



1

Preguntamos a un grupo de especialistas en inteligencia artificial sobre las principales incógnitas que despierta su uso y desarrollo. Nos han señalado cuáles son los mayores desafíos éticos, de seguridad, privacidad y regulación, haciendo hincapié en la necesidad de supervisión humana y señalando el avance exponencial que seguiremos viendo en el progreso de la IA.

2

¿Cómo describir el estado actual de la inteligencia artificial y qué avances consideramos que veremos en el futuro inmediato?

3

¿Qué desafíos éticos ve asociados al avance de la IA y cómo creemos que debemos abordarlos?

¿Cómo afecta la IA a la privacidad y la seguridad de los datos de las personas?

El desarrollo de la inteligencia artificial en las últimas décadas ha sido muy rápido debido al crecimiento de la capacidad computacional, la disponibilidad de grandes conjuntos de datos y avances en algoritmos. Esto ha llevado a la creación de múltiples aplicaciones y sistemas de IA en diversas áreas. Se trata de un campo en constante crecimiento con un gran potencial para transformar diferentes aspectos de nuestras vidas. Por este motivo, hemos contactado con una gran variedad de expertos en IA de diferentes ámbitos de trabajo. Hemos preguntado a Ramón López de Mántaras, uno de los pioneros de la IA en España; Nerea Luis, cofundadora de T3ch-Fest; Jordi Torres, profesor del grado de IA en la UPC; Idoia Salazar, cofundadora de OdiselA; Óscar Corcho, director académico de AI4Gov y Elena Gil, directora en Telefónica Tech. Es importante abordar este tema desde diferentes puntos de vista, explicando de manera comprensible las implicaciones, beneficios y desafíos de la IA. Con ello buscamos ayudar a contrarrestar la desinformación y los mitos sobre la inteligencia artificial, brindando información precisa basada en la experiencia de los entrevistados. “Lo fundamental es la necesidad de involucrar a la sociedad en su conjunto en la toma de decisiones sobre la IA para garantizar que esta tecnología sea utilizada de manera responsable”, remarca Jordi Torres en este reportaje.

CON LAS RESPUESTAS DE RAMÓN LÓPEZ DE MÁNTARAS, NEREA LUIS MINGUEZ, IDOIA SALAZAR, JORDI TORRES, ÓSCAR CORCHO Y ELENA GIL LIZASOLA.

Spanish experts address the ethical, security and regulatory challenges of artificial intelligence
CHALLENGES AND OPPORTUNITIES OF THE AI

We asked a group of specialists in artificial intelligence about the main questions raised by its use and development. They pointed out the major ethical, security, privacy and regulatory challenges, emphasizing the need for human oversight and noting the exponential advance we will continue to see in the progress of AI.

Keywords: artificial intelligence, cybersecurity, privacy, regulation, oversight, appropriate technology.

abordan los retos éticos, de seguridad y regulación de la inteligencia artificial

DESAFÍOS

Y

OPORTUNIDADES

DE LA IA



Palabras clave: Inteligencia artificial, ciberseguridad, privacidad, regulación, supervisión, tecnología adecuada.



Ramón
López de
Mántaras

ES EL PIONERO DE LA INTELIGENCIA ARTIFICIAL EN ESPAÑA. PROFESOR INVESTIGADOR DEL CENTRO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS (CSIC) Y DIRECTOR DEL INSTITUTO DE INVESTIGACIÓN DE INTELIGENCIA ARTIFICIAL (IIIA)

“LA INTELIGENCIA ARTIFICIAL NO ES NI TAN INTELIGENTE, NI TAN ARTIFICIAL”

1 Este boom del último año sobre todo viene dado por lo que llamamos “inteligencia artificial generativa”. Esta inteligencia artificial generativa no es nueva. Cuando empezó a desarrollarse parecía interesante, pero no que fuera a explotar de esta manera ni que se iban a meter las grandes empresas tecnológicas como ha ocurrido. En mi opinión, esto tiene aspectos positivos y aspectos negativos.

Yo sigo pensando que las grandes preguntas científicas que hay detrás de la inteligencia artificial, como si podrá haber inteligencias artificiales verdaderamente generales, y el avance hacia una inteligencia artificial general que verdaderamente entienda el mundo no lo veo por ninguna parte. Muchos defienden que la inteligencia artificial generativa es la antesala para conseguir una inteligencia artificial general, pero yo no lo pienso en absoluto. La generalidad de la inteligencia artificial generativa es muy limitada. Con estos sistemas preentrenados no estamos ganando generalidad. Estas inteligencias artificiales no nos acercan

a la respuesta a la pregunta de si habrá inteligencias artificiales que verdaderamente entiendan. Por ello me interesan más bien poco. A mí lo que me interesa es hacer ciencia.

2 Desde hace mucho, tenemos una serie de desafíos éticos como, por ejemplo, que un sistema de inteligencia artificial sea una caja negra y que no se sepa cómo funciona y que no tenga capacidad explicativa. Plantea una serie de cuestiones no sobre su uso ni sobre su impacto, sino sobre la prudencia necesaria a la hora de desplegarlo.

Con la inteligencia artificial generativa nos aparecen problemas, digamos, “nuevos” que hace unos años no se veían tan claramente y es el tema de generar todas las falsedades que se hacen y esa facilidad con la que puedes falsificar personas o hacer pasar por personas, ya sea de manera textual o audiovisual. Si seguimos con la inteligencia artificial generativa le veo difícil solución.

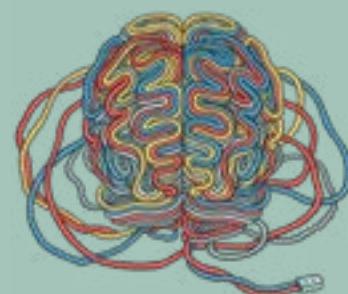
La manipulación se acentúa mucho más con esta capacidad de generar fakes. Llegará el día en el que en un juicio se presen-

tará un vídeo o una imagen de alguien cometiendo un delito infraganti y no se podrá utilizar porque no habrá manera de saber si es verdadero o falso. Esto es muy grave porque hace explotar la convivencia social e incluso la democracia.

Además, otro problema ético que se presenta son los millones de personas en países del sur global que llaman, como Filipinas o Pakistán, que están horas y horas al día limpiando lo que hacen los modelos de lenguaje por 0,50€/hora e, incluso, algunos trabajos que no se llegan ni a pagar. También modelando contenido o entrenando estas inteligencias artificiales generativas o etiquetando imágenes indicándole que esto es un árbol o esto es una persona o esto es una farola. Todo esto hecho a mano por una persona. Como digo, no estamos a la vuelta de la esquina de una inteligencia artificial general porque cada vez necesitamos más intervención humana. He llegado a la conclusión de que la inteligencia artificial no es ni tan inteligente, ni tan artificial.

3 Esto ya era un problema incluso sin la inteligencia

artificial. Hace ya bastantes años que parece que mucha gente ha renunciado a la privacidad. Yo creo que fue en el momento que aparecen los móviles y los sistemas que hacen recomendaciones o un perfil de tus gustos y preferencias que están diseñados para que te enganches y que además son supuestamente gratis. Hay que moverse desde la sociedad civil y organizarse para enfrentarse a todo lo que se nos está viniendo encima con esto de la inteligencia artificial generativa. Por primera vez en mi vida, viendo la balanza, soy más pesimista que optimista porque, ¿en qué mal momento salió la inteligencia artificial generativa? A nivel científico no sirve de nada prácticamente y a nivel práctico no veo muy claro quién está ganando con esta IA.



Nerea
Luis
Mingueza

DOCTORA EN CIENCIAS DE LA COMPUTACIÓN Y COFUNDADORA DE T3CHFEST. TRABAJA COMO INGENIERA EN INTELIGENCIA ARTIFICIAL EN SINGULAR

“LA RECOGIDA DE DATOS PARA EL ENTRENAMIENTO DE LOS MODELOS ES UN PRIMER DESAFÍO”

1 Vivimos un momento en el que la inteligencia artificial se está integrando en las herramientas que conocemos tales como aquellas relacionadas con la ofimática, los gestores de email, de videollamadas o el diseño gráfico. Esto hará que rápidamente aprendamos a sacarle su potencial en las tareas cotidianas y que los mismos modelos de IA aprendan y personalicen su tono, estilo y uso de la palabra en función de los documentos previos que hayamos escrito. Sin duda la personalización será una pieza clave para simplificar los prompts. Por otro lado, en la generación de imágenes, técnicas como Gaussian Splatting están dando muy buenos resultados para recrear escenas 2D > 3D. Sin duda se convertirá en otra de las tendencias próximas.

2 Hay un primer desafío alrededor de la recogida de datos para el entrenamiento de los modelos “grandes” o “fundacionales” (como la familia GPT o Palm2). En algunos casos se están viendo avances como en Adobe Firefly, donde solo entrenan con imágenes libres, pero aún hay mucha ca-

suística alrededor de cómo pedir permiso o cómo reportar sobre las fuentes de datos que se utilizan para entrenar a los modelos.

Otro punto es la propiedad intelectual de los textos e imágenes generadas, tanto para reclamar autoría como para reclamar un mal uso de las herramientas de IA si se utilizan para generar desinformación o deepfakes.

3 Existen multitud de servicios web ahora mismo que utilizan por detrás modelos de inteligencia artificial generativa. Generar imágenes, hacer preguntas sobre un PDF, resumir reuniones... En todos estos servicios, a la hora de usarlos, tenemos que preocuparnos de saber si los datos que volcamos los utilizamos para “reentrenar” o si por el contrario solo los mantienen un tiempo online con mero fin consultivo. De esta forma sabremos a qué riesgo nos estamos exponiendo cuando los utilizamos.

Poco a poco van saliendo licencias de uso personal, para empresas y entornos en la nube donde se respeta la privacidad y la seguridad de los datos, eso sí.



Jordi
Torres

CATEDRÁTICO DE LA UNIVERSIDAD POLITÉCNICA DE CATALUÑA (UPC). ES AUTOR DE MÁS DE 200 PUBLICACIONES DE INVESTIGACIÓN, Y LIBROS SOBRE COMPUTACIÓN DE ALTO RENDIMIENTO Y DE INTELIGENCIA ARTIFICIAL

“LA HUMANIDAD SE BENEFICIARÁ ENORMEMENTE DE LA INTELIGENCIA ARTIFICIAL ACTUAL”

1 A pesar de que actualmente la población en general tiene la percepción de que la inteligencia artificial (IA) ha alcanzado competencias casi humanas, la realidad es que los algoritmos en que se basan las IA no tienen inteligencia, sino solo “habilidades sin comprensión”, en el sentido de que son algoritmos que pueden llegar a ser muy hábiles realizando tareas específicas, pero sin comprender absolutamente nada de lo que están haciendo, ni cuentan con el llamado “sentido común” de la inteligencia humana que la hace tan diversa, rica, espontánea y maravillosamente imprevisible.

A pesar de que los sistemas actuales de IA todavía carecen de capacidades importantes, estamos inmersos en una carrera para crear sistemas de IA cada vez más potentes. Estos últimos años este progreso de las IA se ha sustentado en una aproximación de “fuerza bruta”, con la creación de algoritmos cada vez más grandes, entrenados en supercomputadores cada vez más potentes, procesando cada vez más y más datos. Pero, aunque este crecimiento exponencial de recursos dedicados a este pro-

pósito no es sostenible por mucho más tiempo, el ritmo del progreso de la IA podría volver a sorprendernos.

De todos modos, de momento, en la comunidad científica se coincide en que para lograr una IA mucho más inteligente que se asemeje a la inteligencia humana se necesitarán innovaciones disruptivas, cuyas características son inciertas y nadie sabe si realmente se producirán algún día.

2 Que no estemos cerca de disponer de una IA próxima a la inteligencia humana no quita que el actual progreso de la IA esté generando nuevos desafíos éticos que debemos abordar. Por un lado, tenemos desafíos en su desarrollo, como la tendencia de los algoritmos de IA a heredar sesgos de los datos de entrenamiento, lo que puede llevar a decisiones discriminatorias o incorrectas. Esto, a su vez, plantea la cuestión de la responsabilidad en caso de fallos de una IA. ¿Deberíamos considerar al creador del algoritmo o al usuario como responsable? Es esencial establecer marcos legales y éticos que traten la cuestión de los errores o daños causados por la IA.

Por otro lado, encontramos desafíos relacionados con su uso. La automatización impulsada por la IA puede llevar a la reestructuración y pérdida de empleos en diversos sectores económicos, lo que plantea interrogantes éticos sobre cómo garantizar una transición justa hacia una economía basada en la IA. Además, existe la preocupación sobre el uso indebido de la IA, como su aplicación en ciberseguridad, la producción de noticias falsas o el desarrollo y uso de armas totalmente autónomas en conflictos militares, lo cual genera inquietudes profundas.

La humanidad se beneficiará enormemente de la IA actual, a pesar de ser limitada, pero no hay duda de que su uso presenta muchos riesgos. Por ello, al igual que ha ocurrido en otras áreas importantes en nuestra vida, es necesario establecer un marco regulador consensuado, en el más amplio sentido del término: leyes, reglas, protocolos de actuación, autorregulación o sensibilización ciudadana, entre otras. A nadie le gusta ser regulado, pero todo lo que puede entrañar un potencial peligro para la sociedad lo debe

de estar. Espero que las aplicaciones específicas de IA estén también reguladas y, en algunos casos, prohibidas si fuese necesario para poder así mitigar sus riesgos.

3 El avance de la IA tiene importantes implicaciones en la privacidad y la seguridad de los datos personales. Esto se debe a que la IA depende en gran medida del análisis de datos para funcionar de manera efectiva lo que resulta en la recopilación masiva de información personal. Por ejemplo, los algoritmos en que se basan los motores de búsqueda en Internet, o simplemente nuestro teléfono móvil, conocen muy bien, gracias al acceso a cantidades masivas de información que generamos, nuestras preferencias y pueden llegar a inferir cómo pensamos.

Esto amplía sustancialmente el poder y el alcance de las empresas o gobiernos que controlan estos datos y pueden socavar nuestras expectativas de privacidad personal. La capacidad de la IA para analizar grandes conjuntos de datos conlleva un riesgo potencial de violación de la privacidad, ya que la información podría ser utilizada de manera in-

Lo desafíos éticos requieren una colaboración entre gobiernos, empresas y desarrolladores



apropiada. Por ejemplo, los sistemas de IA que utilizan el reconocimiento facial pueden identificar a las personas de manera precisa, una tecnología útil en ciertos contextos, pero que plantea preocupaciones sobre la invasión de la privacidad y un posible uso indebido. Además, los sistemas de IA son susceptibles a ataques, con el riesgo de robo de datos y la generación de información falsa por parte de algoritmos maliciosos.

En resumen, estos desafíos éticos o de privacidad son abordables, pero requieren una colaboración decidida entre gobiernos, empresas y desarrolladores de IA para establecer regulaciones sólidas que promuevan un uso y desarrollo responsables de la IA. Sin embargo, lo más fundamental es la necesidad de involucrar a la sociedad en su conjunto en la toma de decisiones sobre la IA para garantizar que esta tecnología sea utilizada de manera responsable, beneficiosa y en línea con los valores humanos. Esto implica el compromiso individual de todos nosotros en adquirir un mayor entendimiento de esta revolución tecnológica, de manera que podamos tomar decisiones informadas.



Idoia Salazar

PERIODISTA,
INVESTIGADORA
Y PROFESORA.
PRESIDENTA Y
CODIRECTORA DEL
ÁREA DE ÉTICA Y
RESPONSABILIDAD
DE ODISEIA

“DEBERÍA DE HABER MÁS HERRAMIENTAS QUE CREEN CONCIENCIA EN EL USUARIO”

1 Yo diría que la inteligencia artificial está teniendo ya, y tendrá aún más en el futuro, un avance exponencial. Cada vez son más las empresas que apuestan por esa tecnología porque no es una opción no usarla. Hay muchísima investigación, muchísima inversión y en los próximos años vamos a ver un aumento significativo en el uso y desarrollo de estos sistemas de inteligencia artificial.

A nivel tecnológico posiblemente se siga avanzando bastante para intentar hacer bancos públicos para que sea más fácil, no solamente para las grandes empresas de los datos, sino también para pymes y poder optar a esta tecnología para usarla para la eficiencia de procesos y otro tipo de cuestiones.

También veo un aumento bastante grande en la implementación de IA generativa a todos los niveles y es algo que va a suponer una revolución también a nivel social que tendremos que aprender a sobrellevar para que el impacto sea positivo.

A nivel de regulación estamos asistiendo también a un avance bastante significativo. Se ha visto la ne-

cesidad de que, a la vez que evoluciona la tecnología, también aumentemos significativamente el esfuerzo para hacer una regulación competente y poder prevenir los riesgos que vienen asociados a esta tecnología. Cada vez hay más empresas que están dispuestas a adoptar estas normas éticas y regulatorias, una vez que entre en vigor la regulación de inteligencia artificial para intentar prevenir posibles impactos negativos.

2 Los desafíos éticos que vienen asociados al avance de esta tecnología son los mismos que tenemos que tener asociados a la ética humana. Uno de los principales es la supervisión humana. Como sabemos, esta tecnología puede actuar de manera autónoma y los humanos siempre tendemos a tener confianza excesiva en ella. Debemos poner ciertas medidas técnicas y de supervisión humana, sobre todo en casos en el que la acción de la IA pueda afectar a los derechos fundamentales. Puede causar un problema si nosotros no realizamos esa supervisión adecuada.

Por otro lado, hay cuestiones relativas a posibles

casos de discriminación si no se ha hecho un eficiente uso, limpieza y revisión de los datos, algo que en temas de transparencia es muy importante.

También las cuestiones de la explicabilidad son fundamentales en la lógica algorítmica de ver cómo se ha tomado la decisión. La transparencia en cuestiones como la IA generativa es muy importante para que se vea que el que ha hecho el output es la máquina y no la persona.

Por supuesto también hay motivos específicos de ciberseguridad que son fundamentales. Es muy importante que todas las empresas empiecen a adoptar estos criterios éticos, que al final serán parecidos a los que van a implementar la regulación de inteligencia artificial, para que consigamos tener desde el diseño un uso adecuado de estos sistemas de inteligencia artificial.

3 En este caso, tenemos que saber que, a pesar de que los sistemas de inteligencia artificial funcionan con datos, la mayoría de los datos que usa son personales, por lo que, obviamente, puede llegar a suponer un



Óscar Corcho

CATEDRÁTICO
DE LA UPM EN EL
DEPARTAMENTO
DE INTELIGENCIA
ARTIFICIAL
Y DIRECTOR
ACADÉMICO DEL
MÁSTER EUROPEO
AI4GOV (ARTIFICIAL
INTELLIGENCE FOR
PUBLIC SERVICES)

“EL PRINCIPAL DESAFÍO ÉTICO ES EL POSIBLE SESGO DE LOS MODELOS QUE SE GENERAN”

1 En estos últimos años se han producido avances muy considerables en la inteligencia artificial, sobre todo en las áreas de aprendizaje automático y en todo lo relacionado con el lenguaje, debido a que cada vez hemos tenido más datos y documentos disponibles para poder entrenar modelos, y que ha sido posible hacerlo con las infraestructuras tecnológicas de las que disponemos.

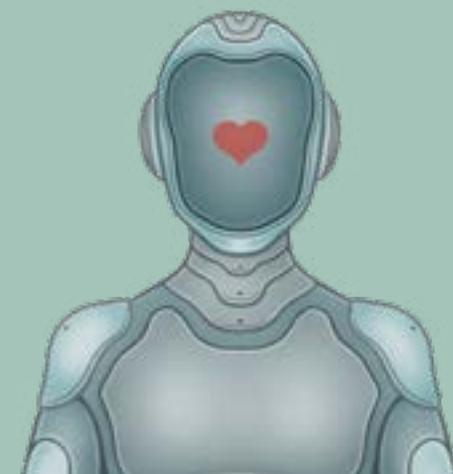
Quizás el avance más significativo de estos últimos años haya sido la aparición de la IA generativa, con los grandes modelos de lenguaje como uno de los pilares básicos.

¿Qué veremos en el futuro inmediato? Bueno, lo vemos ya... Una gran cantidad de aplicaciones que hacen uso de estos modelos y de estas capacidades para poder hacer nuestro trabajo aún más eficiente. Pero no debemos olvidarnos también del “Good old AI”. Se siguen utilizando técnicas tradicionales para resolver una gran cantidad de problemas que ya eran conocidos.

2 El principal desafío ético es el del posible sesgo de los modelos que se generan, basado fundamentalmente en el sesgo que siempre

puede existir en los datos que se utilizan para el entrenamiento. Necesitamos investigación en técnicas cada vez más robustas para explicar los resultados obtenidos por modelos que actúan como cajas negras, y una buena regulación que no necesariamente frene los avances, pero que nos genere seguridad en que no se producen estos sesgos.

3 No creo que este sea un problema que se vea agravado por la IA puesto que ya éramos conscientes de los posibles riesgos en este sentido cuando se puso el foco en la regulación de protección de datos, por ejemplo. Por supuesto, si cualquiera de los modelos que usamos se ha entrenado con datos personales, o los usuarios no son conscientes de que están entregando datos personales al interactuar con sistemas como ChatGPT, podemos poner en riesgo esa privacidad.



CUESTIONARIO

1. ¿Cómo describiría el estado actual de la inteligencia artificial y qué avances considera que veremos en el futuro inmediato?

2. ¿Qué desafíos éticos ve asociados al avance de la IA y cómo cree que debemos abordarlos? 3. ¿Cómo afecta la IA a la privacidad y la seguridad de los datos de las personas?



Elena Gil
Lizasoain

DIRECTORA DE LA
UNIDAD DE NEGOCIO
DE INTELIGENCIA
ARTIFICIAL Y BIG
DATA EN TELEFÓNICA
TECH, ES LICENCIADA
EN CIENCIAS
ECONÓMICAS
Y EMPRESARIALES EN
LA UNIVERSIDAD DE
DEUSTO Y MBA POR
EL MIT SLOAN SCHOOL
OF MANAGEMENT

“LA IA TODAVÍA CARECE DE LA VERDADERA COMPRENSIÓN HUMANA”

1 La inteligencia artificial se encuentra en una fase de aceleración y democratización, impulsada recientemente por los avances de la IA generativa y modelos de lenguaje de gran escala (LLM). Es una realidad y una parte esencial de la actividad diaria de las personas, así como un motor de innovación en muchos sectores. En el futuro más inmediato, veremos avances significativos en la aplicación de la IA en ámbitos como la medicina personalizada de precisión, automatización inteligente de tareas o el desarrollo sostenible de la movilidad y las ciudades. Sin embargo, aún queda mucho por investigar y desarrollar ya que, a pesar de la impresionante capacidad de la IA para detectar patrones y automatizar tareas específicas en un dominio concreto, todavía carece de la verdadera comprensión humana.

Como ha ocurrido hasta ahora, adoptaremos nuevas herramientas de IA en nuestro día a día de manera natural. La inteligencia artificial lleva muchos años estando presente en nuestras vidas como cuando usamos un GPS, utilizamos un asistente virtual o compramos en una tienda

online. La principal diferencia, ahora, es que se están amplificando sus capacidades y la formación es una herramienta esencial para adaptarse a esta nueva realidad. Aparecen nuevos perfiles con formación específica en esta tecnología y los empleos preexistentes se verán transformados. Es una tendencia imparable y, como sociedad, debemos prepararnos para extraer el máximo valor de la IA de manera responsable, adaptando la regulación y la fuerza de trabajo.

2 Como toda tecnología disruptiva, la adopción de la IA plantea desafíos que deben ser abordados. Existen desafíos en el desarrollo de las herramientas de IA como el respeto a la privacidad, la protección de la propiedad intelectual de los datos de entrenamiento, la eliminación de sesgos o, incluso, el consumo energético asociado.

Adicionalmente, se plantean retos éticos relacionados con la desigualdad, desinformación, suplantación de la identidad o el fraude. Resolverlos requiere un enfoque multidisciplinario y es responsabilidad de todos. Es necesario desarrollar un

marco que promueva, a nivel global, el desarrollo de la IA responsable y que ponga un énfasis especial en la transparencia y la rendición de cuentas.

3 La inteligencia artificial tiene múltiples efectos para la privacidad y seguridad. Por un lado, ha incrementado la vulnerabilidad por la recopilación y almacenamiento de más información sobre las personas, como el robo de datos y la suplantación de la identidad. Sin embargo, también se ha producido una mayor concienciación de la importancia de proteger la privacidad de la información y, tanto las empresas como los gobiernos, están invirtiendo y utilizando de manera creciente IA para protegerse. Esto incluye el cifrado de datos, la autenticación de usuarios y el monitoreo de redes y sistemas para detectar posibles amenazas. La normativa europea está siendo pionera en materia de protección de la privacidad y seguridad de los datos de las personas.