URÍA
MENÉNDEZ

Principales conclusiones de la guía de la AEPD para la “Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial”

Marzo de 2020

Objetivo y alcance

La guía de la Agencia Española de Protección de Datos (la "**AEPD**") para la *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción* (la "**Guía**") tiene como objetivo ser una "*primera aproximación*" para la adecuación al Reglamento General de Protección de Datos (el "**RGPD**") de productos y servicios que incluyan componentes de inteligencia artificial ("**IA**"). La AEPD no pretende, por tanto, realizar un repaso exhaustivo a lo establecido en el RGPD, pero sí abordar las dudas planteadas en el marco de protección de datos personales y señalar los aspectos más relevantes en la relación IA-RGPD que deben ser tenidos en cuenta desde el diseño e implementación de tratamientos que incluyan IA.

En particular, en la Guía, la AEPD facilita criterios sobre:

- La **información** que debe proporcionarse a los interesados cuyos datos se tratan en un sistema de IA.
- Los **roles** que asumen los distintos actores que participan en la creación, desarrollo y utilización de un sistema de IA.
- Las **medidas técnicas y organizativas** a aplicar por parte de los responsables del tratamiento para mitigar o eliminar determinados riesgos inherentes al uso de sistemas de IA.

Cuestiones y definiciones generales sobre la tecnología IA

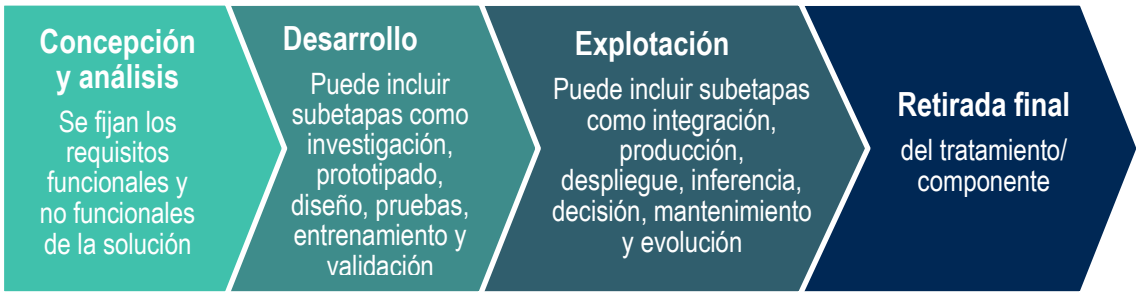
Análisis de la IA débil. La AEPD realiza una serie de consideraciones preliminares sobre la definición, concepto y principales características y tipologías de la IA (p. ej., *machine learning*). En este sentido, distingue tres categorías: las IA fuertes, generales y débiles, y señala que la Guía se centra en tratar la adecuación al RGPD de tratamientos que incorporan componentes de IA-débil (i. e., soluciones tecnológicas capaces de resolver un problema concreto y acotado).

Afectación (o no) de datos personales. La AEPD reconoce que pueden distinguirse aquellos componentes de IA que tratan datos de personas físicas (p. ej., modelo de perfilado de *marketing*) de otros que no implican tratamientos de datos personales (p. ej., modelos de predicción meteorológica).

Respecto de las soluciones que tratan datos personales, se identifican distintos tipos, tales como las que pueden hacer predicciones de un sujeto, las que realizan una evaluación sobre el estado actual de este o las que deciden la ejecución de un conjunto de acciones. También se señala que, en la toma de decisiones, la IA puede adoptar dos roles: (i) el de una ayuda para el proceso de decisión para que un ser humano tome la decisión final, y (ii) el de toma y ejecución (autónoma) de la decisión.

La Guía destaca que, en general, el componente IA no va a estar aislado, sino que va a formar parte de un tratamiento más amplio y que puede relacionarse con otras tecnologías o tratamientos, como *big data*, IoT, 5G, etc., y que hay tener en cuenta todos los elementos humanos y tecnológicos para determinar las implicaciones que tiene la utilización de un tratamiento que incorpora una solución de IA.

Ciclo de vida de las soluciones de IA. La Guía identifica las principales fases del ciclo de vida de una solución de IA:



Retos éticos. Por último, la AEPD destaca los retos éticos y no solamente jurídicos que pueden surgir del uso de esta tecnología, como los sesgos discriminatorios o los riesgos de aceptación de los resultados sin espíritu crítico, y la necesidad de evaluar esta dimensión ética en todo el ciclo de vida de la solución de IA.

Identificación de tratamientos y roles

Hechas las consideraciones preliminares, la Guía se centra en analizar los tratamientos de datos personales que se pueden realizar dentro de una solución de IA, así como en identificar los roles (responsable vs. encargado) de los distintos participantes en el desarrollo y/o gestión de una IA en atención a las distintas fases del ciclo de vida de la IA.

¿QUÉ TRATAMIENTOS DE DATOS PERSONALES SE GENERAN USANDO IA?

Entrenamiento	definición / búsqueda y obtención del conjunto de datos / preprocesamiento (tratamiento de datos no estructurados, limpieza, balanceo, selección, transformación) / *splitting* o partición del conjunto de datos para verificación / trazabilidad y auditoría.
Validación	puede implicar tratamiento de datos si se utilizan datos reales para determinar la bondad del modelo.
Despliegue	puede implicar comunicación de datos si la solución de IA se distribuye a terceros incluyendo datos de personales o si existe una forma de obtenerlos.
Explotación	• Inferencia: cuando se usen datos del interesado para obtener un resultado, cuando se usen datos de terceros con el mismo propósito o cuando datos e inferencias del interesado se almacenan. • Decisión: la decisión sobre un interesado es un tratamiento de datos. • Evolución: si se usan datos y resultados de los interesados para refinar el modelo de IA.
Retirada	puede tener efectos en tratamientos de supresión local, centralizada o distribuida de datos, así como sobre la portabilidad del servicio.

La AEPD recuerda que no todas las soluciones de IA tratan datos personales en todas las etapas de su ciclo de vida, pero avisa de que, si se considera que no se tratan datos personales (p. ej., porque estos se han eliminado o anonimizado), **hay que demostrar que estos procesos han sido realmente efectivos y evaluar cuál es el riesgo de reidentificación**.

¿QUÉ ROLES TIENEN LAS PARTES INVOLUCRADAS EN EL USO DE LA IA?

En la etapa de explotación del sistema de IA podemos encontrar varias situaciones:

1. El propietario de la IA da acceso a la solución a los interesados sometidos al tratamiento de IA (p. ej., herramienta de IA que hace una evaluación psicológica de los usuarios de una red social).

2. El propietario de la IA traspasa los derechos de uso de la solución de IA a un tercero.

3. Una entidad decide los fines de un tratamiento y contrata al propietario de la IA para que ejecute una de las fases de dicho tratamiento (responsable-encargado).

Tendrá la consideración de **responsable del tratamiento** la entidad que tome la decisión de utilizar, en el marco de un tratamiento de datos personales, una solución de IA, al ser la entidad que "*determina los medios y fines del tratamiento*".

La decisión de adoptar una solución técnica basada en IA es tomada por parte del responsable del tratamiento y, por tanto, tiene a su cargo la toma de decisión de seleccionar una solución tecnológica u otra. Por ello, los responsables deben ser diligentes al elegir la solución de IA y conocer su funcionamiento y los tratamientos de datos que se van a desarrollar dentro de dicha solución, sin que puedan escudarse en la falta de información o desconocimiento técnico. Por tanto, en ningún caso resulta aceptable para la AEPD trasladar la responsabilidad al propio sistema.

¿QUÉ ROLES INTERVIENEN EN LAS DISTINTAS ETAPAS DEL CICLO DE VIDA DE UN COMPONENTE DE IA?

Dejando de lado y sin entrar a evaluar los posibles modelos de redes cooperativas y modelos de corresponsabilidad, la AEPD determina que los roles de los distintos participantes en las diferentes etapas de un ciclo de vida de una IA podrían resumirse en los siguientes:

ETAPA	RESPONSABLE	ENCARGADO
Desarrollo / Entrenamiento	• Entidad que define los fines de la IA y decide qué datos van a utilizarse para entrenar a la IA.	• Tercero contratado para el entrenamiento o desarrollo de la IA si el contratante determina los términos y características principales del tratamiento (indiferente si el contratante proporciona los datos o si el contratado los recopila por sí mismo).
Validación	• Entidad que desarrolla la IA si toma decisiones para sus propios fines sobre los datos a utilizar para su entrenamiento.	
Despliegue	• Entidades que venden y compran la solución de IA que contiene datos personales (comunicación de datos), a excepción de que la solución se venda a una persona física para su uso particular (aplicación de la excepción doméstica).	• Entidad que proporciona la solución de IA a un tercero para su uso en el marco de una prestación de servicios y no trata los datos con fines propios.
Inferencia / perfilado	• Entidad que trata los datos incluidos dentro de la solución de IA para sus propios fines, a excepción de que el tratamiento se lleve a cabo por una persona física en una actividad exclusivamente personal o doméstica.	• Entidad que proporciona la solución de IA a un tercero para su uso en el marco de una prestación de servicios y no trata los datos con fines propios.
Decisión	• Entidad que toma decisiones automatizadas sobre los interesados para sus propios fines.	• Entidad que proporciona la solución de IA a un tercero para su uso en el marco de una prestación de servicios y no trata los datos con fines propios.

ETAPA	RESPONSABLE	ENCARGADO
Evolución	• Entidad que comunica a una tercera entidad datos de los usuarios. • Entidad que determina la evolución de la IA.	• Entidad contratada que trata los datos en el sistema de IA con la finalidad de prestar un servicio.

Cumplimiento de la normativa: *accountability*, legitimación, información y derechos de los interesados

La AEPD determina una serie de condiciones que deben cumplirse para garantizar que los tratamientos llevados a cabo a través de una solución de IA respeta la normativa de protección de datos. Destacan los siguientes elementos clave:

Accountability → Legitimación → Información → Derechos de los interesados

Llama especialmente la atención que la AEPD dispone, dentro de la Guía, que una solución de IA no será madura y, por tanto, no podrá cumplir con el principio de *accountability*, transparencia y legalidad en el caso de que no garantice ciertos elementos técnicos, como (i) la precisión, exactitud y medidas de error, (ii) la predictibilidad del algoritmo, o (iii) la consistencia de los resultados del proceso de inferencia.

¿SOBRE QUÉ BASE LEGAL PUEDEN REALIZARSE LOS TRATAMIENTOS DE DATOS EN UNA IA?

Las bases jurídicas que habitualmente legitimarán un tratamiento en una solución de IA son:

Más habituales

- Necesidad de tratar los datos para la ejecución de un contrato o para la aplicación de medidas precontractuales (p. ej., establecer un contrato con un interesado para el uso de sus datos en el entrenamiento de la IA).
- Interés legítimo.
- Consentimiento, teniendo en cuenta que la retirada posterior del consentimiento no afectará al tratamiento desarrollado hasta ese momento. La retirada del consentimiento no tendrá un efecto retroactivo con relación a los resultados obtenidos en un tratamiento ya realizado.

Excepcionales

- Protección de intereses vitales.
- Razones de interés público establecidas en el derecho de la UE o de los Estados miembros.
- Cumplimiento de obligaciones legales determinadas en una norma de la UE o de los Estados miembros.

Si se utilizan **conjuntos de datos de terceros** para el entrenamiento de la IA, el responsable del tratamiento deberá demostrar diligencia debida en comprobar la legitimidad de la fuente de datos adquirida (p. ej., que el contrato con el tercero que proporciona el conjunto incluya cláusulas en las que se reclamen evidencias y compromisos de dicha legitimidad).

Por lo que respecta a las categorías especiales de datos, la AEPD recuerda que el RGPD determina que las decisiones basadas únicamente en un tratamiento automatizado no se basarán en categorías especiales de datos, salvo si media el consentimiento del interesado o si el tratamiento es necesario por razones de un interés público esencial.

Por último, la AEPD destaca la posibilidad de que los datos personales pueden tratarse para finalidades distintas para las que se recogieron inicialmente, siempre que se cumplan todos los requisitos para un tratamiento lícito de los datos.

¿CÓMO CUMPLIR CON EL PRINCIPIO DE TRANSPARENCIA? ¿SOBRE QUÉ DEBO INFORMAR A LOS INTERESADOS?

La AEPD hace especial hincapié en el principio de transparencia y dispone que este es un aspecto crítico, dado que la información que se proporcione a los interesados debe permitirles conocer el impacto del uso de las soluciones de IA en el tratamiento de sus datos. Además, esa información debe proporcionarse no solo en un momento puntual, sino de forma dinámica a lo largo de todo el tratamiento de datos. Por ello, la AEPD dispone en la Guía ciertos criterios sobre cómo ser transparente con los

interesados que excederían del concepto de "transparencia" habitual, lo que comporta un esfuerzo adicional que deben realizar los responsables del tratamiento que traten datos en una solución de IA.

Información que debe proporcionarse durante la etapa de entrenamiento.- La AEPD destaca la necesidad de informar claramente al interesado sobre la posibilidad de ser reidentificable a partir de los datos del modelo.

Certificación.- Uno de los problemas clave es la confidencialidad para preservar la propiedad industrial. Pueden utilizarse mecanismos de certificación para aumentar la transparencia en los tratamientos de datos.

Primera capa informativa.- Los interesados deberán ser informados sobre el tratamiento de sus datos de conformidad con los artículos 13 y 14 del RGPD, y la información debe adaptarse a las peculiaridades de cada etapa del ciclo de vida de la IA en la que se realice el tratamiento.

La AEPD destaca que en la primera capa informativa debe incluirse, además de la información especificada en el artículo 11 de la LOPDGDD, información sobre la elaboración de perfiles o la toma de decisiones automatizadas. Se debe informar sobre el derecho a oponerse al tratamiento, la lógica aplicada y las consecuencias previstas para el interesado.

Por lo que respecta a la información sobre la lógica aplicada, la AEPD interpreta que debe proporcionarse al interesado información que le haga consciente del tratamiento que se lleva a cabo con sus datos y le proporcione certeza y confianza sobre los resultados que van a adoptarse a través de los algoritmos utilizados. Como ejemplo, la AEPD destaca que la información que podría tener relevancia para el interesado sería la siguiente:

- Detalle de los datos utilizados para la toma de decisión e importancia de cada uno de ellos.
- Calidad de los datos de entrenamiento y los patrones utilizados.
- Perfilados utilizados e implicaciones.
- Valores de precisión o error según la métrica adecuada para medir la bondad de la inferencia.
- Existencia o no de supervisión humana cualificada.
- Referencia a auditorías y certificación sobre el sistema.

- Si el sistema de IA contiene datos de terceros, prohibición de tratar dicha información sin legitimación y consecuencias de realizarlo.

Personal del responsable.- Los responsables deben proporcionar información clara y formación específica a su personal sobre las limitaciones de la IA. Además, si el sistema de IA se utiliza como una herramienta de ayuda a la toma de decisiones, deberán adoptarse medidas (de información, formación y auditorías) para gestionar el riesgo de que el personal solo transmita la decisión adoptada de forma automática sin llevar a cabo ningún tipo de intervención o cuestionamiento sobre dicha decisión.

Delegado de Protección de Datos ("DPD").- Aunque en algunos supuestos pueda no ser obligatorio, la AEPD destaca que el nombramiento de un DPD puede ser útil en empresas que utilizan sistemas de IA o que desarrollen soluciones de IA.

¿CÓMO ASEGURAR QUE SE RESPETAN LOS DERECHOS DE LOS INTERESADOS?

La AEPD hace especial hincapié en que los responsables del tratamiento deben establecer mecanismos y procedimientos para atender las solicitudes de ejercicios de derecho desde la misma fase del diseño del tratamiento o sistema de IA.

Además, si en el momento de la distribución de la solución de IA esta contiene datos personales, la AEPD indica que el responsable del tratamiento deberá:

- Suprimir los datos o, por el contrario, justificar la imposibilidad de hacerlo.
- Determinar la base jurídica para llevar a cabo la comunicación de datos a terceros e informar a los interesados sobre ello.
- Demostrar que se han ejecutado las medidas de privacidad por defecto y desde el diseño.
- Realizar una evaluación de impacto en función de los riesgos.

La AEPD también destaca la importancia de implementar, dentro del sistema de IA, la posibilidad de bloquear los datos relativos al proceso de inferencia (como mínimo, las entradas y resultados obtenidos) necesarios para atender un recurso o reclamación de los interesados.

La AEPD también incluye una serie de precisiones respecto a tres derechos:

- **Derecho de rectificación.** Podrán tratarse datos inexactos de forma excepcional siempre que el tratamiento se lleve a cabo con el fin de anonimizar y evitar la reidentificación de los interesados y dicho tratamiento de datos inexactos no tenga efectos sobre el interesado.
- **Derecho de portabilidad.** El responsable del tratamiento deberá evaluar y documentar si el tratamiento de datos que lleva a cabo está sujeto al derecho de portabilidad de los datos.
- **Derecho a no ser objeto de decisiones individuales automatizadas.** El responsable deberá proporcionar siempre la opción de que un operador humano pueda ignorar el algoritmo utilizado. Para ello, la AEPD recomienda documentar las incidencias o cuestionamientos de las decisiones automáticas que se reciban de interesados para poder valorar y detectar las situaciones en las que sea necesaria la intervención humana.

Por último, la Guía destaca que solo podrán tomarse decisiones individuales automatizadas sobre **datos de menores** cuando sea imprescindible para proteger el bienestar del menor y cuando se implementen las garantías adecuadas.

Gestión de riesgos

En la Guía, la AEPD destaca la importancia de implementar las medidas técnicas y organizativas adecuadas y proporcionales para garantizar un nivel de seguridad adecuado. Esas medidas serán determinadas a través de un análisis basado en el riesgo que consta de dos fases diferenciadas:

(i) una **primera fase** en la que se identifican las amenazas y se evalúa el nivel de riesgo intrínseco existente, y

(ii) una **segunda fase** en la que se implementan las medidas técnicas y organizativas adecuadas y proporcionales para eliminar o mitigar dicho riesgo.

1. IDENTIFICACIÓN DE LAS AMENAZAS

El responsable del tratamiento deberá ser capaz de identificar aquellos riesgos adicionales que puedan derivar de la novedad de la IA y de su modo de implementación. Dado que se deberán realizar evaluaciones de impacto de protección de datos ("**EIPD**") cuando exista una elaboración de perfiles o se lleven a cabo decisiones individuales automatizadas, a través de dichas evaluaciones también se concretarán las medidas específicas y concretas para la gestión del riesgo.

2. DETERMINACIÓN DE LAS MEDIDAS TÉCNICAS Y ORGANIZATIVAS

El responsable del tratamiento deberá determinar las medidas técnicas y organizativas necesarias para eliminar o mitigar el riesgo tras la realización del pertinente análisis. En relación con el uso de sistemas de IA, la AEPD destaca las siguientes medidas:

Principio de transparencia	Ver las medidas establecidas en el punto "*¿Cómo cumplir con el principio de transparencia? ¿Sobre qué debo informar a los interesados?*".

Principio de exactitud	**Factores**.- Existen tres factores que influyen en la exactitud de los datos: • La introducción de errores en los sistemas que se produzcan por la propia implementación del sistema de IA, por elementos externos (p. ej., a través de lectores biométricos) y, por último, por errores de programación o diseño. • La producción de errores contenidos dentro del conjunto de datos de entrenamiento o validación. • La evolución sesgada del modelo de IA. **Información biométrica.-** Debe darse especial importancia a la exactitud de los datos biométricos (p. ej., reconocimiento facial, huellas dactilares, voz, etc.) dada su especialidad.
Principio de minimización	**Limitaciones.-** La minimización consiste en: • Limitar la extensión de los tipos de datos a tratar a los necesarios en cada fase. • Limitar el grado de detalle o precisión de la información. • Limitar el número de interesados de los que se tratan datos. • Limitar la accesibilidad de las distintas categorías de datos al personal o usuario final en todas y cada una de las fases. **Técnicas.-** Existen varias técnicas de minimización de datos, entre las que destacan (i) la supresión de datos no estructurados o información no necesaria recogida durante el preproceso de la información, (ii) la agregación de datos y la (iii) anonimización y seudonimización. **Inclusión de datos de terceros en la solución de IA distribuida.-** En el caso de que la solución de IA se distribuya con datos de terceros, (i) deberá evaluarse de manera formal qué datos personales se contendrían dentro de dichas soluciones (i. e., en qué medida dichas personas son identificables), y (ii) deberán aplicarse medidas técnicas, organizativas o jurídicas para minimizar la identificación o extensión de la identificación al mínimo número de datos posibles.

Seguridad	**Amenazas específicas.-** Existen determinadas amenazas específicas para los sistemas de IA. Algunas de las que destaca la AEPD son las siguientes: • Acceso y manipulación del conjunto de datos de entrenamiento. • Inclusión de troyanos o puertas traseras en el código o en las herramientas de desarrollo. • Manipulación de la API de usuario para acceder al modelo y manipular los parámetros, filtrado o ataques a integridad o disponibilidad. • Filtrado o acceso a los *logs* resultado de las inferencias generadas en la interacción con los interesados. ***Logs* o registros de actividad.-** Su recogida forma parte de las estrategias para cumplir con el principio de *accountability*. Los *logs* deberán proporcionar las siguientes evidencias: • Determinar quién y bajo qué circunstancias accede a los datos personales que pudieran estar incluidos en el modelo. • Proporcionar trazabilidad en cuanto a la actualización de los modelos de inferencia, las comunicaciones del API del usuario con el modelo y la detección de intentos de abuso o intrusión. • Proporcionar trazabilidad para permitir la gobernanza en la comunicación de datos entre los intervinientes en la solución de IA. • Proporcionar seguimiento de los parámetros de calidad de la inferencia cuando la IA se utilice para la toma de decisiones o en procesos de ayuda a la toma de decisiones. **Base legal para la creación y mantenimiento de los *logs*.-** La AEPD destaca la posibilidad de tratar los datos personales contenidos en este tipo de ficheros con base en varias bases jurídicas establecidas en el artículo 6 del RGPD, entre las que destaca (i) la existencia de un interés legítimo en garantizar la seguridad de la red y de la información y (ii) el cumplimiento de obligaciones legales (p. ej., derivadas de la Ley 10/2010, de 28 de abril, de prevención del blanqueo de capitales y de la financiación del terrorismo).

EIPD	**Alto riesgo y ponderación.-** La AEPD hace especial hincapié en que el tratamiento de datos en el contexto de una solución de IA puede conllevar un "alto riesgo" para los derechos y las libertades de los interesados. Por ello, destaca la obligación de los responsables de valorar la posibilidad de llevar a cabo dicho tratamiento utilizando algún otro tipo de solución que, alcanzando la misma funcionalidad con un margen de rendimiento aceptable, tenga un riesgo menor para los derechos y libertades de los interesados. En conclusión, el responsable deberá llevar a cabo dicha ponderación y podrá emplear la solución de IA si se derivan más beneficios y ventajas para el interés general y la sociedad que perjuicios sobre los derechos y libertades de los interesados.
Auditoría	Debe llevarse a cabo una auditoría del tratamiento durante todo su ciclo de vida con el fin de determinar la adecuación a la normativa de protección de datos. Además, deberá determinarse que la solución de IA se utiliza para el propósito para el que fue diseñada y, si se ha adquirido de un tercero, analizar si pueden estar llevándose a cabo funciones colaterales que puedan tener repercusiones normativas (p. ej., que el sistema de IA realice tratamientos paralelos distintos a los gestionados por parte del responsable, como funciones analíticas).

Transferencias internacionales

Existen múltiples escenarios en los que se llevan a cabo flujos transfronterizos de datos en el marco de un sistema de IA, como por ejemplo que el desarrollo o despliegue del sistema se base en servicios en la nube, que se comuniquen datos a terceros para la evolución del modelo o que se distribuya el componente con datos de terceros. En este sentido, la AEPD destaca la importancia de que dichas transferencias cumplan en todo momento con las obligaciones del RGPD, en particular con el deber de información y con las garantías del capítulo V del RGPD.

Abogados de contacto

Leticia López-Lapuente
Socia
+ 34 91 586 01 31
leticia.lopez-lapuente@uria.com

Reyes Bermejo
Asociada Coordinadora
+34 96 352 91 91
reyes.bermejo@uria.com

Laia Reyes
Asociada *senior*
+34 91 586 01 31
laia.reyes@uria.com

Carolina Moniz Pina
Asociada *senior*
+34 91 586 01 31
carolina.monizpina@uria.com

Lidia Gimeno Rodríguez
Asociada *junior*
+34 91 586 01 31
lidia.gimeno@uria.com

BARCELONA
BILBAO
LISBOA
MADRID
PORTO
VALENCIA
BRUXELLES
LONDON
NEW YORK
BOGOTÁ
CIUDAD DE MÉXICO
LIMA
SANTIAGO DE CHILE
BEIJING

www.uria.com