

# European Union Artificial Intelligence Act

Trustworthy AI Playbook for Enterprises

*June 2024*

AI & Partners

For more information on this publication, visit <https://www.ai-and-partners.com/>.

## About AI & Partners

**‘AI That You Can Trust’** - Your trusted advisor for EU AI Act Compliance. Unlock the full potential of artificial intelligence while ensuring compliance with the EU AI Act by partnering with AI & Partners, a leading professional services firm. We specialise in providing comprehensive and tailored software solutions for companies subject to the EU AI Act, guiding them through the intricacies of regulatory requirements and enabling responsible and accountable AI practices. At AI & Partners, we understand the challenges and opportunities that the EU AI Act presents for organisations leveraging AI technologies. Our team of seasoned experts combines in-depth knowledge of AI systems, regulatory frameworks, and industry specific requirements to deliver strategic guidance and practical solutions that align with your business objectives.

To find out how we can help you, email [contact@ai-and-partners.com](mailto:contact@ai-and-partners.com) or visit <https://www.ai-and-partners.com>.

## Business Integrity

AI & Partners defends and extends the digital rights of users at risk around the world. By combining direct technical support, comprehensive policy engagement, global advocacy, grassroots professional services, regulatory interventions, and participating in industry groups such as AI Commons, we fight for fundamental rights in the artificial intelligence age.

AI & Partners’ publications do not necessarily reflect the opinions of its clients, partners and/or stakeholders.

Foreword	(Slide 4)
Executive Summary	(Slide 5)
Introduction	(Slide 6)
The Business Case for TAI and the Role of Stakeholders	(Slide 7)
Enterprise Engagement in TAI	(Slide 8)
The Obstacles Ahead and How to Successfully Beat Them	(Slide 9)
Conclusion	(Slide 10)

## — ‘Trustworthy AI’ – Reliable. Ethical. Transparent.

When AI & Partners embarked on its European Union (EU) artificial intelligence (AI) Act (EU AI Act) preparations in 2021, the broad relevance of trustworthy artificial intelligence (TAI) to enterprises was still under discussion. The role of enterprises in expediting the adoption of TAI, particularly given other expectations placed on them in the areas of climate change, equity, diversity and inclusion, and cybersecurity, remains also unclear.

With the release of the EU AI Act proposal on **21 April 2021**, both the opportunities and risks of AI came into sharper focus. In this sense, the imperative of promoting the uptake of TAI while also ensuring the protection of health, safety, and fundamental rights gained more stakeholder confidence.

### What, then, is the specific function of enterprises in advancing the adoption of TAI?

As espoused by Metcalfe’s Law, the value of a network is proportional to the square of the number of connected users of the system. This adage emphasizes that, as the number of users of general-purpose AI (GPAI) increases, the more valuable the AI ecosystem, including actors interacting within this, becomes. A growing social network requires strong regulatory pillars to provide a robust, secure, and safe substrate on which to build. On a cross-sectoral basis, enterprises are already struggling with risks such as data breaches, privacy loss, job loss, ethical challenges, misinformation and disinformation. Regulation can help address these burdens.

To protect their organisations’ stability both now and in the future, enterprises should tackle both these immediate issues as well as the deeper implications of AI. This starts with establishing strong governance frameworks and clear principles, processes, procedures and practices to integrate TAI standards into all applications.

If successful, the term “**TAI**” will resonate globally, as high-quality, trustworthy and safe AI becomes the norm. Just as we don’t distinguish between “businesses” and “businesses that don’t fail”, the qualifier “trustworthy” will become an unspoken expectation. Today, we are at the embryonic stages of this era, as the EU AI Act, together with new AI laws and regulations, emerge to ensure that all AI applications are trustworthy.

This circles us back to our original question, for which there is an unequivocal answer: **large enterprises can and should exercise the influence afforded by their resources (e.g. operational, technological, financial, etc.) to promote the use of TAI in their organisations, in their work with clients and partners, and in the ecosystem at large.**

This white paper offers a playbook for how enterprises can accelerate the adoption of TAI, as loosely recognised by the EU AI Act, to help drive value. While it is not a comprehensive view of all levers available, we believe it provides useful guidance, theoretical examples and concrete steps for engaging with other stakeholders.

## — Solidifying firms' success in frontier technology transformation

### Risk Reduction and Growth Driver

- **It enhances safety and reliability:** Trustworthy AI systems are designed to be robust and secure, minimizing errors and vulnerabilities. This ensures consistent performance and reduces the risk of harm, fostering a safer environment for users and operators (Recital 27).
- **It aids ethical and fair decision-making:** By adhering to principles like non-discrimination and transparency, trustworthy AI promotes fairness and accountability. This builds public trust and encourages wider adoption, driving growth in various sectors (Recital 27).
- **It boosts innovation and market confidence:** Trustworthy AI frameworks provide clear guidelines and standards, reducing legal uncertainties. This encourages investment and innovation, as businesses feel more confident in developing and deploying AI technologies (Article 15).

### Engagement Methods

- **Conduct inclusive impact assessments:** Enterprises should involve relevant stakeholders, including civil society organizations and independent experts, in conducting impact assessments for high-risk AI systems. This ensures comprehensive risk evaluation and fosters trust among affected groups (Recital 96).
- **Hold balanced advisory forums:** Establish advisory forums with balanced representation from industry, SMEs, academia, and civil society. This diverse input helps shape AI strategies that are both innovative and ethically sound (Recital 165).
- **Promote voluntary codes of conduct:** Encourage the creation of voluntary codes of conduct that incorporate ethical guidelines and best practices. Engaging stakeholders in this process ensures that AI systems are developed responsibly and sustainably (Recital 165).

### Improvement Areas

- **Enhanced Data Governance:** There is a need for improved data governance frameworks to ensure high-quality, unbiased data for training AI systems (Recital 27).
- **Ethical and Inclusive Design:** Further development is required in creating AI systems that adhere to ethical guidelines, promoting diversity, non-discrimination, and fairness (Recital 20).
- **AI Literacy and Public Awareness:** Increasing AI literacy among all stakeholders, including developers, users, and the general public, is crucial (Recital 20).



## TAI Dimensions and Principles

TAI encompasses several key dimensions and principles aimed at ensuring ethical and reliable AI systems. These include *human agency and oversight*, which ensures AI serves people and respects human dignity, and technical robustness and safety, which focuses on resilience and minimizing unintended harm. **Privacy** and **data governance** ensure compliance with privacy laws and data integrity, while **transparency** promotes traceability and explainability of AI operations.

Additionally, **diversity**, **non-discrimination**, and **fairness** are crucial, ensuring AI systems avoid biases and promote equal access. Societal and environmental well-being emphasizes sustainable and beneficial AI development, monitoring long-term impacts. Lastly, accountability ensures clear mechanisms to address AI-related issues, fostering trust and ethical AI deployment. These principles collectively contribute to the design of *coherent, trustworthy, and human-centric AI systems*.



— Representing both an opportunity and incentive

## Risk drivers affecting value creation

### Bias in Data

Data sets may contain biases that lead to unfair or discriminatory outcomes, affecting the reliability and fairness of AI systems.

### Data Privacy

Inadequate data governance can lead to breaches of personal data, compromising user privacy and violating regulations.

### Data Quality

Poor-quality data can result in inaccurate AI predictions and decisions, undermining the system's effectiveness and trustworthiness.

### System Failures

Technical malfunctions or software bugs can cause AI systems to fail, leading to operational disruptions and potential harm.

### Security Vulnerabilities

AI systems may be susceptible to cyberattacks, which can compromise their integrity and lead to unauthorized access or manipulation.

### Human Oversight

Insufficient human oversight can result in AI systems making critical decisions without proper review, increasing the risk of errors.

### Transparency Issues

Lack of transparency in AI decision-making processes can erode trust and make it difficult to understand or challenge AI outcomes.

### Discrimination

AI systems may inadvertently perpetuate or exacerbate existing biases, leading to discriminatory practices against certain groups.

### Ethical

Ambiguities in accountability can make it hard to identify who is responsible for AI-related harms, affecting redress mechanisms.



— Cross-organisational and wider collaboration required

## Prioritising high impact areas

**Step 1:** Develop TAI commitments and apply its principles and practices to internal operations.

**Step 2:** Conduct TAI due diligence on the business.

**Step 3:** Engage with shareholders, and the broader ecosystem.

### Step 1

The first step for enterprises looking to integrate TAI across business functions is to become knowledgeable on AI/TAI and apply it to their own operations.

### Step 2

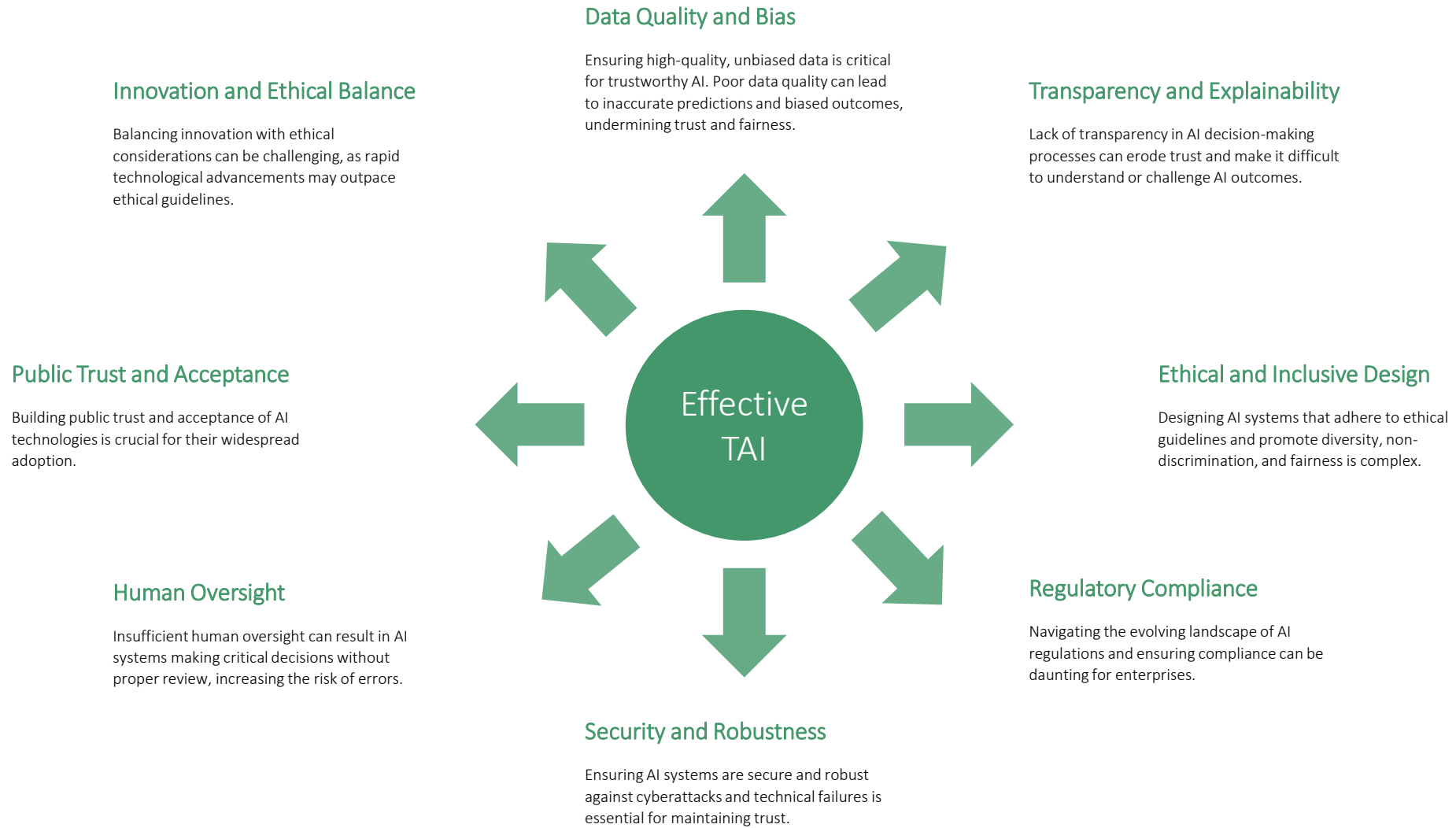
Enterprises should perform proper due diligence to assess how they and strategic partners are positioned to meet TAI principles.

### Step 3

Working with shareholders, and other players can extend enterprises' influence and help them maximize the value of their AI activities by building bespoke plans to accelerate the adoption of TAI.



# Emerging strategies for effective Trustworthy AI Engagement



## — Revolutionary potential of AI becoming clearer



### Challenges surmountable with collaborative efforts

- Enterprises play a pivotal role in this endeavour. By driving TAI, they not only help mitigate risks but also unlock opportunities for sustainable growth and innovation. This playbook serves as a strategic guide, offering insights and actionable strategies for enterprises to drive the integration of TAI in AI development and deployment.
- For enterprises aiming towards the elixir of long-term value creation, driving TAI is a strategic business decision. Through collaboration, education and proactive engagement, enterprises can help steer the AI revolution towards a future where technology operates with trustworthiness at its core, benefiting enterprises, shareholders, consumers and society at large.

— Thank you!



Amsterdam - London - Singapore



Email

[contact@ai-and-partners.com](mailto:contact@ai-and-partners.com)



Phone

+44(0)7535 994 132



Website

<https://www.ai-and-partners.com/>



Social Media

LinkedIn: <https://www.linkedin.com/company/ai-&-partners/>

Twitter: [https://twitter.com/AI and Partners](https://twitter.com/AI_and_Partners)

## — Disclaimer

This Presentation may contain information, text, data, graphics, photographs, videos, sound recordings, illustrations, artwork, names, logos, trade marks, service marks, and information about us, our lines of services, and general information may be provided in the form of documents, podcasts or via an RSS feed (“the Information”).

Except where it is otherwise expressly stated, the Information is not intended to, nor does it, constitute legal, accounting, business, financial, tax or other professional advice or services. The Information is provided on an information basis only and should not be relied upon. If you need advice or services on a specific matter, please contact us using the contact details for the relevant consultant or fee earner found on the Presentation.

The Presentation and Information is provided “AS IS” and on an “AS AVAILABLE” basis and we do not guarantee the accuracy, timeliness, completeness, performance or fitness for a particular purpose of the Presentation or any of the Information. We have tried to ensure that all Information provided on the Presentation is correct at the time of publication. No responsibility is accepted by or on behalf of us for any errors, omissions, or inaccurate information on the Presentation. Further, we do not warrant that the Presentation or any of the Information will be uninterrupted or error-free or that any defects will be corrected.

Although we attempt to ensure that the Information contained in this Presentation is accurate and up-to-date, we accept no liability for the results of any action taken on the basis of the Information it contains and all implied warranties, including, but not limited to, the implied warranties of satisfactory quality, fitness for a particular purpose, non-infringement, compatibility, security, and accuracy are excluded from these Terms to the extent that they may be excluded as a matter of law.

In no event will we be liable for any loss, including, without limitation, indirect or consequential loss, or any damages arising from loss of use, data or profits, whether in contract, tort or otherwise, arising out of, or in connection with the use of this Presentation or any of the Information.