



AI
AI & Partners

Amsterdam - London - Singapore

EU AI Act

Code of Conduct

Deployers of AI Systems

Version 1.0

Last Updated: 17 August 2024

Disclaimer

Codes of Conduct (CoCs) are publications published by AI & Partners B.V. (AI & P). It should be noted that the sole purpose of these publications is to provide guidelines and standards of behaviour for organizations under the European Union Artificial Intelligence Act (EU AI Act). CoCs serve as a framework for promoting ethical conduct, defining acceptable behaviours, and guiding decision-making processes. CoCs are designed to ensure consistency, integrity, and accountability in the actions and interactions of those subject to them.



Contents

- Introduction** 2
- Purpose and Scope** 2
- Definitions** 4
- General Principles** 5
- Compliance Framework** 6
- Transparency and Accountability** 7
- Data Governance and Privacy** 9
- Human Oversight and Responsibility** 10
- Robustness and Safety** 11
- Non-discrimination and Fairness** 12
- Environmental Sustainability** 14
- AI Literacy and Education** 15
- Inclusive and Diverse Design** 16
- Assessing and Preventing Negative Impacts** 17
- Stakeholder Engagement** 19
- Governance Mechanisms** 20
- Reporting and Enforcement** 21
- Conclusion** 23





Introduction

In an era defined by unprecedented technological advancement, the integration of artificial intelligence (AI) systems has become increasingly ubiquitous across various sectors and industries. The transformative potential of AI holds promise for enhancing efficiency, productivity, and innovation, yet it also brings forth ethical considerations and challenges that must be addressed to ensure responsible deployment and use.

The Code of Conduct for Deployers of AI Systems serves as a foundational framework for guiding the ethical and responsible deployment of AI technologies. As the cornerstone of ethical AI practices, this Code outlines the principles, requirements, and guidelines that deployers of AI systems must adhere to in order to foster trust, transparency, and accountability in the development and utilization of AI technologies.

At its core, this Code is designed to support the voluntary application of specific requirements outlined in the EU AI Act. By providing clear guidance and standards, it empowers deployers to navigate the complex landscape of AI deployment while upholding ethical principles and legal obligations. Through voluntary compliance with this Code, deployers demonstrate their commitment to ethical AI practices and contribute to the advancement of a responsible AI ecosystem.

The scope of this Code is broad, encompassing all deployers of AI systems, regardless of the level of risk associated with their AI applications. Whether deploying AI systems in high-stakes environments or in more routine operations, all deployers are subject to the principles and requirements outlined herein. This inclusivity ensures that ethical considerations are embedded throughout the AI lifecycle, from development and deployment to monitoring and evaluation.

Key aspects covered in this Code include transparency, data governance, human oversight, non-discrimination, environmental sustainability, and stakeholder engagement. Each of these areas plays a critical role in ensuring that AI systems are developed and deployed in a manner that respects fundamental rights, promotes fairness, and mitigates potential risks and harms.

The principles outlined in this Code are underpinned by a commitment to accountability and continuous improvement. Deployers are expected to establish robust compliance frameworks, conduct impact assessments, and implement mechanisms for ongoing monitoring and evaluation. By embracing transparency and accountability, deployers can identify and address potential biases, errors, and negative impacts associated with AI systems.

Furthermore, this Code emphasizes the importance of stakeholder engagement and collaboration throughout the AI deployment lifecycle. By involving civil society organizations, academia, industry representatives, and other stakeholders, deployers can ensure that diverse perspectives are considered and that AI systems are developed and deployed in a manner that reflects societal values and concerns.

Purpose and Scope

The purpose of this Code of Conduct for Deployers of AI Systems is to provide clear guidance and standards for deployers to ensure the responsible and ethical deployment of AI technologies. It serves as a foundational framework to support the voluntary application of specific requirements outlined in the EU AI Act, thereby fostering trust, transparency, and accountability in the development and use of AI systems.





This Code applies to all deployers of AI systems, regardless of the level of risk associated with their AI applications. Whether deploying AI systems in critical sectors such as healthcare and finance or in more routine operations, all deployers are subject to the principles and requirements outlined herein. By applying these standards uniformly across diverse applications and contexts, this Code aims to promote consistency and coherence in the ethical deployment of AI technologies.

The scope of this Code is comprehensive, covering various aspects of AI deployment to ensure that ethical considerations are embedded throughout the AI lifecycle. Key areas addressed in this Code include:

- **Transparency:** Deployers are expected to provide clear information on the development, capabilities, limitations, and potential impacts of AI systems, ensuring transparency in their deployment and use.
- **Data Governance and Privacy:** AI systems must adhere to data protection laws and principles, with deployers implementing measures to safeguard the privacy and confidentiality of data used by AI systems.
- **Human Oversight and Responsibility:** Deployers must ensure human oversight over AI systems, particularly in high-risk applications, to prevent autonomous decision-making that could result in harm or discrimination.
- **Robustness and Safety:** AI systems must be designed and deployed in a manner that ensures their robustness, reliability, and safety, with deployers conducting thorough testing and validation to identify and mitigate associated risks.
- **Non-discrimination and Fairness:** Deployers must implement measures to prevent bias and discrimination in AI systems, ensuring that they do not discriminate against individuals or groups based on protected characteristics.
- **Environmental Sustainability:** Deployers should consider the environmental impact of AI systems and strive to minimize their energy consumption and carbon footprint through energy-efficient programming and design techniques.
- **Stakeholder Engagement:** Deployers are encouraged to engage with stakeholders throughout the AI deployment lifecycle, including civil society organizations, academia, and industry representatives, to ensure transparency, accountability, and inclusivity.

By addressing these key areas, this Code aims to promote responsible AI deployment practices that respect fundamental rights, promote fairness, and mitigate potential risks and harms associated with AI technologies. Deployers are encouraged to voluntarily adopt and implement the principles and requirements outlined in this Code to ensure the responsible development and deployment of AI systems, contributing to the advancement of a responsible and sustainable AI ecosystem.





Definitions

AI Act Term	AI Act Definition
AI System	A machine-based system that is designed to operate with varying levels of autonomy and that can, for explicit or implicit objectives, generate outputs such as predictions, recommendations, or decisions, that influence physical or virtual environments.
AI Washing	A practice whereby trustworthy-related statements, declarations, actions, or communications do not clearly and fairly reflect the underlying trustworthy profile of an AI solution, service, platform, product or other asset. This practice may be misleading to consumers, investors, or other market participants.
Authorised Representative	Any natural or legal person located or established in the EU who has received and accepted a mandate from a Provider to carry out its obligations on its behalf.
Deployer	A natural or legal person, public authority, agency, or other body using an AI system under its authority.
Distributor	Any natural or legal person in the supply chain, not being the Provider or Importer, who makes an AI System available in the EU market.
General-Purpose AI Model (“GPAI”)	Means an AI model, including where such an AI model is trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable of competently performing a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications, except AI models that are used for research, development or prototyping activities before they are placed on the market;
Importer	Any natural or legal person within the EU that places on the market or puts into service an AI system that bears the name or trademark of a natural or legal person established outside the EU.
Operator	A general term referring to all the terms above (Provider, Deployer, Authorised Representative, Importer, Distributor, or Product Manufacturer).
Product Manufacturer	A manufacturer of an AI System that is put on the market or a manufacturer that puts into service an AI System together with its product and under its own name or trademark.
Provider	A natural or legal person, public authority, agency, or other body that is or has developed an AI system to place on the market, or to put into service under its own name or trademark.





Trustworthy AI	<p>Defined through a set of principles aimed at ensuring that AI systems are developed and used in a manner that is ethical, respects fundamental rights, and is aligned with societal values.</p> <p>These principles, as outlined in the references provided, include:</p> <ul style="list-style-type: none"> • Human Agency and Oversight: AI systems should serve people, respect human dignity, personal autonomy, and can be overseen and controlled by humans. • Technical Robustness and Safety: AI systems should be resilient and secure, minimizing unintended harm and ensuring reliability. • Privacy and Data Governance: Development and use of AI should comply with privacy and data protection rules, ensuring data quality and integrity. • Transparency: AI systems should be transparent, providing traceability and explainability, making users aware of AI interaction, and informing deployers and affected persons about their rights. • Diversity, Non-discrimination, and Fairness: AI development and use should promote equal access, gender equality, cultural diversity, and prevent discriminatory impacts. • Societal and Environmental Well-being: AI systems should benefit society and the environment, contributing positively to societal challenges. • Accountability: There should be mechanisms in place to ensure responsibility and accountability for AI systems and their outcomes.
----------------	---

General Principles

The General Principles outlined in this Code of Conduct for Deployers of AI Systems serve as foundational values that underpin responsible and ethical deployment practices. Adherence to these principles is essential for fostering trust, transparency, and accountability in the development and use of AI technologies. The following principles guide deployers in their efforts to ensure the responsible deployment of AI systems:

- **Transparency:** Deployers must prioritize transparency throughout the entire lifecycle of AI systems. This includes providing clear and understandable information about the capabilities, limitations, and potential impacts of AI systems to relevant stakeholders. Transparency enhances trust and enables stakeholders to make informed decisions about the deployment and use of AI technologies.
- **Accountability:** Deployers are accountable for the decisions, actions, and outcomes of AI systems under their control. They must accept responsibility for any harms caused by AI systems and implement mechanisms for identifying, addressing, and remedying biases, errors, or adverse effects. Accountability ensures that deployers are held responsible for the ethical and responsible use of AI technologies.



- **Fairness:** Deployers must ensure that AI systems do not produce or perpetuate unfair biases or discriminate against individuals or groups based on protected characteristics such as race, gender, ethnicity, or disability. Fairness requires the identification and mitigation of biases in AI systems to ensure equitable outcomes for all stakeholders.
- **Respect for Fundamental Rights:** Deployers must uphold and respect the fundamental rights enshrined in national and international law. This includes protecting the rights to privacy, freedom of expression, and non-discrimination in the deployment of AI systems. Respecting fundamental rights ensures that AI technologies are used in a manner that upholds human dignity and equality.
- **Ethical Considerations:** Deployers must consider the ethical implications of AI deployment and ensure that AI systems align with ethical standards and values. This includes respecting cultural norms, societal values, and ethical guidelines in the development and use of AI technologies. Ethical considerations guide deployers in making decisions that prioritize the well-being and interests of individuals and society.
- **Continuous Improvement:** Deployers should strive for continuous improvement in the development and deployment of AI systems. This includes regularly assessing and refining processes, practices, and technologies to enhance the effectiveness, efficiency, and ethical robustness of AI deployment. Continuous improvement fosters a culture of learning and adaptation that supports responsible AI deployment practices.
- **Collaboration and Engagement:** Deployers should actively collaborate and engage with stakeholders, including civil society organizations, academia, industry representatives, and affected communities. Collaboration facilitates the exchange of knowledge, perspectives, and insights, enabling deployers to address complex challenges and ensure that AI deployment reflects diverse stakeholder interests and concerns.

Compliance Framework

The Compliance Framework outlined in this Code of Conduct for Deployers of AI Systems serves as a structured approach to ensure adherence to regulatory requirements and ethical standards throughout the deployment lifecycle of AI systems. It provides deployers with a systematic methodology for assessing, monitoring, and evaluating compliance, thereby promoting responsible and ethical deployment practices. The following components constitute the Compliance Framework:

- **Regulatory Requirements Assessment:** Deployers must conduct a comprehensive assessment of relevant regulatory requirements, including those outlined in the EU AI Act and other applicable laws and regulations. This assessment ensures that AI deployment practices align with legal obligations and regulatory standards governing the development and use of AI technologies.
- **Ethical Standards Evaluation:** In addition to regulatory requirements, deployers should evaluate compliance with ethical standards and guidelines governing AI deployment. This includes consideration of ethical principles such as fairness, transparency, accountability, and respect for fundamental rights, as well as adherence to industry-specific ethical codes and frameworks.
- **Impact Assessment:** Deployers are required to conduct impact assessments to evaluate the potential effects of AI systems on various stakeholders, including individuals, communities, and society as a whole. Impact assessments help identify and mitigate risks, including social, economic, and ethical implications, associated with AI deployment.





- **Risk Management Plan:** Based on the findings of impact assessments, deployers must develop a risk management plan to address identified risks and vulnerabilities associated with AI deployment. This plan should outline strategies for risk mitigation, including preventive measures, corrective actions, and contingency plans to minimize adverse effects on stakeholders.
- **Compliance Monitoring Mechanisms:** Deployers should implement mechanisms for ongoing monitoring and evaluation of compliance with regulatory requirements and ethical standards. This includes regular audits, assessments, and reviews of AI deployment practices to ensure continued adherence to legal and ethical obligations.
- **Documentation and Record-keeping:** Deployers must maintain comprehensive documentation and records related to AI deployment activities, including regulatory compliance assessments, impact assessments, risk management plans, and monitoring reports. Documentation facilitates transparency, accountability, and regulatory compliance verification.
- **Training and Awareness Programs:** Deployers should provide training and awareness programs to educate relevant stakeholders, including employees, contractors, and partners, about regulatory requirements, ethical standards, and compliance expectations related to AI deployment. Training programs enhance awareness and understanding of compliance obligations and promote a culture of ethical conduct within organizations.
- **Stakeholder Engagement and Consultation:** Deployers should engage with stakeholders, including affected individuals, communities, civil society organizations, and regulatory authorities, throughout the compliance process. Stakeholder engagement facilitates collaboration, feedback, and consultation, enabling deployers to address stakeholder concerns and perspectives in compliance activities.
- **Continuous Improvement Initiatives:** Deployers should implement continuous improvement initiatives to enhance the effectiveness and efficiency of compliance efforts. This includes feedback mechanisms, performance evaluations, and process optimization to identify areas for improvement and address emerging challenges in AI deployment compliance.
- **Reporting and Disclosure Obligations:** Deployers must fulfill reporting and disclosure obligations related to compliance with regulatory requirements and ethical standards. This includes timely and accurate reporting of compliance activities, findings, and outcomes to regulatory authorities, stakeholders, and the public as required by law or industry standards.

By implementing the Compliance Framework outlined in this Code of Conduct, deployers can ensure that AI deployment practices are conducted in a responsible, ethical, and compliant manner. Compliance with regulatory requirements and ethical standards promotes trust, transparency, and accountability in the development and use of AI technologies, ultimately contributing to the responsible advancement of AI innovation for the benefit of individuals and society.

Transparency and Accountability

Transparency and accountability are fundamental principles in the deployment of AI systems, ensuring that stakeholders have access to clear information about AI technologies and holding deployers responsible for the outcomes of AI applications. This section outlines the key considerations and guidelines related to transparency and accountability in the development, deployment, and use of AI systems.





- **Clear Communication:** Deployers should provide clear and accessible information about AI systems, including their capabilities, limitations, and potential impacts. This information should be easily understandable by stakeholders, including end-users, regulators, and affected communities, to promote transparency and facilitate informed decision-making.
- **Disclosure of Decision-Making Processes:** Deployers must disclose the decision-making processes employed by AI systems, including the algorithms, data inputs, and criteria used to make decisions or predictions. Transparent disclosure enhances accountability and enables stakeholders to assess the fairness and reliability of AI-driven decisions.
- **Risk Communication:** Deployers should communicate the potential risks and uncertainties associated with AI systems openly and transparently. This includes disclosing known limitations, biases, and uncertainties in AI technologies to enable stakeholders to make informed decisions about their use and mitigate potential harms.
- **Algorithmic Transparency:** Deployers must ensure transparency in the design and operation of AI algorithms, providing insights into how algorithms function and how they may impact individuals and society. Transparency in algorithmic decision-making helps build trust, fosters accountability, and enables stakeholders to understand and address biases or errors in AI systems.
- **Accountability for Harms:** Deployers are accountable for any harms caused by AI systems, including errors, biases, or unintended consequences. They must establish mechanisms for identifying and addressing harms promptly, including remediation, compensation, and corrective actions, to uphold accountability and mitigate adverse impacts on affected parties.
- **Bias Detection and Mitigation:** Deployers should implement measures to detect and mitigate biases in AI systems, ensuring fairness and equity in decision-making processes. This includes regular audits, bias assessments, and algorithmic fairness evaluations to identify and address discriminatory outcomes and ensure accountability for bias mitigation efforts.
- **Error Reporting and Correction:** Deployers must establish processes for reporting and correcting errors or inaccuracies in AI systems promptly. Transparent reporting of errors and corrective actions enhances accountability, facilitates learning from mistakes, and promotes continuous improvement in AI deployment practices.
- **External Audits and Reviews:** Deployers should facilitate external audits and reviews of AI systems by independent third parties to assess compliance with transparency and accountability requirements. External scrutiny enhances credibility, promotes trust among stakeholders, and validates the integrity and reliability of AI technologies.
- **Documentation and Documentation:** Deployers must maintain comprehensive documentation of AI systems, including algorithms, data sources, training procedures, and performance metrics. Documentation facilitates transparency, accountability, and reproducibility of AI deployment practices, enabling stakeholders to evaluate and validate the integrity of AI technologies.
- **Ethical Oversight:** Deployers should establish ethical oversight mechanisms, such as ethics boards or committees, to review and assess the ethical implications of AI deployment decisions. Ethical oversight promotes accountability, ensures alignment with ethical principles, and safeguards against the misuse or abuse of AI technologies.

By adhering to the principles of transparency and accountability outlined in this section, deployers can foster trust, promote responsible AI deployment practices, and uphold ethical standards in the



development and use of AI technologies. Transparency and accountability are essential for building public confidence, mitigating risks, and maximizing the societal benefits of AI innovation.

Data Governance and Privacy

Data governance and privacy are critical considerations in the deployment of AI systems, ensuring the lawful and ethical processing of data while safeguarding individuals' privacy rights. This section outlines the key principles and guidelines related to data governance and privacy that deployers of AI systems should adhere to.

- **Compliance with Data Protection Laws:** Deployers must ensure that AI systems adhere to applicable data protection laws and regulations, such as the General Data Protection Regulation (GDPR) in the European Union. This includes obtaining explicit consent for the collection, processing, and storage of personal data and implementing measures to protect individuals' privacy rights.
- **Lawful and Ethical Data Processing:** AI systems should process data in a lawful, ethical, and transparent manner, respecting individuals' rights and freedoms. Deployers must ensure that data processing activities are conducted for legitimate purposes and with appropriate legal justifications, such as consent or legitimate interests.
- **Data Minimization and Purpose Limitation:** Deployers should practice data minimization by collecting only the data necessary for the intended purposes of AI systems. They must also adhere to the principle of purpose limitation, ensuring that data is only used for specified, explicit, and legitimate purposes and not further processed in a manner incompatible with those purposes.
- **Data Quality and Accuracy:** Deployers are responsible for ensuring the quality, accuracy, and reliability of data used by AI systems. This includes implementing measures to maintain data integrity, correctness, and currency throughout the data lifecycle to mitigate the risk of errors, biases, or inaccuracies in AI-driven decision-making.
- **Privacy by Design and Default:** Deployers should integrate privacy by design and default principles into the development and deployment of AI systems, embedding privacy protections and safeguards into the design, architecture, and functionality of AI technologies. Privacy-enhancing features, such as data anonymization, encryption, and access controls, should be implemented by default to minimize privacy risks and protect individuals' personal data.
- **Data Security and Confidentiality:** Deployers must implement robust data security measures to protect against unauthorized access, disclosure, alteration, or destruction of data used by AI systems. This includes encryption, access controls, data masking, and secure storage practices to safeguard sensitive information and ensure data confidentiality.
- **Transparency and Accountability in Data Processing:** Deployers should ensure transparency and accountability in data processing activities, providing clear information to individuals about how their data is being used, processed, and shared by AI systems. Transparency fosters trust and enables individuals to exercise their privacy rights, such as the right to access, rectify, or erase their personal data.
- **Data Retention and Disposal:** Deployers must establish policies and procedures for the retention and disposal of data used by AI systems, ensuring that data is retained only for as long as necessary to fulfill the purposes for which it was collected and in compliance with legal requirements. Data should be securely deleted or anonymized when it is no longer needed, minimizing the risk of unauthorized access or misuse.





- **Cross-Border Data Transfers:** Deployers should assess and manage the risks associated with cross-border data transfers, ensuring that data transferred to third countries outside the European Economic Area (EEA) is adequately protected and processed in accordance with EU data protection standards. Adequate safeguards, such as standard contractual clauses or binding corporate rules, should be implemented to protect individuals' personal data when transferred internationally.
- **Data Governance Framework:** Deployers must establish a comprehensive data governance framework to oversee and manage data-related activities throughout the AI deployment lifecycle. This includes roles and responsibilities for data stewardship, data governance policies, procedures for data risk management, and mechanisms for ensuring compliance with data protection requirements.

By adhering to the principles of data governance and privacy outlined in this section, deployers can ensure the lawful, ethical, and responsible processing of data in AI systems, while upholding individuals' privacy rights and fostering trust in AI technologies. Data governance and privacy are essential components of responsible AI deployment practices, promoting transparency, accountability, and compliance with legal and ethical standards.

Human Oversight and Responsibility

Human oversight and responsibility are essential aspects of the deployment of AI systems, particularly in high-risk applications where autonomous decision-making could have significant consequences. This section outlines the key principles and guidelines related to human oversight and responsibility that deployers of AI systems should adhere to.

- **Ensuring Human Control:** Deployers must ensure that AI systems are subject to human oversight and control throughout their lifecycle. Human oversight involves human intervention in the decision-making process, allowing humans to review, intervene, and override AI-generated decisions when necessary. This is particularly important in critical applications such as healthcare, finance, and criminal justice, where human judgment and expertise are indispensable.
- **Accountability for AI Decisions:** Deployers are accountable for the decisions and actions of AI systems under their control. They must take responsibility for the outcomes produced by AI technologies and address any harm or adverse effects caused by AI-generated decisions. This includes implementing mechanisms for accountability, such as audit trails, transparency reports, and recourse mechanisms for affected individuals.
- **Mitigating Risks of Bias and Discrimination:** Human oversight is crucial for mitigating the risks of bias and discrimination in AI systems. Deployers must actively monitor AI systems for biases and errors, conduct regular audits and evaluations, and implement corrective measures to address any disparities or unfair outcomes. Human intervention can help identify and rectify biased decision-making processes, ensuring fairness and equity in AI-driven decision-making.
- **Ensuring Ethical Use of AI:** Human oversight is necessary to ensure the ethical use of AI technologies and prevent misuse or abuse of AI systems. Deployers must establish ethical guidelines and principles for the development and deployment of AI systems, ensuring alignment with ethical norms, societal values, and human rights standards. Human oversight enables ethical considerations to be integrated into AI decision-making processes, promoting responsible and ethical AI deployment practices.





- **Monitoring System Performance and Reliability:** Human oversight involves monitoring the performance and reliability of AI systems in real-world settings. Deployers should establish mechanisms for ongoing monitoring, evaluation, and validation of AI systems to assess their accuracy, robustness, and safety. Human intervention is essential for identifying system failures, errors, or unexpected behaviors, enabling timely intervention and corrective action to maintain system integrity and performance.
- **Training and Expertise:** Deployers must ensure that individuals responsible for overseeing AI systems possess the necessary training, expertise, and skills to effectively monitor, evaluate, and intervene in AI-driven decision-making processes. This includes providing training and education on AI technologies, ethical principles, and relevant regulatory frameworks to enhance human oversight capabilities and ensure informed decision-making.
- **Transparency and Communication:** Human oversight involves transparent communication with stakeholders about the role of humans in AI decision-making processes. Deployers should provide clear information on the extent of human oversight, the limitations of AI systems, and the potential risks and uncertainties associated with AI-generated decisions. Transparent communication fosters trust, accountability, and confidence in AI technologies, enabling stakeholders to understand and accept the role of humans in AI deployment.
- **Continuous Improvement and Adaptation:** Human oversight requires continuous improvement and adaptation to evolving technological advancements, regulatory requirements, and societal expectations. Deployers should establish mechanisms for feedback, learning, and adaptation to enhance human oversight practices and address emerging challenges and opportunities in AI deployment. Continuous improvement ensures that human oversight remains effective and relevant in safeguarding against risks and promoting responsible AI deployment practices.

By adhering to the principles of human oversight and responsibility outlined in this section, deployers can ensure that AI systems are subject to human control, accountability, and ethical oversight throughout their lifecycle. Human oversight plays a crucial role in mitigating risks, ensuring fairness, and promoting responsible and ethical deployment of AI technologies. Deployers must prioritize human oversight to uphold trust, transparency, and accountability in AI decision-making processes, thereby advancing the responsible development and use of AI systems.

Robustness and Safety

Robustness and safety are paramount considerations in the development and deployment of AI systems to ensure reliability, resilience, and trustworthiness. This section outlines the key principles and guidelines related to robustness and safety that deployers of AI systems should adhere to.

- **Design for Reliability:** AI systems must be designed with a focus on reliability to ensure consistent and accurate performance across diverse scenarios and conditions. Deployers should employ robust design principles, such as redundancy, fault tolerance, and error handling mechanisms, to mitigate the risks of system failures and ensure continuous operation.
- **Thorough Testing and Validation:** Deployers must conduct thorough testing and validation of AI systems to identify and mitigate potential risks and vulnerabilities. This includes testing the system under various inputs, conditions, and edge cases to assess its performance, accuracy, and safety. Rigorous validation processes are essential to uncover any issues or deficiencies in the AI system before deployment, ensuring its reliability and safety in real-world applications.
- **Risk Assessment and Mitigation:** Deployers should conduct comprehensive risk assessments to identify potential hazards and risks associated with AI systems. Risk mitigation strategies should





be implemented to address identified risks and ensure the safety and reliability of AI technologies. This includes implementing safeguards, fail-safe mechanisms, and contingency plans to mitigate the impact of system failures or errors.

- **Adherence to Safety Standards:** AI systems must adhere to safety standards and best practices to ensure compliance with regulatory requirements and industry norms. Deployers should follow established safety guidelines, such as those outlined by relevant regulatory bodies or industry consortia, to address safety-critical aspects of AI system design, deployment, and operation.
- **Continuous Monitoring and Maintenance:** Deployers must establish mechanisms for continuous monitoring and maintenance of AI systems to detect and address emerging risks and vulnerabilities. This includes real-time monitoring of system performance, anomaly detection, and proactive maintenance to ensure the ongoing reliability and safety of AI technologies throughout their lifecycle.
- **Transparency and Explainability:** Deployers should ensure transparency and explainability in AI systems to enhance trust and understanding among users and stakeholders. Transparent AI systems provide insights into their decision-making processes, enabling users to assess their reliability and safety. Explainable AI techniques, such as interpretable machine learning models and decision-making algorithms, facilitate understanding of AI-generated outcomes and enable users to identify potential risks or biases.
- **User Training and Education:** Deployers must provide adequate training and education to users and stakeholders to ensure safe and responsible use of AI technologies. Training programs should cover topics such as system capabilities, limitations, and safety guidelines to empower users to make informed decisions and mitigate risks associated with AI deployment.
- **Ethical Considerations:** Deployers should integrate ethical considerations into the design and deployment of AI systems to ensure alignment with societal values and ethical norms. Ethical AI principles, such as fairness, transparency, and accountability, should guide decision-making processes to minimize the risks of unintended consequences or harm to individuals or communities.
- **Collaboration and Knowledge Sharing:** Deployers should collaborate with industry peers, researchers, and regulatory authorities to share knowledge, best practices, and lessons learned in ensuring the robustness and safety of AI systems. Collaborative efforts facilitate the development of common standards, guidelines, and frameworks for enhancing the reliability, resilience, and safety of AI technologies across different domains and applications.

By adhering to the principles of robustness and safety outlined in this section, deployers can ensure the reliability, resilience, and trustworthiness of AI systems in diverse applications and scenarios. Robust and safe AI technologies inspire confidence among users and stakeholders, fostering trust, acceptance, and adoption of AI-driven solutions to address societal challenges and drive innovation. Deployers must prioritize robustness and safety to uphold the integrity and reliability of AI technologies and promote their responsible and ethical deployment in various domains and industries.

Non-discrimination and Fairness

Ensuring non-discrimination and fairness is essential in the development and deployment of AI systems to uphold human rights, promote equality, and mitigate biases. This section delineates the principles and guidelines that deployers of AI systems must adhere to in order to prevent discrimination and promote fairness.



- **Algorithmic Fairness:** Deployers must prioritize algorithmic fairness to ensure that AI systems do not perpetuate or amplify biases against individuals or groups based on protected characteristics such as race, gender, ethnicity, or disability. They should implement measures to detect and mitigate bias in AI algorithms, including bias testing, fairness-aware training, and algorithmic debiasing techniques.
- **Diverse and Representative Data:** Deployers should utilize diverse and representative datasets to train AI systems, ensuring that the data used is inclusive and encompasses the diversity of the population. By incorporating diverse perspectives and experiences into the training data, deployers can mitigate biases and enhance the fairness of AI systems' decision-making processes.
- **Bias Detection and Mitigation:** Deployers must implement mechanisms for detecting and mitigating biases in AI systems throughout the development lifecycle. This includes regular audits and assessments to identify bias in training data, algorithms, and decision outputs, as well as implementing corrective measures to address any identified biases and ensure fairness in AI-driven outcomes.
- **Explainability and Transparency:** Deployers should prioritize explainability and transparency in AI systems to enable users to understand the rationale behind AI-generated decisions and outcomes. Transparent AI systems facilitate the identification and mitigation of biases, allowing users to assess the fairness and reliability of AI-driven processes.
- **Fairness in Decision-Making:** Deployers must ensure that AI-driven decision-making processes are fair and equitable, avoiding unjustified discrimination or disparate impact on individuals or groups. This includes implementing measures to prevent discriminatory outcomes in areas such as hiring, lending, and criminal justice, where AI systems are increasingly being deployed.
- **Continuous Monitoring and Evaluation:** Deployers should establish mechanisms for continuous monitoring and evaluation of AI systems to assess their fairness and non-discriminatory behavior in real-world applications. This includes ongoing analysis of system performance and decision outcomes to identify and address any emerging biases or discriminatory patterns.
- **User Empowerment and Redress:** Deployers must empower users to challenge and address instances of discrimination or unfair treatment resulting from AI systems. This includes providing avenues for users to report concerns or grievances related to discriminatory AI-driven decisions and implementing mechanisms for redress and corrective action.
- **Ethical Review and Oversight:** Deployers should establish ethical review boards or oversight committees to assess the potential ethical implications of AI systems and ensure compliance with non-discrimination and fairness principles. Ethical review processes should include considerations of fairness, equity, and social justice to prevent the deployment of discriminatory AI technologies.
- **Community Engagement and Consultation:** Deployers should engage with affected communities and stakeholders throughout the AI development and deployment lifecycle to ensure that AI systems are designed and implemented in a manner that reflects community values and priorities. Community engagement facilitates the identification of potential biases and discriminatory impacts, enabling deployers to address concerns and promote fairness and equity in AI applications.

By adhering to the principles of non-discrimination and fairness outlined in this section, deployers can mitigate biases, promote equity, and uphold human rights in the development and deployment of AI systems. Non-discriminatory and fair AI technologies contribute to a more just and equitable society,



fostering trust, inclusivity, and acceptance of AI-driven solutions across diverse communities and stakeholders. Deployers must prioritize non-discrimination and fairness to ensure the responsible and ethical deployment of AI systems and mitigate the risks of discriminatory outcomes and social harm.

Environmental Sustainability

As the deployment of AI systems becomes more pervasive across various industries and sectors, it is crucial for deployers to consider the environmental impact of these technologies. This section of the Code outlines the principles and guidelines for promoting environmental sustainability in the development, deployment, and use of AI systems.

- **Energy Efficiency:** Deployers should prioritize energy-efficient programming and design techniques to minimize the energy consumption of AI systems. This includes optimizing algorithms and models for reduced computational resources, as well as leveraging energy-efficient hardware infrastructure for AI training and inference tasks.
- **Carbon Footprint Reduction:** Deployers must strive to minimize the carbon footprint associated with AI systems by adopting sustainable practices throughout the AI lifecycle. This includes using renewable energy sources for AI infrastructure, reducing unnecessary computational overhead, and optimizing resource utilization to minimize energy-intensive operations.
- **Efficient Training and Deployment:** Deployers should employ techniques for efficient AI model training and deployment to reduce energy consumption and environmental impact. This includes optimizing training workflows, minimizing data processing overhead, and utilizing distributed computing resources to improve energy efficiency during model development and deployment phases.
- **Lifecycle Assessment:** Deployers should conduct lifecycle assessments of AI systems to evaluate their environmental impact across various stages, including development, deployment, and disposal. This includes assessing the energy consumption, carbon emissions, and resource utilization associated with AI technologies to identify opportunities for improvement and optimization.
- **Green AI Practices:** Deployers should adopt green AI practices that prioritize environmental sustainability in AI development and deployment. This includes promoting research and innovation in eco-friendly AI technologies, such as energy-efficient algorithms, hardware optimization, and sustainable computing practices.
- **Eco-Design Principles:** Deployers should integrate eco-design principles into the development and deployment of AI systems to minimize environmental impact. This includes considering environmental factors such as energy consumption, resource utilization, and carbon emissions during the design phase and implementing measures to mitigate environmental risks and promote sustainability.
- **Environmental Impact Assessment:** Deployers must conduct environmental impact assessments of AI projects to evaluate their potential environmental consequences and identify strategies for mitigating negative impacts. This includes assessing the direct and indirect environmental effects of AI systems, such as energy consumption, carbon emissions, and resource depletion, and implementing measures to minimize environmental harm.
- **Continuous Improvement:** Deployers should continuously monitor and evaluate the environmental performance of AI systems and implement measures for continuous improvement and optimization. This includes setting environmental sustainability goals,





tracking key performance indicators related to energy efficiency and carbon footprint, and implementing strategies to achieve environmental sustainability targets.

- **Stakeholder Collaboration:** Deployers should collaborate with stakeholders, including environmental organizations, policymakers, and industry partners, to promote environmental sustainability in AI deployment. This includes sharing best practices, exchanging knowledge and expertise, and collectively addressing environmental challenges associated with AI technologies.

By adhering to the principles of environmental sustainability outlined in this section, deployers can minimize the environmental impact of AI systems and contribute to a more sustainable and eco-friendly future. Environmental sustainability practices in AI deployment not only reduce energy consumption and carbon emissions but also promote responsible and ethical use of technology for the benefit of both present and future generations. Deployers must prioritize environmental sustainability to ensure the responsible development and deployment of AI systems and mitigate the environmental risks associated with these technologies.

AI Literacy and Education

In an era dominated by rapidly advancing technologies, promoting AI literacy and education is essential to ensure that individuals involved in the development, operation, and use of AI systems possess the necessary knowledge and understanding to navigate the complexities of artificial intelligence. This section of the Code emphasizes the importance of AI literacy and education and provides guidelines for fostering a comprehensive understanding of AI technologies and their implications.

- **Foundational Knowledge:** Deployers should prioritize providing individuals with foundational knowledge about AI, including its basic principles, algorithms, and applications. This includes educating stakeholders on key concepts such as machine learning, neural networks, and data analytics to build a solid understanding of AI technologies.
- **Ethical Considerations:** AI literacy and education initiatives should also cover ethical considerations related to AI development and deployment. This includes discussing ethical principles such as fairness, transparency, accountability, and privacy to ensure that individuals are equipped to make informed decisions about AI technologies.
- **Real-world Applications:** Deployers should provide practical examples and case studies to illustrate real-world applications of AI across various industries and sectors. This includes showcasing how AI is being used to solve complex problems and improve decision-making processes in fields such as healthcare, finance, transportation, and education.
- **Risk Awareness:** AI literacy and education efforts should also focus on raising awareness about the potential risks and challenges associated with AI technologies. This includes discussing issues such as bias, discrimination, security vulnerabilities, and job displacement to ensure that individuals understand the multifaceted nature of AI and its impact on society.
- **Continuous Learning:** Deployers should encourage individuals to engage in continuous learning and professional development to stay abreast of the latest advancements in AI technologies. This includes providing access to resources such as online courses, workshops, conferences, and academic programs to support ongoing education and skill development in the field of AI.
- **Interdisciplinary Collaboration:** AI literacy and education initiatives should promote interdisciplinary collaboration and knowledge sharing across different domains and fields of expertise. This includes fostering partnerships between AI experts, domain specialists,





policymakers, educators, and other stakeholders to facilitate cross-disciplinary dialogue and collaboration.

- **Accessible Learning Materials:** Deployers should ensure that AI literacy and education materials are accessible to individuals from diverse backgrounds and educational levels. This includes providing resources in multiple languages, formats, and modalities to accommodate different learning preferences and accessibility needs.
- **Community Engagement:** AI literacy and education efforts should extend beyond traditional educational settings to engage with communities and the general public. This includes organizing outreach events, public lectures, and educational campaigns to raise awareness about AI and empower individuals to participate in discussions about its societal implications.
- **Youth Engagement:** Deployers should prioritize engaging with youth and students to cultivate interest and enthusiasm for AI technologies from an early age. This includes integrating AI education into school curricula, extracurricular activities, and youth programs to inspire the next generation of AI innovators and leaders.
- **Empowerment and Participation:** Ultimately, AI literacy and education initiatives should empower individuals to actively participate in the development, deployment, and governance of AI technologies. This includes equipping individuals with the knowledge, skills, and confidence to contribute to discussions about AI policy, regulation, and ethics and to advocate for responsible and ethical AI practices.

By promoting AI literacy and education, deployers can empower individuals to navigate the complex landscape of AI technologies responsibly and ethically. AI literacy and education initiatives play a critical role in fostering a well-informed and engaged society that can harness the potential of AI to address global challenges and create a more inclusive and equitable future. Deployers must prioritize AI literacy and education to ensure that individuals are prepared to leverage AI technologies for the benefit of humanity while mitigating potential risks and challenges.

Inclusive and Diverse Design

Inclusivity and diversity are crucial considerations in the design and development of AI systems to ensure that they serve the needs of all individuals and communities, regardless of their backgrounds or circumstances. This section of the Code emphasizes the importance of inclusive and diverse design practices and provides guidelines for incorporating inclusivity and diversity into the AI development process.

- **Understanding Diversity:** Deployers should recognize and appreciate the diversity of perspectives, experiences, and needs among potential users of AI systems. This includes considering factors such as age, gender, ethnicity, culture, language, socio-economic status, and accessibility requirements to ensure that AI systems are inclusive and accessible to all individuals.
- **User-Centered Design:** Inclusive and diverse design begins with a user-centered approach that prioritizes understanding the needs, preferences, and challenges of diverse user groups. Deployers should actively engage with end-users and stakeholders from diverse backgrounds throughout the design process to gather insights and feedback that inform the development of AI systems.
- **Accessibility Considerations:** Deployers should integrate accessibility considerations into the design of AI systems to ensure that they are usable and accessible to individuals with disabilities. This includes incorporating features such as alternative input methods, screen





readers, voice commands, and text-to-speech functionality to accommodate diverse accessibility needs.

- **Cultural Sensitivity:** AI systems should demonstrate cultural sensitivity and awareness to avoid perpetuating stereotypes or biases that may marginalize or exclude certain groups. Deployers should conduct cultural assessments and consult with cultural experts to identify and address potential cultural biases or sensitivities in AI systems.
- **Language Diversity:** AI systems should support language diversity by providing multilingual capabilities and language localization features to accommodate users who speak different languages. Deployers should prioritize language diversity in AI interfaces, content, and communication channels to ensure equitable access for all users.
- **Gender-Inclusive Design:** Deployers should adopt gender-inclusive design practices that recognize and respect the diversity of gender identities and expressions. This includes avoiding gender stereotypes and biases in AI systems and providing options for users to specify their preferred gender identity or pronouns.
- **Ethical Considerations:** Inclusive and diverse design should be guided by ethical principles that prioritize fairness, equity, and respect for human dignity. Deployers should consider the potential ethical implications of AI design decisions on diverse user groups and take proactive measures to mitigate bias, discrimination, and harm.
- **Representation and Diversity in Data:** AI systems rely on training data to learn and make decisions, making it essential to ensure that training datasets are representative and diverse. Deployers should strive to collect and use data that reflects the diversity of the target population and includes diverse perspectives, experiences, and demographics.
- **Bias Detection and Mitigation:** Deployers should implement measures to detect and mitigate bias in AI systems, including algorithmic bias, data bias, and user interface bias. This includes using bias detection tools, conducting bias audits, and involving diverse stakeholders in the review and validation of AI systems.
- **Continuous Monitoring and Evaluation:** Inclusive and diverse design is an ongoing process that requires continuous monitoring and evaluation to ensure that AI systems remain inclusive and responsive to evolving user needs and societal dynamics. Deployers should establish mechanisms for gathering user feedback, monitoring system performance, and iterating on design improvements over time.

By prioritizing inclusive and diverse design practices, deployers can create AI systems that are accessible, equitable, and inclusive for all users. Inclusive and diverse design not only enhances the usability and effectiveness of AI systems but also fosters trust, engagement, and empowerment among diverse user groups. Deployers must commit to embracing inclusivity and diversity as core principles of AI design to address the needs and aspirations of all individuals and communities.

Assessing and Preventing Negative Impacts

As AI technologies become more integrated into various aspects of society, it is essential for deployers to assess and mitigate the potential negative impacts that AI systems may have on individuals, communities, and society as a whole. This section of the Code emphasizes the importance of proactive assessment and prevention of negative impacts associated with AI deployment and provides guidelines for deployers to identify, address, and mitigate such impacts.





- **Vulnerability Assessment:** Deployers should conduct comprehensive vulnerability assessments to identify individuals or groups that may be disproportionately impacted by AI systems. This includes assessing the potential impact on vulnerable populations, such as persons with disabilities, marginalized communities, and economically disadvantaged groups, and taking proactive measures to mitigate any negative effects.
- **Accessibility Considerations:** Deployers must ensure that AI systems are accessible to all individuals, including those with disabilities. This involves designing AI systems with accessibility features such as screen readers, voice commands, and alternative input methods to accommodate diverse user needs and promote inclusivity.
- **Fairness and Bias Mitigation:** Deployers should assess AI systems for fairness and bias to prevent discrimination or unfair treatment of individuals based on protected characteristics such as race, gender, ethnicity, or disability. This includes implementing measures to detect and mitigate bias in data, algorithms, and decision-making processes to ensure equitable outcomes for all users.
- **Transparency and Explainability:** Deployers must ensure transparency and explainability in AI systems to enable users to understand how decisions are made and why specific outcomes occur. This involves providing clear explanations of AI system functionalities, decision-making processes, and potential impacts to promote trust and accountability among users.
- **Privacy Protection:** AI systems must adhere to data protection laws and principles to safeguard the privacy and confidentiality of user data. Deployers should implement measures such as data anonymization, encryption, and access controls to protect sensitive information from unauthorized access or misuse.
- **Security Measures:** Deployers must implement robust security measures to protect AI systems from cyber threats, data breaches, and malicious attacks. This includes conducting regular security audits, implementing encryption protocols, and training personnel on cybersecurity best practices to mitigate security risks and ensure the integrity and reliability of AI systems.
- **Ethical Guidelines:** Deployers should adhere to ethical guidelines and principles in the design, development, and deployment of AI systems to promote responsible and ethical use of AI technologies. This includes considering the potential ethical implications of AI applications on individuals and society and integrating ethical considerations into decision-making processes.
- **Continuous Monitoring and Evaluation:** Assessing and preventing negative impacts is an ongoing process that requires continuous monitoring and evaluation of AI systems' performance and outcomes. Deployers should establish mechanisms for gathering user feedback, monitoring system behavior, and evaluating the social, ethical, and environmental impacts of AI deployment to identify and address any negative effects in a timely manner.
- **Collaboration and Stakeholder Engagement:** Deployers should collaborate with stakeholders, including civil society organizations, academic institutions, and government agencies, to assess and address the negative impacts of AI deployment collectively. This involves engaging in dialogue, sharing best practices, and coordinating efforts to mitigate potential harms and promote responsible AI deployment practices.

By prioritizing the assessment and prevention of negative impacts, deployers can ensure that AI systems are developed and deployed in a manner that minimizes harm and maximizes benefits for individuals and society. Deployers must commit to integrating these guidelines into their AI deployment processes to foster trust, transparency, and accountability in the development and use of AI technologies.





Stakeholder Engagement

Stakeholder engagement is a crucial aspect of responsible and ethical deployment of AI systems. This section of the Code emphasizes the importance of actively involving various stakeholders throughout the AI deployment lifecycle to ensure transparency, accountability, and inclusivity. By engaging with stakeholders, deployers can gain valuable insights, address concerns, and build trust in AI technologies among diverse communities and interest groups.

- **Definition of Stakeholders:** Stakeholders in the context of AI deployment include but are not limited to civil society organizations, academia, industry representatives, government agencies, regulatory bodies, end-users, affected communities, and individuals impacted by AI technologies. Deployers must identify and engage with relevant stakeholders to ensure comprehensive and inclusive participation in the AI deployment process.
- **Early Engagement:** Deployers should initiate stakeholder engagement early in the AI development lifecycle to gather input, solicit feedback, and incorporate diverse perspectives into the design and deployment of AI systems. Early engagement allows deployers to identify potential concerns, address ethical considerations, and proactively mitigate risks associated with AI technologies.
- **Transparency and Information Sharing:** Deployers must be transparent about their AI deployment activities and provide stakeholders with clear and accessible information about the purpose, capabilities, limitations, and potential impacts of AI systems. This includes sharing relevant data, methodologies, and decision-making processes to facilitate informed discussions and decision-making among stakeholders.
- **Two-Way Communication:** Stakeholder engagement should involve two-way communication channels that enable meaningful dialogue, active listening, and mutual exchange of ideas between deployers and stakeholders. Deployers should create opportunities for stakeholders to express their views, raise concerns, and contribute to the development of AI systems in a collaborative and participatory manner.
- **Diverse Representation:** Deployers should strive for diverse representation among stakeholders to ensure that the interests, perspectives, and needs of all stakeholders are adequately represented in the AI deployment process. This includes actively seeking input from marginalized communities, underrepresented groups, and individuals who may be disproportionately impacted by AI technologies.
- **Continuous Engagement:** Stakeholder engagement is an ongoing process that requires continuous communication, collaboration, and feedback throughout the AI deployment lifecycle. Deployers should establish mechanisms for regular updates, consultations, and engagement activities to keep stakeholders informed and involved in decision-making processes related to AI deployment.
- **Conflict Resolution:** Deployers should be prepared to address conflicts, disagreements, and divergent interests among stakeholders in a constructive and transparent manner. This may involve facilitating mediation, arbitration, or consensus-building processes to resolve disputes and reach mutually acceptable solutions that uphold the principles of fairness, transparency, and accountability.
- **Capacity Building:** Deployers should invest in capacity building initiatives to empower stakeholders with the knowledge, skills, and resources needed to actively participate in discussions and decision-making processes related to AI deployment. This includes providing



training, educational materials, and opportunities for skill development to enhance stakeholders' understanding of AI technologies and their implications.

- **Ethical Considerations:** Stakeholder engagement should be guided by ethical principles such as respect for autonomy, dignity, and diversity. Deployers must ensure that stakeholder engagement processes are conducted ethically, with full respect for the rights, interests, and well-being of all individuals and communities involved.
- **Documentation and Transparency:** Deployers should document stakeholder engagement activities, including meeting minutes, feedback received, decisions made, and actions taken, to ensure transparency, accountability, and traceability in the AI deployment process. This documentation should be made publicly available to stakeholders and regulatory authorities as appropriate.

By prioritizing stakeholder engagement, deployers can build trust, foster collaboration, and promote responsible and ethical deployment of AI systems that benefit society as a whole. Deployers must commit to implementing robust stakeholder engagement practices and integrating stakeholder feedback into their AI deployment strategies to ensure inclusive, transparent, and accountable decision-making processes.

Governance Mechanisms

Effective governance mechanisms are essential for overseeing compliance with the principles and requirements outlined in this Code of Conduct for Deployers of AI Systems. This section emphasizes the importance of establishing clear roles, responsibilities, and processes to ensure accountability, transparency, and ethical conduct in the development and deployment of AI technologies.

- **Role of the AI Office:** Deployers should designate an AI Office or similar governing body responsible for overseeing compliance with the Code and related regulatory requirements. The AI Office serves as the central point of contact for AI-related matters, providing guidance, support, and resources to stakeholders involved in AI deployment activities.
- **Responsibilities of the AI Office:** The AI Office is tasked with various responsibilities, including but not limited to:
 - Developing and implementing AI governance policies, procedures, and best practices.
 - Conducting regular assessments and audits to ensure compliance with the Code and relevant regulations.
 - Providing training, education, and awareness programs on responsible AI deployment practices.
 - Facilitating stakeholder engagement and collaboration to address emerging challenges and opportunities in the AI landscape.
 - Monitoring technological advancements and regulatory developments to adapt governance mechanisms accordingly.
 - Serving as a point of contact for reporting concerns or grievances related to AI deployment and coordinating with regulatory authorities as necessary.
- **Composition of the AI Office:** The AI Office should comprise multidisciplinary teams with expertise in areas such as AI ethics, law, technology, data privacy, and stakeholder engagement. The composition of the AI Office may include representatives from relevant departments within the organization, external advisors, and independent experts to ensure diverse perspectives and comprehensive coverage of AI-related issues.



- **Decision-Making Processes:** The AI Office should establish transparent and accountable decision-making processes to guide AI deployment activities. Decisions related to AI development, deployment, risk management, and mitigation strategies should be based on ethical considerations, stakeholder input, and best available evidence.
- **Oversight and Review Mechanisms:** The AI Office is responsible for overseeing the implementation of governance mechanisms and conducting regular reviews to assess their effectiveness. This includes monitoring key performance indicators, conducting risk assessments, and soliciting feedback from stakeholders to identify areas for improvement and refinement.
- **Collaboration with Regulatory Authorities:** The AI Office should collaborate with regulatory authorities and industry stakeholders to ensure alignment with regulatory requirements and industry standards. This may involve participating in regulatory consultations, sharing best practices, and contributing to the development of industry guidelines and frameworks for responsible AI deployment.
- **Continuous Improvement:** Governance mechanisms should be subject to continuous improvement and adaptation in response to changing technological, regulatory, and societal landscapes. The AI Office should foster a culture of learning, innovation, and continuous improvement to enhance the effectiveness and relevance of governance mechanisms over time.
- **Transparency and Accountability:** The AI Office should operate with transparency and accountability, providing regular updates and reports on AI deployment activities, compliance efforts, and outcomes achieved. Transparency fosters trust among stakeholders and demonstrates a commitment to responsible AI deployment practices.
- **Ethical Considerations:** Governance mechanisms should be guided by ethical principles such as fairness, justice, autonomy, and respect for human rights. The AI Office must ensure that decisions and actions are consistent with ethical norms and values, prioritizing the well-being and dignity of individuals and communities affected by AI technologies.
- **Integration with Organizational Governance:** Governance mechanisms for AI deployment should be integrated into broader organizational governance structures and processes. This includes aligning AI governance with existing policies, procedures, and risk management frameworks to ensure consistency and coherence in decision-making across the organization.

By establishing robust governance mechanisms, deployers can effectively manage risks, uphold ethical standards, and promote responsible and ethical deployment of AI systems in alignment with regulatory requirements and societal expectations. The AI Office plays a central role in driving these efforts, serving as a catalyst for ethical innovation and responsible AI deployment practices.

Reporting and Enforcement

The effectiveness of any code of conduct relies heavily on the mechanisms put in place for reporting violations and enforcing compliance. In the context of AI systems deployment, it is imperative that deployers establish robust procedures for reporting concerns or grievances related to the deployment of AI systems and cooperate fully with regulatory authorities in investigations and enforcement actions. This section outlines the key components of reporting and enforcement mechanisms within the framework of this Code of Conduct.



- **Establishment of Reporting Channels:** Deployers should establish clear and accessible channels through which individuals can report concerns or grievances related to the deployment of AI systems. These channels should be easily accessible to all stakeholders, including employees, customers, and members of the public, and provide options for anonymous reporting to encourage transparency and accountability.
- **Confidentiality and Protection:** Deployers must ensure the confidentiality and protection of individuals who report concerns or grievances related to AI deployment. Whistleblowers should be safeguarded against retaliation or victimization, and their identities should be kept confidential unless disclosure is required by law or necessary for the investigation process.
- **Investigation Process:** Upon receiving a report of a potential violation or grievance, deployers should promptly initiate an investigation process to assess the validity of the reported concerns and determine appropriate actions. Investigations should be conducted impartially, thoroughly, and transparently, with due regard for the rights of all parties involved.
- **Enforcement Actions:** If a violation of the Code or related regulatory requirements is substantiated through the investigation process, deployers must take appropriate enforcement actions to address the violation and prevent recurrence. Enforcement actions may include corrective measures, disciplinary actions against responsible parties, implementation of additional safeguards or controls, or, in severe cases, termination of AI deployment activities.
- **Cooperation with Regulatory Authorities:** Deployers are obligated to cooperate fully with regulatory authorities in investigations and enforcement actions related to AI deployment. This includes providing access to relevant data, documentation, and personnel, as well as complying with any directives or recommendations issued by regulatory authorities to address identified violations or deficiencies.
- **Transparency in Reporting:** Deployers should maintain transparency in reporting concerns, investigations, and enforcement actions related to AI deployment. Regular updates on the status and outcomes of investigations should be provided to stakeholders, ensuring accountability and building trust in the deployment process.
- **Continuous Improvement:** Reporting and enforcement mechanisms should be subject to continuous improvement based on lessons learned from reported concerns, investigations, and enforcement actions. Deployers should analyze trends, identify systemic issues, and implement corrective measures to strengthen the effectiveness of reporting and enforcement processes over time.
- **Training and Awareness:** Deployers should provide training and awareness programs to educate employees and stakeholders about reporting procedures, ethical obligations, and the importance of compliance with the Code. By promoting a culture of accountability and transparency, deployers can encourage proactive reporting and facilitate early detection and resolution of potential issues.
- **Documentation and Record-Keeping:** Deployers should maintain accurate and comprehensive records of reported concerns, investigations, and enforcement actions for documentation and audit purposes. These records should be securely stored and accessible only to authorized personnel involved in the reporting and enforcement process.
- **Public Accountability:** Deployers should demonstrate public accountability by disclosing information about reported concerns, investigations, and enforcement actions in accordance with applicable legal and regulatory requirements. Transparency in reporting builds trust with stakeholders and enhances the credibility of the AI deployment process.



By establishing robust reporting and enforcement mechanisms, deployers can effectively address concerns, prevent violations, and uphold the principles of responsibility, transparency, and accountability in the deployment of AI systems. These mechanisms play a crucial role in fostering trust among stakeholders and ensuring the ethical and responsible use of AI technologies in accordance with regulatory requirements and societal expectations.

Conclusion

The Code of Conduct for Deployers of AI Systems presented herein serves as a critical framework for guiding the responsible and ethical deployment of AI technologies. Through its comprehensive set of principles, requirements, and guidelines, this Code aims to promote trust, transparency, and accountability in the development and use of AI systems while fostering adherence to regulatory standards, particularly those outlined in the EU AI Act.

Deployers of AI systems play a pivotal role in shaping the ethical landscape of AI deployment, and this Code empowers them with the necessary guidance to navigate complex ethical considerations. By voluntarily adopting and implementing the principles and requirements outlined in this Code, deployers can demonstrate their commitment to responsible AI deployment practices and contribute to the establishment of a trustworthy AI ecosystem.

The principles articulated in this Code underscore fundamental values such as transparency, accountability, fairness, and respect for fundamental rights. Deployers are called upon to ensure transparency in the development, deployment, and use of AI systems, providing clear information on their capabilities, limitations, and potential impacts. Moreover, they must uphold accountability for any harms caused by AI systems and implement mechanisms for identifying and addressing biases and errors.

Data governance and privacy are paramount considerations in AI deployment, and deployers are obligated to adhere to data protection laws and principles to ensure the lawful and ethical processing of data. Likewise, human oversight and responsibility are essential to prevent autonomous decision-making that could result in harm or discrimination, particularly in high-risk applications.

The Code also emphasizes the importance of robustness and safety in AI systems, as well as the need to mitigate risks associated with AI deployment through thorough testing and validation. Furthermore, deployers must ensure that AI systems do not discriminate against individuals or groups based on protected characteristics, and they should strive to promote inclusivity and diversity in AI design and development processes.

Engagement with stakeholders throughout the AI deployment lifecycle is critical to ensuring transparency, accountability, and inclusivity. Deployers are encouraged to collaborate with civil society organizations, academia, industry representatives, and other stakeholders to address societal concerns and promote responsible AI deployment practices.

Effective governance mechanisms are essential for overseeing compliance with the Code, and this section outlines the roles and responsibilities of the AI Office and regulatory authorities in this regard. Deployers are encouraged to establish mechanisms for reporting concerns or grievances related to AI deployment and to cooperate fully with regulatory authorities in investigations and enforcement actions.



In conclusion, the Code of Conduct for Deployers of AI Systems provides a robust framework for promoting responsible and ethical deployment practices in the development and use of AI technologies. By adhering to the principles and requirements outlined in this Code, deployers can contribute to building trust, fostering transparency, and ensuring accountability in the responsible development and deployment of AI systems.

