El tratamiento de datos en entornos digitales: los desafíos jurídicos de la inteligencia artificial (IA)

PID_00270349

Alessandro Mantelero

Tiempo mínimo de dedicación recomendado: 2 horas

Universitat Oberta de Catalunya

Alessandro Mantelero

Profesor titular de Derecho Privado y de *Data Ethics and Protection* del Politécnico de Turín. Es experto científico del Consejo de Europa para la protección de datos y para el Consejo de Europa (*Guidelines on personal data in a world of Big Data*, 2017; *Report on Artificial Intelligence and Data Protection: Challenges and Possible Remedies*, 2019; *Guidelines on Artificial Intelligence and Data Protection*, 2019). Es miembro coeditor de la revista *Computer Law and Security Review* y miembro del consejo editorial de las revistas *European Data Protection Law Review* e *IDP: Revista d'Internet, Dret i Política*.

El encargo y la creación de este recurso de aprendizaje UOC han sido coordinados por la profesora: Mònica Vilasau Solana (2020)

Primera edición: febrero 2020

Av. Tibidabo, 39-43, 08035 Barcelona
Realización editorial: FUOC

Índice

Introducción

La inteligencia artificial,[1] a pesar de la atención que está despertando recientemente, no constituye un tema nuevo en la investigación científica y en el debate sobre sus posibles consecuencias sociales. Si bien no es este el lugar para una amplia reflexión sobre el tema, ni para examinar la variedad de aplicaciones concretas que se suelen incluir en esta categoría, es necesario delimitar el objeto de la investigación y preguntarse por qué en los últimos años se ha abierto un debate a distintos niveles sobre este tema.

(1)En adelante IA, según el acrónimo inglés más común.

Lecturas recomendadas

W. S. McCulloch y otros (1943). «A Logical Calculus of the Ideas Immanent». *Nervous Activity. Bulletin of Mathematical Biophysics* (vol. 5, págs. 115-133).

A. M. Turing (1950). «Computing Machinery and Intelligence». *Mind* (vol. 49, págs. 433-460).

La definición de inteligencia artificial (IA)

No hay una definición unitaria de inteligencia artificial (IA). El término *inteligencia artificial* fue acuñado originalmente por John McCarthy, un informático estadounidense conocido como el padre de la IA. Con el término *inteligencia artificial* se suelen describir los sistemas informáticos que son capaces de aprender de sus propias experiencias y resolver problemas complejos en diferentes situaciones, habilidades que antes pensábamos que eran exclusivas del ser humano.

Esta es la definición de IA proporcionada por el Consejo de Europa:

«AI is actually a young discipline of about sixty years, which brings together sciences, theories and techniques (including mathematical logic, statistics, probabilities, computational neurobiology and computer science) and whose goal is to achieve the imitation by a machine of the cognitive abilities of a human being. Specialists generally prefer to use the exact names of the technologies actually used (which today are essentially machine learning) and are sometimes reluctant to use the term "intelligence" because the results, although extraordinary in some areas, are still modest compared to the stated ambitions.»

Council of Europe. «What's AI?» [en línea]. <www.coe.int/en/web/artificial-intelligence/what-is-ai>

Lectura recomendada

J. McCarthy; M. L. Minsky; N. Rochester; C. E. Shannon (1955). «A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence» [en línea]. <www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>

Al mismo tiempo, para abordar este debate desde una perspectiva reguladora, es necesario preguntarse qué formas de aplicación de la IA son razonables esperar en los próximos años, con el fin de definir correctamente el objeto de la regulación. En este sentido, hay que rechazar los escenarios apocalípticos o de ciencia ficción que prefiguran una inteligencia artificial comparable a la humana y que plantean preguntas sobre la subjetividad jurídica que parece carecer de fundamento con referencia al futuro cercano.

Este renovado interés por la IA se debe, como en el pasado, a una nueva fase de desarrollo de la investigación en este campo. Esta fase se refiere a los perfiles tecnológicos necesarios y funcionales para una aplicación concreta de la IA, más que a los fundamentos teóricos. En los últimos diez años, distintos factores han confluido para crear un entorno completamente nuevo y extremadamente fértil respecto la IA. Cabe destacar, entre estos factores, las transferencias de datos cada vez más rápidas y potentes a través de la red, una capacidad de almacenamiento significativamente mayor y la posibilidad de utilizar recursos computacionales potencialmente ilimitados gracias a la computación en la nube. Además, debe tenerse en cuenta la transformación progresiva en datos de incluso los eventos más insignificantes de la vida cotidiana de la mayoría de los individuos.

Lecturas recomendadas

V. Mayer-Schönberger; K. Cukier (2013). *Big Data. A Revolution That Will Transform How We Live, Work and Think* (pág. 78). Londres: John Murray.

M. Lycett (2013). «Datafication: making sense of (big) data in a complex world». *European Journal of Information Systems* (vol. 22, núm. 4, págs. 381-386).

El aprendizaje automático

La inteligencia artificial incluye formas diferentes de aprendizaje automático. El aprendizaje automático puede describirse como un conjunto de técnicas y herramientas que permiten que las computadoras puedan aprender nociones o entender hechos, creando algoritmos matemáticos basados en datos acumulados a gran escala que parece que pueden razonar independientemente de la intervención humana y también construir nuevos algoritmos.

El aprendizaje profundo es un modo de aprendizaje automático. Algunos tipos de aprendizaje profundo se basan en la llamada «red neuronal», que utiliza un conjunto conocido de datos de entrenamiento que ayudan a los algoritmos de autoaprendizaje a realizar una tarea. Esto está condicionado a que la propia red pueda determinar la respuesta correcta para resolver la tarea. Un ejemplo de aplicación de estas redes neurales fue el programa AlphaGo, que derrotó a uno de los más grandes campeones del mundo de Go.

En el pasado, las teorizaciones de la IA habían chocado con barreras tecnológicas que limitaban sus posibles aplicaciones. Ahora, los desarrollos recientes mencionados aquí han liberado este potencial. Esto ha llevado a un cambio, permitiendo nuevas formas de gestión de datos destinadas a extraer nuevos conocimientos, incluso de naturaleza predictiva. El *big data* y el aprendizaje automático son los productos más recientes de este proceso de desarrollo.

Lectura recomendada

The Norwegian Data Protection Authority (2018). *Artificial Intelligence and Privacy Report* [en línea]. <www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf>

En este sentido, debemos tener en cuenta las aplicaciones concretas de estas tecnologías, para entender qué tipo de IA poblará las aplicaciones de los próximos años, confirmando que todavía estamos muy lejos de la llamada «IA general», es decir, de un modelo de IA capaz de enfrentarse a cualquier tipo de problema y contexto, de modo similar a lo que le sucede a la mente humana. En este momento y en los próximos años, de hecho, solo nos hallaremos ante formas de IA capaces de realizar una o más tareas específicas, ciertamente con una gama variada de aplicaciones (desde el reconocimiento de imágenes, pasando por la traducción de textos, hasta las aplicaciones de juegos), pero siempre con un enfoque específico.

Al enfrentarse al tema de la regulación de la IA, es necesario enfocar el debate sobre las aplicaciones concretas y que se están desarrollando, sin considerar los escenarios más extremos sobre la interacción hombre-máquina. En este sentido, los algoritmos de IA actuales ya tienen un impacto en el uso de

datos personales y plantean interrogantes acerca de la idoneidad de las regulaciones existentes para abordar los problemas planteados por estos nuevos paradigmas.

Lecturas recomendadas

CNIL (2017). *How Can Humans Keep the Upper Hand? The Ethical Matters Raised by Algorithms and Artificial Intelligence. Report on the Public Debate Led by the French Data Protection Authority (CNIL) as Part of the Ethical Discussion Assignment Set by the Digital Republic Bill* [en línea]. <www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf>

N. Bostrom (2016). *Superintelligence paths, dangers, strategies*. Oxford: Oxford University Press.

R. Kurzweil (2016). *The singularity is near: when humans transcend biology*. Londres: Duckworth.

Además, el impacto de la IA no se refiere solamente al procesamiento de datos en sí mismo, sino a las finalidades aplicativas y a los principios y valores que deben guiar la implementación de la IA en casos concretos. De hecho, está surgiendo una tendencia hacia una sociedad tecnocrática impulsada por el mercado que conduce a la monetización de los datos personales, a formas de control social y soluciones «económicas y rápidas» de toma de decisiones tanto a gran escala (por ejemplo, en el contexto de las ciudades inteligentes) como a pequeña escala (por ejemplo, medicina de precisión, asistentes personales, dispositivos domésticos inteligentes, etc.). Esta tendencia plantea desafíos significativos para la protección de la autodeterminación individual y presenta problemas críticos para los modelos tradicionales centrados únicamente en la protección de la información personal.

Lecturas recomendadas

S. Speikermann (2016). *Ethical IT Innovation. A Value-Based System Design Approach* (pág. 152). Boca Raton: CRC Press.

A. Mantelero (2018). *Ciudadanía y gobernanza digital entre política, ética y derecho*. En: T. Quadra Salcedo; J. L. Piñar Mañas. *Sociedad Digital y Derecho*. Madrid: Boletín Oficial del Estado.

La bulimia de los datos, la complejidad del procesamiento y una lógica profundamente centrada en la medición de los fenómenos y las interacciones sociales pueden socavar el uso democrático de la información e imponer una especie de dictadura de datos en la que los modelos algorítmicos se desarrollan, aplican y utilizan para la toma de decisiones sin sentido crítico. Para evitar que las consecuencias adversas de la IA prevalezcan sobre los beneficios, es necesario preservar y reafirmar la centralidad del ser humano respecto al desarrollo tecnológico en general y, específicamente, en relación con la IA.

Lectura recomendada

J. L. Piñar Mañas (2018). *Derecho e innovación tecnológica. Retos de presente y futuro*. Madrid: CEU Ediciones.

Esto significa reafirmar el predominio de los derechos fundamentales en este ámbito. En este sentido, el derecho a la protección de los datos puede convertirse en el punto de partida para el desarrollo de una IA que no esté impulsada por el mero interés económico o por la eficiencia de los procesos, sino que sea capaz de combinar innovación y protección de los derechos individuales e intereses colectivos.

Lectura recomendada

A. Rouvroy (2016). *Of Data and Men»: Fundamental Rights and Liberties in a World of Big Data* [en línea]. <rm.coe.int/CoERMPublicCommonSearch Services/DisplayDCTMContent?documentId =09000016806a6020>

Desde esta perspectiva, el desarrollo de la IA –que se centra necesariamente en los datos personales cuando se refiere a aspectos individuales y sociales– debe basarse en una lectura crítica y actualizada de los principios de proporcionalidad, responsabilidad y transparencia, así como en formas adecuadas de gestión del riesgo y en formas de participación activa de las partes interesadas.

1. Los desafíos de la IA

Como ha sucedido en otras ocasiones, el cambio de paradigma introducido primero por el *big data* y luego por la IA se ha encontrado con un marco regulatorio europeo sobre procesamiento de datos que es parcialmente inadecuado para proporcionar respuestas oportunas.

Lectura recomendada

A. Mantelero (2016). «Personal data for decisional purposes in the age of analytics: from an individual to a collective dimension of data protection». *Computer Law & Sec. Review* (vol. 32, págs. 238-255).

La creciente concentración del poder de información, la oscuridad frecuente de su uso y el impacto del uso de los datos en los procesos de toma de decisiones a gran escala parecen ser abordados solo parcialmente por el legislador europeo, que, en el Reglamento (UE) 2016/679, sigue siendo principalmente fiel a un modelo de protección centrado en los derechos y en el papel de la persona interesada, reafirmando principios como los de trasparencia y minimización, que son difíciles de conciliar con la naturaleza de la IA.

Si, por un lado, el nuevo reglamento muestra estos límites, por otro, marca un punto de inflexión en el sentido del análisis y la gestión del riesgo relacionado con el tratamiento. Se desplaza el enfoque de la normativa desde la autodeterminación del interesado hacia una mayor centralidad de la responsabilidad del titular del tratamiento y de aquellos involucrados en el tratamiento de los datos.

En este sentido, la responsabilidad y el principio de *accountability*, cuyo objetivo es demostrar el cumplimiento de los principios y las obligaciones establecidos por el Reglamento, constituyen el núcleo del nuevo marco europeo de protección de datos y son un elemento útil para abordar los posibles efectos negativos del uso de la IA.

Sin embargo, la transición actual hacia este diferente modo de tratar la protección de datos, centrado en el riesgo, parece todavía incompleta. Siguen existiendo elementos del modelo anterior centrados en la dimensión individual del interesado y no ha habido ningún replanteamiento de la arquitectura reguladora, cuyos principios fundadores se remontan a la década de los noventa y, en cierta medida, a la de los setenta.

En este contexto, las disposiciones del artículo 22 del Reglamento y su papel potencial en el contexto de la IA son particularmente interesantes en relación con el debate doctrinal.[2] A este respecto, sin embargo, cabe señalar que la importancia de esta norma puede sobrestimarse, ya que el número de casos en los que se adoptan decisiones de manera exclusivamente automatizada que producen «efectos jurídicos» para la persona afectada o que pueden afectar «de

(2)Véase art. 35, Reg. (UE) 2016/679.

manera similar y significativa a su persona» sigue siendo limitado. En muchos casos, de hecho, los sistemas de IA son herramientas de análisis para apoyar las decisiones, que serán tomadas por las personas con poder de decisión.

Lecturas recomendadas

M. Veale; L. Edwards (2018). «Clarity, surprises, and further questions in the Article 29 Working Party draft guidance on automated decision-making and profiling». *Computer Law & Security Rev.* (vol. 34, núm. 2, págs. 398-404).

L. Edwards; M. Vale (2017). «Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for». *Duke Law & Technology Review* [en línea]. <papers.ssrn.com/sol3/papers.cfm?abstract_id=2972855>

L. A. Bygrave (2001). «Minding the Machine: Article 15 of the EC Data Protection Directive and Automated Profiling». *Computer Law & Security Rev.* (vol. 17, núm. 1, pág. 17).

W. Schreurs; M. Hildebrandt; E. Kindt-M. Vanfleteren (2010). «Cogitas, Ergo Sum. The Role of Data Protection Law and Non-discrimination Law in Group Profiling in the Private Sector». En: M. Hildebrandt; S. Gutwirth (eds). *Profiling the European Citizen. Cross-Disciplinary Perspective* (pág. 241). Dordrecht: Springer.

Por lo tanto, el problema más relevante se refiere al grado efectivo de libertad que caracteriza la autodeterminación de los responsables de la toma de decisiones con respecto a las sugerencias proporcionadas por los algoritmos. En este contexto, debemos preguntarnos si la presencia de un sujeto que adopta decisiones termina siendo más la razón formal de la exclusión de la aplicación del art. 22 que una garantía efectiva de autonomía de juicio con respecto a la máquina.

Lectura recomendada

Article 29 Data Protection Working Party (2017). *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679* (pág. 10).

También cabe señalar que el derecho a obtener la intervención humana no parece ser particularmente decisivo con respecto a los problemas relacionados con el uso de la IA. Por una parte, de hecho, dicha protección no va a operar cuando se persiguen fines contractuales o el tratamiento de datos se basa en el consentimiento del interesado,[3] lo que reduce la protección proporcionada por la norma. Por otra parte, la debilidad de los responsables de la toma de decisiones frente a las sugerencias ofrecidas por los algoritmos se repite en este caso y plantea dudas sobre la eficacia concreta de la «intervención humana».

[3] Véase art. 22, par. 2, a) y c), Reg. (UE) 2016/679.

En términos de análisis de riesgo, el RGPD contiene normas importantes que también podrían aplicarse para evaluar el impacto del uso de los datos en el contexto de las aplicaciones de IA. Sin embargo, el nuevo reglamento no contiene indicaciones específicas para definir un modelo de análisis de riesgos que sea capaz de tener en cuenta el impacto más amplio que tiene el uso de los datos en el contexto de la IA; un impacto que va más allá de la simple protección de los datos. Esto se ve confirmado por los modelos de DPIA, que todavía se centran principalmente en la seguridad de las informaciones y la calidad de los datos, más que en las consecuencias sobre una pluralidad de derechos y la dimensión colectiva del uso de los datos.

Lectura recomendada

A. Mantelero (2016). *Op. cit.*

Además, por su propia naturaleza, la legislación sobre los datos personales no puede proporcionar respuestas completas a las preguntas sobre el impacto social de la IA y las cuestiones éticas que surgen del uso de la IA. Por esta razón,

se han propuesto soluciones más innovadoras, como las directrices sobre *big data* adoptadas por el Consejo de Europa en 2017 y las directrices sobre la IA y el tratamiento de datos que el Consejo de Europa adoptó en enero de 2019, ambas destinadas a identificar nuevas maneras de regular el impacto de los algoritmos en la sociedad.

Lecturas recomendadas

Guidelines on the protection of individuals with regard to the processing of personal data in a world of Big Data [en línea]. <rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000016806ebe7a>

A. Mantelero (2017). «Regulating Big Data. The guidelines of the Council of Europe in the context of the European data protection framework». *Computer Law & Sec. Review* (vol. 33, págs. 584-602).

Guidelines on Artificial Intelligence and Data Protection [en línea]. <www.coe.int/en/web/data-protection/-/new-guidelines-on-artificial-intelligence-and-personal-data-protection>

Independent High-Level Expert Group on Artificial Intelligence (2019). *Ethics Guidelines for Trustworthy AI* [en línea]. <ec.europa.eu/futurium/en/ai-alliance-consultation>

Council of Europe-Committee of Experts on Internet Intermediaries (MSI-NET) (2018). *Study on the Human Rights Dimensions of Automated Data Processing Techniques (in Particular Algorithms) and Possible Regulatory Implications* [en línea]. <rm.coe.int/algorithms-and-human-rights-en-rev/16807956b5>

European Data Protection Supervisor-Ethics Advisory Group (2018). *Towards a digital ethics* [en línea]. <edps.europa.eu/sites/edp/files/publication/18-01-25_eag_report_en.pdf>

Parlamento Europeo (2017). *European Parliament resolution of 14 March 2017 on fundamental rights implications of big data: privacy, data protection, non-discrimination, security and law-enforcement (2016/2225(INI))* [en línea]. <www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P8-TA-2017-0076+0+DOC+XML+V0//EN&language=EN>

European Union Agency for Fundamental Rights (FRA) (2018). *Big Data: Discrimination in Data-Supported Decision Making* [en línea]. <fra.europa.eu/en/publication/2018/big-data-discrimination>

Las tecnologías destinadas a un uso masivo e intensivo de los datos, como la IA, representan un desafío para la aplicación de varios principios tradicionales con respecto a la protección de datos, haciéndolos más borrosos, menos claros o más difíciles de aplicar. En este sentido, varios autores han señalado la debilidad, por ejemplo, de la eficacia del consentimiento informado, como una herramienta para una verdadera autodeterminación informativa. Esto es aún más evidente frente a las formas de creación de perfiles basadas en algoritmos de IA, que son muchas veces oscuros y complejos, o en prácticas ocultas de *nudging* que socavan la noción de control de la información por parte de los interesados.

Lecturas recomendadas

Ex multis, **M. Hildebrandt** (2016). *Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology*. Edward Elgar Publishing.

S. Barocas; H. Nissenbaum (2015). «Big Data's End Run around Anonymity and Consent». En: J. Lane y otros (dir.). *Privacy, big data, and the public good: frameworks for engagement*. Cambridge: Cambridge University Press.

D. K. Citron; F. Pasquale (2014). «The Scored Society: Due Process For Automated Predictions». *Wash. L. Rev.* (vol. 89, págs. 1-33).

A. Mantelero (2014). «The future of consumer data protection in the E.U. Rethinking the "notice and consent" paradigm in the new era of predictive analytics». *Computer Law & Sec. Rev.* (vol. 30, págs. 643-660).

I. S. Rubinstein (2013). «Big Data: The End of Privacy or a New Beginning?». *International Data Privacy Law* (vol. 3, núm. 2, págs. 74-87).

También debe señalarse que la noción de autodeterminación, en el contexto de la IA, no puede considerarse limitada al mero uso de datos, sino que asume una mayor importancia en relación con una libertad de elección más general, tanto con respecto a soluciones centradas en el uso de la IA, como a la pretensión de poder utilizar una versión de dispositivos y servicios equipados con IA que pueda ser no «inteligente». Este tipo de opción que excluye la IA no se refiere solo a la dimensión individual y al uso de dispositivos/servicios específicos por parte del usuario, sino también a la libertad más amplia de una comunidad para decidir sobre el papel que la IA debe desempeñar en la configuración de la dinámica social, el comportamiento colectivo y las decisiones que afectan a grupos de individuos.

Lecturas recomendadas

Office of the Privacy Commissioner of Canada (2016). *The Internet of Things: An Introduction to Privacy Issues with a Focus on the Retail and Home Environments* [en línea]. <www.priv.gc.ca/en/opc-actions-and-decisions/research/explore-privacy-research/2016/iot_201602/#heading-0-0-2-15>

Asilomar AI Principles (2017). <futureoflife.org/ai-principles>

La doctrina ha intentado proporcionar una respuesta a estas demandas apoyando la necesidad de fortalecer el papel de la transparencia en el tratamiento de datos o, en cuanto al consentimiento, sugiriendo formas más flexibles, como el consentimiento amplio y el consentimiento dinámico. Aunque ninguna de estas soluciones puede, por sí sola, proporcionar una respuesta exhaustiva a la crisis del consentimiento en el contexto de la IA, en algunas áreas, estas soluciones –solas o combinadas– pueden fortalecer la autodeterminación individual efectiva.

Lecturas recomendadas

Ex multis, **L. Edwards y otros** (2017). «Slave to the Algorithm? Why a "Right to an Explanation" Is Probably Not the Remedy You Are Looking For». *Duke Law and Technology Review* (vol. 16, núm. 1, págs. 18-84).

A. Selbst y otros (2017). «Meaningful Information and the Right to Explanation». *International Data Privacy Law* (vol. 7, núm. 4, págs. 233-242).

S. Wachter y otros (2017). «Why a right to explanation of automated decision - making does not exist in the General Data Protection Regulation». *International Data Privacy Law* (vol. 7, núm. 2, págs. 76-99).

M. Sheehan (2011). «Can Broad Consent be Informed Consent?». *Public Health Ethics* (vol. 3, págs. 226-235).

J. Kaye y otros (2015). «Dynamic consent: a patient interface for twenty-first century research networks». *European Journal of Human Genetics* (vol. 23, núm. 2, págs. 141-146).

Otro problema importante de las aplicaciones de la IA está relacionado con los sesgos potenciales que pueden afectar a estos sistemas. Si por un lado los sistemas de IA pueden contribuir a reducir o eliminar los sesgos que pueden afectar a las personas en el proceso de toma de decisiones, por otro lado, es posible que ellos mismos estén viciados por prejuicios o elementos desviados que llevan a conclusiones erróneas en el proceso automatizado de toma de decisiones. Tanto los modelos deterministas de IA como los de aprendizaje automático

utilizan datos preexistentes como base para extraer más informaciones (análisis de *big data*) o crear y «entrenar» modelos de aprendizaje automático. De ahí la centralidad del tema del sesgo en la selección y el uso de la información.

El sesgo potencial puede afectar a aspectos diferentes: la metodología de investigación, el objeto de la investigación (este es el caso, por ejemplo, del sesgo social debido al sesgo en las series históricas de datos o a la falta de representación adecuada de algunas categorías), las fuentes de datos (sesgo en los procesos de selección de datos) o el propio comportamiento del autor de la actividad de investigación.

Lectura recomendada

M. Veale; R. Binns (2017). «Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data>» [en línea]. *Big Data & Society* (vol. 4, núm. 2). <doi.org/10.1177/2053951717743530>

La presencia de sesgo puede afectar negativamente al desarrollo y a la aplicación de algoritmos, con un impacto aún mayor en el caso de algoritmos de aprendizaje automático, donde los sesgos pueden influir tanto en el diseño de estos como en el desarrollo. De ahí la necesidad de abordar un enfoque desde el diseño para evitar «potential hidden data biases and the risk of discrimination or negative impact on the rights and fundamental freedoms of data subjects, in both the collection and analysis stages». Sin embargo, este enfoque no se limita a la simple privacidad o protección de datos desde el diseño, sino que debe tener en cuenta cada vez más la amplia variedad de derechos, libertades e intereses potencialmente afectados por el uso de soluciones centradas en la inteligencia artificial.

Lectura recomendada

A. Mantelero (2019). *Report on Artificial Intelligence and Data Protection: Challenges and Possible Remedies* (pág. 9).

Por lo tanto, ante la complejidad de los sistemas de IA y la pluralidad de posibles impactos, tanto respecto a las personas como respecto a la sociedad, se hace evidente el riesgo de dejar la tarea de la realización de estos sistemas a los desarrolladores de software. En consecuencia, no es posible dejar la evaluación y gestión de estos impactos a las decisiones de los desarrolladores ni a la visión del mundo que ellos o las compañías a las que pertenecen tienen.

Lectura recomendada

AI Now Institute (2017). *AI Now 2017 Report* (pág. 18) [en línea]. <assets.contentful.com/8wprhhvnpfc0/1A9c3ZTCZa2KEYM64Wsc2a/8636557c5fb14f2b74b2be64c3ce0c78/_AI_Now_Institute_2017_Report_.pdf>

De ahí la necesidad de un desarrollo más unánime y participado de la IA, que vea un papel activo de las diferentes categorías de partes interesadas y tenga en cuenta conocimientos diferentes.[4] De este modo se pueden cerrar las brechas cognitivas y experienciales de una formación técnica, que es adecuada para desarrollar el software, pero no proporciona medidas suficientes y adecuadas para interpretar las consecuencias sociales del uso de los algoritmos de IA.

(4)Véase art. 35, par. 9, Reg. (UE) 2016/679.

Lectura recomendada

Article 29 Data Protection Working Party (2018). *Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is «likely to result in a high risk» for the purposes of Regulation 2016/679*.

Como se mencionó, otra manera de reducir el posible sesgo de aplicación de la IA es recurrir a formas colaborativas de evaluación de riesgos centradas no solo en la seguridad y la calidad de los datos, sino también en la participación activa de grupos potencialmente afectados por las aplicaciones de IA. Asimismo, deben participar las partes interesadas capaces de contribuir a la identificación y eliminación de los sesgos y los potenciales impactos negativos del uso de datos. Este enfoque, dirigido al desarrollo responsable de soluciones basadas en IA, tiene como objetivo evitar factores de distorsión que pueden afectar a los conjuntos de datos o los algoritmos.

En un contexto necesariamente caracterizado por un nivel significativo de complejidad y una transparencia limitada de los sistemas, el recurso a formas de análisis preliminares de riesgos y planificación responsable parece, de hecho, prometer una mayor eficacia de la que se puede esperar de cualquier remedio *a posteriori*, adoptado una vez que se hayan producido efectos discriminatorios. De la misma manera, la respuesta centrada en la gestión del riesgo parece más eficaz que aquella basada en la transparencia de los algoritmos. Y ello porque esta transparencia parece difícil de lograr y, a menudo, no resulta muy efectiva porque existe una falta de interés y de conocimiento por parte de los afectados.

Lecturas recomendadas

A. D. Selbst (2017). «Disparate Impact in Big Data Policing». *Georgia Law Review* (vol. 52, núm. 1, págs. 109-195).

R. Brauneis; E. P. Goodman (2018). «Algorithmic Transparency for the Smart City». *Yale J. L. & Tech.* (vol. 20, pág. 131).

Este enfoque *ex ante* debería, por lo tanto, inducir una reflexión más profunda sobre los conjuntos de datos utilizados para la creación y el entrenamiento de los algoritmos con el fin de evitar consecuencias desfavorables derivadas, por ejemplo, del denominado sesgo histórico resultante del uso de conjuntos de datos preexistentes. Este riesgo se agrava cuando los conjuntos de datos no se crean *ad hoc* para el desarrollo de la aplicación específica, sino que consisten en bases de datos creadas para diferentes propósitos o disponibles en el mercado, lo que no permite a los desarrolladores un conocimiento preciso de los criterios de composición de la base de datos.

Sesgo histórico

Esto es el sesgo que se genera cuando una serie histórica determinada presenta una tendencia adversa respecto a algunas categorías, por lo que el desarrollo del algoritmo basado en el uso de datos históricos tiene una alta probabilidad de adquirir y mantener esta tendencia como elemento caracterizador.

Cómo reducir los riesgos

A este respecto, con el fin de reducir estos riesgos, se ha sugerido introducir formas de trazabilidad de las fuentes, desarrollo y uso de los conjuntos de datos, que se utilizan por este fin, a lo largo de su ciclo de vida.

Siempre teniendo en cuenta las soluciones enfocadas en el diseño de desarrollo de IA, es posible introducir pruebas precisas en la fase de entrenamiento de los algoritmos antes de su aplicación *in vivo*. Sin embargo, en algunos casos, el uso de una menor cantidad de datos necesario en la fase de entrenamiento no permite a los algoritmos de aprendizaje automático predecir los efectos distorsionadores que se producen cuando estos se utilizan a gran escala. En este sentido, a diferencia de lo que ocurre en el contexto del uso de datos personales con fines estadísticos, cabe señalar que los sesgos de las aplicaciones de IA pueden originarse en las mismas soluciones técnico-informáticas adoptadas, y no solo en el comportamiento o los errores de los científicos de datos.

Lecturas recomendadas

Guidelines on the protection of individuals with regard to the processing of personal data in a world of Big Data, op. cit.

M. L. Cummings y otros (2018). *Chatham House Report. Artificial Intelligence and International Affairs Disruption Anticipated*. Londres: Chatham House. The Royal Institute of International Affairs.

R. Caruana y otros (2015). «Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission». En: *Proceedings of the 21st Annual SIGKDD International Conference on Knowledge Discovery and Data Mining* (págs. 1721-1730).

Además, en el contexto de la IA, la evaluación del sesgo potencial también puede ser controvertida. La multiplicidad de las variables involucradas y la clasificación de las personas en grupos que no corresponden necesariamente a las categorías tradicionales que se utilizan en los casos de discriminación provocan que sea más difícil identificar los prejuicios potenciales.

Lecturas recomendadas

J. Donovan y otros (2018). *Algorithmic Accountability: A Primer* [en línea]. <datasociety.net/output/algorithmic-accountability-a-primer>

A. Mantelero (2016). *Op. cit.*

Finalmente, debemos considerar la línea argumentativa dirigida a disminuir el posible sesgo, y los consecuentes efectos perjudiciales, de la IA sobre la base de la falibilidad que ya afecta a la toma de decisiones por parte de seres humanos. En este sentido, se argumenta que la falibilidad humana puede ser reducida recurriendo a la IA, eliminando comportamientos negligentes o irracionales. Sin embargo, cuatro argumentos diferentes se oponen a esta conclusión.

En primer lugar, las soluciones de IA están destinadas a una aplicación en serie, de lo que se deduce que, como en el caso de la responsabilidad por productos defectuosos, las soluciones cualitativamente deficientes afectan inevitablemente a una pluralidad de personas que comparten connotaciones iguales o similares, mientras que el error humano en el proceso de toma de decisiones afecta necesariamente solo al caso específico que está decidido.

En segundo lugar, aunque hay áreas en las que los índices de error de la IA son cercanos o inferiores a los índices de error humano (por ejemplo, el reconocimiento de imágenes), en muchos otros contextos hay márgenes de error excesivos e inaceptables que hacen que sea indeseable la transición desde la falibilidad humana a la falibilidad algorítmica.

En tercer lugar, no debe pasarse por alto la dimensión sociocultural del error humano en términos de aceptabilidad social. A nivel político y social, es mucho más difícil admitir la falibilidad de los algoritmos en los que se basan precisamente las soluciones para evitar decisiones afectadas por sesgos u otras formas de perjuicios que la humana.

Lectura recomendada

Council of Europe. *Guidelines on Artificial Intelligence and Data Protection, op. cit.*

Finalmente, incluso admitiendo la posibilidad de comparar las decisiones humanas con las que puede tomar un algoritmo, esta misma comparación resulta ser metodológicamente difícil. En particular, no se puede hacer una simple comparación cuantitativa basada en el número de errores y sus consecuencias (por ejemplo, el número medio de víctimas causadas por automóviles con conductor humano en comparación con el mismo número en el caso de vehículos autónomos). De hecho, al evaluar las consecuencias de la IA y las decisiones humanas, es necesario tener en cuenta la distribución de los efectos (es decir, las personas afectadas negativamente que pertenecen a diferentes categorías, las distintas condiciones en las que se produjo el daño, la gravedad de las consecuencias, etc.). Además, este tipo de enfoque cuantitativo, que considera positivamente la solución que ofrece el menor impacto en términos de error,

parece estar en contraste con el enfoque precautorio, que requiere la adopción de políticas activas de prevención de riesgos, más que una mera reducción de daños.

2. Los límites de la transparencia

El recurso a obligaciones de transparencia a menudo se invoca como una herramienta para resolver, al menos en parte, cuestiones críticas mencionadas en los párrafos anteriores. Desde este punto de vista, la transparencia haría que los sujetos interesados fueran más conscientes, mitigando los límites que afectan a la autodeterminación individual con respecto al tratamiento de datos en el contexto de la IA. La transparencia ayudaría entonces a aclarar los propósitos de dichos tratamientos y, por último, pero no menos importante, serviría para prevenir cualquier sesgo que pudiera ser una fuente de prejuicios potenciales, especialmente en términos de discriminación.

Antes de reflexionar sobre cómo, en la práctica, estas expectativas pueden materializarse solo en parte con respecto a la adopción de políticas de transparencia, cabe destacar que la propia noción de transparencia en el contexto de los algoritmos tiene contornos a menudo indefinidos, lo que dificulta incluso invocar una eficacia operacional significativa.

La noción de transparencia puede, de hecho, tener diferentes significados. Puede interpretarse de modo que los interesados sean conscientes del hecho de que las soluciones de IA se utilizan en la interacción con ellos mismos, o puede entenderse como trasparencia del procesamiento de los datos que los conciernen. También puede implicar la descripción de la lógica de los algoritmos utilizados para este propósito o, finalmente, llegar a un acceso directo a la estructura del algoritmo y, cuando corresponda (casos de aprendizaje automático), a los conjuntos de datos utilizados para la fase de entrenamiento del algoritmo.

Lectura recomendada

R. Binns y otros (2018). «It's Reducing a Human Being to a Percentage, in Perceptions of Justice» [en línea]. *Algorithmic Decisions*, ArXiv:1801.10408 [Cs] (págs. 1-14). <doi.org/10.1145/3173574.3173951>

Examinando aquí brevemente los diferentes aspectos de la noción de transparencia en el contexto de la IA, cabe señalar que, si bien es importante tener un control público sobre los modelos automatizados de toma de decisiones, la primera perspectiva, que consiste en el simple conocimiento sobre el uso de la IA, es útil para que los interesados tomen conciencia del contexto, pero es poco eficaz a la hora de abordar los riesgos de un posible uso ilegítimo de datos y de la IA.

Lectura recomendada

D. Reisman y otros (2018). *Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability* [en línea]. <ainowinstitute.org/aiareport2018.pdf>

En el otro extremo, el acceso a la estructura del algoritmo y, cuando corresponda, a los datos utilizados para el entrenamiento sin duda será efectivo para detectar posibles sesgos o el uso ilegítimo de la información. Sin embargo, una noción tan amplia de transparencia con frecuencia entraría en conflicto con los derechos de propiedad intelectual y también podría implicar cuestiones relacionadas con la competencia y la libertad de empresa.

Además, si estos obstáculos no existieran o pudieran superarse (por ejemplo, a través de un acceso restringido a autoridades de control o judiciales, como suele ser el caso), en cualquier caso se daría el hecho de que la complejidad de los modelos adoptados constituye a menudo un desafío considerable para las habilidades cognitivas humanas, lo que implica que los procesos implementados por la máquina sean inexplicables.

Finalmente, cabe señalar que hay casos en los que el uso de la transparencia puede suponer otras dificultades. Se puede revelar como un obstáculo para lograr los objetivos propios de la Administración pública en el ejercicio de sus competencias. Por ejemplo, en las tareas relacionadas con la prevención del delito (piénsese en cómo la transparencia de los algoritmos de los sistemas de policía predictivos podría socavar su eficacia). En otros casos, el acceso al algoritmo entra en conflicto con las obligaciones de seguridad del responsable del tratamiento respecto a los datos personales de los interesados distintos de aquellos que solicitan el acceso, exponiendo dicha información a posibles riesgos en términos de seguridad informática.

Por estas razones, la solución intermedia entre las dos analizadas anteriormente, que consiste en el conocimiento de la lógica del algoritmo utilizado, parece ser la más practicable. Aun así, la transparencia de los algoritmos se puede interpretar más o menos estrictamente. Proporcionar información sobre los datos de entrada y los resultados esperados, describir las variables utilizadas y el peso que se les atribuye, o explicar el modelo analítico utilizado, son los diversos modos que puede asumir la noción de transparencia con respecto a la lógica de los algoritmos de IA.

El aprendizaje supervisado y sin supervisión

El aprendizaje automático generalmente comienza con información seleccionada que contiene *patterns* o similitudes, después el software de aprendizaje automático identifica los *patterns* que se encuentran en esta información y, finalmente, genera un modelo que puede reconocer los *patterns* que surgen en datos nuevos.

Hay varios modos de aprendizaje, que pueden utilizar datos con etiqueta o no. Los datos etiquetados son datos que incluyen informaciones sobre el contenido de cada dato (en el caso de imágenes, las etiquetas pueden ser, por ejemplo, la raza, el género o el sexo de un perro que está en la imagen). Estos datos etiquetados se utilizan en el aprendizaje supervisado, donde las etiquetas hacen posible la supervisión mediante la cual el algoritmo se entrena con un «histórico» de datos y aprende a asignar las etiquetas a los nuevos datos que se le facilitan. En estos casos se entrena una máquina de IA con un resultado esperado y si la salida generada por la máquina es incorrecta, el proceso se repite sobre el mismo conjunto de datos hasta que esta no tenga más errores. Es evidente que las características de los datos etiquetados son decisivas para una categorización correcta y para producir resultados correctos.

En el caso del aprendizaje sin supervisión se utilizan datos que no se han etiquetado previamente, y el sistema tiene que agrupar los datos que sean similares. En este caso se utiliza un conjunto de datos en el que el algoritmo tiene que buscar agrupamientos basados en similitudes, pero nada garantiza que estas tengan algún significado o utilidad.

Independientemente de los algoritmos o métodos utilizados para el aprendizaje automático, el resultado será un modelo que puede ser alimentado con nuevos datos para producir el tipo de resultado deseado. Esto puede ser, por ejemplo, una clasificación o un grado de probabilidad.

Sin embargo, incluso en este caso, los modelos de análisis complejos que se centran en la IA, como los algoritmos de aprendizaje profundo, constituyen un desafío para esta noción de transparencia, entendida como una explicación de la lógica de los algoritmos, dado que los sistemas no determinísticos dificultan el suministro de información detallada sobre la lógica existente detrás del tratamiento de los datos.

Además, la naturaleza dinámica de muchos algoritmos contrasta con la propia noción de transparencia, que por su propia naturaleza es estática, una «fotografía» de una situación en un momento dado. De hecho, la actualización y modificación continua de los algoritmos de IA provocan que sea difícil su evaluación sobre la base de un análisis que requiere tiempo para su realización y cuyos resultados, por lo tanto, se pueden lograr cuando el algoritmo ya ha asumido una estructura diferente.

Esta última observación también resalta cómo la mera transparencia no es suficiente para aclarar el impacto de los algoritmos y cómo, en cambio, el acceso a los algoritmos de IA debe ser seguido por un análisis preciso, y a menudo complejo, para verificar los riesgos de posibles prejuicios. Por lo tanto, es necesario disponer de recursos adecuados en términos de tecnologías, tiempo y competencias, que no permiten obtener los beneficios de la transparencia de manera inmediata y en todas las situaciones y por parte de todos los interesados.

Ciertamente, se puede alegar que el periodismo de investigación, las asociaciones que protegen los intereses de categorías específicas (por ejemplo, minorías o trabajadores), así como el papel de las autoridades de control, podrían tener un papel más relevante y reducir estos problemas. Si bien esto es cierto, también lo es que el uso de soluciones de IA es cada vez menos costoso y, en consecuencia, más generalizado y capilar. De ahí la dificultad de poder imaginar que estas diferentes entidades y formas de vigilancia pueden supervisar y controlar de manera eficaz todas las aplicaciones de IA.

Estos factores limitan las soluciones operativas centradas en los mecanismos de auditoría o en la intervención humana en el proceso de la toma de decisiones. En respuesta a estos problemas, la investigación más reciente tiene como objetivo desarrollar algoritmos capaces de supervisar otros algoritmos, permitiendo la detección automática de los sesgos potenciales. En este sentido, aunque la intención de desarrollar tecnologías responsables es apreciable, existen dudas en cuanto al hecho de que la introducción de una especie de supervisor algorítmico se traduzca en una reducción de la complejidad de la gobernanza de los datos, especialmente asumiendo la perspectiva de las partes interesadas.

Lectura recomendada

M. Veale; R. Binns (2017). *Op. cit.*

Sin embargo, los asuntos críticos resaltados aquí con respecto al papel de la transparencia no deben entenderse como un intento de debilitar los argumentos a favor de una mayor transparencia o de disminuir este principio desde una

perspectiva de salvaguardia de la autodeterminación de los interesados, sino más bien como una exhortación a comprender plenamente la complejidad de este principio, incluso a nivel operativo.

En todo caso, debe tenerse en cuenta cómo la transparencia es solo una parte de la solución a los desafíos de la IA y tiene varias limitaciones que deben abordarse por completo. Asimismo, el algoritmo, cuya transparencia se discute habitualmente, es solo uno de los componentes de una aplicación de IA, mientras que otros, igualmente relevantes, son los conjuntos de datos utilizados para crear y entrenar el algoritmo y para los datos utilizados en la posterior fase de procesamiento y análisis. Por lo tanto, independientemente de la transparencia del algoritmo, estos conjuntos de datos pueden producir resultados distorsionados si están afectados por sesgos.

Lectura recomendada

M. Ananny; K. Crawford (2016). «Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability» [en línea]. *New Media & Society*. <doi.org/10.1177/1461444816676645>

Finalmente, tanto con respecto a los algoritmos como a los datos, surge el problema del uso descontextualizado de estas herramientas y recursos. Por ello, no es raro que las formas intensivas de análisis de datos se centren en información descontextualizada, sin referencias al contexto de origen útiles tanto para comprender plenamente los resultados del tratamiento como para comprender sus utilidades en términos de la aplicación concreta. De manera análoga, los modelos algorítmicos originalmente creados para el análisis de un problema dado se pueden utilizar posteriormente para distintos propósitos en diferentes contextos. O bien, modelos elaborados sobre la base de datos históricos de una población se pueden usar con respecto a una población diferente, sin que se haya verificado cuidadosamente la efectividad inalterada de los propios modelos en el caso de uso descontextualizado.

Información descontextualizada

Esto puede ser el caso, por ejemplo, de un sistema algorítmico que evalúa el rendimiento de los estudiantes de diferentes escuelas, sin tener en cuenta las variables socioeconómicas de las familias de origen.

Lectura recomendada

Donovan y otros (2018). *Ob. cit.* (pág. 7).

3. El papel del análisis de riesgos

Dadas las limitaciones que afectan tanto a la autodeterminación de las personas en relación con el tratamiento de datos como a la transparencia de los procesos algorítmicos, las normas de protección de datos consideran cada vez más el papel de la evaluación de riesgos como un remedio preventivo para evitar consecuencias perjudiciales.

La evaluación de riesgos, en el contexto del procesamiento de datos personales, no solo es una herramienta para prevenir posibles perjuicios a los derechos y libertades de los interesados, sino que también desempeña un papel importante en la dinámica centrada en la confianza de los usuarios que siempre ha caracterizado el desarrollo de tecnologías. En este sentido, la presencia de un entorno de IA «seguro», desde el punto de vista técnico y jurídico, puede mejorar la confianza de los usuarios de dichas aplicaciones y aumentar su disposición a utilizarlas.

Lectura recomendada

A. Mantelero (en prensa). «Comment to Articles 35 and 36» [en línea]. En: M. Cole; F. Boehm (dirs.). *GDPR Commentary* Edward Elgar Publishing. <papers.ssrn.com/sol3/papers.cfm?abstract_id=3362747>

En este sentido, las preferencias de los usuarios pueden basar su confianza más correctamente en un análisis de los riesgos efectivos y de las medidas que se han adoptado para hacer frente a estos riesgos que en campañas de marketing o reputación de marca. Por esta razón, el uso de algoritmos en el contexto del tratamiento de datos personales y el uso creciente de tecnologías que hacen un uso intensivo de datos han llevado a una mayor atención a los posibles efectos adversos del procesamiento de datos.

Lecturas recomendadas

A. Mantelero (2016). «Children online and the future EU data protection framework. Empirical evidences and legal analysis». *International Jour. Tech. Policy & Law* (vol. 2, núm. 2/3/4, págs. 169-181).

European Data Protection Supervisor - Ethics Advisory Group, *op. cit.*

Grupos de expertos y académicos han ido más allá de la esfera tradicional de protección de datos para considerar el impacto del uso de datos en los derechos fundamentales y los valores sociales y éticos, tanto a nivel individual como colectivo. Si, por un lado, esta extensión de la evaluación de riesgos se debe a los efectos de las aplicaciones centradas en la IA, que va mucho más allá del entorno del tratamiento de datos personales, por otro lado, la evaluación del respeto por los derechos humanos y los valores éticos y sociales hace más difícil llevar a cabo la evaluación de impacto del uso de datos.

Lecturas recomendadas

Access Now (2018). *The Toronto Declaration: Protecting the rights to equality and non-discrimination in machine learning systems* (vol. 2, núm. 2/3/4, págs. 169-181) [en línea]. <www.accessnow.org/cms/assets/uploads/2018/08/The-Toronto-Declaration_ENG_08-2018.pdf>

A. Mantelero (2017). «From group privacy to collective privacy: towards a new dimension of privacy and data protection in the big data era». En: L. Taylor y otros (dir.). *Group Privacy New Challenges of Data Technologies*. Cham: Springer International Publishing.

Además, mientras que, en el campo de la seguridad y la gestión de la información, los criterios y valores de referencia (por ejemplo, la integridad de los datos) se centran en la tecnología y, por lo tanto, pueden generalizarse en varios contextos, la situación es diferente cuando se toman como referencia los valores éticos y sociales. Estos últimos son, de hecho, necesariamente específicos y contextuales y difieren de una comunidad a otra. De ello se deduce que será más difícil identificar un modelo de valor de referencia para una evaluación del riesgo que vaya más allá de la protección de datos personales y también incluya otros derechos fundamentales, como el derecho a la no discriminación, así como cuando se tengan en cuenta las consecuencias éticas y sociales del uso de la IA.

Lectura recomendada

D. Wright (2011). «A framework for the ethical impact assessment of information technology». *Ethics Inf. Technol* (núm. 13, pág. 201).

La naturaleza contextual de la evaluación inherente al respeto de los valores éticos y sociales conduce, por lo tanto, a reevaluar soluciones operativas que permitan una mayor granularidad del análisis y una mayor cercanía a las comunidades de referencia. De ahí el creciente interés en el papel de los comités de expertos o comités de ética. Estos comités, que en algunos casos ya existen en la práctica, deberían determinar los valores específicos que deben protegerse en relación con el uso específico de los datos, proporcionando indicaciones más detalladas y contextuales al responsable del tratamiento de datos para llevar a cabo una evaluación de riesgos que sea exhaustiva.

Lectura recomendada

Council of Europe. *Guidelines on the protection of individuals with regard to the processing of personal data in a world of Big Data, op. cit.*

A este respecto, debe tenerse en cuenta que el mayor esfuerzo resultante de una evaluación de riesgos más amplia no solo se debe a la naturaleza de los derechos y libertades potencialmente afectados por la aplicación de la IA y por sus importantes consecuencias sociales, sino que también representa una oportunidad para el desarrollo de innovación responsable, así como una ventaja competitiva potencial para las empresas.

De hecho, incrementar la confianza de los usuarios en los productos y servicios de IA ofrece a las empresas la oportunidad de responder mejor a las crecientes preocupaciones de los consumidores sobre el uso de sus datos y de soluciones de IA en general. Del mismo modo, una mayor consideración de las consecuencias de la adopción de sistemas de IA por parte de los organismos públicos solo puede tener un impacto positivo en términos de la confianza de los ciudadanos en la acción de la Administración pública y en la prevención de decisiones afectadas por sesgos en la toma de decisiones.

Lectura recomendada

IEEE (2019). *Ethically Aligned Design. A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems First Edition Overview* [en línea]. <ethicsinaction.ieee.org>

Certificaciones, códigos de conducta y estándares

En esta perspectiva, cabe destacar que las certificaciones, los códigos de conducta y los estándares también pueden desempeñar un papel importante. Estas diferentes herramientas ayudan en la rendición de cuentas y proporcionan orientación sobre la integridad de los datos y del sistema.

Finalmente, cabe destacar que la evaluación de las consecuencias del uso de los sistemas de IA, dado el impacto que suelen tener en una pluralidad de personas y en comunidades enteras, no debe llevarse a cabo sin la participación

[5]Véase art. 35, par. 9, Reg. 2016/679.

activa de estas comunidades. Este es el tema del enfoque participativo en la evaluación de riesgos, ya considerado por el legislador europeo con respecto al procesamiento de datos personales,[5] pero aún más relevante si queremos crear un modelo más amplio de análisis de riesgos, que incluya las consecuencias sociales.

Lecturas recomendadas

Article 29 Data Protection Working Party, *op. cit.*

Un enfoque participativo también puede ser útil para una mejor comprensión de los diversos intereses y valores en juego, con respecto a aplicaciones específicas de IA. En este sentido, la participación de los interesados también representa un objetivo de la evaluación de impacto, ya que contribuye a reducir el riesgo de subrepresentación de algunos grupos o categorías de personas y también es potencialmente capaz de resaltar aspectos críticos que se subestiman o ignoran en la evaluación abstracta realizada por el responsable del tratamiento de datos.

Lectura recomendada

United Nations Office of the High Commissioner for Human Rights (2016). *Frequently asked questions on a human rights-based approach to development cooperation*. Nueva York / Génova: United Nations.

Obviamente, la participación de las partes interesadas no debe verse como una forma en que quienes toman las decisiones (los responsables del tratamiento de datos en este caso) pueden eludir sus responsabilidades como sujetos que tienen que gestionar todo el proceso. Más bien, el propósito de una evaluación participativa de los efectos de largo alcance de la toma de decisiones algorítmicas es inducir a los responsables del tratamiento a adoptar soluciones de codiseño para el desarrollo de aplicaciones de IA, involucrando activamente a los grupos potencialmente afectados.

Resumen

Con estos temas de reflexión que se han tratado en las páginas anteriores, se quería llevar a cabo un análisis de los problemas críticos que plantea el recurso creciente a la IA con respecto a la protección de datos personales y, en general, con respecto a los derechos y libertades de las personas, así como a las posibles consecuencias sociales.

Se trata de un tema de investigación que es en gran medida nuevo y abierto, que se está alejando de las formas de protección y medidas diseñadas exclusivamente para la protección de datos personales, que parecen cada vez más limitadas para abordar de manera adecuada los aspectos más críticos de la IA y de las consecuencias que conlleva la sociedad algorítmica. Prueba de este esfuerzo para mirar más allá son las diversas iniciativas desarrolladas a nivel internacional por el Consejo de Europa, el EDPS, el Parlamento Europeo, la Comisión Europea y muchas otras entidades, con el fin de dirigir nuestra mirada hacia un escenario más amplio de criterios y valores sobre los que construir un futuro marco regulatorio. Un marco que también incluye las consecuencias sociales del uso de la IA y que evalúa en un sentido amplio, sin limitarse solo al derecho a los datos personales, el impacto de las aplicaciones de esta tecnología en los derechos y libertades de las personas.

Aunque todavía queda mucho camino por recorrer antes de que se puedan establecer modelos comunes y consolidados, el debate científico sobre estos temas y las iniciativas mencionadas antes ya han marcado la dirección principal.

Bibliografía

Access Now (2018). *The Toronto Declaration: Protecting the rights to equality and non-discrimination in machine learning systems* [en línea]. <https://www.accessnow.org/cms/assets/uploads/2018/08/The-Toronto-Declaration_ENG_08-2018.pdf>

ACM (2018). *ACM Code of Ethics and Professional Conduct* [en línea]. <https://www.acm.org/code-of-ethics>

AI Now Institute (2016). *The AI Now Report. The Social and Economic Implications of Artificial Intelligence Technologies in the Near-Term* [en línea]. <https://ainowinstitute.org/AI_Now_2016_Report.pdf>

AI Now Institute (2017). *AI Now 2017 Report* [en línea]. <https://assets.contentful.com/8wprhhvnpfc0/1A9c3ZTCZa2KEYM64Wsc2a/8636557c5fb14f2b74b2be64c3ce0c78/_AI_Now_Institute_2017_Report_.pdf>

AI Now Institute (2018). *Litigating Algorithms: Challenging Government Use of Algorithmic Decision Systems* [en línea]. <https://ainowinstitute.org/litigatingalgorithms.pdf>

Ananny, M.; Crawford, K. (2016). «Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability» [en línea]. *New Media & Society*. doi: <https://doi.org/10.1177/1461444816676645>

Artificial Intelligence Index. Annual Report 2017 [en línea]. <http://aiindex.org/2017-report.pdf>

Asilomar AI Principles 2017 [en línea]. <https://futureoflife.org/ai-principles/>

Axon AI and Policing Technology Ethics Board [en línea]. <https://www.axon.com/axon-ai-and-policing-technology-ethics>

Barocas, S.; Nissenbaum, H. (2015). «Big Data's End Run around Anonymity and Consent». En: L. Lane; V. Stodden; S. Bender; H. Nissenbaum (eds). *Privacy, big data, and the public good: frameworks for engagement*. Cambridge: Cambridge University Press.

Barocas, S.; Selbstr, A. D. (2016). «Big Data's Disparate Impact». *California Law Review* (vol. 104, núm. 3, págs. 671-732).

Barse, E. L.; Kvarnstrom, H.; Jonsson, E. (2003). «Synthesizing Test Data for Fraud Detection Systems». En: *19th Annual Computer Security Applications Conference. Proceedings* (págs. 384-394). doi: <https://doi.org/10.1109/CSAC.2003.1254343>

Binns, R. y otros (2018). «It's Reducing a Human Being to a Percentage» [en línea]. *Perceptions of Justice in Algorithmic Decisions*. ArXiv:1801.10408 [Cs], (págs. 1-14). doi: <https://doi.org/10.1145/3173574.3173951>

Bostrom, N. (2016). *Superintelligence paths, dangers, strategies*. Oxford: Oxford University Press.

Boyd, D.; Crawford, K. (2012). «Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon». *Information, Communication, & Society* (vol. 15, núm. 5, págs. 662-679).

Brauneis, R.; Goodman, E. P. (2018). «Algorithmic Transparency for the Smart City». *Yale J. L. & Tech* (vol. 20, págs. 103-176).

Bray, P. y otros (2015). *International differences in ethical standards and in the interpretation of legal frameworks SATORI Deliverable D3.2* [en línea]. <http://satoriproject.eu/work_packages/legal-aspects-and-impacts-of-globalization/>

Brundage, M. y otros (2018). «The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation» [en línea]. <https://maliciousaireport.com/>

Burrell, J. (2016). «How the machine "thinks": Understanding opacity in machine learning algorithms». *Big Data & Society* (vol. 3, núm. 1). doi: <https://doi.org/10.1177/2053951715622512>

Burt, A.; Leong, B.; Shirrell, S. (2018). «Beyond Explainability: A Practical Guide to Managing Risk in Machine Learning Models». *Future of Privacy Forum*.

Bygrave L. A. (2001). «Minding the Machine: Article 15 of the EC Data Protection Directive and Automated Profiling». *Computer Law & Security Rev.* (vol. 17, núm. 1).

Calo, R. (2013). «Consumer Subject Review Boards: A Thought Experiment» [en línea]. *Standford Law Review* (vol. 66). <http://www.stanfordlawreview.org/online/privacy-and-big-data/consumer-subject-review-boards>

Caruana, R.; Lou, Y.; Gehrke, J.; Koch, P.; Sturm, M.; Elhadad, N. (2015). «Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission». En: *Proceedings of the 21st Annual SIGKDD International Conference on Knowledge Discovery and Data Mining* (págs. 1721-1730).

Citron, D. K.; Pasquale, F. (2014). «The Scored Society: Due Process For Automated Predictions». *Washington Law Review* (vol. 89, págs. 1-33).

CNIL - LINC (2017). *La Plateforme d'une Ville Les Données Personnelles Au Coeur de La Fabrique de La Smart City* [en línea]. <https://www.cnil.fr/sites/default/files/atoms/files/cnil_cahiers_ip5.pdf>

CNIL (2017). *How Can Humans Keep the Upper Hand? The Ethical Matters Raised by Algorithms and Artificial Intelligence* [en línea]. Report on the Public Debate Led by the French Data Protection Authority (CNIL) as Part of the Ethical Discussion Assignment Set by the Digital Republic Bill (pág. 14). <https://www.cnil.fr/sites/default/files/atoms/files/cnil_rapport_ai_gb_web.pdf>

Conseil National du Numérique (2015). *Ambition numérique: pour une politique francaise et europeéenne de la transition numérique* [en línea]. <https://cnnumerique.fr/nos-travaux/ambition-numerique>

Council of Europe (2017). *Guidelines on the protection of individuals with regard to the processing of personal data in a world of Big Data* [en línea]. <https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000016806ebe7a>

Council of Europe-Committee of experts on internet intermediaries (MSI-NET) (2018). *Study on the Human Rights Dimensions of Automated Data Processing Techniques (in Particular Algorithms) and Possible Regulatory Implications* [en línea]. <https://rm.coe.int/algorithms-and-human-rights-en-rev/16807956b5>

Cummings, M.; Roff, L.; Heather, M.; Cukier, K.; Parakilas, J.; Bryce, H. (2018). *Chatham House Report. Artificial Intelligence and International Affairs Disruption Anticipated* [en línea]. Londres: Chatham House. The Royal Institute of International Affairs. <https://www.chathamhouse.org/sites/default/files/publications/research/2018-06-14-artificial-intelligence-international-affairs-cummings-roff-cukier-parakilas-bryce.pdf>

Diakopoulos, N. (2013). *Algorithmic Accountability Reporting: on the Investigation of Black Boxes (Tow Center for Digital Journalism).*

DNA Web Team (2018, abril). «Google drafting ethical guidelines to guide use of tech after employees protest defence project» [en línea]. *DNA India*. <http://www.dnaindia.com/technology/report-google-drafting-ethical-guidelines-to-guide-use-of-tech-after-employees-protest-defence-project-2605149>

Donovan, J.; Matthews, J.; Caplan, R.; Hanson, L. (2018, abril). *Algorithmic Accountability: A Primer* [en línea]. <https://datasociety.net/output/algorithmic-accountability-a-primer/>

Doshi-Velez, F. y otros (2017). *Accountability of AI Under the Law: The Role of Explanation* [en línea]. <https://cyber.harvard.edu/publications/2017/11/AIExplanation>

Edwards, L.; Vale, M. (2017). «Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For». *Duke Law and Technology Review* (vol. 16, núm. 1, págs. 18-84).

European Commission (2018). *The European Artificial Intelligence Landscape* [en línea]. <https://ec.europa.eu/digital-single-market/en/news/european-artificial-intelligence-landscape>

European Commission - European Group on, Ethics in Science and, & New Technologies (2018). *Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems* [en línea]. <https://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf>

European Data Protection Supervisor (2016). *Opinion 8/2016. EDPS Opinion on coherent enforcement of fundamental rights in the age of big data*.

European Data Protection Supervisor - Ethics Advisory Group (2018). *Towards a digital ethics* [en línea]. <https://edps.europa.eu/sites/edp/files/publication/18-01-25_eag_report_en.pdf>

European Economic and Social Committee (2017). *The Ethics of Big Data: Balancing Economic Benefits and Ethical Questions of Big Data in the EU Policy Context* [en línea]. <https://www.eesc.europa.eu/en/our-work/publications-other-work/publications/ethics-big-data>

European Parliament (2017). *European Parliament resolution of 14 March 2017 on fundamental rights implications of big data: privacy, data protection, non-discrimination, security and law-enforcement (2016/2225(INI))* [en línea]. <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P8-TA-2017-0076+0+DOC+XML+V0//EN&language=EN>

European Union Agency for Fundamental Rights (FRA) (2018). *#BigData: Discrimination in Data-Supported Decision Making* [en línea]. <http://fra.europa.eu/en/publication/2018/big-data-discrimination>

Executive Office of the President, and National Science and Technology Council - Committee on Technology (2016). *Preparing for the Future of Artificial Intelligence* [en línea]. Washington D.C. <https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf>

Federal Ministry of Transport and Digital Infrastructure (2017). *Ethics Commission Automated and Connected Driving* [en línea]. <https://www.bmvi.de/SharedDocs/EN/publications/report-ethics-commission.pdf?__blob=publicationFile>

Gama, J. y otros (2013). «A survey on concept drift adaptation» [en línea]. *ACM Computing Surveys* (vol. 1, núm. 1). <http://www.win.tue.nl/~mpechen/publications/pubs/Gama_ACMCS_AdaptationCD_accepted.pdf>

Goodman, B.; Flaxman, S. (2016). «EU Regulations on Algorithmic Decision-Making and a "right to Explanation"» [en línea]. arXiv:1606.08813 [cs, stat]. <http://arxiv.org/abs/1606.08813>

Hildebrandt, M. (2016). *Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology*. Cheltenham: Edward Elgar Publishing.

IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems (2016). *Ethically Aligned Design: A Vision For Prioritizing Wellbeing With Artificial Intelligence And Autonomous Systems* (Version 1) [en línea]. IEEE. <https://standards.ieee.org/industry-connections/ec/autonomous-systems.html>

Information Commissioner's Office (2017). *Big Data, Artificial Intelligence, Machine Learning and Data Protection* [en línea]. <https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf>

ITU (2017). *AI for Good Global Summit Report 2017* [en línea]. <https://www.itu.int/en/ITU-T/AI/Documents/Report/AI_for_Good_Global_Summit_Report_2017.pdf>

Kaye, J. y otros (2015). «Dynamic consent: a patient interface for twenty-first century research networks» [en línea]. *European Journal of Human Genetics* (vol. 23, núm. 2, pág. 141). <https://www.nature.com/articles/ejhg201471>

Kurzweil, R. (2016). *The singularity is near: when humans transcend biology*. Londres: Duckworth.

Linnet, T.; Floridi, L.; van der Sloot, B. (eds). (2017). *Group Privacy: New Challenges of Data Technologies*. Nueva York: Springer International Publishing.

Lipton, Z. C. (2018). «The Mythos of Model Interpretability. In Machine Learning, the Concept of Interpretability Is Both Important and Slippery» [en línea]. *ACMQueue* (vol. 16, núm. 3). <https://queue.acm.org/detail.cfm?id=3241340>

Lomas, N. (2017). «DeepMind now has an AI ethics research unit. We have a few questions for it...» [en línea]. *TechCrunch*. <http://social.techcrunch.com/2017/10/04/deepmind-now-has-an-ai-ethics-research-unit-we-have-a-few-questions-for-it/>

Lycett, M. (2013). «Datafication: making sense of (big) data in a complex world». *European Journal of Information Systems* (vol. 22, núm. 4, págs. 381-386).

Mantelero, A. (2014). «The future of consumer data protection in the E.U. Rethinking the "notice and consent" paradigm in the new era of predictive analytics». *Computer Law and Security Review* (vol. 30, núm. 6, págs. 643-660).

Mantelero, A. (2017). «Regulating Big Data. The guidelines of the Council of Europe in the Context of the European Data Protection Framework». *Computer Law & Sec. Rev.* (vol. 33, núm. 5, págs. 584-602).

Mantelero, A. (2018). «AI and Big Data: A blueprint for a human rights, social and ethical impact assessment» [en línea]. *Assessment. Computer Law & Security Review*. doi: <https://doi.org/10.1016/j.clsr.2018.05.017>

Mayer-Schönberger, V.; Cukier, K. (2013). *Big Data. A Revolution That Will Transform How We Live, Work and Think*. Londres: John Murray.

McCulloch, W. S.; Pitts, W. H. (1943). «A Logical Calculus of the Ideas Immanent in Nervous Activity». *Bulletin of Mathematical Biophysics* (vol. 5, págs. 115-133).

Office of the Privacy Commissioner of Canada (2016). *The Internet of Things: An Introduction to Privacy Issues with a Focus on the Retail and Home Environments* [en línea]. <https://www.priv.gc.ca/en/opc-actions-and-decisions/research/explore-privacy-research/2016/iot_201602/#heading-0-0-2-15>

O'Neil, C. (2017). *Weapons of math destruction*. Londres: Penguin Books.

Palm, E.; Hansson, S. O. (2006). «The case for ethical technology assessment (eTA)». *Technological Forecasting & Social Change* (vol. 73, núm. 5, págs. 543, 550-551).

Polonetsky, J.; Tene, O.; Jerome, J. (2015). «Beyond the Common Rule: Ethical Structures for Data Research in Non-Academic Settings». *Colorado Technology Law Journal* (vol. 13, págs. 333-367).

Raso, F. y otros (2018). *Artificial Intelligence & Human Rights: Opportunities & Risks* [en línea]. <https://cyber.harvard.edu/sites/default/files/2018-09/2018-09_AIHumanRightsSmall.pdf?subscribe=Download+the+Report>

Reisman, D.; Schultz, J.; Crawford, K.; Whittaker, M. (2018). *Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability* [en línea]. <https://ainowinstitute.org/aiareport2018.pdf>

Rossi, F. (2016). *Artificial Intelligence: Potential Benefits d Ethical Considerations' (European Parliament: Policy Department C: Citizens' Rights and Constitutional Affairs 2016)* [en línea]. Briefing PE 571.380. <http://www.europarl.europa.eu/RegData/etudes/BRIE/2016/571380/IPOL_BRI(2016)571380_EN.pdf>

Rouvroy, A. (2016). *«Of Data and Men»: Fundamental Rights and Liberties in a World of Big Data* [en línea]. Estrasburg: Consejo de Europa. <https://pure.unamur.be/ws/portalfiles/portal/13278394/Report_Big_Data.pdf>

Rubinstein, I. S. (2013). «Big Data: The End of Privacy or a New Beginning?». *International Data Privacy Law* (vol. 3, núm. 2, págs. 74-87).

Schreurs W.; Hildebrandt M.; Kindt E.; Vanfleteren M. (2010). «Cogitas, Ergo Sum. The Role of Data Protection Law and Non-discrimination Law in Group Profiling in the Private Sector». En: M. Hildebrandt-S; Gutwirth (eds). *Profiling the European Citizen. Cross-Disciplinary Perspective* (pág. 241). Dordrecht: Springer.

Selbst, A. D. (2017). «Disparate Impact in Big Data Policing». *Georgia Law Review* (vol. 52, núm. 1, págs. 109-195).

Selbst, A. D.; Powles, J. (2017). «Meaningful Information and the Right to Explanation». *International Data Privacy Law* (vol. 7, núm. 4, págs. 233-242).

Sheehan, M. (2011). «Can Broad Consent be Informed Consent?». *Public Health Ethics* (vol. 3, págs. 226-235).

Spiekermann, S. (2016). *Ethical IT Innovation. A Value-Based System Design Approach*. Boca Raton: CRC Press.

Szegedy, C.; Zaremba, W.; Sutskever, I.; Bruna, J.; Erhan, D.; Goodfellow, I.; Fergus, R. (2013). *Intriguing properties of neural networks* [en línea]. <https://arxiv.org/abs/1312.6199>

Tene, O.; Polonetsky, J. (2012). «Privacy in the Age of Big Data. A Time for Big Decisions». *Stanford Law Review Online* (vol. 64 págs. 63-69).

The Danish Institute for Human Rights (2016). *Human rights impact assessment guidance and toolbox* [en línea]. <https://www.humanrights.dk/business/tools/human-rights-impact-assessment-guidance-and-toolbox>

The IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems (2016). *Ethically Aligned Design: A Vision For Prioritizing Wellbeing With Artificial Intelligence And Autonomous Systems* (Version 1) [en línea]. <http://standards.ieee.org/develop/indconn/ec/autonomous_systems.html>

The Norwegian Data Protection Authority (2018). *Artificial Intelligence and Privacy Report* [en línea]. <https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf>

Turing, A. M. (1950). «Computing Machinery and Intelligence». *Mind* (vol. 49, págs. 433-460).

UK Department for Digital, Culture, Media & Sport. *Data Ethics Framework - GOV.UK* [en línea]. <https://www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework>

United Nations Office of the High Commissioner for Human Rights (2006). *Frequently asked questions on a human rights-based approach to development cooperation*. Nueva York / Génova: United Nations.

United Nations (2011). *Guiding Principles on Business and Human Rights: Implementing the United Nations «Protect, Respect and Remedy» Framework*. United Nations Human Rights Council (UN Doc. HR/PUB/11/04).

Veale M.; Binns R. (2017). «Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data» [en línea]. *Big Data & Society* (vol. 4, núm. 2, 2053951717743530). doi: <https://doi.org/10.1177/2053951717743530>

Veale M.; Edwards L. (2018). «Clarity, surprises, and further questions in the Article 29 Working Party draft guidance on automated decision-making and profiling». *Computer Law & Security Review* (vol. 34, núm. 2, págs. 398-404).

Veale, M.; Binns, R.; Edwards, L. (2018). «Algorithms That Remember: Model Inversion Attacks and Data Protection Law» [en línea]. *Philosophical Transactions of the Royal Society*. doi: <https://doi.org/10.1098/rsta.2018.0083>

Villani, C. (2018). *For a Meaningful Artificial Intelligence towards a French and European Strategy* [en línea]. <https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf>

Wachter, S.; Mittelstadt, B.; Loridi, L. (2017). «Why a right to explanation of automated decision - making does not exist in the General Data Protection Regulation» [en línea]. *International Data Privacy Law* (vol. 7, núm. 2, págs. 76-99). <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2903469>

Walker, S. M. (2009). *The Future of Human Rights Impact Assessments of Trade Agreements* [en línea]. Utrecht: G. J. Wiarda Institute for Legal Research. <https://dspace.library.uu.nl/bitstream/handle/1874/36620/walker.pdf?sequence=2>

White House (2015). *Consumer Privacy Bill of Rights Act* [en línea]. §103(c). Administration Discussion Draft 2015. <https://obamawhitehouse.archives.gov/sites/default/files/omb/legislative/letters/cpbr-act-of-2015-discussion-draft.pdf>

World Economic Forum (2018). *How to Prevent Discriminatory Outcomes in Machine Learning* [en línea]. <http://www3.weforum.org/docs/WEF_40065_White_Paper_How_to_Prevent_Discriminatory_Outcomes_in_Machine_Learning.pdf>

Wright, D. (2011). «A framework for the ethical impact assessment of information technology». *Ethics and Information Technology* (vol. 13, págs. 199, 201-202).

Wright, D.; De Hert, P. (ed). (2012). *Privacy Impact Assessment*. Dordrecht: Springer.

Wright, D.; Mordini, E. (2012). «Privacy and Ethical Impact Assessment». En: D. Wright; P. De Hert (eds). *Privacy Impact Assessment* (págs. 397-418). Dordrecht: Springer.