DIPUTACIÓ DE VALÈNCIA

Protecció de Dades i Seguretat de la Informació

Boletín protección de datos

Boletín del Departamento de protección de datos y Seguridad de la Información de la Diputación Provincial de Valencia

Boletín N.º 36| Junio 2023

SISTEMAS DE INTELEIGENCIA ARTIFICIAL (IA) Y PROTECCIÓN DE DATOS DE CARÁCTER PERSONAL

Í N D I C E

Os invitamos a trasladarnos aquellas temáticas que resulten de vuestro interés para los próximos boletines informativos. Estas peticiones deberán dirigirse a:

Diputación de Valencia

Dpto. de Protección de Datos y Seguridad de la Información

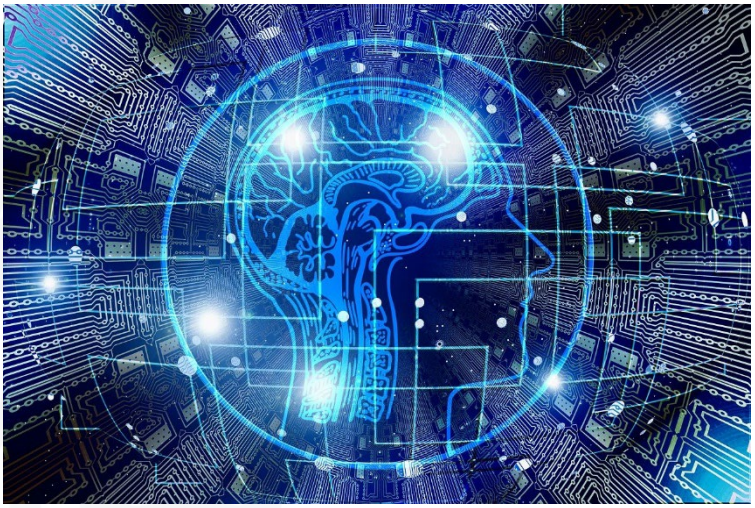
Pl. de Manises, 4 46003 Valencia

email: dpdsi@dival.es

SUSCRIPCIONES

Si deseas suscribirse a nuestro Boletín informativo accede al siguiente enlace

INTRODUCCIÓN



Actualmente la protección de los datos personales se ve desafiada por el rápido desarrollo y el veloz despliegue de la Inteligencia Artificial (en adelante, "IA"), pues su utilización implica necesariamente el tratamiento de datos masivos, dentro de los cuales se incluyen diferentes categorías de datos personales.

La mayoría de las aplicaciones de IA requieren grandes volúmenes de datos para aprender y tomar decisiones inteligentes. En ese sentido, los datos son necesarios no solo para que la IA alcance su máximo potencial, sino también para que ésta pueda evitar sesgos o errores al momento de realizar un tratamiento.

El rápido avance de la tecnología y las herramientas de IA han traído cambios que posibilitan el procesamiento de millones de datos en diferentes partes del mundo y por diferentes actores a una velocidad inimaginable. Pero ¿tenemos respuestas para todos los retos que suponen en privacidad el uso de estos sistemas?

Cabe preguntarse cuáles son los escenarios en los que la IA pueda generar un riesgo al titular de datos personales. En el presente Boletín trataremos de dar respuesta a algunas cuestiones suscitadas a propósito de la afectación de este tipo de sistemas de IA y las diferentes implicaciones que tienen. Así, nos preguntamos si los sistemas de IA ¿podrían ser un medio seleccionado por un responsable para implementar operaciones de datos personales en un tratamiento, pero no constituir un tratamiento en sí mismo? y ¿cuán importante es el principio de exactitud en la recogida de datos y su posterior procesamiento?, así como ¿qué implicaciones tiene en las decisiones automatizadas en este tipo de sistemas?

SISTEMA VS. TRATAMIENTO IA

Un sistema de Inteligencia Artificial (IA), o varios sistemas de IA, podría ser un medio seleccionado por una organización para implementar operaciones de datos personales en un tratamiento, pero no constituye un tratamiento en sí mismo. Es importante entender que la finalidad última de un tratamiento es diferente de los medios seleccionados para implementarlo. Será el responsable quien determine si los resultados de un sistema de IA implicarían una decisión automática o determinará que se incluya una supervisión humana que tome la decisión final. Por ello, tal como dispone la AEPD en alguna de sus publicaciones, las decisiones automatizadas no están en la naturaleza del sistema de IA, sino que son una opción elegida por el responsable.

Teniendo en cuenta la premisa, sobre que un sistema de IA en sí mismo es sólo un medio para el tratamiento de datos personales y no un fin último, pongamos como ejemplo el marco de un tratamiento de datos con fines de selección de personal de una bolsa de empleo.

Con el ejemplo seleccionado, el tratamiento se define por su **finalidad** (reclutamiento), **alcance** (datos de empleados potenciales afectados, duración del proceso de selección, etc.), **contexto** (regulación laboral, situación social, estatus de dicho trabajo en la sociedad, etc.) y **naturaleza** (cómo se implementará el tratamiento).

Tras decidir la implementación del tratamiento de datos, el responsable abordará el diseño del conjunto de operaciones necesarias para la consecución de su finalidad, algunas de la cuales requerirán procesar datos y otras no. Entre dichas operaciones pueden figurar las siguientes:

- publicar la oferta laboral,
- disponer de un procedimiento para orientar y recoger las solicitudes,
- filtrar y seleccionar las solicitudes que cumples con los requisitos,
- hacer y publicar la selección final,
- contratar al finalmente elegido.

Será el responsable - la Administración Pública oportuna- quien determine el medio de implementación de las diferentes operaciones. Estos podrán ser automatizados o manuales, en las

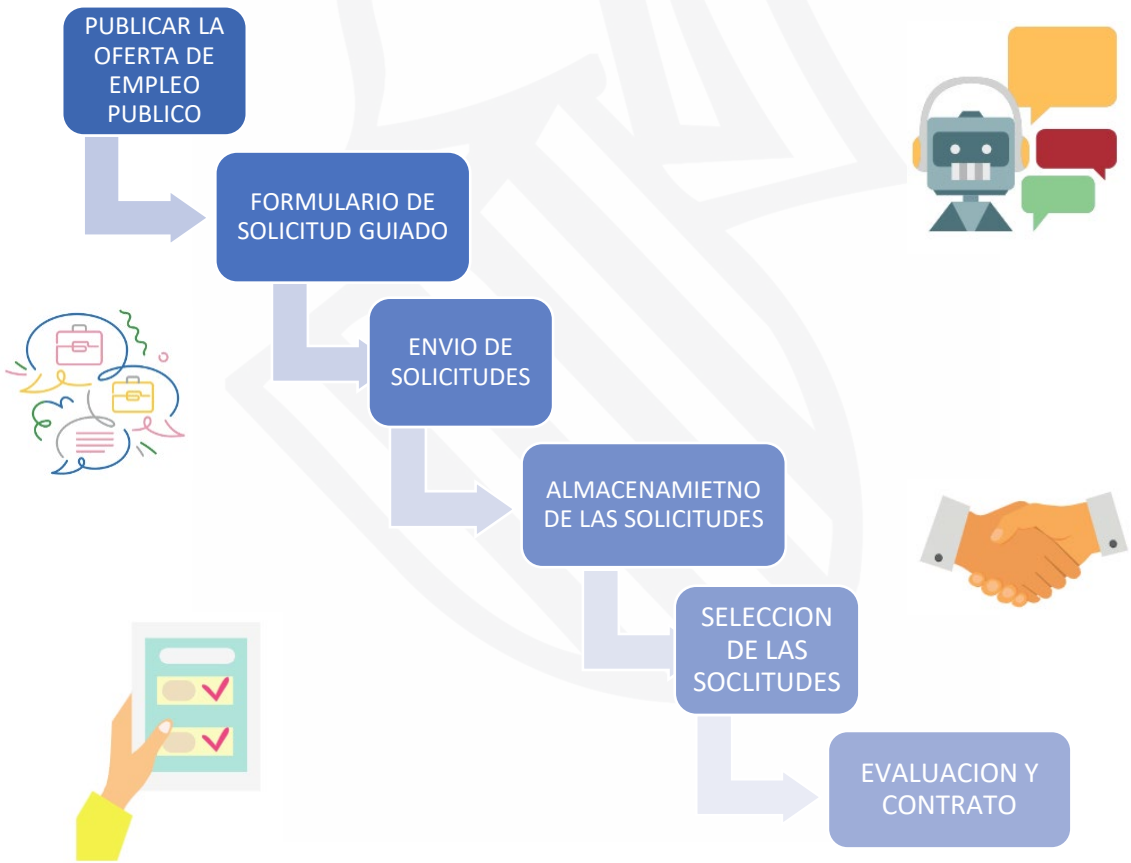
propias instalaciones o en la nube, por sistemas móviles, mediante la externalización de algunas operaciones a los encargados, etc.

En este marco, la AEPD pone varios ejemplos de cómo el responsable puede decidir implementar las distintas operaciones del tratamiento propuesto.

Así, señala la posibilidad de que el responsable decida que el procedimiento para guiar a los candidatos a completar el formulario de solicitud donde incluirían los requisitos a implementar se implemente mediante un chatbot (un sistema de IA) dentro de la propia web de la Administración; que decida utilizar un sistema de IA para la selección automática de las solicitudes que si cumplen las condiciones de la oferta, de acuerdo con los criterios necesarios para el desarrollo del puesto determinados por el responsable (la Administración), quien también decidirá si dicha selección será supervisada por un humano o no, y por tanto si dicha operación se configura como una toma de decisiones automatizada.

El responsable también podría implementar la evaluación de los candidatos mediante otro sistema de IA que realice y evalúe las pruebas para los candidatos previamente seleccionados e incluso decidir que la selección final de los candidatos se realice por los resultados del sistema de IA o que hubiera una supervisión humana de dicha evaluación.

En consecuencia, la AEPD sostiene que el responsable tiene en su mano la determinación de configurar una operación como una toma de decisiones automatizada, y lo ilustra del siguiente modo en el ejemplo propuesto:



Por tanto, en este ejemplo, el responsable decide utilizar hasta tres sistemas de IA diferentes en el marco de una misma actividad de tratamiento, pero cada uno de estos sistemas constituye un medio para implementar distintas operaciones y no un tratamiento en sí mismo.

Así, entendemos que son operaciones dentro de un tratamiento, porque cada operación aislada no podría legitimarse si no se incluye en un tratamiento amplio con una finalidad última y legítima.

Al estudiar este proceso la conclusión que alcanzamos es que es una decisión del responsable sobre el uso de los sistemas de IA pueda conducir a la toma de decisiones automatizadas. En consecuencia, **el hecho de que la toma de decisiones sea automatizada no es una característica del sistema IA en sí**. A la hora de diseña el tratamiento, **es el responsable quien determinará si incluye una operación adicional de supervisión humana sobre los resultados producidos por el sistema IA**.

Finalmente, y teniendo en cuenta lo señalado por la AEPD hallamos que un tratamiento de datos personales podría ser implementado por diferentes sistemas de IA simultáneamente, y también, estos sistemas de IA podrían implementarse en local o en la nube, podrían involucrar encargados del tratamiento, etc.

Los sistemas IA formarán parte de la naturaleza del tratamiento cuando hayan sido incluidos en algunas de las operaciones necesarias para llevar a cabo el tratamiento con su finalidad explícita. En tal caso, se podrían generar riesgos específicos para los derechos y libertades de los interesados que deben ser evaluados y gestionados.

"EN CONSECUENCIA, EL HECHO DE QUE LA TOMA DE DECISIONES SEA AUTOMATIZADA NO ES UNA CARACTERÍSTICA DEL SISTEMA IA EN SÍ. A LA HORA DE DISEÑA EL TRATAMIENTO, ES EL RESPONSABLE QUIEN DETERMINA SI INCLUYE UNA OPERACIÓN ADICIONAL DE SUPERVISIÓN HUMANA SOBRE LOS RESULTADOS PRODUCIDOS POR EL SISTEMA IA".

PRINCIPIO DE EXACTITUD EN LOS TRATAMIENTOS IA

En un tratamiento de datos que incorpora un sistema de IA resulta necesario evaluar la exactitud de los datos que se van a tratar. Puesto que en ocasiones el comportamiento de un algoritmo puede verse comprometido por la inexactitud de los datos de ejecución del mismo, no solo por los datos utilizados en su desarrollo.

Se hace necesario evaluar la fidelidad de los datos de entrada, ya que los mismos podrían introducir sesgos resultando comprometidos no solo el rendimiento del algoritmo, sino todo el tratamiento de datos. Con ello debe se debe procurar introducir salvaguardas que se ocupen de evitar las posibles inexactitudes de los datos de entrada y proteger asi de su posible impacto; con estrategias “desde el diseño” en la ejecución del tratamiento.

Con respecto al **principio de exactitud**, podemos identificar varios puntos clave derivados de este ejemplo desde el punto de vista del RGPD:

- La falta de definición de los datos de entrada a un algoritmo podría dar lugar a errores o sesgos que no forman parte del algoritmo en sí.
- El principio de exactitud debe aplicarse tanto en los datos de entrada, como de salida e incluso en los datos intermedios de todo tratamiento.
- Debe establecerse una definición precisa de cada dato de entrada (semántica) "por diseño" y documentarse adecuadamente.
- El impacto de cada dato de entrada en el resultado final debe evaluarse "por diseño", para cada fin específico, mediante la realización de un análisis del algoritmo implementado, mediante las pruebas de verificación de los requisitos y mediante las pruebas de validación en el contexto de la operación.

- De recopilarse de manera manual los datos de entrada los interesados (y los que recopilan datos) deben conocer y comprender la semántica de los datos y el impacto de su respuesta.

- Cabe la posibilidad de que los datos de entrada de un algoritmo se recopilan de fuentes distintas, ya sean bases de datos, sensores como cámaras, lectores de huellas dactilares. Desde la etapa de recopilación hasta la ejecución del algoritmo, Los datos pueden sufrir varias transformaciones que también forman parte del tratamiento. Todas esas operaciones forman parte del tratamiento, junto con el algoritmo.

- Por último, "Se adoptarán todas las medidas razonables para que se supriman o rectifiquen sin dilación los datos personales que sean inexactos con respecto a los fines para los que se tratan" (artículo 5.1.d) del RGPD), es decir, para cada fin específico y durante el funcionamiento del tratamiento.

En conclusión, hay que tener en cuenta que incluso el algoritmo más preciso puede fallar porque otro elemento del tratamiento no funciona correctamente, y en cambio un algoritmo menos preciso puede funcionar lo suficientemente bien si durante el diseño se incluyen las garantías adecuadas para evitar gestionar posibles errores.

"PODEMOS CONCLUIR QUE UNA IMPLEMENTACIÓN ADECUADA DEL PRINCIPIO DE EXACTITUD (ARTÍCULO 5.1.D) DEL RGPD) ES ESENCIAL EN LA EJECUCIÓN DEL ALGORITMO Y EN EL DESEMPEÑO DE TODO EL TRATAMIENTO."

INTELIGENCIA ARTIFICIAL Y LAS DECISIONES AUTOMATIZADAS

El Reglamento General de Protección de Datos (artículo 22.1) garantiza a los interesados, el derecho a no ser objeto de decisiones individuales automatizadas, incluida la elaboración de perfiles, que produzca en ellos efectos jurídicos o les afecten significativamente. Además, prevé la intervención o supervisión humana sobre la decisión automática para salvaguardar los derechos, las libertades y los intereses legítimos de los interesados.

Dejar claro que, este derecho se aplica única y exclusivamente a la toma de decisiones totalmente automatizadas y que produzca efectos jurídicos o, en su caso, afecten significativamente a los titulares de los datos tratados; es decir, que afecte a sus derechos o a su esfera jurídica, como, por ejemplo, la denegación de subvenciones públicas o de entrada a un país, o la cancelación de un contrato o afecte significativamente a cualquiera de éstos.

Más ejemplos al respecto, serían a desestimación en un proceso de selección, la denegación de un crédito o seguro, o la aplicación de precios distintos a un mismo colectivo. No es necesario que afecte a un gran número de personas. De acuerdo con lo anterior, vemos que la casuística puede ser muy diversa:

- Cabe tratamiento automatizado de datos sin elaboración de perfiles
- Elaboración de perfiles como apoyo a procesos de decisión automatizada
- Elaboración de perfiles como base de procesos de decisión completamente automatizada
- Procesos de decisión automatizada con o sin elaboración de perfiles previa.

Pues bien, debemos mencionar la inminente entrada en vigor del Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de Inteligencia Artificial (en adelante, RUEIA) el cual plantea criterios de coherencia normativa en la interpretación de este artículo 22 RGPD. Los artículos 5 y 14 RUEIA establecen, respectivamente, la prohibición de la IA solo en ciertos casos y la supervisión humana únicamente en supuestos de sistemas de IA de alto riesgo. En el resto de los usos de la IA no se prevé ni reserva de humanidad ni supervisión humana.

En el décimo punto de la Exposición de Motivos de la citada norma se indica que los Estados miembros deben estar facultados para especificar la aplicación de las disposiciones a través de su propia normativa nacional, para lo cual el Reglamento concede a los Estados un margen de maniobra en la concreción de dicha normativa interna, incluso en el tratamiento de datos personales.

En relación con la legislación española sobre IA y la toma de decisiones administrativas, cabe hacer mención del artículo 23 de la Ley 15/2022, de 12 de julio, integral para la igualdad de trato

y la no discriminación, que establece los principios de buena administración y diligencia debida como límites de las decisiones administrativas automatizadas.

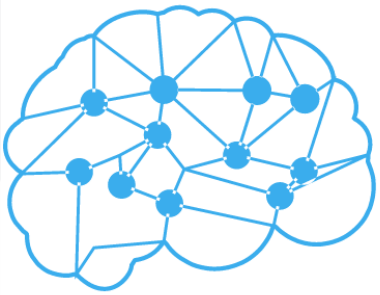
De conformidad con este estándar jurídico, las Administraciones públicas deberán tener en cuenta criterios de minimización de sesgos, transparencia y rendición de cuentas, y, siguiendo las recomendaciones de la U.E., promoverán *"el uso de una IA ética, confiable y respetuosa con los derechos fundamentales"*. Con todo, las operaciones de elaboración de perfiles realizadas por las autoridades públicas deberían ser legales, proporcionadas y necesarias respecto a los fines de dichas operaciones.

La mención a la necesidad de seguir avanzando en la rendición de cuentas de los sistemas de IA con toma de decisiones algorítmicas resulta oportuna, en aras de la protección de los derechos y libertades de los ciudadanos. Puesto que, a pesar de las innegables ventajas del empleo de sistemas de decisiones automatizadas basadas en algoritmos en términos de eficacia y economía, no pueden soslayarse los riesgos que conllevan estas decisiones para los derechos e intereses de la ciudadanía.

PROTECCION DE DATOS Y LA DIMENSIÓN ÉTICA

La perspectiva ética de la IA, como una parte de la "ética digital", es uno de los aspectos que más inquietud despierta. Es preciso ser vigilantes tanto sobre la legitimidad ética de los tratamientos como de los efectos inesperados de estos. Asimismo, debe considerarse el posible impacto colateral de dichos tratamientos en un entorno social, más allá de las limitaciones concebidas inicialmente de propósito, de duración en el tiempo y de extensión.

Es decir, hay que analizar la solución IA per se, pero también en el marco del tratamiento en el que se integra, y las relaciones de dicho tratamiento con el entorno en varios aspectos:



- En el aspecto cultural, con su escala de valores.
- En el contexto en el que se despliega el servicio, con sus requisitos de calidad.
- En los aspectos que se derivan de la interconexión masiva de componentes en la sociedad de la información.

Los sistemas algorítmicos requieren evaluaciones adecuadas del impacto en la privacidad, que incluyan también consideraciones sociales y éticas de su utilización y un empleo innovador del enfoque de privacidad desde la etapa de concepción. Los actores de la IA deben asumir la responsabilidad de la concepción y la aplicación de los sistemas de IA de manera que se garantice la protección de la información personal durante todo el ciclo de vida del sistema de IA.

Un aspecto crítico de los sistemas de IA es el de la posible existencia de sesgos ("bias" en inglés). Los sesgos son particularmente graves cuando, por ejemplo, derivan en discriminaciones de un de un grupo en favor de otro. Pero, existe otro tipo de sesgo que puede ser aún más preocupante y que es el del **sesgo en la interpretación de los resultados de la IA**.

Este sesgo humano consiste en aceptar, sin espíritu crítico, los resultados de una IA como ciertos e inamovibles.

Como ya mencionamos en el epígrafe anterior, puede ocurrir que, en algunas ocasiones, los seres humanos decidan depender de los sistemas de IA por razones de eficacia. Pero, la decisión de ceder el control en contextos limitados seguirá recayendo en los seres humanos, ya que estos pueden recurrir a los sistemas de IA en la adopción de decisiones y en la ejecución de tareas. Sin embargo, un sistema de IA nunca podrá reemplazar la responsabilidad final de los seres humanos y su obligación de rendir cuentas. Por regla general, las decisiones de vida o muerte no deberían cederse a los sistemas.

Son diferentes problemáticas las que deberán abordarse de cara a disponer de una protección homogénea sobre el uso de los sistemas de IA. Actualmente, estas son algunas de las cuestiones que los Estados Miembros deberán resolver:

1. Que sea posible atribuir la responsabilidad ética y jurídica, en cualquier etapa del ciclo de vida de los sistemas de IA, así como en los casos de recurso relacionados con sistemas de IA, a personas físicas o a entidades jurídicas existentes.

2. Deberían establecer marcos de evaluación del impacto, como evaluaciones del impacto ético, para determinar y analizar los beneficios, los problemas y los riesgos de los sistemas de IA, así como medidas adecuadas de prevención, atenuación y seguimiento de los riesgos, entre otros mecanismos de garantía.

3. Velar por que los mecanismos de gobernanza de la IA sean inclusivos, transparentes, multidisciplinarios y multilaterales.

4. Procurar elaborar estrategias de gobernanza de datos que garanticen la evaluación continua de la calidad de los datos de entrenamiento para los sistemas de IA, en particular la idoneidad de los procesos de recopilación y selección de datos, y que prevean medidas adecuadas de seguridad y protección de los datos, así como mecanismos de retroalimentación para aprender de los errores y compartir las mejores prácticas entre todos los actores de la IA.

MATERIAL COMPLEMENTARIO

- Reglamento (UE) 2022/868 del Parlamento Europeo y del Consejo de 30 de mayo de 2022 relativo a la gobernanza europea de datos y por el que se modifica el Reglamento (UE) 2018/1724 (Reglamento de Gobernanza de Datos). Consulta la Directiva en este enlace.
- Publicación guía de la AEPD sobre "Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción. Consulta la Publicación en este enlace.
- Publicación de la AEPD, sobre "Mapa de referencia para tratamientos que incluyen Inteligencia Artificial". Consulta este enlace.
- Propuesta de reglamento del parlamento europeo y del consejo por el que se establecen normas armonizadas en materia de inteligencia artificial (ley de inteligencia artificial) y se modifican determinados actos legislativos de la unión. Consulta este enlace.
- Carta de los derechos fundamentales de la Unión Europea. Consulta este enlace.
- Tratado de funcionamiento de la Unión Europea. Consulta este enlace.
- Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de Inteligencia Artificial). Consulta este enlace.

NOTICIAS

- **Las autoridades de la Red Iberoamericana de Protección de Datos Personales inician una acción coordinada en relación con el servicio ChatGPT**

La Red Iberoamericana de Protección de Datos (RIPD), foro constituido por 16 autoridades de protección de datos de 12 países de la región cuyo objetivo es promover y garantizar el derecho fundamental a la protección de datos personales y cuya secretaría permanente ostenta a la Agencia Española de Protección de Datos, va a coordinar una acción sobre el servicio ChatGPT.

Consulta la publicación en este enlace.

- **La AEPD inicia de oficio actuaciones de investigación a OpenAI, propietaria de ChatGPT**

La Agencia Española de Protección de Datos (AEPD) inició de oficio actuaciones previas de investigación a la empresa estadounidense OpenAI, propietaria del servicio ChatGPT, por un posible incumplimiento de la normativa

Consulta la publicación en este enlace.

- **La autoridad de control de protección de datos de Italia bloquea ChatGPT. Recopilación ilegal de datos personales. Ausencia de sistemas de verificación de la edad de los menores.**

El Garante per la Protezione dei Dati Personalipara la protección de datos personales ordenó, con efecto inmediato, la limitación temporal el tratamiento de datos de usuarios italianos. Alemania, Francia e Irlanda han solicitado informacion al Garante por si también toma estas medidas.

Consulta la publicación en este enlace.