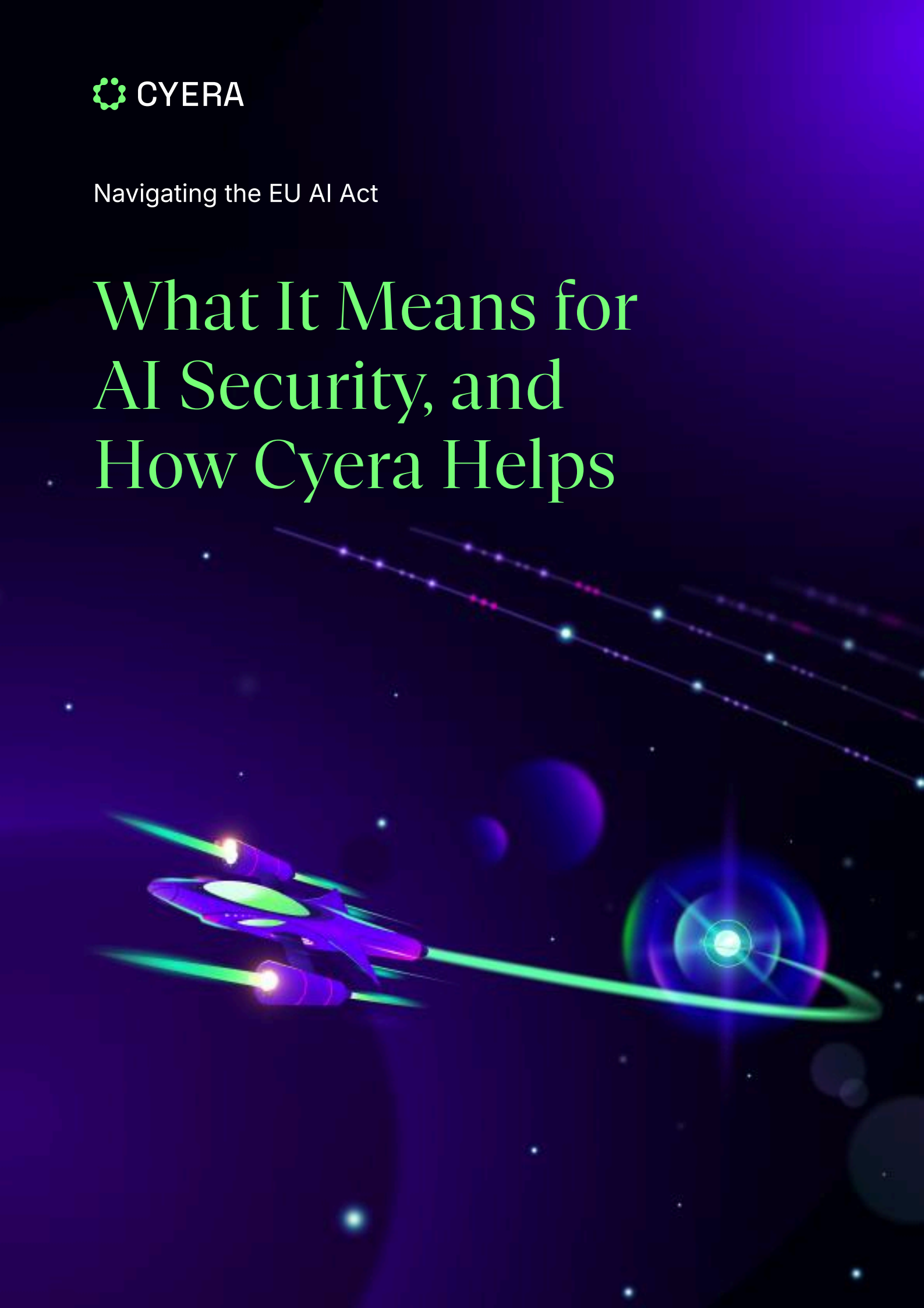
CYERA

Navigating the EU AI Act

What It Means for AI Security, and How Cyera Helps



Overview of the EU AI Act

The European Union's Artificial Intelligence Act takes a risk-based approach to regulating AI systems. It seeks a balance between the benefits of AI innovation and the potential harms to individuals and communities

The Act divides AI systems into four risk tiers:

- Unacceptable
- High
- Limited
- Minimal

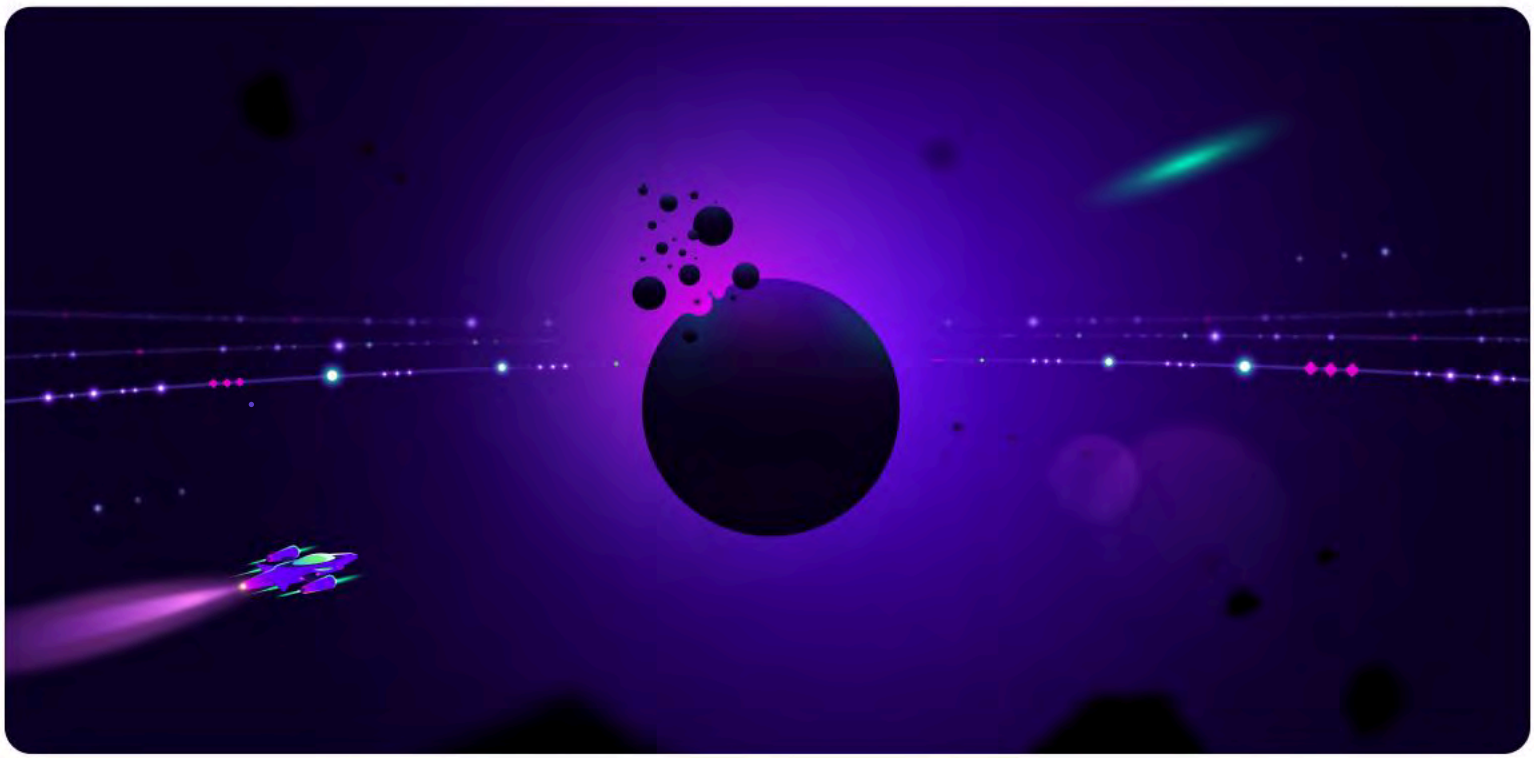
AI systems that pose an unacceptable risk of harm are strictly prohibited. These include systems designed to:

- Distort user behavior by deploying "subliminal, manipulative, or deceptive techniques," or that take advantage of users' specific vulnerabilities relating to age, disability, or socioeconomic status
- Assign individuals a "social score" based on their behavior or traits, and then use that score to justify detrimental treatment
- Assess individuals' risk of committing a crime;
- Scrape the internet or CCTV footage in order to build facial recognition databases
- Infer the emotional state of an individual in a workplace or educational setting, except for medical or safety purposes
- Leverage biometric data to infer sensitive personal information about individuals (such as their sexual orientation, political or religious beliefs, trade union membership, or other categories of sensitive data as defined by the GDPR)
- Deploy real-time remote biometric identification (RBI) of individuals in public places for law enforcement purposes (except when searching for missing, kidnapped, or trafficked persons; identifying suspects in serious crimes; or preventing "specific, substantial, and imminent" threats to life or physical safety).

High, limited, and minimal risk AI systems are permitted under the Act, but each risk tier is subject to different degrees of scrutiny. Minimal risk systems are essentially unregulated, and include things like spam filters and video games. By contrast, providers and deployers of limited risk systems must inform users that they are interacting with an AI system or its outputs. Examples of limited risk systems include those that interact directly with users such as chatbots, systems that generate synthetic images or audio such as deepfakes, or that generate synthetic text designed to inform the public on matters of public interest.

But by far the heaviest regulatory burden falls on providers and deployers of high risk AI systems. It's therefore very important for providers and deployers of AI products to understand whether those products fall within one of the categories designated as high risk, and the specific steps they have to take before they can lawfully place their products on the market. High risk AI systems include all of the following:

- **Biometric systems that perform RBI** (unless solely for the purpose of confirming an individual's claimed identity), that categorize people based on inferences about sensitive or protected attributes such as race, sexual orientation, etc., or that are used for emotion recognition.
- **Any system intended to be used as a safety component** in the management of critical digital, transportation, or energy infrastructure.
- **Systems used in educational or vocational training** settings when they are intended to do things like determine access to educational resources, evaluate learning outcomes, or monitor for prohibited behavior.
- **Systems deployed in employment settings** when used for purposes such as recruiting or evaluating job candidates; making decisions about promotion, termination, or the allocation of responsibilities to individual employees; or to monitor and evaluate employee performance.
- **Systems used to limit or condition individuals' access** to essential public or private benefits and services, including public assistance programs, healthcare services, credit, insurance, and emergency or other first response services.
- **Any system used by law enforcement** to assess the likelihood that a person will become the victim of a crime; as a polygraph or similar tool; to determine the reliability of evidence in a criminal investigation; to predict a person's risk of offending or re-offending; or otherwise in the course of detecting, investigating, or prosecuting criminal offenses.
- **Systems used by government authorities** for migration, asylum, and border control management.
- **Systems designed to influence the administration of justice or democratic processes,** including those used by judicial authorities (or private forums such as arbitration panels) to interpret facts or laws, or apply the law to a given set of facts; and any system designed to influence voting behavior or the outcome of an election.



Hitting the Ground Running

Before diving into the specific requirements for providers and deployers of high risk AI systems, it's helpful to take a step back and consider that the AI Act grows out of a regulatory context that includes other laws like the General Data Protection Regulation (GDPR) and the Digital Services Act. Compliance with these regulations is practically a prerequisite for compliance with the AI Act.

To take a specific example, the Digital Services Act already prohibits "dark patterns," the kind of deceitful or manipulative design techniques that are intended to distort users' behavior and impair their ability to make free and informed decisions. Organizations that take this responsibility seriously will therefore be on the lookout for potential dark patterns in their product designs, and will be that much quicker to spot situations in which AI has hit upon problematic strategies for driving engagement or boosting sales.

More generally, a sweeping framework like the GDPR embraces certain core values, including

- **Transparency:**
  Organizations must notify data subjects of the kinds of data they are collecting, and the uses to which they will be put (and in some cases must get their consent).
- **Minimization:**
  Organizations must not collect or keep more data than necessary to carry out their legitimate business purposes.
- **Purpose limitation:**
  Organizations must not use the data they've collected for purposes other than those they have disclosed to data subjects.

Organizations with mature data governance and privacy policies are already training their workforces to recognize and correct potential deviations from these norms. That makes it much easier for them to anticipate and respond to situations where an AI system infers and attempts to act upon sensitive personal information whose collection was not authorized, that is not necessary to fulfill a legitimate business purpose, or that exceeds the scope of the permission granted by the data subject. Furthermore, organizations with experience conducting data privacy impact assessments (DPIAs) will find it much easier to perform the conformity assessments required by the AI Act.

Cyera's Role In Simplifying EU AI Act Compliance

Strong data governance doesn't just support compliance with the broader context of EU regulations, it also lays a strong foundation for compliance with the AI Act specifically. And it's at this stage that Cyera is an indispensable tool.

Cyera's AI-Native Data Security Platform gives organizations visibility into their data, where it resides, who has access to it, and what they're doing with it. Cyera's agentless DSPM discovers data across the cloud and on-prem, and classifies them according to pre-defined categories aligned with regulatory frameworks like GDPR, as well as customized categories created by the organization. It even learns to develop and suggest new categories of data classification based on the kind of data an organization collects and how it uses them. With Cyera, administrators can see when employees or contractors are feeding personal data into unmanaged applications and services (Shadow IT).

Furthermore, Cyera's Omni DLP can monitor and track data flowing to third parties such as vendors or suppliers, who may also be using AI-powered apps, and then analyze the existing DLP vendor's policies to auto-suggest ways to make them more effective.

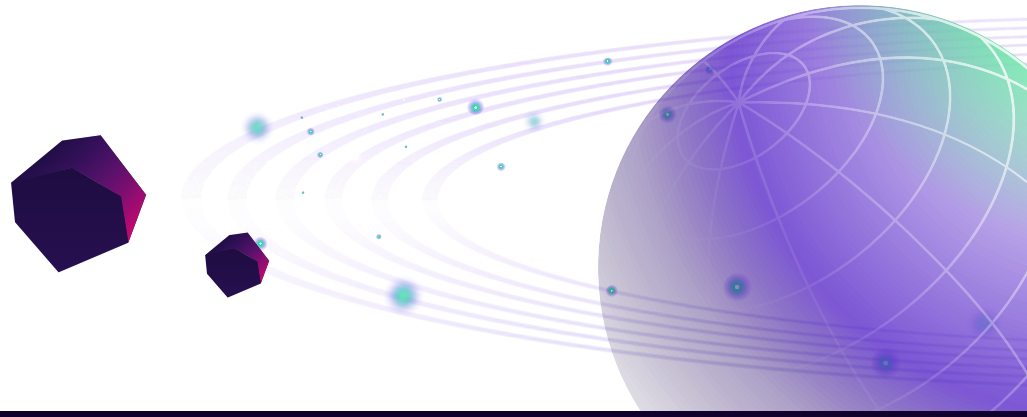
Cyera's AI Security Posture Management (AISPM) detects and classifies AI tools deployed across SaaS, IaaS, PaaS, and end-user systems—including unsanctioned browser extensions or Shadow AI usage—helping you determine whether AI systems fall within the high-, limited-, or minimal-risk categories under the EU AI Act. Meanwhile, Cyera's AI Runtime Protection redacts or blocks unauthorized use of personal or sensitive categories in prompts, logs, or AI outputs.

Having this kind of visibility and control greatly simplifies the task of ensuring adherence to those core values of transparency and purpose limitation. Cyera can be configured to alert when data migrates out of dedicated environments or network segments. It can also detect and alert on access misconfigurations that may result in personal data being improperly used or disclosed. And it can detect encryption gaps and orchestrate remediation.

When it comes to minimization, Cyera is absolutely essential. Already many Cyera customers are saving tens of thousands of dollars every month from reduced data storage costs. Organizations that deploy Cyera's Data Security Posture Management (DSPM) solution are discovering petabytes of data they didn't realize they had, and which can be safely deleted either because they are redundant, obsolete, or trivial (ROT), or otherwise violate their data retention policies. This doesn't just cut costs, it also reduces the organization's attack surface and supports compliance with GDPR and other regulations.

The bottom line is that without Cyera, many organizations simply don't know what data they have, so they can't be sure it isn't finding its way into environments or applications that would violate data subjects' privacy rights. For the same reason, they also can't confidently deploy AI systems, since they can't guarantee those systems aren't ingesting personal information from data stores they're unaware of, or generating sensitive personal data that isn't being properly classified and protected.

But with Cyera, organizations will have done much of the work to prepare for the AI Act's implementation, allowing them to hit the ground running as each section of the Act comes into effect.



How Cyera Supports Providers and Deployers of High Risk AI Systems

Chapter III of the AI Act sets out specific requirements for high risk systems and their providers and deployers. Some of these requirements involve high-level determinations of the risks inherent in a given AI system and the organization's plans to mitigate those risks. Others relate to specific types of documentation that must be maintained and used for completing the conformity assessment for the system. Still others set out the manner in which providers and deployers will interface with the appropriate authorities to ensure ongoing compliance.

While those strategic plans and technical details generally fall outside of Cyera's purview, Cyera can directly support compliance with many important requirements of Chapter III. This section will address how Cyera assists providers and deployers of high risk AI systems with respect to risk management, data governance, cybersecurity, quality management, post-market monitoring, and incident response.

Risk Management

Article 9 requires the establishment of a risk management system for all high risk AI systems. The purpose of the risk management system is to identify, evaluate, and mitigate any reasonably foreseeable risks to health, safety, or fundamental rights that could arise out of the AI system, even when used in accordance with its intended purpose.

The Act envisions an iterative process by which appropriate risk mitigation strategies and techniques are implemented, tested, and refined.

Cyera assists organizations in assessing the overall risks to their data estates, including risks to training data sets.

Cyera's Data Risk Assessment service provides security officials with a virtual, CISO-led evaluation of the organization's data security posture relative to more than 30 controls from frameworks such as ISO 27001 and NIST CSF. The service provides actionable intelligence that can help organizations immediately shrink their attack surface, gain greater visibility into potential threats, and develop a plan for improving their security posture going forward, including timelines and milestones.

Cyera's AI Risk Assessment service offers a similar evaluation with respect to AI-specific frameworks such as the NIST AI Risk Management Framework and ISO 42001. The AI Risk Assessment and Breach Readiness services now include AI-specific tabletop exercises, helping organizations assess data-related vulnerabilities in AI training, deployment, and output monitoring.

Data Governance

Article 10 requires developers of high risk AI systems to implement data governance and management policies appropriate to the intended uses of those systems.

Training, validation, and testing data sets must be representative, relevant, accurate, and complete.

Providers of high risk AI systems may process special categories of personal data, but only to the extent necessary to identify and mitigate harmful biases in the operation of their AI systems.

Cyera's agentless DSPM identifies and classifies your organization's data across the cloud and on-prem. This gives you visibility into the location, type, and sensitivity of data, as well as who is using the data and for what purposes, allowing organizations to quickly identify and confirm the segregation of training, validation, and testing data sets, and making it easier to audit those data sets to ensure they meet the AI Act's requirements for representativeness, relevance, accuracy, and completeness.

Cyera can classify data according to pre-set identifiers aligned with regulatory frameworks like the GDPR, as well as custom rules crafted by the organization. Its AI-native design even allows it to independently learn data categories unique to each organization.

Cyera's AI SPM identifies AI tools and agents in use in your IT environment, and AI Runtime Protection can prevent sensitive data leakage by monitoring LLM prompts and outputs.

Cybersecurity

Article 15 requires providers and deployers of high risk AI systems to implement security controls to prevent unauthorized parties from altering the use, outputs, or performance of those systems.

Security controls must be able to prevent, detect, and respond to attempts to compromise training data sets or components of the AI model, and prevent attacks (like prompt injection) that cause the model to make a mistake.

Cyera's DSPM, Identity Access, Omni DLP, and AI Guardian products directly assist in maintaining secure access controls.

Cyera's AI-native DSPM discovers organizational data across cloud and on-prem resources, and classifies them by type and sensitivity.

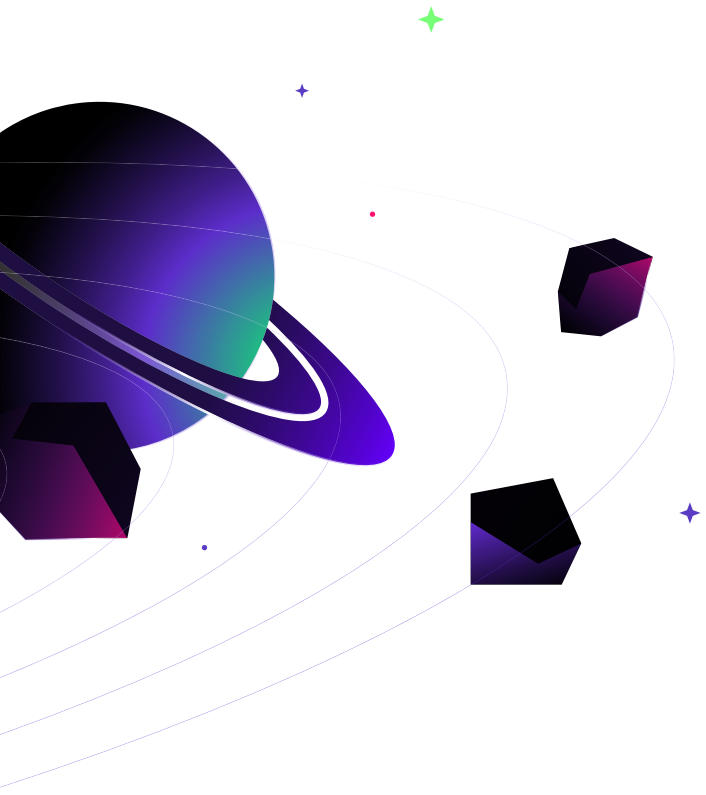
Cyera's Identity Access catalogues entities that have access to your organization's data, whether human or non-human, internal or external, and can even discover stale identities that still have access to sensitive data, allowing administrators to quickly determine which users' access exceeds the principle of least privilege, and take steps to remediate any misconfigurations found.

Cyera's Omni DLP protects sensitive data in transit to ensure compliance with regulations like the GDPR.

Cyera's AI Runtime Protection actively monitors AI usage across tools, detecting prompt injection, output leakage, or unapproved inference. It blocks unsafe prompts and redacts regulated outputs in real-time, preventing AI misuse and supporting Article 15's requirement to protect AI system integrity.

Cyera logs data events, emits data activity telemetry, and integrates with SIEM tools for centralized logging, alerting, and forensics

Cyera can be configured to alert on the detection of anomalous or suspicious activity such as the sharing of large numbers of sensitive records. It also integrates with various SIEM tools and can support automated remediation of access misconfigurations.



Quality Management

Article 17 requires providers of high risk AI systems to implement a quality management system that lays out the provider's strategy for complying with the AI Act. The operation of the quality management system includes documenting the risk management system discussed above, as well as the post-market monitoring and incident response plans discussed below.

Furthermore, it requires providers to implement "systems and procedures for data management, including data acquisition, data collection, data analysis, data labelling, data storage, data filtration, data mining, data aggregation, data retention and any other operation regarding the data that is performed before and for the purpose of the placing on the market or the putting into service of high-risk AI systems."

Cyera discovers and classifies data according to identifiers used in dozens of common regulatory frameworks like the GDPR, that come out of the box as part of the solution, as well as customized categories created by the organization. It even learns new categories for types of data based on patterns of usage in the organization.

Organizations that deploy Cyera in their environments will find it much easier to manage and label their data, and comply with internal policies or external regulations regarding data collection, storage, and retention.

Cyera also provides broad visibility into AI related data flows and policy enforcement in supported channels. It supports automated remediation workflows for AI policy violations and captures relevant events that support lifecycle traceability and auditability.

Post-Market Monitoring

Article 72 requires providers of high risk systems to implement a post-market monitoring system that "shall actively and systematically collect, document and analyse relevant data which may be provided by deployers or which may be collected through other sources on the performance of high-risk AI systems throughout their lifetime, and which allow the provider to evaluate the continuous compliance of AI systems with" the other requirements of the Act.

Cyera captures AI related events such as prompt history, user role, content type, and remediation actions. These logs feed directly into AI-specific MCP dashboards and SIEM tools for long-term risk monitoring.

Incident Response

Article 73 requires providers of high risk systems to report any serious incidents arising out of the operations of their AI systems to the relevant authorities of the member states where the incidents occur.

The Cyera platform generates alerts on policy violations, and can classify them by criticality. It also integrates with SIEM tools to support automated remediation workflows.

Moreover, Cyera's AI Runtime Protection enables real-time alerts on AI-related data exposure or behavioral anomalies, ensuring fast detection and mitigation.

Cyera further supports incident response by helping to determine the data blast radius from a security incident.

Event and audit logs generated by Cyera can be used to support the documentation and reporting requirements of the AI Act.

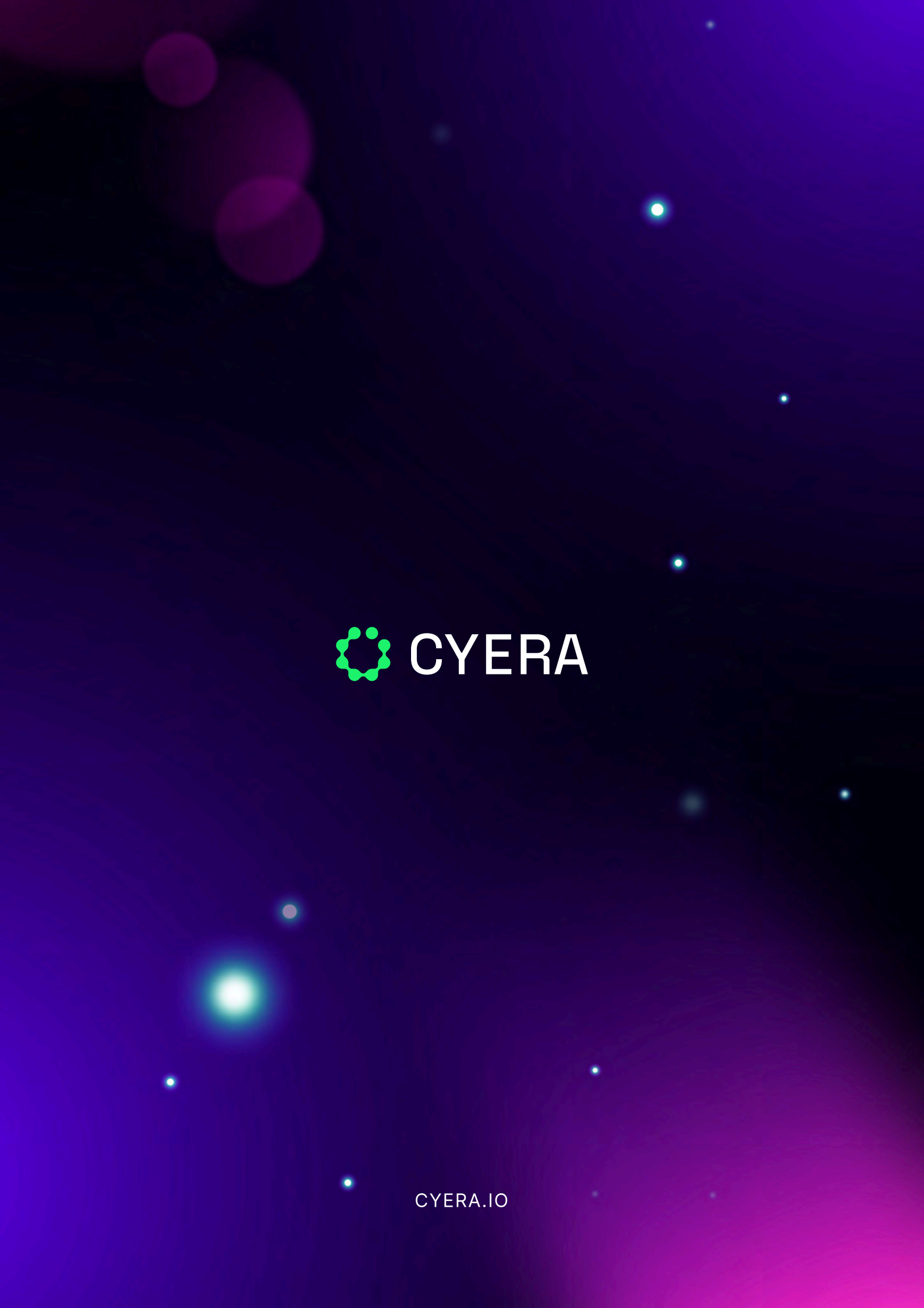
General Purpose AI Models

The Act defines a general purpose AI model (GPAI) as any AI model 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." Examples of GPAI models include large language models (LLMs) like ChatGPT, or AI image generators like Dall-E and Midjourney.

GPAI models and systems based on those models may or may not be incorporated into other high risk AI systems, but either way the providers of GPAI models are required to provide some additional documentation. They must make publicly available a sufficiently detailed summary of the data used to train their models. They must also put in place a policy for ensuring compliance with EU copyright laws.

Cyera can assist providers of GPAI models by identifying and classifying training datasets across connected sources. But a key differentiator for Cyera is its deep contextual awareness, making it possible to scour unstructured data sets and identify protected IP - including copyrighted works - with speed, scale, and precision.

To learn more about how Cyera can help comply with EU AI Act mandates visit www.cyera.com



CYERA.IO