Presentado el código de buenas prácticas con respecto a la inteligencia artificial

Enlace al documento

01/09/2025

La Comisión Europea ya ha recibido la versión final del Código de buenas prácticas para la inteligencia artificial (IA) de uso general (es decir, aquellos modelos de IA[1] que pueden aplicarse a una amplia gama de propósitos, sin estar limitados a uno concreto) el cual es aplicable desde el pasado 2 de agosto.

Este documento, cuya adhesión es voluntaria, fue elaborado por trece expertos contando con las aportaciones de más de 1.000 entidades que dieron respuesta a una consulta pública, incluyendo empresas, académicos, especialistas en seguridad, organizaciones civiles, etc. Por el momento se han adherido al Código de Conducta 26 proveedores, entre los cuales se encuentran empresas como Amazon, Google, IBM, Microsoft o OpenAI, además de alguna entidad europea de investigación.

El objetivo de este Código es ayudar a los proveedores[2] de modelos de IA de uso general a prepararse para cumplir con las obligaciones impuestas por el Reglamento de IA de la UE (‘AI Act’, en inglés) , de aplicación desde el 2 de agosto de 2026.

Aquellos proveedores que opten por no adherirse podrán, como ha informado la Comisión Europea demostrar su cumplimiento del Reglamento de IA por otros

[1] Según IBM, un modelo de IA es “un programa que ha sido entrenado con un conjunto de datos para reconocer determinados patrones o tomar decisiones sin más intervención humana”. Son, en definitiva, el resultado de aplicar un conjunto de datos a un algoritmo -procedimiento de toma decisiones descrito en lenguaje matemático o pseudocódigo, que se aplica a un conjunto de datos para lograr una función o propósito determinados-, pudiendo llegar a optimizar su toma de decisiones a lo largo del tiempo si son capaces de llevar a cabo procesos de machine learning, sea de forma supervisada o no.

Con respecto a la definición, cabe destacar que el Reglamento de IA (‘AI Act’) diferencia en su Considerando nº. 97 **este concepto del de los ‘sistemas de IA’**. En concreto, destaca que “requieren que se les añadan otros componentes, como, por ejemplo, una interfaz de usuario, para convertirse en sistemas de IA” y que “los modelos de IA suelen estar integrados en los sistemas de IA”. A este respecto, el Servicio de Estudios del Parlamento Europeo (‘EPRS’, por us siglas en inglés) ha aclarado recientemente que ‘sistemas de IA’ son sistemas basados en máquinas (‘machine-based’, en inglés) capaces de inferir resultados a partir de las entradas de los usuarios, mientras que los modelos de IA [de uso general] son modelos entrenados con una gran cantidad de datos, capaces de realizar una amplia gama de tareas

[2] El art. 3.3 del Reglamento de IA, define ‘proveedor’ [de sistemas y modelos de IA] como “una persona física o jurídica, autoridad pública, órgano u organismo que desarrolle un sistema de IA o un modelo de IA de uso general o para el que se desarrolle un sistema de IA o un modelo de IA de uso general y lo introduzca en el mercado o ponga en servicio el sistema de IA con su propio nombre o marca, previo pago o gratuitamente”

medios, por ejemplo informando a la Oficina de IA de la Unión Europea de las medidas que para ello hayan implementado para mitigar los riesgos inherentes a sus modelos de IA. Asimismo, advierte la Comisión Europea de que estos proveedores podrán ser sujetos a un mayor número de solicitudes de información por parte de la Oficina de IA, al no poder conocer de forma detallada en qué forma estos cumplen con los requerimientos impuestos por el Reglamento de IA, precisamente por no haberse estos adherido al Código de Conducta.

La Oficina de IA de la Unión Europea comenzará a hacer cumplir este Código en 2026 para nuevos modelos de IA, en 2027 para modelos ya existentes, garantizando seguridad y transparencia en el mercado europeo. La Oficina de IA, creada en el seno de la Comisión Europea como centro de conocimientos especializados en IA, con el objetivo de asegurar una IA segura y fiable para Europa.

Entre sus funciones se encuentran el apoyo al Reglamento de IA e impulso de su implementación, fomentar el desarrollo y uso de una IA fiable mediante políticas y acciones que aprovechen sus beneficios sociales y económicos, el fomento de la cooperación internacional en materia de IA y, por último, la cooperación con instituciones, expertos y partes interesadas.

En esta etapa el documento es una propuesta, por lo que una vez aprobado el texto por la Comisión Europea y los Estados miembros de la UE e informado al Parlamento Europeo, los proveedores que se adhieran voluntariamente al Código podrán demostrar el cumplimiento de la Ley con mayor facilidad, beneficiándose de menos trámites y mayor seguridad jurídica.

El Código de buenas prácticas se estructura en **tres capítulos**:

1) **Transparencia**,
2) **Copyright**: aplicables a todos los proveedores de IA de uso general,
3) **Seguridad y Protección**: dirigido a los desarrolladores de modelos de IA más avanzados que impliquen riesgos sistémicos).

1)Transparencia:

El Capítulo de Transparencia del Código ofrece un Modelo de Formulario de Documentación del Modelo de IA fácil de usar, que permite a los proveedores recopilar la información que deben proporcionar en un solo lugar, y así poder ofrecer un mayor nivel de transparencia en línea con lo dispuesto en el Capítulo IV del Reglamento de IA.

Este Capítulo recoge las siguientes **tres medidas** que las entidades firmantes se comprometen a implementar a fin de cumplir con sus obligaciones de

transparencia según el Artículo 53.1, puntos a) y b), y los Anexos XI y XII correspondientes del Reglamento de IA:

- **Recoger y actualizar la documentación relativa a los modelos de IA**: los proveedores firmantes se comprometen a recoger, al menos, la información referida en el Modelo de Formulario de Documentación ('Model Documentation Form', en inglés), diseñado para permitirles recopilar fácilmente la información requerida por las disposiciones mencionadas del Reglamento de IA en un único lugar. Estos deberán actualizar este documento para que este refleje los cambios relevantes que tengan lugar en su modelo de IA de uso general.
- **Proveer información relevante**: los proveedores firmantes ofrecerán en sus respectivas páginas web -o por otros medios apropiados en el caso de que no tengan- la información de contacto de la Oficina Europea de IA y de los proveedores de servicios derivados de su modelo de IA de uso general, así como cualquier otra información necesaria. Asimismo, colaborarán con la Oficina Europea de IA ofreciéndole la información que esta -de oficio o a través de las autoridades nacionales competentes- les requiera en virtud de los artículos 91 o 75.3 del Reglamento de IA, la cual podrá o no estar comprendida dentro del ámbito del Modelo de Formulario de Documentación, en su versión más actualizada.

  Por último, los proveedores firmantes ofrecerán a los proveedores de servicios derivados de su modelo de IA la información del Modelo de Formulario de Documentación relativa a estos últimos. Asimismo, les ofrecerán información adicional en caso de que estos así lo necesiten para poder comprender las posibilidades y limitaciones del modelo de IA de uso general gestionado por los proveedores firmantes.
- **Asegurar la calidad, integridad y seguridad de la información**: los proveedores firmantes garantizarán que la información documentada sea controlada en cuanto a calidad e integridad, conservada como prueba del cumplimiento de las obligaciones del Reglamento de IA, y protegida de alteraciones involuntarias. En este contexto, se anima a los firmantes de a seguir los protocolos y normas técnicas establecidos para la preparación, actualización y control de la calidad y seguridad de la información y registros.

2) Copyright:

El **Capítulo de Copyright** del Código proporciona a los proveedores soluciones prácticas para establecer una política que cumpla con la legislación de derechos de autor de la Unión Europea, en cumplimiento con lo dispuesto en el art. 53.1.c) del Reglamento de IA. A este respecto, el Código de buenas

prácticas distingue **cuatro medidas** que los proveedores firmantes deberán cumplir en este Capítulo, siendo:

- **Desarrollar, mantener actualizada e implementar una política de copyright**: esta política deberá alinearse con la normativa de la UE en materia de copyright y derechos derivados. Los proveedores firmantes deberán describir en un solo documento dicha política, así como sus medidas.
- **Reproducir y extraer solamente contenidos protegidos por copyright accesibles legalmente al utilizar restreadores web**: los proveedores firmantes cuyos modelos de IA utilicen rastreadores web se asegurarán de que sus acciones no dejan sin efecto las medidas tecnologícas establecidas a fin de evitar accesos no autorizados a contenidos protegidos por derechos de copyright. Además, deberán excluir de entre las páginas web analizadas por dichos restreadores webs, aquellas que, siendo su actividad la publicación de contenidos, hayan sido condenadas en repetidas ocasiones por órganos jurisdiccionales o autoridades públicas en la Unión o en el Espacio Económico Europeo por vulnerar derechos de copyright y otros derechos relacionados.
- **Identificar y atenerse a las reservas de derechos al utilizar ratreadores web**: los proveedores firmantes se comprometen a respetar las reservas de derechos de los titulares al usar rastreadores web[3] para minería de texto[4] y datos[5] o entrenamiento de modelos de IA. Deben utilizar protocolos legibles por máquina como robots.txt[6] y otros estándares técnicos apropiados. Este compromiso no limita los derechos de los titulares de expresar reservas ni exime a los firmantes del cumplimiento de la ley de la UE sobre derechos de autor. También deben proporcionar información clara sobre sus prácticas de rastreo y permitir

[3] Según Cloudflare, un rastreador web es un tipo de bot que "accede, descarga o indexa contenido de todo Internet". Es una herramienta frecuentemente utilizada por motores de búsqueda como Google o Bing para encontrar páginas relevantes que puedan mostrar en los resultados de búsqueda.

[4] Según IBM, la minería de texto "es el proceso de transformar texto no estructurado en un formato estructurado para identificar patrones significativos y nuevos conocimientos", siendo posible utilizarla "para analizar vastas colecciones de materiales textuales para capturar conceptos clave, tendencias y relaciones ocultas".

[5] Según IBM, la minería de datos es "el uso del machine learning y el análisis estadístico para descubrir patrones y otra información valiosa a partir de grandes conjuntos de datos". Esta clase de procesos son mencionados -junto con la minería de texto- en el Considerando nº. 105 del Reglamento de IA ('AI Act') como algunos de lás técnicas a emplear en el proceso de desarrollo y entrenamiento de modelos de IA, si bien advierte de los problemas que podrían surgir en relación a aquellos textos y datos protegidos por derechos de copyright, en cuyo caso los proveedores deberían obtener la correspondiente licencia.

[6] Según el Centro de la Búsqueda de Google, los archivos.txt "sirven principalmente para gestionar el tráfico de los rastreadores a tu sitio, aunque también suelen usarse para que Google no rastree determinados archivos". En otras palabras, son una herramienta empleada por los desarroladores de páginas webs para evitar una sobrecarga de los servidores provocado por el acceso masivo rastreadores web a la página web o archivo en cuestión.

notificaciones automáticas sobre actualizaciones. Finalmente, se alienta a los proveedores de motores de búsqueda a garantizar que el cumplimiento de estas reservas no afecte negativamente la indexación de contenidos.

- **Mitigar el riesgo de resultados infractores de los derechos de copyright**: los proveedores firmantes deben tomar medidas técnicas y contractuales para evitar que sus modelos de IA generen contenidos que infrinjan los derechos de copyright y derechos afines. Esto incluye evitar la reproducción ilegal de contenido protegido y advertir o prohibir expresamente su uso indebido, incluso si el modelo se distribuye como software libre. Estas obligaciones aplican tanto si el modelo se usa internamente como si se cede a terceros.
- **Designar un punto de contacto y permitir la presentación de reclamaciones**: los proveedores firmantes deberán designar un punto de contacto a efectos de la comunicación electrónica con los titulares de derechos, así como ofrecer información accesible acerca del mismo. Además, deberán poner a disposición, de dichos titulares y sus representantes autorizados, mecanismos que los permitan presentar reclamaciones en relación a los incumplimiento de los proveedores de las obligaciones derivadas de este Capítulo de forma suficientemente precisa y adecuadamente sustanciada. Los proveedores firmantes gestionarán las reclamaciones dirigidas a ellos con diligencia, sin arbitrariedad y en un plazo de tiempo razonable.

3) Seguridad y protección:

El Capítulo de Seguridad y Protección, proporciona a los desarrolladores de modelos de IA más avanzados prácticas de vanguardia con los cuales mitigar los riesgos sistémicos que dichos modelos acarrean. Estos modelos están regulados en la Sección 3ª del Capítulo V del Reglamento de IA.

En concreto, para cumplir con las obligaciones derivadas del presente Capítulo, los proveedores firmantes deberán adoptar **diez compromisos**:

1º compromiso: Marco de Seguridad y Protección

- **Crear un Marco de Seguridad y Protección**: los proveedores firmantes deben elaborar un Marco de Seguridad y Protección (en adelante, 'el Marco') que describa cómo evalúan y mitigan los riesgos sistémicos asociados con sus modelos de IA. Este documento incluirá puntos clave como los momentos de evaluación durante el ciclo de vida del modelo, criterios para aceptar o no un riesgo sistémico, niveles de riesgo, medidas de mitigación, estimaciones sobre futuros desarrollos, y cómo influyen actores externos en el proceso. También se establecerá cómo se

distribuyen las responsabilidades y cómo se actualizará este Marco. El Marco deberá estar confirmado antes de lanzar modelos al mercado y deberá ser actualizado por los proveedores si es necesario.

- **Implementar el Marco**: Los firmantes deben evaluar y mitigar de forma continua los riesgos sistémicos de sus modelos de IA a lo largo de todo su ciclo de vida. Esto incluye evaluaciones periódicas, monitoreo tras la comercialización e intervención ante incidentes graves. Además, deben llevar a cabo un proceso completo de evaluación de riesgos antes de lanzar un modelo al mercado o cuando cambien ciertas condiciones, incluyendo un análisis individualizado de cada uno de los riesgos sistémicos identificados. Si estos últimos se consideran inaceptables, se deberán implementar medidas de seguridad y mitigación de dicho riesgo, y repetir el proceso de evaluación. De todo este proceso deberá ser informada la Oficina Europea de IA.
- **Actualizar el Marco**: los proveedores firmantes revisarán el Marco toda vez que el nivel de protección que este pueda estar ofreciendo haya disminuido o pueda disminuir, el cual se llevará a cabo empleando criterios de, en primer lugar, adecuación del Marco a los riesgos sistémicos propios de su modelo de IA y, en segundo lugar, adherencia del proveedor al Marco en sus operaciones. Estos actualizarán el Marco en consonancia con las conclusiones extraídas de su revisión

2º compromiso: Identificación de Riesgos Sistémicos

Los proveedores firmantes identificarán los riesgos sistémicos que su modelo de IA implica para las partes interesadas, así como el grado de aceptabilidad de dichos riesgos.

- **Proceso de identificación de riesgos sistémicos**: Los signatarios deben identificar los riesgos sistémicos asociados a un modelo de IA mediante un proceso estructurado que incluye, en primer lugar, compilar posibles riesgos basados en tipos predefinidos y datos relevantes (incluidos incidentes pasados); en segundo lugar, analizar la naturaleza y origen de esos riesgos; por último, determinar cuáles de ellos son sistémicos.
- **Escenarios de riesgos sistémicos**: los firmantes desarrollarán escenarios de riesgo sistémico apropiados, incluyendo el número y nivel de detalle de estos escenarios de riesgo sistémico, para cada riesgo identificado.

3º compromiso: Análisis de Riesgos Sistémicos

Los proveedores firmantes deberán analizar cada uno de los riesgos sistémicos detectados en su análisis, a fin de determinar si estos resultan o no aceptables.

El análisis de los riesgos sistémicos se dividirá en cuatro fases, las cuales podrán superponerse en ocasiones, siendo estas las siguientes:

- **Recopilación de información independiente del modelo de IA**: los proveedores firmantes reunirán información en relación con el riesgo sistémico -si bien independiente del modelo de IA en cuestión- proveniente de diferentes fuentes (p. ej. búsquedas web, análisis de mercado, revisiones de datos, predicción de tendencias futuras…).
- **Realización de evaluaciones de modelos de IA**: los proveedores firmantes llevarán a cabo evaluaciones punteras del modelo en las modalidades pertinentes para el riesgo sistémico a fin de evaluar las capacidades, propensiones, asequibilidades y/o efectos del modelo. Además, se asegurarán de que dichas evaluaciones del modelo se diseñen y realicen utilizando métodos que sean apropiados para el modelo y el riesgo sistémico en cuestión.
- **Delineación y delimitación del riesgo sistémico**: los proveedores firmantes llevarán a cabo la modelización del riesgo sistémico para el riesgo sistémico, utilizando para ello, los métodos más avanzados y la información recopilada en el primer paso.
- **Estimación y valoración del riesgo sistémico**: para la estimación del riesgo sistémico, los proveedores firmantes emplearán métodos conformes con el estado del arte y tendrán en cuenta la información recopilada en el marco de su análisis de riesgos sistémicos (2º Compromiso). El riesgo podrá ser expresado en diferentes formatos (p. ej. puntuación de riesgo, probabilidad) empleando indicadores cuantitativos o cualitativos.
- **Monitoreo posterior a la entrada en el mercado**: una vez el modelo de IA haya sido introducido en el mercado, los proveedores firmantes llevarán a cabo un análisis adicional evaluando el rendimiento del modelo de IA. Para ello, podrán emplear *feedback* proveniente de los usuarios, proveerles de canales de reporte de incidencias o establecer métodos de evaluación colectiva, entre otros. Además, los proveedores firmantes darán acceso gratuito a una versión actualizada del modelo de IA -tanto con las salvaguardas de seguridad introducidas como sin ellas- y sus *chains-of-thought* (método que consiste en entrenar a los modelos de IA para que proporcionen una explicación paso a paso del razonamiento subyacente a sus respuestas) a un número adecuado de evaluadores externos independientes.

4º compromiso: Determinación de la aceptabilidad de los riesgos sistémicos

Los proveedores firmantes especificarán criterios de evaluación de los riesgos sistémicos detectados. En base a los mismos, estos decidirán si continuar o no con el desarrollo, implantación en el mercado y/o con el uso del modelo de IA.

- **Criterios de aceptación del riesgo sistémico**: los proveedores firmantes deben definir y justificar cómo evaluarán si los riesgos sistémicos de un modelo de IA son aceptables. Para ello, establecerán niveles de riesgo medibles basados en capacidades del modelo u otros criterios apropiados; usarán estos niveles o criterios para decidir si los riesgos individuales y el riesgo global son aceptables; justificarán que estos métodos garantizan un nivel de riesgo aceptable; aplicarán un margen de seguridad que contemple incertidumbres, mejoras futuras del modelo, errores de evaluación y posibles fallos en las mitigaciones de seguridad.
- **Decisión de si proceder o no en base a la determinación de la aceptabilidad de los riesgos sistémicos**: los proveedores firmantes solo podrán desarrollar, comercializar o usar un modelo de IA si sus riesgos sistémicos se consideran aceptables. Si estos no lo son, o pronto podrían dejar de serlo, deben detener o limitar la disponibilidad de su modelo de IA, aplicar medidas que mitiguen los riesgos y, por último, repetir el proceso de evaluación de riesgos sistémicos y su aceptabilidad.

5º compromiso: Mitigaciones de seguridad

Los proveedores firmantes se comprometen a aplicar medidas apropiadas que mitiguen los riesgos sistémicos detectados en su análisis, de determinarse que estos no son aceptables.

En concreto, se destacan las siguientes **medidas de seguridad apropiadas**, las cuales deberán adoptar los proveedores de modelos de IA:

- filtrar y limpiar los datos de entrenamiento, por ejemplo, los datos que podrían dar lugar a propensiones indeseables del modelo, como trazas infieles de la cadena de pensamiento;
- supervisar y filtrar las entradas y/o salidas del modelo;
- cambiar el comportamiento del modelo en aras de la seguridad, por ejemplo, ajustando el modelo para rechazar determinadas solicitudes o proporcionar respuestas inútiles;
- organizar el acceso al modelo, por ejemplo, limitando el acceso a la API a usuarios verificados, ampliando gradualmente el acceso en función de la supervisión posterior a la comercialización, y/o no haciendo que los parámetros del modelo estén disponibles públicamente para su descarga inicialmente;
- ofrecer herramientas para que otros actores las utilicen con el fin de mitigar los riesgos sistémicos derivados del modelo;
- técnicas que proporcionen garantías cuantitativas de seguridad de alta fiabilidad en relación con el comportamiento del modelo;

- técnicas que permitan ecosistemas seguros de agentes de IA, como identificaciones de modelos, protocolos de comunicación especializados o herramientas de seguimiento de incidentes; y/o
- otras medidas de mitigación de la seguridad emergentes, como las destinadas a lograr la transparencia en el razonamiento de la cadena de pensamiento o a defenderse de la capacidad de un modelo para subvertir sus otras medidas de mitigación de la seguridad.

6º compromiso: Mitigaciones de protección

Los proveedores firmantes se comprometen a mantener medidas de ciberseguridad adecuadas para sus modelos de IA e infraestructura durante todo su ciclo de vida, a fin de garantizar que los riesgos sistémicos detectados se mantengan en niveles aceptables. Los modelos cuyas capacidades sean inferiores a las de al menos otro modelo disponible al público están exentos. Estas medidas de seguridad deben mantenerse hasta que los parámetros del modelo se hagan públicos o se eliminen de forma segura.

A estos efectos, los proveedores de IA definirán un **objetivo de protección** que especifique las amenazas que sus medidas de mitigación enfrentan.

Para alcanzar dicho objetivo, además, implementarán **medidas de mitigación apropiadas** para alcanzar el objetivo de seguridad. Si estos se desvían de las medidas de seguridad establecidas, deberán establecer medidas de mitigación alternativas que contribuyan a alcanzar el mismo objetivo.

7º compromiso: Informes del Modelo de Seguridad y Protección

Los proveedores firmantes se comprometen a presentar un Informe de Seguridad y Protección del Modelo de IA (en adelante, 'el informe') a la Oficina de IA antes de lanzarlo al mercado, detallando el modelo y sus procesos y medidas de evaluación y mitigación de riesgos sistémicos. También se comprometen a mantener el informe actualizado y a notificar dichas actualizaciones a la Oficina de IA.

Si ya se ha proporcionado información relevante en otros informes anteriores, puede hacerse referencia a ellos en el Informe del Modelo. Un único informe puede cubrir varios modelos si sus evaluaciones de riesgo están interrelacionadas. Las pymes y pequeñas empresas tecnológicas pueden presentar informes menos detallados para ajustarse a sus limitaciones de recursos.

Los proveedores firmantes incluirán en el informe una **descripción de su modelo de IA y de su conducta** incluyendo:

- Una descripción general de la arquitectura, capacidades, propensiones y funcionalidades del modelo de IA, cómo se desarrolló, su método de entrenamiento y datos, y en qué se diferencia de otros modelos que han puesto en el mercado.
- Una descripción del uso actual y esperado del modelo, incluyendo su uso en el desarrollo, supervisión y evaluación de otros modelos.
- Una descripción de las versiones del modelo que están o estarán disponibles en el mercado y/o en uso, señalando diferencias en las mitigaciones y riesgos sistémicos.
- Una especificación de cómo se espera que opere el modelo de IA, detallando los principios que debe seguir, la forma en la que prioriza principios e instrucciones, temas sobre los que debe rechazar instrucciones y el *prompt* del modelo de IA.

- **Justificación de razones para proceder**: los proveedores firmantes justificarán en el informe la aceptabilidad de los riesgos sistémicos detectados, los supuestos hipotéticos en los que estos dejarían de serlo y una descripción acerca de cómo la decisión de proceder con el desarrollo, introducción en el mercado y/o uso fue tomada, y si la opinión al respecto de los asesores externos fue tomada en cuenta.
- **Descripción de los resultados de la identificación y análisis de riesgos sistémicos**: los proveedores firmantes deben incluir en el Informe del Modelo:
  - El proceso de identificación de riesgos según tipos específicos,
  - Incertidumbres y suposiciones sobre el uso e integración del modelo,
  - Resultados del modelado de riesgos sistémicos,
  - Justificación y estimaciones de riesgos, comparando escenarios con y sin mitigaciones,
  - Resultados de evaluaciones del modelo, métodos usados, tareas, puntuaciones y comparación con estándares humanos,
  - Muestras aleatorias de entradas y salidas para facilitar la interpretación independiente,
  - Detalles sobre el acceso y recursos para evaluadores internos y externos,
  - Justificación si se utiliza el concepto de "modelo igual o más seguro",
  - Descripción de todas las mitigaciones de seguridad implementadas, cómo cumplen los requisitos y sus limitaciones,
  - Descripción del objetivo de seguridad, las mitigaciones implementadas, su alineación con estándares internacionales, y justificación de cualquier desviación,
  - Descripción general de las técnicas y recursos que se usarán para el desarrollo futuro del modelo en los próximos seis meses,

diferencias esperadas en capacidades y propensiones, y nuevas o actualizadas mitigaciones de seguridad que se planean implementar.

- **Informes externos**: los proveedores firmantes incluirán asimismo en el informe las evaluaciones realizadas por expertos externos independientes en relación al modelo de IA, respetando la confidencialidad y permitiendo en todo momento a dichos expertos decidir acerca de a publicación de sus hallazgos. En caso de no haber mediado dicha evaluación externa,
- **Cambios sustanciales en los riesgos sistémicos**: los proveedores firmantes incluirán en el informe información acerca de cómo el desarrollo, introducción en el mercado y/o uso del modelo de IA resultan en cambios sustanciales en el escenario de riesgos sistémicos detectados.
- **Actualizaciones del informe**: el informe será actualizado a medida que ellos sea necesario para reflejar los cambios en el grado y justificación de la aceptabilidad de los riesgos sistémicos detectados. Los proveedores firmantes actualizarán sus respectivos informes en un tiempo razonable después de haber tenido conocimiento de las circunstancias que motivan dichas actualizaciones.
- **Notificación del informe**: los proveedores firmantes darán acceso a su informe a la Oficina Europea de IA al tiempo de introducir en el mercado un modelo de IA de IA. Si el modelo de IA es un modelo actualizado, los proveedores darán a la Oficina Europea de IA acceso a un nuevo informe actualizado en consonancia.

8º compromiso: Asignación de responsabilidades por riesgos sistémicos

Los proveedores firmantes establecerán de manera clara los responsables de la gestión de los riesgos sistémicos a través de los diferentes niveles organizativos, y asignarán a dichos responsables recursos suficientes para llevar a cabo sus funciones.

Los proveedores de modelos de IA **definirán de forma clara** las responsabilidades de gestión de riesgos sistémicos derivados de sus modelos en todos los niveles de la organización. Estas incluyen:

- Supervisión de riesgos sistémicos: Supervisar los procesos y medidas de evaluación y mitigación.
- Responsabilidad por riesgos sistémicos: Gestionar los riesgos y responder a incidentes graves.
- Apoyo y monitoreo: Apoyar y hacer seguimiento a las medidas de mitigación y evaluación.
- Garantía ante riesgos sistémicos: Proveer aseguramiento interno (y externo, si corresponde) sobre la adecuación de los procesos al órgano de supervisión o a otro ente independiente.

Estas responsabilidades deben asignarse según la estructura de gobernanza y complejidad organizacional, distribuyéndolas entre:

- El órgano de supervisión o un ente independiente (como un consejo o junta).
- La dirección ejecutiva.
- Los equipos operativos relevantes.
- Proveedores internos de garantía (como auditoría interna), si los hay.
- Proveedores externos de garantía (como auditores externos), si los hay.

- **Asignación de recursos apropiados**: los proveedores firmantes asegurarán que aquellas personas sobre las cuales recaigan las responsabilidades mencionadas dispongan de recursos humanos, financieros, de acceso a información y computacionales suficientes para llevar a cabo sus tareas.
- **Promoción de una cultura saludable del riesgo**: los proveedores firmantes propocionarán entre aquellas personas sobre las cuales recaigan las responsabilidades mencionadas una cultura saludable del riesgo que les permita tomar mayor conciencia de los riesgos sistémicos presentes en la actividad del proveedor y su modelo de IA. Ejemplos de esta cultura son implementarla desde los más altos niveles de dirección, permitir la comunicación clara y el desacuerdo en la toma de decisiones relacionadas con los riesgos sistémicos, establecer incentivos, conceder un grado de independencia suficiente a aquellas personas sobre las cuales recaigan las responsabilidades mencionadas, o la realización de encuestas sobre el grado de comodidad de los empleados al reportar riesgos sistémicos del modelo de IA a sus superiores.

9º compromiso: Reporte de incidentes graves

Los proveedores firmantes implementarán procesos apropiados y medidas para monitorizar, docmuentar y reportar a la Oficia Europea de IA todos aquellos incidentes que tengan lugar a lo largo del ciclo de vida de sus respectivos modelos de IA, así como aquellas medidas correctivas al respecto.

- **Métodos para la identificación de incidentes graves**: los proveedores firmantes deberán considerar los métodos propuestos en relación con el monitoreo postereior a la entrada en el mercado del modelo de IA (3º compromiso) para hacer seguimiento de información relevante sobre incidentes graves. Además, deberán revisar otras fuentes de información como informes policiales, medios de comunicación o redes sociales; y facilitar que reporten información relevante sobre incidentes graves, ya sea directamente al proveedor o a la Oficina Europea de IA. Además, deberá informar a terceros sobre los canales disponibles para hacerlo.

- **Información relevante para el seguimiento, documentación y reporte de incidentes graves**: los proveedores firmantes deben registrar, documentar y reportar a la Oficina Europea de IA y, cuando corresponda, a las autoridades nacionales competentes, la siguiente información sobre incidentes graves, en la medida de su conocimiento y cumpliendo con las leyes de la Unión Europea aplicables (p. ej. fechas de inicio y fin del incidente, daños causados, personas/grupos afectados, modelo de IA implicado, acciones de respuesta realizadas o previstas, etc.).
- **Plazos para informar de incidentes graves**: cunado al menos uno de sus modelos de IA esté implicado en un incidente grave, los proveedores firmantes presentarán a la Oficina Europea de IA y, si corresponde, a las autoridades nacionales competentes en la materia la mencionada información dentro de los siguientes plazos:
    - En los casos de interrupción grave e irreversible de infraestructura crítica: en un plazo de 2 días desde que tengan conocimiento del hecho.
    - En los casos de violación grave de ciberseguridad: en un plazo de 5 días.
    - En los casos de muerte de una persona: en un plazo de 10 días.
    - En los casos de daño grave a la salud, violación de derechos fundamentales o daño grave a bienes o al medio ambiente: en un plazo de 15 días.

  Si el incidente grave no se resuelve de inmediato, los signatarios deben actualizar la información en informes intermedios cada 4 semanas, hasta completarla. Una vez resuelto el incidente, deben presentar un informe final con toda la información exigida, dentro de 60 días. En caso de múltiples incidentes similares dentro del mismo periodo, estos pueden incluirse en el informe del primer incidente, siempre respetando los plazos de dicho primer caso.
- **Periodo de conservación**: los proveedores firmantes conservarán toda aquella documentación que evidencia el cumplimiento por su parte de este Compromiso durante al menos un plazo de 5 años desde la fecha de documentación del incident o la fecha en la que el incidente tuvo lugar.

10º compromiso: Documentación adicional y transparencia

Los proveedores firmantes se comprometen a documentar su cumplimiento con los compromisos contenidos en este Capítulo y a publicar versiones resumidas de su Marco y informes del modelo de IA cuando sea necesario.

- **Documentación adicional**: los proveedores firmantes elaborarán y actualizarán, a fin de poder ofrecérsela a la Oficina Europea de IA, documentación relativa a la descripción detallada de su modelo de IA, de cómo este está integrado dentro de los modelos de IA, de cómo sus

difeentes componentes se retroalimentan, de las evaluaciones del modelo hechas en cumplimiento del presente Capítulo y, por último, de las medidas de mitigación de riesgos sistémicos planteadas. Toda esta documentación será conservada durante un plazo mínimo de 10 años, a contar desde la introducción al mercado del modelo de IA.

- **Transparencia pública**: los proveedores firmantes deberán publicar, si es necesario para evaluar o mitigar riesgos sistémicos, una versión resumida de su Marco y de sus Informes del Modelo (incluidas actualizaciones), por ejemplo, en sus sitios web. Esta publicación debe excluir información que pueda comprometer la eficacia de las medidas de seguridad o que sea comercialmente sensible.En el caso de los Informes del Modelo, la publicación debe incluir descripciones generales de los resultados de la evaluación de riesgos sistémicos y de las medidas de seguridad aplicadas. No será necesaria la publicación del Marco ni del Informe del Modelo si todos los modelos del signatario son considerados "igualmente seguros o más seguros".