THALES
Building a future we can all trust

AI ACT

El ~~futuro~~ presente de la regulación de la IA en Europa

Enrique Morey
Cybersecurity GRC Spain Manager

Cyber Solutions by Thales

DEFINICIÓN DE IA SEGÚN LA UE

«Sistema de IA» significa un sistema basado en una máquina que está diseñado para funcionar con diferentes niveles de autonomía y que puede presentar capacidad de adaptación tras su implementación y que, con fines explícitos o implícitos, deduce, a partir de las entradas que recibe, cómo generar salidas tales como predicciones, contenidos, recomendaciones o decisiones que pueden influir en entornos físicos o virtuales.

PLAN DE IMPLANTACIÓN DEL AI ACT

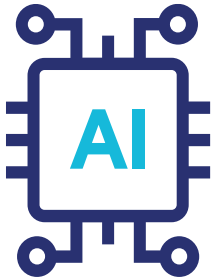
VARIABLES CLAVE EN EL AI ACT

> Qué rol desempeñamos

> Qué nivel de riesgo supone el uso de la IA en nuestro rol

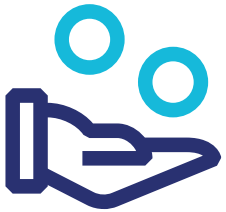
> Qué controles aplican en cada caso

ROLES DE LAS PARTES INTERESADAS



Proveedor

Desarrollan la IA y se la venden a terceros.



Distribuidor

Compran sistemas de IA a un proveedor o importador y las venden a terceros.

Importador

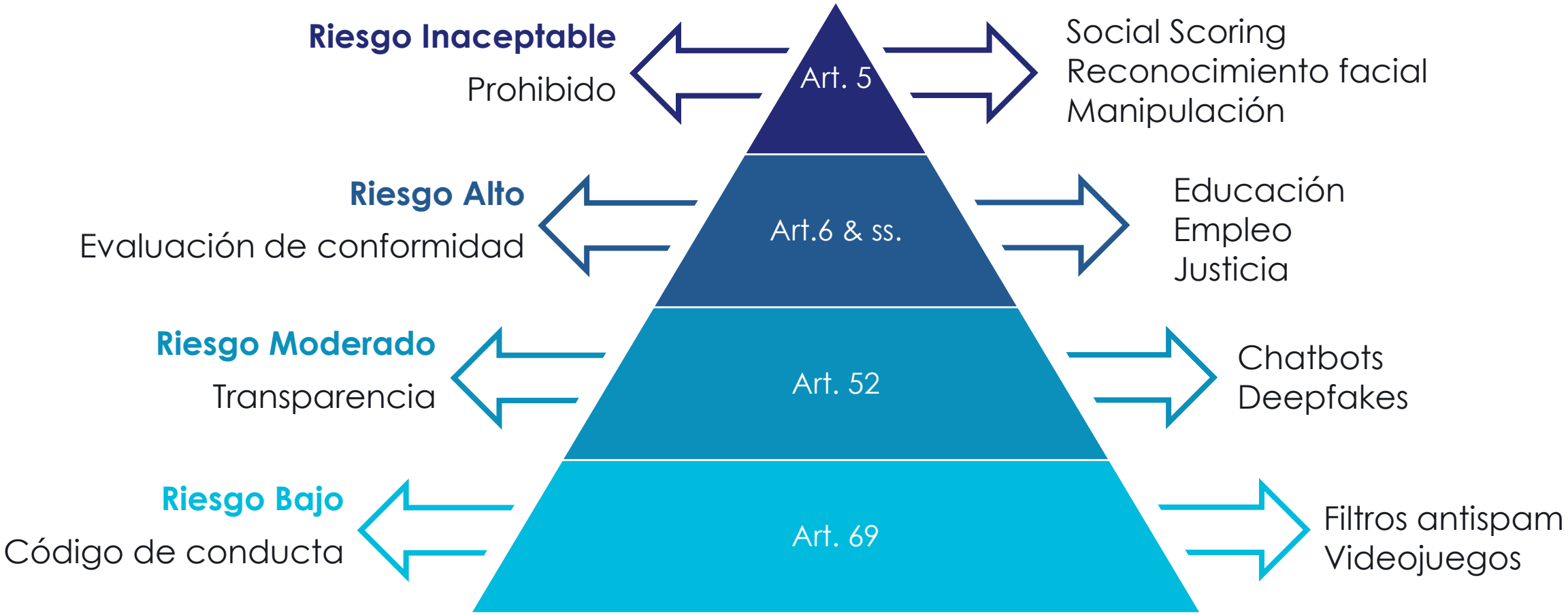
Compran sistemas de IA fuera de la UE y los venden dentro de la UE.

Implementador

Usan sistemas IA y los despliegan dentro de la UE.

Aplica a las partes que operan dentro de la UE.

CATEGORIZACIÓN DE IA BASADA EN RIESGO



EL ENFOQUE BASADO EN EL RIESGO DEL AI ACT

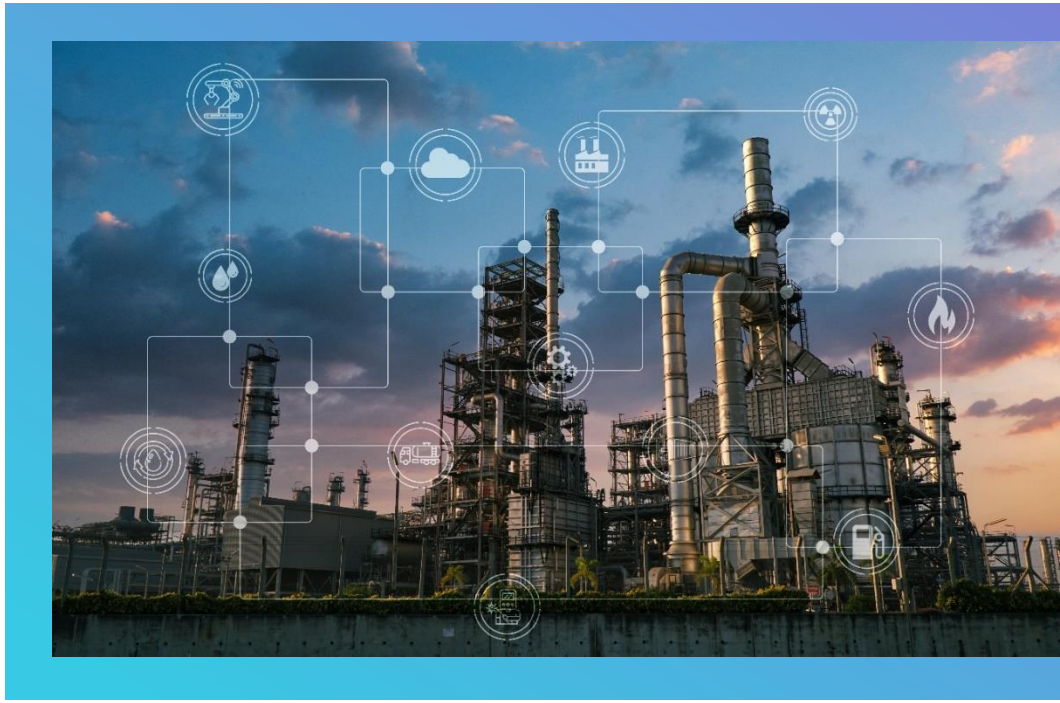
Riesgo Inaceptable

- Sistemas que asignan puntuaciones de crédito social basadas en comportamientos o afiliaciones políticas.
- Identificación biométrica remota en tiempo real (RBI) para la aplicación de la ley en espacios públicos.
- Uso de técnicas subliminales o engañosas para distorsionar comportamientos y perjudicar la toma de decisiones.
- Explotación de comunidades vulnerables
- Uso de la biometría para deducir atributos sensibles (raza, opiniones políticas, etc.).
- Perfiles criminalizantes basados en rasgos de personalidad.
- Inferencia de emociones en lugares de trabajo o instituciones educativas.

EL ENFOQUE BASADO EN EL RIESGO DEL AI ACT

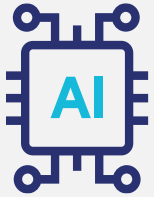
Riesgo Alto

- Infraestructuras críticas que pueden poner en peligro la vida (energía, transporte y agua).
- Formación que puede determinar el acceso a la educación (Educación)
- Componentes de seguridad de productos
- Gestión de trabajadores
- Servicios públicos y privados esenciales (servicios bancarios y evaluación crediticia)
- Aplicación de la ley (vigilancia y procesos de decisión automatizados)
- Gestión de la migración, el asilo y el control de fronteras
- Justicia y procesos democráticos



EL ENFOQUE BASADO EN EL RIESGO DEL AI ACT

Riesgo Alto - Responsabilidades



Proveedores

- Gestión de riesgos
- Gobernanza de datos
- Documentación del cumplimiento normativo
- Registros
- Instrucciones
- Supervisión humana
- Seguridad y precisión
- Sistema de gestión de calidad

Implementadores

- Supervisión humana
- Seguir las instrucciones
- Monitorización y reporte

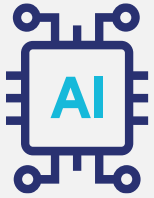
EL ENFOQUE BASADO EN EL RIESGO DEL AI ACT

Riesgo moderado

- Herramientas de diagnóstico: sistemas de IA que ayudan en el diagnóstico médico, pero que no son decisivos. Por ejemplo, rayos X o resonancias magnéticas.
- Sistemas de reclutamiento automatizados
- Aplicaciones de análisis de sentimientos: sistemas que analizan la opinión pública en redes sociales o comentarios sobre productos.
- Sistemas de crédito y puntuación: algoritmos que analizan datos financieros para proporcionar sugerencias de crédito, pero sin ser la única base para las decisiones de concesión de crédito.
- Asistencia en la toma de decisiones: herramientas que componen análisis o informes para ayudar en la toma de decisiones en entornos como las finanzas o las operaciones.

EL ENFOQUE BASADO EN EL RIESGO DEL AI ACT

Riesgo Moderado - Responsabilidades



Proveedores

- Seguridad y precisión
- Código de conducta
- Transparencia y responsabilidad
- Indicar a los usuarios sobre interacciones con la IA

Implementadores

- Indicar a los usuarios sobre interacciones con la IA
- Seguir las instrucciones

Importadores y Distribuidores

- Asegurar la transparencia
- Reportar a autoridades relevantes

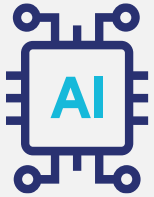
EL ENFOQUE BASADO EN EL RIESGO DEL AI ACT

Riesgo Bajo

- Filtros de spam
- Sistemas de recomendación de productos
- Chatbots
- Filtros sencillos de imagen o vídeo
- Juegos sencillos
- Algoritmos de detección de spam
- Herramientas educativas con personalización limitada
- Herramientas de mejora de imagen

EL ENFOQUE BASADO EN EL RIESGO DEL AI ACT

Riesgo Bajo - Responsabilidades



Proveedores

- Documentación
- Compartir información
- Política de Copyright
- Transparencia en el Entrenamiento del dato

Implementadores, Importadores y Distribuidores

Asegurarse de:

- Funcionalidad y seguridad
- Tener la documentación e instrucciones
- Uso ético y responsable
- Reportar a las autoridades fallos de seguridad

¿CÓMO PUEDEN PREPARARSE LAS EMPRESAS PARA EL AI ACT?

ANÁLISIS DE NECESIDADES

¿Cómo se está utilizando actualmente la IA en la organización? ¿Qué cualificación tiene el equipo? ¿Qué tipo de formación es necesaria?

PLAN DE FORMACIÓN

El AI ACT afirma que las empresas necesitan competencia en IA para utilizar los sistemas de IA de forma segura y eficaz. La formación debe reflejarlo.

NOMBRAR A UN ESPECIALISTA EN IA

No es obligatorio, pero es recomendable contar con alguien que tenga conocimientos detallados sobre sistemas de IA y cierta experiencia en materia legal.

CONOCER LOS STANDARDS DE LA IA

Las organizaciones deben conocer la norma técnica ISO/IEC 42001. Las empresas pueden utilizarla para crear sistemas de gestión de la IA que se ajusten a los nuevos requisitos reglamentarios.

Los sectores de riesgo moderado y alto deben mejorar desde ya su postura en materia de ciberseguridad

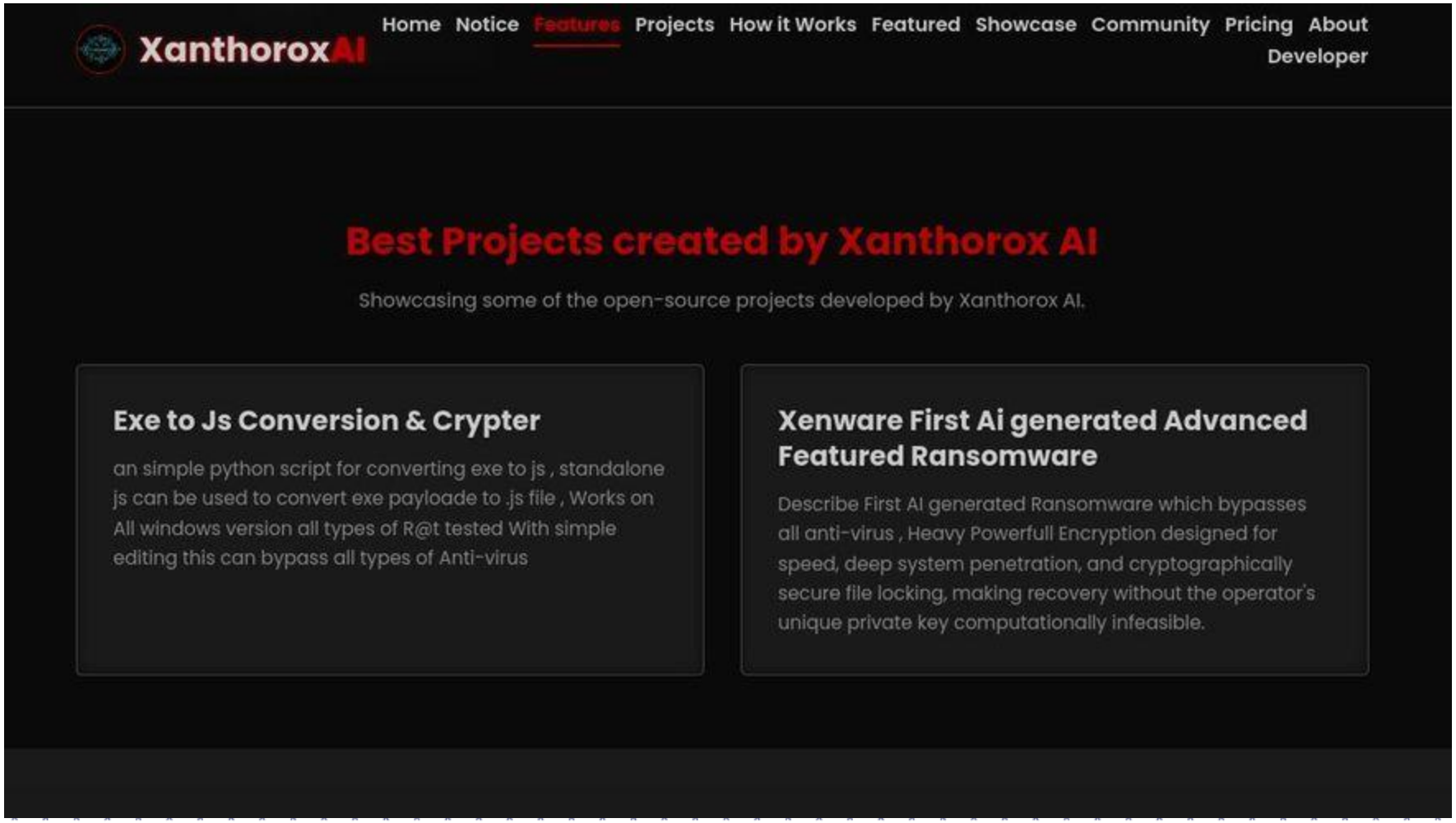
AI ACT- MULTAS

Cualquier empresa activa en IA debe respetar las nuevas normas

- **Lanzamiento de sistemas de IA prohibidos. Las multas pueden alcanzar el 7 % de la facturación global anual o 35 millones de euros (lo que sea mayor).**
- **Violaciones de IA de alto riesgo y GPAI. 3 % de la facturación global anual o 15 millones de euros.**
- **Suministro de información incorrecta. El 1 % de la facturación global anual o 7,5 millones de euros.**

LA IA EN EL MUNDO DE LA CIBERSEGURIDAD

LOS CIBERCRIMINALES NO NECESITAN CUMPLIR CON EL AI ACT



LA IA EN EL PANORAMA ACTUAL DE LA CIBERSEGURIDAD

AI EN LOS CIBERATAQUES

Phishing automatizado y personalizado
La IA se utiliza para crear correos electrónicos de phishing muy convincentes, con un lenguaje adaptado al perfil de la víctima.

Malware autoadaptable
El malware con IA puede modificar su comportamiento para evitar ser detectado por los antivirus tradicionales.

Ataques Basados en Deepfakes
Los deepfakes de voz y vídeo se utilizan para engañar a los sistemas de verificación o manipular a las personas en ataques de ingeniería social.

IA EN CIBERSEGURIDAD

Detección de amenazas en tiempo real
El aprendizaje automático supervisado y no supervisado ayuda a detectar malware, ransomware y actividades sospechosas.

Respuesta automatizada a incidentes
Las plataformas SOAR utilizan IA para automatizar las respuestas a incidentes, reduciendo el tiempo de reacción.

Análisis predictivo
La IA anticipa posibles vectores de ataque basándose en comportamientos históricos y en threat intelligence.

Autenticación inteligente
Los sistemas de autenticación biométrica y conductual utilizan la IA para reforzar la seguridad.

Y PARA FINALIZAR, SI SOY PROVEEDOR Y/O IMPLEMENTADOR, ¿QUÉ HAGO?

Todos los proveedores GPAI

Documentación técnica
Redactarla, incluido el proceso de formación y ensayo, y los resultados de la evaluación.

Documentación para proveedores posteriores
Elaborar información y documentación para integrar el modelo GPAI en su propio sistema de IA, para comprender capacidades y limitaciones.

Derechos de autor
Establecer una política de respeto de la Directiva sobre derechos de autor.

Transparencia en el entrenamiento
Publicar un resumen suficientemente detallado sobre el contenido utilizado para el entrenamiento del modelo GPAI.

Proveedores GPAI con riesgo sistémico

+ **Evaluar el riesgo de su sistema**
Realizar evaluaciones de modelos, incluida la realización de pruebas contradictorias para identificar y mitigar el riesgo sistémico.

+ **Mitigar los riesgos sistémicos**
Evaluar y mitigar los posibles riesgos sistémicos, incluidas sus fuentes.

+ **Comunicación a autoridades**
Rastrear, documentar y notificar los incidentes graves y las posibles medidas correctivas a la Oficina de AI y a las autoridades competentes.

+ **Medidas y controles**
Garantizar un nivel adecuado de protección de la ciberseguridad, y estar preparadas para auditorías e inspecciones.

GRACIAS

cds.thalesgroup.com