Breve análisis del texto final del Reglamento de Inteligencia Artificial

Desde la Federación de Consumidores y Usuarios CECU hemos llevado a cabo un primer análisis del texto final del Reglamento de Inteligencia Artificial (Reglamento de IA), que ha sido filtrado por la presidencia belga el pasado 26 de enero y que se ha sometido a votación del COREPER el 02 de febrero.

En el presente documento hemos recogido las demandas que desde CECU[1] hemos venido proclamando desde mediados de 2022 -para lograr una mejor protección de las personas consumidoras en el Reglamento de IA- y las hemos comparado con el texto acordado (valorando según lo pedido y lo finalmente recogido en la norma).

Criterios de valoración:

× **Negativo**

– **Ni positivo ni negativo**

✓ **Positivo**

1. Ampliación del ámbito de aplicación del Reglamento e imponerse una serie de principios y obligaciones horizontales (equidad, transparencia, rendición de cuentas, entre otros) que apliquen a todos los sistemas de IA:

× Definición de sistema de IA (art. 3.1):

Desde CECU requerimos una definición amplia de sistemas de IA que incluyera enfoques de aprendizaje automático, basados en la lógica, el conocimiento o estadísticas.

En la versión acordada, la definición ha sido limitada a "*(...) un sistema basado en una máquina diseñado para funcionar con diferentes niveles de autonomía y que pueden mostrar adaptabilidad después del despliegue y que, por razones explícitas u objetivos implícitos, infiere, a partir de los insumos que recibe, cómo generar resultados, tales como predicciones, contenidos, recomendaciones o decisiones que puedan influir física o entornos virtuales*".

Según se explica en el informe de la presidencia belga, la definición se habría modificado para alinearla más estrechamente el trabajo de la OCDE. Y aclara que "*la definición no pretende abarcar sistemas de software o enfoques de programación*

[1] Recomendaciones CECU para el Reglamento de IA (agosto 2022): https://cecu.es/publicaciones/inteligencia-artificialia-propuesta-de-regulacion-de-la-union-europea-ue/ ; Recomendaciones CECU para los trílogos del Reglamento de IA (julio 2023) https://cecu.es/wp-content/uploads/2023/07/Recomendaciones-CECU-trilogos-AI-Act-1.pdf

tradicionales más simples, que se basan en las reglas definidas únicamente por personas naturales para ejecutar operaciones automáticamente".

Al incluirse en la definición a sistemas que "p*ueden exhibir adaptabilidad después del despliegue*" se limita el alcance del Reglamento porque es una característica intrínseca a algunos, pero no a todos los sistemas de IA. Esto dejaría afuera de la regulación a los sistemas de IA basados en reglas (que siguen un algoritmo predeterminado sin la capacidad de autoaprender y adaptarse a partir de nuevas entradas de datos después del despliegue).

- Principios horizontales:

Al ser un Reglamento que adopta un enfoque basado en el riesgo (es decir, apunta a regular mayormente a aquellos sistemas de IA que se clasifiquen como de "alto riesgo"), desde CECU pedimos que exija una serie de principios horizontales a todos los sistemas de IA empleados en la UE, incluyendo los de riesgo medio y bajo (por ejemplo, transparencia, equidad, no discriminación).

En algún punto de las negociaciones el Parlamento Europeo logró introducir algunos principios horizontales –aunque de aplicación voluntaria- como la supervisión humana, la robustez técnica y la seguridad, la privacidad y la gobernanza de datos, la transparencia, etc. Estos principios fueron eliminados en el texto final del Reglamento.

- Alcance limitado:

No se han introducido reglas sustantivas para sistemas de IA distintos de los de alto riesgo, excepto el artículo 52 (obligaciones de transparencia para proveedores y usuarios de determinados sistemas de IA y modelos de IA de propósito general) y el artículo 69 (códigos de conducta). Esto significa que los asistentes virtuales o los juguetes integrados con IA no estarán regulados adecuadamente.

2. Clasificación clara y ampliación de la lista de sistemas de alto riesgo (Anexo III):

- Clasificación de sistemas de alto riesgo (art. 6):

Desde CECU hemos pedido que se establezca un claro sistema de clasificación de riesgos, sin lagunas ni discrecionalidad por parte de las empresas. De esto dependía la efectividad de la norma, ya que sobre los sistemas de alto riesgo recaen todas o la mayoría de las obligaciones que impone.

Sin embargo, el art. 6 refiere a que los sistemas enumerados en el Anexo III no se considerarán de alto riesgo si no suponen un riesgo significativo de daño, a la salud, a la seguridad o a los derechos fundamentales de las personas naturales, incluso sin influir materialmente en el resultado de la toma de decisiones. Y explica que tal será el caso si cumple uno o más de los siguientes criterios: (a) el sistema de IA está destinado a

realizar una tarea procesal limitada; (b) el sistema de IA está destinado a mejorar el resultado de una actividad humana previamente realizada; (c) el sistema de IA está destinado a detectar patrones de toma de decisiones o desviaciones de patrones anteriores de toma de decisiones y no pretende reemplazar ni influir en la evaluación humana realizada previamente, sin una revisión humana adecuada; o (d) el sistema de IA está destinado a realizar una tarea preparatoria para una evaluación relevante para la finalidad de los casos de uso enumerados en el anexo III.

El proveedor que considere que su sistema no es de alto riesgo debe documentar esta autoevaluación de manera previa a que el sistema sea puesto en el mercado o en el servicio. Y deberá registrarse en la base de datos del art. 60.

✓ Por su parte, los sistemas que realicen perfilado o *profiling* se considerarán de alto riesgo siempre.

Más allá de ello, desde CECU consideramos que el texto acordado presenta varias lagunas y fallas desde la perspectiva de protección de las personas consumidoras al permitir tal evaluación y no establecer grandes criterios de supervisión.

– Listado de sistemas de IA de alto riesgo del Anexo III

El Reglamento de IA se focaliza en gran medida en los sistemas que se consideran de "alto riesgo" (sobre los que establece obligaciones especiales como la gestión de riesgos, gobernanza de datos, documentación técnica, transparencia e información para el implementador, supervisión humana, etc.), dejando a muchos sistemas de IA prácticamente sin regulación, como los juguetes integrados con IA o los asistentes virtuales. En el texto acordado la lista de sistemas de alto riesgo se ha ampliado – aunque no lo suficiente- y se han incluido algunos pedidos del movimiento europeo de consumidores, como:

- Los sistemas de IA destinados a evaluar la solvencia de personas físicas o establecer su puntaje crediticio, excepto los sistemas de IA utilizados para detectar fraude financiero (Y se ha eliminado las excepciones que se querían introducir para las PYMES).
- Los sistemas de IA destinados a usarse para evaluar riesgos y fijar precios en relación con personas físicas en caso de seguros de vida y salud.
- Los sistemas de IA utilizados para el reconocimiento de emociones, que no estén prohibidos.

3. Ampliación de la lista de prácticas prohibidas (art. 5):

✗ Reconocimiento biométrico en espacios de acceso público:

Desde CECU se pedía la prohibición total del reconocimiento biométrico remoto por parte de autoridades públicas y entidades privadas en espacios de acceso público, tanto en tiempo real como ex post. Ello, por considerar podría dar lugar a la vigilancia

masiva biométrica generalizada, afectando el derecho a la privacidad de las personas consumidoras, entre otros.

Lamentablemente, en el texto acordado la prohibición se limitó al reconocimiento biométrico en espacios de acceso público en tiempo real por parte de autoridades públicas de aplicación de la ley y, además, con tres excepciones relacionados con delitos graves.

- Reconocimiento de emociones:

Desde CECU requeríamos que se establezca una amplia prohibición sobre el reconocimiento de emociones e incluya garantías adecuadas para cualquier excepción limitada que se pretenda introducir, porque suelen ser muy invasivos y presentan muchas inexactitudes.

En la versión final solo se prohibieron los sistemas de IA utilizados para inferir las emociones de una persona física en el lugar de trabajo y en las instituciones educativas. Todos los demás usos del reconocimiento de emociones son de alto riesgo. No resulta totalmente positivo porque los usos del reconocimiento de emociones por parte de las personas consumidoras no están prohibidos.

✓ Técnicas subliminales que causan distorsión y daño del comportamiento

Originalmente esta prohibición hacía referencia a sistemas de IA que tuvieran el "propósito" de causar un "daño físico o psicológico". Desde CECU habíamos requerido la eliminación de ese requisito de "intención" que sería muy difícil de darse en la práctica o de probar y que se incluyeran todo tipo de daños, considerando por ejemplo también el daño económico.

En el texto acordado esta prohibición se amplió al referirse al despliegue de técnicas subliminales más allá de la conciencia de una persona o técnicas deliberadamente manipuladoras o engañosas, que tengan el objetivo o "el efecto" de distorsionar materialmente el comportamiento de una persona, causando un daño significativo.

✓ Sistemas de IA que afecten grupos vulnerables

En la propuesta original esta prohibición estaba enfocada a determinados grupos vulnerables, como los niños o las personas con discapacidad. Y también hacía referencia a sistemas de IA que tuvieran el "propósito" de causar un "daño físico o psicológico"

En la versión final se hace referencia a sistemas de IA que exploten vulnerabilidades relacionadas con la edad, la discapacidad o la situación social o económica específica con el objetivo o efecto de distorsionar el comportamiento de alguien de una manera que cause un daño significativo.

✓ Calificación social

Originalmente el Reglamento de IA establecía la prohibición de la calificación social por parte de autoridades pública. Desde CECU requerimos su ampliación a entidades privadas porque entendemos que resulta una práctica demasiado degradante y discriminatoria para las personas consumidoras. En el texto final esta prohibición fue ampliada a entidades privadas.

– Categorización biométrica

Se introdujo una nueva prohibición de los sistemas de categorización biométrica que infieren atributos sensibles (raza, política, orientación sexual, sindicato), excepto el etiquetado o filtrado de conjuntos de datos biométricos adquiridos legalmente, como imágenes, basados en datos biométricos o la categorización de datos biométricos en el área de aplicación de la ley

✓ Extracción o *scrapping* de imágenes faciales de Internet o de imágenes CCTV

Se introdujo la prohibición de crear de bases de datos de reconocimiento facial mediante extracción no dirigida de imágenes faciales de Internet o de imágenes de CCTV.

4. Las personas consumidoras deben tener una serie de derechos reconocidos, tales como el derecho a acceder a la justicia y a la reparación, incluido el daño colectivo

✓ Reconocimiento de derechos en favor de las personas consumidoras

Desde los inicios de los debates alrededor de este Reglamento, desde CECU requerimos el reconcomiendo de derechos en favor de las personas consumidoras. En el texto final:

- Se ha introducido el derecho a presentar a reclamación ante la autoridad pertinente (art. 68b)
- × Se ha eliminado el derecho a acudir a remedio judicial contra las autoridades nacionales de supervisión.
- Se ha introducido el derecho a ser informado y recibir explicaciones del rol que ha tenido un sistema de IA (de los de alto riesgo) en un proceso de toma de decisiones (art. 68c)
- Se ha incluido el Reglamento en la Directiva de Acciones de Representación, que permite los reclamos colectivos (art. 68d)

5. Requerir un estudio de impacto de derechos fundamentales para todos los sistemas de IA de alto riesgo de forma previa a que se pongan en marcha

- Estudio de impacto de derechos fundamentales (art. 29a)

El Parlamento Europeo había introducido este requisito para todos los sistemas de IA de alto riesgo en verano de 2023. Esta decisión fue apoyada desde CECU.

Aunque este requisito se mantiene en el texto final, se ha visto limitado. El Estudio solo deberá ser realizado por implementadores de sistemas de alto riesgo que sea organismos públicos u operadores privados que prestan servicios públicos y operadores que implementen únicamente los sistemas de IA de evaluación crediticia y de seguros que enumera el Anexo III. Por su parte, los implementadores solo tendrán que mitigar los riesgos después de que se hayan materializado, lo cual resulta preocupante.

6. Se regulen adecuadamente los modelos base/fundacionales y la IA generativa (como ChatGPT o Bard) con el foco de proteger a las personas de sus daños, de manera que sean más fiables, seguros y de mayor calidad

- ✓ Reglas insuficientes para los modelos de propósito general –ex modelos base o fundacionales-, sistemas de IA de propósito general e IA Generativa (art. 52a en adelante)

Hay algunos aspectos positivos, pero también faltan algunas disposiciones claves para garantizar la rendición de cuentas por parte de las empresas. Por ejemplo, no hay reglas para los sistemas de IA de propósito general (en adelante, "IA-PG"), solo modelos de IA-PG.

Las reglas están escalonadas: algunas se aplican a los modelos de IA de propósito general, mientras que se aplican reglas más estrictas a los modelos de IA con riesgos sistémicos.

Designación del modelo de IA-PG con riesgo sistémico:

- El criterio principal para determinar si un modelo de IA-PG tiene un riesgo sistémico son las capacidades de alto impacto evaluadas sobre la base de herramientas técnicas apropiadas (por ejemplo, FLOPS -la cantidad de computación utilizada para el entrenamiento- superiores a 10^25).
- Si un modelo IA-PG cumple este requisito, deberá notificarlo a la Comisión en un plazo de dos semanas.
- La Comisión Europea puede designar un modelo IA-PG en determinadas circunstancias.
- Es necesario publicar la lista de modelos IA-PG con riesgo sistémico.

Las reglas que se aplicarán a todos los proveedores del modelo son:

- Elaborar y mantener actualizada la documentación técnica; Compartir la documentación técnica necesaria con los desarrolladores posteriores;
- Disponer de medidas para respetar la ley de derechos de autor;
- Publicar un resumen detallado del contenido utilizado para entrenar el modelo (plantilla que elaborará la Oficina de IA);
- Puede basarse en códigos de prácticas para demostrar el cumplimiento de estas obligaciones hasta que se publique un estándar armonizado.

Las reglas que se aplicarán a los modelos de IA de propósito general con riesgo sistémico:

- Realizar una evaluación del modelo (incluida la realización de pruebas adversas internas y/o externas);
- Evaluar y mitigar posibles riesgos sistémicos;
- Realizar un seguimiento de los documentos y reportar a la oficina de IA información relevante sobre incidentes graves;
- Garantizar un nivel adecuado de ciberseguridad;
- Puede basarse en códigos de prácticas para demostrar el cumplimiento de estas obligaciones hasta que se publique un estándar armonizado.

Códigos de práctica:

- La oficina de IA facilitará la elaboración de códigos de prácticas para garantizar la conformidad de los modelos IA-PG con el Reglamento.
- Tendrá en cuenta los enfoques internacionales.
- La sociedad civil puede apoyar el proceso de redacción de códigos de conducta.

7. Los estándares armonizados deben utilizarse únicamente para definir requerimientos técnicos, no para definir o aplicar principios legales o relacionados con los derechos fundamentales y su desarrollo debe garantizar mayor participación de la sociedad civil

– Desarrollo de estándares técnicos

Hay que tener en cuenta que la mayoría de las obligaciones que dispone el Reglamento dependen del desarrollo de estándares técnicos. Y que su art. 40 refiere a que los sistemas de IA de alto riesgo y los modelos de propósito general que estén de conformidad con estándares armonizados se presumirán que cumplen con la norma.

Por eso, desde CECU ponemos de resalto que los procesos europeos de estandarización están dominados por la industria, por lo que hay que involucrar a otros actores de la

sociedad civil en desarrollar estándares de IA, ya que esta tecnología puede influir en nuestros derechos fundamentales. Más allá de que seguimos considerando que los estándares técnicos no deben involucrarse en materia de derechos fundamentales, lo cierto en la práctica será muy difícil trazar esa línea. De ahí la necesidad de ampliar la pluralidad de voces en la elaboración de estándares técnicos.