eBook

Step-by-Step Guide to AI Compliance

TrustArc

Introduction

Managing Artificial Intelligence (AI) risk involves a comprehensive approach that includes identifying, assessing, mitigating, and monitoring risks of using personal or sensitive information throughout the AI lifecycle.

Within an organization, AI can be applied across many use cases, including your own deployment and use within business processes, engineering, or HR hiring practices. Third party systems and vendors may also utilize personal information from your employees or customers within their AI for marketing or talent acquisition.

Managing AI risk within your organization (business processes, systems, and third-party vendors) involves the same basics and steps when managing your overall privacy risk program. This includes understanding what regulatory requirements and standards apply to your business, managing and determining risk activities, tracking and reporting for ongoing risk and compliance monitoring, and demonstrating compliance with your customers and partners.

AI & Privacy Regulations

Several privacy regulations have specific requirements for using AI, addressing various aspects such as data protection, transparency, accountability, and the rights of individuals. Here are some of the key regulations and their requirements:

EU AI Act (AIA)

- **Risk-based framework:** Prohibits certain unacceptable-risk AI practices; imposes detailed requirements on high-risk AI systems; establishes transparency obligations for certain AI systems and AI-generated or manipulated content; and creates obligations for providers of general-purpose AI models.
- **High-Risk Systems:** Provider obligations may include risk management, data governance, technical documentation, logging, human oversight, accuracy, robustness, cybersecurity, conformity assessment, and post-market monitoring. Deployer obligations are different and may include following instructions for use, ensuring human oversight, monitoring operation, retaining logs, and conducting data protection or fundamental-rights assessments where applicable.
- **Transparency:** Certain systems must inform individuals when they are interacting with AI, and certain synthetic or manipulated content must be identifiable. Article 50 obligations should be described separately from high-risk system requirements.

Colorado AI Act

- **High-risk AI systems:** Colorado's Act applies primarily to AI systems used to make, or substantially contribute to, consequential decisions involving areas such as employment, housing, education, lending, insurance, health care, essential goods or services, legal services, criminal justice, and essential government services.
- **Developer obligations:** Developers must use reasonable care to protect consumers from known or reasonably foreseeable risks of algorithmic discrimination. They must provide deployers with information about intended uses, foreseeable limitations, training data, performance and explainability testing, data governance, intended outputs, mitigation measures, and appropriate use and monitoring.
- **Deployer obligations:** Deployers must implement a risk-management policy, conduct an impact assessment before deployment and at least annually thereafter, and reassess after an intentional and substantial modification. They must address transparency, monitoring, safeguards, and algorithmic-discrimination risks.
- **Consumer rights:** For consequential decisions, deployers must provide notice, the purpose and nature of the decision, contact information, information about applicable opt-out rights, and—when the decision is adverse—reasons for the decision, relevant data and sources, an opportunity to correct inaccurate data, and an appeal opportunity.
- **Affirmative defense:** An affirmative defense may be available where the developer or deployer can demonstrate compliance with an internationally recognized framework, including the NIST AI Risk Management Framework or ISO/IEC 42001. This is not an automatic safe harbor; the party asserting the defense bears the burden of proving that the statutory conditions are satisfied.

EU General Data Protection Regulation (GDPR)

- **Automated decision-making:** Article 22 generally protects individuals from decisions based solely on automated processing, including profiling, where the decision produces legal effects or similarly significant effects. Exceptions may apply where the processing is necessary for a contract, authorized by law, or based on explicit consent, subject to required safeguards. Human involvement must be genuine, meaningful, timely, and capable of changing the outcome; a person merely rubber-stamping an AI output may not be sufficient.
- **Transparency and individual rights:** Organizations must provide meaningful information about the logic involved, the principal factors affecting the outcome, the significance and envisaged consequences of the processing, and available rights. They must also address access, rectification, erasure, objection, and human-review mechanisms where applicable.
- **Core GDPR principles:** AI processing must have a lawful basis and comply with fairness, transparency, purpose limitation, data minimization, accuracy, storage limitation, security, privacy by design, and restrictions on special-category data.
- **DPIAs:** A DPIA is required where AI-related processing is likely to result in a high risk to individuals' rights and freedoms—for example, large-scale profiling, extensive monitoring, use of sensitive data, or significant automated decision-making. AI use alone does not automatically trigger a DPIA.

United Kingdom GDPR and Data Protection Act 2018

- **Automated Decision-Making:** The UK GDPR restricts decisions based solely on automated processing, including profiling, where the decision produces legal effects or similarly significant effects on an individual. Organizations must provide appropriate safeguards, including information about the decision, an opportunity for the individual to make representations, meaningful human intervention, and the ability to contest the decision.
- **Sensitive Personal Data:** Stricter conditions apply when significant automated decisions are based solely on special-category personal data. Organizations should confirm that an applicable legal exception, such as explicit consent or a legal authorization, applies before using such data in this way.
- **Lawful Basis and Data Protection Principles:** AI systems that process personal data must have a lawful basis and comply with UK GDPR principles, including fairness, transparency, purpose limitation, data minimization, accuracy, storage limitation, security, and data protection by design and default.

AI & Privacy Regulations

Several privacy regulations have specific requirements for using AI, addressing various aspects such as data protection, transparency, accountability, and the rights of individuals. Here are some of the key regulations and their requirements:

United Kingdom GDPR and Data Protection Act 2018

- **Transparency and Individual Rights:** Organizations should provide meaningful information about the purpose of AI processing, the principal factors affecting an automated decision, the significance and likely consequences of the processing, and how individuals can exercise their rights. Organizations must also consider access, rectification, erasure, objection, and restriction rights where applicable.
- **Data Protection Impact Assessments:** A Data Protection Impact Assessment (DPIA) is required where AI-related processing is likely to result in a high risk to individuals' rights and freedoms, such as large-scale profiling, extensive monitoring, use of sensitive personal data, or significant automated decision-making.
- **Regulatory Framework:** The Data Protection Act 2018 supplements the UK GDPR and contains UK-specific data-protection provisions. The Data (Use and Access) Act 2025 amended the UK automated-decision framework, including requirements relating to notice, representations, human intervention, and the ability to challenge significant automated decisions.

South Korea Basic AI Act

- **Risk-Based Framework:** The Act, effective January 22, 2026, distinguishes between generative AI and "high-impact AI." High-impact AI includes systems that may significantly affect human life, physical safety, or fundamental rights, including certain systems used in health care, medical devices, biometric identification, recruitment, loan screening, education, transportation, and public services.
- **Transparency:** Businesses incorporating high-impact AI or generative AI into products or services must provide advance notice that the product or service is AI-powered. Providers of generative AI must identify outputs as artificially generated, including through human-readable or machine-readable notices. AI-generated deepfakes must be clearly labeled as AI-generated.
- **High-Impact AI Governance:** Businesses providing products or services that use high-impact AI must assess whether the system qualifies as high-impact AI and implement safeguards that may include a risk-management plan, measures addressing safety and reliability, explainability mechanisms including an overview of relevant training data, human supervision and intervention, user-protection measures, and documentation of safety and reliability measures.
- **Fundamental-Rights Impact Assessments:** Before incorporating high-impact AI into a product or service, businesses must assess the potential impact on users' fundamental rights. Businesses may also be required to publish information about their risk-management policies, output criteria, training-data overview, user safeguards, and the person responsible for AI oversight. Relevant documentation must be retained for the required period.

South Korea Basic AI Act (Cont'd)

- **Large-Scale AI Systems:** AI business operators whose systems exceed the applicable computational threshold for large-scale AI training must establish lifecycle risk-management and monitoring measures to identify, assess, and mitigate AI risks.
- **Foreign Businesses:** Foreign AI business operators that meet applicable user or revenue thresholds and do not have a business address or office in South Korea may be required to appoint and report a domestic representative.
- **Accountability:** Organizations should maintain records demonstrating AI risk assessments, transparency notices, human oversight, user-protection measures, and the controls used to manage safety, reliability, bias, and fundamental-rights risks. The Act imposes administrative penalties for certain failures, including failure to provide required AI disclosures or appoint a required domestic representative

California Consumer Privacy Act (CCPA)

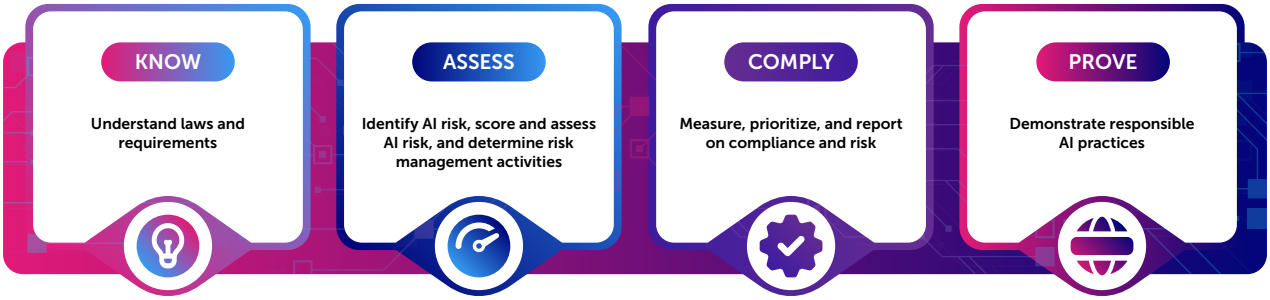
- **Existing privacy rights:** The CCPA gives consumers rights to opt out of the sale or sharing of personal information and to limit certain uses of sensitive personal information. These rights may be relevant to AI training, profiling, behavioral advertising, and related data practices, but they are not a single, general AI opt-out right.
- **ADMT rules:** California's finalized ADMT regulations apply beginning January 1, 2027. Businesses using ADMT for significant decisions must provide a pre-use notice and, subject to exceptions, offer consumers a right to opt out. A business may avoid the opt-out obligation for certain significant decisions if it provides a meaningful appeal to a human reviewer with authority to overturn the result. Consumers also have rights to access information about the specific purpose, logic, output, and outcome of ADMT used concerning them.
- **Risk assessments:** Separate California privacy risk-assessment requirements may apply to certain high-risk processing activities, including some uses involving ADMT, extensive profiling, sensitive personal information, or AI training.

Federal Trade Commission (FTC)

- **FTC Act § 5 / FTC enforcement:** The FTC may challenge unfair or deceptive AI-related practices, including unsupported claims about accuracy or capability, misleading privacy or consent representations, inadequate security, and unfair uses of biometric or algorithmic systems. The FTC has pursued AI-related conduct involving facial-recognition bias, false accuracy claims, and misleading representations about data use. The FTC does not impose a single, comprehensive algorithmic-fairness regime applicable to all AI systems. Organizations should instead assess AI practices under the FTC Act's prohibitions on unfair or deceptive acts or practices, together with applicable sectoral and state laws.

AI Risk Management & Compliance Steps

Privacy and governance programs are dynamic in nature due to the constantly evolving regulatory landscape and volume of personal data for organizations. Steps for AI risk and compliance involve continuous, ongoing management. Generally, the steps evolve around continuous risk identification, management, mitigation, and demonstration.



KNOW

1 Understand what AI laws and compliance requirements apply to your business

A good starting place is to understand how your business uses AI by asking some basic questions like:

What type of AI are you using (e.g., large language model (LLM), Foundation model)?

How does your AI system process data, particularly if it uses personal information to train and/or make decisions?

What are the potential impacts of the AI processing on individuals, like customers and employees?

Key AI stakeholders within your business should be able to explain how the AI system works and *how AI-driven decisions are made*.

AI compliance is more than just complying with AI regulations, such as the EU AI Act. You must ensure that the data you are using to train your AI model is collected and processed in accordance with applicable privacy rules. If your jurisdiction has not yet implemented an AI law or regulation, make sure your AI data processing complies with relevant privacy laws and regulations.

TIP

When it comes to AI compliance, you don't need to reinvent the wheel. Take time to revisit your current privacy management program and incorporate AI into your existing privacy policies and procedures. Look at things like the following to make sure that they appropriately address the use of AI technologies:

- *Privacy notices*
- *Developer disclosures (e.g., training systems, purpose, benefits, evaluations)*
- *Processes for handling individual rights requests (e.g., DSARs) Employee acceptable use policies*
- *Data retention and data security policies*

2 Identify and understand risky activities in your business

Determine if your organization's processing would be considered "high-risk processing." Look to relevant AI and privacy legislation to see if your processing meets the definition of "high-risk", as well as regulator and industry guidance to understand their concerns regarding risky processing activities.

To truly understand risky activities, look beyond just how personal information is being processed. Be sure to consider things like whether you provide appropriate transparency and explainability, and whether you engage with stakeholders in your AI processing to determine if these activities also create a risk to the business.

3 Score and assess AI risk

Implement a system or methodology for classifying risks associated with AI use. For example, you can determine the levels of risk based on the sensitivity of the data processed by the AI system and its intended uses.

Where required by AI or privacy legislation, conduct the appropriate assessment required to identify potential risks to individuals and consequences for the organization. This may include:

- Security risk assessment
- Data protection impact assessment/privacy impact assessment (DPIA/PIA)
- Algorithmic impact assessment
- Other assessments that measure the impact of the processing on individuals and society at large

It's important to remember that in most jurisdictions that require a data protection impact assessment (DPIA), the use of a new technology to process personal data may trigger the need to conduct a DPIA due to potential "high-risks" associated with AI.

Leverage your business's existing risk assessment procedures to incorporate AI. You can also use external risk assessment frameworks, such as the NIST AI Risk Framework, to ensure your risk assessment captures vital elements of AI processing.

4 Determine risk management activities

Develop and document a plan for how your organization will mitigate AI risks. Prioritize risks based on the likelihood that they will occur and the severity of the impact on individuals and the business should they occur.

Identify who in the business is responsible for determining and implementing the controls necessary to mitigate the identified risks or who is responsible for accepting the risk (e.g., business unit leader, process owner, etc.).

Risk mitigation measures should include human oversight and intervention processes that enable a human to intervene in the AI processing at any time to review, assess, and override the AI model outputs.

COMPLY

5 Track and measure AI risk and compliance

Keep up-to-date on existing and emerging privacy laws and AI compliance requirements to ensure that your use of AI, and the processing of any associated personal information, complies with privacy and security requirements.

Sign up for compliance alerts and customizable reports to help stay on top of developments. Appoint dedicated individual(s) or team to communicate important updates and changes in the privacy and AI landscape to internal AI teams and stakeholders so that they can properly assess any impacts on the organizations' use of AI.

6 Ongoing reporting on risk and compliance

Develop and implement plans to continually monitor the functionality and behavior of the AI system after it has been deployed. Regularly review the AI system's outputs to ensure it continues to function the way it is intended. For example, monitoring can be done by:

- Setting up detection mechanisms to alert you when a processing anomaly or change in expected behavior occurs
- Running and reviewing reports on system outcomes at regular intervals

Figure out ahead of time what actions the business will take if the AI system does not perform the way it is intended. For example, the next steps could be to:

- Update and retrain the AI system so that it performs according to its intended use (to the extent the organization has control over the inputs used for systems procured from a third party)
- Cease the use of the AI system where its outputs are likely to result in decisions that produce adverse legal or other detrimental effects (in case of a third-party system, notify the developer that the system does not perform according to its intended use)

PROVE

7 Verify responsible AI practices with a third party

Engage a third party auditor or certification provider to verify you are doing things compliantly. Revisit or review your compliance regularly and especially when a new AI technology or related process is introduced.

Third party assessors should operate independently of the organization involved in the system's design and development to avoid the presumption of bias.

Best Practices of AI Governance Programs

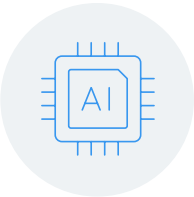
At their core, successful governance programs ensure responsible development, deployment, monitoring, and use of AI technologies. Good governance starts from the top-down of an organization, with appropriate leadership oversight being a key factor.

Establish an AI Risk Committee

Establishing oversight with a specific body or committee to review AI risk use cases can promote improvements and change the cross-functional governance of your organization. Additionally, for organizations in specific industries such as tech, appointing a Responsible AI lead, AI officer, or AI champion to establish an internal interdisciplinary AI board or committee may ensure appropriate attention and oversight.

TIP

For "high-risk" AI use cases, implement specific turn-key processes for ethics review by your organization's governing body or committee. As part of your oversight body or committee, action items should include the maintenance and retention of AI design documentation, developing incident response and recovery plans for AI systems and establishing escalation pathways to send AI-related issues to your oversight body.



Develop and adopt AI-specific policies

To enable successful outcomes of your program, reduce legal risk, and enable compliance, develop and adopt specific AI policies and procedures on how to develop, deploy, or sell AI. Draft policies using and applying privacy and security by design principles as your foundation, as many AI and global privacy and security regulations and standards evolve around the same principles.

Prioritize transparent communication and training

Communicate disclosures in a simple and non-legalese language for employees and customers alike. In particular, provide an accessible privacy policy disclosing essential information to AI end users for any applicable systems. Disclosing the implementation and operations of the use of personal information by AI systems, fosters transparency and awareness. For employees, appropriate training on AI governance measures and ethical considerations. Internal procedures and training promoting the use of smaller, higher quality data sets and clean and curated data sets before model training, can minimize an organization's risk and level of reactiveness.

Leverage human-centric design and values

If your organization is deploying AI, ensure AI models operate using accurate information by considering the completeness, recency, and structure of data. Implement interventions to detect and mitigate sources of bias within the system. Design AI systems to be adaptive to real-world conditions and resilient against systematic errors. Leveraging human-centric design and values can ensure the system will be used in the best interest of humanity. Establish accountability and integrity by developing recovery plans for security incidents involving the AI system.

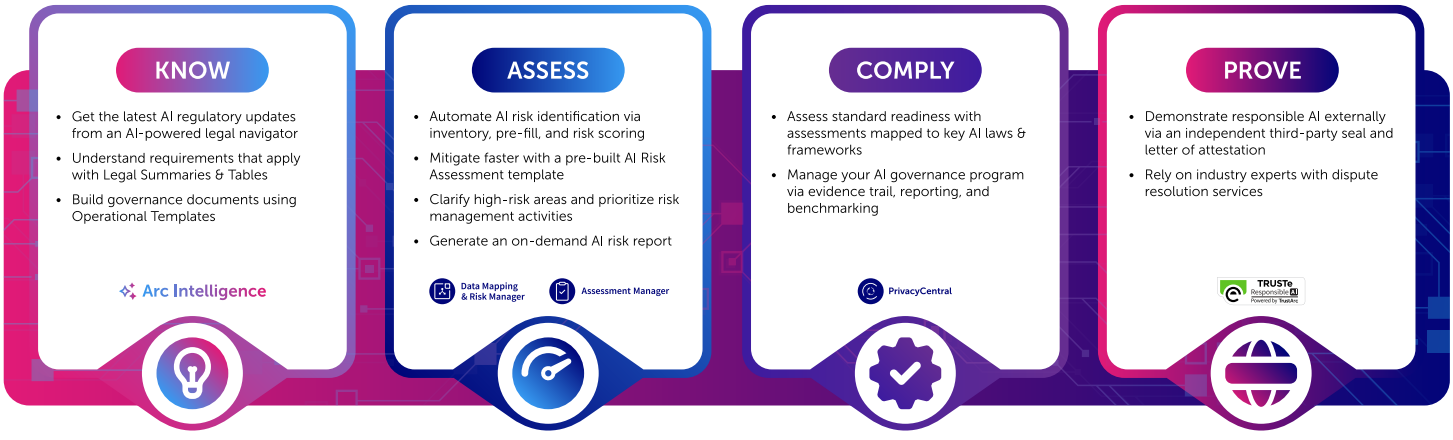
Document AI processes and procedures

An elevated, mature, and proactive governance program involves having policies, training, and recovery plans in place to support accountability and show due diligence. Organizations with a mature AI governance program have robust, well-documented, and repeatable processes and procedures in place. They proactively work to regularly review controls and other elements of their program to ensure continued improvement and full optimization.

Future Proof Your Organization With Streamlined Management

With the evolving and growing number of AI and privacy regulations and the dynamic nature of organizations, purpose-built technology can help you streamline risk management and prioritization for cost savings, speed, and scale.

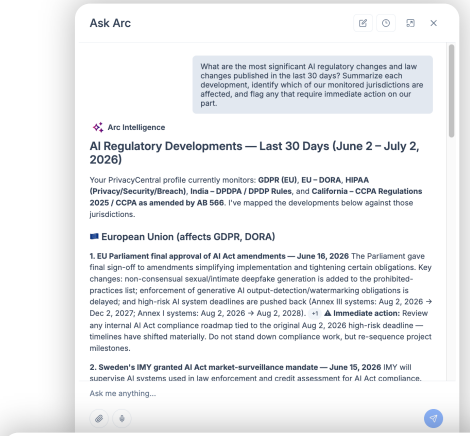
Take out the compliance guesswork of what requirements apply to you, easily identify and mawnage AI risk continuously, and demonstrate your organizations' Responsible AI use with a well-known third party to elevate your credibility and brand trust. TrustArc's AI Risk Governance bundle provides one holistic solution and platform designed to elevate your AI risk, compliance, and overall program governance.

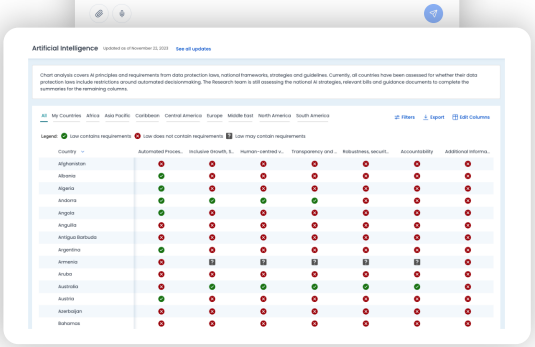


Managed Services

- Get a senior consultant to help your organization get started, including: establish and build your internal governance team, draft a charter, plan strategic development, establish operational best practices, and review AI frameworks (e.g., NIST AI RM, EU AIA, TRUSTe Responsible AI) in relation to your organization

Arc Intelligence & Nymity Research





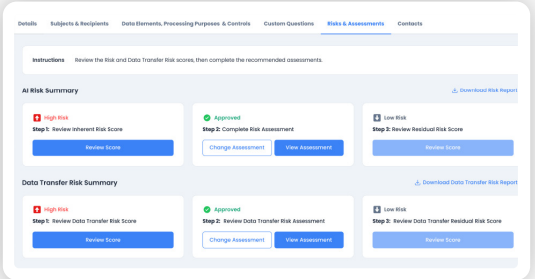
- Quickly get the latest AI regulatory information from an AI-powered legal navigator on AI regulatory developments, requirements, best practices, and 200+ pre-built operational templates

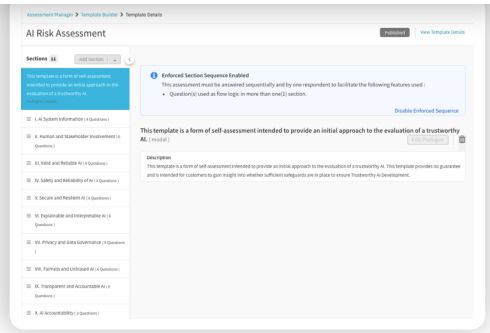
- Understand and compare requirements that apply to your organization in one place with Legal Summaries and Tables

- Start building out AI governance documents using over a dozen of AI-specific Operational Templates (e.g., Testing AI Systems, Procurement Guide for AI Systems, AI Training & Awareness Checklist, Responsible AI Checklist)

Data Mapping & Risk Manager

Assessment Manager



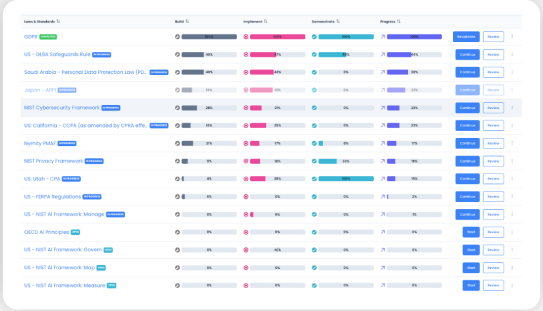


- Intelligently automate AI risk identification through inventory management, pre-fill capabilities, and risk scoring

- Manage and mitigate potential risk faster utilizing a pre-built AI Risk Assessment and risk reporting to understand the biggest risk areas to prioritize

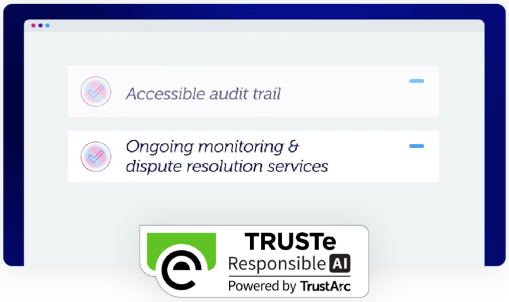
PrivacyCentral

- Assess organizational compliance readiness through readiness assessments mapped to key AI laws and frameworks
- Manage your AI governance program through logging and tracking evidence of compliance, reporting, and benchmarking



TRUSTe Responsible AI Certification

- Demonstrate appropriate responsible use of AI and applicable data protection practices externally through an independent third party assurance seal, letter of attestation, and utilize an industry-recognized dispute resolution service.



Elevate your AI Risk Governance

One holistic solution with flexible options

Get started

About TrustArc

As the leader in data privacy, TrustArc automates and simplifies the creation of end-to-end privacy management programs for global organizations. TrustArc is the only company to deliver the depth of privacy intelligence, coupled with the complete platform automation, that is essential for the growing number of privacy regulations in an ever-changing digital world. Headquartered in San Francisco, and backed by a global team across the Americas, Europe, and Asia, TrustArc helps customers worldwide demonstrate compliance, minimize risk, and build trust. For additional information visit TrustArc.com.

TrustArc