Trusted AI – Digital Innovation

KPMG Switzerland

Amidst the evolving Generative AI landscape, businesses face an unregulated "Wild West". Our pioneering service offering, Trusted AI, harnesses AI's potential to help organizations navigate complex and legally binding regulations. Providing cutting-edge AI-powered tools, we enable businesses to achieve compliance while upholding ethical standards, allowing confident adoption of AI's transformative capabilities within a responsible regulatory framework.

Regulatory AI Perspective

Risks and Baselines
The underlying drivers for AI regulation are the need to comply with existing laws and regulations, and the business-centric approach that aims to build trust, mitigate risks, and promote sustainable growth in the rapidly evolving AI landscape. The EU AI Act holds significance for Switzerland, as FINMA is likely to follow with this regulation.

External Requirements and Regulations
[1] The EU AI Act is imminent and represents a significant step towards regulating AI technologies within the European Union. This comprehensive legislation aims to establish clear guidelines and rules for AI development and address concerns related to ethics, security and fundamental rights. With the EU AI Act on the horizon, banking and insurance companies operating in the EU or in Switzerland, will need to prepare to comply with the new legal and regulatory framework to ensure responsible, trustworthy and accountable use of AI systems. Due to heightened market pressures, the EU AI Act will be in effect starting from January 2024, and will include a two-year transition period for implementation.

Intrinsic Drivers
Companies endeavour to embrace responsible and trusted AI practices to build trust with their customers and stakeholders, fostering a positive brand reputation.
By adhering to ethical AI principles and mitigating bias, companies can avoid potential legal and financial risks associated with discriminatory or harmful AI outcomes.

Trusted AI also facilitates long-term sustainability and growth by ensuring that AI systems align with societa values and contribute positively to the well-being of communities and the environment.

Risks of doing nothing

01 **Potential Legal and Regulatory Violations**
Companies may face legal and regulatory sanctions if AI systems lead to discrimination, privacy breaches, or being non-compliant.

02 **Reputational Damage**
Lack of responsible AI can result in negative publicity, eroding customer trust and damaging the company's reputation.

03 **Bias and Discrimination**
Unchecked AI solutions can perpetuate bias can lead to unfair treatment and discrimination against certain groups resulting in societal backlash and being disposed to reputational damage.

04 **Economic Loss**
AI systems with unintended consequences or unpredictable behavior can lead to financial losses, as errors, inefficiencies, or security breaches impact operations, customer satisfaction, and revenue.

[1] Approach Definition: https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0001.02/DOC_2&format=PDF

Definition of AI Systems

Various definitions of AI systems exist, reflecting the complexity and diversity of artificial intelligence technologies. We present two important definitions from the EU AI Act and the OECD. These jurisdictions attempt to characterize AI systems while considering their societal impact to guide the way for trusted AI development and deployment.

Definition of AI Systems
What precisely constitutes AI systems under the EU AI Act? Presently, the directive defines it as follows, though this definition remains subject to potential changes.

But other jurisdictions define AI-systems from various angles.

EU AI Act Definition [2]
"(1) An artificial intelligence system (AI system) is a machine-based system designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments;"

The definition is taken from the latest released draft of the EU AI Act dated 26th January 2024.

A certain similarity of the EU AI Act definition with the OECD Definition (see below) is obvious.

OECD Definition [3]
The OECD Directorate for Science, Technology, and Innovation (STI) conducts various activities to study the impact of information and communication technologies (ICTs) on sustainable economic growth and social wellbeing. The OECD defines an AI system as "a machinebased system that is capable of influencing the environment by producing recommendations, predictions or other outcomes for a given set of objectives. It uses machine and/or human-based inputs/data to:

1) perceive environments;
2) abstract these perceptions into models; and
3) use the models to formulate options for outcomes. AI systems are designed to operate with varying levels of autonomy."

Other Definitions
ISO (International Organization for Standardization) and IEC (International Electrotechnical Commission) form the global standardization system, with national bodies participating in developing International Standards through specialized technical committees, fostering harmonization and innovation.

According to the ISO/IEC 22989:2022 3.1.4 [4] standard, an AI system is defined as an "engineered system that generates outputs such as content, forecasts, recommendations or decisions for a given set of human-defined objectives".

What is trusted AI and why is it needed

Trusted AI refers to the ethical and responsible development, deployment, and use of artificial intelligence technologies. According to the published ethical guidelines by the European Commission, trustworthy AI should be lawful, ethical, and robust.

Companies in the financial sector, such as banks, insurers, but also software service providers and vendors need responsible AI to build trust with customers, stakeholders and society by ensuring that AI systems are fair, transparent and do no harm. It helps companies avoid legal and reputational risks, promotes innovation in socially beneficial ways, and ensures that AI is consistent with ethical principles and contributes to long-term sustainability and positive impacts on communities and the environment.

How KPMG can help
KPMG's Trusted AI offering is a set of frameworks, controls, processes, tools, and platforms to ensure AI systems are being designed and deployed in a trustworthy and ethical manner so that companies can accelerate value. We understand trusted AI is a complex business, regulatory, and technical challenge, and we are committed to helping clients put it into practice.

---

² EU AI Act: https://artificialintelligenceact.eu/the-act
³ OECD AI Definition: https://www.oecd-ilibrary.org/docserver/2448f04b-en.pdf?expires=1689962519&id=id&accname=guest&checksum=65753939 D1B3615FC89CF0A8EE506E4E
⁴ ISO/IEC 22989:2022: https://webstore.iec.ch/publication/77839

Fairness
The fairness category refers to the fact that AI systems should be committed to unbiased process and outcome, avoiding discrimination (e.g. gender, race, age or other factors) against any individual or group.

Reliability
The reliability category refers to the fact that AI systems should be reliable and stable. This implies that the systems should perform consistent and accurate under different conditions and circumstances.

Integrity
The Integrity category refers to the fact that AI systems should have accuracy, consistency and be trustworthy. This means that the systems should not use any incorrect information or manipulated data throughout its lifecycle.

Safety
The Safety category refers to the fact that AI systems should be safe and reliable. This means that the systems should not harm human or the environment.

Explainability
The explainability category refers to the fact that AI systems should be clear, transparent and comprehensible. This means that the decisions and processes of the systems should be understandable and explainable to the users and stakeholders.
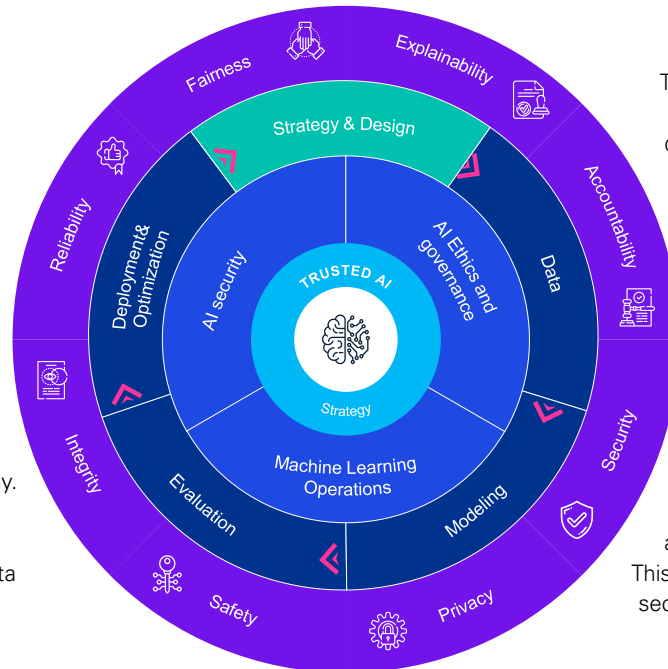
Accountability
The accountability category refers to the fact that responsibilities for the development and use of AI systems should be clearly defined. This means that the developers and users of the systems can be held accountable for their actions and decision across the development lifecycle and outcomes.

Security
The security category refers to the fact that AI systems should be protected against unauthorized access, exposure and cyber threats. This means that the systems should be secure and robust against external and internal threats.

Privacy
The privacy category refers to the fact that the users' privacy of AI systems should be safeguarded. This means that the systems should not access, collect or pass on any personal and sensitive data.

Our eight core principles guide our approach to trusted AI across the AI/ML lifecycle

Fairness
Ensure models are free from bias and equitable.

Explainability
Ensure AI can be understood, documented, and open for review.

Accountability
Ensure mechanisms are in place to drive responsibility across the lifecycle.

Security
Safeguard against unauthorized access, corruption, or attacks.

Privacy
Ensure compliance with data privacy regulations and consumer data usage.

Safety
Ensure AI does not negatively impact humans, property, and environment.

Data integrity
Ensure data quality, governance, and enrichment steps embed trust.

Reliability
Ensure AI systems perform at the desired level of precision and consistency.

What we offer

With our Trusted AI Solution Suite, we provide comprehensive support across the entire AI lifecycle, starting with model selection based on specific use cases and data analysis. Our services also extend to assessing data for ethics, bias, sustainability, and privacy to ensure responsible and transparent AI development. In addition, we emphasize risk management in accordance with regulatory standards. This enables companies to navigate the evolving AI landscape with confidence and compliance and avoid potential pitfalls.

AI Governance Health Check:
- Assess your AI compliance and models against EU AI Act standards
- Identify actionable steps for compliance
- Modular services for tailored support:
  – Assessments
  – Technical model reviews
  – Assurance & certification

Trusted AI Model & Algorithm Assessment:
- Evaluate AI models for explainability, bias, fairness, and accuracy.
- Enhanced AI practices for model improvement and best practices establishment
- Receive a status summary and actionable recommendations.

AI Governance Framework for Financial Institutions:
- Essential for mitigating risks to ensure ethical AI use, and regulatory compliance.
- Build trust, enhance decision-making, gain a competitive edge, and foster long-term sustainability.
- Accelerate implementation, comply with industry best practices, and ensure unbiased assessments.
- Foster transparency, accountability, and ethical AI use while optimizing resources and enabling organizational change.

Benefits

The choice of a service provider is not just a decision but a reflection of your organization's commitment to excellence, ethics, and trustworthiness. We understand the values that define your industry – integrity, trust, stability, customer centricity, and innovation. Our Trusted AI Solution Suite goes beyond being a solution; it's an embodiment of these values.

Risks of doing nothing

01 Enhanced Trust and Transparency
Trusted AI fosters trusts among stakeholders and clients by bringing transparency into the decisionmaking.

02 Reduced Bias and Discrimination
Trusted AI aims to minimize biases and discrimination in AI algorithms to ensure fair and equitable outcomes.

03 Improved Decision Making
Reliable AI systems lead to make better informed decision and improves operational efficiency.

04 Legal and Ethical Compliance
Trusted AI guides financial institutions to meet regulatory requirements and adhere to ethical.

05 Competitive Advantage
Organizations that prioritize Trusted AI can gain a competitive advantage by attracting clients, employees, and investors who value ethical and trusted practices.

We are the ideal partner to accomplish and materialize your vision. Our experts will help you design, deploy, and evaluate AI systems that are trusted, ethical, and fair, empowering your employees and businesses while benefiting customers and society. Our commitment to accuracy, ethical practices, and reliability in the financial sector aligns seamlessly with your goals. We look forward to the opportunity to support and enhance your mission with our expertise.

Contacts

KPMG AG
Badenerstrasse 172
P.O. Box
CH-8036 Zurich

kpmg.ch

Alexander Lacher
Partner, Insurance Regulation & Compliance

+41 58 249 33 66
alacher@kpmg.com

Bobby Zarkov
Partner, Digital Innovation
Financial Services

+41 58 249 65 91
bzarkov@kpmg.com

Thierry Kellerhals
Director, Digital Innovation
Financial Services

+41 58 249 44 84
tkellerhals@kpmg.com