

Elementos para una cultura europea de desarrollo de herramientas de inteligencia artificial: el *Libro blanco sobre la inteligencia artificial* y las *Directrices éticas para una IA fiable*

A European Union culture for developing of the AI-based products: *The White Paper on Artificial Intelligence* and the *Ethics Guidelines for Trustworthy AI*

VALENTINA COLCELLI¹  Y LAURA BURZAGLI² 

RESUMEN

El objetivo del trabajo es abordar el problema ético/jurídico del desarrollo de las tecnologías que utiliza la inteligencia artificial, utilizando la vía metodológica trazada por la Unión Europea en sus *Directrices éticas para una IA fiable*. Así, el artículo se centra en la descripción de algunos elementos que caracterizan la estrategia elegida por la Unión Europea para abordar la cuestión de la relación entre los aspectos ético-jurídicos y la tecnología relacionada con el desarrollo de la inteligencia artificial. Seguidamente, expone un ejemplo práctico de la sinergia entre desarrolladores de aplicaciones tecnológicas y juristas, que se ha llevado a cabo en un instituto del Consejo Nacional de Investigación de Italia.

Palabras clave: Inteligencia artificial fiable, Unión Europea, Libro blanco sobre la inteligencia, Directrices éticas para una inteligencia artificial fiable, aspectos ético-jurídicos de la inteligencia artificial, datos.

¹ Investigadora del Consejo Nacional de Investigación de Italia IFAC Institute - Via Madonna del Piano, 10 50019 Sesto Fiorentino, Florencia (Italia) Correo: valentina.colcelli@cnr.it

² Investigadora del Consejo Nacional de Investigación de Italia IFAC Institute - Via Madonna del Piano, 10 50019 Sesto Fiorentino, Florencia (Italia) Correo: l.burzagli@ifac.cnr.it

ABSTRACT

The article goal is to explain the European Union approach to artificial intelligence based on principles of democracy, the rule of law and fundamental rights incorporated on the development of the AI-based products. So, the article is focused on the description of the main elements of the European Union strategy lunched to face the issues of the legal and ethical issues related to IA. Thus, the framework mentioned above will be used to describe a scientific experiment realized in an Institute of the Italian National Research Council (CNR), that is an application requiring the adoption of innovative Artificial Intelligence techniques, where the ethical component to be faced is not limited to the aspect of privacy, but, much more important, to the intelligibility of the reasoning process.

Keywords: Trustworthy artificial intelligence, European Union, White paper on artificial intelligence, Ethics guidelines for trustworthy artificial intelligence, legal and ethical issues related to artificial intelligence, data.

1. Introducción

El artículo se centra en la descripción de algunos elementos que caracterizan la estrategia elegida por la Unión Europea (UE) para abordar la cuestión de la relación entre los aspectos ético-jurídicos y la tecnología relacionada con el desarrollo de la inteligencia artificial (IA). Seguidamente, expone un ejemplo práctico de la sinergia entre desarrolladores de aplicaciones tecnológicas y juristas, que se ha llevado a cabo en un instituto del Consejo Nacional de Investigación de Italia (CNR).

El objetivo es abordar el problema ético/jurídico del desarrollo de las tecnologías que utiliza la IA, utilizando la vía metodológica trazada por la UE en sus *Directrices éticas para una IA fiable*.³ Debería desarrollarse una nueva cultura para esta última, precisamente a partir del documento mencionado, redactado por un grupo de expertos de alto nivel sobre IA (AI HLEG) creado por la Comisión Europea.⁴ La IA requiere una nueva metodología y necesita, en gran medida, un enfoque multidisciplinario de las cuestiones críticas o de los posibles riesgos asociados a su desarrollo y aplicación.

Desde el punto de vista del análisis jurídico, es habitual analizar los efectos de la acción de las tecnologías de IA adoptando un enfoque reactivo, que a menudo considera las medidas resarcitorias o las de cesación como los instrumentos que deben utilizarse en los supuestos de vulneración de los derechos de las partes interesadas.⁵ Cabe recordar que los sistemas de IA ya funcionan en un mundo de leyes y que, en varios niveles de la UE y de los ordenamientos de los Estados miembros, ya son aplicables a la IA, si bien sigue siendo necesario construir una normativa específica (Comisión Europea, 2020, p. 10) —como recuerda el *Libro blanco sobre la inteligencia artificial— también dirigida a crear certificaciones de calidad en la materia* (Comisión Europea, 2019).

De hecho, a los sistemas de IA les son de aplicación las normas de los tratados internacionales sobre derechos humanos, los tratados fundacionales de y de su funcionamiento, las normas de la Carta de Derechos Fundamentales o algunas normas de derecho, derivados todos de la UE, como obviamente

³ “Esta es la razón por la que, el 25 de abril de 2018, la Comisión estableció una estrategia sobre IA que abordaba los aspectos socioeconómicos junto con un aumento de la inversión en investigación, innovación y capacidad en materia de IA en toda la UE”. Para más información, véase Comisión Europea (2020, p. 10).

⁴ La Comisión creó, además, un grupo de expertos de alto nivel que, en abril de 2019, publicó *Directrices para una IA fiable*. Para más información, véase Comisión Europea (2019).

⁵ Véase, por ejemplo, la sentencia del Consejo de Estado italiano 2019/2270. El Consejo de Estado italiano abordó por primera vez el uso de algoritmos e IA para la sincronización de medidas administrativas que afectan los derechos de los ciudadanos. Aclara las condiciones bajo las cuales se pueden utilizar los procesos automatizados de toma de decisiones en los procedimientos administrativos. Para un comentario, podemos hacer referencia a Colcelli y Burzagli (2020).

el Reglamento sobre protección de datos,⁶ pero también el relativo a la libre circulación de datos no personales,⁷ las directivas contra la discriminación,⁸ la denominada directiva de máquinas,⁹ la directiva sobre responsabilidad del productor,¹⁰ la directiva sobre derecho del consumidor,¹¹ y las directivas sobre salud y seguridad en el trabajo.¹² Muchos de estos reglamentos apoyan la fase de industrialización del producto y su presencia en el mercado interno (Hacker, 2020).

Por otro lado, también es cierto que ningún sistema jurídico será nunca lo suficientemente reactivo como para adaptar su reglamentación al avance continuo y constante de los sistemas de IA (Hacker, 2018). El recurso a los principios generales y a los derechos fundamentales europeos permite, por tanto, una adaptación constante del ordenamiento, así como un enfoque holístico y sistémico de todos los actores y de todos los procesos pertenecientes al contexto sociotécnico del sistema de IA durante todo su ciclo de vida.

En este contexto, la UE ha considerado necesario establecer una serie de acciones que definan los valores sobre los que basarse para promover el desarrollo y la difusión de la llamada IA «fiable», con el objetivo de garantizar un marco normativo claro, basado en la credibilidad y la excelencia, para hacer frente a este reto, que ya no puede postergarse, y así evitar la fragmentación de la acción legislativa nacional.

La Comisión Europea ha adoptado su propia estrategia sobre IA dentro de la más general estrategia europea de datos. En esta perspectiva dirigida a construir un mercado interno cualificado en lo que respecta a las tecnologías y el uso de los datos, el *Libro blanco sobre inteligencia artificial* y las *Directrices éticas para una IA fiable* son un componente de ella y representan las opciones estratégicas que permiten un desarrollo seguro y fiable de la IA, respetando plenamente los valores y los derechos de los ciudadanos de la UE.

Los principales elementos del *Libro blanco* son la creación de un marco estratégico mediante una asociación entre el sector público y el privado para movilizar recursos. Esto, con el fin de lograr un «ecosistema de excelencia», la creación de un futuro marco reglamentario para la IA en Europa para establecer un «ecosistema de confianza» único, y el desarrollo de una IA que pueda considerarse fiable.

⁶ Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, del 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos).

⁷ Reglamento (UE) 2018/1807 del Parlamento Europeo y del Consejo, del 14 de noviembre de 2018, relativo a un marco para la libre circulación de datos no personales en la UE.

⁸ Directiva 2000/43/CE: Aplicación del principio de igualdad de trato de las personas independientemente de su origen racial o étnico; Informe de la Comisión al Consejo y al Parlamento Europeo - Aplicación de la Directiva 2000/43/CE, del 29 de junio de 2000, relativa a la aplicación del principio de igualdad de trato de las personas independientemente de su origen racial [COM(2006) 643 final del 30 de octubre de 2006]; Recomendación del Consejo, del 9 de diciembre de 2013, relativa a la adopción de medidas eficaces de integración de los gitanos en los Estados miembros (DO C 378 de 24.12.2013, pp. 1-7); Directiva 2000/78/CE del Consejo, del 27 de noviembre de 2000, relativa al establecimiento de un marco general para la igualdad de trato en el empleo y la ocupación («Directiva de igualdad en el empleo»). [COM (2014) 2 final del 17 de enero de 2014].

⁹ La denominada «Directiva de Máquinas» ha sufrido una importante evolución desde que se aprobó la Directiva 89/392/CEE del Consejo, del 14 de junio de 1989, relativa a la aproximación de las legislaciones de los Estados miembros sobre máquinas.

¹⁰ Directiva 85/374/CEE del Consejo, del 25 de julio de 1985, relativa a la aproximación de las disposiciones legales, reglamentarias y administrativas de los Estados miembros en materia de responsabilidad por los daños causados por productos defectuosos.

¹¹ Directiva (UE) 2019/771 del Parlamento Europeo y del Consejo, del 20 de mayo de 2019, relativa a determinados aspectos de los contratos de compraventa de bienes, por la que se modifican el Reglamento (CE) 2017/2394, la Directiva 2009/22/CE, y se deroga la Directiva 1999/44/CE, Directiva 2011/83/UE del Parlamento Europeo, y del Consejo, del 25 de octubre de 2011, sobre los derechos de los consumidores, por la que se modifican la Directiva 93/13/CEE del Consejo y la Directiva 1999/44/CE del Parlamento Europeo y del Consejo, y se derogan la Directiva 85/577/CEE del Consejo y la Directiva 97/7/CE del Parlamento Europeo y del Consejo.

¹² Directiva 89/391/CEE del Consejo, del 12 de junio de 1989, relativa a la aplicación de medidas para promover la mejora de la seguridad y de la salud de los trabajadores en el trabajo.

La idea principal es un algoritmo de IA que incorpora los principios de democracia, Estado de derecho y derechos fundamentales desde el momento de su diseño.

Para que un sistema de IA sea comprensible, predecible y controlable *ex post*, debe ser capaz de garantizar el principio de transparencia y el Estado de derecho *ex ante*, es decir, en el momento del diseño algorítmico.

Las recomendaciones sobre la futura evolución de las políticas relativas a la IA y las cuestiones éticas, jurídicas y sociales que le atañen, incluyen desafíos sociojurídicos relacionados con los requisitos que todo sistema de estas características debe cumplir para ser fiable (Floridi, et al., 2018).

Con el fin de aplicar la metodología concebida por la UE para la creación de una IA fiable, se presenta, en el tercer apartado de este artículo, una actividad multidisciplinar en la que se analiza una aplicación tecnológica de asistencia a un usuario final según aspectos éticos/legales. Se trata de una aplicación llamada *Mai più soli* (Nunca más solos), que se encuentra en fase de desarrollo en el Instituto de Física Aplicada «Nello Carrara» del Consejo Nacional de Investigación de Italia.

Este ejemplo de actividad adquiere un valor particular. Mientras que en el pasado un proyecto de investigación científica asumía como objetivos específicos solo una lista de elementos técnicos, el enfoque más reciente e innovador introduce, desde las etapas iniciales, una descripción más amplia del problema, en la que también se tienen en cuenta los componentes sociales y legales. Desde este punto de vista, la búsqueda de una solución tecnológica resulta a menudo más compleja, ya que puede requerir el desarrollo de nuevas estrategias. Sin embargo, abre las puertas a un producto real en lugar de un prototipo de laboratorio.

2. *Directrices éticas para una IA fiable*

Un marco regulador europeo sólido que garantice una IA fiable protegerá a todos los ciudadanos europeos y contribuirá a crear un mercado interior sin fricciones de cara al desarrollo, la adopción futura de la IA y reforzará los cimientos industriales de Europa en el sector (Comisión Europea, 2020, p. 13).

El primer paso para la creación de un ecosistema de confianza dedicado a la IA consiste en la adopción de un enfoque de ella que identifique una vía metodológica para abordar el problema ético en el ámbito científico y tecnológico, la vía metodológica trazada por la UE en sus *Directrices éticas para una IA fiable*. La vía trazada es un ejemplo de sinergia entre desarrolladores de aplicaciones tecnológicas, juristas y expertos en ética, realizado con la intención de abordar el problema ético y jurídico en el desarrollo de tecnologías que explotan la IA desde el momento de su programación. Así:

La fiabilidad de la IA se apoya en tres componentes que deben satisfacerse a lo largo de todo el ciclo de vida del sistema:

- a) La IA debe ser lícita, es decir, cumplir todas las leyes y reglamentos aplicables.
- b) Ha de ser ética, de modo que se garantice el respeto de los principios y valores éticos.
- c) Debe ser robusta, tanto desde el punto de vista técnico como social, ya que los sistemas de IA, incluso si las intenciones son buenas, pueden provocar daños accidentales (Comisión Europea, 2020, p. 2).

Como se ha indicado anteriormente, la confianza en la IA depende de la seguridad de sus procesos, datos y resultados, así como de un diseño robusto que le permita hacer frente a posibles fallos y ataques malintencionados. Los sistemas de IA deberían incluir un mecanismo de apagado a prueba de fallos que, además, posibilite la reanudación del funcionamiento del sistema tras un apagado forzado (por ejemplo, un ataque).

Además, el testeo ofrece orientaciones sobre cómo lograr una IA fiable por medio de siete requisitos que, idealmente, deben cumplir los sistemas, los cuales serán expuestos a continuación.

- *Acción y supervisión humana.* Sobre este punto, es importante mencionar que:

La supervisión humana ayuda a garantizar que un sistema de IA no socave la autonomía humana o provoque otros efectos adversos. El objetivo de una IA fiable, ética y antropocéntrica solo puede alcanzarse garantizando una participación adecuada de las personas en relación con las aplicaciones de IA de riesgo elevado (Comisión Europea, 2020, p. 25).

- *Solidez técnica y seguridad.* Aquí, se incluye la capacidad de resistencia a los ataques y la seguridad, un plan de repliegue y la seguridad general, precisión, fiabilidad y reproducibilidad.
- *Gestión de la privacidad y de los datos.* En este punto, se considera el respeto de la privacidad, la calidad y la integridad de los datos, así como el acceso a estos.
- *Transparencia.* Con respecto a esto, se incluye la trazabilidad, la explicabilidad (Wachter *et al.*, 2017) y la comunicación.
- *Diversidad, no discriminación y equidad.* Se incluye la ausencia de sesgos injustos, la accesibilidad y el diseño universal, así como la participación de las partes interesadas.
- *Bienestar ambiental y social.* Se considera la inclusión de la sostenibilidad y el respeto del medio ambiente, el impacto social, la sociedad y la democracia.
- *Rendición de cuentas.* Se incluye la auditabilidad, la minimización de efectos negativos y la notificación de estos, la búsqueda de equilibrios y las compensaciones.

Para aplicar estas orientaciones se pueden utilizar tanto métodos técnicos como de otro tipo. Por otra parte, está:

Una serie de métodos técnicos para garantizar la fiabilidad de la IA que se pueden incorporar en las fases de diseño, desarrollo y utilización de un sistema. [...] Los requisitos de una IA fiable deben «traducirse» en procedimientos (o en la imposición de restricciones sobre estos) que deben integrarse en la arquitectura de los sistemas. Esto puede lograrse a través de un conjunto de normas de tipo «lista blanca» (comportamientos o estados) que el sistema debería seguir en todo momento, restricciones («lista negra») sobre determinados comportamientos o estados que el sistema jamás debería transgredir y combinaciones de ambas, o garantías demostrables más complejas sobre el comportamiento del sistema. El control del cumplimiento de dichas restricciones por parte del sistema durante su funcionamiento puede llevarse a cabo a través de un proceso separado. [...] La idea de que el cumplimiento de las normas puede introducirse en el diseño del sistema de IA es clave en este método. Las empresas son responsables de identificar los efectos de sus sistemas de IA desde el principio, así como las normas que deberían cumplir dichos sistemas para evitar efectos negativos (privacidad desde el diseño o la seguridad desde el diseño). [...] Para que un sistema sea fiable,

hemos de ser capaces de comprender por qué se comportó de una determinada manera y por qué ofreció una interpretación específica (IA explicable). [...] Para verificar y validar el tratamiento de los datos, debe llevarse a cabo un seguimiento minucioso de la estabilidad, solidez y funcionamiento del modelo subyacente, tanto durante la formación como durante el despliegue, dentro de unos límites predecibles y correctamente entendidos. [...] El ensayo y la validación del sistema deben tener lugar lo antes posible, garantizando que el sistema se comporte según lo previsto a lo largo de todo su ciclo de vida y, en especial, tras su despliegue. Dicha fase de ensayo y validación debería abarcar todos los componentes de un sistema de IA, incluidos los datos, los modelos previamente formados, los entornos y el comportamiento del sistema en su conjunto. Su diseño y ejecución deberían correr a cargo de un grupo de personas lo más diverso posible. Se deberían desarrollar múltiples parámetros que engloben las categorías sometidas a verificación, con el fin de obtener diferentes perspectivas. Existe la posibilidad de definir indicadores adecuados de calidad del servicio para los sistemas de IA con objeto de garantizar que sea posible saber si dichos sistemas se han ensayado y desarrollado teniendo presentes las consideraciones relativas a la seguridad. Estos indicadores podrían incluir parámetros para evaluar los ensayos realizados y la formación de algoritmos, así como los parámetros tradicionales de funcionalidad del *software*, *su rendimiento*, *usabilidad*, *fiabilidad*, *seguridad* y *facilidad de mantenimiento* (Comisión Europea, 2020, p. 26).

También existen métodos no técnicos para garantizar y mantener la fiabilidad de la IA. Estos métodos también deberían someterse a una evaluación constante en cuanto a normativa (por ejemplo, en la legislación sobre seguridad de los productos y en los marcos de responsabilidad) y códigos de conducta.

Las organizaciones y partes interesadas pueden adoptar estas directrices y adaptar sus cartas de responsabilidad empresarial, sus indicadores clave de rendimiento (KPI), sus códigos de conducta o sus documentos internos de política para contribuir a los esfuerzos conducentes a la creación de una IA fiable.

- *Normalización*. Las normas, como las relativas al diseño, la fabricación o las prácticas empresariales, pueden funcionar como un sistema de gestión de la calidad para los usuarios de la IA, los consumidores, las organizaciones, las instituciones de investigación y los gobiernos, ofreciendo a todos estos agentes la capacidad de reconocer y fomentar una conducta ética a través de sus decisiones de compra. [...] Entre los ejemplos existentes en la actualidad se encuentran, por ejemplo, las normas ISO o la serie de normas IEEE P7000, aunque en el futuro podría resultar adecuado crear un sello de «IA fiable» que, a partir de las normas técnicas especificadas, confirme, por ejemplo, que el sistema cumple los requisitos de seguridad, solidez técnica y explicabilidad.
- *Certificación*. Puede tenerse en consideración a aquellas organizaciones que puedan acreditar ante el público que un sistema de IA es transparente, responsable y equitativo. Así:

La etiqueta voluntaria permitirá a los agentes económicos interesados mostrar que los productos y servicios provistos de IA que ofrecen son fiables. Además, permitirá a los usuarios distinguir fácilmente si los productos y servicios en cuestión respetan ciertos referentes objetivos y normalizados a escala de la UE, que van más allá de las obligaciones legales aplicables normalmente. Ello contribuirá a incrementar la confianza de los usuarios en los sistemas de IA y fomentará una adopción generalizada de esta tecnología. Esta opción conlleva la creación de un nuevo instrumento jurídico para establecer un marco de etiquetado voluntario para los desarrolladores y/ o implementadores de los sistemas de IA que no se consideren de alto riesgo. Si bien la participación en el sistema de etiquetado debe ser voluntaria, una vez que el desarrollador o implementador opte por usar la etiqueta, todos los requisitos

serán vinculantes. La combinación de estas imposiciones *ex ante* y *ex post* debe garantizar que se cumplan todos los requisitos (Comisión Europea, 2020, p. 29).

Además, las Directrices facilitan una lista para que las empresas comprueben, en la práctica, si se cumplen los requisitos. Se trata de una lista no exhaustiva para la evaluación de la fiabilidad de la IA con el fin de poner en práctica un mecanismo fiable.

3. Desarrollo de la aplicación *Nunca más solos* a la luz de las Directrices éticas para una IA fiable

La construcción de una IA fiable (lícita, ética y robusta) conlleva la necesidad de incorporar elementos éticos/jurídicos desde el momento en que se diseña el algoritmo, vigilando su desarrollo para evitar distorsiones de distinta naturaleza en el momento operativo y final. La investigación para el desarrollo de la aplicación *Mai più soli* partió del análisis de un problema social, con la identificación de los aspectos críticos de esta condición (Mitrou, 2018). Los grandes cambios que se han producido en nuestra sociedad en el ámbito social y tecnológico en los últimos años, han provocado nuevas necesidades de los usuarios y, por lo tanto, han dado lugar a la ampliación de la búsqueda de servicios de asistencia a los usuarios, por ejemplo, los destinados a las personas mayores o, más en general, a las personas con limitaciones de la actividad, en los que el componente crítico está representado por la red de relaciones humanas, que se hace cada vez más débil y limitada.

En este contexto, el diseño de un servicio relacionado con la soledad adquiere especial relevancia para la calidad de vida de las personas mayores, aunque el problema no se limite a ellas.

Mai più soli es un sistema basado en técnicas de IA que relaciona perfiles de usuarios con sugerencias; un algoritmo específico selecciona la mejor sugerencia entre un conjunto de opciones relacionadas con las diferentes actividades de la red de personas y, tras la elección del usuario, un *machine learning* tiene en cuenta también la reacción (confirmación o rechazo) del usuario para mejorar la calidad del servicio.

Para la primera implementación de la aplicación *Mai più soli* se ha elegido un contexto específico: una comunidad de vecinos, ya que representa un límite espacial en el que se mapean las relaciones. El proyecto prevé que un conjunto de herramientas tecnológicas adquiera los datos de un grupo de personas y, a cambio, ofrezca una sugerencia al usuario, sugerencia que estará estrechamente relacionada con la condición de las otras personas del grupo.¹³

¹³ Las personas contratadas para la actividad del proyecto/investigación deben ser informadas sobre el objeto de la investigación, su finalidad, el uso de los datos personales recogidos, su conservación y protección. En caso de entrevistas o participación en eventos:

- La persona será informada con antelación de las actividades de investigación y de los fines científicos de la entrevista y del tratamiento de los datos personales; se le invitará a evitar comentarios ofensivos o discriminatorios.
- En particular, se informará a la persona de que su entrevista puede ser publicada o divulgada al público, si procede.
- Se informará a la persona de que tendrá derecho a retirar su consentimiento, siempre que ello no comprometa las actividades de investigación ya realizadas.
- La persona interesada debe dar su consentimiento expreso por escrito a la recogida y el tratamiento de datos personales y a la publicación de los datos o de los resultados obtenidos a partir de ellos.

Para ello, los investigadores proporcionarán la hoja de informativa a la persona participante. La persona leerá y, si está de acuerdo, firmará el formulario de consentimiento informado. La hoja informativa y los formularios de consentimiento informado deben ser traducidos a la lengua materna de las personas que participan en el proyecto. Asimismo, con arreglo al artículo 10, apartado 2, del Anexo 1 de las Normas éticas para el tratamiento con fines estadísticos o de investigación científica, "el personal responsable de la recogida deberá cumplir las disposiciones contenidas en estas normas de conducta y las instrucciones recibidas. En particular, deberá:

- a) Revelar su identidad, su función y el propósito de la recogida, también mediante la documentación apropiada.

La idea innovadora es el estudio de una herramienta automática que toma una decisión sobre la mejor sugerencia para un usuario específico, decisión que depende de lo que estén haciendo todas las personas de un grupo específico, en este caso, la comunidad de vecinos.

Por lo tanto, el planteamiento consiste en considerar no a una sola persona, sino a un grupo de usuarios. En otras palabras, se trata de diseñar un servicio de asistencia en el ámbito colectivo, aunque cada sugerencia se refiera solo a un usuario específico. Este enfoque se utiliza porque el usuario recibe una sugerencia que tiene en cuenta todas las posibilidades disponibles dentro del grupo.

Para crear un primer grupo de personas, se han identificado siete perfiles diferentes: una señora mayor, una pareja de ancianos, una pareja con un niño, una pareja de ancianos con un nieto joven, una chica, un joven médico y un señor mayor. A cada uno le corresponde un perfil estático, en el que sus preferencias se establecen con un valor inicial, por ejemplo, si les gusta o no andar, si les gusta la música, si tienen enfermedades específicas o limitaciones de la actividad.

Se han creado una serie de condiciones iniciales, por ejemplo, si están trabajando en la cocina o viendo la televisión. Cuando la aplicación esté operativa, se recibirá información sobre la condición del usuario a través de un conjunto de sensores y se procesará dicha información para obtener una sugerencia apropiada.

Se ha creado también una lista de sugerencias con el fin de seleccionar, por el momento dentro de un conjunto limitado pero definido, la mejor sugerencia en cada condición.

También hay disponible una interfaz para tabletas, que permite al usuario recibir una sugerencia basada en su condición y que puede aceptar o rechazar. Tras un número limitado de sugerencias proporcionadas por el sistema, el usuario puede seleccionar la sugerencia entre todas las ideas que están almacenadas en el sistema.

En ese contexto, el sistema de IA puede entrañar cuestiones éticas con respecto a la participación de personas y, por supuesto, cuestiones jurídicas relacionadas con el tratamiento de datos personales y la utilización de datos no personales.

4. Conclusión: Reflexiones sobre los requisitos éticos/legales para iniciar el desarrollo de *Nunca más solos*

Las normas, en materia de privacidad y ética, deben incorporarse desde el momento del diseño, por lo que los desarrolladores de la herramienta automática de la aplicación *Nunca más solos* deben ser conscientes, en primer lugar, de las cuestiones jurídicas y éticas que se derivan de la aplicación, y deben encontrar una solución tecnológica que respete los derechos fundamentales de los usuarios.

b) Proporcionar la información contemplada en el artículo 13 del Reglamento y en el artículo 6 del presente código, así como cualquier otra aclaración que permita al interesado responder de manera adecuada e informada, evitando todo comportamiento que pueda adoptar la forma de artificios y presiones indebidas.

c) No realizar, al mismo tiempo, para los mismos interesados actividades de recopilación de datos personales por cuenta de más de un responsable, salvo autorización expresa.

d) Corregir con prontitud los errores e inexactitudes de la información obtenida durante la recogida.

e) Asegurar una diligencia particular en la recopilación de las categorías especiales de datos previstas en el artículo 9, apartado 1, y de los datos previstos en el artículo 10 del Reglamento”.

A continuación, resumimos los aspectos éticos/jurídicos que se han tenido en cuenta en la fase inicial de desarrollo del proyecto.

En primer lugar, el desarrollador debe ser consciente de la naturaleza jurídica de los datos que tanto la herramienta como los sensores recogen y procesan. La herramienta automática *Nunca más solos* puede almacenar y tratar sistemas complejos de datos, no solo personales, sino también diferentes informaciones como, por ejemplo, las relacionadas con diversos tipos de hábitos.¹⁴ Este aspecto, en primer lugar, subraya la diferencia entre la garantía de la protección de los datos personales y la de la privacidad, habida cuenta de que la primera está incluida en la segunda, pero que esta va más allá de la protección asegurada de los datos personales. La cuestión planteada conlleva naturalmente la aplicación del Reglamento (UE) 679/2016 sobre el tratamiento de datos, si bien los desarrolladores deben tener en cuenta también los llamados datos no personales cuando estos tienen un carácter mixto. La UE esboza un marco para la libre circulación de los datos no personales (Reglamento (UE) 2018/1087).¹⁵ El Reglamento califica también los llamados datos mixtos, compuestos por datos personales y no personales. Para este tipo de datos, la Comunicación de la Comisión Europea «Orientaciones sobre el Reglamento relativo a un marco para la libre circulación de los datos no personales en la UE» recuerda que: a) el Reglamento de libre circulación de datos no personales se aplica a la parte de datos no personales del conjunto del datos; b) la disposición de libre circulación del Reglamento general de protección de datos se aplica a la parte de datos personales del conjunto de datos; c) si las partes de datos personales y de datos no personales están «inextricablemente ligadas», los derechos y las obligaciones de protección de datos derivados del Reglamento general de protección de datos se aplicarán completamente a todo el conjunto de datos mixtos,¹⁶ incluso cuando los datos personales representen solo una pequeña parte del conjunto de datos. Los conjuntos de datos mixtos representan la mayoría de los conjuntos de datos utilizados en la economía de datos y son los más comunes de entre los que se derivan de actividades tecnológicas, como el internet de las cosas (por ejemplo, objetos que se conectan digitalmente), la IA y las tecnologías que permiten el análisis de macrodatos. Por lo tanto, la mayoría de los datos reunidos y tratados en la aplicación *Nunca más solos* pertenecerán a la categoría c).

El peligro de que la información —más aún los datos mencionados anteriormente— pueda ser conocida por los miembros del grupo que interactúan entre ellos (la comunidad de vecinos) abre la primera cuestión relacionada con la robustez del sistema, que también surge con respecto a la cuestión de dónde se almacenan los datos, por cuánto tiempo en cumplimiento del principio de minimización y con respecto a la técnica de seudonimización adoptada. La aplicación del Reglamento (UE) 679/2016 remite a la identificación del responsable del tratamiento de los datos personales y, por lo tanto, a la información clara y sencilla que debe facilitarse a los usuarios del producto, de modo que los interesados den su consentimiento al tratamiento de sus datos personales¹⁷ para las finalidades específicas del uso de su información dentro

¹⁴ Los datos personales deben recopilarse para los fines previstos en el contexto de actividades de investigación específicas que se refieran “a un proyecto de investigación establecido con arreglo a las correspondientes normas metodológicas y éticas relacionadas con el sector, de conformidad con las buenas prácticas”. Para más información, véase Comisión Europea (2016, p. 27).

¹⁵ Actualmente, esta dimensión debe considerarse reforzada por la Estrategia Europea de Datos publicada por la Comisión Europea el 19 de febrero de 2020.

¹⁶ Dicho concepto es el resultado de una amplia reflexión llevada a cabo por la Comisión Europea para profundizar en el ámbito de aplicación del Reglamento General de Protección de Datos y del Reglamento europeo sobre la libre circulación de datos no personales, en el contexto de la llamada Estrategia Europea de Datos. Reglamento (UE) 2018/1807 del Parlamento Europeo y del Consejo, del 14 de noviembre de 2018, relativo a un marco para la libre circulación de datos no personales en la UE (DO L 303 del 28 de noviembre de 2018, p. 59).

¹⁷ El consentimiento debe ser libre, específico, informado e inequívoco en todos los casos; no se permite el consentimiento tácito o presunto (las casillas ya marcadas en un formulario no deben constituir consentimiento); debe expresarse mediante una «declaración o acción positiva inequívoca» (para más información, véanse los considerandos 39 y 42 del Reglamento), y el consentimiento debe poder demostrar que el interesado ha dado su consentimiento a una operación de tratamiento determinada. El consentimiento no tiene que ser necesariamente «documentado por escrito»; la «forma escrita» no es necesaria, si bien es una forma idónea de configurar el carácter inequívoco y «explícito» del consentimiento en el caso de los datos denominados «sensibles» en la anterior Directiva 95/46/CE. Para las categorías especiales de datos

del grupo y que si se espera que comparta alguna información (incluso datos que no sean de carácter personal), esta debe ser debidamente identificada.¹⁸

En todo caso, se ha tenido que prestar especial atención a la información que pudiera contener categorías especiales de datos personales¹⁹ con respecto a la autorización para el tratamiento.

Entre los ocho perfiles establecidos hay una pareja con un niño. Los desarrolladores deberán asegurarse de que la información de la pareja no incluya datos relativos al menor, a menos que —en su caso— tomen las medidas adecuadas para garantizar que los padres den el consentimiento apropiado como sujetos que ejercen la responsabilidad parental.

El sistema utiliza sensores para recoger información y datos. Los desarrolladores tendrán que demostrar/evaluar que las actividades de la herramienta no son invasivas y no producirán ningún tipo de dolor o estrés a los participantes.

Este aspecto pone de relieve el hecho de que, si la herramienta se dirige a personas mayores, los desarrolladores tendrán que definir quién es una «persona mayor» también por referencia a las normas jurídicas y si esta o las personas con movilidad limitada también pueden ser personas vulnerables. Estas últimas deberán definirse asimismo por referencia a la legislación europea.

El ejemplo que se ha presentado sigue siendo reactivo más que proactivo, pero representa un paso en la dirección adecuada hacia una definición más correcta de los requisitos técnicos, cuyo enfoque multidisciplinario del problema (Leikas, *et al.*, 2019, p. 18) muestra cómo los conocimientos tecnológicos por sí solos ya no son suficientes para abordar el desarrollo de las herramientas de IA.

Bibliografía citada

Cippitani Roberto (2018): “Genetic research and exceptions to the protection of personal data”, en Cippitani Roberto, Arnold Rainer, Colcelli Valentina (editores), *Genetic information and individual rights. Serie Law & Science, vol 1.* (Regensburg, Universität Regensburg) pp. 54-79. <https://isaac.ifac.cnr.it/wp-content/uploads/2021/02/Genetic-Information-OK-1.pdf>

personales (artículo 9 del Reglamento General de Protección de Datos), el consentimiento debe ser «explícito». El Reglamento General de Protección de Datos contiene algunas novedades relativas a los datos sensibles que ahora se definen como «categorías especiales de datos personales» y el tratamiento automatizado. El consentimiento debe ser también «explícito» para las decisiones basadas en este último tipo de tratamiento (incluida la elaboración de perfiles; artículo 22 del Reglamento General de Protección de Datos). El consentimiento al tratamiento de datos debe darse generalmente para todas las actividades de tratamiento realizadas con el mismo o los mismos fines. Cuando el tratamiento tenga varios fines, debe darse el consentimiento al tratamiento de datos personales para todos ellos (considerando 32 del Reglamento General de Protección de Datos). Con frecuencia, no es posible determinar totalmente la finalidad del tratamiento de los datos personales con fines de investigación científica en el momento de su recogida. Por consiguiente, debe permitirse a los interesados dar su consentimiento solamente para determinadas áreas de investigación o partes de proyectos de investigación, en la medida en que lo permita la finalidad perseguida (considerando 33 del Reglamento General de Protección de Datos). Para más información, véase Cippitani (2018).

¹⁸ La información debe cumplir la legislación en materia de protección de datos, incluida la proporcionalidad y la limitación de la finalidad claramente detallada, así como no poner en manos del responsable de tratamiento un poder discrecional indebido sobre cómo cumplir la obligación legal - Dictamen 3/2019 sobre las preguntas y respuestas acerca de la relación entre el Reglamento sobre ensayos clínicos y el Reglamento general de protección de datos.

¹⁹ Datos personales relativos al origen racial o étnico, opiniones políticas, convicciones religiosas o filosóficas o afiliación sindical, y el tratamiento de datos genéticos, datos biométricos para la identificación de una persona física, datos relativos a la salud o datos relativos a la vida sexual o la orientación sexual de una persona física. A la expresión «datos relativos a la salud» se le dará una interpretación amplia de modo que comprenda «la información relativa a todos los aspectos de la salud de una persona» según lo determinado desde hace tiempo por el Tribunal de Justicia de la UE: Tribunal sobre la Directiva 95/46/CE, TJUE del 6 de noviembre de 2003, C-101/01 (Lindqvist), apartado 50.

- Colcelli Valentina, Burzagli Laura (2020): “Public administration and technology in the time of artificial intelligence: The Italian case of Council State judgment n. 2019/2270”, en Arnold, R y Danėlienė (editores), *The Concept of Democracy as developed by Constitutional Justice* (Vilnius, Constitutional Court of the Republic of Lithuania) pp. 137-148. <https://www.lrkt.lt/data/public/uploads/2020/10/the-concept-of-democracy-web.pdf>
- Comisión Europea (2016): *Directrices del antiguo Grupo de Trabajo del Artículo 29, del 10 de abril de 2018, sobre el consentimiento en el sentido del Reglamento (UE) 2016/679, WP259 rev. 1, 17*. bit.ly/3dTUeNI.
- Comisión Europea (2019): *Directrices para una IA fiable*. bit.ly/2RYbh8D.
- Comisión Europea (2020): *Libro blanco sobre la inteligencia artificial: Un enfoque europeo orientado a la excelencia y la confianza* (Bruselas) bit.ly/3gMTZ8S.
- Floridi Luciano, et alii; (2018): AI4People-An Ethical Framework for a Good AI Society: Opportunities, risks, principles, and recommendations, *Minds and Machines*, N° 28: pp. 689-707. bit.ly/3u0v71b.
- Hacker, Philipp (2018): Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies Against Algorithmic Discrimination Under EU Law, *Common Market Law Review*, N° 55: pp. 1143-1186.
- Hacker, Philipp (2020): AI Regulation in Europe. bit.ly/2S02i6S.
- Jaana Leikas*, Raija Koivisto et al., (2019): Ethical Framework for Designing Autonomous Intelligent Systems, *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 5, No 1: p. 18. bit.ly/3sWLN6.
- Mitrou, Lilian (2018): Data protection, artificial intelligence and cognitive services: Is the general data protection regulation (GDPR) “Artificial Intelligence-Proof”? bit.ly/32VHeRb.
- Wachter, Sandra, Mittelstadt, Brent y Floridi, Luciano (2017): Why a right to explanation of automated decision-making does not exist in the general data protection regulation, *International Data Privacy Law*, vol. 7, N° 2: pp. 76–99.

Normas citadas

- Directiva (UE) 2019/771 del Parlamento Europeo y del Consejo del 20 de mayo de 2019 relativa a determinados aspectos de los contratos de compraventa de bienes, por la que se modifican el Reglamento (CE) n.º 2017/2394 y la Directiva 2009/22/CE y se deroga la Directiva 1999/44/CE Directiva 2011/83/UE del Parlamento Europeo y del Consejo, del 25 de octubre de 2011, sobre los derechos de los consumidores, por la que se modifican la Directiva 93/13/CEE del Consejo y la Directiva 1999/44/CE del Parlamento Europeo y del Consejo y se derogan la Directiva 85/577/CEE del Consejo y la Directiva 97/7/CE del Parlamento Europeo y del Consejo.
- Directiva 2000/43/CE: Aplicación del principio de igualdad de trato de las personas independientemente de su origen racial o étnico.
- Directiva 2000/78/CE del Consejo, del 27 de noviembre de 2000, relativa al establecimiento de un marco general para la igualdad de trato en el empleo y la ocupación («Directiva de igualdad en el empleo») [COM(2014) 2 final del 17 de enero de 2014].
- Directiva 89/392/CEE del Consejo, del 14 de junio de 1989.
- Directiva 89/391/CEE del Consejo, del 12 de junio de 1989, relativa a la aplicación de medidas para promover la mejora de la seguridad y de la salud de los trabajadores en el trabajo.

Directiva 85/374/CEE del Consejo, del 25 de julio de 1985, relativa a la aproximación de las disposiciones legales, reglamentarias y administrativas de los Estados miembros en materia de responsabilidad por los daños causados por productos defectuosos.

Recomendación del Consejo, del 9 de diciembre de 2013, relativa a la adopción de medidas eficaces de integración de los gitanos en los Estados miembros (DO C 378 de 24.12.2013, pp. 1-7).

Reglamento (UE) 2018/1807 del Parlamento Europeo y del Consejo, de 14 de noviembre de 2018, relativo a un marco para la libre circulación de datos no personales en la Unión Europea.

Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, del 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos).

Jurisprudencia citada

Tribunal sobre la Directiva 95/46/CE, TJUE 6.11.2003, C-101/01 (Lindqvist), apartado 50.