AI & Partners

Amsterdam - London - Singapore

The European Union Artificial Intelligence Act

Trustworthy AI Uptake

AI Procurement Guidelines for Enterprises

August 2024

About

AI & Partners

Amsterdam - London - Singapore

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 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.

This report aims to service business and procurement leaders who want to guarantee that their organizations are at the forefront of Trustworthy AI uptake under the EU AI Act by offering practical guidance for trustworthy procurement and implementation of AI solutions across all sectors. This report adapts similar guidelines produced by the WEF.

Contents

AI & Partners

Amsterdam - London - Singapore

Executive Summary	(Slide 4)
Introduction	(Slide 5 – 8)
Acquisition Framework	(Slide 9)
Business Strategy	(Slide 10 – 11)
Commercial Strategy	(Slide 12)
Data Strategy	(Slide 13)
Ethics and Sustainability	(Slide 14 – 20)
GRC	(Slide 21 – 23)
Appendix	(Slide 24)

Offering practical guidance for trustworthy procurement and implementation of AI

AI & Partners

Amsterdam - London - Singapore

The demand for artificial intelligence ("AI") products and serivices in enterprises has grown exponentially in the past few years, driven by optimised data availability, advanced algorithms and increased processing power. While the use, development, marketing and deployment of AI products & services delivers significant value, it is required to approach it carefully and avoid its potentially harmful, and even unsafe, implications. AI & Partners has released this comprehensive guide for enterprises – across all industries – to facilitate the process of identifying, selecting and implementing AI products & services securely, ethically, and in a trustworthy manner.

This report is a practical toolkit that will help navigate the challenges of AI procurement under EU AI Act through a structured framework. It is directional, not prescriptive and both firm-/industry-agnostic, rather than problem specific.

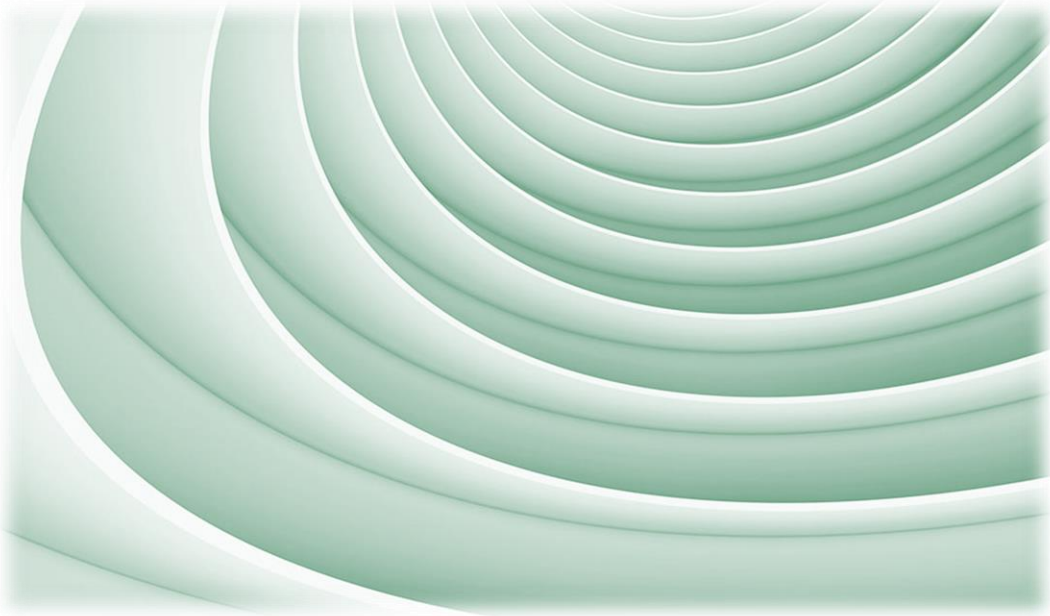
It outlines the requirement for AI products and services to closely match business goals, ethical standards and the EU AI Act's requirements, as well as the importance of stakeholder collaboration and an enterprise-wide evaluation process. It highlights five key considerations – business strategy, commercial strategy, data strategy, ethics and sustainability, and governance, risk and compliance ("GRC") – against which AI products and services (both embedded and non-embedded) can be assessed for trustworthy AI acquisition, with procurement as the coordinator driving the implementation of this framework.

Holistically, the report provides practical advice on:

- Assessing AI products and services ethical standards and EU AI Act compliance;
- Matching AI products and services with business and commercial goals; and
- Assessing their potential impact.

From automating mundane/repetitive tasks, reducing errors or risks and optimizing pricing to identifying new opportunities, making accurate forecasts and enabling investment decisions, AI products and services can assist enterprises perform a wide range of tactical and strategic activities to improve efficiency and support growth. With considered and trustworthy procurement, enterprises can utilise the power of AI to improve their productivity and gain a competitive edge.

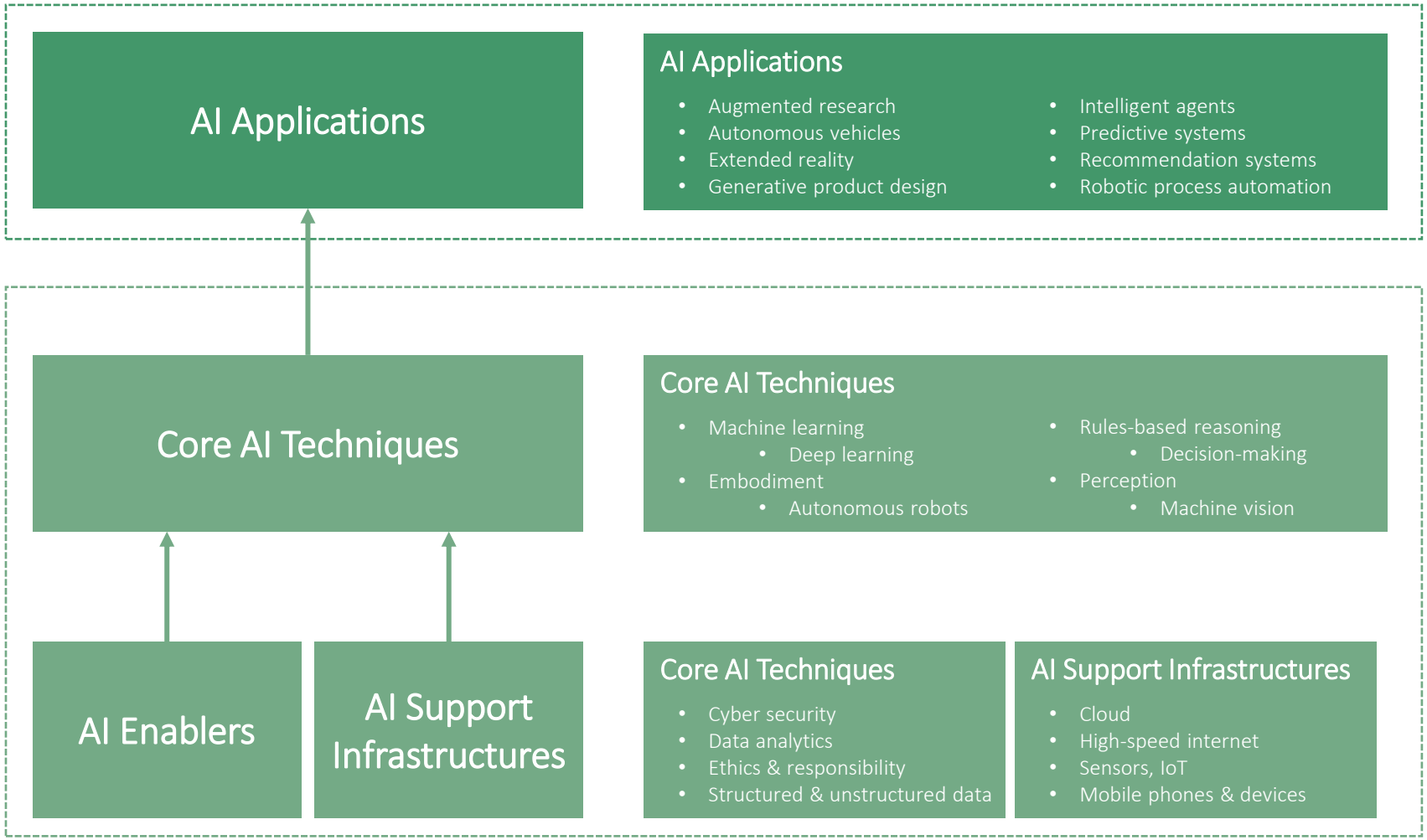
The customizable framework in this AI procurement guide targets open doors for enterprises looking to capitalise on trustworthy AI's explosive power.



Understanding AI: Tools, techniques, and enablers

AI & Partners

Amsterdam - London - Singapore



Core AI techniques and models are used to process, solve and learn utilising intended application within the organization. Usually, core AI techniques are aggregated with other technologies to drive the desired outcomes.

Core AI techniques are both mathematical and statistical models and frameworks that are deployed to process large swathes amounts of data, make decisions, learn about outcomes, contain results and use them as an extra data point to improve future decisions.

AI enablers and support infrastructure are the key foundation technologies that an AI ecosystem needs to be successful. Improvements in these technologies bolster the overall effectiveness and efficiency of AI products and services.

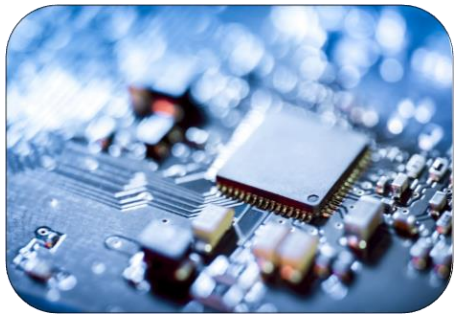
A key enabler in trustworthy AI uptake factor

AI & Partners

Amsterdam - London - Singapore

Data Collection and Cleaning

Stakeholder requirements across different business functions



System implementation and monitoring

Risk aspects

Project strategy, feasibility, and scalability

Integration with current IT infrastructure

Requirement for cross-functional collaboration

AI & Partners
Amsterdam - London - Singapore

Team	Business/End-User	IT and Cybersecurity	AI Centre of Excellence/Data Team	Data Management	Procurement
Role	• Assessment and ranking of use cases • Process definitions • Workflow design • Defining business needs • Defining expected results • Exception management • Change management	• Architecture and integration design • Infrastructure assessment and design • Security assessment and design • Application management and support • Platform deployment and scalability assessment • Software and licence management	• AI and cognitive capability assessment • Tool adaptability and configuration of needs • Process modelling • Service support capabilities • Testing and debugging • GRC • User interface (“UI”)/ user experience (“UX”) management • Change management • AI ethics and sustainability	• Data needs and consumption requirements • Data governance framework (Art. 10) • Data integrity requirements	• Align with the business team on requirements • Collaborate with IT and cybersecurity, AI centre of excellence and data management teams to define the scope of the AI project • Assess the market to identify and shortlist suitable supplier(s) • Coordinate with all teams to enable the implementation of the AI product or service • Support the business team to monitor and measure relevant key performance indicators (“KPIs”)

Usage of ethical guidelines across standard sourcing process

AI & Partners

Amsterdam - London - Singapore

Project Scoping and due-diligence

- Prioritize business use case for AI deployment
- Define business outcome criteria and objectives
- Document as-is processes and current gaps that the AI product or service will help solve
- Establish business governance prerequisites
- Establish risk management requirements

Market intelligence

- Undertake high-level market analysis on available AI product or services and AI product or service providers
- Understand data complexity plus AI product or service specifications and identify various supplier types that could help achieve the business goal
- Qualify suppliers that would best fit requirements

Sourcing exception

- Draft scope requirements for the identified AI application
- Involve internal functions to identify must-have vs good-to-have functionalities and features
- Identify what would separate one AI product or service from the other in each of the RFx questionnaires
- Refer sample questions in the guidebook to ensure all focus areas and details are covered
- Execute RFx for responses from suppliers
- Prepare supplier evaluation matrix based on criteria across areas of consideration

Evaluation and negotiations

- Perform objective evaluations through supplier scorecard-based questionnaire review
- Do total cost of ownership ("TCO") analysis
- Conduct supplier risk assessment

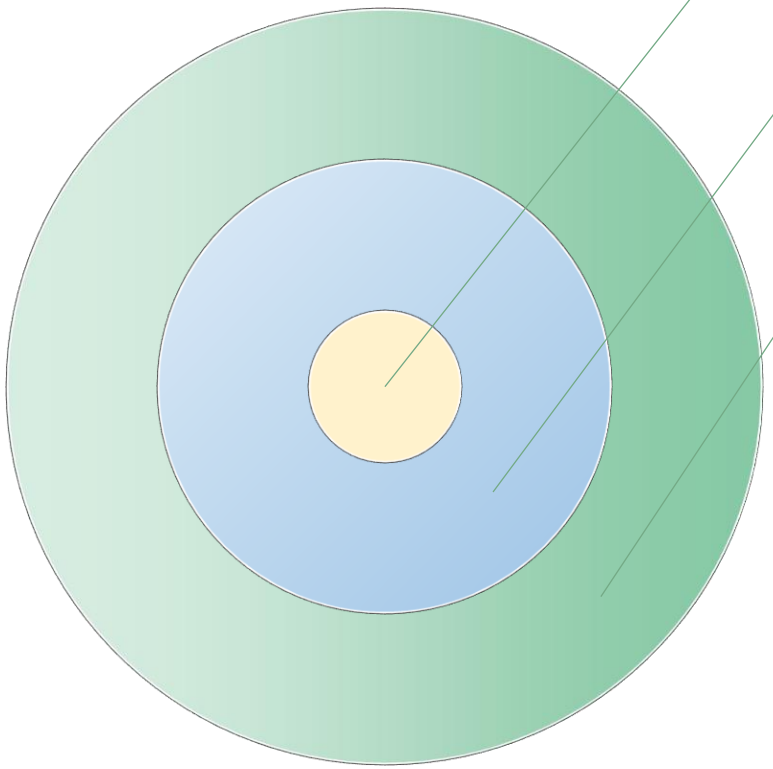
Contracting and implementation

- Select suppliers
- Contract negotiations: ensure all requirements are met through various terms and conditions
- Consider outcome-based commercial model
- Draw up a detailed project implementation plan
- Create a plan for continuous improvement
- Plan for continuous retraining of algorithms
- Develop a framework for business governance

Key considerations for trustworthy AI acquisition

AI & Partners

Amsterdam - London - Singapore



Ethics

Strategy

- Business
- Commercial
- Data

Infrastructure

- GRC
- Sustainability

Questions to ask to achieve business goals

AI & Partners
Amsterdam - London - Singapore

Specifications	Key Considerations
Is there an inherent feedback loop in the system to adjust to changes in the external environment? An AI product or service can become a self-learning system, provided a sufficient and feedback loop is embedded into it.	• Does the AI product or service have an automatic feedback/retraining loop, or is there a human in the loop? • How does the AI product or service measure business outcomes and user satisfaction? How do those insights get delivered and/or feed into any product changes or updates?
How effectively can the AI product or service be updated to accommodate change in requirements? AI is a highly dynamic system, so the supplier should integrate with the buyer to redesign the AI product or service in case business requirements change. This would be applicable especially if data complexity is high. **A steady state is when the system is acting in the intended manner, and all high-severity incidents are resolved*	• How feasible is it to adapt or customize the AI product or service during the initial and steady state*? Does the AI product or service update its behaviour based on newly ingested data? • How will the supplier help in course correction if the AI model or prediction delivers unexpected results? • How is the service tested and monitored for model or performance drift over time? • Does the supplier provide KPIs for monitoring any performance drifts that may prompt retraining of the model (if there are unexpected changes)? • Does the supplier provide documentation detailing how the AI system can be reconfigured or adapted if the results are not delivering the goals?
Is the supplier a thought leader in this domain? All new technologies require partnerships with suppliers that are thought leaders in this domain as their vision steers the AI product or service to meet future requirements.	• How much does the supplier invest in R&D in the domain of their AI product or service and company requirement? • How does the supplier product development stay in line with market trends? • Does the supplier publish reports offering industry best practices and actionable insights in terms of optimizing AI outcomes? • Does the supplier organize in-person/virtual networking events for its clients to increase awareness of existing AI product or services? • How does the supplier AI product or service differentiate itself from its competitors? • Is the supplier investing in upskilling and training its talent pool to stay relevant with state-of-the-art technology?

Questions to ask to achieve business goals

AI & Partners
Amsterdam - London - Singapore

Specifications	Key Considerations
How will the AI product or service deliver on expected business outcomes? Understanding how the AI product or service will help enterprises meet their business goals.	• Is the supplier able to understand your business goals and explain how their capabilities will help achieve them? • Are the AI product or service's capabilities outlined clearly, and are there demonstrated use cases they can reference? • How much custom development will be required for the AI product or service to meet your requirements? • What level of guarantee does the AI product or service provider give for the process and business outcomes? • Could the AI product or service scale up to meet increased demand/productivity? • If your business slows down, is the AI product or service flexible enough to adjust to the changed needs?
Is there transparency about what the AI product or service can and cannot achieve? Often, businesses are not fully aware of the internal working mechanism of the AI model being used. It is the responsibility of the AI product or service provider to educate potential buyers and be transparent about the capabilities and limitations of their offering.	• Does the supplier explain the techniques applied in the AI system, including the use of algorithms and associated software libraries for the algorithms? • Is the supplier able to articulate the workings of the AI product or service in an easy-to-understand manner? • Does the supplier dedicate the required resource(s) to train/educate your team about the AI product or service? • Does the AI product or service offer explainable results and transparency in the decision-making process? • Does the supplier recognize and describe any limitations of the AI system for the problem you want addressed?
Can a non-AI product or service deliver the same outcomes? Like any other technology, AI is not a magic bullet for all problems. It is essential that businesses explore alternative technology/ AI product or services before deciding on an AI product or service.	• Can you justify why the use of AI/ML is the optimal approach to meeting the specified business goals? • Can any other cost-effective, non-AI technology or process be used to achieve the expected business outcomes?

Questions to ask to achieve commercial goals

AI & Partners

Amsterdam - London - Singapore

Specifications	Key Considerations
What is the expected value of the AI product or service to be delivered? Commitment to business case requirements will be a key ingredient in a supplier partnership.	• In what ways is the supplier committed to achieving the business case objectives? • Are there any dependencies/assumptions made by the supplier on achieving these objectives? • Are there ways in which the value delivered is optimized during the steady state*? • Will the AI product or service require any co-development/co-innovation? How will you benefit from the potential gains of the supplier in case of co-innovation?
Do you understand all the costs involved in the purchase and maintenance of the AI system? Determining the viability of an AI product or service for a longer or shorter term can be done by calculating the TCO, which includes all direct and indirect costs throughout the lifetime of the AI product or service – from acquiring, building, running to retiring. **Much of the cost of AI projects is internal; the process of getting data ready to enable an AI solution or product is time consuming and expensive.*	• Can the supplier provide a breakdown of one-time costs based on project milestones/ measurable deliverables? • Has the supplier provided a resource-wise effort estimation required for each of these milestones? • Are the recurring cost components clearly defined? Is there clarity on the frequency of the recurring costs? • What would be the cost of changes required in the existing systems/infrastructure?* • Has the supplier provided the hourly/daily rates of consulting/technical resources to be involved in the project? • What are the relevant skills/years of experience in the AI domain of the resources the supplier plans to deploy? • Does the supplier offer different pricing options to accommodate the gradual expansion of the AI system deployment in the client environment? • What are the discounts/rebates offered by the supplier?
How can organizations mitigate the investment risks of the AI product or service? There are multiple unknowns when organizations invest in new technologies and the team should evaluate possibilities to mitigate investment risks.	• Is the supplier able to delink the discovery and execution phases to restrict upfront investment while the business case is being tested? • Can the AI product or service/implementation be in phases and gradually widened to avoid large upfront costs? For e.g. implementation of a low-cost model or proof of concept before full-scale implementation? • Is the supplier able to commit to an outcome-based pricing model? What would be the KPIs that can be linked to outcomes? • Is the supplier financially solvent? • Is the supplier currently going through or has planned for an acquisition/merger? • Who will own the AI model in the event of insolvency or ownership transfer? (e.g. code escrow, data escrow, model escrow)

Questions to ask to achieve data goals

AI & Partners

Amsterdam - London - Singapore

Specifications	Key Considerations
Do you have a clear and defined as-is and to-be data strategy? To maximize business value from implementation of the AI product or service, organizations need to assess, identify gaps and adopt a long-term data strategy.	• What type of data platform does your organization currently have in place? How will the new AI implementation impact your existing data platform? • What is the enterprise data structure strategy if you are planning to migrate to the cloud and use cloud-based AI capabilities? What data will go on the cloud? Are you looking at single or multi cloud providers? Do you want to prioritize modularity or flexibility? • Is there an existing data culture across the organization? • Is senior leadership involved in establishing and communicating data strategies? • How are different parts of the organization incentivized to share and reuse data? Are employees being trained in data quality management? • Does the data-based feedback need to be looped back into business decisions and how? • Is there a specific team for data collection, validation, storage, governance, security and accountability structures across the entire data supply chain? • Who is responsible for ethical data use in the enterprise?
What are the different sources of data to be considered? For organization-specific contextual learning process, internally available data is ideal. Otherwise, external sources of data (synthetic data) can be considered.	• What are the different kinds of data expected to be ingested by the AI product or service? • Can the AI product or service meet existing and new enterprise objectives based on internally available data? • Do you know where and how to collect the data internally? Is the available internal data ready to be consumed is it accurate, complete, consistent and up to date? • Will external sources of data be needed? What is the cost of acquiring/generating synthetic data? • Can the external data provider ensure complete, relevant, unbiased and timely data? Who is responsible for ensuring the quality, usability and reliability of third-party data? • Should there be different contractual provisions for the exchange of different categories of internal data and related data models outside your organization? For example, you may want to limit/ encrypt sharing of data/data models built on personal identifiable information (PII)-related data. • Is the collection of additional data necessary in the future for optimal AI performance?
Does the supplier have satisfactory data management practices? The data for the AI model should meet data quality standards defined by the governance team. Furthermore, data ownership and management, right from storage to extraction, need to be efficient, secure and adhere to regulatory requirements.	• Does the supplier have adequate data quality assurance processes and frameworks around storage, management and transfer of data? • Who will have the ownership and be accountable for the data and derivative models? –Does the supplier have the standard data privacy/security frameworks for its industry? • How will the supplier follow data privacy practices for sensitive data that falls outside of the General Data Protection Regulation (GDPR)? • Where is the data collected for the AI model stored? What are the security measures to prevent a data breach? • Who has access to the stored data? Elaborate on the liability in case of data breach.

Seven core principles for ethical AI

AI & Partners

Amsterdam - London - Singapore

Design

1. Empowering Human Beings

- **User-Centric Design**: AI models should be designed with user empowerment in mind, ensuring that users can make informed decisions based on the AI's outputs. This involves creating intuitive interfaces and providing clear, understandable information about the AI's functionality and limitations.
- **Respect for Fundamental Rights**: The design process must incorporate safeguards to protect users' fundamental rights, such as privacy and non-discrimination. This includes implementing data protection measures and ensuring that the AI does not perpetuate biases.

2. Oversight Mechanisms

- **Human-in-the-Loop (HITL)**: Design AI systems to require human intervention at critical decision points. This ensures that humans can review and validate AI outputs before any significant action is taken.
- **Human-on-the-Loop (HOTL)**: Implement monitoring systems that allow humans to oversee AI operations continuously. This enables real-time intervention if the AI system behaves unexpectedly or produces questionable results.

Use

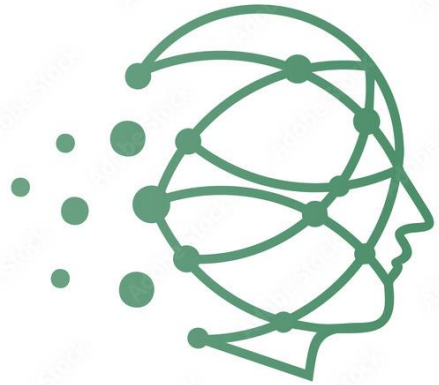
1. Empowering Human Beings

- **Informed Decision-Making**: Users should be provided with comprehensive information about the AI system's capabilities, limitations, and the context in which it operates. This helps users understand the AI's recommendations and make informed decisions 3.
- **Training and Competence**: Users assigned to oversee AI systems should receive adequate training to understand the system's functionality and potential risks. This ensures they are competent to intervene when necessary.

2. Oversight Mechanisms

- **Monitoring and Feedback**: Users should continuously monitor the AI system's performance and provide feedback to improve its accuracy and reliability. This includes reporting any anomalies or unexpected behaviours to the provider.
- **Decision-Making Authority**: Users should have the authority to override AI decisions when they conflict with human judgment or ethical considerations. This ensures that human values and ethical standards are upheld.
- **Stop Mechanisms**: AI systems should include mechanisms, such as a "stop" button, that allow users to halt the system's operation in case of emergencies or malfunctions.

AI systems should empower human beings, allowing them to make informed decisions and fostering their fundamental rights. Proper oversight mechanisms need to be ensured, which can be achieved through human-in-the-loop, human-on-the-loop, and human-in-command approaches.



Seven core principles for ethical AI

AI & Partners

Amsterdam - London - Singapore

Design

1. Resilience and Security

- **Robust Design**: AI models should be designed to withstand errors, faults, and inconsistencies. This includes implementing technical solutions such as redundancy and fail-safe mechanisms to ensure the system can safely interrupt its operation in the presence of anomalies.
- **Cybersecurity Measures**: Incorporate robust cybersecurity measures to protect against unauthorized access and manipulation. This includes measures to prevent, detect, respond to, and control attacks such as data poisoning, model poisoning, and adversarial examples.

2. Safety and Accuracy

- **Accuracy and Reliability**: AI models must be designed to achieve high levels of accuracy and reliability. This involves rigorous testing and validation to ensure the system performs consistently and meets the required performance metrics.
- **Fallback Plans**: Develop and integrate fallback plans to ensure the system can handle unexpected situations without causing harm. This includes mechanisms to safely interrupt operations if the system operates outside predetermined boundaries.

Use

1. Monitoring and Maintenance

- **Continuous Monitoring**: Regularly monitor the AI system's performance to ensure it remains accurate, reliable, and secure. This includes tracking system outputs and identifying any deviations from expected behaviour.
- **Incident Reporting**: Establish procedures for reporting and addressing incidents involving the AI system. This ensures that any issues are promptly identified and resolved to maintain system safety.

2. User Training and Awareness

- **User Training**: Ensure that users are adequately trained to understand the AI system's capabilities, limitations, and safety measures. This includes training on how to respond to system anomalies and implement fallback plans.
- **Transparency**: Provide clear and comprehensive information to users about the AI system's performance metrics, including accuracy, robustness, and cybersecurity levels. This helps users make informed decisions and use the system appropriately.

AI systems need to be resilient and secure. They need to be safe, ensuring a fallback plan in case something goes wrong, as well as being accurate, reliable, and reproducible. This is essential to minimize and prevent unintentional harm.

Seven core principles for ethical AI

AI & Partners
Amsterdam - London - Singapore

Design

1. Ensuring Privacy and Data Protection

- **Data Minimization and Protection by Design**: AI models should be designed to adhere to the principles of data minimization and data protection by design and by default. This involves collecting only the data necessary for the intended purpose and implementing privacy-preserving techniques such as anonymization and encryption.
- **Compliance with Data Protection Laws**: The design process must ensure compliance with Union data protection laws, such as the General Data Protection Regulation (GDPR). This includes obtaining necessary consents, ensuring data subject rights, and implementing appropriate technical and organizational measures to protect personal data.

2. Data Governance Mechanisms

- **Quality and Integrity of Data**: Implement robust data governance practices to ensure the quality and integrity of the data used for training, validation, and testing AI models. This includes verifying the accuracy, completeness, and representativeness of the data sets.
- **Legitimized Access to Data**: Ensure that data access is legitimized and controlled. This involves setting up access controls and audit trails to monitor who accesses the data and for what purpose.

Use

1. Maintaining Privacy and Data Protection

- **Ongoing Compliance**: Continuously monitor and ensure that the AI system's use complies with data protection regulations. This includes regular audits and assessments to identify and mitigate any privacy risks.
- **User Awareness and Consent**: Inform users about how their data is being used by the AI system and obtain their consent where necessary. This includes providing clear and transparent information about data processing activities.

2. Data Governance in Practice

- **Data Management Practices**: Implement data management practices that ensure the ongoing quality and integrity of the data. This includes regular data cleaning, updating, and validation to maintain data accuracy and relevance.
- **Bias Detection and Mitigation**: Continuously monitor the AI system for potential biases in the data and implement measures to detect, prevent, and mitigate these biases. This helps ensure that the AI system operates fairly and does not perpetuate discrimination.

Besides ensuring full respect for privacy and data protection, adequate data governance mechanisms must also be ensured, taking into account the quality and integrity of the data, and ensuring legitimized access to data.

Seven core principles for ethical AI

AI & Partners
Amsterdam - London - Singapore

Design

1. Transparent Data and System Design

- **Traceability Mechanisms**: Implement traceability mechanisms to ensure that every decision made by the AI system can be traced back to its source. This includes logging data inputs, processing steps, and outputs.
- **Documentation**: Maintain comprehensive technical documentation that includes a general description of the AI model, its intended tasks, acceptable use policies, and the architecture and number of parameters. This documentation should be kept up to date and made available to relevant stakeholders, including downstream providers and regulatory authorities.

2. Explainability and User Awareness

- **Explainable AI**: Design AI systems to provide explanations for their decisions in a manner that is understandable to the intended users. This involves developing models that can generate human-readable explanations for their outputs.
- **User Information**: Ensure that users are informed that they are interacting with an AI system. This includes providing clear information about the AI system's capabilities, limitations, and the context in which it operates

Use

1. Meeting Transparency Obligations

- **Information Disclosure**: Providers and deployers of AI systems must ensure that the AI system is accompanied by instructions for use that include concise, complete, correct, and clear information relevant to the deployers. This includes details about the system's characteristics, capabilities, and limitations.
- **Marking AI Outputs**: For AI systems generating synthetic content, such as audio, image, video, or text, ensure that the outputs are marked in a machine-readable format and detectable as artificially generated or manipulated.

2. User Interaction and Feedback

- **User Awareness**: Inform users at the time of interaction that they are engaging with an AI system, unless it is obvious to a reasonably well-informed person. This helps users understand the nature of their interaction and manage their expectations accordingly.
- **Feedback Mechanisms**: Implement feedback mechanisms that allow users to report issues or provide feedback on the AI system's performance. This helps in continuously improving the system and maintaining transparency.

The data, system, and AI business models should be transparent. Traceability mechanisms can help achieve this. Moreover, AI systems and their decisions should be explained in a manner adapted to the stakeholder concerned. Humans need to be aware that they are interacting with an AI system and must be informed of the system's capabilities and limitations.

Seven core principles for ethical AI

AI & Partners

Amsterdam - London - Singapore

Design

1. Avoiding Unfair Bias

- **Bias Mitigation**: Implement strategies to identify and mitigate biases during the design phase. This includes using diverse and representative datasets to train AI models, ensuring that the data reflects various demographics and does not perpetuate historical biases.
- **Algorithmic Fairness**: Develop and test algorithms to ensure they do not produce discriminatory outcomes. This involves using fairness metrics and conducting regular audits to detect and correct biases.

2. Inclusive Design

- **Accessibility**: Design AI systems to be accessible to all users, including those with disabilities. This involves adhering to accessibility standards and guidelines, such as the Web Content Accessibility Guidelines (WCAG) and relevant EU directives.
- **Stakeholder Involvement**: Engage a diverse group of stakeholders, including representatives from vulnerable and marginalized communities, throughout the design process. This ensures that the AI system addresses the needs and concerns of all potential users.

Use

1. Ensuring Fairness in Deployment

- **Continuous Monitoring**: Regularly monitor the AI system's performance to ensure it continues to operate fairly and does not produce biased outcomes. This includes tracking the system's impact on different demographic groups and making necessary adjustments.
- **User Feedback**: Establish mechanisms for users to provide feedback on the AI system's performance. This helps identify any issues related to fairness and allows for timely interventions.

2. Promoting Diversity and Inclusion

- **Training and Awareness**: Provide training for users and operators of AI systems on the importance of diversity, non-discrimination, and fairness. This includes educating them on how to use the system responsibly and recognize potential biases.
- **Inclusive Practices**: Encourage the adoption of inclusive practices in the use of AI systems. This involves ensuring that the system is used in a way that promotes equal access and opportunities for all users, regardless of their background or abilities.

Unfair bias must be avoided, as it could have multiple negative implications, from the marginalization of vulnerable groups to the exacerbation of prejudice and discrimination. Fostering diversity, AI systems should be accessible to all, regardless of any disability, and involve relevant stakeholders throughout their entire lifecycle.

Seven core principles for ethical AI

AI & Partners

Amsterdam - London - Singapore

Design

1. Sustainability and Environmental Friendliness

- **Sustainable Design**: AI models should be designed with sustainability in mind. This includes optimizing algorithms to reduce energy consumption and using energy-efficient hardware. The design process should consider the environmental impact of the entire lifecycle of the AI system, from development to deployment.
- **Environmental Impact Assessment**: Conduct thorough environmental impact assessments during the design phase to identify and mitigate potential negative effects on the environment. This includes evaluating the carbon footprint and resource usage of AI systems.

2. Societal Impact Consideration

- **Inclusive Design**: Ensure that AI systems are designed to be inclusive and accessible to all, including vulnerable and marginalized groups. This involves engaging with a diverse group of stakeholders, including representatives from various communities, to understand their needs and concerns.

Use

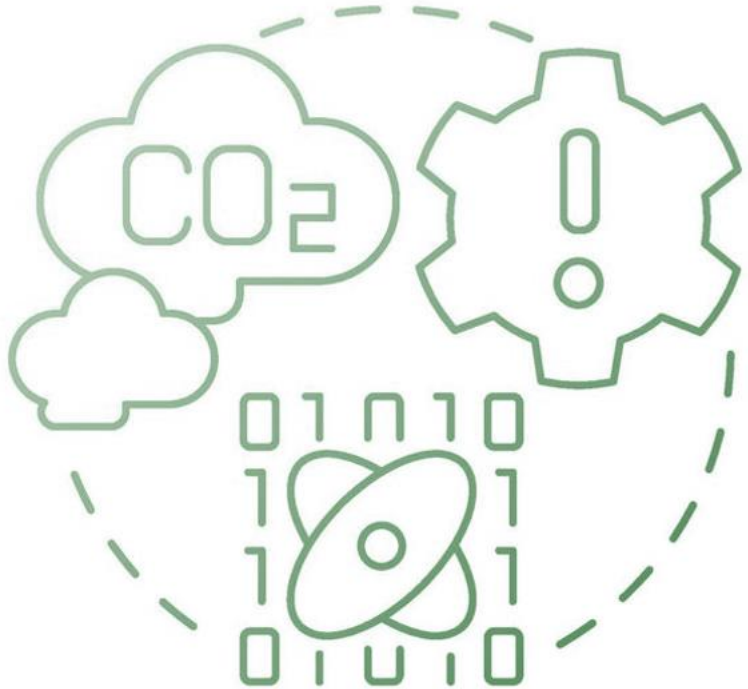
1. Monitoring and Assessing Impact

- **Continuous Monitoring**: Regularly monitor the societal and environmental impact of AI systems after deployment. This includes tracking the system's performance and its effects on different communities and the environment.
- **Impact Reporting**: Implement mechanisms for reporting the societal and environmental impact of AI systems. This includes providing transparent reports to stakeholders and regulatory bodies, detailing the AI system's performance and its impact on society and the environment.

2. Promoting Positive Outcomes

- **AI for Social Good**: Encourage the use of AI systems for socially and environmentally beneficial outcomes. This includes developing AI solutions that address societal challenges such as healthcare, education, and environmental conservation.
- **Resource Allocation**: Allocate resources, including public and Union funding, to support the development and deployment of AI systems that promote societal and environmental well-being. This includes funding projects that aim to increase accessibility for persons with disabilities, tackle socio-economic inequalities, and meet environmental targets.

AI systems should benefit all human beings, including future generations. It must hence be ensured that they are sustainable and environmentally friendly. Moreover, they should take into account the environment, including other living beings, and their social and societal impact should be carefully considered.



Seven core principles for ethical AI

AI & Partners

Amsterdam - London - Singapore

Design

1. Responsibility and Accountability Mechanisms

- **Clear Accountability Framework**: Establish a clear accountability framework that outlines the responsibilities of all stakeholders involved in the design, development, and deployment of AI systems. This includes defining roles and responsibilities for management and staff.
- **Compliance Strategy**: Develop a comprehensive strategy for regulatory compliance, including adherence to conformity assessment procedures and management of modifications to the AI system.

2. Auditability

- **Documentation and Record-Keeping**: Maintain detailed documentation of the AI system's design, development, and deployment processes. This includes keeping records of data acquisition, data processing, algorithm design, and system testing.
- **Regular Audits**: Implement regular internal and external audits to assess the AI system's compliance with regulatory requirements and ethical guidelines. This helps ensure that the system operates as intended and adheres to established standards.

Use

1. Continuous Monitoring and Reporting

- **Post-Market Monitoring**: Set up a post-market monitoring system to continuously evaluate the AI system's performance and impact. This includes collecting and analysing data on the system's operation and identifying any issues that may arise.
- **Incident Reporting**: Implement procedures for reporting serious incidents involving the AI system. This ensures that any issues are promptly identified and addressed, maintaining the system's safety and reliability.

2. Accessible Redress Mechanisms

- **User Feedback and Complaints**: Provide mechanisms for users to submit feedback and complaints about the AI system. This helps identify and resolve issues related to the system's performance and impact on users.
- **Redress Procedures**: Establish clear and accessible procedures for redress, ensuring that users can seek remedies for any harm caused by the AI system. This includes providing information on how to file complaints and the steps involved in the redress process.

Mechanisms should be put in place to ensure responsibility and accountability for AI systems and their outcomes. Auditability, which enables the assessment of algorithms, data, and design processes, plays a key role therein, especially in critical applications. Moreover, adequate and accessible redress should be ensured.

Questions to ask to manage GRC

AI & Partners
Amsterdam - London - Singapore

Specifications	Key Considerations
What is the target demographic for data collection for the model? (Article 10 of EU AI Act) As AI modelling depends on data, it is imperative to consider the impact of data use on the target, especially in terms of privacy, consent and regulations related to individuals.	• Is the data being collected from vulnerable demographic groups of the target countries? • If personal data is being used, is it being collected in compliance with the data protection and privacy laws of the country (GDPR, Health Insurance Portability and Accountability Act ("HIPAA"), regional laws etc.)? • Does the supplier have the informed consent of the individuals whose data has been collected, i.e. have the individuals been provided all the necessary information? • Can individuals withdraw their consent to the data collected? If so, will the collected data be withdrawn from the AI model? • What are the relevant PII categories for the data collection process? • Is the content moderated? (e.g. for sexuality/violence)
How has the supplier accounted for managing cybersecurity risks? (Article 15 of EU AI Act) An important concern to be addressed early on is cybersecurity failures, given its potential to have a serious impact on the output of the AI product or service.	• What are the proactive measures the supplier has taken to detect, and tackle cyberattacks • How does the supplier minimize the effect of an attack? • Does the supplier actively perform vulnerability management to address common and frequent threats?
Has the supplier reviewed the potential geopolitical risks that arise from operating in different physical locations? For a company looking to embed a disruptive technology like AI into its systems, the impact of geopolitics must be considered.	• Has the supplier accounted for the geopolitical risks associated with collecting data from certain disputed territories? • Have the risks of storing or processing data in unstable regions been considered? • Will AI pose any risk if used in such territories? e.g. can it heighten instability in regional politics, affect peace and security? • Are data collectors at any physical risk during the process of data collection?

Questions to ask to manage GRC

AI & Partners

Amsterdam - London - Singapore

Specifications	Key Considerations
Have the risks related to the project been defined clearly? Due to the uncertainty of AI/ML work, an experienced AI modelling supplier is expected to identify possible risks and ways to mitigate them.	• Is the scope for the AI defined clearly in terms of deliverables/outcomes to be achieved? • How does the supplier manage unsupported content types? • How does the supplier define hard performance metrics with AI? • To what extent is the AI product or service reproducible? • Will the AI model be covered by intellectual property policy? Who has legal ownership of source data, models and resell rights?
Is the supplier compliant with the EU AI Act when building an AI model? AI offers great value to businesses, but it comes with a strategic risk for all stakeholders. Governments and institutions are actively taking measures to prevent the misuse of the technology and to build trust in AI tools.	• Has the supplier proactively prepared to ensure compliance with the EU AI Act? Does the supplier provide an explainability statement outlining the critical dimensions of the AI product or service?
How does the supplier prepare for audits and compliance requirements? A risk management process that captures the policies, processes, procedures and practices across the organization involved in the development, testing, deployment, use and auditing of AI systems should be in place. It must be implemented effectively as well as be transparent.	• Does the supplier conduct mandatory conformity assessments? At what frequency? • Has the supplier clearly defined the systems in place for internal audits? What are the audit artifacts that it can share with you? • How does the supplier ensure that compliance is met on its side as well as the buyer's side after implementation of the AI model? • If access to legal support is limited (in case of a smaller buyer), how can the supplier assist you in ensuring compliance? • Has the AI model been assessed for its performance with algorithm assessment tools, model cards etc., to prevent biases and undesirable outputs?

Questions to ask to manage GRC

AI & Partners
Amsterdam - London - Singapore

Specifications	Key Considerations
Is the supplier implementing international standards and certifications in the model? Standards and certifications can form a part of the initial guiding principles of governance to help AI developers and its users in their journey to build a responsible AI model.	• Does the supplier follow any AI governance standards set by international standard bodies (such as the International Organization for Standardization (ISO) and the Institute of Electrical and Electronics Engineers Standards Association (IEEE) and others) to ensure that best practices are being followed? • Will the AI system be accredited by any recognized institute that provides a calibrated conformity assessment of the model?
What are the organizational practices recommended by the supplier? Organizations should establish practices that can characterize the AI model specifications paying special attention to attributes such as accuracy, bias, consistency, transparency, interpretability and fairness.	• Is the risk-based approach developed by the supplier based on the AI model and the industry in which it is being implemented? • Has the supplier conducted an AI impact assessment of the buyer organization at an early stage of the procurement process?
Do the contractual agreements include all compliance-related factors? A good agreement will not only involve basic contractual terms and project management details but also have protective measures against non-compliance.	• Can the supplier develop capacity in the form of new contract requirements? • Is there a supplier compliance statement that organizations can include in their master service agreements? • Are there contractual agreements surrounding restricted use or prohibited forms of use? • Has the buyer developed well-defined KPIs and compliance metrics to track performance during the AI life cycle? • Does the supplier offer support beyond contractual agreements to assist in governance, maintenance and change management?

Questions to ask for first steps

AI & Partners
Amsterdam - London - Singapore

Specifications	Key Considerations
1. Business strategy	• What kind of assurance or warranty does the supplier offer regarding the process and business outcomes? • How is the proposed AI/ML product or service an optimal approach to meet requirements? Can the supplier provide case studies to support their response? • Does the AI product or service measure business results and user satisfaction? If so, how are these insights used for product modifications or upgrades? • Does the supplier provide key performance indicators (KPIs) for monitoring performance drifts that may prompt retraining of the model?
2. Data strategy	• What data quality assurance processes and frameworks (storage, management, transfer etc.) does the supplier follow? • In the case of external data ingestion, how does the supplier ensure complete, relevant, unbiased and timely data? Who is responsible for ensuring the quality, usability and reliability of third-party data?
3. Ethics and sustainability	• What methods are used to train the AI model? Do the training methods uphold the principles of ethics (fairness, interpretability, privacy, security etc.)? • Can the supplier identify possible sources of bias? What are the checks in place within the model to prevent biases from creeping in? • Does your AI product or service address new and emerging ethical risks such as misinformation, over-reliance and loss of skills?
4. GRC	• Does the supplier comply with data-related regulations (e.g. General Data Protection Regulation (GDPR), California Consumer Privacy Act (CCPA), Health Insurance Portability and Accountability Act (HIPAA), Children's Online Privacy Protection Act (COPPA) etc.)? • Does the supplier follow any AI governance standards and best practices set by international standards organizations (such as the International Organization for Standardization (ISO) and the Institute of Electrical and Electronics Engineers Standards Association (IEEE))?

Contact us today – we're happy to help!

AI & Partners

Amsterdam - London - Singapore

AI & Partners

Amsterdam - London - Singapore

Email

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

Disclaimer

AI & Partners

Amsterdam - London - Singapore

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.