bora

INFORMATION SECURITY BUZZ

RELEASE DATE:
12/6/2024

AI Governance Regulations Tracker: 2025 Edition

What Businesses Need to Know About Responsible and Trustworthy AI Development and Use

Contents

Introduction

In an era where artificial intelligence shapes everything from business innovation to daily life, governments worldwide are racing to create frameworks that protect individuals and societies while enabling progress. The urgency for AI regulation has never been greater.

With AI's potential to influence economies, shift competitive landscapes, and even impact democracy itself, nations are devising unique, often contrasting, governance strategies that reflect their cultural and political values.

From the European Union's pioneering risk-based AI Act, which sets global benchmarks for ethical AI, to China's assertive measures designed to maintain national security and societal values, and the United States' fragmented but rapidly evolving regulatory landscape, a patchwork of policies is emerging. As companies operate across borders, they face a challenging, high-stakes environment where compliance and competitive advantage are inextricably linked.

This guide, jointly developed by the experts at Bora cybersecurity marketing and Information Security Buzz, dives into the latest AI regulations across the globe, exploring the key elements, contrasting strategies, and the practical implications for businesses striving to innovate responsibly. Discover how diverse regulatory approaches impact AI's deployment and shape the future of AI on the world stage.

The European Union AI Act

The European Union's Artificial Intelligence Act (AI Act) is a pioneering regulatory framework established to oversee the development, deployment, and use of artificial intelligence (AI) within the EU.

Adopted in May 2024 and entered in force in 1 August 2024, it represents the world's first comprehensive AI legislation, aiming to ensure that AI technologies are safe, ethical, and respect fundamental rights.

Scope and Objectives

The AI Act applies to providers and users of AI systems within the EU and to entities outside the EU, provided that their AI systems or outputs are used in the EU. Its primary objectives are to mitigate risks associated with AI, promote innovation, and establish a unified legal framework across member states.

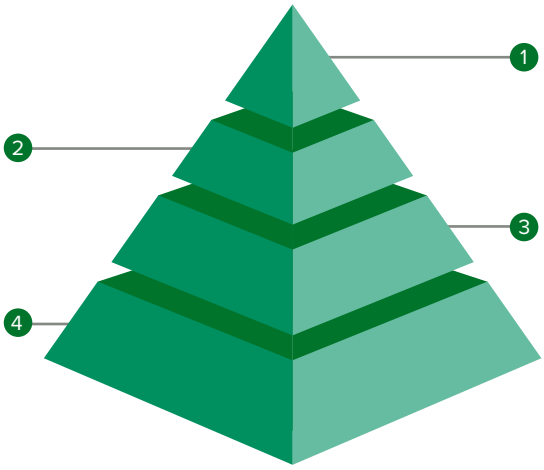
Risk-Based Classification

The Act adopts a risk-based approach, categorizing AI systems into four levels:

- **Unacceptable Risk:** AI applications deemed to pose significant threats to safety or fundamental rights are prohibited. This includes systems that manipulate human behavior or enable social scoring by governments.
- **High Risk:** AI systems used in critical areas such as healthcare, education, employment, and law enforcement are subject to stringent requirements, including risk assessments, data governance measures, and human oversight. Before deployment, high-risk AI systems must undergo conformity assessments to certify compliance with EU standards.
- **Limited Risk:** These systems must adhere to transparency obligations, informing users that they are interacting with AI. Examples include chatbots and AI-generated content.
- **Minimal or No Risk:** AI applications like spam filters or AI-powered games fall under this category and are largely unregulated.

High Risk: Most regulated AI systems, as these have the potential to cause significant harm if they fail or are misused, e.g. if used in law enforcement or recruiting.

Minimal Risk: All other AI systems, e.g. a spam filter, which can be deployed without additional restrictions.



Unacceptable Risk: Highest level of risk prohibited in the EU. Includes AI systems using e.g. subliminal manipulation or general social scoring

Limited Risk: Including AI systems with a risk of manipulating or deceit, e.g. chatbots or emotion recognition systems. Humans must be informed about their interaction with AI.

Figure 1: The EU AI Act Risk Pyramid.
Source: https://www.trail-ml.com/

Obligations and Compliance

Providers of high-risk AI systems are required to implement comprehensive risk management systems, ensure high-quality datasets, maintain detailed documentation, and facilitate human oversight. Depending on the severity of the violation, non-compliance can result in substantial fines, up to €35 million or 7% of global annual turnover.

Compliance Timeline

2 February 2025 (Six months after entry into force):

- Prohibitions on unacceptable risk AI become effective.

2 August 2025 (12 months after entry into force)

- Obligations for providers of general-purpose AI models commence.
- Member states must appoint competent authorities.
- Annual reviews of the list of prohibited AI systems by the European Commission.

2 February 2026 (18 months after entry into force)

- The European Commission implements post-market monitoring regulations.

2 August 2026 (24 months after entry into force)

- Obligations for high-risk AI systems, particularly t hose listed in Annex III (such as biometric systems, critical infrastructure, education, and employment), become effective.
- Member states must establish rules on penalties and set up at least one operational AI regulatory sandbox.
- The European Commission reviews and possibly amends the list of high-risk AI systems.

2 August 2027 (36 months after entry into force)

- Obligations for high-risk AI systems not listed in Annex III but intended as safety components of products come into effect.
- High-risk AI systems that must undergo third-party conformity assessments under existing EU laws (for example, toys, medical devices, and civil aviation security) are also covered.

By the End of 2030:

- Obligations for AI systems that are components of large-scale information technology systems established by EU law in areas like freedom, security, and justice (for instance, Schengen Information System) come into effect.

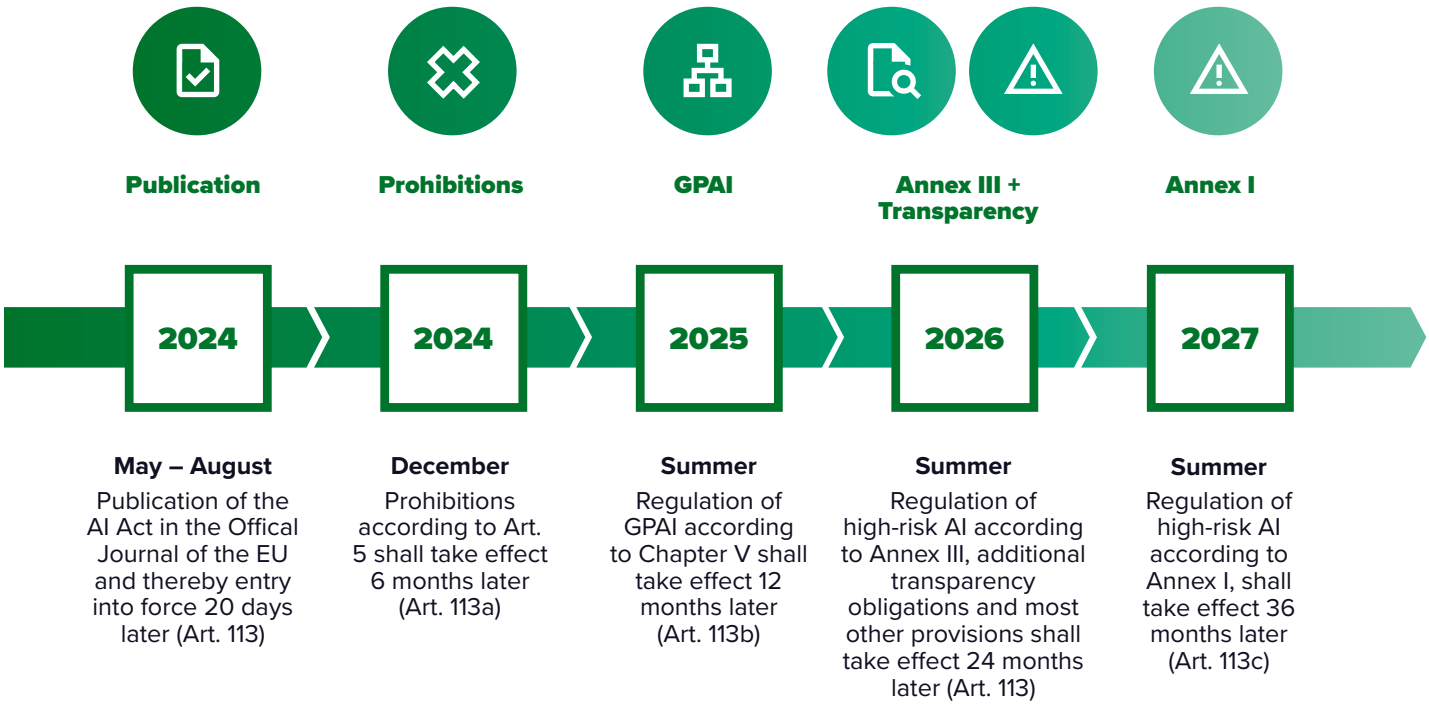


Figure 2: EU AI Act Compliance Timeline.
Source: https://www.pwc.de/en/risk-regulatory/responsible-ai/navigating-the-path-to-eu-ai-act-compliance.html

What is the connection between the AI Act, DSA, and DMA?

The EU AI Act, Digital Services Act (DSA), and Digital Markets Act (DMA) form an integrated regulatory ecosystem addressing different yet interconnected aspects of digital governance.

The **AI Act** establishes a risk-based framework for AI, emphasizing transparency, accountability, and human oversight, especially for high-risk AI applications like biometric surveillance and credit scoring.

The **DSA** focuses on creating a safer digital environment, mandating accountability, transparency in content moderation, and systemic risk assessments, particularly for very large online platforms (VLOPs).

Meanwhile, the **DMA** targets fair competition, regulating "gatekeeper" platforms to prevent anti-competitive practices and ensuring data access and interoperability for smaller businesses.

Together, these acts address overlapping areas, such as transparency and risk management, especially for AI-powered recommendation engines and content moderation algorithms. Businesses face unified compliance obligations, as they must adhere to technology-specific (AI Act) and broader digital service regulations (DSA and DMA). By aligning on transparency, accountability, and competitive fairness, the AI Act, DSA, and DMA create a comprehensive framework that encourages innovation while protecting user rights and fostering a fair digital market across the EU.

How does the AI Act complement GDPR?

The EU AI Act and the General Data Protection Regulation (GDPR) work together to protect personal data and ensure ethical AI use across Europe. While each legislation addresses different facets of data governance and privacy, they share a lot of commonalities, and they complement each other.

Both frameworks share a risk-based approach. The AI Act categorizes AI systems by risk level, with stringent requirements for high-risk applications. This aligns with GDPR's goal to protect personal data, ensuring that high-risk AI systems implement robust safeguards to mitigate potential risks to privacy.

Transparency and accountability are central to both regulations. GDPR requires organizations to inform individuals about how their data is used and grants rights to access, rectify, or delete it. The AI Act, on the other hand, builds on this by mandating that users are informed when they interact with an AI system, ensuring transparency in AI-driven interactions. Additionally, the Act's logging and traceability requirements for high-risk systems enhance GDPR's accountability principles, providing an auditable record of how AI systems process personal data.

The AI Act also strengthens human oversight for high-risk AI, aligning with GDPR's protection against fully automated decision-making that significantly affects individuals. This ensures that critical AI-based decisions, especially those involving personal data, have a "human in the loop" to prevent biases or errors, reinforcing individual rights.

Overall, the AI Act and GDPR in tandem create a comprehensive compliance landscape, compelling businesses to prioritize both data protection and responsible AI use, enhancing privacy, accountability, and trust across the EU.

The AI Act's impact on European and Global businesses

The EU AI Act has significant implications for both European and global businesses, setting a new standard for AI governance that impacts compliance, innovation, and competitive dynamics.

1

High compliance standards and extraterritorial scope

The AI Act mandates rigorous compliance for businesses deploying AI within the EU, especially those using high-risk applications. The Act's extraterritorial nature means that any business, regardless of location, must comply if it markets AI products or services in the EU. This affects global tech companies, including AI providers and businesses using AI for functions like biometric identification, recruitment, or credit scoring. It incentivizes non-EU companies to align with these standards, potentially influencing AI practices worldwide.

2

Innovation with guardrails

For European businesses, the Act aims to enable responsible innovation by encouraging the safe development and deployment of AI. For instance, AI sandboxes are provided to help companies experiment with AI in a controlled, compliant environment. However, stringent requirements might also limit rapid AI experimentation, particularly for smaller businesses facing compliance costs.

3

Competitive advantage in ethical AI

Businesses that comply with the AI Act may benefit from increased consumer trust, especially in data-sensitive sectors like finance, healthcare, and human resources. Companies proactively aligning with the Act's ethical standards may gain a competitive edge, positioning themselves as leaders in responsible AI, appealing to a global market that values privacy and transparency.

The United States

Although the US is at the vanguard of AI innovation, its regulatory approach appears to be disjointed, with a hotchpotch of laws (e.g. city laws, state laws) and non-binding guidelines.

While the country lacks a comprehensive federal law governing AI development, deployment, and use, a few AI-related Acts are in place. However, they mainly address specific administrative issues within the federal government, with limited impact outside the public sector.

Executive Order 14110 "Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence"

Executive Order 14110, titled "Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence," was signed by President Joe Biden on October 30, 2023. This directive establishes a comprehensive national strategy for AI governance, focusing on promoting innovation while mitigating associated risks. Key objectives include enhancing competition in the AI industry, safeguarding civil liberties, and maintaining U.S. leadership in AI technology. The order mandates federal agencies to appoint Chief AI Officers and develop guidelines for AI deployment, emphasizing transparency, accountability, and ethical standards. It also calls for the development of watermarking systems for AI-generated content to address concerns like misinformation and intellectual property theft. By implementing these measures, the order aims to ensure that AI technologies are developed and utilized in ways that are safe, secure, and aligned with democratic values.

Memorandum on Advancing the U.S. Leadership in AI

On October 24, 2024, President Joe Biden issued a memorandum titled "Advancing the United States' Leadership in Artificial Intelligence; Harnessing Artificial Intelligence to Fulfill National Security Objectives; and Fostering the Safety, Security, and Trustworthiness of Artificial Intelligence." This directive outlines a comprehensive strategy to integrate artificial intelligence (AI) into national security functions while ensuring its ethical and secure development.

The memorandum emphasizes three primary objectives:

1. **Leadership in Safe AI Development:** The U.S. aims to lead globally in creating AI systems that are safe, secure, and trustworthy. This involves collaboration with industry, academia, and civil society to promote and protect the foundational capabilities essential for AI advancement.
2. **Harnessing AI for National Security:** The memorandum directs federal agencies to adopt AI technologies responsibly to achieve national security goals. This includes enhancing intelligence operations, defense capabilities, and cybersecurity measures through AI integration.
3. **Mitigating AI Misuse:** It stresses the importance of preventing the misuse of AI technologies, both domestically and internationally, to protect democratic values and human rights.

The memorandum also includes a classified annex addressing sensitive national security issues related to AI. By implementing these measures, the U.S. seeks to balance AI innovation with ethical considerations, maintaining its competitive edge while safeguarding national and global security interests.

The memorandum serves as a follow-up to Executive Order 14110. The executive order established a comprehensive national strategy for AI governance, emphasizing safe, secure, and trustworthy development and use of artificial intelligence. The subsequent memorandum builds upon this foundation by outlining specific strategies to integrate AI into national security functions while ensuring ethical and secure development. Together, these documents reflect a coordinated effort by the U.S. government to lead in AI innovation while addressing associated risks and ethical considerations.

NIST AI Risk Management Framework

The NIST AI Risk Management Framework (AI RMF) is a voluntary guideline developed by the National Institute of Standards and Technology to assist organizations in managing risks associated with artificial intelligence. Released in January 2023, it offers a structured approach to enhance the trustworthiness of AI systems by focusing on principles such as transparency, fairness, and accountability.

The NIST AI RMF is centered around four core principles, also called "functions," which guide organizations through managing AI risks:

- **Govern:** Establish policies, accountability structures, and transparency practices around AI to guide ethical and responsible AI use.
- **Map:** Understand and analyze potential risks by mapping out AI system goals, capabilities, stakeholders, and potential impacts.
- **Measure:** Develop metrics and testing methodologies to assess how well the AI system aligns with intended outcomes and to detect potential biases, vulnerabilities, and security issues.
- **Manage:** Implement controls and continuously monitor AI systems to mitigate identified risks, adapting as new risks arise.

The framework is designed to be adaptable, allowing organizations to tailor its use to their specific needs and contexts. It emphasizes the importance of continuous monitoring and assessment to address the evolving nature of AI technologies and their associated risks.

NIST has also released the "Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile" (NIST AI 600-1). This profile serves as a companion to the NIST AI RMF and focuses specifically on the unique risks associated with generative AI technologies. The Generative AI Profile identifies twelve primary risks unique to or exacerbated by generative AI, including:

- Confabulation (generation of false or misleading content)
- Data privacy concerns
- Harmful bias or homogenization
- Information security vulnerabilities

To address these risks, the profile provides over 200 suggested actions aligned with the AI RMF's core functions. These actions guide organizations in implementing effective risk management practices for generative AI systems.

The Executive Order 14110 directs federal agencies to align their AI-related efforts with the principles outlined in the NIST AI RMF, ensuring a cohesive national strategy for AI governance. This alignment underscores the framework's pivotal role in shaping AI policies and practices across the United States.

Algorithmic Accountability Act

The U.S. Algorithmic Accountability Act is proposed legislation aimed at enhancing transparency and accountability in automated decision-making systems. It mandates that companies assess the impacts of AI systems they develop, use, or sell, focusing on identifying and mitigating potential biases and discriminatory outcomes. The Act requires businesses to conduct impact assessments, particularly for high-risk AI applications affecting areas like employment, housing, and credit. These assessments must evaluate the systems' accuracy, fairness, privacy, and security.

The Federal Trade Commission (FTC) is tasked with enforcing compliance, ensuring that companies adhere to these standards to protect consumers from harm caused by flawed or biased algorithms. By promoting responsible AI development and deployment, the Act seeks to safeguard individuals' rights and foster trust in automated systems.

As of November 2024, the bill remains under consideration in Congress and has not yet been enacted into law.

New York City Bias Audit Law

The New York City Bias Audit Law, officially known as Local Law 144 of 2021, mandates that employers and employment agencies using automated employment decision tools (AEDTs) conduct annual independent bias audits. Effective from July 5, 2023, this law aims to identify and mitigate discriminatory outcomes in AI-driven hiring and promotion processes. Employers must also provide candidates with prior notice of AEDT usage and publicly disclose audit results. Non-compliance can result in penalties, underscoring the city's commitment to fairness and transparency in employment practices.

EU vs. US – A Comparative Analysis

Governance Element	European Union (EU)	United States (US)
Core Legislation	EU AI Act	Lacks comprehensive federal law; relies on a mix of state, city, and non-binding guidelines (e.g., EO 14110, NIST AI RMF)
Approach to Regulation	Centralized, risk-based framework with strict, uniform rules across all member states	Fragmented and sector-specific, with varied approaches across federal, state, and local levels
Risk Categorization	Four levels: Unacceptable, High, Limited, and Minimal risk	Limited to specific applications, such as high-risk AI (e.g., NYC Local Law 144 for employment tools)
High-Risk AI Requirements	Conformity assessments, transparency, human oversight, and regular documentation for high-risk applications	Varies by law (e.g., annual bias audits for AEDTs under NYC Local Law 144, impact assessments for ADS under Algorithmic Accountability Act)
Transparency & Accountability	Requires providers to disclose information about AI use, especially for high-risk applications	Relies on impact assessments and independent audits for certain sectors but lacks overarching requirements
Data Protection	Integrated with GDPR, ensuring AI systems respect data privacy, consent, and user rights	Data protection varies; no federal-level requirement aligning AI regulation with data privacy standards like the EU's GDPR
Compliance & Penalties	Penalties up to 35 million euros or 7% of global annual turnover for non-compliance	Penalties depend on specific state or city laws; FTC enforces compliance under certain acts, with penalties focused on non-compliance with audits and transparency
Impact on Businesses	High for companies due to strict risk management and human oversight requirements	Varied impact based on jurisdiction and industry; guidelines like the NIST AI RMF encourage, but do not mandate, best practices

The UK

The UK government has adopted a context-driven, proportionate regulatory approach, using existing sectoral laws to implement guardrails for AI systems.

Pro-Innovation Approach to AI Regulation White paper

The UK's "Pro-Innovation Approach to AI Regulation" white paper, published in March 2023, outlines the government's strategy to foster artificial intelligence (AI) development while ensuring safety and public trust. This approach emphasizes flexibility and adaptability, aiming to balance innovation with appropriate oversight.

The UK's approach is based on five key principles:

1. **Safety, Security, and Robustness:** AI systems should function reliably and be resilient to risks.
2. **Appropriate Transparency and Explainability:** AI operations must be understandable to users and stakeholders.
3. **Fairness:** AI should not perpetuate bias or discrimination.
4. **Accountability and Governance:** Clear responsibility structures are essential for AI deployment.
5. **Contestability and Redress:** Mechanisms should exist to challenge and rectify AI-driven decisions.

These principles are designed to be interpreted and applied by existing regulators across various sectors. They allow for context-specific regulation without imposing rigid, one-size-fits-all rules. This sectoral approach leverages the expertise of regulators familiar with their respective industries, promoting a nuanced application of AI oversight.

The government plans to support regulators in developing tailored guidance and tools to apply these principles effectively. This includes establishing a central function to monitor AI developments, facilitate coordination among regulators, and ensure a coherent regulatory landscape. Additionally, the white paper proposes the creation of regulatory sandboxes to allow businesses to test AI innovations in a controlled environment, fostering experimentation while maintaining safety standards.

Following up, the UK government has established the AI Safety Institute (AISI). Established to advance AI safety, AISI conducts research and builds infrastructure to test advanced AI models for potential threats, aiming to place its operations on a statutory footing.

AI Regulation Bill (Draft)

The UK's Artificial Intelligence (Regulation) Bill, introduced in the House of Lords on November 22, 2023, aims to establish a comprehensive framework for AI governance. Key provisions include the creation of an AI Authority responsible for overseeing AI regulation and ensuring safety, transparency, fairness, accountability, and contestability in AI applications.

Like the EU AI Act, the bill mandates that businesses developing, deploying, or using AI systems conduct thorough testing and comply with applicable laws. The bill also introduces a system for categorizing AI systems based on risk levels, with corresponding regulatory requirements for each category. Higher-risk AI applications would face stricter scrutiny.

As of November 14, 2024, the bill has undergone multiple readings and is progressing through the legislative process.

EU vs. UK – A Comparative Analysis

Governance Element	European Union (EU)	United Kingdom (UK)
Core Legislation	EU AI Act	Pro-Innovation Approach to AI Regulation Whitepaper; AI Regulation Bill (Draft)
Approach to Regulation	Centralized, risk-based framework with uniform rules across all member states	Context-driven, proportionate, and sector-specific approach, aiming to support innovation while addressing risks
Risk Categorization	Four levels: Unacceptable, High, Limited, and Minimal risk	Proposed risk-based framework in draft legislation, with higher-risk AI applications facing stricter scrutiny
High-Risk AI Requirements	Conformity assessments, transparency, human oversight, and regular documentation for high-risk applications	Draft AI Regulation Bill suggests high-risk AI systems must meet compliance standards for transparency, fairness, and accountability
Transparency & Accountability	Mandatory disclosures for high-risk applications	Emphasis on a flexible, principles-based framework with sectoral variation; different sectors may have tailored regulatory measures
Data Protection	Integrated with GDPR to ensure AI systems respect data privacy, consent, and user rights	Guided by the UK's Data Protection Act; AI systems must align with existing data protection and privacy standards
Compliance & Penalties	Penalties up to 35 million euros or 7% of global annual turnover for non-compliance	Proposed penalties for non-compliance in draft legislation; exact figures to be determined based on sector and risk
Impact on Businesses	High, with extensive risk management, human oversight, and stringent requirements for high-risk AI	Moderate, with a focus on promoting innovation; encourages companies to implement AI in a responsible manner based on sector-specific risks

China

China is a prominent leader in AI innovation, actively working to advance various AI applications while subtly taking the lead in shaping the regulatory framework for AI.

In 2022, the country enacted and implemented three distinct regulatory measures at national, regional, and local levels. This progress continued into 2023 when the government intensified its efforts by introducing national legislation specifically aimed at addressing deepfake technology and generative AI.

Deep Synthesis Management Provisions

China's Deep Synthesis Management Provisions, also known as the "Deepfake Law," were implemented in January 2023 to regulate the use of deep synthesis technology.

This regulation covers many deep synthesis applications, including text, image, audio, and video generation. It mandates deep synthesis activities that adhere to Chinese laws, regulations, and ethical standards. Moreover, explicit consent is required for individuals' personal information to be used in deep synthesis content, and the law prohibits deepfakes that harm reputation or privacy.

It also prohibits creating and disseminating deepfake content that endangers national security, harms the public interest, or violates social order. To prevent misinformation, unambiguous labeling of deepfake content is demanded. It also gives authorities the power to supervise and manage deep synthesis activities.

Algorithmic Recommendation Management Provisions

China's "Internet Information Service Algorithmic Recommendation Management Provisions," effective March 1, 2022, and developed by the Cyberspace Administration of China (CAC), target recommendation algorithms, emphasizing transparency, user rights, and the prevention of discriminatory practices. These rules aim to ensure that algorithmic decisions are fair and accountable.

Key aspects include:

- **Content Regulation:** Providers must ensure that algorithms do not disseminate illegal information or engage in activities harmful to national security or public interest. They are required to promote positive energy and uphold mainstream values.
- **User Rights:** Users have the right to be informed about the principles and purposes of algorithmic recommendations. Providers must offer options to disable personalized recommendations and facilitate user control over algorithmic outputs.
- **Transparency and Accountability:** Providers are obligated to disclose the basic principles, purposes, and main operating mechanisms of their algorithms. They must establish mechanisms for manual intervention and user choice, ensuring transparency in information filtering and recommendation processes.
- **Algorithm Filing:** Services capable of influencing public opinion or social mobilization are required to file their algorithms with the Cyberspace Administration of China within ten working days of service provision. This filing includes detailed information about the algorithm's purpose, data sources, and operational mechanisms.

Interim Measures for the Management of Generative AI Services

China's "Interim Measures for the Management of Generative Artificial Intelligence Services," effective August 15, 2023, establish guidelines for AI services accessible to the public. Key provisions include:

- **Content Alignment:** AI-generated content must adhere to "Core Socialist Values" and avoid undermining state authority.
- **Data Management:** Providers are responsible for the legality and accuracy of training data, ensuring it doesn't infringe on intellectual property rights.
- **User Accountability:** Users must register with real identities, and providers should implement measures to prevent misuse.
- **Transparency:** AI-generated content must be clearly labeled to inform users.
- **Security Assessments:** Services with public opinion influence or social mobilization capabilities require security evaluations and algorithm filings.

These measures aim to balance AI innovation with national security and public interest considerations.

Global AI Governance Initiative

China's Global AI Governance Initiative, launched in October 2023, outlines a comprehensive framework for the development and regulation of artificial intelligence (AI) on a global scale. The initiative emphasizes a people-centered approach, advocating for AI technologies that benefit all of humanity while upholding principles of mutual respect, equality, and mutual benefit.

It also highlights the importance of incorporating ethical principles into AI development and use, including ensuring that AI systems are developed responsibly, align with human values, promote social good, and avoid harmful impacts.

The initiative advocates for robust measures to ensure the safety and security of AI systems, involving addressing risks related to data protection, cybersecurity, and the potential misuse of AI technologies. It also stresses the need for transparency in AI operations and decision-making processes. It calls for accountability mechanisms to ensure that AI systems operate responsibly and stakeholders are held accountable for their actions.

The initiative promotes the cooperation between nations to tackle the challenges and opportunities posed by AI while fostering an open and inclusive environment for technological advancement. Finally, it supports the creation of global governance mechanisms, with the United Nations playing a central role in establishing standardized regulations that guide the development and deployment of AI.

EU vs. US vs. China – A Comparative Analysis

Governance Element	European Union (EU)	United States (US)	China
Core Legislation	EU AI Act	No comprehensive federal law; relies on a mix of state, local laws and non-binding guidelines (e.g., NIST AI RMF, Algorithmic Accountability Act)	Multiple national-level regulations, including the Global AI Governance Initiative, Deep Synthesis Management Provisions, and Algorithmic Recommendations
Approach to Regulation	Centralized, risk-based framework with uniform rules across member states	Fragmented, sector-specific approach with guidelines varying across federal, state, and local levels	Centralized, strict regulatory framework with a focus on national security, societal ethics, and control over AI technology development
Risk Categorization	Four levels: Unacceptable, High, Limited, and Minimal risk	Limited categorization within specific laws (e.g., high-risk for employment tools under NYC Local Law 144)	Risk-based approach with stringent controls on high-risk and generative AI applications
High-Risk AI Requirements	Conformity assessments, transparency, human oversight, and regular documentation for high-risk applications	Varies by state and sector (e.g., impact assessments for ADS under the Algorithmic Accountability Act)	Strict regulation on high-risk applications; requires transparency, labeling, and supervision, especially for generative and deepfake technologies
Transparency & Accountability	Requires providers to disclose information on AI use, especially for high-risk applications	Relies on impact assessments and independent audits in certain sectors but lacks overarching requirements	High transparency and accountability requirements, including mandatory labeling of AI-generated content and algorithmic auditing
Data Protection	Integrated with GDPR, ensuring AI systems respect privacy, consent, and user rights	Data protection is fragmented; lacks federal-level alignment with privacy standards like GDPR	Strict data protection laws align with AI governance, including cybersecurity laws and requirements for informed consent for deep synthesis usage
Compliance & Penalties	Penalties up to 35 million euros or 7% of global annual turnover for non-compliance	Penalties vary across state/local laws; FTC enforces compliance in some areas with penalties for non-compliance	Severe penalties for non-compliance; regulatory bodies can impose fines, suspend services, or enforce content takedown
Impact on Businesses	High, due to stringent requirements on risk management and human oversight	Varies by jurisdiction and industry; compliance is encouraged but often voluntary	High impact, especially on foreign businesses; strict control and compliance mandates aligned with national security interests
Human Rights Considerations	Strong focus on fundamental rights; bans certain applications (e.g., social scoring, manipulative AI)	Human rights addressed in guidelines but not consistently enforceable across federal and state levels	Emphasis on ethical AI aligned with national interests, with a strong focus on social stability, safety, and alignment with core socialist values

Canada

In response to AI's growing influence, Canada is developing frameworks that balance innovation with protecting privacy, security, and human rights.

The AI and Data Act (AIDA)

The Artificial Intelligence and Data Act (AIDA) is a proposed Canadian legislation introduced in June 2022 as part of Bill C-27, the Digital Charter Implementation Act. Its primary objective is to establish a regulatory framework for the responsible design, development, and deployment of AI systems in Canada. AIDA aims to ensure that AI technologies are safe, non-discriminatory, and uphold Canadian values and human rights.

Key provisions of AIDA include:

- **Risk-Based Approach:** Similarly to the EU AI Act, AIDA adopts a risk-based framework, focusing on AI systems that pose the highest potential for harm. It categorizes AI systems based on their impact, with high-impact systems subject to stricter requirements.
- **Obligations for High-Impact AI Systems:** Entities responsible for high-impact AI systems must implement measures to identify, assess, and mitigate risks of harm and biased outputs. They are also required to maintain records and ensure transparency in their AI operations.
- **Oversight and Enforcement:** The Act grants authority to designated officials to oversee compliance, conduct audits, and enforce regulations. Non-compliance can result in penalties, including fines and other corrective measures.

The timeline for AIDA's enactment will depend on the progression of Bill C-27 through the remaining legislative stages, including further readings, potential Senate review, and receiving Royal Assent. The Act and its regulations are not expected to come into force before 2025 at the earliest.

Directive on Automated Decision-Making

The Directive on Automated Decision-Making, implemented by the Government of Canada, establishes guidelines for the responsible use of automated decision systems within federal institutions. Effective since April 1, 2019, the directive aims to ensure that such systems are transparent, accountable, and fair, aligning with core administrative law principles.

Key components of the directive include:

- **Algorithmic Impact Assessment (AIA):** Departments must assess the potential impact of automated decision systems using the AIA tool, which evaluates risks and determines the necessary level of oversight.
- **Transparency:** Departments are required to provide clear information to the public about the use of automated decision systems, including publishing the results of AIAs and ensuring explanations are available for decisions made by these systems.
- **Quality Assurance:** The directive mandates rigorous testing and monitoring to ensure data quality and system performance, aiming to prevent unintended biases and errors.
- **Recourse Mechanisms:** It ensures that individuals affected by automated decisions have access to recourse options, allowing them to challenge and seek review of such decisions.

The directive applies to all federal departments using automated systems developed or procured after April 1, 2020, for administrative decision-making processes.

Brazil

Brazil has unveiled an AI Strategy and a proposed AI Bill to date.

National Strategy for AI (EBIA)

In April 2021, Brazil launched the National Strategy for Artificial Intelligence (EBIA in Portuguese), aligning with the OECD AI Principles. Brazil's AI Strategy outlines initiatives to back research projects focused on ethical AI solutions, establish technical standards to promote ethical applications, and develop ways to limit algorithmic bias. It also outlines the need to define parameters for human intervention in high-risk automated decision-making scenarios and enforce codes of conduct to enhance traceability and protect legal rights.

Additionally, Brazil intends to promote data sharing in compliance with its data protection law, the LGPD, and establish an AI observatory to measure impact and distribute open-source codes for detecting discriminatory trends.

AI Bill

Brazil also has a proposed comprehensive AI Bill (Bill No. 2338/2023) aiming to establish general rules for the development, implementation, and responsible use of AI systems in Brazil. The bill follows a risk-based approach, emphasizing protecting fundamental rights and ensuring safe and reliable AI systems. The proposed AI Bill would:

- Prohibit specific "excessive risk" systems.
- Establish strict requirements for high-risk systems.
- Require reporting obligations for significant security incidents.
- Guarantee various individual rights, such as explanation, nondiscrimination, rectification of identified biases, and due process mechanisms.

The bill also outlines governance structures, transparency measures, and accountability mechanisms for AI system providers and operators.

It is unclear when Brazil's Proposed AI Regulation will come into effect and what its final text will entail. Before the president approves it, it must still be scrutinized and voted on in the Federal Senate and the House of Representatives, so the details remain subject to change. There is currently no expected date for the subsequent developments in the legislative procedure.

Singapore

Singapore has developed several voluntary AI governance frameworks to guide businesses in AI's responsible and ethical use. These include:

Model AI Governance Framework

The Model AI Governance Framework, introduced by Singapore in 2019 and updated in 2020, provides organizations with practical guidelines for the responsible development and deployment of AI systems. It emphasizes key principles such as transparency, fairness, and accountability, offering actionable recommendations to ensure AI technologies are used ethically and effectively.

The framework is structured around four key areas:

1. **Internal Governance Structures and Measures:** Organizations are encouraged to establish robust internal frameworks to oversee AI deployment and ensure alignment with ethical standards and regulatory requirements.
2. **Human Involvement in AI-Augmented Decision-Making:** The framework emphasizes the importance of human oversight in AI-driven decisions, advocating for mechanisms that allow human intervention, especially in critical scenarios.
3. **Operations Management:** It provides guidance on effectively managing AI operations, including data management, model monitoring, and continuous evaluation, to maintain system integrity and performance.
4. **Stakeholder Interaction and Communication:** The framework underscores the necessity of transparent communication with stakeholders, ensuring that AI processes and decisions are explainable and understandable to all affected parties.

Model AI Governance Framework or Generative AI

Singapore introduced the Model AI Governance Framework for Generative AI in May 2024. It provides organizations with comprehensive guidelines for responsibly developing and deploying generative AI systems.

Building upon the earlier Model AI Governance Framework, this updated version addresses the unique challenges posed by generative AI technologies. It outlines nine key dimensions: accountability, data management, model development, performance monitoring, transparency, fairness, safety, security, and robustness. By focusing on these areas, the framework aims to foster a trusted AI ecosystem that balances innovation with ethical considerations, ensuring that generative AI systems are developed and used in a manner that benefits society while safeguarding individual rights.

AI Verify

AI Verify is an AI governance testing framework and software toolkit developed by Singapore's Infocomm Media Development Authority (IMDA) and the Personal Data Protection Commission (PDPC). Launched in May 2022, it enables organizations to assess their AI systems against internationally recognized principles such as transparency, fairness, and accountability. The toolkit offers standardized tests and process checks to validate AI performance, promoting responsible AI deployment. In June 2023, Singapore established the AI Verify Foundation to harness global open-source contributions, enhancing AI testing tools and fostering a trusted AI ecosystem.

South Korea

South Korea has established regulatory frameworks for AI that focus on transparency, data protection, and the ethical use of AI technologies, aligning with the country's goal of responsible innovation.

Act on the Promotion of AI Industry and Framework for Establishing Trustworthy AI

South Korea's proposed AI Act, formally known as the "Act on the Promotion of AI Industry and Framework for Establishing Trustworthy AI," was introduced to establish a legal foundation for AI governance and industry promotion. Passed by the Science, ICT, Broadcasting, and Communications Committee in February 2023, the act reflects South Korea's ambition to lead in AI innovation while ensuring technologies align with societal values and international norms.

The key features of the proposed AI Act include:

1. **Ethical Guidelines:** The act incorporates principles to ensure AI technologies are fair, transparent, and non-discriminatory, addressing concerns about bias and misuse.
2. **Risk-Based Approach:** It categorizes AI applications based on their risk levels, with high-risk AI systems subjected to stricter oversight and safety requirements.
3. **Support for AI Industry:** It seeks to boost South Korea's AI sector by providing funding, research opportunities, and infrastructure development to encourage innovation.
4. **Trustworthy AI Development:** The act emphasizes the need for reliable and explainable AI systems, ensuring users can understand and trust AI decisions.
5. **Interagency Cooperation:** A framework for collaboration among government, private entities, and academia is proposed to enhance AI research, policy-making, and compliance.

OECD AI Principles

The OECD AI Principles, established in 2019 and updated in 2024, provide a framework for the responsible development and deployment of artificial intelligence (AI) systems.

These principles aim to promote innovative and trustworthy AI that respects human rights and democratic values.

The framework comprises five values-based principles:

1. **Inclusive Growth, Sustainable Development, and Well-being:** AI should benefit people and the planet by driving inclusive growth, sustainable development, and well-being.
2. **Human Rights and Democratic Values:** AI systems should be designed to respect the rule of law, human rights, democratic values, and diversity, and they should include appropriate safeguards to ensure a fair and just society.
3. **Transparency and Explainability:** AI systems should be transparent and transparently disclosed to ensure that people understand AI-based outcomes and can challenge them.
4. **Robustness, Security, and Safety:** AI systems must function robustly, securely, and safely throughout their life cycles, and potential risks should be continually assessed and managed.
5. **Accountability:** Organizations and individuals developing, deploying, or operating AI systems should be held accountable for their proper functioning in line with the above principles.

Additionally, the OECD provides five recommendations for policymakers to facilitate the implementation of these principles:

- Invest in AI research and development.
- Foster a digital ecosystem for AI.
- Shape an enabling policy environment for AI.
- Build human capacity and prepare for labor market transformation.
- International cooperation for trustworthy AI.

These principles and recommendations serve as a foundation for international cooperation and interoperability, guiding countries in harnessing AI's potential while mitigating associated risks.

A Global Commitment to Responsible AI for a Safer Global Economy

The global commitment to harmonizing innovation with accountability is reflected in the changing landscape of AI regulation.

There is a trend toward establishing governance that promotes the ethical deployment of AI while mitigating risks as nations forge their own unique paths, from the EU's comprehensive AI Act and China's robust control measures to the US's expanding frameworks.

Global AI governance initiatives are revolutionizing the operations of companies in the contemporary interconnected business environment. Flexible governance frameworks are essential for business executives to navigate a complex web of regulations across jurisdictions. Supply chain partnerships are influenced by compliance with these standards, necessitating transparent data management and robust data management to guarantee responsible AI utilization.

Investing in risk management and ethical training becomes indispensable as accountability for AI systems increases. Collaboration across industries is essential to addressing compliance challenges and sharing best practices. Adhering to global standards can harmonize operations, thereby reducing long-term compliance costs and nurturing a competitive advantage.

Companies are positioned as leaders in responsible AI deployment by proactively aligning AI strategies with these initiatives, which not only ensures compliance but also enhances trust. This strategic alignment is essential for maintaining competitiveness and enabling market expansion in a landscape that is becoming increasingly concerned with ethical and transparent AI practices.

Businesses can improve their competitive advantage by implementing a variety of strategic approaches that capitalize on global AI governance initiatives:

- **Regulatory Compliance and Trust:** By aligning with global AI regulations such as the EU AI Act, businesses can assure compliance, avoid legal penalties, and build trust with consumers and partners. This trust is essential for the preservation of a positive brand reputation and consumer loyalty.
- **Risk Management:** Businesses can mitigate risks associated with AI deployment, including data privacy concerns and biases, by establishing effective AI governance frameworks. This proactive risk management can prevent costly errors and improve operational efficiency.
- **Innovation and Efficiency:** AI governance frameworks promote structured innovation by outlining explicit guidelines for the ethical use of AI. This structure enables businesses to safely experiment with new AI technologies, thereby promoting innovation and assuring adherence to ethical standards.
- **Strategic Alignment:** Businesses can optimize the return on AI investments and achieve significant business results by incorporating AI governance into their strategic planning. This ensures that AI initiatives align with broader business objectives.

In general, the implementation of AI governance initiatives not only guarantees compliance but also establishes businesses as pioneers in responsible AI deployment, thereby improving their competitive position in the global market.

Want more information?

About Bora

Bora is a global B2B cybersecurity marketing agency for enterprises that want to build brand awareness, sales pipeline and nurture opportunities in the rapidly evolving information security sector. For enquiries, contact info@welcometobora.com

About Information Security Buzz

Information Security Buzz is an independent cybersecurity publication featuring the latest industry news and expert insights. For media or promotional inquiries, contact

Managing Editor:
editor@informationsecuritybuzz.com

General enquiries:
isb@informationsecuritybuzz.com