comes to research. All authors have read this manuscript and agreed with the changes and results described in it.

## Availability of Data and Materials

Not applicable.

## Declarations

## Ethics Approval and Consent to Participate

Not applicable.

## Consent for Publication

Not applicable.

## Competing Interests

The authors declare that they have no competing interests.

**Journal of Global Knowledge and Innovation (JGKI)**
Volume 1, Issue 1, October-2024
**ISSN 3079-1995**
**DOI:** https://doi.org/10.5281/zenodo.13885892

## Referances

[1] M. R. A. S. S. B. M. M. H. M. M. &. S. F. H. Sarkar, "Deep Learning Methods for Chest X-Ray Imaging-Based COVID-19 Pneumonia Detection.," *Journal of Image Processing and Intelligent Remote Sensing,* vol. 4, no. 5, pp. 55-63, 2024.

[2] F. A. F. L. &. H. C. H. Batarseh, "A survey on artificial intelligence assurance.," *Journal of Big Data, 8(1), 60.,* 2021.

[3] E. &. W. A. W. Glikson, "Human trust in artificial intelligence: Review of empirical research.," *cademy of Management Annals,* vol. 14, no. 2, pp. 627-660., 2022.

[4] N. D. S. J. C. M. d. P. M. L. H.-V. E. &. H. F. Díaz-Rodríguez, "Connecting the dots in trustworthy Artificial Intelligence: From AI principles, ethics, and key requirements to responsible AI systems and regulation.," *Information Fusion,* 2023.

[5] C. A. S. &. P. D. Sanchez Hernandez, "An explainable artificial intelligence (xAI) framework for improving trust in automated ATM tools.," 2021.

[6] M. A. E. A. Ö. K. S. I. T. S. I. &. B. I. Ozkan-Ozay, "A Comprehensive Survey: Evaluating the Efficiency of Artificial Intelligence and Machine Learning Techniques on Cyber Security Solutions.," *IEEE Access.,* 2024.

[7] D. Ajish, "The significance of artificial intelligence in zero trust technologies: a comprehensive review. Journal of Electrical Systems and Information Technology," *Springer.com,* 2024.

[8] S. A. T. E.-S. S. M. K. A.-M. J. M. C. R. .. &. H. F. Ali, "Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence.," *Information fusion. https://doi.org/10.1016/j.inffus.2023.101805,* 2023.

**Journal of Global Knowledge and Innovation (JGKI)**
Volume 1, Issue 1, October-2024
**ISSN 3079-1995**
**DOI:** https://doi.org/10.5281/zenodo.13885892

**Received: 13 August 2024** **Published: 3 October 2024**