



EIXO 3 – ABORDAGENS INOVADORAS NA AUDITORIA DO SETOR PÚBLICO: CASOS E LIÇÕES APRENDIDAS

AI AUDITING – A NECESSARY TOOL TO STRENGTHEN RESPONSIBLE AI AND PROVIDE ASSURANCE

Dirk Brand

Independent legal consultant, special counsel at Swart Law and an Extraordinary Senior Lecturer at the School of Public Leadership, Stellenbosch University. He holds the following degrees: BComm, LLB, LLM (European Union Law) and LLD (Constitutional Law). He was the leader of the LegalAlzers team who won the First St.Gallen Grand Challenge in Switzerland on the application of the EU AI Act (July 2023). He is also a guest lecturer at the Hochschule Kehl, Germany, the Law Faculty, University of Verona, Italy, the SSML Pisa, Italy and West University Timisoara, Romania. He serves as a Policy Fellow at the Centre for Artificial Intelligence and Digital Policy (CAIDP) in Washington D.C. Much of his current work focuses on AI ethics, AI policy and AI assessment.

McElory Hoffmann

CEO and co-founder of Praelexis, a machine learning company that empowers its clients to become data-driven decision makers. He is also an extraordinary senior lecturer in Computer Science at Stellenbosch University. He has a BEngSci (Cum laude) in Computer Engineering and a PhD in Applied Mathematics, from Stellenbosch University. As a qualified computer engineer, mathematician and data scientist, McElory has discovered the different ways in which AI can solve a business's unique challenges. He is fascinated by how technology can make the world a better place and has devoted his life to helping others utilise AI to create a sustainable future.

Johan Van der Merwe

Data Science Strategist at Praelexis. He has the following degrees: BA, B.Th. and M. Th. in Systematic Theology, as well as an MBA and a PhD in Ethics, all from Stellenbosch University. His current focus areas include business model design and business strategy, as well as data strategy and agile methodology. He is also a Research Fellow at Stellenbosch University.

Aloisio Dourado Neto

Bachelor's degree in electrical engineering from UFBA, a master's in applied computing focused on data analysis, and a Ph.D. in Computer Science with a focus on AI and Computer Vision from UnB. A public servant at TCU since 2010, he has been working with data analysis for auditing since 2014. Currently, he is the Deputy Chief Auditor for Data Analysis and Innovation at the Secretariat of Strategic Information for Control, where he is responsible for the Court's Data and Analytics Governance.



Google Scholar:

<https://scholar.google.com.br/citations?user=g3v0lqQAAAAJ&hl=en>.

MOTIVATION

The rapid increase in the development and application of AI in both the private and the public sector, justifies specific rules to ensure the responsible development and use of AI. The unfolding AI regulatory landscape in many countries, raises various questions about assessing issues such as fairness, non-discrimination, transparency and accountability. Setting standards and stipulating requirements, for example for high-risk AI, create some normative frameworks for the development and deployment of AI systems. While this is crucial, it is evident that this is not sufficient, and that AI algorithms should also be audited to contribute to assurance.

Beyond its technical and compliance-driven objectives, AI auditing is a fundamental instrument for advancing citizen-centric public governance. By systematically assessing AI systems for core ethical norms, an audit provides tangible assurance that automated public services do not perpetuate societal biases or create inequitable outcomes. This process is crucial for safeguarding fundamental human rights in an era of digital transformation. Therefore, this audit framework should be viewed not merely as a technical assessment, but as a critical mechanism to foster public trust, ensure accountability, and guarantee that the deployment of AI in the public sector ultimately serves the citizen's best interest.

METHODOLOGY

We conducted a comprehensive research project for INTOSAI in the form of a desktop study, to develop an AI audit framework and methodology for the use of Supreme Audit Institutions (SAIs). This article is based on the research report which we submitted in December 2024.



From a society perspective there is a need to develop policies and rules that put appropriate guardrails in place to ensure the application of ethical standards and protection of human rights in the context of AI. While national AI policies are developing in countries in all parts of the world, legislative processes are much slower. The reality thus is that there are not yet so many dedicated and comprehensive laws on AI in the world. The model comprehensive law is the EU AI Act, and South Korea and Brazil have similar approaches to AI regulation.

Having good legislation in place does not necessarily mean that users can rest assured that all AI models are safe and reliable. Effective implementation is obviously important. The risk-based approach of the EU AI Act for example places duties on providers and deployers of AI to ensure that it conforms to the set criteria in the AI Act, aimed at strengthening the responsible development and use of AI, which support the protection of human rights, democracy and the rule of law. Reflecting on the requirements for ethical or responsible AI, various questions about issues such as fairness, transparency and accountability are raised. How could AI systems be assessed appropriately taking into account these and other ethical concerns? Appropriate legislation is crucial, but international standards such as ISO 42001 on AI governance developed by the International Standards Organisation (ISO) are equally important as a normative framework.

Is it then sufficient to have detailed legal requirements and standards? No, there should be a form of assessment as well. The EU AI Act provides for conformity assessments, which is a form of audit, although the concept of AI audit is not used in the Act. How could an AI audit be designed? The assessment of an AI system can include technical, legal and ethical requirements, although the lines between these requirements are often blurred. The nature of AI algorithms creates challenges to design an audit framework which would enable an audit team to assess the performance and legal compliance of AI systems. It is therefore recommended that an AI audit team should include a diverse range of skills, thus ideally a multidisciplinary team that include both legal and technical expertise. In the public sector it is possible and likely that SAIs will in future have an extended mandate that will include auditing



of AI systems used by public sector institutions. It is therefore important to develop a comprehensive AI audit framework that could be used by SAIs to provide assurance. The auditing process should not be treated as a procedural formality to be carried out ex post facto only, and leading to reactive, punitive measures. Instead, an “ethics by design” approach should position AI auditing as a catalyst for innovative and meaningful development practices. If a provider or a developer of AI knows the audit requirements, it can guide them in the development and deployment of AI systems. Other key role players in an organisation, such as the internal risk officer and chief AI officer could use an AI audit framework as a guide in the responsible development and deployment of AI systems. There could be different types of AI audits, namely:

- Compliance audit (e.g. conformity assessment),
- Performance audit,
- A combination of performance and compliance audit, and
- AI impact assessment.

Providing assurance is, like the case of financial audits, crucial in determining the scope and objectives of an AI audit. In view of the nature of AI, there should also be an AI lifecycle approach. This means that the AI audit should deal with the whole AI lifecycle, which includes the data sources.

RESULTS

In our research we reviewed a wide range of publications related to AI auditing from practice, academia and public institutions. It was evident that a comprehensive and generally accepted AI audit framework for public sector institutions does not exist. The result of the research project is a comprehensive **AI audit framework**, including an **AI Audit Guide** with detailed questions, that consists of the following key features, namely:

Purpose: describe clearly what the purpose of the audit is.



Scope: determine which AI systems should be audited, and if it is general AI or generative AI.

Risk: determine the risk profile of the AI system. If the EU AI Act applies, the risk classification in the Act must be used.

Methodology: clarify what AI standards and legislation will be used as normative framework and apply it to the AI audit. The evaluation of the following key ethical norms should be the primary focus of the audit methodology, namely:

- Privacy and data governance
- Bias, fairness, non-discrimination
- Transparency and explainability
- Human oversight
- Technical robustness and safety

The following steps should be implemented by the audit team:

- 1) Obtain the technical documentation, risk management plan, conformity assessment and any other relevant documentation in support of the audit.
- 2) Engage the auditee in the audit process to get responses on all the selected questions in the AI Audit Guide. During this stage it is important to get input from key stakeholders in the organisation, such as the legal unit, data science office and risk management office.
- 3) Assess the risk management and mitigation measures.
- 4) Develop an AI audit report that includes findings and recommendations, and which should acknowledge the input of the auditee.
- 5) Present the AI audit report to the auditee.

The assessment in accordance with the AI Audit Guide should be done by applying a scoring scale from 1 – 5, with 1 being poor/critical and 5 being excellent.



Auditing generative AI such as LLMs is more complex and requires a different methodology. In such a case the primary focus should be on auditing the downstream application developed from the LLM.

The AI audit team could consider using a range of software tools to assist the audit process. AI systems, including generative AI applications, could be used for example in automated documentation and reporting, bias identification and mitigation and risk identification and assessment. An AI Audit assistant (AI agent) could be a useful tool for auditors to manage large volumes of documents, such as reports, opinions and policies included in the database of a Supreme Audit Institution.

CONCLUSION

AI auditing is still fairly new and evolving, but it will eventually become standard practice that will make a significant contribution to building trust in ensuring responsible AI development and deployment. We recommended that SAIs adopt the proposed AI audit framework and ensure that appropriate skills are developed to support multidisciplinary AI audit teams for the public sector.