Universidad de
SanAndrés

Universidad de San Andrés

Departamento de Derecho

Abogacía

Inteligencia artificial y transparencia: un vínculo esencial para la protección de los datos personales en la era de la IA generativa

Autora: Giuliana Jimena Bustamante

Legajo: 30031

Mentor: Marina Bericua

Argentina

Agosto de 2023

Agradecimientos

A mi mamá, Fabiana, por siempre creer en este proyecto y apostar todo por mí. Gracias por tu amor y por siempre impulsarme a seguir mis sueños. A mi hermana, por ser mi confidente, mi mejor amiga, mi motivación y apoyo incondicional en este camino. A mi mami Mirian, por ser mi guía y mi otra mamá, gracias por siempre inspirarme a ser mejor cada día.

A mi esposo Tommaso, mi compañero de vida, por siempre alentarme y ser un pilar en esos momentos difíciles. Haberte encontrado en este camino es una de las dichas más grandes de mi vida.

A familia que, a pesar de la distancia, siempre me alentaron, apoyaron y me acompañaron durante este camino. Nada sería posible sin mi querida tribu.

A mis amigos, por su amor y apoyo en esta etapa de mi vida. A Ana, por ser mi amiga, mi compañera de largas jornadas de estudio y una fuente de inspiración.

A Marina, por haberme transmitido su pasión por la transformación tecnológica del futuro y por siempre haberme inspirado como profesional.

A la Universidad de San Andrés, por haberme dado la oportunidad que transformó mi vida.

Abstract

El vertiginoso avance de la Inteligencia Artificial (IA) está dando forma a nuevas dinámicas en la vida cotidiana y las interacciones sociales. Dada la velocidad de su evolución y su capacidad para recopilar grandes volúmenes de datos, surge una preocupación acerca de la protección de la privacidad debido al procesamiento de información personal. Esta convergencia entre IA y la proliferación de información subraya el rol fundamental de la transparencia algorítmica en este panorama de desarrollo.

Centrándose en el rol de la IA en la toma de decisiones, este trabajo explora el equilibrio entre el avance de tecnológico y la privacidad de los datos personales, haciendo hincapié en la importancia de la transparencia algorítmica, como uno de los principios desarrollados para la protección de los datos personales. El análisis concluye analizando los desafíos que implica la transparencia algorítmica en la protección de datos, considerando su aplicación en los procesos de IA generativa.

Índice

1. Introducción 6

1.1. Pregunta 8

1.2. Metodología 9

2. Derecho a la privacidad 9

2.1. La protección de datos personales 12

2.2. Principio de transparencia en el tratamiento automatizado de datos 15

2.3. Regulación 18

2.3.1. Unión Europea: Reglamento General de Protección de Datos 18

2.3.2. Argentina: Ley de Protección de Datos Personales Nº 25.326 21

3. Inteligencia artificial 22

3.1. IA: Tratamiento inteligente de datos y toma de decisiones automatizadas 24

3.2. El estado regulatorio de la IA 26

3.2.1. Unión Europea: *IA Act* 28

3.2.2. Argentina: CONVENIO 108+ y el proyecto de reforma de la Ley Nº 27.326 29

3.3. Guías y recomendaciones legales para la IA responsable 33

3.3.1 *OCDE* 34

3.3.2. *UNESCO* 35

4. Transparencia algorítmica 35

4.1. ¿Es la transparencia un valor fundamental en la protección de datos personales en el contexto de los sistemas de IA? 40

4.2. Desafíos y críticas al principio de transparencia algorítmica 43

4.3. Caso de análisis: Azure OpenAI Services 46

5. Conclusión 51

Referencias bibliográficas 55

1. Introducción

La creciente presencia de la inteligencia artificial (IA) está revolucionando nuestra vida cotidiana, transformando nuestras interacciones y cómo llevamos a cabo nuestras actividades como miembros de una comunidad. Mientras tanto, nuestro mundo está experimentando una especie de "Big Bang" de la información, con el universo de datos duplicándose cada dos años, y quintillones de bytes de datos generados a diario (Marr, 2018). A medida que la IA evoluciona, amplifica su capacidad para utilizar información personal de maneras que pueden comprometer los intereses de privacidad, llevando el análisis de información personal a niveles de poder y velocidad nunca antes vistos. Esta convergencia entre la IA y la explosión de información redefine las implicaciones de la privacidad en nuestra sociedad contemporánea, enfatizando la importancia crucial de la transparencia algorítmica como un pilar fundamental en este nuevo paradigma de la tecnología.

La IA es una rama de la informática que se enfoca en crear sistemas que pueden aprender y realizar tareas que suelen implicar un proceso de cognición humana, y ofrece la oportunidad de innovar y promover el desarrollo de la civilización. Desde la industria a la educación, vemos como los algoritmos de IA forman parte intrínseca de los diversos tipos de tecnologías, aplicaciones y artefactos que utilizamos a diario. Sin embargo, definir el término de la inteligencia artificial resulta un desafío, dado que este modelo puede adquirir diversos significados.

Según la definición establecida por la Organización para la Cooperación y el Desarrollo Económico (OCDE), la IA es un "sistema basado en la máquina que puede hacer predicciones, recomendaciones o tomar decisiones, influyendo en entornos reales o

virtuales, sobre ciertos objetivos definidos por los humanos". Esta tecnología se define como un sistema que usa información producida por máquinas o humanos para identificar entornos reales y/o virtuales; es decir, recoge información abstracta y juega con inferencias modelo para formular opciones de información o acción (OCDE, 2019).

Esta tecnología representa grandes oportunidades económicas para el sector privado y público, ya que ofrece un aumento de la productividad en el tratamiento de datos y la toma de decisiones, basándose en grandes volúmenes de información que alimentan los algoritmos de la IA. Frente a este contexto y otras aplicaciones de la AI en procesos que avanzamos como sociedad, surge la preocupación de impulsar su desarrollo de manera ética y responsable, para producir impactos positivos para la comunidad. Una serie de posibles riesgos y consecuencias son asociados a la IA, por lo que se produce la tensión entre su desarrollo e implementación y los mecanismos de regulación legales y éticos necesarios para contener sus efectos negativos.

Por su parte, si se analiza que la tecnología de la IA depende necesariamente del consumo de datos para su entrenamiento, se puede afirmar que el desarrollo de su potencial se encuentra sujeto a riesgos asociados a la privacidad y protección de datos personales.

Por ello, en el presente trabajo se analizará particularmente la aplicación de la IA en la “automatización inteligente de los procesos”, que involucra el tratamiento automatizado de datos (en adelante, “tratamiento”) y la toma de decisiones automatizadas (en adelante, “ADM”) basados en modelos de IA, considerando que, a medida que la transformación digital avanza y las empresas empiezan a implementar IA en sus actividades, es necesario abordar los desafíos legales y éticos que implica su incorporación considerando los

principios en torno a la protección de la privacidad de los datos personales.

A lo largo del desarrollo, se considerarán los aspectos que comprende el impacto del principio de transparencia para la protección de datos personales y la evolución conceptual de la "transparencia algorítmica" en el marco de los procesos automatizados[1]. A estos fines se reconoce que, por la composición de estos sistemas, en alguna de sus etapas, hay problemas de transparencia en la recolección y tratamiento de datos, en tanto la IA es considerada un una "black box" o "caja negra" de contenido inaccesible para los usuarios (Sartor, 2020) y, por lo tanto, es necesaria la implementación de medidas de transparencia que colaboren en la protección de los derechos de las personas.

Finalmente, se intentará arribar a un análisis sobre la relevancia de la transparencia como un valor fundamental para la protección de datos, en el contexto de la automatización de los procesos y su intersección con el modelo de inteligencia artificial generativa (IAG), considerando las opiniones contrarias sobre este concepto.

1.1. Pregunta

¿Es el principio de transparencia un valor fundamental para la protección de datos personales en el contexto de la automatización de la toma de decisiones y el tratamiento de datos? ¿Cuáles son sus desafíos frente a los modelos automatizados de inteligencia artificial generativa?

[1] Es importante tener en consideración que, generalmente, la enumeración de los principios éticos de los marcos más importantes no asocia directamente la transparencia algorítmica con la privacidad de datos mencionando la privacidad como un principio diferente dentro del marco ético.

1.2. Metodología

Para el desarrollo de este trabajo se proponen tres etapas. En la primera etapa, se brindará un análisis conceptual sobre la privacidad de datos y el desarrollo del principio de transparencia de los marcos éticos de inteligencia artificial como una garantía esencial de los principios contenidos en las principales normativas de privacidad. Para ello, se considerará un breve análisis del Reglamento General de Privacidad de los Datos (RGPD) desarrollado por la Unión Europea y la Ley de Protección de Datos Personales N.º 25.326 de Argentina (LPDP).

En la segunda etapa, se desarrollará una breve introducción sobre el concepto y las implicancias de la Inteligencia Artificial y su aplicación. Asimismo, en este bloque, se observará cuáles son las propuestas regulatorias desarrolladas recientemente en la materia analizando, particularmente, los lineamientos de la UNESCO y OCDE para el desarrollo de una "IA responsable".

Finalmente, en la última sección se explorará el concepto de "transparencia algorítmica", así como, sus implicancias positivas para la protección de la privacidad de datos personales y los desafíos regulatorios que le propone la opacidad técnica de los algoritmos. Para reflejar estos conceptos, se brindará un análisis acerca de la transparencia en el producto de IAG, desarrollado por la compañía Microsoft: Azure OpenAI Services.

2. Derecho a la privacidad

El derecho a la privacidad es una garantía esencial reconocida en diversos acuerdos internacionales y en marcos constitucionales como un derecho humano básico. Es de vital importancia para preservar la dignidad de las personas y constituye un pilar fundamental en

cualquier sociedad que se rige por principios democráticos. Sin embargo, por su espectro de garantías, la determinación del concepto considerando que puede variar según el contexto de análisis y las variables culturales, legales y técnicas.

El desarrollo de la privacidad es difuso, mientras no se conoce cuál es el origen preciso de esta garantía. Sin embargo, a los fines del presente trabajo se considerará la influencia conceptual propuesta por el derecho angloamericano. Esta contribución fue fundamental para el reconocimiento del derecho a la privacidad, puesto que señaló la necesidad de salvaguardar a los individuos de cualquier tipo de intromisión injustificada del poder público en su esfera de actuación privada, en el contexto de construcción de la democracia y el surgimiento de una sociedad liberal (Basterra, 2016).

En este ámbito, un análisis conceptual de relevancia fue el artículo denominado “The Right to Privacy” de Samuel D. Warren y Louis D. Brandeis, dos juristas estadounidenses que sentaron las bases doctrinarias para la normativa de protección de datos posterior. La inquietud de Warren y Brandeis surgió a partir de los peligros que suponía para la intimidad de las personas el desarrollo de la sociedad tecnológica a comienzos del siglo XX (Saldaña, 2012). La rápida propagación de información mediante avances tecnológicos como los teléfonos y las cámaras fotográficas, sumado al auge de la prensa, ponía en riesgo la difusión descontrolada de información privada, revelando aspectos íntimos en las páginas de los diarios para complacer el interés morboso mediante la intromisión indebida en la esfera privada (Warren y Brandéis, 1890).

Ambos autores, en este contexto, desarrollaron el concepto de privacidad como una garantía para proteger la información de las personas, frente a la amenaza de que los medios de comunicación o

cualquier otro agente tuviera la posibilidad de registrar y difundir la imagen, pensamiento u otros aspectos de la esfera personal (Saldaña, 2012). Para garantizar el derecho a la privacidad, debían dar a los sujetos la posibilidad de consentir sobre el uso de su información personal, mientras el derecho a la privacidad se interpreta como una presunción a favor del control individual sobre sus datos. No obstante, esta presunción podría encontrarse sujeta a varias limitaciones vinculadas a la libertad de expresión y la disponibilidad pública de la fuente de información (Dreiffus, 1999).

En Argentina, el derecho a la privacidad se incluye dentro de la Constitución Nacional en su artículo 19 que establece que “Las acciones privadas de los hombres que de ningún modo ofendan al orden y a la moral pública, ni perjudiquen a un tercero, están sólo reservadas a Dios y exentas de la autoridad de los magistrados. Ningún habitante de la Nación será obligado a hacer lo que no manda la ley, ni privado de lo que ella no prohíbe”. Aunque el término "privacidad" no es mencionado específicamente en la Constitución Argentina, la jurisprudencia ha sostenido en repetidas ocasiones que este artículo garantiza su existencia y protección. Por su parte, en relación a la protección de la información personal, el Artículo 43 de la Constitución Nacional, establece que “toda persona podrá interponer esta acción para tomar conocimiento de los datos a ella referidos y de su finalidad, que consten en registros o bancos de datos públicos, o los privados destinados a proveer informes, y en caso de falsedad o discriminación, para exigir la supresión, rectificación, confidencialidad o actualización de aquellos. [...]”. De la misma manera, mediante la ratificación de tratados internacionales de derechos humanos, tales como el Pacto Internacional de Derechos Civiles y Políticos y la Convención Americana sobre Derechos Humanos, Argentina ha suscripto una serie de obligaciones que disponen que “nadie será

objeto de injerencias arbitrarias o ilegales en vida privada, su familia, su domicilio o su correspondencia, ni de ataques ilegales a su honra y reputación" (Art. 12, DUDH).

Considerando este marco regulatorio, es posible establecer que, en Argentina, una primera aproximación conceptual al derecho de privacidad implica reconocer que los sujetos tienen el derecho de decidir qué información personal revelar, quién puede acceder a ella y cómo se utiliza, es decir, se garantiza la protección de la intimidad personal, la inviolabilidad del domicilio, la confidencialidad de las comunicaciones y la protección de los datos personales (Basterra, 2016).

El análisis del concepto de derecho a la privacidad cobra relevancia a los efectos de presente trabajo, dado que permite a los sujetos resistir cualquier intromisión en su esfera privada y la difusión sus datos que, por su naturaleza, deben ser protegidos de la mirada pública, independientemente de que el hecho difundido sea verdadero o falso, ya que en ambos casos puede dañarse la intimidad (IX Jornadas Nacionales de Derecho Civil, 1983).

2.1. La protección de datos personales

Desde la llegada del internet y las nuevas tecnologías, las agencias regulatorias han enfrentado dificultades para establecer un marco de protección a la privacidad de los datos personales, en un contexto caracterizado por el impacto de la rápida difusión de los datos. En este contexto, el derecho a la protección de la privacidad de los datos personales es un concepto modificado repetidamente para evitar las injerencias a la vida privada de las personas y ajustar el entorno y la estructura del manejo de los datos personales en el ecosistema digital (Lu, 2022).

Históricamente, la preocupación por la privacidad estuvo asociada con el surgimiento de nuevos fenómenos que afectan a la estructura social. Tal es el caso de lo sucedido, por ejemplo, con la aparición de las cámaras fotográficas que proponían el riesgo de la difusión de las imágenes de las personas, sin el consentimiento necesario para hacerlo. Como resultado de estas tecnologías u otras similares, surgieron mecanismos regulatorios para la protección de los sujetos de las ofensas particulares y la violación de su esfera de privacidad. Así, con la llegada de diferentes tecnologías, el derecho a la privacidad y el derecho a la protección de datos han evolucionado en sus mecanismos de protección. A continuación, se analizará cómo la doctrina distingue entre tres generaciones de evolución del derecho a la privacidad.

Los primeros desarrollos sobre la protección de datos personales estuvieron enfocados en establecer el alcance de las libertades individuales frente a la intervención de los poderes públicos, exigiendo limitaciones en favor de la protección de los derechos individuales (Visintini, 2020). Entre estos derechos se observan, por ejemplo, la protección del espacio personal, la autonomía, la libertad y la protección frente a la intromisión estatal, asociados a las primeras reglas para proteger la privacidad en los espacios físicos, como la inviolabilidad del domicilio y la correspondencia personal. En esta primera generación surgieron regulaciones de gran relevancia, tales como la Declaración Universal de los Derechos Humanos, adoptada por la Asamblea General de Naciones Unidas en diciembre de 1948, la cual sentó las bases de una primera aproximación hacia la protección de datos personales.

Por su parte, la segunda generación se vio influenciada por la expansión de la globalización y el intercambio transfronterizo de datos.

En este contexto, surgen los conceptos en torno a la protección de datos respecto a la recopilación, uso y almacenamiento de información relativa a los sujetos. Así, surgen acuerdos que garantizan, por primera vez, que la información personal estuviera protegida en el marco de intercambios internacionales y el desarrollo de nuevos modelos de gobierno y sus influencias en los derechos económicos y culturales.

En esta etapa surgen conceptos como la "autodeterminación informativa" (informational privacy) desarrollada por Alan Westin, que reconoce a los individuos una expectativa legítima de que su información personal sea tratada de forma confidencial y segura por los terceros con acceso a la misma. Westin defendía que las personas deberían poder tomar decisiones informadas sobre la recopilación, el almacenamiento y el uso de información personal; en tanto la autodeterminación informativa es fundamental para equilibrar los intereses comerciales y gubernamentales con los derechos individuales de privacidad (Westin, 1967).

Finalmente, en la tercera generación, se incorpora al conjunto preexistente de garantías, la libertad informática como un nuevo derecho orientado a tutelar la identidad digital de los sujetos. Esta garantía se concreta mediante la posibilidad de acceder y controlar las bases de información que contienen datos de relevancia de la identidad personal del individuo (Rodríguez, 1997). Dicho concepto abarca tanto aspectos propios de la individualidad, como la información personal, los rasgos, los comportamientos y las asociaciones que definen colectivamente el carácter distintivo de una persona (Acquisti, 2020). Las normativas de protección de datos buscan garantizar la preservación de la identidad digital indemne del uso, acceso y tratamiento indebido de datos.

A partir de esta tercera generación es donde surgen los primeros principios éticos vinculados al tratamiento de datos. Sin embargo, a los fines del presente trabajo, nos centraremos en explicar la relevancia del principio de transparencia, la explicabilidad y el consentimiento para el tratamiento automatizado de los datos, como un elemento fundamental para la protección de la privacidad de datos.

2.2. Principio de transparencia en el tratamiento automatizado de datos

Este principio contenido en casi todas las leyes asociadas con la protección de datos implica la incorporación de medidas de control para la recogida, el almacenamiento, el tratamiento y el intercambio de datos personales, así como la concesión a las personas del derecho a controlar su propia información y el modo en que se utiliza. En la actualidad, la transparencia es un principio utilizado con frecuencia para la protección de la identidad digital, puesto que incluye el cumplimiento de otras garantías como el consentimiento informado, la minimización de datos, el derecho al olvido, la seguridad y protección de la información y la regulación sobre las transferencias internacionales de datos por medios digitales (Acquisti, 2020).

El término “transparencia”, se relaciona con múltiples conceptos y funciones que se pretenden implementar para cumplir determinados objetivos en la protección de la privacidad. De esta manera, la transparencia puede hacer referencia a la explicabilidad, la interpretabilidad, la apertura del código, la visibilidad y la accesibilidad (Weller, 2017; Felzmann, 2020). A pesar de su gran relevancia, su conceptualización ambigua genera ciertas tensiones entre la transparencia como un ideal normativo dentro de las reglas de privacidad de datos, y su aplicación concreta sobre el tratamiento de datos (Felzmann, 2020). Es por esta razón, desde el análisis académico de la transparencia, se propone que su aplicación no sea

únicamente un deseo regulatorio sino un elemento constitutivo de la privacidad.

Por ello, para discutir sobre la transparencia como valor fundamental hay que establecer que este principio no es un deber de proveer información sobre el tratamiento de datos, que, aunque es el concepto que le regula la privacidad de datos reciente, no abarca los significados y valores de la transparencia respecto a sus efectos de protección. Para ello, desde el análisis político de la transparencia, Meijer (2014) propuso un análisis de este concepto bajo tres nociones: la transparencia como virtud, relación y como sistema (Felzmann, 2020).

En primer lugar, la transparencia se considera como una acción valiosa en tanto consiste en la acción de los agentes o sistemas de exhibir sus operaciones, comportamiento y consideraciones. En esta dimensión, es posible comparar que este concepto de la transparencia como un "valor", es la definición que proponen los actuales reglamentos de privacidad, como el RGPD y las guías de Unesco, donde se impone la obligación a los responsables de informar a los usuarios sobre el tratamiento de datos al que se someten. Sin embargo, según explican Meijer (2014) y Felzmann (2020), esta concepción del principio de transparencia no es eficiente hasta que se lo vincula con la segunda dimensión de su significado: la transparencia relacional.

La transparencia relacional implica que haya una relación entre la apertura de los responsables del tratamiento y la comprensión y recepción de los usuarios involucrados. En este sentido, vemos que las regulaciones sobre transparencia establecen garantías para que los usuarios puedan reclamar la supervisión del mecanismo detrás del tratamiento de datos. Pero, como se analizará después, esta

dimensión del concepto de transparencia se ve afectada por las características técnicas de la inteligencia artificial.

En tercer lugar, la perspectiva de este principio como “transparencia sistémica” implica comprender cuales son las reglas y mecanismos legales que disponen la aplicación práctica de ciertas medidas para asegurar la transparencia en los procedimientos. En esta dimensión del concepto, se observa la relación con las normas que predisponen la auditoria y la rendición de cuentas de los sistemas involucrados en el tratamiento de datos, de acuerdo a la categorización de riesgos que corresponda.

Según Felzmann (2020), la integración de estas tres perspectivas es fundamental para comprender el significado del principio de transparencia y su relevancia dentro de los procesos de tratamiento automatizado de datos (TAD) y la toma de decisiones automatizadas (TDA). Entender a la transparencia como un valor fundamental, implica reconocer la relevancia de no solo pensar a los conceptos de “transparencia desde el diseño” de los sistemas, sino también incorporar como un requisito ex ante del desarrollo de cualquier tipo de tecnología involucrada en el tratamiento de datos y las organizaciones que llevan a cabo estas acciones (Tielenburg 2018) sobre el tratamiento de los datos personales.

Como será analizado a continuación el objeto y contenido del principio de transparencia y, otros principios “hermanos” del concepto como, la explicabilidad y la interpretabilidad, han sido tratados por diversos conjuntos reglamentarios, dentro de los cuales la transparencia como adquiere significado y consideraciones diversas.

2.3. Regulación

En esta era de digitalización y conectividad digital, la protección de datos personales se ha convertido en una preocupación creciente para los ciudadanos y las autoridades en todo el mundo. Todos los países a lo largo de las últimas décadas han promulgado regulaciones específicas para abordar este desafío y salvaguardar las garantías de sus ciudadanos, aún con las implicaciones transfronterizas que el fenómeno tecnológico implica. En este trabajo se analizará, el marco regulatorio europeo, como precursor en términos de protección de datos y transparencia en el ámbito digital, y la situación regulatoria en Argentina, representada por la Ley Nacional de Protección de Datos Personales (Ley Nº 25.326) y modificatorias, comparando sus enfoques, similitudes y diferencias.

2.3.1. Unión Europea: Reglamento General de Protección de Datos

El Reglamento General de Protección de Datos (RGPD) es un conjunto de normas y regulaciones de la Unión Europea que establece un marco unificado para la protección de datos personales de los ciudadanos de las naciones miembros de la Unión Europea. Adoptado en abril de 2016 y aplicable desde mayo de 2018, el RGPD permitió fortalecer y armonizar la protección de datos en toda la UE y garantizar que las organizaciones que manejan datos personales cumplan con los más altos estándares de seguridad disponibles al momento.

Este Reglamento desarrolló una serie de principios impulsados por la necesidad de fortalecer la protección de datos personales en la era digital y promover un abordaje unificado en toda la Unión Europea. El RGDP es uno de los marcos regulatorios pioneros en protección de datos en la era digital, y establece principios orientados a proteger al individuo y sus datos personales. Estos principios son: a) licitud,

lealtad y transparencia; b) limitación de la finalidad; c) minimización de datos; d) exactitud; e) limitación del plazo de conservación y; f) integridad y confidencialidad. Todos integran un sistema que permite que el tratamiento de los datos personales se haga de forma controlada y segura.

Según el artículo 12 del RGPD, la transparencia implica que todo tratamiento de datos personales debe ser lícito y leal, para que las personas físicas conozcan el tratamiento y las finalidades para las que se someten sus datos. Para ello, el RGPD establece ciertas pautas de cumplimiento que las organizaciones deben realizar de acuerdo a los estándares de transparencia, como el conocimiento de los riesgos, las normas, las salvaguardias y los derechos relativos al tratamiento de datos personales, así como de los mecanismos de ejercicio de estos derechos.

Para su análisis, utilizaremos una definición propuesta por el investigador Felzmann (2020) respecto de los elementos que componen la concepción informativa de la transparencia en el RGPD. Este autor sostiene que el principio de transparencia puede ser analizado según el momento en que se pretenda aplicar sus efectos. Por ello, por un lado, el elemento prospectivo de la transparencia que considera las implicancias del principio ex ante del tratamiento de datos. En otras palabras, la transparencia prospectiva, garantiza que las personas tengan el derecho a ser informadas sobre el tratamiento de datos al que se someten antes de involucrarse en el mismo. Según esta normativa, el principio implica que los responsables brinden a los interesados, información expresada en lenguaje sencillo, claro y simple de comprender; que deberá estar disponible en escrito o por medios electrónicos, con facilidades de acceso para los titulares de la información (RGPD, artículos 5 y 13). Asimismo, la legislación obliga

a que los responsables del tratamiento establezcan, la cantidad y calidad de los datos procesados, el tiempo de las actividades de procesamiento, la razón y el propósito del procesamiento (Felzmann et al., 2019).

Por otra parte, la transparencia retrospectiva refiere a la aplicación de la transparencia, luego de realizado el tratamiento de la información. Por lo tanto, constituye la posibilidad de que los sujetos interesados soliciten un informe acerca de las implicancias técnicas (cómo) que llevaron a una determinada decisión o predicción concreta (porque). En este sentido, el RGPD otorga a los interesados la capacidad de solicitar intervención humana en el sistema del tratamiento automatizado para obtener una decisión alcanzada tras dicha evaluación e impugnar los fundamentos de la decisión (Articulo 71, RGPD). Repetidamente, este punto ha sido discutido entre expertos legales respecto a si existe un derecho a la explicabilidad de los sistemas (Volg et al., 2019). El ámbito de la inteligencia artificial, observamos que este debate es particularmente controversial, en tanto es necesario definir si las regulaciones sobre privacidad van a incluir la explicabilidad como un requisito para la Inteligencia Artificial (Miller, 2019).

De esta manera, observamos como los reguladores europeos han considerado a la transparencia como un principio fundamental para la protección de la privacidad de los sujetos (Felzmann et al., 2019). En este primer análisis se puede señalar que, según lo mencionado, el principio de transparencia se limitó a la función informativa del principio, dejando de lado su asociación con otros valores que subyacen a la transparencia y le agregan valor como un principio autónomo y fundamental para la protección de los datos personales.

2.3.2. Argentina: Ley de Protección de Datos Personales Nº 25.326

La Ley de Protección de Datos Personales N° 25.326 (LPDP), promulgada en octubre del 2000 es la actual regulación relativa a la protección y el tratamiento de los datos personales en Argentina. Esta ley se desarrolló para proteger los datos personales individuales con la constitución de principios y pautas regulatorias que las entidades públicas y privadas deben seguir al recopilar, procesar, almacenar y transmitir información. Asimismo, otorgo un marco regulatorio a la modificación constitucional realizada en 1994 que incorporo la acción de habeas data a la Constitución Nacional.

La LPDP, junto con sus modificaciones (Decreto 1558/2001), se aplica a todas las personas físicas y jurídicas, que se encuentren involucradas en el tratamiento de datos personales en el territorio argentino. Entre los principios legales establecidos para la protección de la privacidad observamos que, como fue analizado en el punto anterior, la normativa argentina reconoce el principio de transparencia en su concepción "informativa" a través de otras consideraciones como el consentimiento, la limitación de la finalidad y el sistema de responsabilidad por el tratamiento. Por tanto, aunque el concepto no se encuentra incluido en la normativa, la LPDP considera estos preceptos desde su concepción, que evoluciona y se refuerza con la incorporación de acuerdos internacionales que desarrollan más el vínculo entre el principio de transparencia y la protección de datos personales (Travieso, 2017).

De esta manera, vemos que ante las exigencias regulatorias de las nuevas tecnologías como la Big Data, la inteligencia artificial y el Internet de las cosas (IoT, por sus siglas en inglés), así como la creciente internacionalización del flujo de datos, el conjunto normativo argentino ha adoptado dos medidas que resultaran de gran relevancia

para las normas de protección de datos futuras: por un lado, la aprobación de la Ley 27.699 de Protección de las Personas con respecto al tratamiento automatizado de datos de carácter personal que aprueba el Convenio 108+; y por otro, el Proyecto de modificación a la LPDP aprobado por el senado y cámara de diputados en fecha 29 de junio de 2023; los cuales serán analizados posteriormente, junto con la normativa de privacidad para la Inteligencia Artificial.

Como se analizará posteriormente, este proceso se encuentra comprendido en los proyectos de ley vinculados a la regulación de privacidad en el marco de la inteligencia artificial.

3. Inteligencia artificial

La llegada de la inteligencia artificial es uno de los mayores avances tecnológicos de las últimas décadas, convirtiéndose en un elemento omnipresente transformador del modo en el que interactuamos con el mundo, desde las tareas cotidianas sencillas hasta los procesos productivos complejos. El término de “inteligencia artificial” ha pasado de estar asociado a la ficción futurista, a ser un descriptor usual en las noticias sobre avances científicos recientes. Este innovador desarrollo impregna la civilización con el mismo impacto que en su momento tuvo la electricidad, en su capacidad de transformar, proponer y permear a todos los procesos productivos, sin distinción de disciplina (Mantegna, 2022).

En sus orígenes, alrededor de los años cincuenta, Alan Turing, un matemático planteo en su artículo “Computing Machinery and Intelligence”, una pregunta que impulsó el desarrollo de esta nueva tecnología: ¿Pueden pensar las máquinas? A partir de ese trabajo se desarrollaron las primeras bases computacionales que defendían la posibilidad de que las computadoras a través de algoritmos podían imitar el pensamiento humano. El desarrollo de estos sistemas fue en

un primer momento experimental, con orientación en la ingeniería, donde se definía a este fenómeno como la habilidad de ciertas máquinas para realizar acciones que imitaban la inteligencia humana. Por lo tanto, establecer un concepto para la “inteligencia artificial” no supone una tarea sencilla, considerando que, por su evolución, no existe una categoría concreta que englobe todas sus funciones.

Si bien el descubrimiento y desarrollo de la IA comienza a mitades del siglo XX, sus primeras aplicaciones estaban limitadas a solo algunas tareas de repetición guiada por algoritmos computacionales. Aun así, podemos observar que, solo unos años después que estuviera en funcionamiento la primera computadora, existían destacados miembros de la ciencia en el área que hacían afirmaciones acerca del gran potencial de IA, estableciendo que “[...] en el mundo se están desarrollando máquinas capaces de pensar, aprender y crear. Su capacidad para hacer lo anterior aumentará rápidamente hasta que –en un futuro previsible- la magnitud de problemas que tendrán capacidad de manejar irá a la par con la capacidad de la mente humana para hacer lo mismo [...]” (Simon, 1957).

Sin embargo, a pesar de su potencial, la IA en sus primeras décadas de existencia, atravesó diversos periodos o “inviernos” en los cuáles, por razones de inversión o interés científico de los gobiernos, los profesionales del campo no obtenían el financiamiento necesario para llevar a cabo investigaciones más profundas sobre sus posibles aplicaciones. No obstante, a fines del siglo pasado, un suceso histórico cambio la historia de la IA para siempre.

En 1997, la IA de Deep Blue de IBM (International Business Machines) venció al campeón mundial de ajedrez Gary Kasparov, siendo esta la primera vez que un hombre es derrotado por la inteligencia artificial de computadoras. Desde allí, se desarrollaron diversos modelos de

aprendizaje sostenidamente, hasta alcanzar el nivel de evolución de los algoritmos actuales.

En la actualidad, la IA se definen como un sistema de software que actúa en una dimensión digital mediante la programación humana para la realización de un objetivo determinado a través de la obtención de datos, la interpretación de los datos estructurados o no estructurados que recopilan, el razonamiento sobre el conocimiento o el procesamiento de la información derivados de esos datos, y decidiendo la acción o acciones óptimas que deben llevar a cabo para lograr el objetivo establecido (OEIAC, 2021).

La IA, como disciplina científica, incluye aplicaciones y modelos técnicos particulares, como el aprendizaje automatizado (donde el deep learning, el reinforcement learning y el supervised learning, son ejemplos), el razonamiento automatizado (que incluye fenómenos como el análisis, tratamiento y razonamiento de los datos, la búsqueda y la toma de decisiones automatizadas) y la robótica (donde se integran las técnicas de software en sistemas ciber físicos) (AI HLEG, 2018)

3.1. IA: Tratamiento inteligente de datos y toma de decisiones automatizadas

El tratamiento automatizado de datos (ADP, por sus siglas en inglés) implica el procesamiento de datos mediante sistemas de aprendizaje profundo de AI aplicados al análisis de texto, imágenes o video, que crean una serie de patrones que luego permiten al razonamiento automatizado desarrollar otras tareas vinculadas como la toma de decisiones (ADM, por sus siglas en inglés). Este sistema, además, funciona con poca o nula interacción humana, basándose en reglas específicas, códigos y configuraciones, y tienen la capacidad de

realizar tareas con mayor velocidad y precisión que los humanos (AI HLEG, 2018).

Una de sus características fundamentales, que lo diferencian de otros sistemas de tratamiento de la información, es la capacidad del ADP de incorporar grandes volúmenes de datos no estructurados y transformarlos en información estructurada que, junto con los datos de inferencia, permite el entrenamiento del razonamiento automatizado del sistema de IA. Esta tecnología es utilizada, por ejemplo, para facilitar ciertos procesos como tareas empresariales complejas como la estructuración de perfiles de consumidores para el posterior análisis de mercado de acuerdo a las categorizaciones realizadas por la automatización de los datos obtenidos en la interacción comercial, entre cientos de otros ejemplos.

Luego de que el ADM facilite la estructuración de los datos para convertirlos en conocimiento, el siguiente paso es se realicen los procesos de razonamiento de la IA, lo que incluye hacer inferencias, que permita asemejar esta información al conjunto de soluciones programadas y optimizarla decisión entre las respuestas disponibles para un problema. El último paso es elegir qué decisión tomar. Aquí es donde el sistema de ADM de un sistema de AI toma relevancia.

La ADM, es de acuerdo a su definición, un proceso de toma de decisiones impulsado de manera autónoma, es decir, sin inferencias humanas para su desarrollo. Desde el punto de vista técnico, estos sistemas incluyen diversos modelos de aprendizaje y procesamiento de datos de IA y, por lo tanto, comprender su mecánica puede ser una tarea compleja. Por ello, en simples palabras es posible establecer que la ADM es un sistema nuclear de la IA, vinculada directamente con el razonamiento automatizado, que toma ciertos datos de entrada que fueron transformados en conocimiento, y emite como respuesta, una

acción a realizar, dado el objetivo programado a alcanzar (Garcia Herrero, 2020). En algunas circunstancias, estas decisiones pueden estar basadas en conocimiento producido luego del procesamiento de datos no estructurados, así como también, la creación de perfiles digitales e inferencias de análisis introducidas en la programación de los algoritmos.

Para la comprensión de estos sistemas proponemos pensar en, por ejemplo, un sistema de IA utilizado por una empresa que quiere comprender cómo se sienten los clientes respecto de sus productos y servicios. Para ello, el sistema de IA de procesamiento del lenguaje natural extraerá las palabras asociadas a un comentario en las publicaciones en redes sociales asociados con su marca, y los procesa para convertirlos en contenido que permite tomar una decisión. Aunque a los humanos nos parezca sencillo comprender que algunas palabras están asociadas a ciertas emociones positivas o negativas, para una máquina, esto no es tan sencillo, puesto que el texto no es más que un simple token numérico dentro de un universo de combinaciones matemáticas.

Considerando el funcionamiento de estos sistemas, a continuación, analizaremos algunas aproximaciones regulatorias asociadas a la inteligencia artificial y la protección de la privacidad de los datos personales.

3.2. El estado regulatorio de la IA

La tecnología, en su conjunto, ha significado siempre un objeto difícil de regulación debido a su rápida evolución y a la complejidad de su impacto en la sociedad. Es un hecho que la misma puede desarrollarse más rápido que los conjuntos regulatorios que la comprenden, lo que provoca un desfasaje entre el avance tecnológico y la regulación aplicable, que desencadena riesgos por ausencia de consideraciones

legales, éticas y sociales. Al mismo tiempo, su naturaleza transfronteriza hace que la regulación sea aún más difícil, puesto que sus efectos toman lugar en diversas jurisdicciones que pueden coincidir o no en la manera de abordar el manejo de riesgos asociados a la IA.

Para llevar adelante la protección de datos en los modelos de IA, las primeras aproximaciones conceptuales, consideraban que era necesaria la colaboración efectiva entre organismos con una visión compartida de que la tecnología debía ser utilizada para el bienestar social. En la actualidad, bajo esta concepción, la discusión acerca de cómo abordar la regulación de la IA se encuentra en la mayoría de las agendas públicas gubernamentales. Entre ellos, organismos mundiales como la ONU (Organización de Naciones Unidas) entienden que las plataformas deberían ser transparentes respecto a la manera que se utiliza la IA y los algoritmos, explicando el tratamiento de datos realizado para arribar a ciertas conclusiones (ONU, 2018).

Al momento de elaboración del presente trabajo, la inteligencia artificial y los modelos que se derivan de la misma se encuentran en un contexto regulatorio complejo y en constante evolución. Aunque aún no existe un sistema regulatorio específico para la inteligencia artificial, varios países y regiones han adoptado marcos normativos que abordan aspectos específicos relacionados con la protección de la privacidad de datos personales. A continuación, se analizará el Reglamento General de Protección de Datos en la Unión Europea, la Ley de Protección de Datos Personales de Argentina y mecanismos de Soft Law que proponen la protección a los datos personales en el contexto de la tecnología de decisiones automatizadas.

Sin embargo, es claro que la creación de marcos éticos para el desarrollo de IA propone una serie de principios integrados para

salvaguardar los derechos de las personas, entre ellos, la privacidad de sus datos personales.

3.2.1. Unión Europea: *IA Act*

En el año 2021, frente a la masividad del uso de estas tecnologías y los riesgos asociados a su falta de regulación, la Unión Europea, como referente en materia de protección de datos, decidió abordar las discusiones necesarias para regular el tratamiento automatizado de datos mediante los sistemas de IA. De estas discusiones surgió el proyecto de Ley de Inteligencia Artificial (Artificial Intelligence Act, "AIA" por sus siglas en inglés) que fue aprobada en junio de 2023[2].

La propuesta regulatoria se enfoca en garantizar que los sistemas de IA sean desarrollados y utilizados con mecanismos de seguridad, transparencia, trazabilidad, igualdad y sustentabilidad. Por lo tanto, establece limitaciones a la autonomía de la IA, en tanto requiere que los sistemas algorítmicos sean intervenidos por la supervisión de humanos, en lugar de la dependencia en la automatización, para evitar que los resultados sean perjudiciales a los derechos y garantías de las personas. Asimismo, esta normativa, adopta nuevos enfoques regulatorios y propone un sistema regulatorio que se adapte a diferentes tecnologías considerando niveles diferentes de riesgo: riesgo inaceptable, alto riesgo, riesgo limitado y riesgo mínimo; cada uno de estos niveles con consideraciones particulares que permiten garantizar la transparencia y la protección de los derechos de las personas (European Parliament, 2023).

[2] El proyecto fue aprobado por el Parlamento Europeo el 14 de junio de 2023. Al momento de investigación del presente trabajo, el proyecto se encuentra en una etapa de negociaciones sobre la forma final de la ley en el Consejo Europeo, junto a los 27 estados miembros de la Unión Europea. La autorización final para su vigencia y aplicabilidad se espera para finales de 2023 o principios del año 2024.

Sin embargo, a pesar de que este conjunto normativo sea un paso importante para el desarrollo de la “IA responsable”, hay quienes consideran que, con el avance de la IAG, esta propuesta desarrollada en 2021 haya quedado obsoleta. Frente a esta situación, el Parlamento Europeo ha comenzado a discutir una serie de anexos y consideraciones regulatorias específicas para la IAG, que proponen introducir requisitos concretos para el desarrollo de modelos fundacionales como GPT. Estos requisitos serían exigibles para todos los modelos de IAG, sin considerar el análisis de riesgo que propone la normativa. Este punto, sin embargo, también es criticado, puesto que contradice el enfoque inicial de la AIA de no regular modelos en particular, sino establecer regulaciones generales capaces de abordar cualquier tipo de tecnología.

La AIA, a pesar de las críticas, es un primer acercamiento regulatorio a los fenómenos de IA que atraviesan y transforman nuestras vidas día a día. Si bien definir límites menoscaba la gobernanza del desarrollo de nuevos modelos, es importante que la regulación pueda contener los efectos negativos del uso de cantidades enormes de datos en sistemas computacionales tan complejos como la IAG.

3.2.2. Argentina: CONVENIO 108+ y el proyecto de reforma de la Ley Nº 27.326

El Convenio 108+ incorporado en diciembre de 2022 por el Decreto Nº 792/2022 que promulgó la Ley Nº 27.699, mediante el cual se determina la versión actualizada del Convenio Europeo 108 para la Protección de las Personas con respecto al tratamiento automatizado de Datos Personales (en conjunto, “Convenio 108”) constituido en 1981 con el fin de proteger la privacidad de los individuos.

A partir de ciertas modificaciones, el tratado se convirtió en un estándar internacional de seguridad jurídica en materia de privacidad

para cualquier Estado que decidiera unirse. Esta normativa permitió robustecer los estándares internacionales en materia de privacidad, favoreciendo el intercambio de datos dentro de la integración global y la protección de la privacidad de los sujetos ante la posibilidad de abusos en el tratamiento de sus datos.

La nueva versión del Convenio 108, también denominada como "Convenio 108+", recoge los desafíos que proponen las nuevas tecnologías y el ADP para la seguridad de la privacidad y propone, entre nuevos conceptos antes no incorporados a la convención original, como la relevancia del "principio de transparencia". Según establece el artículo 8 de la Convención 108+, se entiende que el ADP debe ser llevado a cabo de forma justa y transparente, con fines específicos, legítimos y no tratarse de forma incompatible a estos.

En este contexto, al igual que en el RGPD, se reconoce a la transparencia como un principio que obliga al responsable del tratamiento a permitir que los interesados comprendan y, por lo tanto, tengan la capacidad de ejercer sus derechos en el contexto de dicho tratamiento de datos, dentro de los cuales se encuentra el consentimiento informado. Por su parte, la ratificación argentina de la Convención 108+, implicó el reconocimiento de los estándares internacionales actualizados y vinculados en temas de privacidad, asociadas al ADP. Por lo tanto, la incorporación del Convenio significo el impulso necesario para ejecutar una propuesta de actualización para la actual Ley de Protección de los Datos Personales, que recoge inspiración de los lineamientos del Convenio 108+ y el Reglamento General de Protección de Datos de la Unión Europea (Szlak, 2022).

El proceso de reforma de la LPDP iniciado en el año 2022, es el resultado de un arduo debate entre ciudadanos, profesionales,

universidades, investigadores y organizaciones privadas y públicas[3]. Su creación se encuentra asociada a la necesidad de actualizar el contenido de la Ley 27.326 para fortalecer la capacidad estatal de regular y gestionar los nuevos desafíos impuestos por las tecnologías recientes en el contexto de una economía digitalizada. Asimismo, como fue mencionado anteriormente, dadas las obligaciones internacionales de Argentina, la modificación era necesaria para estar alineada con las Guías y Convenciones suscritas en materia de protección de datos personales.

Esta reforma incorpora cambios inspirados en conjuntos normativos como el RGPD sobre el tratamiento automatizado de los datos, en tanto considera nuevas definiciones para los términos como la transferencia internacional de datos, la anonimización, la elaboración de perfiles, la toma de decisiones automatizadas, la transparencia y la evaluación de impacto de la protección de datos.

En materia de transparencia, el proyecto de reforma reconoce nuevas bases para el consentimiento y para los deberes de transparencia respecto a la automatización del tratamiento de datos personales. Esta modificación propone reconocer al principio de transparencia y sus diferentes fases (informativa, relacional y sistémica) de acuerdo al siguiente análisis.

En primer lugar, la concepción del consentimiento informado amplía sus implicancias, en función de proyectar mayor seguridad e interés legítimo del responsable del tratamiento en obtener el asentimiento del

[3] Se reconoce que, a pesar de la participación de las asociaciones empresariales interesadas, la versión final del proyecto de actualización de la LPDP no recoge algunos de los comentarios aportados en el proceso de consulta pública. Al momento de investigación del presente trabajo, el Proyecto de reforma se encuentra en el Congreso de la Nación y ha sido sometido a una primera sesión informativa de Comisiones.

titular de datos. En segundo lugar, el proyecto establece que la información debe ser accesible a los titulares de los datos, sin limitaciones e incluir una base legal específica para el tratamiento de los datos y el plazo de conservación. Además, el Proyecto reconoce a los titulares de datos, el derecho a oponerse al tratamiento de sus datos personales en fuentes automatizadas; el derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado o semiautomatizado de datos personales cuando dicha decisión pueda tener efectos discriminatorios; el derecho a la portabilidad de los datos; y el derecho a solicitar la limitación del tratamiento (Fernández, 2023).

En este contexto, surge la primera conexión entre las decisiones basadas únicamente en el tratamiento automatizado o semiautomatizado de datos personales y el principio de transparencia que permite que los titulares de los datos tengan información sobre los procedimientos y criterios utilizados en dicha decisión. De esta manera, al establecer ciertos deberes de información para los responsables del tratamiento, se favorece no solo el principio de transparencia en su espectro informativo, sino también, poniendo especial atención en la relación entre los agentes responsables y los usuarios afectados.

En tercer lugar, las consideraciones de la transparencia sistémica en la protección de datos también se encuentran presente en el nuevo proyecto regulatorio, en tanto considera que existe el derecho a solicitar una auditoría para verificar que el procedimiento no esté incumpliendo la ley ni afectando ninguna garantía de los sujetos.

Finalmente, analizamos que, en el proyecto de reforma, se incorpora el concepto de “transparencia desde el diseño”, en tanto los responsables del tratamiento se encontrarán obligados a aplicar desde el diseño y por defecto en sus productos, medidas de privacidad (o de

transparencia) y realizar evaluaciones de impacto sobre la protección de datos personales cuando se cumplan determinadas condiciones (Fernández, 2023).

De esta manera vemos que el proyecto regulatorio en Argentina, inspirado por las bases normativas del RGPD de la Unión Europea, comienzan a considerar al principio de transparencia como un elemento transversal a todo el proceso involucrado en la IA. De esta manera, el término “transparencia” abandona su concepto original sobre la función informativa, para mutar hacia el concepto de transparencia desde el diseño, transparencia sistémica y transparencia relacional.

Sin embargo, a pesar de estos avances en el concepto, aun el proyecto de normativa no ensaya, ni siquiera en los debates que se llevaron a cabo con anterioridad a la confección de su versión presentada al Congreso, cómo se asegurará el cumplimiento práctico del principio de transparencia como elemento fundamental en la protección de la privacidad en relación con los modelos complejos de IA, es decir, en relación con las nuevas tecnologías de inteligencia artificial generativa.

3.3. Guías y recomendaciones legales para la IA responsable

Así como fueron analizados conjuntos normativos particulares, cabe mencionar aquellos desarrollos privados que sentaron las bases para el desarrollo de principios y derechos para la protección de la privacidad en los sistemas de inteligencia artificial. A continuación, haremos un breve recorrido por los conjuntos normativos internacionales que intervienen en diversas jurisdicciones para asegurar la unificación del enfoque regulatorio de estas nuevas tecnologías, dadas sus características transfronterizas.

3.3.1 *OCDE*

La "Guía de Recomendaciones sobre Inteligencia Artificial" de la OCDE fue el primer marco regulatorio supranacional desarrollado para la promoción y regulación de IA de acuerdo a principios de transparencia y respeto a los derechos humanos y los valores democráticos. Por lo tanto, constituyeron los primeros indicios regulatorios para algunos de los gobiernos que suscribieron sus principios e indicaciones.

En este contexto, la transparencia fue entendida para la OCDE en su Artículo 1.3, como un concepto relevante dentro del desarrollo de una IA confiable. La guía resalta una serie de deberes que corresponden a los agentes de IA para asegurar su compromiso y cumplimiento en el desarrollo de productos basados en estas tecnologías. Entre estos, establece que: (i) es necesario brindar información que permita comprender el funcionamiento de los sistemas de IA; (ii) informar y advertir a los interesados sobre las interacciones basadas en IA; (iii) brindar apoyo a aquellos que hayan sido afectados por una decisión de IA a comprender el proceso involucrado en el resultado; y (iv) permitir que aquellos que se hayan sentido afectados por una decisión automatizada, impugnen el resultado basándose en la información brindada sobre el funcionamiento del sistema de IA para la recomendación o decisión (OCDE, 2019).

Si bien esta concepción se limitó al concepto informativo del principio de transparencia, fue un primer paso para establecer ciertas medidas de protección para el tratamiento automatizado de datos. Algunos de estos preceptos sirvieron como inspiración para otras guías internacionales para la IA, leyes nacionales y recomendaciones realizadas por agentes privados como empresas relevantes en el sector digital.

3.3.2. *UNESCO*

Desde el surgimiento de la inteligencia artificial, por su parte, la UNESCO ha desarrollado diversos lineamientos que promueven y protegen los derechos humanos en el ecosistema digital. Dentro de estos, la UNESCO pone gran atención a la relevancia al aumento de la transparencia como un medio óptimo para que las empresas de Internet rindan cuentas de sus operaciones, cuyo impacto repercute en ámbitos cada vez más amplios de la vida social (Puddephatt, 2021).

Asimismo, han definido que la transparencia es una garantía que permite asegurar otros derechos universales como la libertad y el acceso a la información. En este sentido, la UNESCO realiza un análisis sobre una doble dimensión que compone a la transparencia. Por un lado, como fue analizado anteriormente, la transparencia se encuentra involucrada con las estructuras de acceso a la información sobre el ADP. Por otra parte, este principio también involucra un deber activo de los agentes responsables de divulgar proactivamente sus acciones y asegurar la construcción de una transparencia relacional con sus usuarios. Para ello, la UNESCO ha desarrollado, en sus más recientes informes, una serie de requisitos de transparencia que pueden ser relevantes para la regulación de los agentes que llevan a cabo el ADP en sus procesos de IA (Puddepahtt, 2021); que, sin embargo, su análisis excede la propuesta del presente trabajo, pero aun así consideramos que los mismos son de vital importancia para asegurar la transparencia en la IA.

4. Transparencia algorítmica

La transparencia, según es analizada por diversos expertos en el campo de la regulación de IA, es uno de los principios fundamentales para el desarrollo y uso de “IA responsable” (Felzmann et al., 2019). De acuerdo a lo considerado por Steffan Larsson (2020), la

transparencia es el principio más común y uno de los cinco más destacados entre aproximadamente 84 directrices técnicas que abordan la IA a nivel mundial (análisis realizado en el año 2020). En la actualidad, vemos como el número de guías y directrices que abordan el desarrollo y uso responsable de la IA ha crecido de manera exponencial y, a su vez, son aún más las recomendaciones normativas que incluyen a la transparencia en sus diferentes dimensiones, es decir, como un valor, un componente relacional y como un sistema integrado de rendición de cuentas (Larsson, 2020).

En la actualidad, definir a la transparencia algorítmica no es una tarea sencilla, puesto que, de acuerdo a lo exhibido anteriormente, tanto el concepto de "transparencia" como el de "inteligencia artificial" implican significados ambiguos que varían de acuerdo el contexto y la aplicación de estos. De esta manera, vincular ambos conceptos dentro de un marco regulatorio en específico cuenta con las mismas dificultades de abstracción que cada elemento por separado.

En la actualidad, dada la gran utilización de la IA en nuestras vidas, las preocupaciones en torno a la privacidad y transparencia se encuentran relacionadas con los procesos de IA y los procesos algorítmicos que la componen para arribar a una determinada decisión o predicción (Algorithm-watch, 2019). Como se analizará posteriormente, estas preocupaciones han guiado la construcción de los preceptos vinculados a la transparencia de los algoritmos, para brindar seguridad en la operación diaria que se realiza con los modelos de IA, particularmente, en aquellos donde se llevan a cabo decisiones que pueden tener graves impactos sobre la salud, la educación, el empleo y el historial crediticio de las personas (Kossow, 2021).

Por otra parte, como se ha mencionado anteriormente, la transparencia es en sí mismo un concepto multifacético que, en

materia de protección de datos, contiene otros principios como la explicabilidad, el acceso a la información, la licitud y la responsabilidad en el ADP. Sin embargo, para analizar su vinculación con el ámbito de la IA, conviene entender que no existe una definición precisa de qué es lo que hace que un sistema algorítmico sea transparente (Kossow, 2021). Aun cuando la regulación actual considera que los riesgos asociados al tratamiento y decisiones automatizadas pueden ser salvados por la transparencia (Burrell, 2016), este principio no es una noción de blancos o negros, por lo que un sistema de IA no puede ser considerado “transparente” u “opaco” en su totalidad.

Por su parte, los recientes lineamientos que consideran la responsabilidad en el uso de la IA abordan la transparencia no solo desde la óptica de que es un principio que inunda de soluciones a los problemas de opacidad de los algorítmicos, sino también, reconoce que existen diferentes grados de transparencia. De esta manera, la transparencia algorítmica dependerá de una combinación entre sus procesos técnicos y la audiencia con la que colabore (público en general, los reguladores, investigadores, compañías).

Un ejemplo de este abordaje es considerar que este principio atraviesa los modelos de IA en diferentes etapas de funcionamiento. Desde la perspectiva informativa, las explicaciones sobre el funcionamiento de un determinado sistema de IA constituyen el cumplimiento de la transparencia intrínseca al modelo, revela cuál es la función de los sistemas, sus fines, inferencias, entre otros. Por otra parte, el principio de transparencia algorítmica en sentido estricto podría considerarse como la acción de los desarrolladores de revelar el conjunto algorítmico que compone el modelo para su auditoría, según corresponda. En tercer y último lugar, la transparencia algorítmica ex post, aplicada sobre los modelos con cierta opacidad y dificultades

para revelar su funcionamiento, se representa mediante la incorporación de técnicas de interpretabilidad que le permiten al usuario comprender su funcionamiento (Gutiérrez, 2021).

A partir de esta distinción realizada por David Gutiérrez (2021) podemos inferir que el concepto de transparencia algorítmica es posible de implementar a sistemas de IA de diferente tipo, considerando la aplicación alternada o concurrente de dos elementos: la transparencia técnica y la interpretabilidad del modelo.

El primer concepto, la transparencia técnica, refiere a una característica ex ante del modelo de IA, es decir, incluir transparencia desde el diseño de los modelos algorítmicos, de manera que se permita la posterior comprensibilidad e interpretación de su funcionamiento global, los elementos individuales y el algoritmo de aprendizaje utilizado para un humano experto o no en el sistema (Cotino Hueso, 2022). De esta manera, los modelos que pudieran ser configurados desde su desarrollo de acuerdo a los requisitos de transparencia técnica, podrían considerarse como sistemas de "IA responsable", puesto que no tienen impedimentos para brindar explicaciones sobre su funcionamiento y resultados.

De acuerdo a esta interpretación, la transparencia algorítmica no puede ser desligada de otros conceptos como la explicabilidad, la descomponibilidad y simulabilidad la transparencia algorítmica en sentido técnico (NIST, 2021). Para estos conceptos, el NIST (Instituto Nacional de Estándares y Tecnología de los Estados Unidos) brinda las siguientes definiciones.

La explicabilidad es un principio que establece que para que un algoritmo sea explicable debe ser capaz de proporcionar evidencias,

razonamientos o aclaraciones que sean capaces de justificar sus conclusiones.

La simulabilidad es precepto que supone que el modelo de AI analizado puede ser replicado por inteligencia humana en un tiempo razonable, considerando los mismos datos y parámetros matemáticos que utilizó el modelo para generar una predicción idéntica o similar;

La descomponibilidad es la capacidad de que un modelo pueda ser dividido de acuerdo a los parámetros, sistemas y algoritmos de aprendizaje que lo componen, para realizar un análisis intuitivo de cómo funciona y porqué arriba a ciertas decisiones o predicciones.

Por otra parte, existe la probabilidad de que, en el desarrollo de modelos de IA, no todos los sistemas puedan ser adaptados de acuerdo a los principios de la transparencia técnica. Estos sistemas serían aquellos que se clasifican como modelos de IA con opacidad para el análisis técnico y el ejercicio de la explicabilidad, acerca de sus implicancias y método automatizado de decisiones. Sin embargo, según reconoce Cotino Hueso (2022), sobre estos sistemas es posible aplicar otras medidas técnicas que permitan la interpretabilidad del modelo, en una suerte de “transparencia prospectiva”.

Para alcanzar la interpretabilidad de los resultados, en cumplimiento del principio de transparencia en algoritmos, según Cotino Hueso (2022), los métodos pueden variar de acuerdo al tipo de sistema que se trate. La pauta general es que cuando el procesamiento en la IA sea mediante un modelo opaco, se intentará descubrir las principales características del modelo o que es lo que se puede inferir del mismo, puesto que, aunque no sea técnicamente transparente, este proceso puede aportar información sobre el funcionamiento de sus algoritmos. Algunos ejemplos de estas técnicas podrían ser la simplificación de los

modelos, la visualización de cada una de las variables utilizadas y la conexión realizada entre las mismas, y el acceso al modelo de aprendizaje involucrado para conocer los parámetros de análisis inferidos para la estructuración de la información.

En este contexto, luego de haber analizado brevemente las implicancias de la ejecución de los principios de la transparencia algorítmica, es posible argumentar que existen aspectos positivos y negativos en torno a su aplicación real sobre los modelos de IA. Por ello, como se desarrollará a continuación, ambas posturas sostienen argumentos que, en la actualidad, son el foco del debate público en torno a la regulación de la IA.

4.1. ¿Es la transparencia un valor fundamental en la protección de datos personales en el contexto de los sistemas de IA?

La transparencia ha sido considerada en diferentes disciplinas como la económica, la política, la sociología y el derecho, como un valor positivo que, a menudo no tiene una definición en concreto (Meijer, 2014; Felzmann, 2020). Por ello, para realizar un análisis acerca de los efectos positivos de la transparencia, únicamente recogeremos sus "fases" como valor informativo y relacional entre los agentes, dejando de lado la relación sistémica de la transparencia.

El principio de transparencia, en su consideración general, se entiende como una garantía que permite la transferencia de información de un agente responsable hacia otro. Por esta razón, vemos como la regulación ha tomado este concepto y lo ha asociado con el deber de informar sobre el funcionamiento de los sistemas involucrados en la toma de decisiones automatizadas. Sin embargo, la transparencia algorítmica va más allá del simple deber de información, puesto que le subyacen otros valores sociales y normativos vinculados a la acción de "informar" (Felzmann, 2019).

Dentro de la consideración informativa, la transparencia es asociada con una herramienta que permite superar las dificultades que genera la asimetría de la información dentro de las relaciones. Las obligaciones impuestas por los lineamientos regulatorios de IA establecen que en favor de utilizar la IA de manera “transparente”, los responsables deben transferir cierta información de la esfera privada de sus desarrollos al público interesado en su funcionamiento, sea este, usuario o no del sistema de IA. Esta apertura y acceso a la información, garantiza la igualdad

En el análisis legal del valor informativo de la transparencia, algunos expertos como Felzmann y Fosch-Villaronga (2020) han propuesto que asegurar el acceso a la información tiene consecuencias positivas dado que fomenta la confianza, facilita la rendición de cuentas y permite un mayor nivel de control sobre los sistemas. De esta manera, vemos que este enfoque permite lograr algunos de los ideales pensados para los sistemas inteligentes de automatización de los procesos, como la explicabilidad y la rendición de cuentas. Ambos conceptos, son principios “hermanos” de la transparencia y colaboran en su efectiva aplicación.

En primer lugar, la explicabilidad en sí misma es un principio valorado y ampliamente estudiado dentro del desarrollo ético de la IA. En conjunto con la transparencia, permite que los sujetos objeto de decisiones automatizadas puedan ejercer su derecho a obtener información sobre la lógica algorítmica utilizada para arribar a los resultados y, en su caso, poder impugnar el contenido de estas.

Esta aplicación se refleja en las actuales normas y directrices públicas, como el RGPD realizado por la Unión Europea, la “Recomendación sobre la Ética de la Inteligencia Artificial” de la UNESCO (2021), las “Recomendaciones para la Inteligencia Artificial Fiable” publicada

mediante la Disposición Nº 2/2023 por la Subsecretaría de Tecnologías de la Información argentina; y reglamentos privados como los “Estándares de IA responsable” propuestos por Microsoft en sus dos versiones, que brinda un enfoque del sector privado involucrado en la protección de los principios generales para la “IA segura”.

La explicabilidad y la transparencia algorítmica, al mismo tiempo implican los conceptos desarrollados como “transparencia prospectiva y retrospectiva” (Felzmann et al., 2019). La transparencia prospectiva se refiere a la obligación de informar a los usuarios sobre el funcionamiento del sistema y el procesamiento de datos antes de la recolección de los datos. En esta etapa, la transparencia se ejecuta de acuerdo al concepto recogido por los conjuntos normativos vigentes, respecto al deber de informar y obtener el consentimiento previo al tratamiento automatizado de datos personales. De esta manera, la transparencia prospectiva se constituye como un sistema de preaviso respecto del uso y tratamiento de datos. Por otra parte, la transparencia retrospectiva, refiere a las explicaciones y justificaciones que proceden luego de la toma de decisiones automatizadas. En este caso, el principio de transparencia en conjunto con la explicabilidad, revelan la obligación de explicar para un caso concreto cómo y por qué se llegó a una determinada decisión, describiendo todo el funcionamiento del procesamiento de datos. En otras palabras, la transparencia implicaría poder realizar un proceso de “auditoría” al sistema, descomponiendo un resultado para comprender la estructura algorítmica y el sistema de pesos dentro del sistema que llevó a un cierto resultado o predicción (Zerilli, 2018).

A su vez, desde la perspectiva informativa de la transparencia, la rendición de cuentas es otro de sus valores positivos, en tanto permite que, a partir de la información obtenida por los requisitos de

transparencia, se facilite el proceso de una rendición de cuentas respecto a la toma automatizada de las decisiones (Felzmann, 2020). En este sentido, la rendición de cuentas presupone que los usuarios tienen la información suficiente para poder comprender los resultados de su rendimiento. En este punto, la transparencia es fundamental para garantizar que en el suministro de "explicaciones", los agentes comprendan las implicancias y consecuencias brindadas por quienes desarrollan los sistemas de IA.

En último lugar, los autores Felzmann y Fosch-Villaronga (2020), proponen que la transparencia es beneficiosa, dado que favorece el fortalecimiento de la confianza entre el ecosistema que ejecuta la IA y quienes se convierten en un usuario de la misma. La confianza como tal, permite que las relaciones se ejecuten bajo un marco de respeto e integridad.

Si se analiza esta última concepción, la transparencia algorítmica fortalece el rol de los actores públicos y privados que llevan a cabo el desarrollo económico de la IA, puesto que, al disponer y acercar información a los agentes, sean usuarios o no, revelan cierta integridad y compromiso en no ocultar las funciones detrás de la toma de decisiones automatizadas. Sin embargo, a pesar de estos argumentos que favorecen a la transparencia como un valor fundamental para la protección de la privacidad en la toma de decisiones automatizadas generadas por IA, existen posiciones que cuestionan su utilidad en estas tecnologías.

4.2. Desafíos y críticas al principio de transparencia algorítmica

A pesar de los beneficios mencionados en el apartado anterior, la transparencia en su función informativa es criticada por algunos problemas en torno a su concepto y aplicación práctica. A continuación, se realizará un breve análisis sobre los principales

puntos de debate sobre la transparencia algorítmica. Por un lado, sus limitaciones respecto al carácter “informativo” de su concepción tradicional. En segundo lugar, se analizarán los problemas relacionados a la aplicación técnica de la transparencia algorítmica en el marco de los desafíos técnicos presentes en los sistemas de inteligencia artificial.

En primer lugar, el análisis conceptual de la transparencia como principio “informativo”, generalmente, tiene una connotación positiva, dado que la misma palabra “transparencia” nos remite a la integridad y la honestidad. Sin embargo, a pesar de su interpretación positiva, algunos expertos en regulaciones privacidad consideran que no es suficiente con solo garantizar la transparencia algorítmica para constituir sistemas de IA responsables.

Una crítica frecuente es que el acceso a la información no implica solamente abrir los códigos algoritmos que componen a un sistema de decisiones automatizadas, sino que, para que verdaderamente se constituya la transparencia, se requiere que esta información sea recibida. Este último requisito implica considerar que la transparencia se realiza solo cuando es posible garantizar que los agentes interesados pueden interpretar y comprender la totalidad del sistema, cuestión que resulta prácticamente imposible en el contexto de la dificultad técnica computacional.

En otras palabras, para llevar a cabo un análisis sobre la transparencia de los sistemas en la IA, en primer lugar, es necesario sortear las dificultades técnicas en torno a las explicaciones de los resultados a los que aborda. Este fenómeno es considerado como un problema de opacidad y explicabilidad de los algoritmos por la complejidad de las tecnologías de la IA.

Los actuales modelos de IA, como el aprendizaje automático, el Deep learning y el GPT (Generative Pre Trained), funcionan como un sistema de redes neuronales entrenadas con grandes volúmenes de datos y configuraciones algorítmicas avanzadas de autoaprendizaje. Estos desarrollos son técnicamente complejos, sin embargo, no resulta imposible brindar una explicación técnica sobre la conexión que realizan las "neuronas" que la componen y el tratamiento realizado sobre los datos para arrojar un resultado en concreto, esta explicación no demuestra un fundamento que sea relevante para el usuario, en tanto es de difícil comprensión para los sujetos que no se encuentran familiarizados con términos de la ciencia computacional de datos; e incluso, en ciertas circunstancias, ni siquiera es claro para los mismos desarrolladores de IA. A este fenómeno, en el último tiempo se lo denomina como la "caja negra" de la IA (Burell, 2016).

Asimismo, la vinculación antes realizada entre el principio de transparencia y la rendición de cuentas también puede ser cuestionada, dados los problemas de problemas de agencia que atraviesan los mecanismos de IA. En otros términos, en los modelos inteligentes, puede ser complejo establecer cuál es el mecanismo detrás de una decisión, cuando los sistemas involucrados en su generación son el resultado de una cadena de procesos en los cuales se involucraron diversos agentes para la recopilación y tratamiento de los datos utilizados en esa decisión.

Finalmente, algunos expertos plantean que la transparencia no refuerza la protección de los datos personales, sino que, por el contrario, puede generar problemas de privacidad de datos personales utilizados para el entrenamiento algorítmico de la IA. En este sentido, el riesgo de visibilizar el mecanismo y la información utilizada para arribar a una determinada una decisión, puede implicar que se revelen

datos personales utilizados para alimentar los algoritmos de aprendizaje automático. Esta información, según Lazmann, se encuentra disponible en la mayoría de los sistemas de la toma de decisiones, puesto que para que los algoritmos sean lo más justo y diverso posible, se incluyen datos de poblaciones vulnerables o categorizadas de manera que se evite la exclusión por sesgos preconfigurados por los datos de entrenamiento. Frente a esto, la transparencia requeriría que se exponga esa información, afectando la privacidad de datos que el mismo principio pretende proteger.

Por ello, ante estas dificultades, algunos expertos consideran que la transparencia algorítmica, en el carácter informacional con el que se la concibe, no es posible y debería buscarse otros principios que aseguren la privacidad de datos en el contexto de las decisiones tomadas por algoritmos, puesto que la transparencia solo funcionara como un mecanismo estático para la disponibilidad de la información (Tielenburg, 2018). Otras posturas neutrales, en cambio, proponen que, frente a las dificultades técnicas, quizás sería ideal limitar el desarrollo de la IA a algoritmos de compresión simple, en los cuales se pueda ejecutar el principio de transparencia y explicabilidad (Adadi y Berrada 2018).

4.3. Caso de análisis: Azure OpenAI Services

Para reflejar lo anteriormente expuesto, se analizará como ejemplo, el reciente desarrollo del producto “Azure OpenAI Services” (en adelante, “Azure” o “Azure OpenAI”) de la compañía Microsoft y como se aplica este análisis de transparencia en su desarrollo y utilización, así como también, revisar cual es la propuesta autorregulatoria de la compañía.

Actualmente, la IA y sus capacidades generativas, están transformando todo el entorno de la toma de decisiones automatizadas, mejorando la eficiencia y productividad de diversos

procesos. Azure, como modelo algorítmico, se compone de otros modelos de IAG (GPT-3, GPT-4 y Codex) que permiten la generación de texto, código e imágenes (Microsoft, 2023). Según explica la compañía encargada de su desarrollo, estos modelos utilizan una arquitectura autorregresiva, es decir, son sistemas predictivos que se basan en la información de entrenamiento para la predicción de la siguiente palabra más probable. Este proceso sucede repetidas veces, con diversas capas de atención y refinamiento del contenido, hasta la generación del texto completo.

De acuerdo con lo expuesto en el sitio "*Transparency Notes for Azure OpenAI*", Microsoft informa, de acuerdo al principio de transparencia, que los modelos que componen este producto se encuentran entrenados con una variedad de datos obtenidos de diferentes sitios web. En primer lugar, como fue comentado durante el auge mediático de ChatGPT-3, este modelo fue entrenado con los datos que provienen de la "Common Crawl" y otros datos que mejoraron la calidad del contenido de preentrenamiento. Por su parte, el modelo mejorado de ChatGPT-3, ChatGPT-4, se entrenó con todos los datos disponibles en la Word Wide Web (en otros términos, todos los datos públicos de internet), algunos datos con licencia de OpenAI e inferencias incorporadas mediante el aprendizaje por refuerzo, es decir, mediante intervención humana en ciertas configuraciones (Microsoft, 2023).

Por esta razón, esta nueva tecnología tiene un gran potencial de aplicación en los procesos de toma de decisiones y la generación de contenido, dada su precisión y capacidad analítica de los datos de entrada. Algunas de sus aplicaciones pueden ser el análisis de datos y estadísticas, interacciones de chat, creación de códigos, búsqueda, asistencia de escritura y resumen de textos, entre otros. Todos estos

puntos, hacen que el producto sea comercialmente atractivo para las empresas, puesto que les permite integrar toda su información y mejorar los procesos desde la aplicación de estos algoritmos de inteligencia artificial generativa. Sin embargo, Microsoft reconoce que, a pesar del potencial de la herramienta propuesta y los esfuerzos que se utilizaron para preservar la transparencia en su proceso de desarrollo, aún existen limitaciones para garantizar el cumplimiento de todos los requisitos impuestos para una “IA responsable”.

En estas circunstancias, Microsoft recomienda a sus usuarios que, por sus riesgos en transparencia, Azure OpenAI no debería ser utilizado en contextos de, por ejemplo, decisiones legales, crediticias y financieras, que pueden generar un menoscabo directo en los derechos y el curso de la vida de los sujetos involucrados bajo la automatización de las decisiones.

Desde el punto de vista de análisis de los desafíos de transparencia en este modelo, en primer lugar, es posible mencionar que dada la concepción informativa de la transparencia y la explicabilidad, vemos que este tipo de inteligencia artificial réplica alguna de las dificultades que tienen otros modelos en brindar una explicación sobre los resultados obtenidos en la automatización de decisiones. Esto significa que aun cuando los desarrolladores de IAG brindan ciertas “explicaciones” sobre cómo, para qué y qué tecnología se involucra, no será suficiente para cumplir con los parámetros informativos relacionales de la transparencia algorítmica. En otras palabras, aun cuando se revele el código, el mecanismo y la fuente de los datos originales de entrenamiento, la explicación en sí misma no satisface los requisitos de transparencia, puesto que el receptor podría no comprender los aspectos técnicos de la explicación.

Por otra parte, su composición técnica no es solo un problema al momento de la explicabilidad para agentes terceros al sistema, sino también para los mismos desarrolladores del modelo Azure OpenAI. En otras palabras, este tipo de IA cuenta con sistemas de aprendizaje profundo y automático, que, junto con otras tecnologías como la Atención, las Redes Adversarias Generativas (GANs) y el modelo Transformer, constituyen un algoritmo prácticamente autónomo de la intervención humana; lo que se traduce en que, incluso los mismos programadores de estos algoritmos, pueden atravesar dificultades para descomponer el proceso que realizo la IAG para obtener ciertos resultados o decisiones. Si bien lo expuesto representa un desafío necesario de ser abordado desde la protección a la privacidad de datos personales, su tratamiento excede los alcances del presente trabajo, por lo cual, se continuará con el desarrollo del segundo punto de interés.

En segundo lugar, otro de los desafíos vinculados al desarrollo de esta nueva tecnología es que la aplicación del principio de transparencia implique necesariamente revelar datos o información sensible utilizada durante el entrenamiento de estos algoritmos. Como se ha debatido públicamente, los modelos GPT-3 y GPT-4 han sido entrenados bajo billones de datos obtenidos de una multiplicidad de sitios disponibles en la web, dentro de los cuales pueden encontrarse comprometidos datos personales que el algoritmo no detecta como tal y los procesa para automatizar las decisiones requeridas por el sistema[4]. Asimismo,

[4] Se entiende que aun así, en la actualidad, se llevan a cabo procesos de anonimización y *data aggregation* en el proceso de recolección y tratamiento de datos para el entrenamiento de los modelos de IAG para garantizar que no se puedan identificar datos personales de los individuos a partir de los resultados generados. Algunas técnicas de anonimizacion puede ser el enmascaramiento, la perturbación, la generalización, la privacidad diferencial y las normas que equilibran la utilización de datos y la preservación de la privacidad.

por sus características técnicas, este tipo de tecnología cuenta con una opacidad algorítmica de distintos niveles, que involucra infinitas capas de procesamiento de datos, en las cuales pueden encontrarse ocultos datos sensibles que ni siquiera el programador del algoritmo, reconoce haber utilizado (Cotino Hueso, 2022).

Sin embargo, a pesar de estos desafíos, y la imposibilidad temporal de los reguladores de acompañar cada uno de los desarrollos tecnológicos en un tiempo adecuado, otros miembros del sector privado como Meta y Microsoft, han propuesto diferentes guías e informes de transparencia que permiten a los agentes realizar un seguimiento sobre el proceso de las decisiones en la IAG (Manguillot, 2022). En este sentido, junto con el lanzamiento de Azure OpenAI Services, la empresa reconoce en su "Guía para la IA responsable" que la transparencia implica que los desarrolladores de IA deben ser abiertos sobre cómo y porqué se utilizan estos sistemas, y ser abiertos sobre sus limitaciones. Además, proponen que el comportamiento de los sistemas inteligentes debe ser comprendido por los usuarios.

En este sentido, observamos que, aún existen limitaciones que impiden la transparencia relacional y sistémica en la utilización de IAG, y estos tienen que ver, fundamentalmente, con la complejidad técnica que involucran estas tecnologías. Aun así, desde una visión en particular de análisis, consideramos que, para asegurar la transparencia informativa, relacional y sistémica, no se requiere que los usuarios o agentes interesados puedan comprender en su totalidad cómo funcionan los modelos algorítmicos que componen la IA, sino que basta con que existan parámetros establecidos por expertos en el área, que sean implementados desde el diseño de nuevos productos y aplicaciones de estas tecnologías. Si bien la transparencia,

explicabilidad y rastreabilidad son conceptos relevantes para la justicia y la protección de los derechos, en otras ocasiones frente a debates similares, la sociedad ha logrado ponderar los beneficios de un "suceso tecnológico" (como el desarrollo de vacunas, la aparición de automóviles, el uso de redes sociales, entre otros ejemplos) por sobre los riesgos de no conocer en su totalidad el funcionamiento y consecuencias de su uso. La pretensión de la transparencia algorítmica clásica que "todos puedan comprender el funcionamiento de los sistemas de IA en los que se involucran" puede ser considerado como únicamente un ideal impracticable en la realidad. Por lo que quizás, la exposición de advertencias para su uso y la educación sobre las limitaciones y riesgos de estos sistemas, podrían ser consideradas como alternativas viables para garantizar la transparencia algorítmica y un uso responsable de la IA.

5. Conclusión

La IA y el uso de modelos funcionales han generado una revolución, con un potencial irrefutable para mejorar nuestras vidas. Con su evolución, predecir, evaluar y decidir en ámbitos de relevancia, como la educación, economía, salud y cuestiones sociales, se convirtió en una tarea algorítmica realizada de manera automática con poca –o nula- intervención humana (Araya Paz, 2021). Esta automatización tiene un gran potencial para colaborar con el avance de la sociedad en general, sin embargo, como fue analizado durante este trabajo, también existen grandes riesgos en su utilización, por lo que resulta importante analizar sus puntos de opacidad y eventuales riesgos, para combatirlos con principios éticos y legales capaces de contener sus efectos.

Durante este trabajo, brindamos un breve análisis sobre cómo, a partir de la concepción original de la privacidad, las normas de protección a

la privacidad de los datos personales han intentado acompañar el desarrollo de nuevas tecnologías, hasta los actuales desafíos vinculados a la IA automatizada. En presencia de esto, pusimos particular atención al desarrollo conceptual del "principio de transparencia algorítmica" que, desde nuestra perspectiva de análisis, constituye el elemento central en las normativas de protección de la privacidad de los datos.

La idea de transparencia en el ecosistema digital es de relevancia dado que, por las implicancias técnicas de la IA y el uso de modelos automatizados, poseen la capacidad de afectar de forma directa a las personas y sus derechos. Frente a ello, la transparencia algorítmica se presenta como una garantía central dentro del núcleo de la protección de la privacidad de los datos, en tanto atraviesa el ejercicio de otros principios y garantías contenidos dentro de los conjuntos normativos como la explicabilidad, la rendición de cuentas, la no discriminación, la licitud y la responsabilidad.

Por lo tanto, luego del análisis realizado acerca de las implicancias positivas y negativas del principio de transparencia algorítmica, concluimos que la transparencia es un valor esencial que, dentro del contexto de la regulación de IA, debe ser considerada como el concepto que permite en sí, la protección de la privacidad de datos de las personas.

Para arribar a esta conclusión, en primer lugar, consideramos que, por la naturaleza técnica de la IA, aplicar la concepción tradicional de los principios asociados a protección de datos personales, puede significar un perjuicio grave para el desarrollo de los modelos algorítmicos, puesto que, como fue analizado, el entrenamiento, uso y perfeccionamiento de estos sistemas dependen necesariamente de grandes volúmenes de información, dentro de la cual, puede

encontrarse incluida, información personal. Por ello, la transparencia constituye un abordaje diferente para el principio de protección de datos tradicional, pero conservando los mismos fines: evitar que el uso de información personal provoque un detrimento en los derechos y libertades de los sujetos.

Tal como hemos aprendido de otras grandes revoluciones tecnológicas en la historia, la solución no radica en restringir estas innovaciones hasta el punto de extinguirlas o perjudicar su progreso. Más bien, la clave está en asegurar que las personas comprendan estos procesos, permitiéndoles utilizar estas tecnologías con conocimiento pleno de las consecuencias y repercusiones que conllevan. Un ejemplo claro de esto se observa en la aparición de los automóviles, donde inicialmente surgieron múltiples cuestionamientos en torno a los riesgos y posibles consecuencias en su uso. Sin embargo, no se requirió que las personas fueran expertas en ingeniería automotriz para permitir su desarrollo; en cambio, con el tiempo, la educación y el conocimiento general sobre su funcionamiento, beneficios y posibles impactos permitieron que esta tecnología se integrará de manera responsable en nuestras vidas.

Así pues, si trasladamos esta analogía al desarrollo de la IA, emerge el principio de transparencia algorítmica, que exige a los desarrolladores proporcionar información sobre el uso responsable de la IA y sus aspectos relacionales. Esto permite que las personas puedan ejercer su derecho a la protección de los datos personales mediante el consentimiento informado acerca del tratamiento automatizado de estos.

Por lo tanto, en conclusión, la transparencia algorítmica no se limita meramente a constituir uno de los principios de la IA responsable, sino que se erige como un elemento esencial, una garantía fundamental

que permite a los individuos ejercer el debido control sobre el tratamiento automatizado de sus datos en el complejo contexto de los nuevos desafíos planteados por esta tecnología. No obstante, es crucial reconocer que esta afirmación no aboga por restringir exclusivamente el desarrollo de la IA a su vertiente transparente, ya que tal enfoque resultaría excesivamente restrictivo y perjudicial para su progresión. Más bien, consideramos que la transparencia representa el principio rector mediante el cual podemos asegurar la salvaguarda adecuada de los datos personales en el uso de estas tecnologías, basándonos en la plena comprensión, el consentimiento informado y la explicabilidad. A través de esta conjunción, logramos capitalizar los beneficios de la IA sin comprometer la integridad de la privacidad y la seguridad individual, promoviendo de esta manera una integración responsable de esta innovación en nuestras vidas.

Referencias bibliográficas

Acquisti, Brandimarte y Loewenstein. 2020. “Secrets and Likes: The Drive for Privacy and the Difficulty of Achieving It in the Digital Age”. Journal of Consumer Psychology 30, num. 4., Disponible en: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3688497

Adadi, A., y Berrada, M.. 2018. “Peeking inside the black-box: A survey on Explainable Artificial Intelligence (XAI)”. En IEEE Access, vol.6, pág. 52138-52160. Disponible en: https://ieeexplore.ieee.org/document/8466590

AlgorithmWatch. 2019. “Automating society: Taking stock of automated decision making in the EU”. Reporte de AlgorithmWatch, primera edición, enero de 2019. Disponible: https:// algorithmwatch.org/

Araya Paz, Carlos. 2021. “Transparencia algorítmica ¿un problema normativo o tecnológico?”. En revista CUHSO (Temuco), vol. 31, no. 2, dic. 2021. Disponible en: https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S2452-610X2021000200306

Basterra, Marcela. 2013. “La Garantía Constitucional de Habeas Data. Lineamientos Generales de la Ley de Protección de Datos Personales”. Disponible en:https://marcelabasterra.com.ar/wp-content/uploads/2016/11/La-Garanti%cc%81a-Constitucional-de-Habeas-Data.-Lineamientos-generales.pdf

Burrell J.. 2016. “How the machine ‘thinks’: Understanding opacity in machine learning algorithms”. Big Data & Society ed. 3. Disponible en: https://www.morgan-klaus.com/readings/opacity-ml.html

Casey B., Farhangi A. y Vogl R.. 2019. “Rethinking Explainable Machines: The GDPR's 'Right to Explanation' Debate and the Rise of Algorithmic Audits in Enterprise”. Berkeley Technology Law Journal, Vol. 34, 2019, Disponible en: https://ssrn.com/abstract=3143325

Cotino Hueso, Lorenzo. 2022. “Transparencia y explicabilidad de la Inteligencia Artificial. Elementos conceptuales, generales y de género”. Transparencia y explicabilidad de la inteligencia artificial. Tirant Lo Blanch. Universitat de València, págs 29-70. Disponible en: https://editorial.tirant.com/es/ebook/transparencia-y-explicabilidad-de-la-inteligencia-artificial-lorenzo-cotino-hueso-9788411471602

Dreyfus, Hubert L. 1999. “Anonymity versus commitment: The dangers of education on the internet”. Ethics and Information Technology. Disponible en: https://doi.org/10.1023/A:1010010325208

European Parliament. 2023. “AI Act: a step closer to the first rules on Artificial Intelligence”. News European Parliament, Press Releases 11-05-2023. Disponible en:https://www.europarl.europa.eu/news/en/press-room/20230505IPR84904/ai-act-a-step-closer-to-the-first-rules-on-artificial-intelligence

Felzmann H., Lutz C., Fosch Villaronga E., Tamò Larrieux A.. 2020. “Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns”. Big Data & Society, num. 6. Disponible en: https://www.researchgate.net/publication/333918635_Transparency_you_can_trust_Transparency_requirements_for_artificial_intelligence_between_legal_norms_and_contextual_concerns/citations

Felzmann H, Fosch-Villaronga E, Lutz C, et al. 2019. “Robots and transparency: The multiple dimensions of transparency in the context of robot technologies”. IEEE Robotics & Automation Magazine num. 26. Disponible en: https://www.researchgate.net/publication/332073237_Robots_and_Transparency_The_Multiple_Dimensions_of_Transparency_in_the_Context_of_Robot_Technologies

Fernandez D. y Barbero J.. 2023. “Se presentó ante el Congreso Nacional argentino un nuevo proyecto de ley para reemplazar la actual Ley de Protección de Datos Personales”. IAPP, 1 de agosto de 2023. Disponible en: https://iapp.org/news/a/se-presento-ante-el-congreso-nacional-argentino-un-nuevo-proyecto-de-ley-para-reemplazar-la-actual-ley-de-proteccion-de-datos-personales/

García Herrero, Jorge. 2020. “Decisiones automatizadas, profiling, Inteligencia artificial ¿Qué son?”. Jorge Garcia Herrero Blog. Disponible en: https://jorgegarciaherrero.com/decisiones-automatizadas-profiling-inteligencia-artificial-que-son/

Gutiérrez David, María Estrella. 2021. “Administraciones inteligentes y acceso al código fuente y los algoritmos públicos. Conjurando riesgos de cajas negras decisionales”. Universidad Complutense de Madrid. Disponible en:http://www.derecom.com/secciones/articulos-de-fondo/item/436-administraciones-inteligentes-y-acceso-al-codigo-

fuente-y-los-algoritmos-publicos-conjurando-riesgos-de-cajas-negras-decisionales

High-Level Expert Group on Artificial Intelligence (AI HLEG). 2018. “The Ethics Guidelines for Trustworthy Artificial Intelligence (AI)”. European Commission, European AI Alliance, AI strategy. Disponible en: https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html

Kettle, Blair. 2022. “In 1957 Herbert Simon Made Four Predictions About the Future of Computing”. Recuperado de LinkedIn: https://www.linkedin.com/pulse/1957-herbert-simon-made-four-predictions-future-computing-kettle/

Kossow N., Windwehr S. y Jenkins M.. 2021. “Algorithmic transparency and accountability”. Transparency International New Zealand. Disponible en: https://knowledgehub.transparency.org/assets/uploads/kproducts/Algorithmic-Transparency_2021.pdf

Larsson S. y Heintz, F. 2020. “Transparency in artificial intelligence”. Internet Policy Review num. 9. Disponible en: https://lucris.lub.lu.se/ws/files/79208055/Larsson_Heintz_2020_Transparency_in_artificial_intelligence_2020_05_05.pdf

Lu, Sylvia. 2022. “Data Privacy, Human Rights, and Algorithmic Opacity”. California Law Review 110 No. 6. Diciembre de 2022: 2087-2147.

Marr, Bernard. 2018. "How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read." Revista Forbes. 21 de mayo de 2018. Disponible en: https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/?sh=467da8d60ba9

Manguillot, Alfred Peris. 2022. “Algoritmos: ¿podemos hacerlos transparentes y trazables en su proceso?”. Transparencia y Explicabilidad de la Inteligencia Artificial. Tirant Lo Blanch. Universitat de València, págs 71-82. Disponible en: https://editorial.tirant.com/es/ebook/transparencia-y-explicabilidad-de-la-inteligencia-artificial-lorenzo-cotino-hueso-9788411471602

Mantegna, Micaela. 2022. “No soy un robot: construyendo un marco etico accionable para analizar las dimensiones de impacto de la Inteligencia artificial”. ARTEficial: creatividad, inteligencia artificial y

derecho de autor. CETyS. Universidad de San Andrés. Pág 24-29. Disponible en: https://repositorio.udesa.edu.ar/jspui/bitstream/10908/19109/1/%5bP%5d%5bW%5d%20-%20Mantegna.pdf

Meijer, Albert. 2014. “Understanding modern transparency”. International Review of Administrative, num. 75. Disponible: https://journals.sagepub.com/doi/10.1177/0020852309104175

Microsoft. 2023. “Transparency Notes for Azure OpenAI”. Disponible en: https://learn.microsoft.com/en-us/legal/cognitive-services/openai/transparency-note?tabs=text

NIST. 2023. “Artificial Intelligence Risk Management Framework”. Disponible en: https://www.nist.gov/itl/ai-risk-management-framework

Observatori d’Ètica en Intel-ligéncia Artificial de Catalunya (OEIAC). 2021. “Inteligencia artificial, ética y sociedad. Una mirada y discusion a través de la literatura especializada y de opiniones expertas”. Universitat de Girona. Disponible en: https://www.udg.edu/ca/Portals/57/OContent_Docs/Informe_OEIAC_2021_cast.pdf

OECD. 2019. “Artificial Intelligence in Society”. OCDE Publishing. Disponible en:https://www.oecd-ilibrary.org/sites/603ce8a2-es/index.html?itemId=/content/component/603ce8a2-es

Pasquale, Frank. 2015. “The Black Box Society: The Secret Algorithms that Control Money and Information”. Harvard University Press. Cambridge. Disponible en: http://www.jstor.org/stable/j.ctt13x0hch.

Puddephatt, Andrew. 2021. “Dejar entrar el sol: transparencia y responsabilidad en la era digital”. UNESCO. Disponible en: https://unesdoc.unesco.org/ark:/48223/pf0000377231_spa

Saldaña Diaz, Maria Nieves. 2012. “The Right to Privacy ". La génesis de la protección de la privacidad en el sistema constitucional norteamericano”. Revista de derecho político Nº 85, UNED, España.

Sartor, Giovanni. 2020. “The impact of the General Data Protection Regulation (GDPR) on artificial intelligence”. Panel for the Future of Science and Technology. European Parliamentary Research Service (EPRS), Scientific Foresight Unit (STOA). Junio de 2020. Disponible en: https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU(2020)641530.

Szlak, Gabriela. 2022. “Tratamiento Automatizado de Datos Personales: Implicancias para la argentina con la aprobación del Convenio 108 modernizado”. Blog de Lerman & Szlak, 22 de diciembre de 2022. Disponible en: https://lermanszlak.com/es/tratamiento-automatizado-de-datos-personales-implicancias-para-la-argentina-con-la-aprobacion-del-convenio-108-modernizado/

Tielenburg, Daan Simon. 2018. “The ‘dark sides’ of transparency: Rethinking information disclosure as a social praxis”. Universidad Utrecht. Disponible en:https://studenttheses.uu.nl/bitstream/handle/20.500.12932/31348/Tielenburg.pdf?sequence=2&isAllowed=y

Travieso, Juan Antonio. 2017. Derecho a la protección de los Datos Personales. Ministerio de Salud, Presidencia de la Nación, Argentina. Disponible en: https://salud.gob.ar/dels/printpdf/98

Visintini G., Busnelli F., Pérez A., Scalzini S., Woolcott-Oyague O., Monje-Mayorca D. 2021. “Vicisitudes del derecho a la privacidad: Cuestiones sobre el tratamiento de datos personales y la responsabilidad civil”. Universidad Católica de Colombia. Disponible en: https://repository.ucatolica.edu.co/entities/publication/1b79ab98-2e70-4f20-b3dd-259151e6200b

Weller, Adrian. 2017. “Transparency: Motivations and Challenges”. En ICML Workshop on Human Interpretability in Machine Learning (WHI), 5ta. edition. Disponible en: https://arxiv.org/pdf/1708.01870.pdf.

Westin, Alan. 1968. “Privacy and Freedom”. Washington and Lee Law Review Vol. 25, Issue 1.

Zerilli J., Knott A., Maclaurin J. y Gavaghan C.. 2019. “Transparency in Algorithmic and Human Decision-Making: Is There a Double Standard?”. En Philosophy & Technology num. 32(8). Disponible en: https://www.researchgate.net/publication/327448136_Transparency_in_Algorithmic_and_Human_Decision-Making_Is_There_a_Double_Standard

Legislación

Argentina. Honorable Congreso de la Nación Argentina. Protección de Datos Personales. Ley No. 25.326. Aprobada el 4 de octubre del 2000. Publicado en el B.O. del 2 de noviembre del 2000. http://servicios.infoleg.gob.ar/infolegInternet/verNorma.do?id=70368

Argentina. Honorable Congreso de la Nación Argentina. Protección de las personas con respecto al tratamiento automatizado de datos de carácter personal. Ley No. 27.699. Aprobada el 9 de noviembre de 2022. Publicada en el B.O. del 30 de noviembre de 2022. https://www.argentina.gob.ar/normativa/nacional/ley-27699-375738/texto

Council of Europe. Convention for the protection of individuals with regard to the processing of personal data. Convention 108+. Aprobado en junio de 2018.https://www.coe.int/es/web/data-protection/convention108-and-protocol

Unión Europea. Parlamento Europeo y Consejo de la Unión Europea. Reglamento General de Protección de Datos (RGPD). Reglamento (EU) 2016/679. Aprobado el 14 de abril de 2016. Aplicado el 25 de mayo de 2018. http://data.europa.eu/eli/reg/2016/679/2016-05-04