

Artificial Intelligence Strategy Framework



SUPREME AUDIT INSTITUTION OF INDIA
लोकहितार्थं सत्यनिष्ठा
Dedicated to Truth in Public Interest

COMPTROLLER AND AUDITOR GENERAL OF INDIA





SUPREME AUDIT INSTITUTION OF INDIA
लोकहितार्थं सत्यनिष्ठा
Dedicated to Truth in Public Interest

Artificial Intelligence Strategy Framework

Supreme Audit Institution of India

Preface



The landscape of public auditing is undergoing a significant transformation driven by emerging technologies. In particular, Artificial Intelligence (AI) has surfaced as a powerful enabler in enhancing audit precision, efficiency, and scope. Recognizing the pivotal role that AI can play in advancing the objectives of public accountability, transparency, and data-driven governance, this document outlines the **Artificial Intelligence Strategy Framework** for the Indian Audit and Accounts Department.

This Strategy Framework seeks to provide structured guidance on the adoption and integration of AI tools and techniques across the audit lifecycle. The Government of India is increasingly using AI in various applications and platforms, and thus it is pertinent for us to have framework and capacity to audit such systems, which is in compliance with the responsible AI principles. This document defines the processes for undertaking such an exercise. It is designed to support audit officials in leveraging AI to augment traditional audit methodologies, enabling smarter data analysis, anomaly detection, risk assessment, and predictive insights. The framework also underscores key considerations related to ethical use, data privacy, model governance, and capacity building.

The document is intended to serve as a reference for field audit teams, functional group coordinators, and policy divisions alike, facilitating a shared understanding of AI-enabled practices within the department. It also reflects the commitment of the Comptroller and Auditor General of India to continuously modernize public auditing processes and align them with global technological standards.

It is expected that this Strategy Framework will act as a catalyst in embedding AI into the core of audit operations, empowering our officials to address the complexities of modern governance and deliver impactful audit outcomes.

A handwritten signature in black ink, appearing to read 'K. Sanjay Murthy'.

(K. Sanjay Murthy)

Comptroller and Auditor General of India

April 2025



Foreword



The public audit landscape is evolving rapidly, driven by the exponential growth of data and technological advancements. Among these, Artificial Intelligence (AI) stands out as a transformative force—one that has the potential to redefine how audits are planned, executed, and evaluated. The Indian Audit and Accounts Department, with its mandate of ensuring transparency and accountability in public finance, must harness these emerging capabilities to stay ahead of the curve.

The Artificial Intelligence Strategy Framework is a forward-looking initiative aimed at equipping audit professionals with the knowledge, tools, and strategies necessary to integrate AI into their day-to-day operations. This framework outlines a structured pathway for the adoption of AI applications in public auditing, from anomaly detection and predictive analytics to intelligent sampling and process automation.

In crafting this strategy, care has been taken to ensure that the principles of audit integrity, ethical AI use, data security, and institutional capacity building are deeply embedded. The framework also reflects our alignment with global best practices and our commitment to leveraging AI for high-impact, value-added audits that respond to the complexity of modern governance.

I am confident that this document will serve as a cornerstone in the Department's journey towards digitally enabled and insight-driven audits. I encourage all officials to actively engage with the framework and become agents of transformation in our shared pursuit of excellence in public auditing.



(S Ramann)

Chief Technology Officer
Office of the CAG of India
April 2025



Contents

| S. No. | Chapter Name | Page No. |
|-------------------|---|-----------|
| Chapter-1: | Artificial Intelligence | 1 |
| 1.1. | Artificial Intelligence - Concepts | 2 |
| 1.2. | AI Ethics | 4 |
| 1.3. | Artificial Intelligence and Audit | 5 |
| Chapter-2: | Audit using AI | 9 |
| 2.1. | Adopting AI in audit | 9 |
| 2.2. | Broad Objectives | 9 |
| 2.3. | Scope of AI in Audit | 9 |
| 2.3.1. | Audit Planning | 9 |
| 2.3.2. | Audit execution | 10 |
| 2.3.3. | Audit Reporting | 11 |
| 2.3.4. | Audit Follow up | 12 |
| 2.4. | Early application of AI in Audit – IA&AD use cases | 13 |
| Chapter-3: | Auditing AI | 17 |
| 3.1. | Auditing AI systems | 17 |
| 3.1.1. | AI Regulatory framework | 17 |
| 3.1.2. | AI Maturity level of the audited entity | 17 |
| 3.1.3. | Ecosystem for AI Auditing | 18 |
| 3.2. | Auditing AI systems - Audit Objectives | 18 |
| 3.2.1. | Whether there are appropriate AI Governance mechanisms in the entity? | 19 |
| 3.2.2. | Whether the AI system has been developed appropriately? | 20 |
| 3.2.3. | Whether the AI system delivers the intended output or not? | 20 |
| 3.2.4. | Whether the AI system is safe and secure? | 21 |
| 3.3. | Auditing AI systems - Audit Process | 22 |
| 3.3.1. | Audit Approach | 22 |
| 3.3.2. | Audit Planning | 22 |
| 3.3.3. | Audit Execution | 22 |
| 3.3.4. | Audit Reporting and Follow up | 23 |

| | | |
|-------------------|---|-----------|
| Chapter-4: | AI- The way forward | 27 |
| 4.1. | Challenges of AI Implementation | 27 |
| 4.2. | Artificial Intelligence – Implementation Strategy | 29 |
| 4.2.1. | Research and Application | 29 |
| 4.2.2. | Training and Capacity Building | 30 |
| 4.2.3. | Infrastructure and resource requirements | 30 |
| Conclusion | | 35 |
| Annexure-1: | Regulatory Frameworks | 39 |
| Annexure-2: | Illustrative checklists | 41 |

CHAPTER 1



Artificial Intelligence



Chapter-1: Artificial Intelligence

Artificial Intelligence is an advanced computer system that can simulate human capabilities, such as analysis, based on a predetermined set of rules.¹ AI has made advancements in recent years, driven by sophisticated algorithms, increased computational power, and the availability of vast data. It has evolved from a theoretical concept to a transformative technology with far-reaching social acceptance and economic implications across diverse domains, in the last decade.

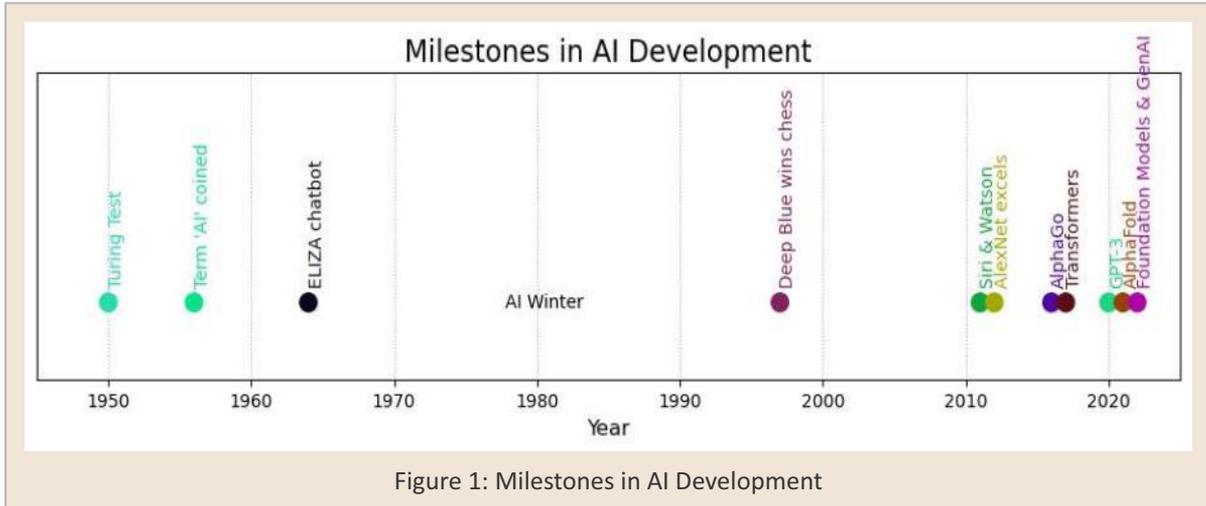


Figure 1: Milestones in AI Development

As AI continues to evolve, its potential to transform industries and drive innovation has assumed immense proportions. For instance, in the present-day healthcare sector, AI algorithms are analysing medical images for early disease detection, accelerating drug discovery etc. The financial sector leverages AI for fraud detection, risk assessment, modelling etc. AI now powers virtual assistants and chatbots, recommendation systems, and self-driving cars. AI systems have over the years rapidly converged and in some cases, even surpassed the human performance and capabilities in image recognition, reading comprehension, language understanding etc. (Refer Image below)

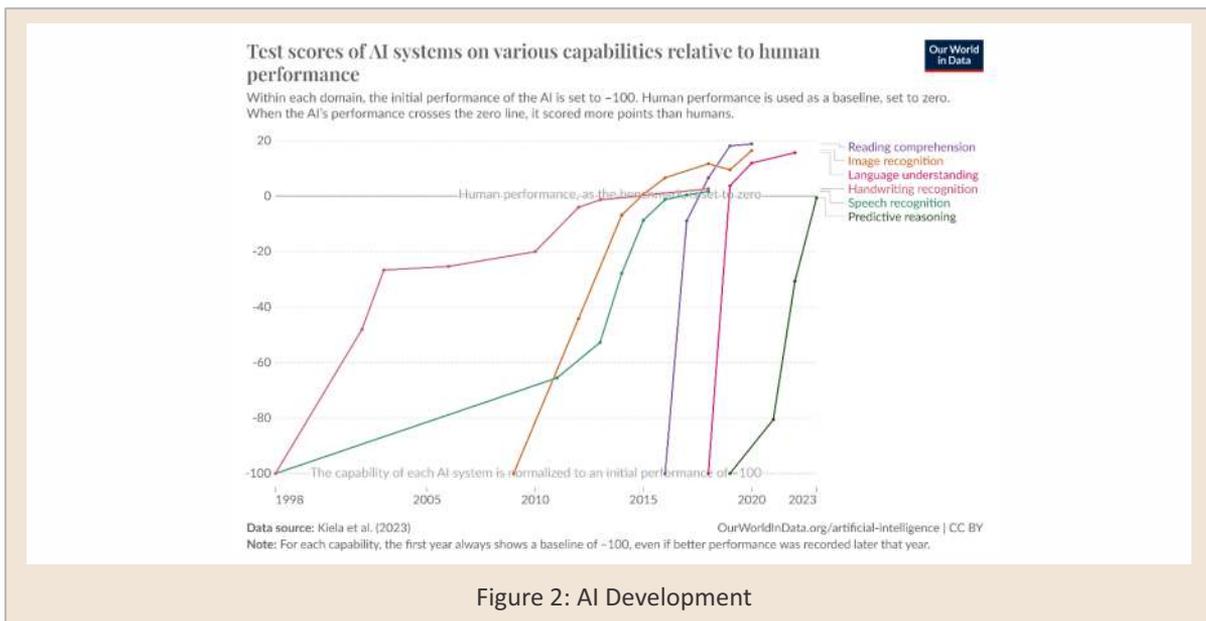


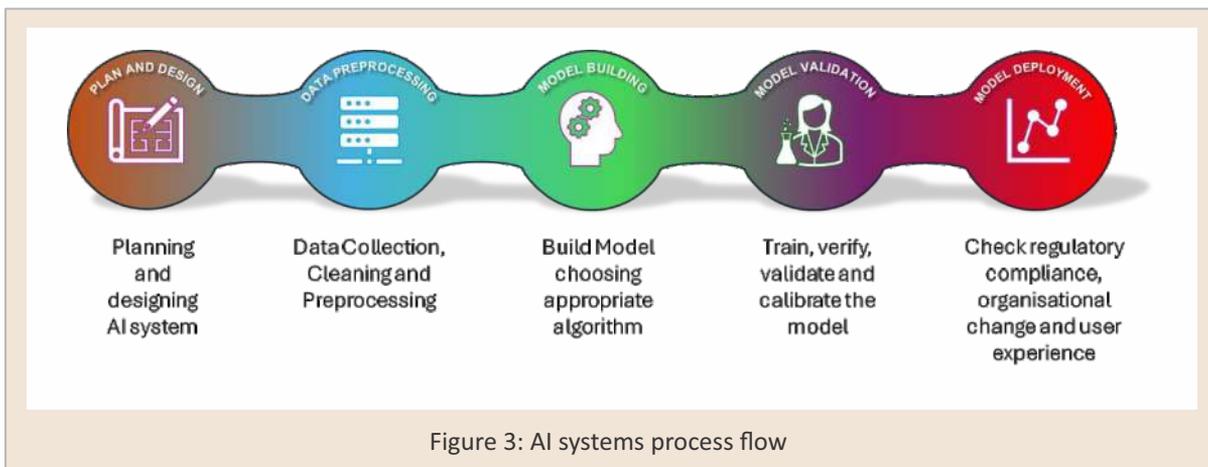
Figure 2: AI Development

¹ <https://www.isaca.org/resources/glossary>

However, at the same time, like with any emerging technology, the adoption of AI has also raised important social and ethical considerations, such as data privacy, algorithmic bias, and job displacement, etc. among others.

1.1. Artificial Intelligence - Concepts

AI technology involves various components, which start with a corpus of a large amount of data, which can be in different forms such as text, images, etc. This data is then fed into machine learning models, which are algorithms designed to learn patterns and relationships. The popular models are Neural Networks, Decision trees, and Support vector machines etc. They are trained using learning algorithms, such as Supervised learning, Unsupervised learning, or Reinforcement learning (explained in 1.1 below). This enables the models to improve their performance on a given task over time. This requires significant computing power, usually in the form of powerful processors (like Graphics Processing Units or GPUs), which is required to train and run these models efficiently. Software libraries and frameworks provide pre-built tools and components for building, training, and deploying AI models. Development and deployment tools help to manage the AI model lifecycle, including data preparation, model training, testing, deployment, and monitoring. Finally, a user interface is devised, such as a graphical user interface (GUI), natural language interface, or API (Application Programming Interface) which allows humans to interact with and utilize the AI system.



Cognitive AI and Generative AI are two approaches within the broader spectrum of artificial intelligence. Cognitive AI focuses on mimicking human thought processes, emphasizing tasks such as perception, reasoning, learning, and problem-solving. Generative AI is oriented towards creating fresh content, often leveraging new methods like deep learning techniques. It excels in tasks involving creative outputs, including generating images, music, text, and even entire virtual environments.

Both these approaches of Artificial Intelligence encompass several key components, each playing a distinct role in enabling machines to exhibit intelligent behavior. These components may work alone or in tandem to create powerful AI systems capable of learning, perceiving, solving problems, reasoning, and understanding human language.

Learning is a core feature of AI, enabling machines to acquire knowledge and improve their performance over time. Machine learning, the engine of AI, empowers systems to evolve and enhance their performance through experience. Unlike traditional programming, where rules are explicitly defined, machine learning algorithms learn from data, adapting and improving over time. This ability to learn is categorized into several approaches:

A) **Supervised Learning:** In supervised learning, models are trained with labeled data, learning to map inputs to desired outputs. This method excels in tasks like image classification, spam detection, and predicting customer churn. In simple terms - it is like teaching a child to identify fruits by showing them pictures and labels

B) **Unsupervised Learning** system would sift through unlabeled data, identifying patterns, grouping similar items, and uncovering hidden structures. This is invaluable for anomaly detection, customer segmentation, and recommendation systems. In simple terms, it is like piecing together a puzzle without knowing the final image.

C) **Reinforcement Learning:** This process mimics how humans learn from rewards and consequences. The Reinforcement learning agents interact with an environment, taking actions and learning from the resulting feedback in multiple iterations. This method is ideal for dynamic environments like robotics, game playing, and self-driving cars.

Deep learning employs artificial neural networks² comprising multiple layers to uncover complex patterns within data. It draws inspiration from the structure of the human brain. Deep learning models excel in intricate tasks including computer vision, facilitating tasks such as object detection, facial recognition, and medical diagnosis. Additionally, in **natural language processing**, they adeptly comprehend and generate human language, contributing to the development of chatbots, machine translation systems, and sentiment analysis tools. Furthermore, in **speech recognition**, deep learning algorithms convert spoken language into text, powering voice assistants, accessibility tools, and real-time transcription services.

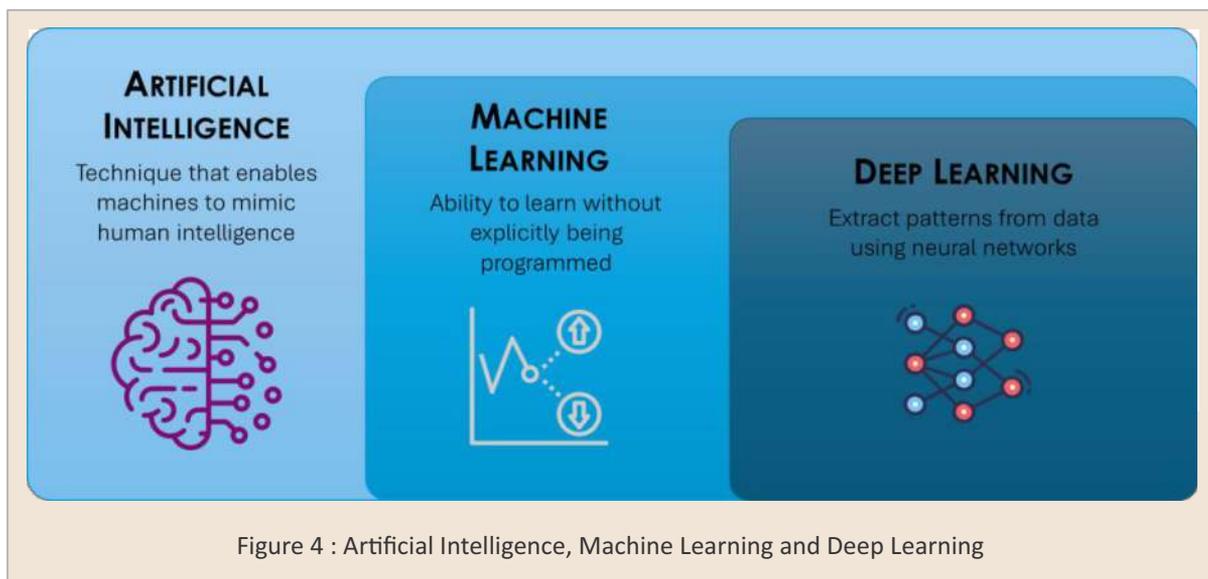


Figure 4 : Artificial Intelligence, Machine Learning and Deep Learning

² According to ISACA's glossary, Neural networks are defined as 'Networks, inspired by the structure of the human brain, that learn by processing data through three layers (input, hidden, and output). They can be trained to match any input to various outputs, including binary ones, making them versatile tools in deep learning (DL) for tasks like image recognition.'

Perception involves the ability of AI systems to sense and interpret the world around them using **sensors** and **data inputs**. It is especially popular in case of Computer vision where it is the ability of machines to understand and interpret visual information from images and videos. It can be used for object detection and recognition with wide application in medical diagnosis, autonomous driving etc. Similarly, in the realm of natural language processing, perception involves the comprehension of human language through techniques such as text analysis and speech recognition. Perception is essential for AI systems to interact effectively with the real world and understand the context of their environment.

Problem-solving involves machines' capacity to analyze complex issues, generate solutions, and make decisions. **Search algorithms**, for instance, are utilized to explore and navigate problem spaces efficiently, as seen in route planning for logistics. **Reasoning mechanisms** enable AI systems to draw logical conclusions or make inferences based on available information and rules, finding applications in expert systems, decision support systems, fault diagnosis, medical diagnosis, and strategy building, among others.

1.2. AI Ethics

Since AI is often used in decision-making, 'Ethics in AI' is vital for assuring that Artificial Intelligence technologies are enhanced and stationed in a way that is organized within a socially beneficial and legal framework.

Key ethical principles:

1. **Fairness** – AI systems must be depicted and deployed to prevent biases, ensuring equitable treatment across different demographic groups.
2. **Transparency** – AI's model and decision-making processes should be transparent to the end users and stakeholders.
3. **Accountability** – Ensuring with ethical standards and regulatory frameworks, the originators and end users of AI systems must be accountable.
4. **Privacy Protection** – AI systems must abide by strict data privacy laws, assuring that sensitive personal data is collected, stored and fabricated in a way that respects individual privacy rights.
5. **Safety and Reliability** – Proper testing of the AI systems should be carried out to ensure that they do not adversely affect individuals or groups.
6. **Human Control** - The AI should accompany human decision-making, rather than replace it completely, especially in crucial areas like healthcare and law enforcement.

AI practices in ethics must coincide with materialized global regulations, such as the Data Protection Bill in India, globally relevant GDPR, to ensure that AI systems are legally compliant while upholding ethical standards. To ensure the ethical components such as fairness, transparency, accountability, privacy and non-discrimination are integrated into the entire AI lifecycle, AI developers should follow clear ethical guidelines' framework.

While designing AI system, the following principles needs to be taken into account:

- Humans in the loop should empower the user and not automate them out,
- AI's behaviour should be explainable
- Avoiding bias and discrimination with mechanisms to identify the same
- Ensuring security and privacy as part of the design
- Ensuring trust, reliability and performance goals as part of the design

1.3. Artificial Intelligence and Audit

The integration of AI technologies into audit processes is expected to cause a change in basic assumptions, revolutionizing the way audits are planned, designed, and conducted by enhancing efficiency, accuracy, and risk management.

With its ability to analyze vast amounts of data, from diverse sources, at unprecedented speeds, AI empowers auditors to delve deeper into the audited entity's electronic records including databases and identify complex patterns, correlations, causalities, and anomalies.

Traditional auditing involves a lot of manual work where the auditors review documents, understand and analyse the numbers, verify transactions, and look for signals. It is time-consuming, labour-intensive, and can sometimes result in missing important details. AI would be of significant value to auditors as it can analyze large amounts of data from various sources, faster than humans and draw richer insights.

Use of AI in audit

With AI, auditors can automate many of the repetitive tasks involved in auditing, freeing up their time to focus on more complex issues. Instances of such benefits are:

- A) Better Risk Assessment by analysing Data and obtaining insights – AI would help to analyse vast amounts of structured and unstructured data, identifying patterns, trends, and anomalies that may be difficult for human auditors to detect. This enables fraud detection, pattern recognition in procurement etc.
- B) Procedure and Tasks Automation – AI can automate repetitive and mundane tasks involved in auditing, such as data entry, calculations, and reconciliations, matching of the key documents with the findings etc. Compliance checks can also be automated.
- C) Repository of knowledge and guidelines - A unified AI based repository of all guidelines and orders, for search and querying using natural language would be amenable to chat bot-based interaction, as well as development of organisation centric Foundation model, trained on such repository. This will enable access to knowledge on various audit subjects.
- D) Report Preparation – AI could be used to generate drafts, which could be generated using the LLM trained and prompted on internal guidelines. AI could be used for dissemination of findings in various languages using NLP Translation and for enabling citizens and stakeholders to interact with reports in a conversational style.



CHAPTER 2

A decorative wavy pattern in dark blue and light beige colors, resembling a stylized wave or cloud, positioned at the bottom of the gold background.

Audit using AI



Chapter-2: Audit using AI

2.1. Adopting AI in audit

The AI technology is percolating into the business processes, in every sector including the public sector. Keeping in pace with the technologies employed in government processes is essential for IA&AD to carry out the constitutional mandate as the watchdog of the government.

Integration of AI into auditing processes promises enhanced accuracy, streamlined operations, and heightened efficiency. By adeptly handling vast and diverse datasets, encompassing both structured and unstructured information, AI can be applied as a versatile tool across multiple domains within IAAD, spanning various phases of the audit lifecycle like risk assessment, fraud detection, continuous auditing, to name a few.

2.2. Broad Objectives

The broad objectives of using AI in Audit are to increase the efficiency and speed of audit processes, to enhance risk assessment, to improve the accuracy of audit findings and fraud detection capabilities.

2.3. Scope of AI in Audit

The scope of application of AI in audit is two-pronged. AI could be used to automate the repetitive audit processes/tasks releasing valuable audit resources for core audit processes. Also, it could be utilised for audit analysis of information.

2.3.1. Audit Planning

A well-planned audit is crucial to ensure profound audit conclusions. AI can enhance the efficiency of the auditors by helping to digest the available information faster and extracting the essence of it, thereby enabling better understanding of the audit entity, which is the cornerstone of a good quality audit.

In the planning phase of an audit, AI powered data analytics, Machine learning algorithms, Natural language processing techniques etc., can be employed to facilitate analysis of data (both structured and unstructured) assess risk factors and enhance the audit planning processes. The application of AI tech may also include speech recognition, grammatical tagging, sentiment analysis, text summarization etc. This can help in understanding the entity better, resulting in more accurate risk assessment.

The figure below provides an indicative list of use of AI in audit planning:

|  Audit Stage - Planning & Monitoring | |
|--|--|
| AI/ML Techniques & Methods that Can be deployed |  <ul style="list-style-type: none">  Clustering - creating clusters on auditee data based on similarity  Summarisation - Computing the impact of certain categories of auditee  Network Analysis - Identification of related auditees across multiple levels of connections  Deviation Detection - Identification of deviations that are significantly different from the normal plan |
| Scope for Automation |  <ul style="list-style-type: none">  Generating risk assessment report  Generating checklist  Generating a summary of policies, SOPs and others in the audit scope  Generating keyword-based analysis  Generate Standard Audit Design Matrix  Generating deviation report on audit plan |

2.3.2. Audit execution

Traditional auditing involves various testing procedures. The auditor does a test of controls and applies substantive analytical procedures. The traditional auditor heavily relies upon Sampling, to draw reasonable assurance. Though the principles remain the same, the auditor modifies the methodologies, when technology is adapted in audit procedures.

Adoption of AI in an audit engagement, is expected to benefit by automation of repetitive auditing tasks, which in turn, would allow auditors to focus their resources on value adding activities.

For instance, Granular transaction level checks could be run across the entirety of data and all the errors, non-compliances could be listed out rather than sample checks. Further pattern analysis of the error transactions could help the auditor to do root cause analysis and impact assessment. The use cases of early application of AI, are given in para 2.4. While there are diverse possibilities of using AI in audit execution, an indicative list is provided in the next page.

| | | |
|---|--|---|
|  | Audit Stage - Execution | |
| AI/ML Techniques & Methods that Can be deployed |  | <ul style="list-style-type: none"> ↪ Classification – for classifying documents and transactions based on segments ↪ Clustering - creating clusters on transaction data based on similarity ↪ Association - Identifying correlations between transactions or actions ↪ Summarization - Computing the impact of certain categories of transactions ↪ Network Analysis - Identification of related entities across multiple levels of connections ↪ Deviation Detection - Identification of transactions which are significantly different from the norm ↪ Prediction - Estimation of quantity/ materiality of transactions Image Analytics - Analysis of visual patterns ↪ Sentiment analysis – to extract sentiment from the text in documents. ↪ Statistical analysis – to evaluate term, phrase or topic trends in transaction data |
| Scope for Automation |  | <ul style="list-style-type: none"> ↪ Generating Samples based on risk parameters ↪ Generating outlier or exception report ↪ Generating Standard insights on data ↪ Generating possible observations based on previous reports |

2.3.3. Audit Reporting

An auditor has to adhere to the reporting standards, while preparing the audit reports. Besides that, Audit reporting is a thoughtful process and involves many finer nuances.

AI would be useful in automating the many manual tasks associated with report preparation, like formatting, tabulations, verifications, grammar checks, typographical errors etc. AI could be used to translate the reports in various languages, enabling easier and faster dissemination.

Going one step further, AI systems could be developed to generate interactive drafts, which would stimulate reader creativity and enhance understanding of the document to the citizens.

An indicative list of applications of AI in Audit reporting is provided below:

|  Audit Stage - Reporting | |
|--|---|
| AI/ML Techniques & Methods that Can be deployed |  <ul style="list-style-type: none"> ↪ Statistical analysis – to evaluate term, phrase or topic trends for reporting ↪ Summarization- summarize reports based on key elements and themes of the report. ↪ Natural Language Processing: system to understand the user's prompts and instructions for generating the report. ↪ Machine Learning- filling content based on pre-defined models and templates. |
| Scope for Automation |  <ul style="list-style-type: none"> ↪ Generate Standard Reports based on Audit Queries ↪ Generate Data visualisations of key issues identifying the risks ↪ Generate management summary on the Report ↪ Generate collated reports out of survey responses ↪ Generate recommendations and its impacts based on predictive analysis |

2.3.4. Audit Follow up

AI can be used to manage the entire gamut of audit follow-up with ease and efficiency. The utility of AI in follow-up of an audit can vary from repetitive tasks like issuing reminders to the audited entities for compliances to analysing the compliances submitted and suggesting a course of action through summarisation techniques and clustering tools. AI can also be used to analyse past observations and assess patterns for future/follow-up audits.

An indicative list of applications of AI in follow-up audit is provided below:

|  Audit Stage - Follow-up | |
|--|---|
| AI/ML Techniques & Methods that Can be deployed |  <ul style="list-style-type: none"> ↪ Clustering- grouping of observations based on nature and risk. ↪ Natural Language Processing- By mining textual feedback, NLP techniques can gauge auditee replies to audit observations. ↪ Classification- anomaly detection in the textual replies to highlight red flags ↪ Predictive analysis- identify key variables or leading indicators that have the greatest influence and suggest the next course of action ↪ Machine Learning- for keyword spotting and giving prominence to actionable inputs |
| Scope for Automation |  <ul style="list-style-type: none"> ↪ Generate standard templated reminders to the auditee like Six monthly statements. ↪ Generate audit replies in translated form in the language of one's choice. ↪ Generate suggested course of action using past trends of similar audit observations and replies. ↪ Generate risk-weighted reports on audit observation for prioritizing follow-up |

2.4. Early application of AI in Audit – IA&AD use cases

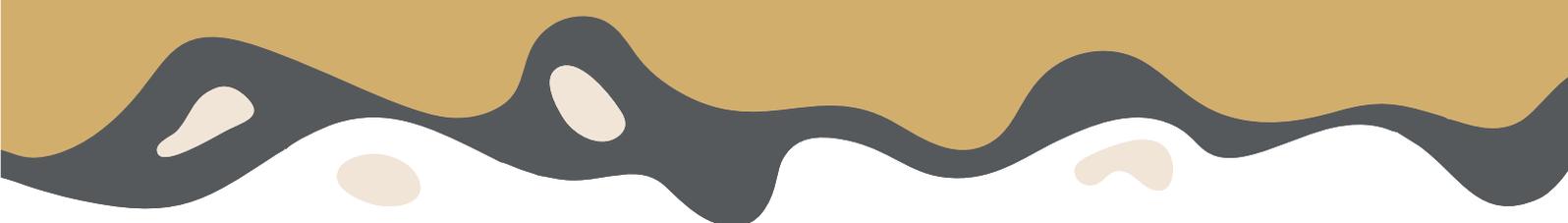
IAAD has experimented with the application of AI in its audits and a few of the interesting use cases are given below:

- ✓ **Use of AI in assessment of plantations executed by the Department of Forest, Odisha State:** In this case, Machine Learning (ML) technique was used to identify the tree species in a plantation. The images of various attributes of tree species like leaves etc. were fed to the system to correctly identify the tree species. Also, AI algorithms were used for qualitative assessment viz Tree height, canopy density, effectiveness of the trenches dug through drainage pattern analysis etc. and for quantitative assessment viz tree counting, species identification, trench counting etc.
- ✓ **Use of AI in identifying circular trading transactions in taxation:** Circular trading normally is used to issue fake invoices in transactions among multiple parties without an actual supply of goods. Using AI algorithms, specific types of circular transactions of up to 8 iterations were identified. The model was trained and tested on the selected e-way bill data set related to taxation and several circular trading transactions/ patterns were discovered by audit.
- ✓ **Use of AI in detecting similar images used by multiple applicants:** With respect to the data related to Digital Literacy programme of Govt of India, the beneficiaries under the scheme were spread across India and training was imparted by designated training centers. The photograph of the beneficiaries was to be uploaded for each training conducted. As manual analysis of the photo images would be tedious, an intelligent AI model was developed to detect cases where same images were used for claiming the training cost, different images of same beneficiaries used for claiming the training cost and non-human images used for beneficiaries.

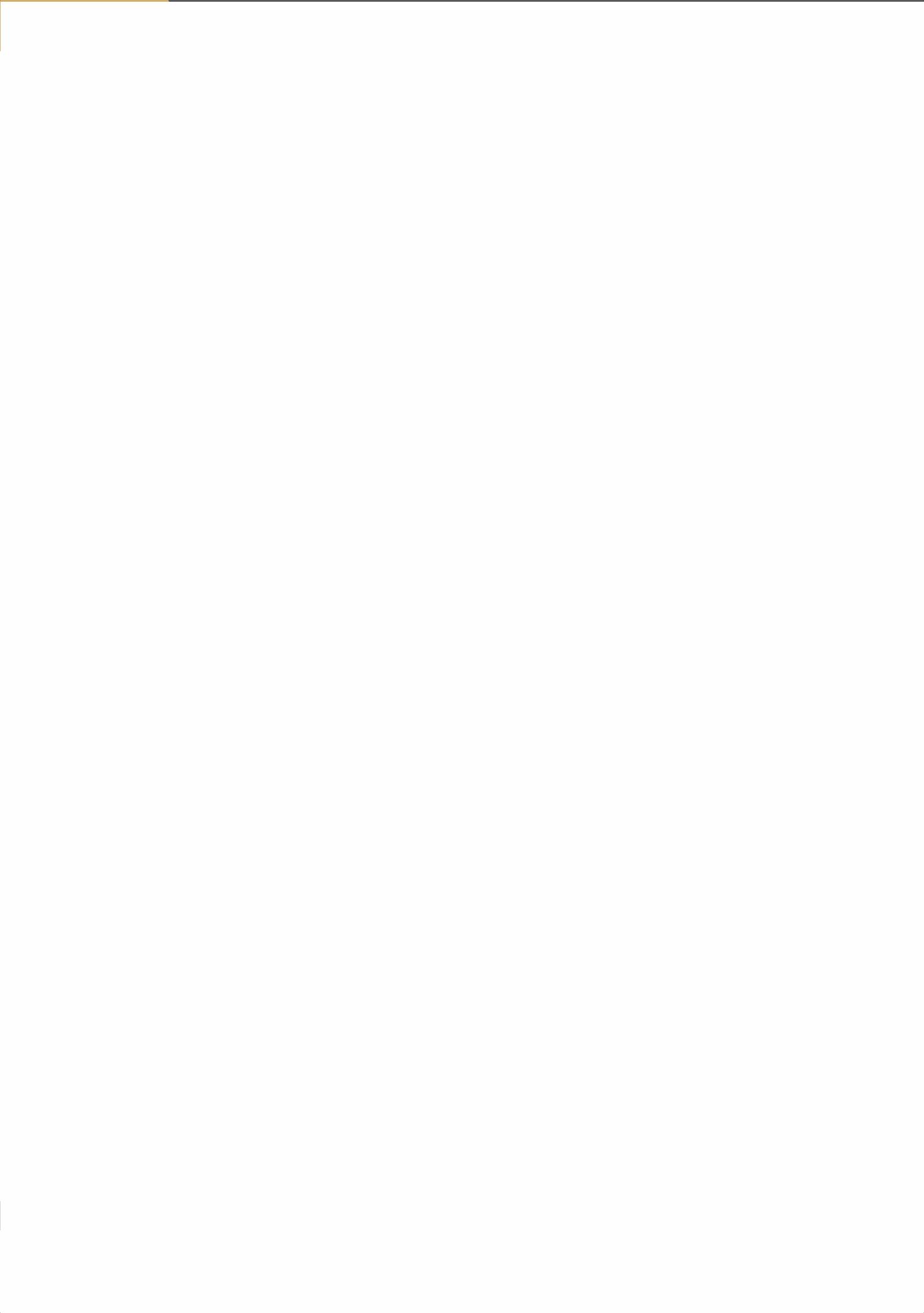
| | | | | |
|--------|-----|-----|-----|----|
| 15919 | 28 | 300 | 1 | 1 |
| 52937 | 147 | 360 | 250 | 23 |
| 55098 | 154 | 357 | 250 | 16 |
| 14563 | 43 | 338 | 1 | 6 |
| 94733 | 286 | 331 | 2 | 4 |
| 49530 | 157 | 315 | 15 | 14 |
| 14430 | 46 | 313 | 1 | 1 |
| | 221 | 290 | 150 | 19 |
| 5844 | | 278 | 15 | 3 |
| | | 275 | 15 | 23 |
| 8137 | | 275 | 1 | 2 |
| | | 253 | 15 | 21 |
| 9674 | | | 0 | 1 |
| | | | 1 | 12 |
| | | | 0 | 4 |
| 6210 | | 2 | 15 | 6 |
| | | | 6 | 2 |
| 7810 | | 3 | 0 | 1 |
| | | | 0 | 1 |
| 9965 | | 4 | 35 | 47 |
| | | | 4 | |
| 78247 | | 3 | | |
| 5251 | | | | |
| 28010 | | | | |
| | 34 | | | |
| 40960 | 190 | | | |
| 105790 | 491 | | | |
| 6448 | 30 | | | |
| 12692 | 60 | | | |
| 5476 | 26 | | | |
| 8096 | 39 | 21 | 15 | |
| 7055 | 34 | 20 | 2 | |
| 14149 | 70 | 202 | 6 | 5 |
| 13168 | 65 | 202 | 2 | 3 |
| 8063 | 40 | 201 | 15 | 5 |
| 19751 | 98 | 201 | 15 | 11 |
| 5225 | 26 | 200 | 150 | 3 |
| 11610 | 58 | 200 | 3 | 2 |
| 5014 | 25 | 200 | 1 | 3 |
| 41947 | 211 | 198 | 6 | 28 |
| 24742 | 126 | 196 | 1 | 3 |
| 21652 | 110 | 196 | 3 | 16 |
| 11551 | 59 | 195 | 25 | 3 |



CHAPTER **3**



Auditing AI



Chapter-3: Auditing AI

Artificial Intelligence (AI) systems are increasingly being deployed in various ways within the Indian governance ecosystem to improve efficiency, enhance decision-making, and provide better services to citizens. Various public service delivery portals employ chatbots for interactive querying by the citizens. (Example: AskDisha of IRCTC)

One can observe that AI models are being used by the authorities in regulatory compliance and monitoring. The Goods and Services Tax Network (GSTN) and Income Tax Department use AI and data analytics to flag discrepancies in tax returns.

Another example of AI in public service delivery is the use of AI for Traffic and Transport Management in Bengaluru city using AI based app ASTraM (Actionable Intelligence for Sustainable Traffic Management).

As the government increasingly incorporates artificial intelligence (AI) into its operations and delivery of services, the need for auditing AI systems grows. The expanding role of AI in government functions necessitates a closer examination of its performance and compliance.

3.1. Auditing AI systems

Developing suitable procedures, processes and practices for audit of AI systems is dependent on the following factors:

3.1.1. AI Regulatory framework

Audit frameworks are intrinsically linked to the regulatory framework in place, in the governance ecosystems. AI is an evolving technology and the regulatory framework in India is yet to mature, as it is across the world (refer Annexure 1). Thus, it is imperative for IA&AD to continuously keep evolving its procedures and practices in sync with the evolving regulatory ecosystem, and international audit frameworks.

To begin with, in IA&AD, the audit of ML models can be structured according to the cross-industry standard process for data mining (CRISP-DM)³ or other IT Audit Frameworks like COBIT as it aligns with the standard development process of ML models.

3.1.2. AI Maturity level of the audited entity

When a technology-based process is implemented in an organisation, the entity passes through different maturity levels. There are many maturity models in place, in different industries. In the IT space, for example, as per the Capability Maturity Model (CMM), the maturity level of an organisation with respect to software development processes, ranges from 1 to 5, viz., Initial, Repeatable, Defined, Managed and Optimised. Similarly, AI's application in governance would range from basic automation of routine tasks to complex data analysis, predictive modelling and automation of complex cognitive and generational tasks. Understanding the AI maturity level within governmental operations is crucial for an auditor to audit the same, effectively.

³ Auditing machine learning algorithms (auditingalgorithms.net)

Systematic auditing of AI based on the AI-technology maturity level of the entity, would allow structured evaluation of data quality, algorithmic accuracy, reliability and enable the auditor to identify vulnerabilities like model drift⁴ and unintended outcomes.

The OECD AI Framework of Classification⁵ provides for a nine-stage maturity level classification. Likewise, an AI maturity model could be adopted for auditing the entities, by IA&AD. This will enable to streamline the AI audit and to focus on materiality and risk-based audit of the entities. An attempt on the above lines, has been made in Annexure-2 which contains checklists to assess and audit the AI model maturity level, document assessment etc.

3.1.3. Ecosystem for AI Auditing

An ecosystem for AI auditing refers to the structured and collaborative framework that enables the systematic evaluation of AI systems to ensure they are ethical, fair, transparent, and compliant with legal frameworks and organizational standards. This ecosystem involves a diverse set of stakeholders, including AI developers, legal experts, ethicists, domain specialists, regulators, and third-party auditors, working together across the lifecycle of AI models—from design and training to deployment and monitoring. It integrates technical tools such as bias detection algorithms, explainability modules, and privacy auditing frameworks with governance mechanisms like documentation standards, model cards, and impact assessments. An effective AI auditing ecosystem also requires strong institutional support, including national guidelines, sector-specific norms, applicable standardized metrics, and independent oversight bodies to ensure accountability. As AI becomes increasingly embedded in decision-making processes, this auditing ecosystem plays a critical role in safeguarding trust, mitigating risk, and aligning AI outcomes with organizational goals.

The AI auditing ecosystem should be multifaceted, involving collaboration across organizations, regulatory bodies, technology providers, and third-party auditors to establish trustworthy and reliable AI systems. The auditing ecosystem should also adhere to the national guidelines and sector-specific norms.

3.2. Auditing AI systems - Audit Objectives

AI systems are technology systems in essence and at a broad level, the audit objectives are similar to that of the audit of IT systems. However, the difference lies in the inherent risks involved in the subject matter, and hence, the audit focus areas and relevant choice of audit methodologies to be employed while auditing AI systems would be very different at the granular level.

The audit focus areas while auditing AI systems would be:

1. AI technology Governance Mechanisms
2. Design, development, deployment of the AI system

⁴ Model Drift is phenomenon where a machine learning model's performance degrades over time as the real-world data it encounters drifts away from the training data it was initially trained on.

⁵ <https://www.oecd.org/publications/oecd-framework-for-the-classification-of-ai-systems-cb6d9eca-en.htm>

3. Operations & Performance
4. Security aspects

The broad audit objectives of audit of AI systems are as follows:

3.2.1. Whether there are appropriate AI Governance mechanisms in the entity?

Sophisticated technologies are generally high risk prone. The ethical and safety risks posed by AI are extremely high compared to the risks posed by a traditional IT system. The area of 'Governance' of IT, has been focused only to the extent of productivity and efficient utilization of resources, in a traditional IT audit. However, with AI, more emphasis must be laid on the 'Governance' of AI systems, considering the high risks.



Illustrative Audit Questions

- ✓ Whether the entity has an AI strategy, policy and necessary frameworks, which are in alignment with the regulatory framework?
- ✓ Whether operational structures have been established with clear roles, responsibilities and line of control for the design, development, deployment and continuous monitoring AI systems, both at the strategic and operational level?
- ✓ Whether the governance structures sufficiently oversee the development of the AI systems, to ensure they are compliant and bias-free?
- ✓ Whether the governance structures sufficiently monitor the operations of the AI systems?
- ✓ Whether the governance mechanisms review the security aspects of the AI systems periodically and exercise sufficient control over it?

The AI governance structure should be aligned to the national societal values and principles, following the AI protocols and frameworks established by the government of India.

The governance of AI systems involves creating a framework that ensures that AI is developed, deployed, and monitored in ways that are ethical, transparent, secure, and aligned with societal values. This governance structure needs to be flexible, adaptable, and responsive to the rapid advancements in AI technology, while also ensuring accountability and mitigating risks such as bias, discrimination, and privacy violations.

One of the vital audit objectives would be to verify if the entity has put appropriate governance mechanisms in place to oversee and monitor the establishment, operations, assessments and continuous improvement of the AI systems.

3.2.2. Whether the AI system has been developed appropriately?

The AI systems deployed may be COTS (Commercial-off-the-shelf) or developed, trained and deployed in-house by the entity themselves. The auditor would be verifying the developmental cycle of the AI system, to see if the relevant checks and measures have been placed at the developmental stage.

As AI models are mostly black box in nature it may lead to lack of trust by users. Therefore, Model Cards⁶ and Documentation are very crucial and should be checked by the auditors. Documentation Review while auditing AI systems will provide comprehensive insights into the design, development, and deployment of AI systems, and would throw light on whether AI systems adhere to relevant laws, regulations, and industry standards, or ethical guidelines for AI development and deployment.

The audit should ensure responsibility in the AI systems by ensuring protocols to handle Fairness and Non-Discrimination, Transparency and Explainability, Accountability, Data Protection, and reliability.

The audit should also ensure that AI systems are cross-collaborative and aligned with the regulatory frameworks outlined by the various government agencies.



Illustrative Audit Questions

- ✓ What was the problem the entity was attempting to solve, whilst deploying AI technology?
- ✓ What was the underlying model adopted by the entity to solve the problem and whether that is the most suitable or appropriate model?
- ✓ Whether appropriate training data has been employed in training the model and was it sufficient to develop an AI model without bias? What kind of controls have been put in place to evaluate data used for training the model?
- ✓ Whether the model developed was tested sufficiently and declared safe before deployment, by the appropriate governance level, within the entity?
- ✓ Whether essential documentation has been done and maintained during the system development cycle?
- ✓ Whether the relevant risks have been addressed by the development team, while developing/deploying the model?

3.2.3. Whether the AI system delivers the intended output or not?

Auditors need to assess whether the AI system produces the intended accurate and reliable results. As Deep Learning algorithms are mostly black box in nature, AI systems could be opaque in nature. However, the auditor is expected to evaluate the performance of the algorithms and the accuracy/consistency of the outputs.

⁶ AI model card is a brief description that provides key details about an AI model, including its intended use, training data, performance benchmarks, and potential limitations or biases.



Illustrative Audit Questions

- ✓ Whether there is appropriate input (data evaluation), processing (review of decision models) and output controls in the AI system to ensure intended results?
- ✓ In the case of AI systems trained on unstructured data, whether the model has picked up bias/spurious co-relations, resulting in unintended outcomes?
- ✓ Is the output of the AI system accurate and reliable?
- ✓ Is the output of the AI system achieving the performance metrics set by the entity?
- ✓ Whether appropriate mechanisms have been put in place by the entity to detect data and model drift?
- ✓ Is there a feedback loop to assess the output of the AI system?
- ✓ Is the output of the AI system and the feedback monitored and verified at appropriate intervals by the operations team?
- ✓ Whether corrective actions based on the results of monitoring activities have been taken by the entity or not?

3.2.4. Whether the AI system is safe and secure?

The security risks associated with AI are extremely high compared to the traditional IT systems. Auditors must ensure that AI systems are secure and safe and handle data as per security standards. Effective auditing of AI model security is essential to build trust in these systems, promote their ethical and responsible use, and ensure compliance with relevant regulations and standards.

This involves assessing data protection measures, encryption protocols, access controls, and other security safeguards. AI security auditing can help identify vulnerabilities, biases, and security risks, and provide insights for improving the design, development, and deployment of AI models.



Illustrative Audit Questions

- ✓ Whether the system complies with the Data Protection laws and regulations, such as the GDPR or CCPA?
- ✓ Whether the entity has implemented security measures to protect AI systems from potential attacks, such as data manipulation and input/output manipulation?
- ✓ Whether the security and safety aspects of the AI systems are continuously monitored, assessed and improved upon?
- ✓ Whether the necessary policies, procedures, structures and practices have been put in place to ensure security incidents are dealt with the utmost priority and seriousness?

3.3. Auditing AI systems - Audit Process

An indicative audit process is suggested below to be followed while auditing AI systems, while keeping in mind the Department's existing audit policies and ISSAI principles of auditing.

3.3.1. Audit Approach

The Audit of AI systems will need a tailored approach that considers both traditional auditing principles and unique characteristics of AI. This audit, like any other audit, will either fall under Compliance audit (CA) or Performance audit (PA) and will have to be conducted within the rules and regulations of a CA or PA. Going forward, IA&AD may come up with a dedicated document detailing the specific AI audit related aspects during audit planning, execution, reporting and follow up stages to enable the field offices to conduct the AI audits efficiently and effectively.

An outline of the audit processes to be followed is discussed in succeeding paragraphs.

3.3.2. Audit Planning

- Audit planning may commence with understanding the AI system, its functionalities, data sources, algorithms etc.
- The key stakeholders of the AI system may be identified along with their assigned roles and responsibilities to understand accountability and oversight mechanisms.
- Conduct risk assessment to identify the risks associated with the AI system and prioritise the identified risks based on their severity.
- Understand the legal and regulatory environment applicable to AI system under audit, identify specific compliance requirements and its adherence by the AI system.
- Formulate clear and specific audit objectives, scope, methodology and sampling procedure, if any.
- Evaluation of AI systems should be guided by qualitative and quantitative metrics ideally suited to the problem statement. Newer metrics must be derived if the current metrics available are not sufficient for completeness of the desired measurement.

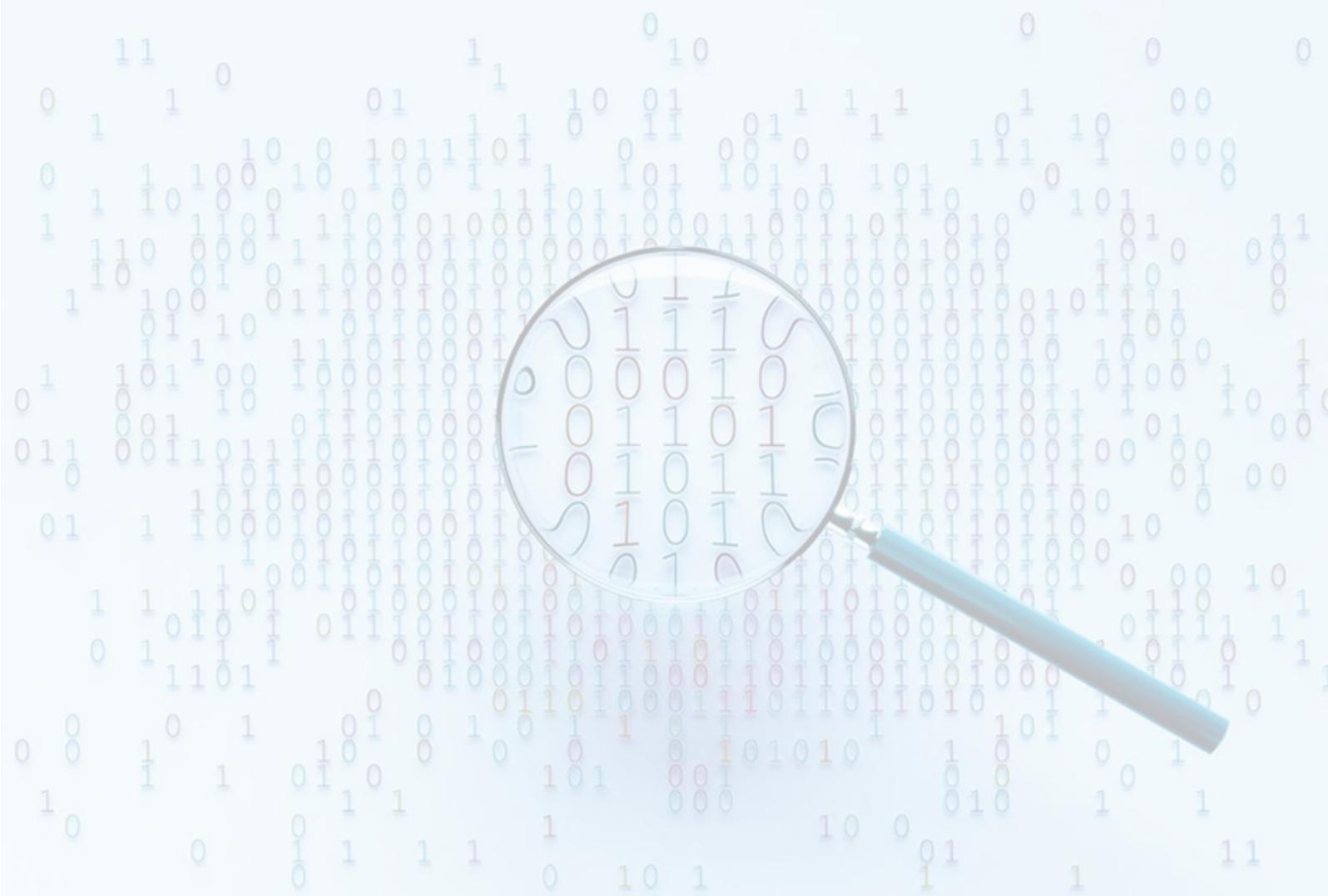
3.3.3. Audit Execution

- Auditee engagement is an important aspect when it comes to audit of AI systems. A detailed entry meeting at the level of the Heads of Department (considering the probable audited entity concerns) clearly explaining the audit objectives, approaches and methodologies, followed by continuous engagement with the entity and sharing of findings to consider auditee's view are crucial steps.
- The auditee should be clearly explained about the process and other granularities of auditing AI systems. The auditee should be kept in confidence if any external experts are used for the audit. Any doubts of auditee should be promptly clarified, and the engagement must be continuous.

- During the audit of AI systems, quality and integrity of data used to train and operate the AI system should be evaluated. Audit must also focus on the accuracy, reliability and performance metrics of AI models audited and assess the interpretability of the AI model.
- Algorithmic fairness may be assessed and implement techniques to mitigate biases in data and algorithms wherever identified.

3.3.4. Audit Reporting and Follow up

- The auditor may establish documentation requirements to document audit findings, methodologies adopted, and evidence collected during the audit process.
- The audit report may comprehensively summarize the audit findings, identifying areas of improvement and providing recommendations to mitigate risks and enhance compliance.
- Like the entry meeting, the Exit conference would be crucial, wherein the findings of audit of AI systems are to be discussed with the audited entity. This Head of Department level engagement should clearly explain the findings and methodology adopted in arriving at the findings so as to give another opportunity to the audited entity to express their explanations and concerns, if any.
- The auditor may implement mechanisms for continuous and ongoing monitoring of AI system's performance, changes in regulatory requirements and technological advancements.
- The auditor may implement mechanisms for continuous and ongoing monitoring of AI systems' performance, changes in regulatory requirements and technological advancements.
- The auditor may review the performance of the AI models across time intervals to identify potential issues such as biases, errors, security risks, and performance degradation that may emerge over time.



CHAPTER 4

AI - The way forward



Chapter-4: AI- The way forward

The integration of artificial intelligence (AI) into the audit process, both for utilizing AI in auditing tasks and auditing AI systems, presents a host of challenges. Firstly, there are technical hurdles associated with implementing AI algorithms for auditing purposes, such as ensuring the accuracy, reliability and replicability of AI-driven analytics in detecting anomalies or identifying patterns within vast datasets. Additionally, auditing AI systems introduces complexities related to understanding the system in general and the biases inherent in AI algorithms in particular, verifying transparency in AI decision-making processes, and assessing the ethical implications of AI-driven auditing practices.

Addressing these challenges is crucial for several reasons. Firstly, ensuring the accuracy and reliability of AI-driven audit processes is essential for maintaining the integrity and trustworthiness of financial reporting and regulatory compliance. Secondly, mitigating biases and ensuring transparency in AI systems is imperative for safeguarding against potential discriminatory or unethical outcomes, as well as maintaining public trust in AI technologies. Finally, addressing these challenges fosters continuous improvement and innovation in AI-driven auditing practices, ultimately enhancing the effectiveness and efficiency of audit processes in an increasingly complex and data-driven governance environment.

4.1. Challenges of AI Implementation

The various challenges foreseen while implementing AI in the department and the ways they could be overcome have been listed below:

The Challenges of Planning for the Future of AI

The rapid advancement of artificial intelligence (AI) technology along with the lack of standardisation in the AI sector, virtually makes long-term planning and forecasting, difficult. The department must evolve an operational structure and system within the department, nimble enough to evolve alongside the technology.

Infrastructure Challenges

Adopting a new technology requires various resources, including appropriate skillsets which, could be a challenge to the department. Various resource constraints viz. infrastructure, tools, budget, software, etc. can also hamper the use of AI in Audit. The department must thoughtfully plan and acquire the necessary resources for implementing AI.

Building Competency

Technically competent skillset is a scarce resource and augmenting the same has always been a challenge for public sector. The department must devise mechanisms through which the skill set could be accessed from the market and developing the internal pool of resources, simultaneously.

Overcoming Resistance to AI in Auditing

Cultural adoption poses a challenge, as introducing AI into traditional audit practices often meets resistance from officials accustomed to traditional methods. The department may gradually overcome the same by adopting appropriate sensitisation and up-skilling measures.

Robustness of audit processes

The complexity of AI algorithms poses significant privacy challenges, including data breaches, adversarial attacks, and AI-assisted hacking, particularly in audit processes where handling sensitive information requires stringent safeguards to maintain data integrity and comply with privacy regulations. Also, there is the risk that in AI audits, bias can arise from various sources such as inherent biases in the modelling process and biases present in the training data used to develop AI models. These biases can impact the objectivity and accuracy of audit outcomes, which the department must be cognisant of while establishing AI driven audit processes.

The department must implement robust model validation procedures to assess the accuracy, reliability, and generalisability of AI models using techniques such as cross-validation, sensitivity analysis, and testing on diverse datasets. The AI models must be validated against historical data, conduct sensitivity analyses, and performance must be assessed under different scenarios. Rigorous validation procedures must be put in place to confirm the integrity of the AI-based audit process.

AI implementation should strive to adhere to global standards by aligning with internationally recognized principles and regulatory frameworks that ensure responsible, ethical, and interoperable use of artificial intelligence across borders. These standards—such as those set by ISO/IEC, OECD, and UNESCO—provide comprehensive guidance on aspects like fairness, transparency, data privacy, human oversight, and accountability. Adhering to these norms would help department to mitigate risks, enhance public trust, and ensure legal compliance.

Auditee Acceptance of AI-Driven Audits

When incorporating AI into the audit processes, the issue of auditee acceptance becomes a critical consideration. Auditee may challenge the validity and reliability of the AI models employed by auditors, potentially disputing the inferences and conclusions drawn from such models and may even leverage their own AI-based analysis to present alternative findings, further complicating the audit process and necessitating robust mechanisms to address these concerns.

To address auditees' concerns about AI in audit processes, auditors must establish transparent and collaborative mechanisms, actively engage with auditees, address their queries, and demonstrate the reliability of AI-based audits. Informing auditees about planned audit process, seeking their suggestions, and addressing their concerns can foster greater trust and acceptance of AI-driven audit findings.

Safety, Security and Privacy issues

The department holds or accesses the auditee data in a fiduciary capacity and appropriate care must be taken in ensuring safety, security and privacy aspects of such data held. The department must also put in place robust systems like regular security assessments and audits and should have secure model development and deployment practices.

Strengthening the existing IT Governance structures in the department

IA&AD had been a pioneer in adopting IT based solutions in the functional realm, and the department has robust governance mechanisms for IT in place. However, considering the changed landscape of risk

and threat pertaining to the AI technology, the department must strengthen its existing structures appropriately, by building additional infrastructure, infusing necessary skillsets and resources.

The AI governance structure should be aligned to the national societal values and principles, following the AI protocols and frameworks established by the government of India.

The governance of AI systems involves creating a framework that ensures that AI is developed, deployed, and monitored in ways that are ethical, transparent, secure, and aligned with societal values. This governance structure needs to be flexible, adaptable, and responsive to the rapid advancements in AI technology, while also ensuring accountability and mitigating risks such as bias, discrimination, and privacy violations.

4.2. Artificial Intelligence – Implementation Strategy

IA&AD must adopt three-pronged strategic approach and the following core areas are to be focussed, in the implementation of AI:

- ✓ Research and Application
- ✓ Training and Capacity building
- ✓ Infrastructure and resource requirements

4.2.1. Research and Application

AI technology is still under evolution and an environment of facilitating and conducting research is essential to keep pace with the changes in the ecosystem. This will allow IA&AD to explore new audit evidencing tools/ techniques, innovation in audit execution and result in efficient and effective audits.

AI Research Collaborations with National premier institutions

Premier institutions/ corporate partners/ government departments/institutes of excellence like IITs/IIMs may be identified for advancing AI research and developing in-house AI models and knowledge sharing.

Central repository of AI models/ Case studies

A central repository which would be the single source for all type of AI related resources, viz. AI Models, tools, Compendiums and periodical dissemination of AI related information to IAAD, may be built by the department. Periodic circulation of the material amongst field offices for information may be done along with support and guidance towards individual requests by the owner of the central repository.

Collaboration with SAI counterparts

It is also necessary for the department to involve the international community of SAIs in the newly emerging field of AI in Audit. Collaborative efforts with other SAIs through bilateral arrangements, multilateral forums like INTOSAI, ASOSAI etc. will not only help create a pool of common resources but also enable standardisation of the processes of audit of AI and usage of AI in audits.

4.2.2. Training and Capacity Building

IA&AD has sufficiently invested in and developed a talent pool in IT audits and data analytics, over the last decades. Taking it forward, IA&AD may identify and develop a selected group of officials, preferably from all levels, and be intensely up skilled in AI technology.

Up-skilling in AI

The department may come up with a systematic training plan to upskill the in-house resources by way of in-house training programs in iCISA and designated RCB&KIs. Training programs of department should include comprehensive modules on AI implementation and audit of AI systems. In addition, various online (self-learning) certification courses may be identified in collaboration with national premier institutions and staff members who finish the certification may be incentivised. The department may devise mechanisms like hackathons and competitions in promoting AI capacity building, fostering creativity, and problem-solving skills among participants.

4.2.3. Infrastructure and resource requirements

Data Management facility

AI technology involves managing huge data sets and thus necessitates suitable high-capacity computing infrastructure. The success of AI implementation in department revolves around strong infrastructure, including energy-efficient computing and data storage at scale. The existing infrastructure at CDMA may be upgraded to an AI facility, with in-built AI/ML capabilities to utilise the datasets available with IA&AD as well as auditee organisations.

Scalable data systems for the department would be critical as growing volumes of digital data across auditee organizations is a reality. IA&AD should be able to handle both horizontal scaling, i.e.- across organizations, and vertical scaling, which is domain specific. This would enable department to manage variety and velocity without compromising on speed, performance and quality.

The CDMA may be the nodal centre for use of AI and to promote AI related activities in IA&AD. The Centre may be made responsible for activities like research projects, to develop AI models, hand hold/monitor execution of AI related audits in field offices, to implement capacity enhancement measures, to act as repository of AI models and audits, and to ensure continuous dissemination of AI related information to IA&AD.

AI Tools

Identification and adoption of suitable tools for implementation and utilisation of AI is a crucial component. The CDMA is developing tools/applications based on AI/Machine learning algorithms to be hosted on CDMA Portal which could be used by field offices by login into the portal. These tools/applications will be both downloadable and web-based, using which IA&AD officials can do their analytical jobs using advanced techniques of AI. These AI tools and libraries can be used at various stages of Audit process viz Audit planning, Audit sampling, Audit execution and Audit reporting.

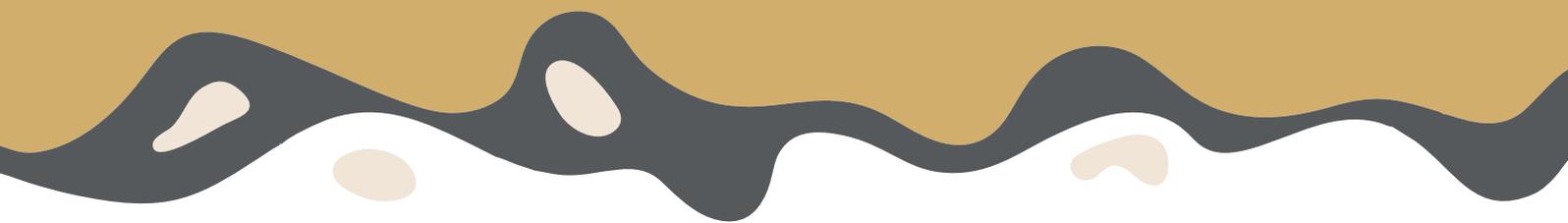
CAG's AI audit tools must furnish clear and tangible takeaways to users without the need for deep technical expertise. The department may come up with comprehensive frameworks and guidelines for using these tools.

AI Talent Recruitment Platform

As stated earlier, technically skilled human resource is vital to implement a robust AI based audit framework in the department. Hence, the department must explore accessing AI skillsets from the market by way of recruitment as consultants/advisors even research interns.



CONCLUSION





Conclusion

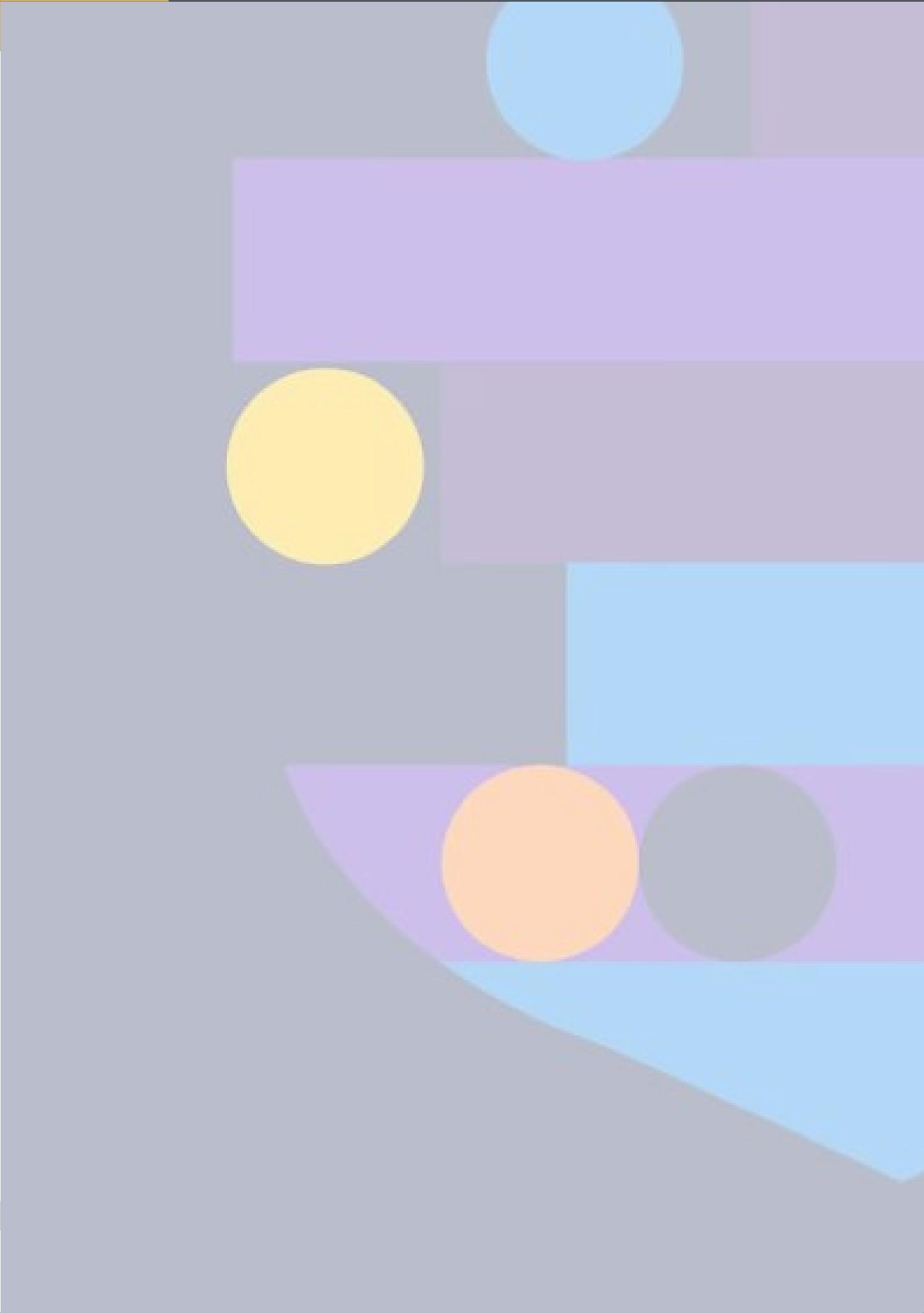
IA&AD has been a pioneer in embracing cutting-edge technologies to enhance the audit process. By use of technologies like GIS analysis, Machine learning, AI and advanced data analysis, IA&AD has not only improved the efficiency but also significantly enhanced the accuracy and depth of its audits. This forward-looking approach not only keeps IA&AD at the forefront of audit innovation but also provides invaluable insights and assurance to stakeholders.

It is beyond doubt that AI is inevitable and hence its adoption in IA&AD is essential and necessary. Keeping up with the technological advancements will ensure that the Department is not outdated and can effectively navigate the complexities.

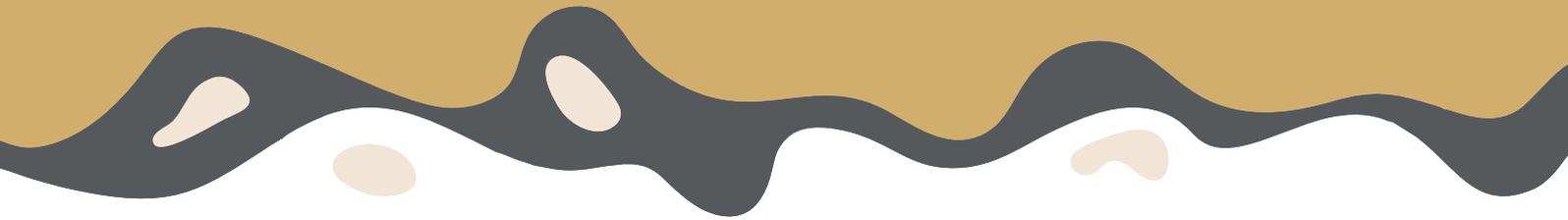
Department may draw up a time bound action plan based on the above-mentioned strategy to achieve the objectives. It is important to delineate tasks, allocate resources and define roles and responsibilities to implement the strategy in a systematic and timely manner.

The tone at the top in adopting the AI strategy would enable percolation of the same in the department. By aligning the long-term strategic plan of the department with this AI strategy, the senior management can effectively inspire the entire workforce through consistent communication and leadership actions. This approach will also ensure clarity and alignment in goals and fosters a sense of purpose and direction among staff members.

AI technologies offer unprecedented capabilities to the auditor, resulting in increased audit effectiveness. With adoption of AI strategy, IA&AD can continue to fulfil its constitutional mandate with more vigour and vibrancy.



ANNEXURE





Annexure-1: Regulatory Frameworks

National perspective based on Acts and regulations issued by the authority within India

Ministry of Law and Justice

1. IT Act 2000 and IT (Amendment) Act 2008
2. The Digital Personal Data Protection Act, 2023

NITI Ayog

3. National Strategy for Artificial Intelligence June 2018 #AIforAll
4. Responsible AI #AIforALL Approach Document for India Part 1 – Principles for Responsible AI FEB 2021
5. Responsible AI #AIforALL Approach Document for India: Part 2 - Operationalizing Principles for Responsible AI AUG 2021

MEITY (Ministry of Electronics and Information Technology)

6. Responsible AI Architect's Guide
7. Report of Committee - A on Platforms and Data on Artificial Intelligence - July 2019
8. Report of Committee - B on Leveraging A.I For Identifying National Missions in Key Sectors – July 2019
9. Report of Committee – C on Mapping Technological Capabilities, Key Policy Enablers Required Across Sectors, Skilling and Re-Skilling, R&D – July 2019
10. Report of Committee – D on cyber security, safety, legal and ethical issues - July 2019

Securities and Exchange Board of India

11. SEBI | Reporting for Artificial Intelligence (AI) and Machine Learning (ML) applications and systems offered and used by Market Infrastructure Institutions (MIIs)
12. Reporting for Artificial Intelligence (AI) and Machine Learning (ML) applications and systems offered and used by Mutual Funds

Department of Biotechnology -Ministry of Science & Technology - Government of India

13. Biological Data Storage, Access and Sharing Policy of India (Draft 2019)
14. The DNA Technology (Use and Application) Regulation Bill - 2019

Indian Council of Medical Research

15. National Ethical Guidelines for Biomedical and Health Research involving Human Participants - 2017
16. National Guidelines for Gene Therapy Product Development and Clinical Trials - 2019

Department of Telecommunications

17. Indian Artificial Intelligence Stack – AI Standardisation Committee – Sep 2020

Global Perspective based on MOU and Agreement entered at international level

ISACA

1. ISACA Publication on Auditing Artificial Intelligence

UNESCO

2. Recommendation on the Ethics of Artificial Intelligence – 2022

GPAI

3. The Global Partnership for Artificial Intelligence -GPAI Reports and Recommendations

Quad countries (Australia, India, Japan, and the United States of America)

4. Quad Principles on Technology Design, Development, Governance, and Use

Indo-U.S. Science and Technology Forum

5. U.S. - India Artificial Intelligence (USIAI) Initiative

ISO

6. ISO/IEC JTC 1/SC 42 – Standard for Artificial intelligence

Annexure-2: Illustrative checklists

A. An illustrative AI maturity level model

|  AI maturity level |  Usage Examples |  Audit Focus |
|--|---|--|
| Experimental (Limited standalone Pilot Projects) | <ul style="list-style-type: none"> ✓ Chatbots for basic citizen inquiries ✓ Simple data models | <ul style="list-style-type: none"> ✓ Data governance, ✓ Algorithm design ✓ Vendor selection |
| Emerging and Functional (Introducing AI into specific workflows) | <ul style="list-style-type: none"> ✓ AI-assisted document processing ✓ Predictive models for budget or demand of public service, etc. ✓ AI based scheduling and flows management | <ul style="list-style-type: none"> ✓ Reliability ✓ Explainability of AI-driven decisions (justifying recommendations) ✓ Performance monitoring |
| Strategic and Integrated (Core functionalities rely heavily on AI to deliver critical services or automate complex tasks) | <ul style="list-style-type: none"> ✓ AI assisted fraud detection or taxation processing ✓ AI assisted routing of traffic, electricity, etc . | <ul style="list-style-type: none"> ✓ Robustness and security ✓ Compliance with evolving AI regulations (e.g., data privacy laws, bias etc.) ✓ Ongoing model governance and model performance |
| Transformative (Widespread adoption of large Standalone AI platforms impacting society and service delivery) | <ul style="list-style-type: none"> ✓ Nationwide AI platform for public services ✓ Departmental or National level Foundation Models | <ul style="list-style-type: none"> ✓ Readiness Assessment, Conformity assessment, Impact Assessment ✓ Public consultation and process chosen ✓ Comprehensive Data and IT assessment ✓ Assessment of full cycle from design, development, testing, deployment to performance. |

⁷ Foundation models are very powerful artificial intelligence systems trained on huge datasets

B. Illustrative check list for document assessment

|  Document Type |  Usage Examples |  Assessment |
|---|--|--|
| Model Card | <ul style="list-style-type: none"> ✓ Model architecture ✓ Training data description ✓ Performance metrics ✓ Compliance to Laws | <ul style="list-style-type: none"> ✓ Algorithm choices, data sources, expected accuracy ✓ Risks: Data bias, limitations in model performance |
| Data Management Documentation | <ul style="list-style-type: none"> ✓ Data governance policies ✓ Data quality reports ✓ Data security protocols | <ul style="list-style-type: none"> ✓ How data is collected, stored, processed, and protected ✓ Risks: Data privacy violations, data breaches, data integrity issues |
| Model Development & Validation Records | <ul style="list-style-type: none"> ✓ Model training logs ✓ Validation dataset results ✓ Code repositories | <ul style="list-style-type: none"> ✓ Methods used to train and evaluate the model ✓ Risks: Overfitting, underfitting |
| Performance Monitoring Reports | <ul style="list-style-type: none"> ✓ Performance dashboards ✓ Anomaly detection logs ✓ Model drift analysis | <ul style="list-style-type: none"> ✓ Real-world performance, system stability, potential degradation ✓ Risks: Model-decay, inaccurate predictions, unexpected system behaviour |

C. Illustrative check list - Evaluate data inputs

|  AI Model and Component |  Assessment |  Evaluation Methods |
|--|--|---|
| Bias and Fairness | <ul style="list-style-type: none"> ✓ Analyse the capacity to identify potential biases or fairness concerns. ✓ Conduct statistical analyses to detect bias in decision outcomes across demographic groups or sensitive attributes. ✓ Assess whether the model perpetuates existing disparities. | <ul style="list-style-type: none"> ✓ Bias Detection: Statistical tests like Chi-square ✓ Fairness Metrics: Demographic parity, ✓ Disparity Impact Analysis |
| Algorithmic Transparency | <ul style="list-style-type: none"> ✓ Evaluate the transparency of the AI algorithm. ✓ Understand how decisions are made, and which factors influence the decision-making process. ✓ Assess interpretability and explainability. | <ul style="list-style-type: none"> ✓ Interpretability Metrics ✓ Explainability Tools ✓ Robustness Tests |
| Model Performance Evaluation and Decision-Making Accuracy and Performance | <ul style="list-style-type: none"> ✓ Functional testing (system meets specified requirements and behaves as expected) ✓ Performance testing ✓ Security testing (penetration tests, data leaking etc.) ✓ Bias and Fairness testing - Robustness testing. ✓ Replicability, accuracy ✓ Consider trade-offs between accuracy and fairness. | <ul style="list-style-type: none"> ✓ Performance Metrics: like Precision, recall, accuracy, F1 score, ROC Curves⁸ and AUC |
| Model Validation and Testing | <ul style="list-style-type: none"> ✓ Examine procedures for validation and testing. Ensure the model meets specified requirements and performs as intended | <ul style="list-style-type: none"> ✓ Validation Techniques ✓ Evaluation Metrics: MSE, RMSE⁹ ✓ Review of Performance Documentation |

⁸ ROC is a graphical representation of the trade-off between true positive rate and false positive rate at different classification

⁹ Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) are measures that quantify the differences between predicted values and observed values

|  AI Model and Component |  Assessment |  Evaluation Methods |
|---|---|---|
| Regulatory Compliance | <ul style="list-style-type: none"> ✓ Verify compliance with legal and regulatory requirements, e.g. data protection laws, anti-discrimination regulations, and industry standards. | <ul style="list-style-type: none"> ✓ Compliance Checklist and Gap Analysis |
| Stress Testing and Scenario Analysis | <ul style="list-style-type: none"> ✓ Test performance under different conditions to assess resilience and limitations. ✓ See if the developer and user have tested for vulnerabilities and failure modes (Mechanism for recovery, human-in-loop¹⁰) | <ul style="list-style-type: none"> ✓ Using noisy data, simulate data drift, to test adverse conditions response. ✓ Robustness & Adaptability Analysis ✓ Identify vulnerabilities or failure modes. |

¹⁰ Human-in-the-loop refers to systems or processes that involve human interaction or oversight at some point, rather than being fully automated.

D. Illustrative check list -Features of data and assessment

|  Feature of data |  Assessment |  Evaluation Methods |
|---|--|---|
| Data Collection Processes | <ul style="list-style-type: none"> ✓ Review the processes and procedures involved in collecting the data. | <ul style="list-style-type: none"> ✓ Review documentation ✓ Interview stakeholders ✓ Assess data collection methods |
| Data Relevance and Completeness | <ul style="list-style-type: none"> ✓ Assess the relevance and completeness of the data. | <ul style="list-style-type: none"> ✓ Examine scope and coverage ✓ Check for completeness and missing values |
| Data Accuracy and Consistency | <ul style="list-style-type: none"> ✓ Evaluate the accuracy and consistency of the data. | <ul style="list-style-type: none"> ✓ Conduct data validation checks ✓ Perform data profiling |
| Data Bias and Fairness | <ul style="list-style-type: none"> ✓ Examine the data for potential biases. | <ul style="list-style-type: none"> ✓ Use statistical techniques |
| Data Privacy and Security | <ul style="list-style-type: none"> ✓ Evaluate the privacy and security measures for protecting data. | <ul style="list-style-type: none"> ✓ Assess data anonymization techniques ✓ Review access controls and encryption methods ✓ Ensure compliance with regulations |
| Data Governance and Documentation | <ul style="list-style-type: none"> ✓ Review data governance policies and procedures. | <ul style="list-style-type: none"> ✓ Assess data governance frameworks ✓ Review data management practices ✓ Examine metadata documentation |



SUPREME AUDIT INSTITUTION OF INDIA
लोकहितार्थ सत्यनिष्ठा
Dedicated to Truth in Public Interest



**Comptroller and
Auditor General of India**
<http://www.cag.gov.in>

