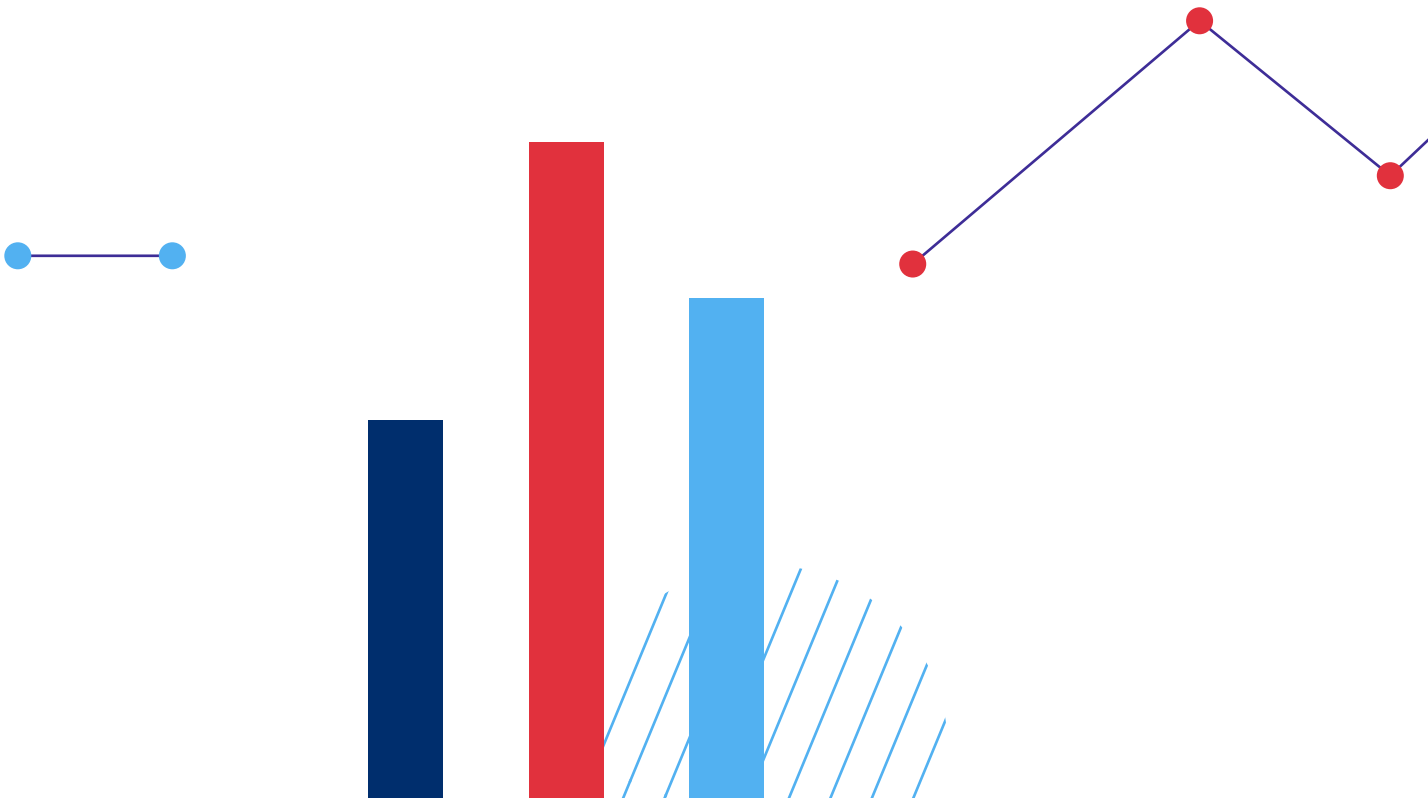
Responsible Artificial Intelligence Framework in Accountancy

By El'fred Boo and Lim Chu Yeong
October 2024



Responsible Artificial Intelligence Framework in Accountancy

Executive Summary

Based on a joint study by the Institute of Singapore Chartered Accountants (ISCA) and National Technological University (NTU), this report presents a framework for the responsible use of Artificial Intelligence (AI) in the accountancy profession. As AI, especially generative AI (GAI), continues to transform the accounting field by automating routine tasks and generating data-driven insights, its growing influence introduces both opportunities and risks. The objective of this study is to review the literature on AI-related risks to provide actionable and practical recommendations to guide the responsible deployment of AI in accountancy.

Introduction and Context

Generative AI (GAI) offers new possibilities for automating tasks and generating complex insights in accounting. However, these advancements bring risks such as over-reliance, data integrity issues, and transparency concerns. We thus propose a Responsible Artificial Intelligence Framework that addresses these challenges by outlining **six key principles** to promote the ethical use of AI in accountancy:

P#1 **Professional Judgement Oversight and Accountability:** Ensuring that AI does not replace human decision-making, but rather acts as a tool that requires constant oversight.

P#2 **Process Robustness and Output Quality:** Safeguarding AI systems from errors and ensuring reliable and reproducible outputs.

P#3 **Data Integrity and Privacy:** Maintaining the accuracy, reliability, and confidentiality of data used in AI systems.

P#4 **Transparency, Traceability, and Explainability:** Providing clarity about how AI decisions are made and ensuring stakeholders understand AI processes.

P#5 **Fairness and Stakeholder Inclusivity:** Preventing biases in AI outputs and ensuring the technology is accessible to all players, large and small.

P#6 **Work-Related Societal and Environmental Effects:** Addressing the broader social and environmental impacts of AI, such as its carbon footprint and potential workforce displacement.

Key Concerns and Response Measures

1. **Over-reliance on AI:** There is a risk that professional accountants may overestimate AI's reliability, leading to poor decision-making. The framework emphasizes the need for continuing professional development to ensure accountants understand AI's limitations and its potential errors, such as "hallucinations" (fabrication of false information). AI outputs must be reviewed and verified, and users should exercise professional judgment at all times.
2. **Algorithmic Robustness:** AI systems must be rigorously tested before deployment to ensure their reliability and accuracy. In cases where AI is repurposed for tasks beyond its original scope, it can produce unintended consequences. Continuous monitoring and updating of AI systems is necessary to ensure they remain robust in dynamic environments.
3. **Data Privacy and Integrity:** AI systems are trained on vast amounts of data, which could potentially include sensitive client information. This framework recommends obtaining client

consent before using data for AI training and employing Privacy Enhancing Techniques (PETs) to anonymize sensitive information. Furthermore, AI systems should be trained on trusted data sets to avoid inaccuracies and biases.

4. **Transparency and Explainability:** Many AI systems, especially those involving neural networks, operate as "black boxes" with limited transparency of their decision-making processes. This poses significant challenges for accountants who rely on AI to make critical financial decisions. To address this, the framework promotes the use of Explainable AI (XAI), which provides more transparency into how AI models arrive at decisions. As XAI is still developing, its reliability remains a key issue and a focus area for future research.

5. **Bias and Inclusivity:** AI systems can perpetuate biases present in the data used for training. The framework calls for a rigorous evaluation of AI algorithms to ensure fairness and inclusivity. In addition, it highlights the potential disparity between large and small accounting firms, with the former having easier access to AI technologies. The report recommends that regulators and professional bodies work towards making AI training and database more accessible to all firms.

6. **Societal and Environmental Impact:** AI's energy-intensive processes contribute to significant carbon emissions. This framework encourages the use of renewable energy sources and suggests that future research could focus on AI systems that track and report carbon emissions. Additionally, concerns about workforce displacement and "de-professionalization" are addressed by highlighting the evolving roles of accountants, emphasizing the importance of upskilling in data science and analytics.

Next Phase: Stakeholder Engagement and Framework Refinement

The next phase of this project involves conducting interviews with a diverse range of stakeholders, including AI developers, regulators, and professional accountants in both practice and business. These interviews aim to identify ethical challenges in applying AI in accountancy and explore practical solutions to overcome them.

The goal of this stakeholder engagement is to refine the Responsible AI Framework, ultimately creating practical guidelines for the ethical and effective deployment of AI in the accounting profession. This will culminate in a set of recommendations tailored to regulators, standard-setters, and accounting professionals to ensure AI is used responsibly and ethically across the industry.

Conclusion

The Responsible AI Framework proposed in this report provides a foundation for addressing the risks and opportunities of AI in accountancy. By focusing on professional accountability, data integrity, transparency and inclusivity, the framework offers a practical approach to managing AI's influence in the profession. As AI technology continues to evolve, ongoing stakeholder engagement and a dynamic framework will ensure that AI serves the interests of both the profession and society at large, fostering trust and innovation in accounting.

The upcoming interviews with key stakeholders in the next phase of this study will update the development of this framework, ensuring it remains relevant and practical as AI's role in accountancy grows. Ultimately, this initiative aims to support regulators, standard-setters, and accounting professionals in navigating the ethical and operational challenges of AI, leading to a more transparent, fair, and responsible use of this transformative technology.

Introduction & Background

The recent emergence of generative artificial intelligence (GAI) has opened up exciting and new possibilities in the way we all work and live. GAI's capacity for large data sets and complex analyses has spurred its development in various sectors and applications, presenting itself as a versatile technology that improves work productivity and reduces costs. GAI is also transforming the accountancy profession via its diverse applications in automating routine work (e.g., reconciliations, bookkeeping, compliance) and generating data-driven insights, analyses and forecasts. For a tool as powerful and versatile as GAI, there are potentials for its abuse and misapplications. There is thus a need for its development and deployment to be guided with a framework to ensure its responsible use for the benefits of the profession and humanity.

The objective of this study is to review the literature to identify key risks arising from the application of GAI in accountancy/finance and response measures to address them. Our review culminates in a framework for the responsible use of AI in accountancy. We value-add by weighing key risk issues in concrete and practical terms within six broad principles. As it is difficult to implement values-based principles, our recommendations go beyond principles and are proposed as actionable practical measures. What we do not cover in this report are information technology (IT) risks that would likewise apply to an AI system that is connected to the web, as these IT or cyber security risks have been commonly and well-discussed elsewhere. Our focus is on unique concerns and risks emerging from AI deployment in accountancy.

We will also highlight the more contentious and challenging issues for further deliberation in the next phase of our study. For this second part of our study, we plan to interview leading AI experts and professional accountants for their views to shed more light on these more intricate issues as we monitor and update the developments in AI technology.

As the field of AI is very dynamic, much of the limitations and constraints discussed in this report might soon be overcome with improved versions of the AI technology. For example, on the issue of AI "black box", OpenAI recently on 12 September 2024, launched its Strawberry series of AI models capable of "truly general reasoning" to solve more challenging problems than previous models by incorporating "chain-of-thought" reasoning, which breaks down complex problems into smaller logical steps (TST 2024b). Thus, the crucial need for us to continue monitoring and watching this space.

We welcome your comments to questions raised in this study, see Framework on pages 18-21.
To contribute your views and comments, please visit: https://forms.office.com/r/nT0pt1qUx5

In response to specific questions, please include the question reference, e.g., **Q1.1a, Q2.1b,** etc.

AI Principles in Accountancy

We reviewed the literature on trustworthy AI and the responsible use of AI to distil **key principles** commonly shared by frameworks from **diverse stakeholders** (e.g., government/national agencies, academic, professional bodies, accounting firms, supranational organisations such as OECD and UNESCO) and identified **six AI Principles** relevant to the accountancy profession:

P#1 Professional Judgement, Oversight & Accountability
P#2 Process Robustness & Output Quality
P#3 Data Integrity & Privacy
P#4 Transparency, Traceability & Explainability
P#5 Fairness & Stakeholder Inclusivity
P#6 Work-Related, Societal and Environmental Effects

Refer to **Appendix 3** for a tabulation of how the above principles that guide the identification of key AI risks and issues are **mapped to the various trustworthy AI frameworks** in the literature. We next discuss key **concerns** (**C1, C2** …) pertaining to AI deployment by the accountants and the corresponding recommended **response** measures to address these risks and concerns (**R1, R2** …).

P#1 Professional Judgement, Oversight & Accountability

Underpinning the responsible use of AI technology is the overriding principle of the exercise of **professional judgment, oversight and accountability**. Under no circumstances should a professional accountant delegate and surrender decision-making responsibility and the exercise of judgment to an AI system without quality assurance and, when required, necessary intervention.

C1.1 One foreseeable risk with the emergence of more advanced AI systems is the gradual **impairment to the accountants' professional judgement and expertise**. Relying on sophisticated AI systems could also lead to **misapplications and overinterpretation of the AI process results** (McKinsey 2023, p.6; Zhang et al. 2023), e.g., overestimating its accuracy and reliability rate. In the extreme scenario, human autonomy in decision-making could be undermined when users relegate and surrender their judgement and decision-making tasks to the AI systems.

R1.1a AI developers have a responsibility to document and highlight the proper use of AI, and its risks and limitations. Accountants, as users of the AI system, should **work closely with the AI developer to train end-users** on the appropriate use of AI. In deploying AI tools, accountants should have a **clear understanding** of how the tools work, the AI system's **risks and limitations** (e.g., algorithm reliability, training data integrity). Accountants also need **to keep up to date** with their knowledge and skills on the use of AI through continuous professional development (ICAEW 2024).

R1.1b The exercise of **professional judgment** does not cease with the use of AI systems. On the contrary, the accountants need to exercise **professional scepticism** as AI systems are prone to errors and "hallucinations" (i.e., fabrication of false facts). The AI system should be viewed as an intelligent collaborating tool, but whose outputs need to be assessed and verified, and not to be taken at face value. In mapping to the NIST model (refer to Appendix 2), accountants exercise their professional judgement at all stages from plan and design to final use.

Shared Responsibility by All

Safe and appropriate AI development and use requires that the accountability and responsibility of ensuring the correct AI use is collectively **shared among the regulators, system developers, and end users, including the accountants/auditors** (OECD 2024; Zhang et al. 2023). They should understand the areas of responsibility of each party in the AI value chain (refer to the various stages of the NIST model in Appendix 1).

Nonetheless in reality, each party in the AI value chain from the AI system developer to its end-users may disclaim responsibility and liability for unforeseen risks arising from the deployment and use of the AI system. The lack of clarity on accountability is demonstrated by a case in 2020 when Hong Kong tycoon Li Kin-kan sued Tyndaris Investments for the use of AI algorithms K1 that caused millions of dollars of losses (Futurism 2019).

In the accounting/audit world, mechanisms should be put in place with audit oversight to ensure responsibility and accountability for AI systems at all stages from plan/design to final use in the NIST model. In certain countries, legislation (e.g., Canada's Artificial Intelligence and Data Act, and the EU's AI Act) requires **documenting the proper functioning of the AI systems throughout their lifecycle** (OECD 2024). Accountants should be trained to understand the implications of developing, deploying or using AI tools on their accounting tasks and to be accountable for their

areas of responsibilities. Relevant departments within the organisation should be identified with responsibilities over quality of data, model training and model selection to work together to put in place good data accountability practices (IMDA & PDPC 2020a).

Human control over AI should allow accountants to be kept in the loop for decisions over the design of the AI model, its application, deployment, and operation. **The appropriate level of human involvement in AI-augmented decision-making should match the nature of tasks at hand**. For example, little or no human intervention is suited for high-volume, routine and structured tasks versus unstructured and unique decisions that could be improved with the involvement of subjective human judgement (IMDA & PDPC 2020b). The extent of giving accountants oversight over when and whether to use AI is based on their objective of using AI and acceptable risks. Auditors who develop and deploy their own AI system need **to seek the relevant regulators' inputs on whether the level of human oversight over AI systems is adequate**. Indeed two U.S. AI startups, OpenAI and Anthropic, have recently agreed to let the U.S. AI Safety Institute (part of the National Institute of Standards and Technology) test their new models before releasing them to the public (Field 2024). AI supervision provided by independent oversight bodies will help ensure responsible use of AI (OECD 2024).

P#2 Process Robustness & Output Quality

A **secure and robust AI system** ensures **algorithm processing integrity, validity and reliability** to produce high quality output. Robustness ensures that the AI system is working as intended in envisaged circumstances (IS0 24368 2022). It is critical that the results of AI systems are reproducible and reliable. A reliable AI system is one that works properly with a range of inputs and in a range of situations. Reproducibility describes whether an AI experiment exhibits the same behaviour when repeated under the same conditions.

C2.1 When an AI system is inadvertently repurposed for **tasks or applications beyond their original scope or intent**, unintended negative consequences could follow (OECD 2024). Such a risk, including "hallucinations", could be commonly envisaged given AI system's characteristics of adaptability and learning versatility. For more complex decision-making (e.g., those involving **ethics in unique situations**), an AI algorithm could potentially oversimplify factors that need to be considered, resulting in **inaccurate, unfair or inappropriate decisions**.

R2.1a Safeguards including **within-scope checking mechanisms and warning-to-users capability to ensure proper AI uses** should be built into the AI system and tested before rolling them out for public uses. AI algorithms to be **rigorously tested in various scenarios** that accurately reflect and address **real-life contexts/circumstances** for their robustness before deploying them. OpenAI and Anthropic, two leading AI startups, took a step further by agreeing to subject their new models to be tested by the U.S. AI Safety Institute before releasing them to the public, following increased concerns in the industry about AI safety and ethics (Field 2024). More recent breakthrough by Meta enabling an AI model to evaluate other AI models reliably offers a glimpse at a possible pathway towards building autonomous, **self-improving AI models** that can learn from their own mistakes, according to Meta researchers (TST 2024c).

Lebovitz et al. (2023) urge users of AI systems to assess the quality of an AI tool by scrutinising the quality of the "ground truth" used to train and validate it. They define ground truth as "information that is known to be true based on objective, empirical evidence" (page 7). **Validating the "correct" answer in the predictive model** is thus an important key step to ensure the quality of the AI model. Moser et al.'s (2023) cautionary remark on AI's limitations is timely when they highlight the danger of what is considered tentative and fluid from an ethical perspective becomes hardwired as the "correct" truth in an AI's algorithm.

R2.1b Apart from subjecting a percentage of AI outputs to **human review**, Lindebaum et al. (2023) suggested setting up a **feedback outlet** for aggrieved users to voice their concerns to be heard by humans, not machines as a safe measure **to correct any potential biases, failures and unintended consequences.**

R2.1c Enhanced transparency requirement would require an AI system to be **transparent about the confidence level or the accuracy/reliability level of its output**. This presents an avenue for future research if existing technology does not permit it.

C2.2 There are concerns raised on the robustness and reliability of AI algorithms and the continuing robustness of AI system in light of dynamic changes in economic, business and regulatory environment. As the **assumptions and relationships between key factors may evolve** over time and are not static,

an AI system trained on past data could result in inferior or poor decisions that are out of sync with the evolving **decision-making context and environment**.

R2.2 Accounting firms should develop competencies (or work with the AI developer) to monitor, review and revise AI systems as the decision-making context and environment evolves, including assessing if there is a need to **upgrade/update the AI algorithm and data.** An auto-update feature built into the AI system can detect error rate exceeding a pre-set threshold that take into account the shifting and evolving decision context and environment (e.g., to match the time frame vs. outdated data).

Continuous Monitoring of AI Robustness

An end-to-end enterprise governance needs to be in place for a responsible AI system to monitor continuing robustness in its results and decision-making, its alignment with the business strategy, its process modifications to improve outputs, its controls to track performance and problems, its consistency and reproducibility of results (PWC 2019). Leading governance practices include a multidisciplinary independent AI advisory board comprising members from ethics, law, philosophy, technology, privacy, regulations and science who report to the Board, AI ethical design standards, inventory of AI projects and impact assessments, AI validation tools, AI awareness training to employees and independent audits.

Impact assessments (e.g., red teaming or forms of Algorithmic Impact Assessment) both prior to and during the development, deployment and use of AI systems can help minimise its negative impact and ensures AI system resilience. These assessments must be proportionate to the risk that the AI systems pose (EU 2024a). IEEE is developing a certification system for AI focused on transparency, bias, etc. (CPA Canada 2019). In Singapore, AI Verify is a governance testing framework and software toolkit with testing reports that helps companies be more transparent about their AI to build trust (AI Verify Foundation 2023).

P#3 Data Integrity & Privacy

Data integrity relates to the completeness, accuracy and reliability of data, while **privacy** relates to keeping information safe from unauthorized access and alteration. As the quality of AI output hinges on the quality of its large volume of data input, data integrity is crucial to ensure the accuracy and reliability of its decision output.

C3.1 Accountants and auditors have a responsibility to safeguard **client data confidentiality** and **integrity**. There is a potential **risk of confidentiality breach** when an audit firm pools data from multiple clients for analysing industry/client risks using third-party AI platform - they could be inadvertently leaked and added to AI training data.

R3.1a Audit firms should seek **clients' permission** before using their data for AI training. Audit firms should consider the viability of training AI system based on segregated sub-sets of data pertaining to each audit client despite AI system thriving on training involving large data sets. Alternatively audit firms could tag data such that they could be used in a more versatile manner according to the level of data confidentiality or sensitivity (e.g., client economic data could be shared but not sensitive remuneration data of top executives). Emerging technologies known as **Privacy Enhancing Techniques (PETs)** can protect personal data by anonymizing sensitive data before using them as training data to develop AI models (MGF for GenAI 2024).

R3.1b Common safeguards of privacy include **limiting data** sourced, collected, used or disclosed to that necessary for accomplishing the intended purposes and tasks, and to authorised parties (ISO 24368 2022).

C3.2 While AI systems need to be trained with vast amount of data, such data harvested from the web are often **contaminated with inaccuracies, errors, biases and falsehoods**. Feeding malicious data into a self-learning AI system could negatively alter its behaviour with ensuing severe consequential effect (EU 2024a). One example of **AI poisoning** was Microsoft's Tay chatbot, which started posting offensive tweets within 24 hours in 2016.

R3.2a Use datasets, and AI system trained with datasets. from **trusted third-party sources that are certified** with appropriate data integrity/protection practices. Else, to require AI developers to document data provenance/lineage for accountability (IMDA & PDPC 2020b).

R3.2b Provide a **reporting hotline to the general public** to flag out inaccurate, biased and gibberish AI outputs. Due protection must be granted to whistle-blowers to promote the reporting of legitimate concerns about the AI system.

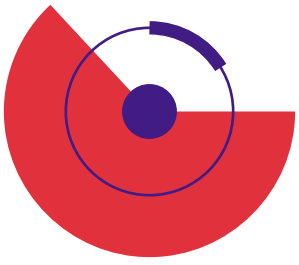
Data Governance

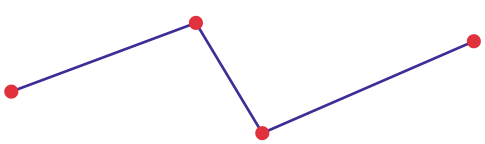
Audit firms must build strong security into creation of algorithms and governance of data, understand intent and context under which model was developed, identify who trained algorithms, know provenance of data and changes made to it, understand how models were served and protected, and maintain continuous review and confirmation of algorithm's effectiveness and accuracy (KPMG 2019).

Procedures to ensure that AI algorithms are sourced, designed and used to treat the individual in a transparent manner and ensuring data privacy are needed. Documentation and testing of processes

and data sets at each step such as planning, training, testing and deployment are needed. **Data protocols** governing data access should be in place within audit firms, outlining who can access data and under which circumstances (EU 2024a).

Understanding the lineage of data means knowing where the data originally came from, how it was collected, curated and moved within the organisation, and how its accuracy is maintained over time. Ensuring data quality means organisations understand and address factors that may affect the quality of data such as its accuracy (IMDA & PDPC 2020a). Common dimensions of privacy include limiting data sourced, collected, used or disclosed to that necessary for accomplishing the intended purposes and tasks, communication of the purpose of processing of personal identifiable information and any sharing of it (ISO 24368 2022). Audit firms must establish clear enterprise-wide policy about AI deployment including the harvesting and use of data and data privacy (KPMG 2019). Consumer privacy must be respected, customer data is not used beyond its intended/stated use and consumers should be given an option to opt out of personal data sharing (Deloitte 2021).
Respecting the confidentiality of client information means not loading confidential information for the training of Generative AI tools. Once uploaded, there is limited visibility and control over who the information is shared with, how it is secured and how long it is retained. This is important for ensuring compliance with data protection laws (ICAEW 2024).





P#4 Transparency, Traceability & Explainability

Transparency is about providing adequate disclosure about an AI system, including its intended uses, functional capabilities, risks and limitations. Organisations are encouraged to provide general information on whether AI is used in their products/services This includes information on what AI is, how AI is used in decision-making in relation to consumers, what are its benefits, why organisation has decided to use AI, how organisation has taken steps to mitigate risks, and the role and extent that AI plays in the decision-making process (NIST 2023). OECD (2024) further suggests easy-to-understand information on the factors and the logic that served as the basis for the prediction or decision.
Traceability involves leaving a documentary/digital trail to allow traceability of an entire AI lifecycle (EU 2024a, 2024b), including tracing an AI output to its data, algorithm and processes involved in generating the output.
Explainability is the ease of understanding to human users on how an AI system arrives at a decision, including the AI technical processes and the reasoning in support of the decision (EU 2024a, 2024b).

C4.1 AI involving neural network analyses operate within a **"black box" and are not easily explained**, particularly those involving more complex analyses and decision-making (TST 2024a). A trade-off between performance and explainability suggests that increasing the complexity of an AI model often contributes to higher accuracy but it reduces explainability (Frasca et al. 2024; Linardatos et al. 2020). As Roi (2022) puts it, "LLMs, by design, produce texts based on ideas generated by others without the user knowing what the exact sources were." Providing transparency, traceability and explainability, however, is crucial for accounting/finance users to justify how, for example, a credit or risk assessment, is arrived at using an AI system.

R4.1a For transparency, accounting firms to disclose the use of AI as a collaborating tool, along with its capabilities, risks, limitations and safeguard measures.

R4.1b The recent emergence of Explainable Artificial Intelligence (XAI) technology aims at overcoming the "black box" AI issue by generating additional explanations on how the model makes predictions (Cerevisiae and Kabasinskas 2024; Ribeiro et al. 2024). While initial evidence is encouraging and illuminating, most XAI methods are still not as stable and warrants close attention when applied to real-world settings (Ribeiro et al. 2024). Accountants whose analyses and decisions are aided by XAI system will need to **review and closely scrutinise the XAI output to ensure its reasonableness and reliability.**

Transparency, Traceability & Explainability – Why Important?

Transparency and explainability are important because people need to understand when they are interacting with an AI system, how it is making its decisions, and how it was designed and tested to ensure that it works as intended. ISO 26000 principle focuses on making sure an organization is transparent in its purposes and processes, while AI-specific principles focus on making sure an AI system is understandable in how it works. This includes disclosing traceability, information about algorithms, training data and user data, how it was collected; disclosing evaluation methods and metrics used to validate how a system works; explaining to stakeholders inputs that were used to reach a decision; explaining to stakeholders, as much as is possible, how an AI system arrived at a decision; notifying stakeholders when a decision about them is made by an AI system; notifying stakeholders when they are interacting with an AI system; consider allowing stakeholders to submit test cases to see how the AI system and application reacts to different situations. Accountability,

transparency and explainability are applied to decisions made at organizational and algorithmic level (ISO 24368, 2022).

To instil trust in AI systems, people must be enabled to look "under the hood" at their underlying models, explore the data used to train them, expose the reasoning behind each decision, and provide coherent explanations to all stakeholders in a timely manner. (PWC 2019). The purpose of being able to explain predictions made by AI is to build understanding and trust. An algorithm deployed in an AI solution is said to be explainable if how it functions and how it arrives at a particular prediction can be explained. When an algorithm cannot be explained, understanding and trust can still be built by explaining how predictions play a role in the decision-making process. There is legal obligation on business to make professional judgement whether system has made a mistake, which requires the explainability of AI-based decisions. Such judgements need traceability of factors that influenced the decision and also transparency concerning inner workings of algorithms behind decisions (Othmar et al. 2022 page 120).

The key issue with unexplainable AI is that it hinders the auditor's ability to understand and rely on AI, as it makes it challenging to document the rationale for the AI's decisions (Mukoko et al. 2020). Accountants should proactively ask questions of developers the AI algorithmic decision-making process and understand inputs and outputs of the AI systems. If the auditor using AI cannot understand its rationale for decisions or actions, how could they rely on the technology without impairing their due professional care? Consider an AI used to perform tasks such as sample selection and risk assessments. If the AI is a 'black-box,' it would be difficult for the auditor using the AI to justify the choice of specific samples or processes for testing. In such situations, auditors may exhibit automation bias and complacency, i.e. less scepticism and trust in the accuracy of the AI system (Parasuraman and Manzey 2010). As such, developing explainable AI (Samek et al. 2017) is a critical step towards ensuring an AI-enabled audit consistent with professional standards. Auditability, which allows an independent verification of the AI algorithms, data and design processes by internal and external auditors, can contribute to the trustworthiness of the AI technology (EU 2024a). ISACA (2018) requires that information should be traceable to originating business events and linked to accountable parties.

Organisations are encouraged to provide general information on whether AI is used in their products and/or services. This includes information on what AI is, how AI is used in decision-making, what are its benefits, why organisation has decided to use AI, how organisation has taken steps to mitigate risks, and the role and extent that AI plays in the decision-making process. (NIST 2023). Organizations need to exhibit transparent disclosure practices (AICPA & CIMA 2024). Organisations need to handle personally identifiable information and data as "contract of trust", give customers clarity and information that they want and inform public what decisions about transparent data mean (KPMG 2019).

AI system should be explicable and transparent in elements of data, system and business models. The data sets and the processes that yield the AI system's decision, including data gathering, data labelling and algorithms should be documented to allow for traceability and transparency. This also applies to the decisions made by the AI system. This enables identification of the reasons why an AI-decision was erroneous which, in turn, could help prevent future mistakes. Traceability facilitates auditability as well as explainability (EU 2024a, 2024b). Developing an intelligent system requires model selection and training. In order for the organisation to provide an account of the decisions, it needs to document how model training and selection processes are conducted, the reasons for which decisions are made, and measures taken to address identified risks. (IMDA & PDPC 2020a)

AI systems should not represent themselves as humans to users who have the right to be informed that they are interacting with an AI system. AI systems must be identifiable, with the option provided to decide against this interaction in favour of human interaction. AI system's capabilities/accuracy and limitations should be communicated to end-users (ISO 24368 2022; EU 2024).
When audit firms declare their use of AI systems, without accompanying disclosure into the actual capabilities and limitations of such systems, an expectation gap can emerge where stakeholders (auditors, clients, shareholders, and the public) have different expectations of the AI-enabled audit. In realizing that the use of AI can result in such ethical dilemmas, the European Commission's High-Level Expert Group on Artificial Intelligence observed there is a need for AI-implementing firms to "provide, in a clear and proactive manner, information to stakeholders (customers, employees, etc.) about the AI system's capabilities and limitations, allowing them to set realistic expectations" (HLEG 2018; Munoko et al. 2020).

Whenever an AI system has a significant impact on people's lives, it should be possible to demand a suitable explanation of the AI system's decision-making process. Such explanation should be timely and adapted to the expertise of the stakeholder concerned (e.g., layperson, regulator or researcher). In addition, explanations of the degree to which an AI system influences and shapes the organisational decision-making process, design choices of the system, and the rationale for deploying it, should be available (hence ensuring business model transparency).

If the public is to trust AI, they need more information on the AI models being used. The audit and assurance profession can help to provide that information and play a vital role in building trust. It can advocate that organisations and boards set the right tone at the top on AI adoption (considering issues such as fairness and transparency) and that they deploy AI models that deliver sustainable, long-term value. ISACA (2018) recommends that auditors focus on transparency in controls and governance through an iterative process. Auditors also play a role in ensuring that organisations are complying with regulation and ethics policies and managing their data appropriately.

Assurance techniques, technical standards and regulation support the development and implementation of trustworthy AI. Assurance techniques to measure, evaluate and communicate the trustworthiness of AI systems across the development and deployment life cycle' include impact assessment, audit, and performance testing, as well as formal verification methods (Department for Science Innovation & Technology 2024). An auditing framework has been developed by Google to ensure accountability in developing and deploying AI systems (Raji et al. 2020).

P#5 Fairness & Stakeholder Inclusivity

In line with the principles of fairness and inclusivity, the lifecycle of an AI system (from development, training to deployment) should be **free from biases**. AI technology should also be equally accessible to all players in the industry.

C5.1 An AI system **trained on incomplete or biased dataset can perpetuate biases** in its decisions, e.g., fraud risk deemed to be higher for certain groups, race, demographics. AI could generate biased results driven by the training data or in-built bias within the algorithm (TST 2024a). It could cause insidious harm to certain stakeholders if left undetected for extended time period.

R5.1 To alleviate any biases, accounting firms should **evaluate AI algorithms regarding its fairness, inclusivity and potential biases**. AI analyses and outputs should also be reviewed for similar purpose.

C5.2 The issue of fairness also applies to how AI technology shapes the competitive environment in the accounting profession. The audit industry is dominated by the Big Four firms with large financial and human resources who can more readily deploy AI technology in their operations. **Easier access to AI can potentially lead to significant gains in efficiency and effectiveness for these large firms, providing them a competitive edge over smaller firms.** On the other hand, would the cost of AI significantly drop, enabling smaller accounting firms to leverage on the technology and become equally competitive? (Munoko et. al. 2020). This issue is consistent with principle of accessibility and universal design (EU 2024a, 2024b).

R5.2 The regulators and industry professional bodies could consider making the competitive landscape more equitable to smaller audit firms. They could **promote and level up AI training to all players**. Combined efforts could also be channelled towards **developing a common AI training database that is reliable and accurate.**

P#6 Work-Related, Societal and Environmental Effects

AI deployment could introduce several unintended consequences to the wider community and society, each of which warrants our attention:

- Increased expectation gap
- Environmental effect from increased emissions
- Work isolation and displacement

Our key message, which is consistent with that from the literature, is this: Contrary to the misconception that AI is a threat to humanity, **human-centred AI is designed to augment human to perform best at what they can humanly deliver** with the assistance from AI (McKinsey & Company 2023). Thus, instead of viewing AI as an autonomous, mysterious black-box, human-centred AI can be user-friendly and accessible to a broad spectrum of users to enable **human and AI to operate synergistically at a higher level of performance with proper safeguards in place**.

C6.1 Given the powerful capabilities of AI technology, financial statement users' expectation of the auditors' responsibilities and capabilities could rise further, **widening the expectation gap** between what the users perceive the auditor should deliver and what the auditors could actually deliver (e.g., better prediction of company going concern and fraud detection capability).

R6.1 Transparency about the AI system's strengths, limitations and risks will help users and other stakeholders to appropriately understand and appreciate AI system's capability and calibrate their expectation accordingly.

C6.2 AI system, which processes large volume of data, is **energy-intensive and emits large volume of greenhouse gases.**

R6.2 Tap from **renewable energy sources**. As responsible deployment of AI system will also need to consider environmental sustainability, a potential area for future research involves using an **AI system in conjunction with other technology (e.g., blockchain) to measure and auto-track carbon emissions.** Potential benefits include reduced errors and increased efficiency across the entire value chain, enabling more timely reporting and spurring higher accountability and efforts towards further reducing carbon footprints.

C6.3 Potential negative effects of AI on the accountancy sector workforce include **the replacement of humans by AI and use of flawed AI in recruitment**, which is inequitable and cause negative social effects. Deskilling of accounting profession from replacement of accounting jobs by AI also discourage people from joining the profession – akin to "de-professionalization" (Munoko et al., 2020).

R6.3 The above two concerns are **myths that need to be debunked**. Organisations should communicate openly on how its **responsible AI, built on principles of fairness and inclusivity**, is used to recruit employees. Accounting professional roles will evolve such that **mundane tasks** (e.g., bookkeeping and reconciliation) **are replaced by skills that require data science knowledge and data analytics expertise**. This will enrich the accounting job scope and increase the attractiveness of the profession.

Responsible Artificial Intelligence Framework in Accountancy

A summary of the key concerns of AI deployment in accountancy and the corresponding response measures is presented below. They form our Responsible AI Framework in Accountancy.

Principles	Concerns/Issues	Response Measures	Questions
P#1 Professional Judgement, Oversight & Accountability	**C1.1** Over-reliance by users on AI, leading to its misapplications and overinterpretations of results.	**R1.1a** AI developer to flag out AI limitations. Users to work with AI developer to train end-users on the appropriate use of AI. **R1.1b** Users to exercise professional judgment and scepticism, viewing AI system as a collaborating tool, whose outputs should be assessed and verified.	"Market beware" model implies that the responsibility lies with users on the appropriate use of AI since AI developers can flag out but cannot be expected to highlight an exhaustive list of AI limitations. **Q1.1a** Is the existing "market beware" model sufficient? **Q1.1b** Suggest alternative feasible measures to counter unintended consequences arising from the misuse of AI.
P#2 Process Robustness & Output Quality	**C2.1** AI may "hallucinate" when repurposed for tasks beyond their original scope or intent. AI may oversimplify complex problems to produce inappropriate decisions.	**R2.1a** Rigorously test AI system before deployment, including validating the "correct" truth. **R2.1b** Test-review outputs and host a feedback channel for aggrieved users. **R2.1c** Provide confidence or accuracy level on AI's output.	**Q2.1a** If AI can reliably provide confidence or accuracy level on its output, what do you think is the threshold acceptable to users? Explain. **Q2.1b** Do you envisage an AI system that could reliably auto-detect and call out an error rate exceeding a pre-set threshold? **Q2.1c** Besides risks such as AI overreliance and loss of

	C2.2 Continuing AI robustness may be compromised in light of dynamic changes in the environment.	**R2.2** Accountants to develop competencies or work with AI developer to monitor and upgrade AI system.	judgement, what other risks should we guard against when an AI system can reliably auto-correct and auto-upgrade itself?
P#3 **Data Integrity & Privacy**	**C3.1** Client data can potentially be leaked into AI training data. **C3.2** Data to train AI system can be contaminated, churning output that can have consequential negative and severe impact.	**R3.1a** Obtain client's permission or use Privacy Enhancing Techniques (PETs) to anonymise personal data before using them as AI training data. **R3.1b** Limit data sourced, collected, used or disclosed to that necessary for accomplishing the intended purposes and tasks. **R3.2a** Use datasets, and AI system trained with datasets, from trusted third-party sources that are certified. Else, to require AI developers to document data provenance/lineage for accountability. **R3.2b** Provide a reporting hotline to the general public to flag out inaccurate, biased and gibberish AI outputs.	**Q3.1** New technology, such Privacy Enhancing Techniques (PETs), anonymises personal data before using them as AI training data. Do you think that audit clients would agree to using their corporate data for AI training if their data is first anonymised using PET? Explain. **Q3.2** Would you be comfortable with the accounting firms and accountants using AI systems that are not trained with certified datasets (on the basis data are harvested on "fair use" basis, market practice, and/or other reasons yet to be clarified in courts of law)? Explain.
P#4 **Transparency, Traceability & Explainability**	**C4.1** AI involving neural network analyses operate within a "black box" and are not easily explained.	**R4.1a** For transparency, accounting firms to disclose the use of AI as a collaborating tool, along with its capabilities, risks, limitations and safeguard measures. **R4.1b** Explainable Artificial Intelligence (XAI) technology	**Q4.1** While XAI (Explainable AI) research efforts are on-going, do you foresee a feasible, reliable and stable model to emerge within the next two years? Explain.

		aims at overcoming the "black box" AI issue by generating additional explanations on how the model makes predictions but its stability is still an issue. Accountants whose analyses and decisions are aided by XAI system will need to review and closely scrutinise the XAI output to ensure its reasonableness and reliability.	
P#5 Fairness & Stakeholder Inclusivity	**C5.1** An AI system trained on incomplete or biased dataset can perpetuate biases in its decisions. **C5.2** Easier access to AI can potentially lead to significant gains in efficiency and effectiveness for large firms, providing them a competitive edge over smaller firms.	**R5.1** To evaluate AI algorithm and its outputs on the issues of fairness, inclusivity and potential biases. **R5.2** The regulators and professional bodies can promote and level up AI training to all players and jointly develop a shared AI training database that is reliable and accurate.	**Q5.1** Would users be able to evaluate AI algorithm and review its outputs for potential biases, even with appropriate training? Explain. **Q5.2** Is the proposal to develop a shared AI training database feasible? Explain and highlight the hurdles that need to be cleared.
P#6 Work-Related, Societal and Environmental Effects	**C6.1** Given powerful capabilities of AI, users' expectation of the auditors' duties and capabilities could rise further, widening the expectation gap. **C6.2** AI system is energy intensive and generates large volume of carbon emissions.	**R6.1** Transparency about the AI system's strengths, limitations and risks can moderate users' expectation. **R6.2** Tap from renewable energy sources. A potential area for future research involves using AI and blockchain to measure, auto-track and report carbon emissions.	**Q6.1** Would transparency/ disclosure about AI's limitations be adequate to moderate users' expectation? Any other effective measures? **Q6.2** Do you think research leveraging on technology (e.g., AI, blockchain) to measure, auto-track and report carbon emissions should be given high priority? What do you think are the facilitating factors

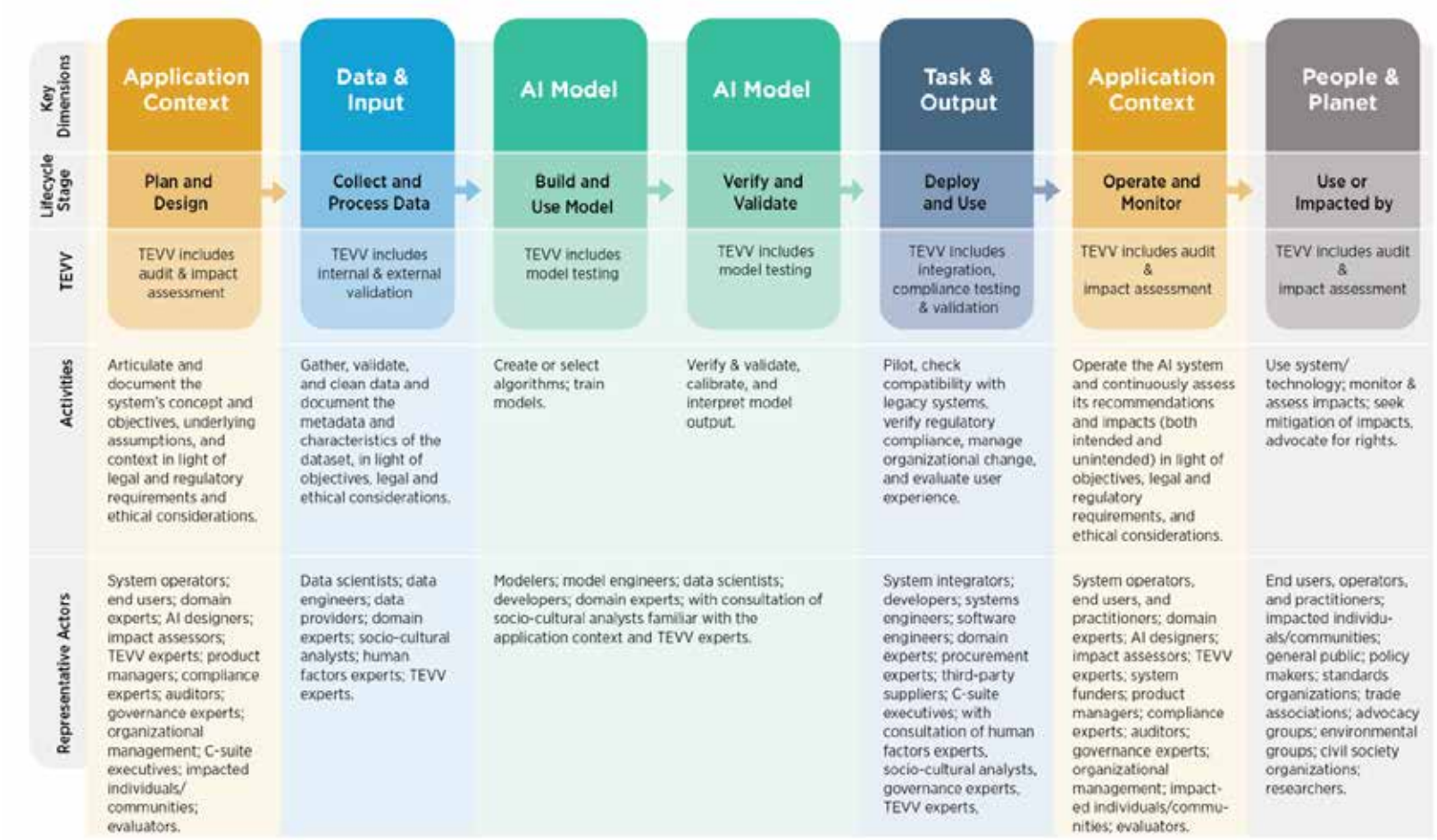
			and potential roadblocks?
	C6.3 Potential negative effects of AI on the accountancy sector workforce include replacement of humans by AI and use of flawed AI in recruitment, which is inequitable and cause negative social effects.	**R6.3** Organisations should communicate openly on how its responsible AI, built on principles of fairness and inclusivity, is used to recruit employees. Accounting professional roles will evolve such that mundane tasks, such as bookkeeping and reconciliation, are replaced by skills that require data science knowledge and data analytics expertise. This will enrich the accounting job scope and increase the attractiveness of the profession.	**Q6.3** Do you envisage AI applications in accountancy to increase the attractiveness of the profession in talent recruitment? Explain.

We welcome your comments and input to the questions raised in this study.

To contribute your views and comments, please visit: https://forms.office.com/r/nT0pt1qUx5

In response to specific questions, please include the question reference, e.g., **Q1.1a, Q2.1b,** etc.

Appendix 1



Key Dimensions	Application Context	Data & Input	AI Model	AI Model	Task & Output	Application Context	People & Planet
Lifecycle Stage	Plan and Design	Collect and Process Data	Build and Use Model	Verify and Validate	Deploy and Use	Operate and Monitor	Use or Impacted by
TEVV	TEVV includes audit & impact assessment	TEVV includes internal & external validation	TEVV includes model testing	TEVV includes model testing	TEVV includes integration, compliance testing & validation	TEVV includes audit & impact assessment	TEVV includes audit & impact assessment
Activities	Articulate and document the system's concept and objectives, underlying assumptions, and context in light of legal and regulatory requirements and ethical considerations.	Gather, validate, and clean data and document the metadata and characteristics of the dataset, in light of objectives, legal and ethical considerations.	Create or select algorithms; train models.	Verify & validate, calibrate, and interpret model output.	Pilot, check compatibility with legacy systems, verify regulatory compliance, manage organizational change, and evaluate user experience.	Operate the AI system and continuously assess its recommendations and impacts (both intended and unintended) in light of objectives, legal and regulatory requirements, and ethical considerations.	Use system/ technology; monitor & assess impacts; seek mitigation of impacts, advocate for rights.
Representative Actors	System operators; end users; domain experts; AI designers; impact assessors; TEVV experts; product managers; compliance experts; auditors; governance experts; organizational management; C-suite executives; impacted individuals/ communities; evaluators.	Data scientists; data engineers; data providers; domain experts; socio-cultural analysts; human factors experts; TEVV experts.	Modelers; model engineers; data scientists; developers; domain experts; with consultation of socio-cultural analysts familiar with the application context and TEVV experts.		System integrators; developers; systems engineers; software engineers; domain experts; procurement experts; third-party suppliers; C-suite executives; with consultation of human factors experts, socio-cultural analysts, governance experts, TEVV experts.	System operators, end users, and practitioners; domain experts; AI designers; impact assessors; TEVV experts; system funders; product managers; compliance experts; auditors; governance experts; organizational management; impacted individuals/communities; evaluators.	End users, operators, and practitioners; impacted individuals/communities; general public; policy makers; standards organizations; trade associations; advocacy groups; environmental groups; civil society organizations; researchers.

Fig. 3. AI actors across AI lifecycle stages. See Appendix A for detailed descriptions of AI actor tasks, including details about testing, evaluation, verification, and validation tasks. Note that AI actors in the AI Model dimension (Figure 2) are separated as a best practice, with those building and using the models separated from those verifying and validating the models.

Source: NIST AI 100-1 AI Risk Management Framework-AIRMF 1.0-Jan2023.pdf

Appendix 2: Mapping of AI Principles in Accountancy to AI Lifecycle

Dimensions	**Application Context**	**Data & Input**	**AI Model**		**Task & Output**	**Application Context**	**People & Planet**
Lifecycle	**Plan & Design**	**Collect/Process Data**	**Build Model**	**Validate Model**	**Deploy/Use**	**Operate & Monitor**	**Use/Impacted by**
Requirements							
P#1 Professional Judgement, Oversight & Accountability	**R1.1a R1.1b**	**R1.1a R1.1b**	**R1.1a R1.1b**	**R1.1a R1.1b**	**R1.1a R1.1b**	**R1.1a R1.1b**	**R1.1a R1.1b**
P#2 Process Robustness & Output Quality	**R2.1a R2.1c R2.2**	**R2.1a R2.2**	**R2.1a R2.1c R2.2**	**R2.1a R2.1c R2.2**	**R2.1b R2.1c R2.2**	**R2.1b R2.2**	**N.A.**
P#3 Data Integrity & Privacy	**N.A.**	**R3.1a,b R3.2a,b**	**R3.1a R3.2a**	**R3.1a R3.2a**	**N.A.**	**R3.2b**	**N.A.**
P#4 Transparency, Traceability & Explainability	**N.A.**	**N.A.**	**N.A.**	**N.A.**	**R4.1a**	**R4.1b**	**R4.1a**
P#5 Fairness & Stakeholder Inclusivity	**R5.1**	**R5.1 R5.2**	**R5.1**	**R5.1**	**R5.1**	**R5.1**	**R5.1**
P#6 Work-Related, Societal and Environmental Effects	**R6.2**	**N.A.**	**N.A.**	**N.A.**	**N.A.**	**N.A.**	**R6.1 R6.2**

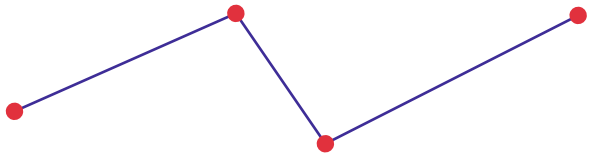
Professional accountants' deployment of the AI technology comes with some obligations, though less than the developers of the AI technology (AI Act 2024).

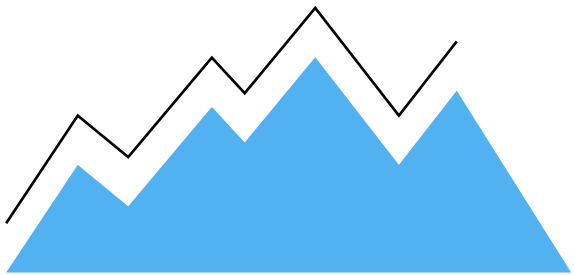
Appendix 3: Mapping of AI Principles in Accountancy to Major Frameworks

Government and National Agencies

AI Principles in Accountancy	Ethics Guidelines for Trustworthy AI (2019) by the EU's HLEG on AI (2024a)	EU AI Act (2024b) http://surl.li/qyfii	NIST (2023)	IMDA & PDPC (2020a, 2020b)	AI Verify Foundation (2023)
P#1 Professional Judgement, Oversight & Accountability	REQUIREMENT #1 Human Agency, Autonomy & Oversight REQUIREMENT #7 Accountability: Auditability, Risk Management	Human Oversight (Article 14) Risk Management System (Article 9) Quality Management System (Article 17)	Accountable and Transparent Page 15	Human intervention, human review, decision-making; human centricity Page 38, 57, 59, 62, 65, 4 Accountability (for data), Responsibility Pages 17, 36, 40, 64, 66, 67	Management and Oversight of AI. Accountability, Human Agency and Oversight.
P#2 Process Robustness & Output Quality	REQUIREMENT #2 Technical Robustness and Safety: Resilience to Attack & Security, General Safety, Accuracy, Reliability, Fall-back plans, Reproducibility	Technical Documentation (Article 11 & Annex IV) Accuracy, Robustness and Cybersecurity (Article 15)	Valid and reliable; Safe; Secure and Resilient Pages 13-15	Safety, reliability, quality, robustness, inaccurate model Page 5, 15-16, 21, 24, 31, 43, 47, 48	Safety and Resilience of AI systems. Safety, Security, Robustness.
P#3 Data Integrity & Privacy	REQUIREMENT #3 Privacy & Data Governance	Data & Data Governance (Article 10)	Privacy-enhanced Page 17	PDPA Privacy Pages 13, 17	Data Governance.
P#4 Transparency, Traceability & Explainability	REQUIREMENT #4 Transparency: Traceability, Explainability, Communication	Record Keeping (Article 12) Transparency & provision of information to user (Article 13)	Accountable and Transparent; Explainable and Interpretable Page 15-16	Decision-making is transparent, explainable and fair; technical explainability, repeatability, traceability, easy to understand communications, auditability Page 15,45-48, 51, 54, 57, 66	Understanding how AI reaches decision. Explainability, Repeatability/ Reproducibility.

					Transparency on use of AI systems.
P#5 Fairness & Stakeholder Inclusivity	REQUIREMENT #5 Diversity, Non-discrimination, Fairness: Avoidance of Unfair Bias, Accessibility & Universal Design, Stakeholder Participation	Fundamental Rights Impact Assessment	Fair with harmful bias managed Page 17	Biased model, dataset Page 24, 36	Fairness/No unintended discrimination.
P#6 Work-Related, Societal and Environmental Effects	REQUIREMENT #6 Societal & Environmental Well-being: Impact on Work and Skills, Impact on Society at large	Fundamental Rights Impact Assessment	Not mentioned	Stakeholder interaction and communication Page 20, 43, 51, 53	Inclusive Growth, Societal and Environmental Well-being





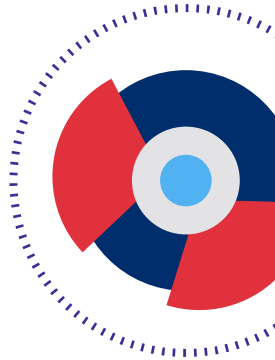
Academic Literature

AI Principles in Accountancy	Munoko et al. (2020)	Zhang et al. (2023)	Othmar et al. (2022)	Huang et al. (2023)	Toth et al. (2022)	Bushmann & Fieseler (2023)
P#1	Accountability, Responsibility gap Table 1 page 212, Figure 2 page 217, 219, Table 5 page 224, 225, 226	Power over user Table 3 page 7, 13-14 Accountability in AI development and using AI Table 3 page 6, 8, 13	Human-machine relationship, humanisation of AI Table 4 Page 118, 123-124 Accountability, normative accountability, designers' accountability, use of experts, public accountability, responsibility and auditability Table 4 Page 118, 121-123	Human rights and Dignity, Freedom and autonomy, Control Table 1 Page 802-803, 805-806 Responsibility & Accountability Table 1 Page 802, 804,806	Less human touch Page 904 Accountability clusters Page 899-902	Respect human autonomy Page 154 Responsible innovation Page 154
P#2	Safety and nonmaleficence Table 3 page 217, 219	Result distortion Table 3 page 6, 11	Human Design Flaws, Cybercrime and fraud, value-laden algorithms Table 4 Page 118-119	Safety, Abuse use of AI Table 1 Page 802, 806	Stealing data from competitors with AI robots Page 904	AI non-malfeasance: "avoiding harm from AI" Page 153-154

P#3	Privacy, Confidentiality, Data Protection Table 1 page 212, Table 3 page 217, 219, 224, 225	Data Security, Privacy and Misuse Table 3 page 6-7	Privacy, Data protection regulation, GDPR, Sensitivity of data Table 4 Page 118-120	Privacy and data protection Table 1 Page 802, 804, 806	Privacy breach through disorderly design Table 2 Page 904	Not mentioned
P#4	Lack of Transparency into AI to perform oversight. Table 3 page 217, 224, Table 5 page 225	Transparency and Trust Table 3 page 6, 10-11	Transparency, Decision-making path, black-boxes Table 4 Page 118, 120-121	Transparency, algorithm explainability Table 1 Page 802-803, 805-806	Not mentioned	Not mentioned
P#5	Algorithm & training data bias, propagate human bias Table 1 page 210, 212, 213, 220	Bias – AI algorithms pass experts' bias to managerial accountants Table 3 page 6, 11	Objectivity, Bias, Neutrality, Discrimination Table 4 page 118-119	Bias, fairness and justice Table 1 Page 802, 804-806	Algorithm injustice Page 904	Justice and fairness; minimize discrimination and bias Page 154
P#6	Unintended consequences for users and other stakeholders; expectation gap between stakeholders; Lack of stakeholder acceptance of AI. Table 2 page 215, 223, 224, Table 5 page 225 Isolation of human Table 3 page 217, 219, 222, 224	Not mentioned	Not mentioned	Not mentioned	Not mentioned	Not mentioned

Academic Literature

AI Principles in Accountancy	Bankins & Formosa (2023)
P#1	Autonomy page 729, 733-736 Explicability (explainability & accountability) page 736-737
P#2	Not mentioned
P#3	Not mentioned
P#4	Not mentioned
P#5	Justice page 736-737
P#6	Task integrity, skill cultivation and use, task significance, belongingness; beneficence & non-maleficence page 729-736



Professional Bodies & Accounting Firms

AI Principles in Accountancy	CPA Canada (2019)	ACCA & ICAANZ (2021)	IESBA & APESB (2023)	AICPA & CIMA (2024)	ICAEW (2024)
P#1	Ethos: Should we build it?	Professional behaviour, professional judgement, professional standards, professional due care, Integrity; human-centred AI; human oversight Tables 1 & 1.1, pages 10, 12, 29 Accountability Tables 1, Page 10, 29, 31, 37,	Prohibition on assuming management responsibilities including certain IT services Page 4	Financial analysis and manipulation; psychological and emotional impact Accountability and responsibility	Professional due care; professional behaviour
P#2	Safety: Are we building it safely?	Safety and robustness Page 29	Frequency of services and provision of insights Page 7	Data reliability	Not mentioned
P#3	Not mentioned	Confidentiality, Ethical use of data (privacy & confidentiality) Tables 1 & 1.1, pages 10, 12, 29	Not mentioned	Privacy and data security	Confidentiality
P#4	Not mentioned	Transparency Page 29, 33	Not mentioned	Transparency	Integrity. Transparent and honest in use of AI.
P#5	Fairness: How do we build it?	Objectivity, Be a Responsible Computer Provider, Managing and reporting of environmental footprint, Rights of individual, employee, consumer; the public interest; Fairness Tables 1 & 1.1, pages 8, 10, 12, 16-19, 21, 29	Technology and Confidentiality Page 7	Bias and discrimination	Objectivity. Bias in use of AI.

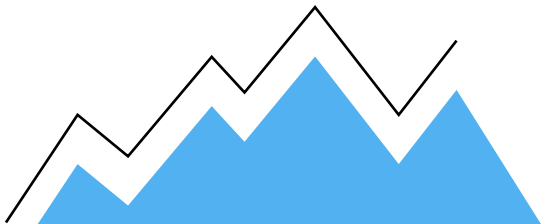
P#6	Justification: What do we build?	Professional competence; Sustainability; Standards and Law Tables 1 & 1.1, pages 8-10, 12, 15, 17, 29	Not mentioned	Intellectual property; Regulatory compliance; Job displacement and economic impact	Professional competence

Professional Bodies & Accountancy Firms

AI Principles in Accountancy	PWC (2019)	Deloitte (2021)	KPMG (2019)	ACCA & EY (2023)
P#1	Ethics and regulation Page 10 Governance. Accountability Page 8	 Responsible/Accountable Pages 2-3	Transforming the workplace Page 4-6 Governance and oversight - accountability Page 8	Not mentioned
P#2	Robustness and security; Governance – consistent, reproducible Page 12	Robust/reliable; Safe/ secure Pages 2-3	Oversight and governance: use of data; Cybersecurity Page 7-8, 10-11	Cybersecurity; Inaccurate and unreliable output Page 11
P#3	Security Page 12	Privacy Pages 2-3	Oversight and governance: standards of privacy Page 7-8	Copyright; Privacy Page 11
P#4	Interpretability and explainability Page 11	Transparent/Explainable Pages 2-3	Governance and explainability; transparency, auditability. Page 6, 12	Transparency Page 11
P#5	Bias and fairness Page 13	Fair and impartial Pages 2-3	Mitigating Bias Page 12	Bias and discrimination Page 11
P#6	Ethics and regulation Page 10	Not mentioned	Transforming the workplace Page 4-6	Disinformation, Environment, Health; Unknown unknowns Page 11

Supranational Organisations

AI Principles in Accountancy	ISO 24368 (2022)	UNESCO (2021)	OECD (2019)
P#1	Professional responsibility and human control of technology; Human centred design Page 10 sect 6.2.4; page 11 section 6.2.8; page 13 sect 6.2.10 Accountability Page 8 sect 6.2.1	Human oversight and determination Page 12-13 Responsibility and accountability Page 14	Principle 1.2 Human rights Principle 1.5 Accountability
P#2	Safety and security Page 11 sect 6.2.7	Safety and security Page 11	Principle 1.4 Robustness, security, safety
P#3	Privacy Page 11 sect 6.2.6	Right to Privacy, and Data Protection. Page 12	Principle 1.2 Privacy
P#4	Transparency and explainability Page 9 sect 6.2.3	Transparency and explainability Page 13	Principle 1.3 Transparency, explainability
P#5	Fairness and non-discrimination Page 9 sect 6.2.2	Fairness and non-discrimination Page 11-12	Principle 1.2 Fairness
P#6	Promotion of human values; Community involvement and development; Respect for rule of law; Respect for international norms of behaviour; environmental sustainability; labour practices Page 10 sect 6.2.5; page 12 sect 6.2.9; page 13-14, sect 6.2.11, 6.2.12, 6.2.13, 6.2.14	Proportionality and do no harm; Sustainability; Awareness and literacy; Multi-stakeholder and adaptive governance Page 11-12, 14	Principle 1.1 People and planet



References

ACCA (The Association of Certified Chartered Accountants), & ICAANZ (Chartered Accountants Australia & New Zealand). (2021). Ethics For Sustainable AI Adoption. Connecting AI and ESG.

ACCA (The Association of Certified Chartered Accountants), & EY (Ernst & Young). (2023). Building The Foundations For Trusted Artificial Intelligence. URL: https://assets.ey.com/content/dam/ey-sites/ey-com/en_gl/topics/ai/ey-pi-ai-regulation.pdf

AI Verify Foundation. (2023). What is AI Verify? URL: https://aiverifyfoundation.sg/what-is-ai-verify

AICPA & CIMA. (2024). Ethics in the World of AI: A CPA's Guide to Managing the Risks. Webcast URL: https://www.aicpa-cima.com/cpe-learning/webcast/ethics-in-the-world-of-ai-a-cpas-guide-to-managing-the-risks

Bankins, S., & Formosa, P. (2023). The Ethical Implications of Artificial Intelligence (AI) For Meaningful Work. *Journal of Business Ethics* 185, 725-740.

Bushmann, A., & Fieseler, C. (2023). Deep Learning Meets Deep Democracy: Deliberative Governance and Responsible Innovation in Artificial Intelligence. *Business Ethics Quarterly* 33(1), 146-179.

Černevičienė, J., Kabašinskas, A. (2024). Explainable artificial intelligence (XAI) in finance: a systematic literature review. *Artif Intell Rev* **57**, 216. https://doi.org/10.1007/s10462-024-10854-8

CPA Canada. (2019). Building Ethical AI solutions: using the ethics funnel and a trusted framework. URL: https://www.cpacanada.ca/business-and-accounting-resources/other-general-business-topics/information-management-and-technology/publications/building-ethical-ai-solutions

Deloitte. (2021). *AI governance for a responsible, safe AI-driven future*. URL: https://www2.deloitte.com/content/dam/Deloitte/us/Documents/risk/us-ai-governance-for-a-responsible-safe-ai-driven-future-final.pdf

Department of Science Innovation & Technology. (2024). *Introduction to AI assurance*. February 2024.

EU (European Commission). (2024a). *Ethical guidelines for trustworthy AI*. URL: https://digital-strategy.ec.europa.eu/library/ethics-guidelines-trustworthy-ai

EU (European Commission). (2024b). AI Act. URL: https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence Impact assessment of the regulation

Field, H. (2024). OpenAI and Anthropic agree to let U.S. AI Safety Institute test and evaluate new models. CNBC. Aug 29, 2024. https://www.cnbc.com/2024/08/29/openai-and-anthropic-agree-to-let-us-ai-safety-institute-test-models.html

Frasca, M., La Torre, D., Pravettoni, G., & Cutica, I. (2024) Explainable and interpretable artificial intelligence in medicine: a systematic bibliometric review. Discov Artif Intell 4(1):15

Futurism. (2019). Investor Sues After an AI's Automated Trades Cost Him $20 Million. URL: https://futurism.com/investing-lawsuit-ai-trades-cost-millions

HLEG (High-level Expert Group on Artificial Intelligence of the European Commission). (2018). *A Definition of AI: Main Capabilities and Disciplines*.

Huang, C., Zhang, Z., Miao, B., & Yao, X. (2023). An Overview of Artificial Intelligence Ethics. *IEEE Transactions on Artificial Intelligence* 4(4), 799-819.

ICAEW (Institute of Charted Accountants England & Wales). (2024). Generative AI and Ethics. URL: https://www.icaew.com/technical/technology/artificial-intelligence/generative-ai-guide/ethics

IESBA (International Ethics Standards Board for Accountants) & APESB (Accounting Professional Ethical and Standards Board). (2023). *Applying The Code's Conceptual Framework To Independence: Practical Guidance For Auditors In Technology-related Scenarios.*
Infocomm Media Development Authority. Personal Data Protection Commission Singapore. (2020a). *Model Artificial Intelligence Governance Framework*. Second Edition.
IMDA (Infocomm Media Development Authority) & PDPC (Personal Data Protection Commission Singapore). (2020b). *Companion to the Model AI Governance Framework – Implementation and Self-Assessment Guide for Organizations.*
ISACA. (2018). *Auditing Artificial Intelligence*. URL: https://transformingaudit.isaca.org/featured-articles/auditing-artificial-intelligence
ISO 24368. (2022). Information Technology – Artificial Intelligence – Overview of Ethical and Societal Concerns.
KPMG. (2019). Ethical AI. Five guiding principles. URL: https://assets.kpmg.com/content/dam/kpmg/es/pdf/2020/01/Informe_Ethical-AI.pdf
Lebovitz, S., Lifshitz-Assaf, H., & Levina, N. (2023). The No. 1 Question to Ask When Evaluating AI Tools. *MIT Sloan Management Review* 6-9.
Linardatos, P., Papastefanopoulos, V., & Kotsiantis, S. (2020). Explainable ai: A review of Machine Learning Interpretability Methods. Entropy 23(1):18
Lindebaum, D., Glaser, V., Moser, C., & Ashraf, M. (2023). When Algorithms Rule, Values Can Wither. *MIT Sloan Management Review* 10-13.
McKinsey & Company. (2023). Human-centred AI: The Power of Putting People First.
MGF for GenAI. (2024). Model AI Governance Framework for Generative AI: Fostering a Trusted Ecosystem. AI Verify Foundation and InfoComm Media Development Authority. 30 May 2024.
Moser, C., Den Hond, F., & Lindebaum, D.v (2023). What Human Lose When We Let AI Decide. *MIT Sloan Management Review* 25-27.
Munoko, I, Brown-Liburd, H.L., & Vasarhelyi, M. (2020). The Ethical Implications of Using Artificial Intelligence. *Journal of Business Ethics* 167, 209-234
NIST (National Institute of Standards and Technology of the US Department of Commerce). (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0).*
OECD. (2019). OECD AI Principles Overview. URL: https://oecd.ai/en/ai-principles
OECD. (2024). *Report on the Implementation of the OECD Recommendation on Artificial Intelligence.*
Othmar Manfred Lehner, Kim Ittonen, Hanna Silvola & Eva Strom. (2022). Artificial intelligence based decision-making in accounting and auditing: ethical challenges and normative thinking. *Accounting, Auditing and Accountability Journal* 35(9), 109-135.
Parasuraman, R., & Manzey, D. (2010). Complacency and Bias in Human Use of Automation: An Attentional Integration. *Human Factors: The Journal of Human Factors and Ergonomics Society.* June 2010.
PWC. (2019). A practical guide to Responsible Artificial Intelligence (AI). URL: https://www.pwc.com/gx/en/issues/data-and-analytics/artificial-intelligence/what-is-responsible-ai/responsible-ai-practical-guide.pdf
Raji, I.D., Smart, A., White, R.N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., Barnes, P. (2020). *Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing*.
Ribeiro, J., Cardoso, L., Santos, V., Carvalho, E., Carneiro, N., & Alves, R. (2024). How Reliable and Stable are Explanations of XAI Methods? Cornell University. https://doi.org/10.48550/arXiv.2407.03108
Rooij, I. V. (2022). Against Automated Plagiarism. https://irisvanrooijcogsci.com/2022/12/29/against-automated-plagiarism/

Samek, W., & Muller, K.R. (2010). Towards Explainable Artificial Intelligence. In: W. Samek et al. (Eds.) *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*. Lecture Notes in Computer Science, vol. 11700, pp. 5-22. Springer, Cham (2019).
The Straits Times (TST). (2024a). Doing AI Right Matters More Than Doing It Fast. 5 June 2024. https://www.straitstimes.com/opinion/doing-ai-right-matters-more-than-doing-it-fast
The Straits Times (TST). (2024b). OpenAI Launches New Strawberry Series of AI Models with Reasoning Abilities. 13 September 2024. https://www.straitstimes.com/tech/tech-news/openai-launches-new-strawberry-series-of-ai-models-with-reasoning-abilities?
The Straits Times (TST). (2024c). Meta Releases AI Model that can Check Other AI Models' Work. 19 October 2024. https://www.straitstimes.com/world/united-states/meta-releases-ai-model-that-can-check-other-ai-models-work?_gl=1*1pvlpk2*_gcl_au*NDMxNDcxNzIuMTcyMzIxOTI5NA.
Toth, Z., Caruana, R., Gruber, T., & Loebbecke, C. (2022). The Dawn of the AI Robots: Towards a New Framework of AI Robot Accountability. *Journal of Business Ethics* 178, 895-916.
UNESCO. (2021). Recommendation on the Ethics of Artificial Intelligence.
URL: https://unesdoc.unesco.org/ark:/48223/pf0000381133/PDF/381133eng.pdf.multi.page=3
Zhang, C., Zhu, W., Dai, J., Wu, Y., & Chen, X. (2023). Ethical impact of artificial intelligence in managerial accounting. *International Journal of Accounting Information Systems* 49, 1-19.