

LA GESTIÓN ÉTICA DE LOS DATOS

Por qué importa y cómo hacer un uso
justo de los datos en un mundo digital

César Buenadicha
Gemma Galdon Clavell
María Paz Hermosilla
Daniel Loewe
Cristina Pombo





Copyright © 2019 Banco Interamericano de Desarrollo. Esta obra se encuentra sujeta a una licencia Creative Commons IGO 3.0 Reconocimiento -No Comercial - Sin Obras Derivadas (CC-IGO 3.0 BY-NC-ND) (<http://creativecommons.org/licenses/by-nc-nd/3.0/igo/legalcode>) y puede ser reproducida para cualquier uso no-comercial otorgando el reconocimiento respectivo al BID. No se permiten obras derivadas.

Cualquier disputa relacionada con el uso de las obras del BID que no pueda resolverse amistosamente se someterá a arbitraje de conformidad con las reglas de la CNUDMI (UNCITRAL). El uso del nombre del BID para cualquier fin distinto al reconocimiento respectivo y el uso del logotipo del BID, no están autorizados por esta licencia CC-IGO y requieren de un acuerdo de licencia adicional.

Note que el enlace URL incluye términos y condiciones adicionales de esta licencia.

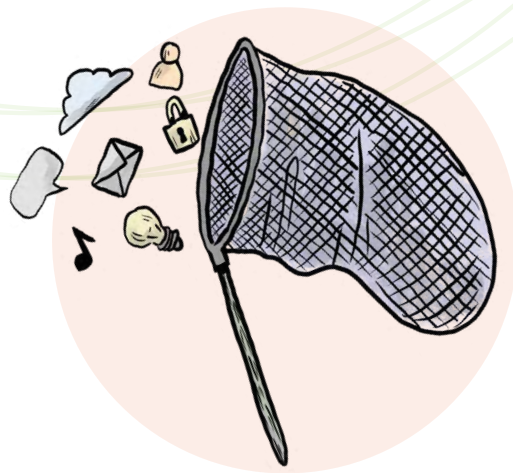
Las opiniones expresadas en esta publicación son de los autores y no necesariamente reflejan el punto de vista del Banco Interamericano de Desarrollo, de su Directorio Ejecutivo ni de los países que representa.



Cada vez más, **actores públicos y privados** se plantean cómo escalar su **impacto a través del uso de la tecnología**. Al mismo tiempo, el **uso y gestión de los datos personales** de millones de personas cada vez preocupa más a los ciudadanos y existe un sentimiento de urgencia sobre la necesidad de **proteger la seguridad y privacidad** de los datos usados.

¿Qué medidas se pueden tomar y cuál es el riesgo de no tomarlas? ¿Cómo puede el sector público gestionar los datos de forma responsable?

Este documento ofrece **marcos de referencia** sobre la **gestión ética de datos y sobre la importancia del consentimiento**, un compendio de **mejores prácticas y una hoja de ruta** con pasos concretos para una gestión responsable de datos por parte del sector público.

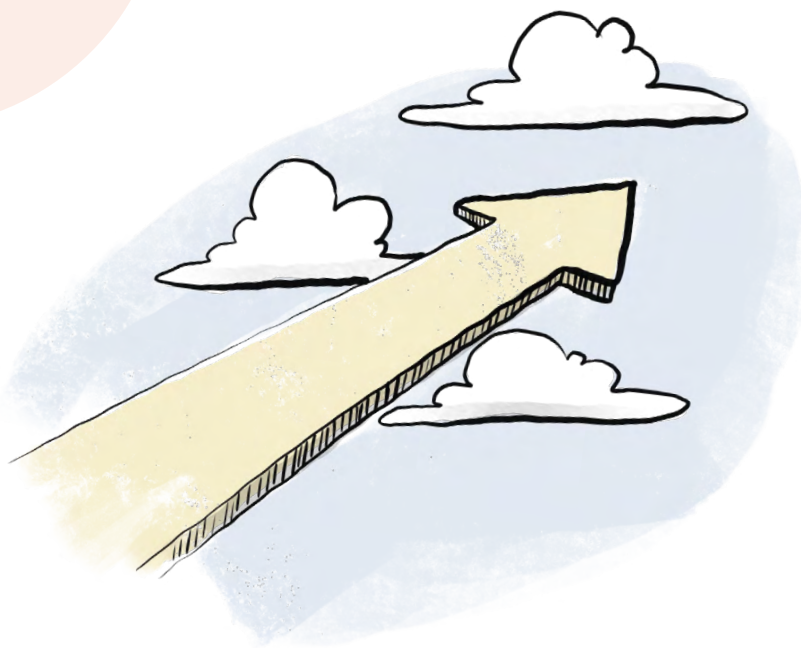




Índice

Introducción	3
1. El potencial del uso de datos en el sector público	6
2. Desafíos en el uso de datos: tipología de riesgos éticos	10
Privacidad	12
Discriminación algorítmica	16
Opacidad	18
3. Marcos de referencia y buenas prácticas en la gestión ética de datos	22
Marcos generales	23
Inteligencia artificial y algoritmos	30
Marcos de gestión ética sectoriales	34
4. Propuesta de criterios para una gestión ética de datos por el sector público	39
Referencias	49

1 El potencial del uso de datos en el sector público



Actualmente vivimos en un universo poblado de datos. Nuestra huella digital va dejando rastro sobre lo que hacemos, dónde lo hacemos, adónde vamos, a quiénes conocemos, qué tenemos, qué nos gusta o cómo nos sentimos. Generamos esta información mientras trabajamos, caminamos, interactuamos, hablamos, protestamos o buscamos información en línea. Las actividades que llevamos a cabo generan datos, y toda esta información es útil para definir los servicios, los productos o la manera cómo funcionan las ciudades. La importancia y alcance de estos datos han sido ampliamente explorados por el

sector privado, sobre todo en los ámbitos del cálculo y análisis de riesgos, y en la personalización de los servicios comerciales (mercadeo).

Paralelamente a esta mayor disponibilidad de datos y a la mejora de las capacidades técnicas para utilizarlos, el sector público también ha comenzado a reivindicar su propio papel, consciente de que esta información puede servir tanto para mejorar la eficiencia en la prestación de servicios sociales, de policía o de transporte público, como para promover la transparencia, la participación ciudadana y la rendición de cuentas.

Ya en 2002, el gobierno de Estados Unidos e IBM había iniciado un esfuerzo colaborativo orientado a gestionar el alto volumen de datos que el gobierno acumulaba. De ahí nacieron IBM InfoSphere Stream e IBM Big Data, dos infraestructuras utilizadas por agencias gubernamentales y organizaciones empresariales para visualizar información procedente de miles de fuentes (Kim, Trimi, y Chung, 2014). En 2012, la National Science Foundation y el National Institute of Health (NIH) de Estados Unidos se unieron para lanzar el Programa de Técnicas y Tecnologías Fundamentales para el Avance de la Ciencia e Ingeniería de Datos (Core Techniques and Technologies for Advancing Big Data Science & Engineering Program) “que apunta a promover los medios científicos y tecnológicos principales para administrar, analizar, visualizar y extraer información útil de conjuntos de macrodatos diversos, distribuidos y heterogéneos” (Kim, Trimi y Chung, 2014).

En América Latina y el Caribe, la ciudad de Río de Janeiro fue pionera en este sentido cuando en 2010 inauguró un centro de operaciones que integraba datos de 30 agencias que operaban en la ciudad (Singer, 2012). En ese momento se trataba de un centro único en el mundo, y se basaba en la idea de que tener acceso a una mayor cantidad de datos provenientes de diferentes agencias relacionados con áreas de gobierno previamente desconectadas permitiría a las autoridades locales formular mejores políticas y tomar mejores decisiones. En particular, los funcionarios de la ciudad estaban interesados en mejorar la respuesta a situaciones de emergencia.

Un experimento similar se llevó a cabo en Nueva York por esa misma época, cuando se estableció una Oficina de Planificación Estratégica y Política (Feuer, 2013). La lógica subyacente era exactamente la misma que había inspirado a su homóloga brasileña: el procesamiento de datos masivos sobre múltiples aspectos de la vida de los ciudadanos para construir modelos predictivos que mejoren la toma de decisiones en el ámbito local. Por ejemplo, la agencia utilizó datos provenientes de la Comisión de Integridad Empresarial para detectar aquellos negocios que tenían más probabilidades de no cumplir con ciertos requisitos de higiene. Con base en esos datos, y a partir de un sistema inteligente para detectar declaraciones de impuestos posiblemente fraudulentas, el Estado de Nueva York destinó mayores recursos a la lucha contra el fraude fiscal, como resultado de lo cual aumentaron los ingresos de las arcas públicas en USD 100 millones, mientras que las transferencias fraudulentas disminuyeron en más de USD 1000 millones en el primer año (Thayer, 2014). Nueva York también utiliza y cruza datos provenientes de diferentes agencias para abordar de manera proactiva el déficit de vivienda y entender mejor sus causas (NYC Center for Innovation through Data Intelligence, s.f).

En los Países Bajos, los datos se utilizan para luchar contra el fraude en las prestaciones sociales. Con base en datos históricos que incluyen información sobre fechas de nacimiento, composición familiar, historial de prestaciones y declaraciones de renta, se establece quién puede estar haciendo un uso fraudulento del sistema. Los Ángeles,

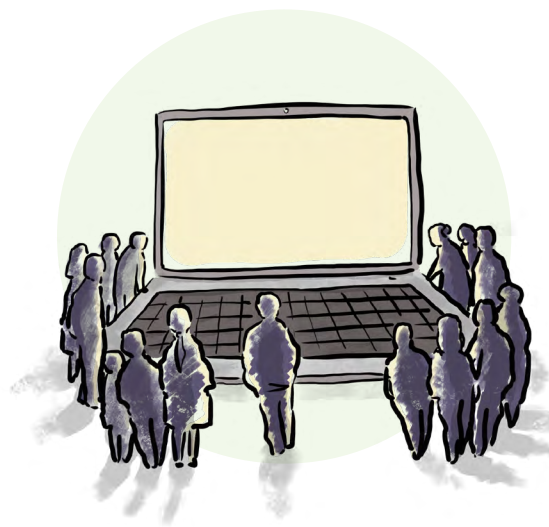
California, lanzó un proyecto parecido en el cual se analizan los datos para detectar el fraude en el uso de los servicios de cuidado infantil (SAS, s.f.).

En el sector de la educación, el distrito de Tacoma en el estado de Maryland, Estados Unidos, cuenta desde 2014 con un modelo para calibrar el porcentaje de deserción escolar presente y futura en los planteles de secundaria. En 2010, solo el 55% de los estudiantes de secundaria del distrito se graduó, una cifra muy baja comparada con el promedio nacional de 81%. El distrito se asoció entonces con Microsoft para usar el análisis de datos con el propósito de entender mejor el desempeño de los estudiantes. Se estudiaron cinco años de datos de estudiantes de secundaria a nivel tanto grupal como individual, y con base en el diagnóstico se implementaron estrategias de intervención específicas que permitieron elevar la proporción de egresados al 82,6% en 2016¹. En Colombia, el Ministerio de Educación Nacional y las Instituciones de Educación Superior implementaron un proceso encaminado a determinar las causas del abandono escolar de los estudiantes universitarios². India y los Países Bajos utilizan modelos similares (Ministerio de Hacienda de Chile, 2017).

Los ejemplos anteriores ilustran la manera en que los datos sirven para mejorar la eficiencia y los diagnósticos en el sector

¹ Harvard Business School, Digital Initiative: <https://rctom.hbs.org/submission/microsoft-helps-tacoma-public-schools-use-data-analytics-to-predict-at-risk-students/>

² IESALC: http://www.iesalc.unesco.org.ve/index.php?option=com_content&view=article&id=1141:s-padies&catid=120:servicios&Itemid=535



público. Pero su uso y proliferación también están permitiendo mejorar la transparencia de la acción gubernamental a través de los datos abiertos. **Un número creciente de gobiernos está publicando datos en sus portales de internet con el objeto de mejorar su transparencia, pero también con la esperanza de que esa disponibilidad estimule el desarrollo de aplicaciones útiles para la ciudadanía, ya sea por parte de las entidades gubernamentales o de los propios ciudadanos.** Uno de estos proyectos actualmente en marcha es “Mejora tu escuela” en México (Young and Verhulst, 2016). La plataforma, creada por el Instituto Mexicano para la Competitividad en 2013, busca ayudar a los padres a elegir la mejor opción de educación para sus hijos, lo que a su vez crea incentivos para que las escuelas tengan un mejor desempeño. La información de la plataforma sobre los planteles educativos combina datos públicos abiertos con otros proporcionados por la ciudadanía a través de reseñas y puntajes.

Cabe notar, sin embargo, que en muchos de los países mencionados estos ejemplos no constituyen un esfuerzo concertado para avanzar hacia un mayor y mejor uso de los datos por parte del sector público. Se trata más bien de proyectos aislados que enfrentan dificultades para darse a conocer y desarrollar su potencial, o de iniciativas que no se aprovechan plenamente debido a las limitaciones en la formación del personal gubernamental o a las dificultades para liderar el desarrollo tecnológico desde el sector público.

De acuerdo con lo anterior, se puede afirmar de manera general que no existen aún estrategias nacionales o regionales que permitan avanzar en la definición de los usos de estos datos en el sector público, de su potencial y sus riesgos, de los estándares de interoperabilidad necesarios y/o de cómo articular la relación de confianza con la ciudadanía que proporciona su información con la esperanza de que se le dé un buen uso. Desafíos como la discriminación algorítmica, la privacidad, la transparencia o los derechos digitales no se han abordado aún de forma consistente como parte de una visión estratégica sobre la tecnología. En este escenario, la irrupción de nuevas herramientas tecnológicas como el aprendizaje automatizado, la predicción algorítmica y la inteligencia artificial corren el riesgo de no ser aprovechadas plenamente en beneficio de la ciudadanía. Asimismo, y como se ha podido observar a partir de los escándalos recientes que involucran a empresas del sector privado intensivas en el uso de datos, es necesario evitar que, sin contar con la experiencia

necesaria, los gobiernos caigan en la tentación de desarrollar sistemas que utilicen los datos de los ciudadanos sin que incorporen paralelamente un enfoque ético y responsable desde su diseño que asegure cuestiones básicas como la privacidad o el consentimiento. **Lo que aquí está en juego no es otra cosa que sentar las bases para un nuevo contrato social que permita una utilización masiva y responsable de los datos por parte de las entidades gubernamentales para proporcionar mejores servicios sociales, al tiempo que se mantiene la confianza de los ciudadanos en que los gobiernos gestionen sus datos de manera responsable.**



2 Desafíos en el uso de datos: tipología de riesgos éticos



En la sección anterior se describieron brevemente algunos de los casos más interesantes sobre el uso de datos en el sector público. Algunos de estos ejemplos son modelos exitosos, aunque también es cierto que la implementación de tales sistemas no ha estado exenta de problemas, que también han dejado enseñanzas valiosas. Aquí se describen entonces dichos desafíos, identificando y tipificando los riesgos subyacentes con el fin de comenzar a pensar en modelos de mitigación que permitan una utilización ética de los datos.

En el caso del sistema para combatir el fraude en los servicios sociales de los Países Bajos arriba mencionado, es importante destacar que se optó porque este sistema no fuera de aplicación automática, sino un elemento más de evaluación de los posibles riesgos a tener en cuenta por parte del personal del sector, que es el que finalmente toma la decisión de abrir una investigación o suspender una prestación. En ese caso se estimó que la mejor forma de aprovechar plenamente las oportunidades que ofrece la disponibilidad de datos en los sistemas de gestión pasaba por combinar su potencial intrínseco con

la intervención humana (Hijink, 2018). Este uso particular de la inteligencia artificial puede permitir que el erario ahorre millones de dólares, aunque también plantea serias dudas de carácter legal y ético, tales como las relacionadas con el consentimiento de uso, la identificación prejuiciada (profiling) o la perpetuación de ciertos sesgos y prejuicios del pasado que puedan estar presentes en los datos con los que se alimenta y entrena al sistema. De hecho, la Autoridad Holandesa de Protección de Datos aún no se ha pronunciado sobre la legalidad de este método de prevención de fraude.

Otro caso pertinente es el de Transport for London (TfL) entidad a cargo de la Oyster Card. TfL tuvo que cambiar sus sistemas para permitir que los usuarios³ pudieran registrarse de forma anónima. En realidad, incluso cuando las personas acceden a dar sus datos, estos se anonimizan pasadas ocho semanas, precisamente para proteger su privacidad (Muller, 2018) y como respuesta a la presión ciudadana.

Los riesgos y desafíos que entraña la sociedad de los datos no afectan solo al sector público. El lema de Silicon Valley, “muévete rápido y rompe cosas” sonaría catastrófico si lo que se rompe son las estructuras sociales o las garantías legales consolidadas en entornos por fuera del ciberespacio. **Empresas globales intensivas en el uso de datos como Uber y Airbnb han sufrido crisis cíclicas de confianza por la manera como gestionan los datos personales de sus clientes.**

³ En este documento se usa el masculino genérico no marcado para incluir a todas las personas, independientemente de su género y/o su sexo.



En el sector de la educación, donde se manejan datos sensibles de personas menores de edad, la preocupación por la privacidad llevó a la administración Obama a promulgar en 2015 una Ley de Privacidad Digital para los Estudiantes. Entre tanto, empresas como Google, Apple y Microsoft, así como otras grandes proveedoras de tecnologías vinculadas al aprendizaje, han firmado el Compromiso con la Privacidad de los Estudiantes para abordar específicamente estos temas. En Estados Unidos, la preocupación por la manera como se recolectan y gestionan los datos del estudiantado se puso de manifiesto en 2014 cuando InBloom, una de las empresas jóvenes más destacadas de la era digital, terminó en la quiebra después de que se la acusara de vulnerar la privacidad tanto de los alumnos como de las escuelas. InBloom ofrecía a los centros educativos un espacio en la nube para almacenar todos los datos resultantes de la relación entre el alumnado y su escuela.

Este fue aprovechado por numerosos distritos, hasta que algunos padres de familia empezaron a preguntar sobre la seguridad y la privacidad de tal información. ¿Podían terminar los historiales académicos de sus hijos en manos de terceros y determinar por esa vía el futuro de aquellos? ¿Cómo se protegían los datos personales delicados, como por ejemplo las dificultades de aprendizaje y otros asuntos familiares relevantes? ¿Acaso InBloom se proponía almacenar esta información de manera permanente? ¿Qué decisiones se estaban tomando sobre estos chicos con base en los datos generados por su comportamiento en línea? El caso de InBloom puso en evidencia el interés comercial por esta cantidad ilimitada de datos personales, así como la ubicuidad de las tecnologías en el aula y la necesidad de ser más cautelosos a la hora de incorporar los sistemas de proveedores externos en entornos educativos.⁴

A la luz de lo expuesto anteriormente, a continuación, se presenta una propuesta de tipología de riesgos que permite clasificar los distintos desafíos éticos a los que las entidades se enfrentan en la gestión de datos, en particular de los de las personas.

⁴ https://blogs.edweek.org/edweek/DigitalEducation/2014/04/inbloom_to_shut_down_amid_growing_data_privacy_concerns.html



Privacidad

El primer riesgo –y el más evidente– al que se enfrentan quienes manejan datos personales es el de la protección de estos y, en un sentido más amplio, el de la privacidad. En realidad, es precisamente esta

última la que ha producido un mayor número de referencias y marcos jurídicos, no solo por el hecho de que se trata de un derecho reconocido en muchas constituciones nacionales, sino especialmente debido a la expedición y entrada en vigor en 2018 del Reglamento General de Protección de Datos (RGPD) de la Unión Europea, que en poco tiempo se ha convertido en una suerte de estándar de referencia global (CNIL, S.f.).

El RGPD establece seis principios básicos sobre el manejo de los datos personales: (i) deben ser tratados de forma lícita, leal y transparente; (ii) se deben recolectar con fines determinados explícitos y legítimos; (iii) deben ser adecuados, pertinentes y limitados a lo necesario dependiendo del uso; (iv) deben ser exactos y estar siempre actualizados; (v) deben mantenerse de forma tal que se permita la identificación de los interesados durante no más tiempo del necesario para los fines del tratamiento; y (vi) deben ser tratados de tal manera que se garantice su seguridad. Sin entrar aquí en los pormenores del reglamento, los cuales se

abordarán en la siguiente sección, uno de los elementos esenciales de este texto es que allí se establece el consentimiento como base de la gestión de datos personales en casi todos los casos. Esto implica que, además de los temas atinentes a la seguridad y a la legalidad, quienes recolecten y gestionen los datos deberán asegurarse siempre de haber informado a sus propietarios (los individuos) y obtener su consentimiento tantas veces como sea necesario si la finalidad del uso que se le dará a sus datos cambia.

Es evidente que las administraciones públicas cuentan con fundamentos jurídicos para el tratamiento de los datos de la población, y que en ciertos casos dotar a la ciudadanía del poder de eliminar su información personal de una base de datos sería irresponsable (su información tributaria, por ejemplo). No obstante, y como se verá más adelante, en el caso del uso de datos de los individuos en modelos predictivos o en políticas de gran impacto sobre derechos fundamentales, el consentimiento ciudadano se ha erigido en un eje fundamental del contrato social que permite generar confianza entre quienes proporcionan los datos y quienes los manejan.

Otro elemento importante es la definición de lo que se entiende por datos personales, lo cual no se limita a nombres y apellidos, sino que además incorpora cualquier elemento que pueda llevar a la identificación de un sujeto determinado. Así, los identificadores únicos de computadores y teléfonos, como también los datos geolocalizados y los biométricos, constituyen datos personales y deben quedar sujetos a protección legal incluso cuando no se

encuentran directamente asociados a un nombre propio. Si se busca compartirlos o abrirlos, será necesario definir un protocolo robusto de anonimización o seudonimización⁵ para evitar su mal uso o su empleo con fines distintos a los expresados cuando se los recolectó en una primera instancia. Es importante insistir en que la anonimización deberá ser robusta, pues, aunque una base de datos haya sido anonimizada, su cruce con otras puede derivar en la reidentificación de algunos individuos (Kupersmith, 2013).

Esta posibilidad fue revelada por un grupo de investigadores a partir de la base de datos de ADN “1000 Genomes Project”, la cual, aunque teóricamente anonimizada, permitió la reidentificación de 50 personas (Gymrek et al., 2013). En 2016, algo similar sucedió en Australia cuando dos programas de atención médica y farmacéutica publicaron registros de pacientes que, en teoría, habían sido anonimizados. Sin embargo, como los métodos utilizados no habían sido lo suficientemente robustos, los investigadores de la Universidad de Melbourne encontraron que era posible reidentificar a los individuos haciendo coincidir partes no cifradas de los datos con información previamente conocida sobre personas específicas, como por ejemplo su género o el tipo de intervenciones médicas a las que habían sido sometidas (Grubb, 2017). Un caso parecido, relacionado

⁵ La anonimización es el tratamiento de datos personales de manera tal que sea imposible vincularlos con la persona a quien se hubiese podido identificar a través de aquellos. La seudonimización, por su parte, es el tratamiento de datos personales de manera tal que no se los pueda atribuir a un individuo determinado sin utilizar información adicional. En este caso se reemplaza el dato identificador (nombre, identificador único, etc.) por un código.

esta vez con la anonimización incorrecta de datos georreferenciados, se presentó cuando la Comisión de Taxis y Limusinas de la ciudad de Nueva York publicó una base de datos de 173 millones de recorridos de vehículos a partir de la cual se podía inferir información sensible como la dirección del domicilio de los conductores, su salario anual y su afiliación religiosa (Metcalf y Crawford, 2016: 9). Estos mismos registros permitieron a los periodistas cruzar las matrículas de los taxis a los que se subían personas famosas y determinar cuál era su destino.

El énfasis en la privacidad no es un asunto menor. En la literatura especializada se ha descrito ampliamente la manera en que **el avance hacia contextos de vigilancia masiva, tipo Gran Hermano, tiene efectos graves en aquellos derechos fundamentales vinculados con la libertad de expresión, reunión y manifestación o a la presunción de inocencia;** en valores como la confianza y la cohesión social; y en procesos humanos importantes como el desarrollo de la identidad.



Uno de los efectos positivos de una gestión responsable de los datos es que esta se traduce de forma automática en una mayor seguridad de tal información y de los sistemas que la almacenan. Los sistemas informáticos están protegidos de diferentes formas contra robos informáticos. **Los cortafuegos, el cifrado, la anonimización y la codificación son algunas de las modalidades de protección de los datos personales,** tanto en los programas como en los equipos informáticos. Aun así, el acceso ilegal o piratería (hacking) es habitual, y cuando la información no está anonimizada, el impacto de estos robos es tan significativo que el RGPD obliga a divulgar públicamente estos hechos. Igualmente prevé multas multimillonarias para aquellos casos en que los responsables de los datos no hayan tomado medidas de seguridad y anonimización acordes con los riesgos existentes.

Los gobiernos no son inmunes a las violaciones de los protocolos de seguridad. En 2008, un pirata informático accedió ilegalmente a las bases de datos gubernamentales y publicó información personal perteneciente a seis millones de chilenos (BBC News, 2008). La información filtrada incluía números de identificación, direcciones y nombres que se publicaron en varios foros de tecnología antes de que la policía pudiera intervenir. El problema en este caso fue la debilidad del cortafuegos utilizado para proteger los datos. Además, el hecho de que estos no estuvieran cifrados y/o anonimizados convirtió el error en una catástrofe.

A menudo las fallas de seguridad son producto de la acción humana.

En México, por ejemplo, un manejo negligente de la información por parte de los partidos políticos condujo a que los datos electorales del 75% de los mexicanos –es decir, más de 93 millones de personas– quedaran a disposición del público en la plataforma de AmazonCloud (Derechos Digitales, 2016). La información publicada consistía en fotografías, identificaciones y direcciones domiciliarias (Derechos Digitales, 2016). En Chile, el Ministerio de Salud sufrió una fisura significativa en la protección de los datos, la cual fue revelada por el Centro de Investigación Periodística (CIPER, 2016). Como resultado de ello, unos 100.000 funcionarios y proveedores externos pudieron acceder a los nombres, direcciones, RUT (registro único tributario), historias clínicas y tratamientos de las personas que padecen de enfermedades mentales, portan el VIH o han solicitado la píldora del día después. Esta fisura en la protección de los datos, tal como lo indica CIPER (2016), dejó cerca de tres millones de archivos expuestos durante meses, a pesar de que contenían información sumamente delicada. Estas filtraciones son habituales en todo el mundo.

Aquí también cabe subrayar que el uso intensivo y masivo de datos, y las disfunciones o errores que se producen ocasionalmente, está dando lugar al surgimiento de nuevos derechos, muchos de ellos cobijados en el concepto de derechos digitales. Específicamente, y en el contexto de la privacidad, el RGPD cubre el derecho al olvido (Artículo 17), el cual permite que los ciudadanos retiren su consentimiento sobre el uso de sus datos y exijan que estos sean eliminados si, por ejemplo, ya no se los requiere para la finalidad original que dio lugar a su obtención.



Discriminación algorítmica

La discriminación es un trato diferente y perjudicial que se da a una persona debido a categorizaciones arbitrarias o irrelevantes. Se la califica de “algorítmica” porque lo que permite que la discriminación se instale y prolifere en los sistemas informáticos es el uso del aprendizaje automatizado y de la inteligencia artificial. **La discriminación**

algorítmica refiere entonces a aquellos procesos a través de los cuales los distintos tipos de discriminación que ocurren en el mundo real son reproducidos en entornos de datos, o a los que surgen exclusivamente

en ellos, como cuando los sistemas de reconocimiento facial producen más errores al procesar rostros no caucásicos. Como bien lo señalan Barocas y Selbst (2016: 671), los datos son “frecuentemente imperfectos” porque pueden reflejar los sesgos de las personas que tomaron las decisiones sobre su recolección. Asimismo, pueden presentar

problemas de insuficiencia, errores, y exceso o déficit de representación de ciertos grupos de la sociedad, todo lo cual podría redundar en una decisión algorítmica equivocada. La falta de actualización oportuna de los datos puede tener efectos en el sujeto de estos, o en terceros. Por ejemplo, si no se actualizan a tiempo los datos de una persona que estaba en un registro de deudores, esta puede resultar perjudicada a la hora de ser evaluada para un empleo o al solicitar un crédito.



Estados Unidos es el país más citado en este aspecto, y también el que más ha empleado sistemas de decisión algorítmica en la prestación de servicios. En el caso de la justicia, por ejemplo, se utilizan algoritmos para asistir a los jueces en el proceso de dictar sentencias (en particular en el cálculo del riesgo de reincidencia de los convictos), para determinar los montos de las fianzas o para conceder la libertad condicional (Angwin et al, 2016). Uno de los proveedores de dichos algoritmos es Northpoint, la empresa estadounidense creadora del sistema COMPAS de seguimiento y evaluación de la actividad policial, otro ámbito donde proliferan los algoritmos. Sin embargo, como lo señala Martin (2018), los cálculos algorítmicos del sistema COMPAS para predecir la reincidencia en crímenes violentos alcanzan márgenes de error del 80%. En un estudio realizado por ProPublica (Angwin et al, 2016) se mostró igualmente que, con el uso de esta herramienta, las posibilidades de que las personas afroamericanas sean consideradas como de mayor riesgo casi duplica las de los blancos, aunque nunca reincidan. En cambio, es mucho más probable que las personas blancas sean calificadas de bajo riesgo, aunque posteriormente cometan un nuevo delito.

Esta discriminación no es más que la consecuencia de reproducir algorítmicamente procesos discriminatorios que ocurren en el mundo real, sin subsanar sus impactos en el momento de programar los algoritmos. La capacidad teóricamente predictiva de estos sistemas no es más que la extrapolación en el futuro de dinámicas

identificadas en el pasado. De esta forma, un algoritmo en el cual se analicen datos de sentencias judiciales y reincidencia en Estados Unidos aprenderá que en el pasado la población afroamericana ha sufrido tasas de encarcelación y delincuencia superiores a las de la población blanca. Si el sistema algorítmico registra el valor “etnia”, asignará automáticamente mayor peligrosidad a una persona afroamericana que a una blanca. Siguiendo una lógica similar, los algoritmos publicitarios de Google asignan ofertas de trabajo a personas con perfiles específicos: dado que históricamente los trabajos bien remunerados han estado copados por hombres (solo el 5% de los presidentes de las empresas de la lista Fortune 500 son mujeres, por ejemplo),⁶ los sistemas de recursos humanos discriminan a las mujeres porque en lugar de entrenarlos con datos que tengan en cuenta las competencias y experiencia de los aplicantes, han sido programados para captar valores sencillos como el género sin corregir los sesgos del pasado⁷.

Es importante recordar que los sesgos existen también en el proceso humano de toma de decisiones. Los juicios de los individuos a menudo pueden resultar afectados por sesgos implícitos que se generalizan en muchas áreas de la vida cotidiana (Saul, 2012). El reto consiste en incorporar los algoritmos, el aprendizaje automatizado y la inteligencia artificial para evitar los sesgos humanos, no para reproducirlos. Infortunadamente abundan

los casos de algoritmos con los cuales se ha contribuido a empeorar los procesos, por lo que cada vez son más numerosas las voces que abogan por la transparencia algorítmica. En Nueva Zelanda, por ejemplo, el servicio de inmigración tuvo que cancelar un piloto de modelos predictivos de riesgo para priorizar las deportaciones de inmigrantes cuando un medio de comunicación descubrió que el sistema discriminaba por nacionalidad (RadioNZ, 2018a y b).

En el contexto de los algoritmos, la proliferación de esta tecnología ha llevado también al surgimiento de nuevos derechos, específicamente, el derecho a

la explicabilidad. El RGPD, por ejemplo, requiere a los organismos que manejan algoritmos que realicen un procesamiento justo y transparente, y que expliquen la manera como los sistemas automatizados toman decisiones, especialmente aquellas que afectan significativamente las vidas individuales (como el acceso a prestaciones sociales, a un trabajo o a un crédito).

Igualmente, cualquier persona que haya sufrido un daño (material o no) debido a la vulneración de su privacidad tiene **derecho a la compensación**. En algunos casos, como en el contexto del control automático de fronteras en la Unión Europea, se han expedido medidas encaminadas a reparar y compensar a los viajeros, permitiéndoles el acceso y la posibilidad de rectificar los datos que los sistemas puedan albergar sobre ellos⁸.

⁶ Fortune: <http://fortune.com/2018/05/21/women-fortune-500-2018/>

⁷ The Guardian: <https://www.theguardian.com/technology/2015/jul/08/women-less-likely-ads-high-paid-jobs-google-study>

⁸ EUR-lex: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52011DC0680>

Opacidad

Como se observó en el apartado anterior, **frente a los errores y riesgos de los sistemas, una de las exigencias más generalizadas es la transparencia.** Esta última noción, así como la evaluación y la rendición de cuentas, se encuentran íntimamente ligadas, dado que para que los ciudadanos logren formarse un juicio sobre un determinado aspecto relativo a la acción gubernamental o pública, es necesario que la información pertinente se encuentre disponible. Es por eso que la falta de transparencia en los sistemas de datos (qué clase de información recogen, cómo la gestionan, cómo la analizan, con quién la comparten, qué decisiones se toman a partir de ella y con base en qué factores) supone un problema de gran calado para la calidad de la democracia y la posibilidad de que la ciudadanía acceda a información valiosa sobre cómo se toman las decisiones que la afectan.

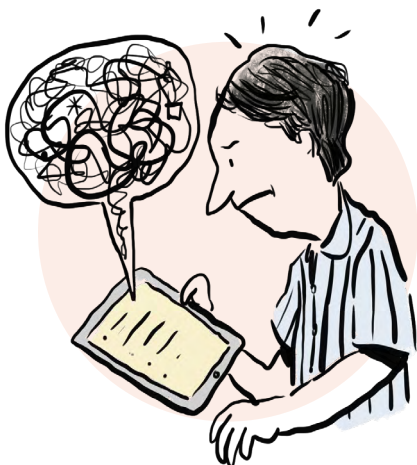


En el ámbito tecnológico, **este riesgo se acentúa debido a que los sistemas informáticos son percibidos por la mayoría de la población como “cajas negras”, es decir, como mecanismos incomprensibles que a menudo bordean lo mágico.** En este contexto, no es posible exigir al ciudadano un dominio pleno de los sistemas que afectan su vida cotidiana o de los servicios a los que tiene acceso.

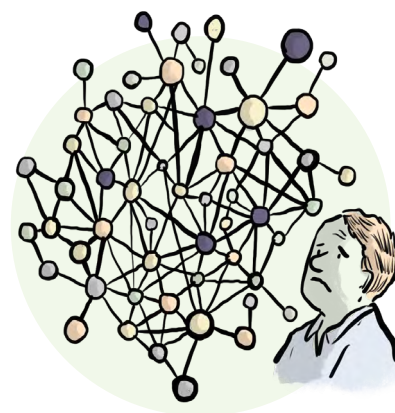
En el contexto de las nuevas tecnologías de datos, se pueden identificar de tres tipos de opacidad (Burrell, 2016):



Opacidad intencional: Es la que ocurre cuando, por ejemplo, un algoritmo no es transparente por motivos relacionados con la propiedad intelectual (el motor de búsqueda de Google, por ejemplo), o cuando la revelación de su funcionamiento pudiera llevar a su irrelevancia (en la detección de evasión de impuestos, por ejemplo). Mientras que en este segundo caso la opacidad respondería al interés público, el uso de un algoritmo opaco por razones de propiedad intelectual para asignar una prestación social, por ejemplo, podría impedir la transparencia y rendición de cuentas del departamento responsable de la implementación.



Opacidad analfabeta: Es la que ocurre debido a la falta de competencias técnicas de las personas potencialmente afectadas para entender el funcionamiento de los algoritmos y los modelos de aprendizaje automático. En este caso, aunque un algoritmo sea explicable, si este es comprensible solo para aquella población dotada de altas capacidades técnicas, la confianza necesaria entre los actores involucrados puede sufrir. En consecuencia, la transparencia formal no sería aconsejable en el desarrollo e implementación de algoritmos de alto impacto para prestaciones o servicios sociales básicos sin que simultáneamente se haga un esfuerzo pedagógico y de transparencia significativo.



Opacidad intrínseca: Surge de la dificultad de explicar los procesos algorítmicos en los que participan redes neuronales complejas. En los casos de aprendizaje automatizado, un sistema algorítmico puede realizar tal cantidad de cálculos que en últimas ni siquiera sus creadores logran explicar cómo llegó el sistema a una solución o cálculo determinado. Esta opacidad es difícilmente justificable cuando los elementos en juego forman parte de la canasta básica de servicios sociales o si estas redes neuronales deben decidir sobre la asignación de oportunidades vitales clave.

La opacidad puede determinar la diferencia entre un error subsanable y una crisis de confianza entre el gobierno y la ciudadanía.

El gobierno francés, por ejemplo, sufrió una crisis de confianza grave cuando implementó un sistema de admisión a las universidades que tomaba en cuenta las preferencias de los estudiantes, las preferencias de otros estudiantes similares y los cupos disponibles

en las instituciones de enseñanza superior. El algoritmo utilizado resultó ser un fracaso pues recomendó a los alumnos plazas que no habían seleccionado; y como además no era transparente, los estudiantes potenciales no entendieron por qué se les recomendaban cupos que no habían solicitado (Bordas, 2018a y 2018b).

Un caso relevante en este contexto es el de la tecnología de reconocimiento facial, actualmente en auge. A pesar de que su uso se está generalizando, la ciudadanía desconoce los índices de falsos positivos o falsos negativos (personas identificadas o no identificadas de manera errónea) de estos sistemas, y de hecho la propia policía metropolitana del Reino Unido ha admitido una tasa de falsos positivos del 98% (Big Brother Watch, 2018), lo que la convierte en una tecnología claramente problemática. Aun así, actualmente se la está implementando en sistemas de fiscalización para los medios de transporte en Chile, por ejemplo (La Segunda, 2018 y El Observador, S.f.).

La Policía de Investigaciones de Chile, junto con una de las municipalidades de la capital, ha puesto en práctica también un sistema de reconocimiento facial para prevenir la delincuencia. A través de cámaras instaladas en lugares públicos, se busca encontrar a personas con órdenes de captura o con antecedentes penales. Según la policía (PDI, 2018), la base de datos cuenta con antecedentes de más de 250.000 personas que han delinquido en la capital. Sin embargo, la institución no indica qué procedimiento seguirá ante un falso positivo, es decir, cuando se identifique a una persona como autora de un delito

que en realidad no cometió. Para aquellos ciudadanos erróneamente identificados no solo no existe información, sino que tampoco está claro cómo se los va a resarcir por los inconvenientes o vulneraciones de derechos causados por el error.

Abordar el tema de la transparencia no se limita entonces al deber de publicar datos sobre un proceso determinado, sino que exige asimismo implementar medidas que permitan a los gobiernos rendir cuentas sobre las decisiones que toman, sus motivaciones y los impactos esperados. Así, la transparencia parte de la capacidad de realzar el valor público de determinada información y de justificar la necesidad de adoptar una cierta medida tecnológica, así como de identificar sus impactos positivos y negativos. Es precisamente la transparencia en la planificación, en la toma de decisiones y en la evaluación del impacto lo que articula un gobierno comprometido con la gestión responsable de la información y los datos.

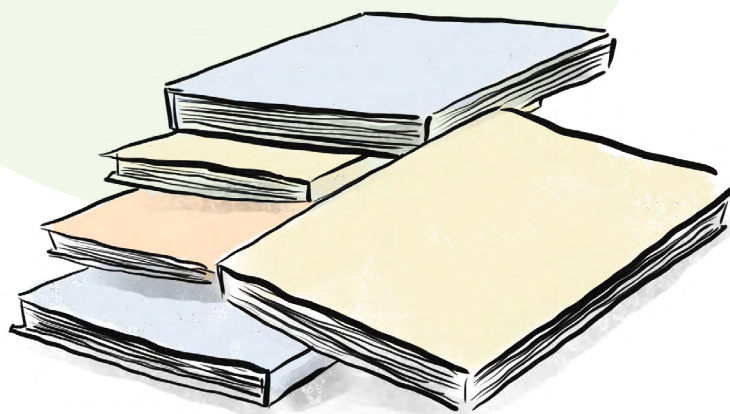
La necesidad de combatir la opacidad de los sistemas de datos es un tema que va más allá del uso de algoritmos, por lo cual debe abordarse en diferentes niveles y en distintos momentos. En el caso de la administración pública, una instancia clave y sensible en la implementación de medidas tecnológicas es la de las licitaciones públicas.

No es de extrañar que, en el momento de comprar nuevas tecnologías a actores privados, aparezcan contratos que no definen con claridad las necesidades técnicas, las limitaciones legales y sociales, y/o las precauciones que se deben tomar para prevenir el mal uso de los datos y sus

impactos sociales negativos. En este sentido, es clave abordar la formación de los equipos humanos del sector público encargados de definir los procesos de licitación de nuevas herramientas tecnológicas, con el fin de evitar que el gobierno quede expuesto y/o que se les dé un uso no permitido a los datos de la ciudadanía.

Por último, son muchas las voces que reivindican la necesidad de planificar metodologías de auditoría algorítmica para poder evaluar de forma periódica cómo van aprendiendo los algoritmos y cuáles son sus impactos. En los casos menos sensibles, esta tarea puede ser asumida por la misma administración o el proveedor de la tecnología. En algoritmos con impactos sobre servicios u oportunidades vitales, es recomendable contar con un consejo externo o un proceso de auditoría independiente.

3 Marcos de referencia y buenas prácticas en la gestión ética de datos



En este documento se ha discutido hasta ahora tanto el potencial como los riesgos del uso de los datos para mejorar la eficiencia y transparencia de la administración y la prestación de servicios públicos.

Precisamente para abordar los desafíos de este campo emergente, numerosos sectores y entidades han elaborado marcos éticos y de buenas prácticas. Estos son tan diversos como la casuística existente en torno al uso de datos. Para facilitar su análisis, acceso y consulta, en esta sección se discuten los marcos de referencia con base en los temas o ámbito de aplicación correspondientes.

Se privilegiarán aquellos ámbitos para los cuales existe un mayor número de referencias en la literatura. Es decir, no se trata de una descripción de los marcos y enfoques deseables, sino de los que ya existen. A partir de estas contribuciones, en la última parte

de este informe se presentará una propuesta específica para el sector social que recoge los mejores aportes y prácticas existentes en el contexto de los servicios de atención a la ciudadanía.

La sección se inicia con un análisis de los marcos más generales y aquellos centrados en la inteligencia artificial y los algoritmos, para finalmente cubrir marcos sectoriales enfocados en la salud y los datos clínicos, el reconocimiento facial y la investigación científica, que son algunos de los ámbitos que más atención han recibido desde una perspectiva ética y de responsabilidad e impacto⁹.

⁹ Si bien es cierto que al ámbito de la automoción y de los vehículos autónomos se le ha prestado mucha atención debido a los dilemas éticos que plantea, se trata de un área que supera los límites del presente informe.

Marcos generales

Desde hace más de tres décadas, el mundo comenzó a tomar conciencia sobre la importancia de proteger los datos personales. Fue así como en 1980, la Organización para la Cooperación y el Desarrollo Económicos (OCDE), con la colaboración del Consejo de Europa, publicó una guía para resguardar la privacidad y los flujos transfronterizos de datos personales a partir de los siguientes principios (OECD, 2013):

- I Establecer límites claros para la obtención de los datos.
- II Determinar la relevancia de los datos para el uso previsto.
- III Definir con claridad el uso que se dará a los datos antes de solicitarlos.
- IV Abstenerse de utilizar los datos para usos distintos al determinado originalmente sin el consentimiento de las personas afectadas.
- V Asegurarse de proteger los datos contra el acceso ilícito o piratería.
- VI Asegurar que los avances, prácticas y políticas sobre el uso de los datos sean abiertos y transparentes.
- VII Garantizar que las personas cuyos datos se han recolectado tengan acceso a los mismos y puedan solicitar bien sea modificaciones o su eliminación definitiva.

Estos principios se encuentran en todas las guías y constituyen quizás el marco de referencia que más ha durado sobre la protección de datos personales. Por su parte, la Asamblea General de las Naciones Unidas adoptó el 14 de diciembre de 1990 la reglamentación de los ficheros computarizados de datos personales que había escrito en 1989 (Asamblea General de Naciones Unidas, 1989). Allí se describen seis principios que los estados miembros deben seguir:



I

Principio de licitud y lealtad: Las informaciones no se pueden obtener o ser utilizadas de manera desleal o ilícita.

II

Principio de exactitud: Los encargados de un fichero deben verificar periódicamente la exactitud y pertinencia de los datos

III

Principio de finalidad: La finalidad de un fichero debe ser conocida, especificada, legítima y pública antes de su creación.

IV

Principio de acceso de la persona interesada: Toda persona cuyos datos están siendo procesados tiene derecho a saber si se está procesando información que le concierne, así como a obtener las rectificaciones o supresiones adecuadas cuando los registros sean ilícitos, injustificados o inexactos.

V

Principio de no discriminación: No se debe registrar información sobre el origen racial o étnico, el color, la vida sexual, las opiniones políticas, las convicciones religiosas, filosóficas o de otro tipo, o sobre la participación en una asociación o la afiliación a un sindicato.

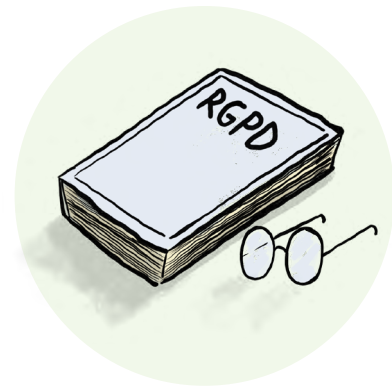
VI

Principio de seguridad: Los ficheros deben ser protegidos adecuadamente contra riesgos naturales y humanos.

El reciente Reglamento General de Protección de Datos-RGPD (CNIL, S.f.) al que se aludió brevemente en la sección anterior, tiene como referente a la Directiva Europea de datos personales publicada en 1995 y a su vez recoge los marcos mencionados en los párrafos precedentes.

El RGPD se elaboró como respuesta al creciente reto que plantea el uso intensivo de datos en la sociedad.

Este integra no solamente todos los principios de la reglamentación de las Naciones Unidas, sino que a su vez crea nuevas herramientas para controlar los datos, confiere un mayor poder a las agencias de protección de datos de cada país, y establece nuevas precauciones y derechos digitales. Con ello busca salvaguardar a los ciudadanos del procesamiento automatizado (algorítmico) de datos, estableciendo para ello los derechos al olvido y a la portabilidad de aquellos, y definiendo un mayor número de requisitos para que el consentimiento pueda considerarse válido, sobre todo en el manejo de información delicada. En general, se trata de una regulación que exige proactividad y garantías a los organismos procesadores y supervisores de datos (los cuales deberán demostrar en ciertos casos la existencia de un Delegado de Protección de Datos en su estructura y realizar estudios de impacto de ciertas actividades que involucran su uso), protege al sujeto de los datos (garantizándole que será notificado de cualquier incumplimiento o acceso indebido a su información), y establece un durísimo régimen sancionatorio en los casos de incumplimiento.



Como resultado de un largo proceso de deliberación y consulta en el ámbito europeo dirigido a generar un marco de referencia útil y relevante para dar forma a la sociedad digital del futuro, **el RGPD se ha constituido desde su entrada en vigor en mayo de 2018 en una especie de estándar global.**

En el nivel nacional, el Reino Unido ha adoptado el uso generalizado de datos y de nuevas tecnologías asociadas con los mismos dentro del contexto europeo. Ello ha provocado que muchas externalidades asociadas con el uso de este tipo de información hayan sido cuestionadas en el debate público, con documentos que abordan el tema no solo desde un punto de vista exclusivamente legal sino desde diversos ángulos. Por ejemplo, el Department for Digital, Culture, Media & Sport del Reino Unido (2018) ha publicado un documento titulado “Data Ethics Framework” en el cual se aborda el uso de datos por parte de las instituciones públicas desde una perspectiva basada en siete principios que coinciden con los establecidos en el RGPD: (i) beneficio claro, (ii) legalidad, (iii) buenas prácticas, (iv) proporcionalidad, (v) limitaciones, (vi) transparencia y (vii) responsabilidad.

En el documento se aborda el tema del consentimiento (que será tratado más adelante) y se señala que la especificidad las circunstancias en las que los datos van a ser transferidos es una condición indispensable para obtener el consentimiento informado. Además, si los datos son utilizados para propósitos distintos a los explicitados al inicio, se debería determinar si existe compatibilidad entre el nuevo propósito y el original. Este marco de referencia incorpora una serie de preguntas concretas que los funcionarios públicos deben responder antes de realizar un proyecto de inteligencia de datos.

En Nueva Zelanda también se ha creado un portal que si bien no constituye un marco ético integral relativo al uso de datos, sí ofrece dar respuesta a preguntas frecuentes relacionadas con su uso (Data Futures Partnership, 2018). Allí se han formulado unas orientaciones generales destinadas a los organismos públicos, a las entidades privadas y a la ciudadanía, donde los principales temas relacionados con el uso de los datos se desarrollan en torno a tres ejes, a saber, valor, protección y elección. A estos se les suma la transparencia como principio transversal.

Valor: necesidad de explicitar qué se va a hacer con los datos, qué efecto positivo va a tener su uso en la vida de las personas (para quién y de qué manera) y quién los va a utilizar.

Protección: seguridad de los datos, privacidad de los usuarios, anonimato y acceso por parte de aquellos a quienes hace referencia.

Elección: consentimiento y posibilidad de elegir qué hacer con los datos en aquellos casos en que estos puedan ser vendidos.

Ya se ha visto cómo el consentimiento es un asunto que siempre sale a relucir al hablar del uso de los datos, dada su relevancia jurídica (RGPD), ética y funcional. No obstante, una cuestión más innovadora que se aborda en el documento neozelandés –y que será cada vez más importante en el futuro– es la de la venta de datos. Este tema también está comenzando a generar debate en el Reino Unido (Perkins, 2018), dado que el gobierno británico tiene previsto invertir grandes sumas en desarrollar sistemas de inteligencia artificial en colaboración con organizaciones del sector privado, lo cual podría conducir a que este último termine lucrando del uso de información perteneciente a la ciudadanía. Según la directriz neozelandesa, la transparencia y el consentimiento informado deben ser la base de este tipo de operaciones. Sin embargo, dado que se trata de guías generales, las especificidades dependerán del sector, la legislación aplicable y otros factores relevantes en cada caso.

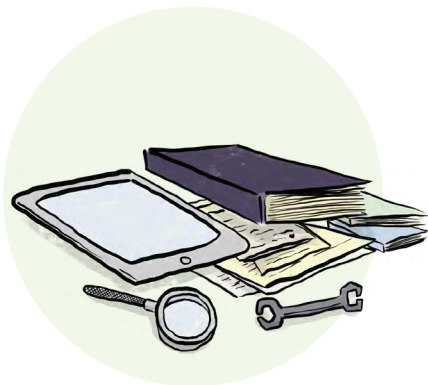
Asimismo, se ha señalado que el valor de este tipo de documentos marco radica sobre todo en que allí se definen cuáles son los aspectos cruciales a tener en cuenta en lo relativo a la utilización de datos por parte de las instituciones públicas. Sin embargo, el contenido es a menudo demasiado general y no aporta mucha información práctica sobre cómo implementar concretamente las recomendaciones de carácter más abstracto. Con el “Data Ethics Framework” (Department for Digital, Culture, Media & Sport, 2018) del gobierno británico, por ejemplo, se intenta suplir este vacío incluyendo referencias a otros códigos y a la legislación pertinente, aunque no está claro que la remisión a otras fuentes facilite la adopción de prácticas concretas por parte de los actores involucrados.

No solo los países están adoptando marcos éticos de referencia. También algunas ciudades, como por ejemplo Ámsterdam, Nueva York y Barcelona, han lanzado iniciativas en este sentido.

Nueva York cuenta con un grupo de trabajo encargado de promover una ley municipal que se prevé esté lista a finales de 2019¹⁰. Ámsterdam está estudiando la posibilidad de empezar a auditar los algoritmos, y Barcelona cuenta con una directiva de datos que incorpora elementos éticos y algorítmicos (Ajuntament de Barcelona, S.f.). Las tres ciudades lanzaron recientemente una Coalición de Ciudades por los Derechos Digitales que tiene precisamente como prioridad la realización de auditorías algorítmicas.

Existen también marcos éticos no gubernamentales sobre este tema. En los Países Bajos, con un manifiesto de profesionales de la región firmado por más de 200 organizaciones se intenta promover una política más ética en el ámbito de los datos. Allí se insiste en la necesidad de que los algoritmos sean desarrollados por y para los ciudadanos de manera abierta, legítima y monitoreada, asegurando la igualdad de las personas y el control sobre sus datos. Los 23 principios de Asilomar (Sterling, 2018) promovidos por más de 100 líderes de opinión, se centran en temas de investigación (objetivos, financiación, valores de equipo y valor social), éticos (seguridad, transparencia, responsabilidad, valores y control humano, privacidad y prosperidad) y de impacto de largo plazo (precaución, responsabilidad, mitigación de riesgos, control del aprendizaje automático y promoción del bien común).

¹⁰ The New York City Council: <http://legistar.council.nyc.gov/LegislationDetail.aspx?ID=3137815&GUID=437A6A6D-62E1-47E2-9C42-461253F9C6D0>



Por su parte la IEEE, una de las mayores organizaciones de profesionales tecnológicos del mundo, publicó en 2016 un documento titulado “Ethically Aligned Design” (IEEE, 2016) sobre la importancia de la ética en el desarrollo de la inteligencia

artificial. Este documento fue realizado con la colaboración de 250 expertos de todo el mundo y tiene como principios generales la integración de los derechos humanos en los avances tecnológicos y la priorización del beneficio máximo para la humanidad y el ámbito natural. Una de las secciones del documento se centra en las metodologías para guiar la investigación y el diseño ético. De todos los documentos existentes hasta el momento, este es el que mejor recoge los principios tradicionales de la robótica responsable, centrados en la importancia de la agencia humana.

Otra organización estadounidense, **AI Now, publicó en 2018 una guía para el uso de algoritmos por parte de las administraciones públicas (AI Now, 2018).**

Allí se insta a estas últimas a que fomenten la transparencia de sus sistemas de decisión automática y la participación en su diseño, a que desarrollen procedimientos para el resarcimiento de las personas afectadas injustamente por una decisión algorítmica, y a que realicen tareas de autoevaluación,

aunque también se les sugiere apoyarse en entes externos independientes que puedan realizar auditorías. En el mismo sentido, la Partnership on AI –un grupo de organizaciones y multinacionales entre las cuales figuran Google, Apple y Amazon– se propone investigar, debatir y compartir conocimientos sobre la inteligencia artificial y su impacto social¹¹.

En las universidades también se han elaborado guías y marcos de referencia para ayudar a los actores que trabajan con algoritmos.

GovEx (en la Universidad Johns Hopkins), la ciudad de San Francisco, Harvard DataSmart y Data Community DC se unieron para proponer una innovadora caja de herramientas –innovadora en el sentido de que exige que se tengan conocimientos técnicos sobre bases de datos para su implementación– que permita a los gobiernos y a otros actores hacer frente a los riesgos éticos de los algoritmos¹². Por su parte, la Universidad de Chicago cuenta con Aequitas, una herramienta de código abierto para auditar algoritmos que permite medir el desempeño de las predicciones y detectar sesgos que puedan afectar a grupos seleccionados (mujeres, minorías étnicas o sexuales, etc.)¹³.

Además de los gobiernos, algunas organizaciones y empresas también han publicado guías y marcos de referencia.

Por ejemplo, Accenture –una de las mayores empresas consultoras del mundo– publicó

¹¹ Partnership on AI: <https://www.partnershiponai.org/>

¹² Ethicstoolkit.ai.: <http://ethicstoolkit.ai/>

¹³ University of Chicago: <https://dsapp.uchicago.edu/aequitas/>

en 2016 un informe sobre la ética de la gestión de datos, y un “libro blanco” con 12 principios afines (Accenture, 2016). El primer principio de este libro blanco establece que el respeto a las personas cuyos datos se van a usar debe ser la prioridad de los profesionales a cargo. Cuando se recolectan datos también es necesario tener en cuenta sus usos secundarios y conocer su origen (la historia, el contexto y los mecanismos). Tal y como lo establecen el reglamento europeo y la reglamentación de las Naciones Unidas, la seguridad de los datos debe ser la adecuada en cada caso. El libro blanco contiene además los principios de minimización de datos y no discriminación, también incluidos en el RGPD. El concepto de transparencia aplica tanto a los métodos de análisis, como al diseño de las prácticas y a las cualificaciones y límites de los expertos. Finalmente, Accenture prescribe una buena gobernanza de los datos y una revisión ética interna y externa de las prácticas. Si bien la mayoría de los principios allí enunciados ya están incluidos en las leyes, cabe notar que Accenture pone más énfasis en la transparencia que el marco jurídico actual (Accenture, 2016).

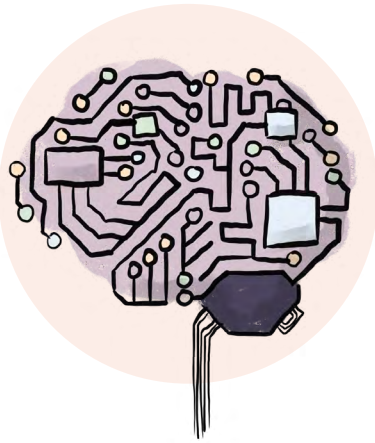
Empresas como Ernst & Young (2017) proponen las leyes existentes como el estándar mínimo, aunque además insisten en la importancia de educar a los usuarios de datos sobre temas de ética. A diferencia de las guías y marcos precedentes, centrados en la rendición de cuentas de los usuarios de los datos, la organización MyData (2018) ha hecho una “declaración de principios MyData” en la que se exige conferir más poder a las personas cuyos datos van a ser utilizados.

Otras empresas relevantes como Microsoft, IBM o Google han formulado recientemente principios éticos, aunque más enfocados en la inteligencia artificial, como se verá en el próximo apartado.

Apple se rige por un código de ética y responsabilidad social desde 2013 que poco tiene que ver con temas de datos¹⁴, mientras que la empresa Axon (especializada en tecnología de armamentos y de seguridad) cuenta con un consejo asesor ético que se reúne periódicamente para valorar el impacto ético de los productos que desarrolla, sobre todo los relacionados con el reconocimiento facial¹⁵.

¹⁴ Apple Inc.: <http://www.appleinc.blogspot.com/2013/09/ethics-and-social-responsibility.html>

¹⁵ The Verge: <https://www.theverge.com/2018/4/26/17285034/axon-ai-ethics-board-facial-recognition-racial-bias>



Inteligencia artificial y algoritmos

En el futuro, seguramente el ámbito donde la gestión ética de los datos –aunada a temas más amplios de la misma índole como la toma de decisiones automáticas– cobrará cada vez mayor relevancia será el de la inteligencia artificial (IA) y el uso de algoritmos. A este respecto se están desarrollando múltiples proyectos e iniciativas en todo el mundo tanto por parte de los gobiernos, como de las empresas, y últimamente de los organismos internacionales. Naciones Unidas ha lanzado diferentes iniciativas relacionadas con la IA, y existen acuerdos sobre estos temas entre los Emiratos Árabes Unidos e India, entre Francia y Canadá, y entre los miembros del G7. Paralelamente a estos avances en el campo de la IA ha surgido una preocupación cada vez mayor por el impacto ético que puedan tener estos sistemas, como se discutió en la sección anterior. Varias instituciones han desarrollado recientemente guías y conformando equipos de trabajo alrededor de estos temas.

La Unión Europea cuenta desde 1991 con un grupo de trabajo sobre ética en la ciencia y las nuevas tecnologías (European Group on

Ethics in Science and New Technologies),¹⁶ cuyo mandato fue renovado en 2016. Además, desde mediados de 2018 existe un grupo de trabajo sobre IA¹⁷ encargado de crear una guía de ética específica con base en principios como la dignidad y autonomía humanas, la responsabilidad en la investigación, la protección de la justicia, la igualdad y la solidaridad. La guía también se enfoca en la necesidad de promover un debate público que derive en acciones específicas, así como de garantizar el estado de derecho y la seguridad e integridad mental y física en la interacción máquina-humano, la protección de los datos y de la privacidad, y la sostenibilidad del ecosistema humano-máquina-medio ambiente. Además, la Convención 108 de 1981 del Consejo de Europa sobre la protección de los individuos en el contexto del procesamiento de datos personales fue reformada en 2018 para actualizarla y relacionarla con los retos digitales actuales, utilizando como base los principios del RGPD y de la OCDE.

Estos esfuerzos se suman a los preceptos ya establecidos en el RGPD, organismo pionero en la regulación de la toma de decisiones a través de medios automatizados. Específicamente explicita el derecho de los individuos a no ser sometidos a una decisión basada puramente en un proceso automatizado si esta ha de tener efectos legales o significativos en la vida de aquellos. Igualmente establece el derecho a la explicabilidad, al cual se aludió anteriormente

¹⁶ European Commission: <https://ec.europa.eu/research/ege/index.cfm?pg=about>

¹⁷ European Commission: <https://ec.europa.eu/digital-single-market/en/news/commission-appoints-expert-group-ai-and-launches-european-ai-alliance>

y que exige que cuando se produzca una decisión algorítmica, la persona afectada tenga acceso a “información significativa sobre la lógica aplicada, así como sobre la importancia y las consecuencias previstas de dicho tratamiento para el interesado” (RGPD, art. 15, 1 h).

Más de 20 países cuentan con estrategias

nacionales de inteligencia artificial. Australia, por ejemplo, se ha comprometido a invertir USD 30 millones en inteligencia artificial y aprendizaje automatizado o *machine learning* (Pearce, 2018). Si bien es cierto que todavía no se ha elaborado un marco ético, el gobierno se ha comprometido a hacerlo en un plazo razonable. Francia, por su parte, ha formulado una estrategia nacional para el desarrollo de la inteligencia artificial, a la que se suman el Informe Villani (2018) y otro publicado por la autoridad de protección de datos francesa (Commission Nationale Informatique et Libertés) (CNIL, 2017) sobre sus implicaciones éticas y políticas. El informe Villani y el del CNIL conforman en su conjunto un marco general de recomendaciones éticas centradas en la necesidad de mejorar la comprensión actual de los algoritmos a través de: (i) apoyo a la investigación sobre temas de rendición de cuentas, (ii) formación del personal a cargo, (iii) elaboración de estudios de impacto y (iv) conducción de auditorías algorítmicas para identificar y mitigar posibles discriminaciones. Asimismo, se señala la necesidad de considerar los temas éticos desde la etapa de diseño.

El Reino Unido cuenta con un documento publicado por la Cámara de los Comunes centrado en el uso de algoritmos en procesos de toma de decisiones (House of Commons,

Science and Technology Committee, 2018). Allí se abordan cuestiones relacionadas con la compartición de datos, la discriminación producida por el uso de algoritmos, la transparencia y la regulación algorítmica. Entre las principales recomendaciones figuran las siguientes:

- ✓ Crear un organismo independiente dedicado al **monitoreo de los temas relacionados con algoritmos** que ayude a encontrar un equilibrio que permita innovar sin erosionar la confianza pública¹⁸.
- ✓ Mejorar la transparencia en las decisiones algorítmicas tomadas por el gobierno para fomentar la participación ciudadana.
- ✓ Mejorar la transparencia y el valor público en los algoritmos de uso público desarrollados por empresas privadas.
- ✓ Expedir **medidas para evitar la discriminación algorítmica**, como por ejemplo incrementar la calidad de los datos utilizados para entrenar los algoritmos (training datasets), asegurar la diversidad dentro de los equipos que desarrollan algoritmos, y explorar formas de control y regulación (auditorías y certificaciones, por ejemplo).
- ✓ Garantizar el derecho a la explicabilidad de los algoritmos que afecten significativamente a las personas y el resarcimiento en caso de errores o fallas.

¹⁸ En 2018 se creó precisamente el Centre for Data Ethics and Innovation.

- ✓ Hacer esfuerzos encaminados a que en las **legislaciones nacionales se llenen los vacíos legales presentes en el RGPD**, específicamente con base en estudios de impacto algorítmico.
- ✓ Insistir en la **importancia de formar al personal relevante** y de **informar al público**.

El equivalente estadounidense es The National Artificial Intelligence Research and Development Strategic Plan (NTRAD, 2016). En este documento se enuncian los criterios generales en función de los cuales se pueden tomar decisiones relativas a la financiación de proyectos de investigación en el ámbito de la IA, tanto dentro de la estructura gubernamental como por fuera de ella. Allí se establecen siete prioridades, entre las cuales figuran cuestiones como la comprensión de los desafíos legales, éticos y sociales que entraña la inteligencia artificial para que el desarrollo tecnológico no se produzca en detrimento de fines sociales y políticos, y la seguridad de los sistemas de IA. Al igual que en los casos de Francia y el Reino Unido, se insiste en la importancia de los datos de entrenamiento como elemento clave para el correcto funcionamiento de los algoritmos y en la necesidad de elaborar estándares adecuados.

Un aspecto interesante que se discute en este documento tiene que ver con el carácter multidisciplinario de los equipos que investigan estos asuntos. Siguiendo el ejemplo de la estrategia británica en materia de diversidad, en Estados Unidos se propone la conformación de equipos transversales

con expertos de varias disciplinas. Con ello se busca abordar las diferentes cuestiones de forma holística y atendiendo a los problemas relacionados con la calidad de los datos y sus sesgos, y con la explicabilidad y transparencia de los sistemas.

En el caso de México, el documento relevante fue desarrollado por la embajada británica en ese país y por dos organizaciones independientes (British Embassy in Mexico, Oxford Insights y Cmind, 2018) con base en 68 entrevistas a expertos locales sobre las posibilidades y retos que plantea la inteligencia artificial en su país. En el informe se aborda una gama variada de temas, y en el ámbito ético se hacen las siguientes recomendaciones generales:

- ✓ **Crear una agencia independiente** que aborde las cuestiones relacionadas con la **inteligencia artificial** tomando como referencia el Instituto Nacional para la Ciencia de Datos y la Inteligencia Artificial de Gran Bretaña.
- ✓ **Asignar un papel preponderante al sector público y al gobierno** en estos temas.
- ✓ **Proteger la privacidad** de los individuos.
- ✓ **Crear un comité ético** sobre inteligencia artificial.

Un punto interesante que se señala en este documento es que, de desarrollar una estrategia propia en relación con la IA,

México sería el primer país latinoamericano en hacerlo. Esto pone de manifiesto tanto la falta de proactividad de los gobiernos de la región en este ámbito como la oportunidad que esto representa.

Canadá es otro de los países que ha decidido abordar la cuestión de la inteligencia artificial de forma proactiva. Prueba de ello es la formulación de la Pan-Canadian Artificial Intelligence Strategy (CIFAR, 2018), a cargo del gobierno. Una de las cuatro metas que allí se proponen es asumir el liderazgo global en el pensamiento relativo a cuestiones económicas, éticas, políticas y legales de la tecnología de IA. Un ejemplo de la voluntad del gobierno canadiense de aprovechar plenamente el uso de los algoritmos y la inteligencia artificial es el desarrollo de una herramienta en línea cuyo propósito es facilitar la identificación de riesgos relativos a los algoritmos a aquellas organizaciones que los utilizan (Government of Canada, 2018).

Otros gobiernos se han propuesto lanzar iniciativas similares, aunque sin éxito hasta el momento. A modo de ejemplo, el gobierno español asignó a un comité de expertos la tarea de crear un documento que abordase la cuestión ética en los ámbitos de los macrodatos y de la IA. Sin embargo, este “libro blanco” (Ministerio de Industria, Comercio y Turismo de España, 2017) no se ha publicado aún y el trabajo parece haberse interrumpido como consecuencia de un reciente cambio de gobierno.

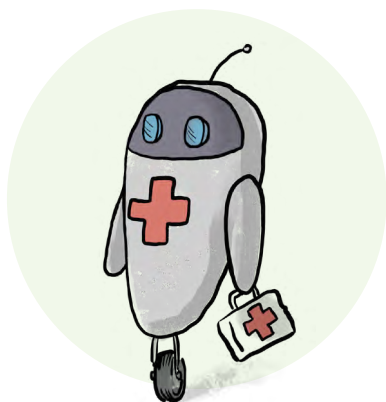
Tal y como se indicó anteriormente, en el sector privado Google y Microsoft cuentan con códigos éticos, mientras que Facebook tiene un equipo dedicado a la ética de

la inteligencia artificial, pero no existe información sobre sus integrantes, principios o resultados.

El código ético de Google se centra en los objetivos de las tecnologías a desarrollar, y plantea la necesidad de que estas produzcan beneficios sociales, no creen y/o refuercen sesgos, sean seguras, permitan rendir cuentas, incorporen elementos de privacidad desde el diseño, se basen en ciencia sólida y limiten sus usos a las aplicaciones beneficiosas. Asimismo, la empresa se compromete a no desarrollar aplicaciones dañinas, armas, tecnologías de vigilancia o tecnologías contrarias a la ley y a los derechos humanos¹⁹. Por su parte, Microsoft se guía por seis principios éticos: (i) equidad y justicia, (ii) confiabilidad, (iii) seguridad y privacidad, (iv) inclusión, (v) transparencia y (vi) rendición de cuentas. Esto se complementa con una política específica para diseñadores de inteligencia artificial conversacional (bots) en la que se les solicita articular su propósito; ser transparentes; incorporar especificidades culturales; prevenir el mal uso de la IA, y aceptar la responsabilidad por su funcionamiento e impacto. Asimismo, se les insta a crear sistemas en los que se pueda confiar y que promuevan la equidad, la justicia y la privacidad, así como la seguridad de los datos y el acceso a ellos²⁰.

¹⁹ Google: <https://www.blog.google/technology/ai/ai-principles/>

²⁰ Microsoft: <https://www.microsoft.com/en-us/ai/our-approach-to-ai>



Marcos de gestión ética sectoriales

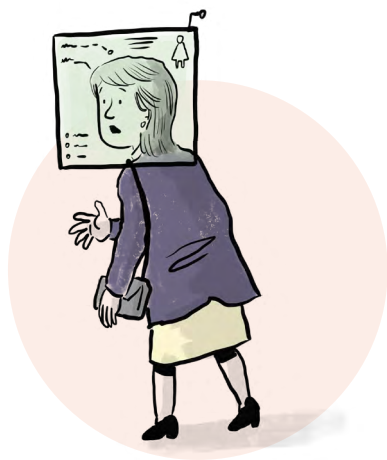
Sector salud y datos clínicos

La salud, un ámbito especialmente delicado y por ende muy regulado, es relevante en el contexto de este informe por dos motivos.

El primero es que **se trata de un área en la cual se lleva años lidiando con cuestiones como el consentimiento informado y la necesidad de abordar dilemas éticos** (de priorización, etc.) en el desempeño de la profesión, por lo que existen importantes enseñanzas provenientes de la praxis médica. El segundo es que **la salud es un sector altamente tecnologizado en el que el uso de dispositivos de monitoreo individual es muy anterior a los móviles**, y donde además se avanza cada vez más hacia un enfoque holístico en el que la disponibilidad de datos clínicos y de otra índole puede contribuir a avanzar la medicina a nivel global y a mejorar la atención personalizada de los pacientes.

En el ámbito internacional existen diferentes guías y marcos de referencia para asegurar un procedimiento ético en la investigación médica. Tales son los casos de las normas

de cooperación de la OCDE de 2018 (Knoppers y Thorogood, 2017) o la política 0070 de la Agencia Europea Médica sobre publicación de datos clínicos. La Organización Mundial de la Salud (OMS) cuenta con una página web dedicada a informar y educar a los profesionales sobre cómo integrar la ética en el tratamiento de brotes de enfermedades infecciosas, en la investigación y en las prácticas de monitoreo que usen datos de salud abiertos. La página incluye publicaciones sobre ética, listas de requerimientos, cursos de aprendizaje, recursos de otras instituciones, modelos e infografías. Entre las publicaciones está la “Guía de la OMS sobre los desafíos éticos en la vigilancia de la salud”, en la cual se incluyen 17 directrices para asegurar que esta se realice de manera ética. Tales directrices cubren temas de transparencia, fines legítimos y definidos para la obtención de datos en el ámbito de la salud pública, su pertinencia, actualidad y adecuación, así como su protección segura. Asimismo, se señala la necesidad de velar por la seguridad de las personas afectadas en relación con la protección del uso de sus datos por fuera de los sistemas de salud pública. Estos principios figuran en muchas de las guías sobre este tema, en especial aquellas en las que se abordan asuntos de monitoreo y vigilancia de la salud.



Seguridad ciudadana: reconocimiento facial

Entre las tecnologías que más polémica han causado en relación con el uso de inteligencia artificial está la del reconocimiento facial.

En el Reino Unido la polémica al respecto surgió a raíz de la utilización de dichos sistemas por parte de la policía británica, que los ha ensayado durante eventos masivos (Brandom, 2018). Una de las cuestiones que han puesto de manifiesto los defensores de los derechos a la privacidad es el hecho de que cuando la policía utiliza sistemas de reconocimiento facial está sometiendo a una gran cantidad de personas a un examen de atributos altamente confidenciales sin requerir su consentimiento (Portal, 2018).

En China la policía también ha utilizado esta tecnología, que ha sido incorporada a las gafas de sol (Vincent, 2018) y desplegada de forma masiva en espacios públicos como parte del esfuerzo del gobierno por instaurar un sistema de calificación del comportamiento social (social scoring). Evidentemente esto ha provocado inquietud

entre los grupos de defensa de la privacidad, que alertan sobre el potencial discriminatorio de estos sistemas y su impacto tanto social como en los derechos fundamentales.

Independientemente de dónde se empleen, estas tecnologías siempre causan grandes controversias. De hecho, su área de aplicación no se limita a las cuestiones de orden público, pues el reconocimiento facial también se usa en el ámbito comercial. Precisamente, para dotar al sector privado de buenas prácticas al respecto, la National Telecommunications and Information Administration (NTIA) de Estados Unidos publicó las Privacy Best Practice Recommendations for Commercial Facial Recognition Use (NTIA, 2016). El documento es breve y en él se reconoce desde el primer momento que el reconocimiento facial es una tecnología que se aplica en múltiples mercados. Por ese motivo ni siquiera pretende prescribir prácticas específicas, sino realizar más bien un mapa general que se adapte a contextos diversos. En este sentido, se insiste en la necesidad de contar con mecanismos de transparencia, fomentar las buenas prácticas en el manejo de los datos, y contar con criterios de limitación de uso y compartición de los mismos. Igualmente se abordan otros temas vinculados a la seguridad, la calidad de los datos y el derecho al resarcimiento.

Finalmente, cabe mencionar que en la Unión Europea existe una opinión consignada en el Article 29 Working Party (2012) relativa al reconocimiento facial en los servicios en línea y móviles que puede ser de interés para resaltar algunos de los principales problemas éticos –y sobre todo legales– que estos sistemas plantean. **En ese documento se ofrecen recomendaciones y buenas prácticas centradas en la necesidad de obtener siempre el consentimiento de la persona a la que corresponden las imágenes;** de que las imágenes se utilicen solo y exclusivamente para los propósitos para los que fueron obtenidas; de garantizar que las plantillas generadas por un sistema de reconocimiento facial no contengan más datos de los necesarios para el fin o los fines previstos, y de insistir no solo en la seguridad de los datos en línea, sino también en la seguridad física del servidor donde se los almacena. Finalmente se insiste en que se debe garantizar el acceso a los datos por parte de las personas afectadas.



Gestión ética de datos en la academia y la investigación

La Dirección General de Investigación e Innovación de la Comisión Europea, en el contexto de su programa de financiación de la actividad de investigación y desarrollo industrial en los estados miembros, cuenta con un robusto sistema de evaluación ética de todos los proyectos que reciben ayudas públicas. Por su impacto en el ecosistema de investigación e innovación, este es un proceso pionero en la UE, y ha obligado a muchas organizaciones y universidades a mejorar sus procesos de certificación ética, extendiéndola a áreas como las ciencias sociales o los proyectos de ingeniería. Las organizaciones que aspiran a conseguir financiación deben abordar 11 áreas de riesgo ético y dar explicaciones y garantías de monitoreo ético efectivo de los proyectos más delicados. Entre las áreas cubiertas figuran el trabajo con humanos (personas que participan en investigaciones, en pilotos, en actividades monitoreadas con fines de investigación, en entrevistas o grupos focales, etc.), la protección de los datos personales (tanto aquellos proporcionados directamente por los participantes como los recabados

en línea o reutilizados a partir de otras investigaciones), y el riesgo de que a los resultados de una investigación no se les dé un uso legítimo.

La Comisión Europea exige que los receptores de financiación cuenten con la validación ética de algún organismo relevante en su país, y también requiere la elaboración y validación de documentos de consentimiento para todas las actividades con humanos y con datos personales. En los casos de proyectos en los cuales se planea trabajar con menores o con datos sensibles, o aquellos para los cuales es necesario realizar un procesamiento masivo de datos, la Comisión exige garantías adicionales.

Quizás lo más destacable del modelo de la Comisión Europea es que se exige a los solicitantes demostrar un buen manejo de los riesgos éticos desde la fase misma de propuesta. Cuando un proyecto recibe la aprobación científica para su financiación, este pasa a un panel de expertos éticos independientes seleccionados por la Comisión que se encarga de valorar el cumplimiento de los requisitos y de identificar los riesgos. Si los estándares no se cumplen durante la formulación de un proyecto, o si este no cuenta con apoyo ético específico y con las certificaciones independientes del caso, es posible que no se inicie o que se cancele su financiación.

Como se indicó anteriormente, la incorporación de este riguroso proceso de evaluación ética en el ámbito europeo por parte del mayor organismo financiador de la región ha llevado a muchas universidades a adoptar medidas similares y a extender

el trabajo de sus equipos de bioética a nuevas disciplinas. Uno de los países más avanzados en este sentido es el Reino Unido. La universidad de Nottingham, por ejemplo, cuenta desde 2012 con un código ético de investigación que incluye temas relacionados con las responsabilidades de los investigadores, el uso de datos personales y de investigación, la propiedad intelectual, los conflictos de interés, y la necesidad de contar con un comité ético interno, entre otros. La universidad sigue los principios éticos Nolan, publicados por el Comité sobre los Estándares en la Vida Pública en 1994. En cambio, la universidad de Bath, también en el Reino Unido, sigue los principios del Acuerdo de Apoyo a la Integridad de la Investigación (2012), desarrollado por Universities UK, la organización representativa de las universidades del país. Las universidades de Manchester y Reading, por su parte, siguen sus propios estándares éticos.

En un estudio reciente publicado en Research Ethics (Vadeboncoeur et al., 2016) se llega a la conclusión de que, pese a que todas las universidades británicas han desarrollado marcos de estándares éticos para llevar a cabo sus investigaciones, estos varían de manera significativa y adolecen de inconsistencias relacionadas con los protocolos a seguir. No obstante, lo anterior, **la inclusión de la evaluación ética como un requerimiento obligatorio para el desarrollo de proyectos que manejan datos personales supone un avance importante en las garantías existentes en este sentido y sitúa al Reino Unido como ejemplo de buena práctica en el ámbito mundial.** Es cierto que algunos de los escándalos recientes

relacionados con el uso ilegítimo de datos por parte de investigadores universitarios –como el famoso caso de Facebook– Cambridge Analytica durante las elecciones presidenciales de Estados Unidos en 2016–, y el uso de datos personales con fines de manipulación política sin el consentimiento de los ciudadanos,²¹ han desvelado el hecho de que estos investigadores mintieron sobre la obtención de la aprobación ética y que llevaron a cabo sus investigaciones de todas maneras al margen de la universidad. Sin embargo, esto no demerita el hecho de que, incluso con las deficiencias notadas, el sistema sí ha conseguido crear una cultura de garantías en relación con el impacto y la ética de los proyectos que utilizan datos personales.

Las iniciativas en este sentido no se limitan a Europa. Otros países también han optado por desarrollar marcos de referencia específicos para promocionar la innovación responsable.

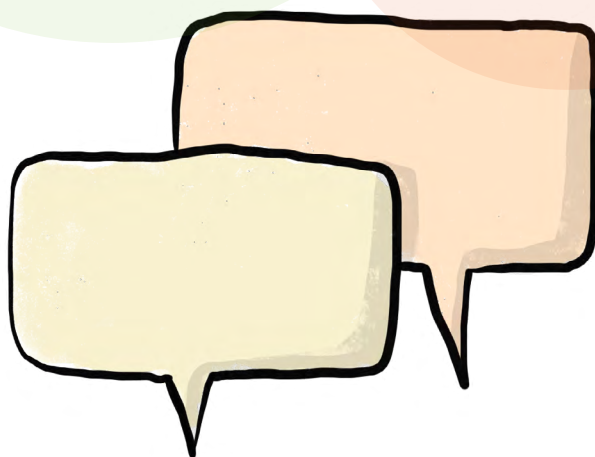
México, por ejemplo, cuenta con una lista de control para ayudar a los científicos a garantizar la ética en sus protocolos, y también con una “Guía nacional para la integración y el funcionamiento de los comités de ética en investigación”.

Australia (ANDS, 2018b; ANDS, 2018c) y Filipinas (Philippine Health Research Ethics Board, 2017) también han elaborado guías de buenas prácticas, específicamente en el ámbito de la investigación médica. Todos estos esfuerzos tienen en común que atribuyen una importancia muy significativa al consentimiento, la transparencia y la calidad de las investigaciones.

Tal como se observó en la sección anterior, a pesar de la abundancia de principios y marcos de referencia, estos no parecen constituir un corpus homogéneo ni detallar prácticas concretas que puedan ser útiles para otros organismos que lidien con cuestiones parecidas. Más allá del RGPD, que sí surge como estándar de referencia global –aunque aún con muchos interrogantes en relación con las prácticas concretas debido a su corto periodo de validez (entró en vigor en mayo de 2018) –, aquellas organizaciones que quieran desarrollar políticas de datos responsables y éticas deberán lidiar con esta diversidad y a su vez procurar llenar los vacíos existentes. No se dispone, por ejemplo, de marcos de referencia éticos en temas tan cruciales y actuales en la sociedad de los datos como la interacción humano-máquina, los desafíos de los sistemas predictivos o los riesgos y elementos a tener en cuenta en la implementación de proyectos basados en datos biométricos.

²¹ Digital Watch Observatory: <https://dig.watch/trends/cambridge-analytica>

4 Propuesta de criterios para una gestión ética de datos en el sector público



Para abordar estas carencias, en la siguiente sección se describen las enseñanzas más valiosas que han dejado los casos menos exitosos de sistemas y programas de datos debido a una valoración incorrecta de los riesgos éticos. El resultado es una aproximación práctica al desarrollo de programas de datos desde la responsabilidad, la transparencia, la eficiencia y el compromiso con los derechos y libertades de la ciudadanía. Los últimos detalles de cualquier proyecto que incorpore datos o algoritmos a la valoración o prestación de servicios en el sector social dependerán evidentemente de la existencia de marcos jurídicos, culturales y sociotécnicos específicos imposibles de cubrir en un documento general como el que aquí se ofrece. Aun así, el marco que se propone a continuación permite que

los promotores de cualquier proyecto se formulen en cada momento las preguntas adecuadas para asegurar la correcta conceptualización e implementación de políticas de datos en el sector social.

A la luz de los riesgos éticos a los que se enfrentan los gobiernos a la hora de usar datos para la toma de decisiones de política pública identificados en la segunda sección (privacidad, opacidad y discriminación), y de los marcos de referencia que están adoptando las instituciones públicas, empresas privadas y organizaciones sin fines de lucro descritos en la tercera sección, a continuación se define una propuesta de marco de criterios para la gestión ética de datos. Esta propuesta está dirigida a los organismos públicos en América Latina y el Caribe.

En esta propuesta se supone que **la gestión ética de datos se debe realizar a lo largo de todo su ciclo de vida**, definidas desde la creación y captura, pasando por el almacenamiento, la transmisión y el análisis, hasta el archivo o eliminación (Faundeen y Hutchison, 2017). A continuación, se ofrece una breve explicación de este ciclo, **dividido en cinco etapas relevantes:**

I. Recolección: En esta etapa el organismo público obtiene los datos a través de distintas formas, sean estas automatizadas o manuales. Esto se puede hacer, por ejemplo, acopiando nueva información mediante encuestas/estudios, extrayendo datos administrativos de sistemas existentes, comprando bases de datos u obteniéndolas de otro organismo porque sus facultades legales así se lo permiten o porque firmó un acuerdo de colaboración²².

II. Almacenamiento: El objetivo de esta etapa es mantener los datos protegidos de manera segura y acceder a ellos cuando sea necesario. Esto incluye un proceso de respaldo de estos para evitar su pérdida en caso de fallas tecnológicas o humanas, virus o acceso ilegal. También considera la seguridad física de la red, de los sistemas y de los archivos mediante perfiles de acceso diferenciados.

III. Análisis: En esta etapa se contemplan las actividades de exploración y evaluación de los datos para extraer información útil destinada a la toma de decisiones. A partir de ello se evalúan hipótesis y se llega a conclusiones con base en análisis estadísticos, visualizaciones, análisis espaciales y modelamiento, entre otras actividades²³.

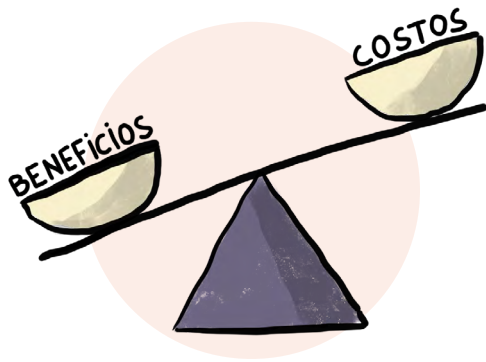
IV. Compartición: Los datos se pueden compartir dentro de la misma institución, con entidades externas como universidades u organismos sin fines de lucro, o con proveedores como las empresas; también se pueden ofrecer a la comunidad en formato de datos abiertos. En esta etapa son claves la anonimización, la privacidad diferencial y la transparencia.

V. Archivo/eliminación: Esta etapa guarda algunas similitudes con la de almacenamiento, en el sentido de proteger los datos archivados y poder acceder a ellos. También contempla la definición de los periodos de retención y los procedimientos para su eliminación.

Basado en lo anterior, se presentan los principales criterios que configuran un marco de referencia para la gestión ética de datos en el sector público y que deben permear todo el ciclo de vida de los datos.

²² USGS: <https://www.usgs.gov/products/data-and-tools/data-management/acquire>

²³ USGS: https://www.usgs.gov/products/data-and-tools/data-management/analyze?qt-science_support_page_related_con=0#qt-science_support_page_related_con



Creación del valor público

En varios de los marcos éticos existentes se destaca la importancia de plantear claramente los beneficios que generará el proyecto²⁴. Esto es importante porque **una de las condiciones para que un proyecto sea ético es que los beneficios superen los costos**. Por ejemplo, en Nueva Zelanda se definieron las tres preguntas²⁵ a las que una organización debe responder para determinar si el uso de datos genera valor: (i) ¿para qué se van a usar los datos? (ii) ¿cuáles son los beneficios y quién se beneficiará? y (iii) ¿quién estará usando los datos? Para determinar el valor público del uso de datos se deberá proceder de la siguiente manera:

²⁴ A Path To Social License:Trusted Data Guidelines, NZ. <https://trusteddata.co.nz/wp-content/uploads/2017/08/Summary-Guidelines.pdf> Data Ethics Framework UK [<https://www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework>]

²⁵ Data Futures Partnership: <https://trusteddata.co.nz/organisations/>

a. Establecer la línea de referencia de desempeño del proceso:

Es importante conocer la situación actual a fin de determinar si esta mejora con el uso de datos. Para ello se requiere especificar la naturaleza del problema. Por ejemplo, cuando se quiso implementar un modelo predictivo de riesgo de vulnerabilidad infantil en el condado de Allegheny en Pensilvania, Estados Unidos, un análisis de las respuestas a denuncias por maltrato entre 2010 y 2016 demostró que se podían hacer mejoras sustanciales. Esto por cuanto se pudo establecer que un 53% de los niños cuyo maltrato se denunció, pero que no fueron visitados por funcionarios de la agencia a cargo, recibieron nuevos maltratos en los siguientes dos años (Chouldechova et al., 2018). En otro caso, en Chicago una red de centros de salud comunitaria decidió que era necesario mejorar los sistemas de detección temprana de la diabetes tipo II. El problema era que, con el sistema actual, solo a un 53% de la población que desarrolla diabetes dos años después de la consulta inicial se le realizan los exámenes de detección temprana²⁶. La diabetes puede generar serias complicaciones para la salud de una persona y es costosa para el sistema público, por lo que su prevención tiene un alto valor público.

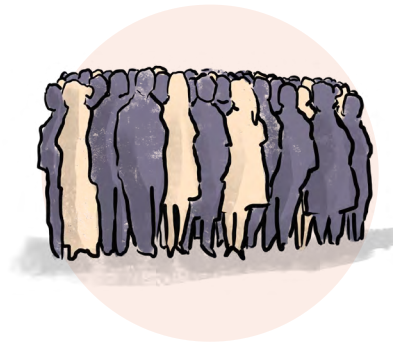
²⁶ The University of Chicago, Data Science for Social Good: <https://dssg.uchicago.edu/project/supporting-proactive-diabetes-screenings-to-improve-health-outcomes/> <https://www.youtube.com/watch?v=SFPVAj8Fefs>

b. Identificar los objetivos del proyecto, las acciones a realizar y el beneficio público que se generará:

Una vez que se conoce la situación de referencia se puede proceder a formular objetivos para el uso de los datos. En el caso de Pensilvania, por ejemplo, se trataba de aumentar el porcentaje de niños de alto riesgo visitados tras una primera denuncia, y reducir las visitas a niños cuyo riesgo de maltrato es bajo. Las acciones por realizar tienen que ver con las competencias legales y con las capacidades técnicas del organismo público; se trata de gestiones concretas que ocurrirán gracias al uso de datos. Para el Departamento de Servicios Humanos de Pensilvania, la acción correspondiente consiste en visitar a una familia tras una denuncia de alguien de su entorno para constatar en terreno si se registran o no señales de maltrato infantil. En el caso de la diabetes, la acción concreta consiste en realizar un examen de sangre al paciente para detectar problemas con el procesamiento de la glucosa. Por último, el beneficio público es el impacto positivo que se logrará con un buen funcionamiento del sistema: menos niños maltratados, personas más saludables, menos gasto público en salud, etc.

c. Diseñar indicadores de desempeño:

A partir de la línea de referencia y de los objetivos fijados, se requiere elaborar indicadores que sirvan para monitorear el progreso que se logra como resultado del proyecto después de su implementación. Así se podrá saber si se están cumpliendo las metas que se establecieron en un comienzo.



Identificación de las personas beneficiadas o afectadas y valoración del impacto

En cada proyecto en el que se pretenda usar datos personales **se debe tener claro cuáles son las personas que se verán beneficiadas o afectadas y cuáles las inequidades que se pueden producir**, prestando especial atención a los grupos minoritarios o vulnerables. Una vez identificados, los funcionarios a cargo del proyecto deberán determinar el impacto que podría tener el proyecto en tales grupos. Por ejemplo, en la provincia de Salta, Argentina, se desarrolló un modelo predictivo de embarazo adolescente a través del cual se identificó a 397 jóvenes con alto riesgo provenientes de barrios de menores ingresos (Ortiz Freuler e Iglesias, 2018). El Ministro de la Primera Infancia indicó que se habían entregado los resultados a las “áreas correspondientes” del gobierno para desarrollar un plan de acción con un “abordaje integral”²⁷ dirigido a las adolescentes, aunque no detalló cómo se informaba sobre el particular a las menores. Esto generó una serie de críticas por parte de organizaciones de la sociedad civil, las cuales denunciaron el riesgo de estigmatización o afectación de la privacidad de las jóvenes.

²⁷ UNO: https://uno.com.ar/tecnologia/como-funciona-el-sistema-para-predecir-embarazos-adolescentes-de-salta-04122018_rJxfqbraiM



Diagnóstico de datos

Con el fin de realizar un diagnóstico correcto de los datos requeridos para un determinado proyecto **es necesario mapear primero los ya disponibles y los deseables, para posteriormente identificar sesgos posibles.**

Con esto se obtendrá idealmente el conjunto de datos necesarios.

a. Mapear datos disponibles y deseables:

En su quehacer diario, los gobiernos generan un volumen muy significativo de datos administrativos. La cantidad y calidad de estos es relevante para los resultados del proyecto. Por eso en la etapa de planificación se requiere identificar todas las bases de datos que se podrían utilizar, tanto dentro como fuera del organismo público, y determinar si es necesario realizar alguna inversión que permita obtener mejores datos. Por ejemplo, se podría decidir que en el caso de un dato que se recolecta diariamente, pero que se registra semanalmente, se cambié la frecuencia de publicación de manera que se cuente con información más oportuna. Otras posibilidades serían capacitar a quienes ingresan datos a los sistemas, instalar sistemas de verificación automática de calidad, firmar acuerdos para obtener datos externos, etc.

b. Identificar sesgos: Las bases de datos adolecen de sesgos que pueden ser producto de decisiones humanas tendenciosas, información insuficiente, o de la falta de representatividad de algún segmento de la población. Para evitar el riesgo de discriminación descrito en la primera sección de este documento, en esta etapa se deben diagnosticar dichos sesgos para poder diseñar el proyecto con las respectivas medidas de mitigación.



Privacidad por defecto/diseño

La privacidad debe incorporarse en cada etapa del ciclo de vida de los datos, y es en la fase de planificación cuando se deben diseñar las acciones relevantes para lograrlo, a la luz de los estándares que hacen parte del reglamento general de protección de datos de la Unión Europea. **Si el proyecto contempla datos personales o sensibles, se deben incorporar las medidas para cumplir con la normativa nacional de protección de datos personales.** Entre los temas relevantes a considerar figuran los siguientes:

a. Consentimiento: Quien obtenga y gestione los datos deberá asegurarse siempre de haber informado a su propietario (el ciudadano), y de solicitar su consentimiento tantas veces como sea necesario si la finalidad del tratamiento cambia.

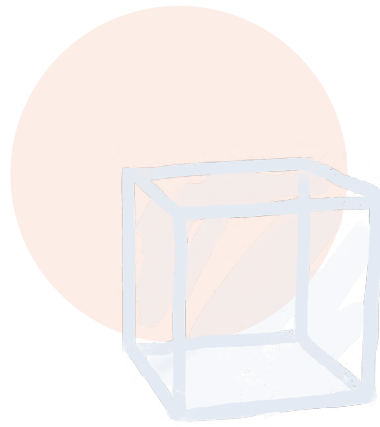
b. Anonimización/seudonimización:

Estas son las medidas técnicas que se deben tomar para asegurar que los datos personales no se atribuyan a una persona identificable. En el caso deseudonimización, se utilizan identificadores artificiales que reemplazan los atributos identificables de un dato.

c. Seguridad de la información: Esto implica considerar medidas, tales como políticas de respaldo (hacer una o varias copias que permitan recuperar la información en caso de pérdida) de la información, medidas de encriptación, accesos diferenciados según el perfil, y protección contra la piratería informática.

d. Estándares por defecto: Los datos personales se deben procesar automáticamente con los más altos estándares, sin que sea necesario tomar medidas adicionales para que eso suceda. Por ejemplo, en caso de que se cree un sistema con perfiles para cada usuario, como condición básica este sistema deberá restringir el acceso ilimitado de otros usuarios a estos perfiles (Zook, et al., 2017).

e. Datos abiertos: Si el proyecto contempla la disponibilidad de datos abiertos sobre comportamiento humano, no es suficiente con anonimizarlos para efectos de su publicación. Dada la posibilidad de reidentificación mediante el cruce con otras bases de datos públicas –por ejemplo, solicitando que algún especialista reidentifique la información–, se debe garantizar que esto no suceda.



Transparencia y rendición de cuentas

A continuación, se lista una serie de medidas para combatir la opacidad a la hora de usar los datos. Esto es particularmente relevante en un escenario de profunda desconfianza de la ciudadanía frente a la labor gubernamental.

a. Delimitar el papel de los datos en la toma de decisiones:

Los algoritmos pueden servir como insumo en la toma de decisiones, pero usarlos para reemplazar este proceso es cuestionable considerando que todos muestran tasas de falsos positivos y falsos negativos. La responsabilidad de las decisiones debe recaer en las personas y en las instituciones.

b. Diseñar estrategias de comunicación y participación de la ciudadanía y los grupos de interés:

En el escenario actual de poca confianza en el gobierno, se requiere integrar la participación de los grupos afectados por el proyecto desde un comienzo. Es importante considerar las aprensiones de dichos grupos y diseñar información clara y útil acerca

del proyecto, por ejemplo, en la forma de respuestas de fácil acceso a las preguntas frecuentes que la ciudadanía pueda tener sobre el mismo.

c. Diseñar medidas de mitigación de sesgos:

Una vez identificados los sesgos, se pueden hacer ajustes en el algoritmo o en las acciones a realizar. Esto con el fin de minimizarlos, o por lo menos de asegurar que no se contribuya a amplificar los que ya existen en la sociedad.

d. Diseñar mecanismos de rectificación y reparación de errores:

Dado que los sistemas pueden cometer errores, se requiere contemplar formas simples de corrección y reparación distintas a un tribunal de justicia. Una simple solicitud de rectificación de los datos debe considerarse en primera instancia.

e. Diseñar mecanismos de monitoreo interno:

La vigilancia del cumplimiento de los indicadores de desempeño se puede integrar a los procesos de control de gestión de la institución. Tras la ejecución del proyecto, este debería incorporarse al plan anual de auditorías de aquella para que se pueda monitorear tanto su impacto ético, como los sesgos en la predicción y ocurrencia de falsos positivos o negativos.

f. Planificar la realización de evaluaciones una vez implementado el proyecto:

Se recomienda obtener los recursos necesarios para llevar a cabo una evaluación de impacto por parte de una entidad independiente. Para ello se deben documentar tanto el proyecto mismo como los procesos del sistema (Zook, et al., 2017), lo cual permitirá realizar una rendición de cuentas efectiva. La realización de evaluaciones a cargo de entidades independientes asegura una mayor independencia de sus resultados, mientras que la incorporación de sus recomendaciones fortalecerá la gestión ética de los datos.

Identificación de buenas prácticas internacionales existentes

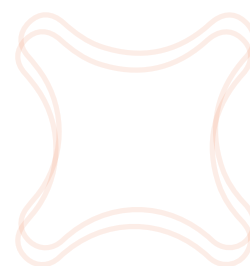
Considerando que el uso de estos avances tecnológicos para abordar los problemas del sector social –y en general aquellos de interés público– son relativamente recientes, es recomendable identificar a aquellos que ya estén desarrollando este tipo de iniciativas de manera exitosa para incorporar sus aprendizajes.

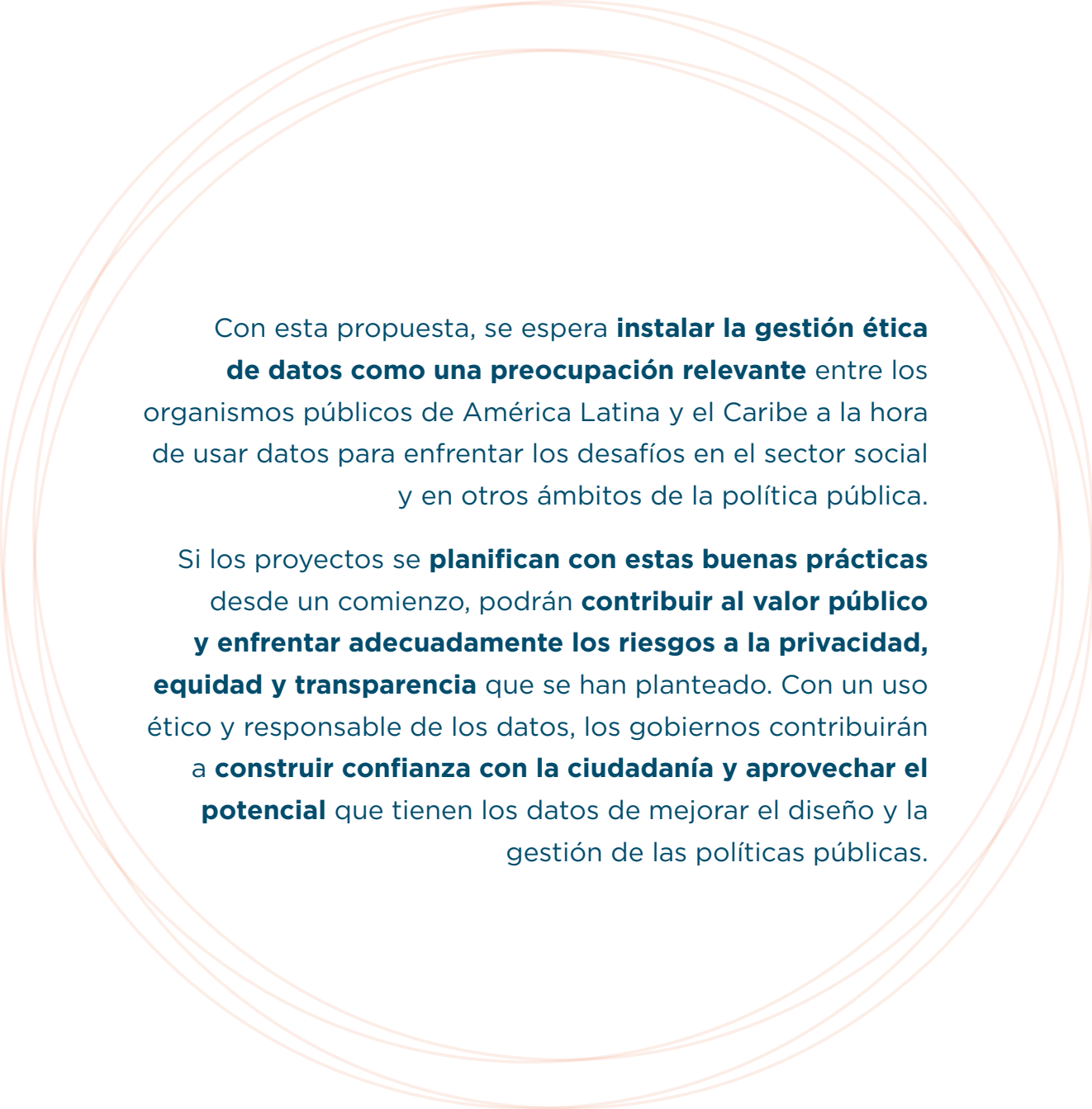
Diseño de gobernanza institucional y definición de capacidades necesarias

El uso de datos no debe ser relegado a las áreas de tecnología de la institución. Se requiere entonces definir una gobernanza del proyecto que incorpore a un equipo diverso y multidisciplinario en donde participen activamente los profesionales del área de gestión correspondiente, cuya experiencia y sensibilidades temáticas les permitan determinar los efectos que este pueda tener en los usuarios.

Diseño de pilotos y pruebas en pequeña escala previamente al despliegue del sistema

Una parte esencial de los procesos de innovación es la iteración, en la medida en esta permite incorporar aprendizajes y mejorar procesos. Para que así sea, es necesario ensayar el proyecto en pequeña escala con el fin de garantizar que su despliegue masivo se pueda realizar de la mejor manera posible.





Con esta propuesta, se espera **instalar la gestión ética de datos como una preocupación relevante** entre los organismos públicos de América Latina y el Caribe a la hora de usar datos para enfrentar los desafíos en el sector social y en otros ámbitos de la política pública.

Si los proyectos se **planifican con estas buenas prácticas** desde un comienzo, podrán **contribuir al valor público y enfrentar adecuadamente los riesgos a la privacidad, equidad y transparencia** que se han planteado. Con un uso ético y responsable de los datos, los gobiernos contribuirán a **construir confianza con la ciudadanía y aprovechar el potencial** que tienen los datos de mejorar el diseño y la gestión de las políticas públicas.

Miranda Ventura y el Gran Big Data



Conoce el cómic de esta publicación en
<https://publications.iadb.org/es>

Referencias

Sección 1

BBC News. 2008. Hacker Leaks 6m Chileans' Records. Obtenido de: <http://news.bbc.co.uk/2/hi/americas/7395295.stm>. Consultado el 10 de septiembre de 2018.

Feuer, A. 2013. The Mayor's Geek Squad. The New York Times. Obtenido de: <https://www.nytimes.com/2013/03/24/nyregion/mayor-bloombergs-geek-squad.html>. Consultado el 6 de septiembre de 2018.

Fortune. S.f. The Share of Female CEOs in the Fortune 500 Dropped by 25% in 2018. Obtenido en: <http://fortune.com/2018/05/21/women-fortune-500-2018/>. Consultado el 21 de enero de 2019.

Young and Verhulst. 2016 "Mexico's Mejora tu Escuela." Open Data Impact Case Studies. Obtenido de: <http://odimimpact.org/case-mexicos-mejora-tu-escuela.html> Consultado el 21 de enero de 2019.

Harvard Business School, Digital Initiative. 2016. Microsoft Helps Tacoma Public Schools Use Data Analytics to Predict At-Risk Students. Obtenido de: <https://rctom.hbs.org/submission/microsoft-helps-tacoma-public-schools-use-data-analytics-to-predict-at-risk-students/>. Consultado el 21 de enero de 2019.

Hijink, M. 2018. Algorithm Predicts Fraud Committed with Social Assistance. Nrc.Nl. Obtenido de: <https://www.nrc.nl/nieuws/2018/04/08/algorithm-voorspelt-wie-fraude-pleegt-bij-bijstandsuitkering-a1598669>. Consultado el 5 de septiembre de 2018.

Kim, G. H., S. Trimi y J. H. Chung. 2014. Big-Data Applications in the Government Sector. Communications of the ACM, 57(3), 78-85. Obtenido de: <http://www.academia.edu/download/41205301/0a85e5328669ac-06ba000000.pdf20160115-19908-1bna8d8.pdf>. Consultado el 8 de septiembre de 2018.

Ministerio de Hacienda de Chile. 2017a. Estudio de uso intensivo de datos en políticas públicas: análisis y recomendaciones estratégicas para la implementación de una política en base a la evidencia internacional. Obtenido de: <http://modernizacion.hacienda.cl/estudios/gestion-de-la-informacion-y-el-conocimiento/estudio-de-uso-intensivo-de-datos-en-politicas-publicas> Consultado el 5 de septiembre de 2018

-----, 2017b. Políticas públicas: análisis y recomendaciones estratégicas para la implementación de una política en base a la evidencia internacional. Obtenido de: <http://modernizacion.hacienda.cl/estudios/gestion-de-la-informacion-y-el-conocimiento/estudio-de-uso-intensivo-de-datos-en-politicas-publicas> Consultado el 8 de septiembre de 2018.

Muller, S. 2018. How Does TfL's Oyster Card Work? Alphr.Com. Obtenido de: <http://www.alphr.com/technology/1002164/how-does-an-oyster-card-work> . Consultado el 6 de septiembre de 2018.

NYC Center for Innovation through Data Intelligence. S.f. Predicting Homeless Shelter Entry - CIDI. Obtenido de: <https://www1.nyc.gov/site/cidi/projects/predicting-homeless-shelter-entry.page>. Consultado el 8 de septiembre de 2018.

SAS. S.f. Uncovering Social Service Fraud Saves Millions, Reinforces Public Trust. Obtenido de: https://www.sas.com/en_us/customers/la-county-dpss.html. Consultado el 8 de septiembre de 2018.

Singer, N. 2012. Mission Control, Built for Cities. The New York Times. Obtenido de: <https://www.nytimes.com/2012/03/04/business/ibm-takes-smarter-cities-concept-to-rio-de-janeiro.html>. Consultado el 6 de septiembre de 2018.

SmartSantanderRA. S.f. Santander Augmented Reality Application. Obtenido de: <http://www.smartsantander.eu/index.php/blog/item/174-smartsantanderra-santander-augmented-reality-application>. Consultado el 6 de septiembre de 2018.

Thayer, M. 2014. Using Data Analytics to Counter Fraud: New York Tax Case Study. Blog. Obtenido de: <https://www.govloop.com/community/blog/using-data-analytics-to-counter-fraud-new-york-state-tax-case-study/>. Consultado el 8 de septiembre de 2018.

Sección 2

Angwin, J., J. Larson, S. Mattu y L. Kirchner. 2016. ProPublica

Machine Bias. Obtenido de: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Consultado el 22 de enero de 2019.

Barocas, S. y A. Selbst. 2016. Big Data's Disparate Impact. *California Law Review* 104: 671-732.

Big Brother Watch. 2018. Face Off the Lawless Growth of Facial Recognition in UK Policing. Obtenido de: bigbrotherwatch.org.uk/wp-content/uploads/2018/05/Face-Off-final-digital-1.pdf. Consultado el 22 de enero de 2019.

Bordas, W. 2018a. Sur Parcoursup, seuls 48% des candidats ont accepté une proposition de formation. Obtenido de: http://etudiant.lefigaro.fr/article/sur-parcoursup-seuls-48-des-candidats-ont-accepte-une-proposition-de-formation_c3fead92-7f61-11e8-86d7-8173784e1d0f/. Consultado el 18 de septiembre de 2018.

------. 2018b. Parcoursup: «moins de 2500» candidats encore en attente, selon Frédérique Vidal. Obtenido de: http://etudiant.lefigaro.fr/article/parcoursup-moins-de-2500-candidats-encore-en-attente-selon-frederique-vidal_1191ec3c-b7f1-11e8-ad2c-a3da0be3bc08/. Consultado el 18 de septiembre de 2018.

Burrell, J. 2016. How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms. *Big Data & Society*. Obtenido de: <https://doi.org/10.1177/2053951715622512>. Consultado el 22 de enero de 2019.

CIPER (Centro de Investigación Periodística, Chile). 2016. Grave falla en la red del Minsal dejó expuesta información confidencial de pacientes. Obtenido de: <https://ciperchile.cl/2016/03/05/grave-falla-en-la-red-del-minsal-dejo-expuesta-informacion-confidencial-de-pacientes/>. Consultado el 22 de enero de 2019.

CNIL(Commission Nationale Informatique & Libertés). S.f. Règlement européen sur la protection des données personnelles. Obtenido en: https://www.cnil.fr/sites/default/files/atoms/files/pdf_6_etapes_interactifv2.pdf. Consultado el 22 de enero de 2019.

Derechos Digitales. 2016. Una filtración que pone en peligro la democracia mexicana. Obtenido de: <https://www.derechosdigitales.org/9907/filtran-datos-de-millones-de-electores-mexicanos/>. Consultado el 22 de enero de 2019.

El Observador. S.f. Microbuses de Olmué cuentan con tecnología de reconocimiento facial para la TNE. Obtenido de: <http://web.observador.cl/microbuses-de-olmue-cuentan-con-tecnologia-de-reconocimiento-facial-para-la-tne/>. Consultado el 22 de enero de 2019.

Grubb, B. 2017. Health Record Details Exposed as 'De-identification' of Data Fails. *The Sydney Morning Herald*. Obtenido de: <https://www.smh.com.au/technology/australians-health-records-unwittingly-exposed-20171218-p4yxt2.html>. Consultado el 22 de agosto de 2018.

Gymrek, M., A. McGuire, D. Golan, E. Halperin e Y. Erlich. 2013. Identifying Personal Genomes by Surname Inference. *Science* 339 (6117): 321-324.

Kroll, J., J. Huey, S. Barocas, E. Felten, J. Reidenberg, D. Robinson y H. Yu. 2017. Accountable Algorithms. *University of Pennsylvania Law Review* 165: 633-705.

Kupersmith, J. 2013. The Privacy Conundrum and Genomic Research: Re-identification and Other Concerns. *Health Affairs Blog*. Obtenido de: <https://www.healthaffairs.org/doi/10.1377/hblog20130911.034137/full/>. Consultado el 22 de enero de 2019.

La Segunda (Chile). 2018. Transportes planea reconocimiento facial para detectar a evasores del Transantiago. Obtenido de: <http://impresa.lasegunda.com/2018/08/29/A/H53ETVDF>. Consultado el 22 de enero de 2019.

Martin, K. 2018. Ethical Implications and Accountability of Algorithms. *Journal of Business Ethics*. Obtenido de: <https://link.springer.com/article/10.1007/s10551-018-3921-3>. Consultado el 22 de enero de 2019.

Metcalf, J. y K. Crawford. 2016. Where are Human Subjects in Big Data Research? The emerging ethics divide. *Big Data & Society*: 1-14.

PDI (Policía de Investigaciones de Chile). 2018. Biometría para identificar a autores de delitos. Obtenido de: <http://www.pdichile.cl/centro-de-prensa/detalle-prensa/2018/05/02/biometria-para-identificar-a-autores-de-delitos>. Consultado el 22 de enero de 2019.

RadioNZ. 2018a. Immigration NZ Using Data System to Predict Likely Troublemakers. Obtenido de: <https://www.radionz.co.nz/news/national/354135/immigration-nz-using-data-system-to-predict-likely-troublemakers>. Consultado el 22 de enero de 2019.

------. 2018b. Immigration Dumps Controversial Deportation Analytical Tool. Obtenido de: <https://www.radionz.co.nz/news/national/361199/immigration-dumps-controversial-deportation-analytical-tool>. Consultado el 22 de enero de 2019.

Saul, J. 2012. Skepticism and Implicit Bias. *Disputatio Lecture* 5(37): 243–263.

Stroud, M. 2014. The Minority Report: Chicago's New Police Computer Predicts Crimes, but is it Racist? Obtenido de: <https://www.theverge.com/2014/2/19/5419854/the-minority-report-this-computer-predicts-crime-but-is-it-racist>. Consultado el 6 de septiembre de 2018.

------. M. 2016. Chicago's Predictive Policing Tool Just Failed a Major Test. Obtenido de: <https://www.theverge.com/2016/8/19/12552384/chicago-heat-list-tool-failed-rand-test>. Consultado el 6 de septiembre de 2018.

The Guardian. 2015. Women Less Likely to be Shown Ads for High-Paid Jobs on Google, Study Shows. Obtenido de: <https://www.theguardian.com/technology/2015/jul/08/women-less-likely-ads-high-paid-jobs-google-study>. Consultado el 22 de enero de 2019.

Sección 3

Accenture. 2016. Universal Principles of Data Ethics: 12 Guidelines for Developing Ethics Codes. Obtenido de: https://www.accenture.com/t20160629T012639Z__w__/us-en/_acnmedia/PDF-24/Accenture-Universal-Principles-Data-Ethics.pdf. Consultado el 22 de enero de 2019.

AI Now. 2018. Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability. Obtenido de: <https://ainowinstitute.org/aiareport2018.pdf>. Consultado el 22 de enero de 2019.

Ajuntament de Barcelona. S.f. Pla Digital de l'Ajuntament de Barcelona. Mesure de Govern de Gestió Ètica i Responsable de Dades: Barcelona Data Commons. Obtenido de: <http://ajuntament.barcelona.cat/premsa/wp-content/uploads/2018/02/Economia-180213-Gesti%C3%B3-%C3%A8tica-dades.pdf>. Consultado el 22 de enero de 2019.

Article 29 Data Protection Working Party. 2012. Opinion 02/2012 on Facial Recognition in Online and Mobile Services. Obtenido de: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2012/wp192_en.pdf. Consultado el 22 de enero de 2019.

Asamblea General de Naciones Unidas. 1989. Principios rectores para la reglamentación de los ficheros computadorizados de datos personales. Obtenido de: https://digitallibrary.un.org/record/79050/files/A_44_606-ES.pdf. Consultado el 22 de enero de 2019.

------. 2018b. Data Sharing Considerations for Human Research Ethics Committees. Obtenido de: https://www.andso.org.au/__data/assets/pdf_file/0009/748737/HREC_Guide.pdf. Consultado el 22 de enero de 2019.

------. 2018c. Publishing and Sharing Sensitive Data. Obtenido de:

https://www.andso.org.au/__data/assets/pdf_file/0010/489187/Sensitive-Data-Guide-2018.pdf. Consultado el 22 de enero de 2019.

Brandom, R. 2018. UK Cops Will Deploy Facial Recognition Scanners for Soccer Championship. Obtenido de: <https://www.theverge.com/2017/4/26/15435620/champions-league-final-cardiff-facial-recognition>. Consultado el 22 de enero de 2019.

British Embassy in Mexico, Oxford Insights y C Minds. 2018. Towards an AI Strategy in Mexico: Harnessing the AI Revolution. Obtenido de: https://docs.wixstatic.com/ugd/7be025_e726c582191c49d2b8b6517a590151f6.pdf. Consultado el 22 de enero de 2019.

CIFAR. 2018. Pan-Canadian Artificial Intelligence Strategy. Obtenido de: <https://www.cifar.ca/ai/pan-canadian-artificial-intelligence-strategy>. Consultado el 22 de enero de 2019.

CNIL (Commission Nationale Informatique & Libertés). 2017. How Can Humans Keep the Upper Hand? The Ethical Matters Raised by Algorithms and Artificial Intelligence. Obtenido de: https://www.cnil.fr/sites/default/files/atoms/files/cnil_rapport_ai_gb_web.pdf. Consultado el 22 de enero de 2019.

Data Futures Partnership. 2017. A Path to Social License: Guidelines for Trusted Data Use. Obtenido de: <https://trusteddata.co.nz/wp-content/uploads/2017/08/Summary-Guidelines.pdf>. Consultado el 22 de enero de 2019.

Department for Digital, Culture, Media & Sport, Reino Unido. 2018. Data Ethics Framework. Obtenido de: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/737137/Data_Ethics_Framework.pdf Consultado el 22 de enero de 2019.

Digital Watch Observatory. 2018. Cambridge Analytica Explained: The Facts, Implications and Open Questions. Obtenido de: <https://dig.watch/trends/cambridge-analytica>

Ernst & Young LLP. 2017. Data Ethics: Digital Dilemmas for the 21st Century Board. Obtenido de: [https://www.ey.com/Publication/vwLUAssets/EY-Data-ethics-digital-dilemmas-for-the-21st-century-board/\\$FILE/EY-Data-ethics-digital-dilemmas-for-the-21st-century-board.pdf](https://www.ey.com/Publication/vwLUAssets/EY-Data-ethics-digital-dilemmas-for-the-21st-century-board/$FILE/EY-Data-ethics-digital-dilemmas-for-the-21st-century-board.pdf). Consultado el 22 de enero de 2019.

Ethictoolkit.ai. S.f. Ethics and Algorithms Toolkit. Obtenido en: <http://ethicstoolkit.ai/>. Consultado el 22 de enero de 2019.

European Commission. S.f. The European Group on Ethics in Science and New Technologies (EGE). Obtenido en: <https://ec.europa.eu/research/ege/index.cfm?pg=about>. Consultado el 22 de enero de 2019.

-----, S.f. Commission Appoints Expert Group on AI and Launches the European AI Alliance. Obtenido de: <https://ec.europa.eu/digital-single-market/en/news/commission-appoints-expert-group-ai-and-launches-european-ai-alliance>. Consultado el 22 de enero de 2019.

EUR Lex. 2016. Regulation (EU) 2016/679 of the European Parliament and of the Council of on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data. Obtenido de: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=EN>. Consultado el 22 de enero de 2019.

Future of Life Institute. 2018. National and International AI Strategies. Obtenido de: <https://futureoflife.org/national-international-ai-strategies/?cn-reloaded=1>. Consultado el 22 de enero de 2019.

Government of Canada. 2018. Algorithmic Impact Assessment (v0.2) - Government of Canada Digital Playbook. Borrador. Obtenido de: <https://canada-ca.github.io/digital-playbook-guide-numerique/views-vues/automated-decision-automatise/en/algorithmic-impact-assessment.html>. Consultado el 22 de enero de 2019.

Google. 2018. AI at Google: Our Principles. Obtenido en: <https://www.blog.google/technology/ai/ai-principles/>. Consultado el 22 de enero de 2019.

House of Commons, Science and Technology Committee, Reino Unido. 2018. Algorithms in Decision-making. Obtenido de: <https://publications.parliament.uk/pa/cm201719/cmselect/cmsctech/351/351.pdf>. Consultado el 22 de enero de 2019.

IEEE. 2016. Ethically Aligned Design. Obtenido de: https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead_brochure.pdf. Consultado el 22 de enero de 2019.

Knoppers, B. y A. Thorogood, A. 2017. Ethics and Big Data in Health. Current Opinion in Systems Biology, 4: 53-57. doi: 10.1016/j.coisb.2017.07.001. Consultado el 22 de enero de 2019.

Microsoft. S.f. Microsoft AI Principles. Obtenido en: <https://www.microsoft.com/en-us/ai/our-approach-to-ai>. Consultado el 22 de enero de 2019.

Ministerio de Industria, Comercio y Turismo, España. 2017. Constituido el Grupo de Sabios sobre Inteligencia Artificial y Big Data. Obtenido de: <https://www.mincotur.gob.es/es-ES/GabinetePrensa/NotasPrensa/2017/Paginas/grupo-expertos-big-data20171114.aspx>. Consultado el 22 de enero de 2019.

MyData.org. 2018. Declaration. Obtenido de: <https://mydata.org/declaration/>. Consultado el 22 de enero de 2019.

NTIA (National Telecommunications and Information Administration). 2016. Privacy Best Practice Recommendations For Commercial Facial Recognition Use. Obtenido de: https://www.ntia.doc.gov/files/ntia/publications/privacy_best_practices_recommendations_for_commercial_use_of_facial_recognition.pdf. Consultado el 22 de enero de 2019.

New York City Council. 2018. Obtenido en: <http://legistar.council.nyc.gov/LegislationDetail.aspx?ID=3137815&-GUID=437A6A6D-62E1-47E2-9C42-461253F9C6D0>. Consultado el 22 de enero de 2019.

NTRAD (National Artificial Intelligence Research and Development Strategic Plan). 2016. Obtenido de: https://www.nitrd.gov/news/national_ai_rd_strategic_plan.aspx. Consultado el 22 de enero de 2019.

OECD (Organisation for Economic Co-operation and Development). 2013. Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data. Obtenido de <https://www.oecd.org/sti/ieconomy/2013-oecd-privacy-guidelines.pdf>. Consultado el 22 de enero de 2019.

Partnership on AI. S.f. Obtenido en: <https://www.partnershiponai.org/>. Consultado el 22 de enero de 2019.

Pearce, R. 2018. Budget 2018: Government Seeks to Boost Australian AI Capabilities. Computerworld. Obtenido de: <https://www.computerworld.com.au/article/640926/budget-2018-government-seeks-boost-australian-ai-capabilities/>. Consultado el 22 de enero de 2019.

Perkins, A. 2018. May to Pledge Millions to AI Research Assisting Early Cancer Diagnosis. The Guardian. Obtenido de: <https://www.theguardian.com/technology/2018/may/20/may-to-pledge-millions-to-ai-research-assisting-early-cancer-diagnosis>. Consultado el 22 de enero de 2019.

Philippine Health Research Ethics Board. 2017. National Ethical Guidelines for Health and Health-related Research. Obtenido de: <http://www.ethics.healthresearch.ph/index.php/phoca-downloads/category/4-neg?-download=98:neghhr-2017>. Consultado el 22 de enero de 2019.

Portal, G. 2018. Facial Recognition Faces Legal Challenge. BBC News. Obtenido de: <https://www.bbc.com/news/uk-44928792>. Consultado el 22 de enero de 2019.

Sterling, B. 2018. The Asilomar AI Principles Blog. Obtenido de: <https://www.wired.com/beyond-the-beyond/2018/08/asilomar-ai-principles-2/>. Consultado el 22 de enero de 2019.

The Verge. 2018. Axon Launches AI Ethics Board to Study the Dangers of Facial Recognition. Obtenido de: <https://www.theverge.com/2018/4/26/17285034/axon-ai-ethics-board-facial-recognition-racial-bias>

Trusted Data Dial. 2018. Obtenido de: <https://trusteddata.co.nz/>. Consultado el 22 de enero de 2019.

University of Chicago, Center for Data Science and Public Policy. S.f. Aequitas. Obtenido de: <https://dsapp.uchicago.edu/aequitas/>. Consultado el 22 de enero de 2019.

Vadeboncoeur, C., N. Townsend, C. Foster y M. Sheehan. 2016. Variation in University Research Ethics Review: Reflections Following an Inter-university Study in England. Research Ethics, 12(4): 217-233. Obtenido de <http://journals.sagepub.com/doi/pdf/10.1177/1747016116652650>. Consultado el 22 de enero de 2019.

Villani, C. 2018. For a Meaningful Artificial Intelligence. Towards a French and European Strategy. Obtenido de: https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf. Consultado el 22 de enero de 2019.

Vincent, J. 2018. Chinese police are using facial recognition sunglasses to track citizens. The Verge. Obtenido de: <https://www.theverge.com/2018/2/8/16990030/china-facial-recognition-sunglasses-surveillance>. Consultado el 22 de enero de 2019.

WHO (World Health Organization). 2011. Standards and Operational Guidance for Ethics Review of Health-Related Research with Human Participants. Obtenido de:

http://apps.who.int/iris/bitstream/handle/10665/44783/9789241502948_eng.pdf?sequence=1. Consultado el 22 de enero de 2019.

-----, 2015. Global Health Ethics Key Issues. Obtenido de: http://apps.who.int/iris/bitstream/handle/10665/164576/9789240694033_eng.pdf?sequence=1. Consultado el 22 de enero de 2019.

Sección 4

Data Futures Partnership. S.f. Transparent Data Use Dial. Obtenido de: <https://trusteddata.co.nz/organisations/>. Consultado el 22 de enero de 2019.

DataONE Public Participation in Scientific Research Working Group. 2013. Data Management Guide for Public Participation in Scientific Research. Obtenido de: <https://www.dataone.org/sites/all/documents/DataONE-PPSR-DataManagementGuide.pdf>. Consultado el 22 de enero de 2019.

Faundeen, J. L. y V. B. Hutchison. 2017. The Evolution, Approval and Implementation of the U.S. Geological Survey Science Data Lifecycle Model. *Journal of eScience Librarianship* 6(2): e1117. Obtenido de: <https://doi.org/10.7191/jeslib.2017.1117>. Consultado el 22 de enero de 2019.

Ortiz Freuler, J. y C. Iglesias. 2018. Algoritmos e inteligencia artificial en Latinoamérica: Un estudio de implementaciones por parte de los gobiernos en Argentina y Uruguay. World Wide Web Foundation. Obtenido de: http://webfoundation.org/docs/2018/09/WF_AI-in-LA_Report_Spanish_Screen_AW.pdf. Consultado el 22 de enero de 2019.

Data Futures Partnership. 2017. A Path to Social License: Guidelines for Trusted Data Use. Obtenido de: <https://trusteddata.co.nz/wp-content/uploads/2017/08/Summary-Guidelines.pdf>. Consultado el 22 de enero de 2019.

Department for Digital, Culture, Media & Sport, Reino Unido. 2018. Data Ethics Framework. Obtenido de: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/737137/Data_Ethics_Framework.pdf. Consultado el 22 de enero de 2019.

University of Chicago. S.f. Bias & Fairness Audit. Obtenido de: <http://aequitas.dssg.io/>. Consultado el 22 de enero de 2019.

University of Chicago, Data Science for Social Good. S.f. Supporting Proactive Diabetes Screenings to Improve Health Outcomes. Obtenidos de: <https://dssg.uchicago.edu/project/supporting-proactive-diabetes-screenings-to-improve-health-outcomes/> <https://www.youtube.com/watch?v=SFPVAj8Fefs>. Consultado el 22 de enero de 2019.

UNO. S.f. Cómo funciona el sistema para prevenir embarazos adolescentes en Salta. Obtenido de: https://uno.com.ar/tecnologia/como-funciona-el-sistema-para-predecir-embarazos-adolescentes-de-salta-04122018_rJxfqbraiM. Consultado el 22 de enero de 2019.

US Geological Survey, Department of the Interior. S.f. Data Management -Acquire. Obtenido de: <https://www.usgs.gov/products/data-and-tools/data-management/acquire>. Consultado el 22 de enero de 2019.

-----, S.f. Data Management -Analyze. Obtenido de: https://www.usgs.gov/products/data-and-tools/data-management/analyze?qt-science_support_page_related_con=0#qt-science_support_page_related_con. Consultado el 22 de enero de 2019.

Zook M., S. Barocas, D. Boyd, K. Crawford, E. Keller y S. P. Gangadharan et al. 2017. Ten Simple Rules for Responsible Big Data Research. *PLoS Comput Biol*,13(3): e1005399. Obtenido de: <https://doi.org/10.1371/journal.pcbi.1005399>. Consultado el 22 de enero de 2019.

AUTORES

César Buenadicha, Gemma Galdon Clavell, María Paz Hermosilla, Daniel Loewe y Cristina Pombo.

Un agradecimiento especial a Laura Bavestrello, Victoria Peuvrelle y a Miguel Valbuena.

DISEÑO GRÁFICO

Los Pájaros Comunicaciones.

