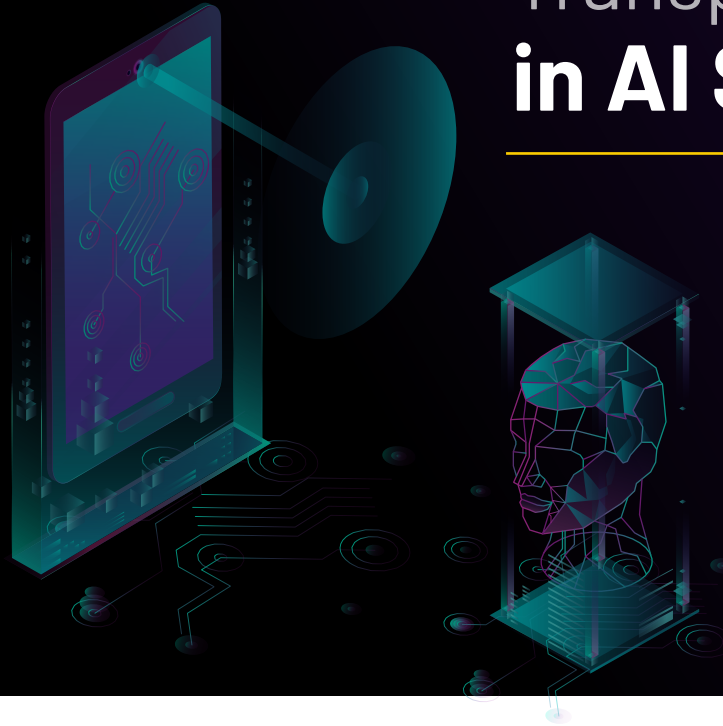


Transparency in AI Systems

Introduction



Transparency in Artificial Intelligence (AI) refers to understanding how, when and why an AI makes decisions.¹ It involves enabling relevant stakeholders, including users, to understand the inner workings of an AI system, from its development to deployment.² AI is often referred to as a “black box” because of the difficulty in determining how it functions or generates outputs. Transparency helps to open and understand this black box by making the data, algorithms, and decision-making processes accessible to stakeholders.³

Regulatory frameworks are also reinforcing the importance of AI transparency. For instance, section 37 of the Nigeria Data Protection Act 2023 and section 35 of the Kenyan Data Protection Act address the important issue of automated decision-making, a key aspect of AI transparency. The regulations provide that where such processing is allowed by law, contract, or consent, data controllers must ensure safeguards like human intervention, the right to contest decisions, and the ability of the data subject to express their point of view. These measures help ensure AI systems remain fair, accountable, and understandable. The Kenyan Data Protection Act further provides more detailed procedural safeguards. It requires data controllers or processors to notify data subjects promptly after an automated decision is made and to respond within a reasonable time to requests for reconsideration or new decisions. Also, Recital 27 of the European Union (EU) AI Act mandates that AI systems should be developed and used in a way that ensures appropriate traceability and explainability.⁴ It also requires that users are made aware when they are communicating or interacting with AI systems and deployers are aware of the system’s capabilities, limitations, and user rights.

Similarly, Principle 1.3 of the Organisation for Economic Cooperation and Development (OECD) AI Principles encourages AI actors to commit to transparency and responsible disclosure. This includes providing meaningful, context-appropriate, and up-to-date information about AI systems.⁵

¹Transparency and Explainability (Principle 1.3) (OECD.AI Policy Observatory) <<https://oecd.ai/en/dashboards/ai-principles/P7>> accessed 25 April 2025

²George Lawton ‘AI Transparency: What Is It and Why Do We Need It? (TechTarget, 10 September 2024) <<https://www.techtarget.com/searchcio/tip/AI-transparency-What-is-it-and-why-do-we-need-it>> accessed 24 April 2025

³Alexander Jonker, Alice Gomstyn and Amanda McGrath ‘What Is AI Transparency? (IBM, 6 September 2024) <<https://www.ibm.com/think/topics/ai-transparency>> accessed 24 April 2025

⁴Key Issues: Transparency Obligations (EU AI Act) <<https://www.euaiact.com/key-issue/5>> accessed 25 April 2025

⁵n 1.

Key Components of AI Transparency



AI transparency extends beyond simply understanding how a system works; it includes broader concepts such as explainability, interpretability, data governance, and accountability. However, despite its importance, transparency is often lacking in real-world applications. A common example is the growing number of digital service providers and fintech companies that promote their use of AI, especially in areas like credit scoring or fraud detection without clearly disclosing the use of AI in their privacy notices.⁶ This lack of clear communication about AI reflects a significant transparency gap and highlights why these broader principles of AI transparency explained below are essential.

- AI explainability is the ability of an AI system to clearly, accurately and simply explain how it reaches its decisions and actions. A clear and accurate explanation gives users an understanding of the AI's decision making process.⁷
- Interpretability focuses on making the internal workings of the AI system understandable to humans. It aims to shed light on how specific outputs are generated based on the system's internal workings.⁸
- Data governance emphasises the quality and sources of the data used to train AI systems. It ensures that stakeholders are informed about how datasets are collected, managed, and validated.⁹
- Transparency also involves accountability which ensures AI systems and the organisations that deploy them are held responsible for the systems' actions, impacts and decisions.¹⁰

⁶Digital Lending: Inside the Pervasive Practice of LendTechs in Nigeria' (Tech Hive Advisory Africa, 23 June 2021)<<https://www.techhiveadvisory.africa/report/digital-lending-inside-the-pervasive-practice-of-lendtechs-in-nigeria>> accessed 28 April 2025.

⁷Hannah Wren 'What Is AI Transparency? A Comprehensive Guide' (Zendesk, 18 January 2024) <<https://www.zendesk.com/blog/ai-transparency/>> accessed 28 April 2025.

⁸ibid.

⁹George Lawton 'AI Transparency: What Is It and Why Do We Need It? (TechTarget, 10 September 2024) <<https://www.techtarget.com/searchcio/tip/AI-transparency-What-is-it-and-why-do-we-need-it>> accessed 24 April 2025

¹⁰Ben Chester Cheong, 'Transparency and Accountability in AI Systems: Safeguarding Wellbeing in the Age of Algorithmic Decision-Making' (2024) 6 Frontiers in Human Dynamics <<https://www.frontiersin.orghttps://www.frontiersin.org/journals/humandynamics/articles/10.3389/fhumd.2024.1421273/full>> accessed 28 April 2025

Benefits of AI Transparency

There are several benefits of implementing transparency in AI systems, including the following;

- **Accountability:** AI transparency ensures accountability. It is easier to identify and hold the right persons responsible for mistakes or wrongs caused by AI as the entire decision-making process becomes more visible and accessible.
- **Trust and Confidence:** Transparency builds trust and confidence in AI systems. When users understand how an AI system works, including the datasets used in its training, they are more likely to trust and adopt AI solutions.¹¹
- **Collaboration and Knowledge Sharing:** Transparent processes encourage collaboration among developers, deployers, and other stakeholders by fostering openness and facilitating knowledge exchange across the AI ecosystem.¹²
- **Informed Decision-Making:** Transparent AI systems enable users and stakeholders to make informed choices, as they can better understand the underlying logic and assess the implications of AI-driven recommendations.¹³

Ultimately, transparency promotes the long-term sustainability of AI systems as it fosters trust, accountability, and adaptability.¹⁴



Practical Ways Companies Can Implement AI Transparency

To implement AI transparency, companies can take these actionable steps;

1. **Stakeholder involvement:** Engage stakeholders throughout the entire AI lifecycle, from development to deployment. This can be achieved through regular and open communication that helps identify stakeholder concerns and provides valuable feedback mechanisms.¹⁵
2. **Notification:** Inform users when they are interacting with or being monitored by an AI system. This could involve clear labelling, which provides clear and visible indicators of involvement of AI systems such as "AI-assisted interaction" or "Powered by AI." This ensures users are aware of AI involvement and can make informed decisions.¹⁶

¹¹ "AI Risk Management: Transparency & Accountability" (Lumenova AI, 28 May 2024) <<https://www.lumenova.ai/blog/ai-risk-management-importance-of-transparency-and-accountability/>> accessed 28 April 2025

¹² AI Transparency: Building Trust in AI' (Mailchimp) <<https://mailchimp.com/resources/ai-transparency/>> accessed 28 April 2025

¹³ Lee Ditmarr 'What Does Transparency Really Mean in the Context of AI Governance?' (OCEG, 8 November 2024) <<https://www.oceg.org/what-does-transparency-really-mean-in-the-context-of-ai-governance/>> accessed 24 April 2025

¹⁴ n 12

¹⁵ Raj Sharma, 'Why transparency is the key to unlocking AI's full potential' (World Economic Forum, 2 January 2025) <<https://www.weforum.org/stories/2025/01/why-transparency-key-to-unlocking-ai-full-potential/>> accessed 25 April 2025

¹⁶ Danielle Draper, 'What's in the Box: Tools that Enhance AI Transparency' (Bipartisan Policy Center, 5 December 2023) <<https://bipartisanpolicy.org/blog/whats-in-the-box-tools-that-enhance-ai-transparency/>> accessed 25 April 2025

3. **Disclosure:** Document and publicly disclose key information about the AI systems. This includes model name and purpose, risk level, model generation and training data, testing accuracy and bias metrics, explainability and fairness metrics and contact information.¹⁷ To ensure accessibility and transparency, Disclosures can be published via official policy pages or through "model cards" which are short, detailed documentation that outlines a model's capabilities, limitations, and ethical considerations.¹⁸
4. **Employee training:** Provide regular training for employees on ethical AI principles, including transparency, fairness, and accountability. Refresher sessions should be conducted periodically to maintain awareness and reinforcement of these principles.¹⁹
5. **Transparent Models:** Companies should consider using simpler models, like decision trees or rule-based systems, whenever possible. These models help make AI systems easier to understand without compromising much on performance.²⁰
6. **Transparency by design and default:** Companies should design models with inherent transparency whenever feasible, as opposed to retrofitting transparency measures after deployment.
7. **Continuous Monitoring:** Companies should monitor AI systems continually. Document any retraining, updates, or changes made to models, and communicate them to stakeholders. Version control tools should be used to track model changes over time.²¹
8. **Audit:** Conduct regular audits to assess the risks inherent in the AI system.²² Upon completion, audit findings should be documented and shared with relevant stakeholders.
9. **Data governance:** Companies should implement proactive data governance practices, including publishing privacy notices on their websites detailing how personal data is collected, processed, and stored.²³ They should also assess the validity of data source, disclose data origins to stakeholders, track data quality metrics, and validate data diversity and fairness.²⁴ Additionally, companies should develop and maintain a comprehensive risk registry that documents potential biases, inaccuracies, and security vulnerabilities within AI systems along with clear, documented risk mitigation steps.²⁵

¹⁷ Alexander Jonker, Alice Gomstyn and Amanda McGrath 'What Is AI Transparency? (IBM, 6 September 2024) <<https://www.ibm.com/think/topics/ai-transparency>> accessed 24 April 2025

¹⁸ ibid

¹⁹ Jeremy Werner, 'The Role of Transparency and Accountability in AI Adoption' (BABL AI, 13 September 2024) <<https://babl.ai/the-role-of-transparency-and-accountability-in-ai-adoption/>> accessed 28 April 2025

²⁰ Narayana pappu, 'AI Explainability 101: Making AI Decisions Transparent and Understandable' (Zendata) <<https://www.zendata.dev/post/ai-explainability-101>> accessed 30 April 2025.

²¹ Lee Ditmarr 'What Does Transparency Really Mean in the Context of AI Governance?' (OCEG, 8 November 2024) <<https://www.oceg.org/what-does-transparency-really-mean-in-the-context-of-ai-governance/>> accessed 24 April 2025

²² Danielle Draper, 'What's in the Box: Tools that Enhance AI Transparency' (Bipartisan Policy Center, 5 December 2023) <<https://bipartisanpolicy.org/blog/whats-in-the-box-tools-that-enhance-ai-transparency/>> accessed 25 April 2025

²³ n 22.

²⁴ Ben Chester Cheong, 'Transparency and Accountability in AI Systems: Safeguarding Wellbeing in the Age of Algorithmic Decision-Making' (2024) 6 Frontiers in Human Dynamics <<https://www.frontiersin.orghttps://www.frontiersin.org/journals/humandynamics/articles/10.3389/fhumd.2024.1421273/full>> accessed 28 April 2025

²⁵ Hannah Wren 'What Is AI Transparency? A Comprehensive Guide' (Zendesk, 18 January 2024) <<https://www.zendesk.com/blog/ai-transparency/>> accessed 28 April 2025

- 10. AI Guidelines:** Companies should develop and implement AI-related guidelines or policies. They may choose to adopt external frameworks, such as the NIST AI Risk Management Framework, or develop internal guidelines tailored to their operations.²⁶

Challenges of Implementing AI Transparency

As stated earlier, transparent AI practices offer many benefits but also raise certain challenges, especially with the advent of generative AI, which has made achieving transparency more challenging due to its complexity.²⁷

One significant challenge is balancing competing interests, such as privacy and intellectual property, with transparency. The more information is disclosed about the inner workings of AI, the easier it may become for competitors or malicious actors to exploit it.²⁸ This risk has made companies reluctant to disclose proprietary algorithms, which hinders transparency efforts. Therefore, transparency regulatory requirements must be carefully crafted to promote openness without compromising individual privacy or proprietary information.²⁹

Additionally, it can be difficult to explain intricate and complex programmes and algorithms to users in a clear and simple manner.³⁰ This can be resolved by developing simplified explanations to make AI more understandable.

Another major challenge to AI transparency is the absence of universally accepted standards and regulations. Without clear and consistent guidelines, companies often struggle to determine the specific steps required to achieve meaningful transparency.³¹ Thus, standardised frameworks are essential as they help organisations align their practices with industry expectations.

Finally, the evolving nature of AI systems adds another layer of complexity. As models grow more sophisticated and change over time through updates, modifications, or retraining on new datasets, it becomes increasingly difficult to trace and explain how they generate their outputs.³² These changes can alter the AI's decision-making processes, making it challenging to maintain consistent transparency over time.³³

²⁶ Jeremy Werner, 'The Role of Transparency and Accountability in AI Adoption' (BABL AI, 13 September 2024) <<https://babl.ai/the-role-of-transparency-and-accountability-in-ai-adoption/>> accessed 28 April 2025

²⁷ George Lawton 'AI Transparency: What Is It and Why Do We Need It?' (TechTarget, 10 September 2024) <<https://www.techtarget.com/searchcio/tip/AI-transparency-What-is-it-and-why-do-we-need-it>> accessed 24 April 2025

²⁸ Alexander Jonker, Alice Gomstyn and Amanda McGrath 'What Is AI Transparency?' (IBM, 6 September 2024) <<https://www.ibm.com/think/topics/ai-transparency>> accessed 24 April 2025

²⁹ n 23.

³⁰ 'AI Transparency: Building Trust in AI' (Mailchimp) <<https://mailchimp.com/resources/ai-transparency/>> accessed 28 April 2025

³¹ *ibid.*

³² 'AI Risk Management: Transparency & Accountability' (Lumenova AI, 28 May 2024) <<https://www.lumenova.ai/blog/ai-risk-management-importance-of-transparency-and-accountability/>> accessed 28 April 2025

³³ Hannah Wren 'What Is AI Transparency? A Comprehensive Guide' (Zendesk, 18 January 2024) <<https://www.zendesk.com/blog/ai-transparency/>> accessed 28 April 2025

Conclusion

Transparency in AI is a continuous journey, not a one-time effort. As AI systems evolve and adapt, companies must regularly monitor, assess, and refine their models to ensure they remain transparent, trustworthy, and aligned with ethical and operational goals. Organisations can build greater stakeholder confidence, navigate regulatory expectations, and drive more responsible AI by embedding transparency into the entire AI lifecycle, from development to deployment and beyond.

Authors:

Ridwan Oloyede
Rodiyyah Bashir