REPORT

AUTOMATING SOCIETY

2020

EDICIÓN EN ESPAÑOL

ALGORITHM WATCH

BertelsmannStiftung

INFORMACIONES LEGALES

Automating Society Report 2020

Febrero de 2021

Disponible en línea en https://automatingsociety.algorithmwatch.org/report2020/spain/

Publicado por
AlgorithmWatch gGmbH
Linienstr. 13
10178 Berlin
Alemania

Bertelsmann Stiftung
Carl-Bertelsmann-Str. 256
33311 Gütersloh
Alemania

Editores
Fabio Chiusi
Sarah Fischer
Nicolas Kayser-Bril
Matthias Spielkamp

Editor de la edición en español
José Miguel Calatayud

Traductor
Santiago Greco

Director de proyecto
Fabio Chiusi

Coordinador de publicaciones
Marc Thümmler

Fumettisti
Samuel Daveti
Lorenzo Palloni
Alessio Ravazzani

Maquetación
Beate Autering

Edición adicional
Leonard Haas

Fecha de actualización: 30 de septiembre de 2020

Índice

Índice 3

La vida en la sociedad automatizada 4

Recomendaciones de políticas públicas 12

Unión Europea 16

España 35

Equipo 53

La vida en la sociedad automatizada. De qué forma se volvieron populares los sistemas de toma de decisiones automatizados y qué hacer

Por Fabio Chiusi

**La fecha de entrega de este informe fue el 30 de septiembre de 2020.
No se ha podido incluir acontecimientos posteriores.**

Era un nublado día de agosto en Londres y los estudiantes estaban enfadados. Cientos de ellos se habían reunido en la Plaza del Parlamento protestando con pancartas y con el apoyo de unos aliados inusuales: sus maestros, y un objetivo aún más inusual: un algoritmo.

Debido a la pandemia de COVID-19, las escuelas en el Reino Unido cerraron en marzo. Con el virus aún activo en toda Europa durante el verano de 2020, los estudiantes sabían que sus exámenes finales iban a ser cancelados y sus evaluaciones, de alguna manera, iban a cambiar. Lo que jamás habrían imaginado, sin embargo, fue que como resultado miles de ellos terminarían con calificaciones inferiores a lo esperado.

Los estudiantes que protestaban sabían cuál era el culpable, como lo demostraban sus carteles y cantos: el sistema de toma de decisiones automatizado (ADM) implementado por la Oficina de Regulación de Calificaciones y Exámenes (Ofqual). Su plan era producir la mejor evaluación basada en datos, tanto para los Certificados Generales de Educación Secundaria como para los resultados de nivel A, de tal manera que "la distribución de calificaciones siga un patrón similar al de otros años, de modo que este año los estudiantes no enfrenten una desventaja sistémica como consecuencia de las circunstancias actuales".

El gobierno quiso evitar el exceso de optimismo[1] que habría resultado basándose solo en el juicio humano, según su propios cálculos: en comparación con la serie histórica, las calificaciones habrían sido demasiado altas. Pero este intento de ser "en la medida de lo posible justo con los estudiantes que no habían podido presentarse a los exámenes este verano" fracasó espectacularmente y, en ese gris día de agosto de protesta, los estudiantes siguieron llegando, cantando y llevando carteles que expresaban una urgente necesidad de justicia social. Algunos estaban desesperados, otros se derrumbaban y lloraban.

"Dejad de robar nuestro futuro", decía una pancarta, haciéndose eco de las protestas de los activistas climáticos de ‚Viernes por el Futuro'. Otros carteles, sin embargo, apuntaban específicamente a los defectos del sistema de ADM de clasificación: "Califica mi trabajo, no mi código postal"; "Somos estudiantes, no estadísticas", decían denunciando los resultados discriminatorios del sistema[2].

Finalmente, un canto surgió de la multitud, uno que ha llegado a definir el futuro de la protesta: "A la mierda el algoritmo". Temerosos de que el gobierno estuviera automatizando de manera casual – y opaca – su futuro, sin importar cuán inconsistente con sus habilidades y esfuerzos, los estudiantes gritaban por el derecho a que sus oportunidades de vida no se vean afectadas indebidamente por un código incorrecto. Querían tener voz y lo que decían debía ser escuchado.

Los algoritmos no son ni "neutrales" ni "objetivos", aunque tendemos a pensar que lo son. Replican las conjeturas y creencias de quienes deciden implementarlos y programarlos. Los seres humanos, por lo tanto, son, o deberían ser, responsables de las elecciones algorítmicas buenas y malas, no es culpa de los "algoritmos" o los sistemas de ADM. La máquina puede dar miedo, pero el espíritu en su interior es siempre humano. Y los humanos son complicados, incluso más que los algoritmos.

De todas formas, los estudiantes que protestaban no eran tan ingenuos como para creer que sus problemas eran únicamente culpa de un algoritmo. De hecho, no estaban cantando contra "el algoritmo" en un arrebato de determinismo tecnológico; estaban más bien motivados por el impulso de proteger y promover la justicia social. En este sentido, su protesta se parece más a la de los luditas. Al igual que el movimiento obrero que aplastó los telares mecánicos y las máquinas de tejer en el siglo XIX, saben que los sistemas de ADM tienen que ver con el poder y no deben confundirse con una tecnología supuestamente objetiva. Es por eso que cantaban "justicia para la clase trabajadora", pedían la renuncia del secretario de Salud y describían al sistema de ADM como" clasismo en su máxima expresión", "clasismo desconsiderado".

Finalmente, los estudiantes lograron abolir el sistema que ponía en riesgo su carrera educativa y sus oportunidades en la vida: en un espectacular cambio de rumbo, el gobierno del Reino Unido desechó el sistema de ADM propenso a errores y utilizó las calificaciones pronosticadas por los maestros.

Pero hay más en esta historia que la victoria de los manifestantes. Este ejemplo destaca cómo los sistemas mal di-

1 "La literatura de investigación sugiere que, al estimar las calificaciones que es probable que obtengan los estudiantes, los maestros tienden a ser optimistas (aunque no en todos los casos)", Escribe Ofqual, cfr. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/909035/6656-2_-_Executive_summary.pdf

2 Cfr. el capítulo del Reino Unido para más detalles.

señados, implementados y supervisados – que reproducen prejuicios humanos y discriminación – no aprovechan el potencial que tienen los sistemas de ADM, como el poder de establecer comparaciones y la equidad inherente.

Más claramente que muchas luchas del pasado, esta protesta revela que no estamos automatizando la sociedad, sino que ya la hemos automatizado y, finalmente, alguien se había dado cuenta.

/ De la Sociedad de la Automatización a la sociedad automatizada

Cuando lanzamos la primera edición de este informe, decidimos llamarla "Automating Society" (Sociedad de la Automatización), ya que los sistemas de ADM en Europa eran en su mayoría nuevos y experimentales, no estaban catalogados y – sobre todo –, eran la excepción más que la norma.

Esta situación ha cambiado rápidamente. Como demuestran claramente los numerosos casos recopilados en este informe a través de nuestra destacada red de investigadores, la implementación de sistemas de ADM ha aumentado enormemente en poco más de un año. Los sistemas de ADM afectan ahora a casi todos los tipos de actividades humanas y, principalmente, a la distribución de servicios que afectan a millones de ciudadanos europeos – así como el acceso a sus derechos.

La tenaz opacidad que circunda el uso cada vez mayor de los sistemas de ADM ha hecho que sea aún más urgente que continuemos aumentando nuestros esfuerzos. Por lo tanto, hemos agregado cuatro países (Estonia, Grecia, Portugal y Suiza) a los 12 que ya analizamos en la edición anterior de este informe, lo que nos lleva a un total de 16 países. Si bien no es exhaustivo, esto nos permite ofrecer una imagen más amplia del escenario de ADM en Europa. Teniendo en cuenta el impacto que estos sistemas pueden tener en la vida cotidiana y cuán profundamente desafían nuestras intuiciones – así como también nuestras normas y reglas – sobre la relación entre la gobernabilidad democrática y la automatización, creemos que este es un esfuerzo esencial.

Esto es especialmente cierto durante la pandemia de COVID-19, un momento en el que hemos sido testigos de la adopción (en su mayoría apresurada) de una gran cantidad de sistemas de ADM con el objetivo de contribuir a asegurar la salud pública a través de herramientas basadas en datos y automatizaciones. Consideramos que este desarrollo era tan importante que decidimos dedicarle un "informe preliminar", publicado[3] en agosto de 2020 en el marco del proyecto ‚Automating Society'.

Incluso en Europa, cuando se trata de implementar sistemas de ADM no hay límite que valga. Basta pensar en algunos de los casos presentados en este informe, que se suman a los muchos – desde la asistencia social hasta la educación, el sistema de salud y el poder judicial – de los que ya informamos en la edición anterior. En las siguientes páginas, y por primera vez, proveemos actualizaciones sobre el desarrollo de estos casos de tres maneras. Primero, a través de relatos periodísticos, luego, a través de secciones de investigación que catalogan diferentes ejemplos y, finalmente, con novelas gráficas. Creemos que estos sistemas de ADM son – y serán cada vez más – tan cruciales en la vida de todos que necesitábamos intentar comunicar cómo funcionan y en qué *nos afectan* realmente, tanto de forma rigurosa como novedosa, para llegar a todo tipo de públicos. Después de todo, los sistemas de ADM tienen un impacto en todos nosotros.

O por lo menos deberían. Hemos visto, por ejemplo, cómo un nuevo servicio automatizado y proactivo distribuye los subsidios familiares en Estonia. Los padres ya ni siquiera necesitan solicitar subsidios: desde el nacimiento, el Estado recopila toda la información disponible sobre cada recién nacido y sus padres, y la centraliza en bases de datos. Como resultado, los padres reciben automáticamente los subsidios si tienen derecho a ellos.

En Finlandia, la identificación de los factores de riesgo individuales con respecto a la exclusión social en los adultos jóvenes se automatiza mediante una herramienta desarrollada por el gigante japonés Fujitsu. En Francia, los datos de las redes sociales se pueden extraer para surtir algoritmos de aprendizaje automático que son a su vez empleados para detectar fraudes fiscales.

Italia está experimentando con la "jurisprudencia predictiva". Un método que utiliza la automatización para ayudar a los jueces a comprender las tendencias de resoluciones judiciales anteriores sobre el tema en cuestión. Y, en Dinamarca, el gobierno intentó monitorear cada clic del teclado y del mouse en las computadoras de los estudiantes durante los exámenes, lo que provocó – igualmente – protestas

3 'Automated Decision-Making Systems in the COVID-19 Pandemic: A European Perspective', https://algorithmwatch.org/en/project/automating-society-2020-covid19/

masivas que comportaron la retirada del sistema, por el momento.

/ Es hora de corregir los errores de los sistemas de ADM

En principio, los sistemas de ADM tienen el potencial de beneficiar la vida de las personas – procesando grandes cantidades de datos, ayudando a la gente en la toma de decisiones y proporcionando utilidades personalizadas.

En la práctica, sin embargo, hemos encontrado muy pocos casos que demostraran de manera convincente un impacto tan positivo.

Por ejemplo, el sistema VioGén, empleado en España desde 2007 para evaluar el riesgo en casos de violencia doméstica, aunque lejos de ser perfecto, presenta "indicadores de rendimiento razonables" y ha ayudado a proteger a muchas mujeres de la violencia.

En Portugal, se ha implementado un sistema automatizado centralizado para impedir el fraude asociado con las recetas médicas que, según se informa, redujo el fraude en un 80% en un solo año. Un sistema similar, en Eslovenia, usado para combatir el fraude fiscal ha sido útil para los inspectores, según las autoridades fiscales.[4].

4 Cfr. el capítulo sobre Eslovenia para más detalles.

El reconocimiento facial, casi ausente en la edición de 2019, se está probando e implementando a un ritmo alarmante en toda Europa.

Al observar el estado actual de los sistemas de ADM en Europa, nos damos cuenta de que los ejemplos positivos con beneficios claros son raros. A lo largo del informe, describimos cómo la gran mayoría de los usos tienden a poner en riesgo a las personas en lugar de ayudarlas. Pero, para juzgar verdaderamente el impacto real positivo y negativo, necesitamos más transparencia sobre los objetivos y más datos sobre el funcionamiento de los sistemas de ADM que son probados e implementados.

El mensaje para los responsables políticos no podría ser más claro. Si realmente queremos aprovechar al máximo su potencial, respetando al mismo tiempo los derechos humanos y la democracia, el momento de dar un paso adelante es ahora, es necesario que estos sistemas sean transparentes y corregir los errores de la ADM.

/ Reconocimiento facial, reconocimiento facial por todos lados

Se están adoptando diferentes herramientas en diferentes países. Sin embargo, una tecnología es por ahora la más común: el reconocimiento facial. Es posiblemente el desarrollo más reciente, rápido y preocupante que se destaca en este informe. El reconocimiento facial, casi ausente en la edición de 2019, se está probando e implementando a un ritmo alarmante en toda Europa. En poco más de un año desde nuestro último informe, el reconocimiento facial está presente en escuelas, estadios, aeropuertos e incluso en casinos. También se utiliza para la vigilancia predictiva, para capturar a los delincuentes, contra el racismo y, con respecto a la pandemia de COVID-19, para reforzar el distanciamiento social, tanto en aplicaciones como a través de videovigilancia "inteligente".

Continúan las nuevas implementaciones de sistemas de ADM, incluso frente a la creciente evidencia de su falta de precisión. Y cuando surgen desafíos, los defensores de estos sistemas simplemente intentan encontrar un modo para sobrepasarlos. En Bélgica, un sistema de reconocimiento facial utilizado por la policía todavía está "parcialmente activo", a pesar de que el órgano de supervisión de la Información Policial lo ha prohibido temporalmente. Y, en Eslovenia, el uso de la tecnología de reconocimiento facial por parte de la policía se legalizó cinco años después de que comenzaran a usarla.

Esta tendencia, si no es desafiada, corre el riesgo de normalizar la idea de ser observado constante y opacamente,

cristalizando así un nuevo *status quo* de vigilancia masiva generalizada. Esta es la razón por la que muchos miembros de la comunidad de libertades civiles habrían recibido con gratitud una respuesta política mucho más agresiva por parte de las instituciones de la UE a esta amenaza[5].

Incluso el acto de sonreír es ahora parte de un sistema de ADM probado en bancos en Polonia: cuanto más sonríe un empleado, mejor es la recompensa. Y no son solo las caras las que están siendo monitoreadas. En Italia, se propuso un sistema de vigilancia de sonido como herramienta contra el racismo para ser utilizado en todos los estadios de fútbol.

/ Las cajas negras siguen siendo cajas negras

Un hallazgo sorprendente en este informe es que, si bien el cambio ocurrió rápidamente con respecto a la implementación de los sistemas de ADM, no ocurrió lo mismo en lo que respecta a la transparencia de estos sistemas. En 2015, el profesor de la Facultad de Derecho de Brooklyn, Frank Pasquale, dijo que una sociedad interconectada basada en sistemas algorítmicos opacos es una "sociedad de cajas negras". Cinco años después, la metáfora, lamentablemente, se mantiene intacta – y se aplica a todos los países que estudiamos para este informe, en todos los ámbitos: no hay suficiente transparencia con respecto a los sistemas de ADM, ni en el sector público ni en el privado. Polonia incluso impone la opacidad, con una ley que introdujo su sistema automatizado para detectar cuentas bancarias utilizadas en actividades ilegales ("STIR"). La ley establece que la divulgación de los algoritmos e indicadores de riesgo adoptados podría implicar hasta 5 años de cárcel.

Si bien rechazamos firmemente la idea de que todos estos sistemas sean intrínsecamente malos – y en su lugar adoptamos una perspectiva basada en la evidencia –, sin duda es algo muy negativo no poder evaluar su funcionamiento e impacto gracias a un conocimiento preciso y basado en los hechos. Aunque solo sea porque la opacidad contribuye considerablemente a impedir la recopilación de pruebas necesarias para formular un juicio informado sobre la implementación de un sistema de ADM.

Si a esto le sumamos la dificultad que encontraron nuestros investigadores y periodistas para acceder a datos significativos sobre estos sistemas, se perfila un escenario preocupante para aquellos que quieran controlarlos y garantizar que su implementación sea compatible con los derechos fundamentales, el estado de derecho y la democracia.

/ Desafiando el status quo algorítmico

¿Qué está haciendo la Unión Europea al respecto? Aunque los documentos estratégicos producidos por la Comisión de la UE, bajo la guía de Ursula Von der Leyen, se refieren a la "inteligencia artificial" en vez de llamarlos sistemas de ADM directamente, expresan intenciones laudables: promover y realizar una "IA confiable" que ponga a "las personas por delante"[6].

Sin embargo, como se describe en el capítulo sobre la UE, el enfoque general de la UE prioriza el imperativo comercial y geopolítico de liderar la "revolución de la IA" antes que asegurarse de que sus productos sean consistentes con las garantías democráticas una vez que son adoptadas como herramientas políticas.

Esta falta de coraje político, que se vuelve evidente en la decisión de descartar las sugerencias de una moratoria sobre las tecnologías de reconocimiento facial en vivo y en lugares públicos (en su paquete de regulación de IA), es sorprendente. Especialmente, en un momento en el que muchos Estados miembros están siendo testigos de un número creciente de desafíos legales - y derrotas - por sistemas de ADM implantados apresuradamente que han impactado de forma negativa en los derechos de la ciudadanía.

Un caso histórico proviene de los Países Bajos, donde activistas de derechos civiles llevaron a los tribunales un sistema automatizado invasivo y opaco, supuestamente para detectar el fraude en la prestación de servicios sociales (SyRI), y ganaron. El sistema no sólo violaba el Convenio Europeo de Derechos Humanos por el tribunal de La Haya (que así lo dictaminó en febrero), y por lo tanto, se detuvo. Sino que el caso también sentó un precedente: según el fallo, los gobiernos tienen una "responsabilidad especial" de salvaguardar los derechos humanos al implementar dichos sistemas de ADM. Proporcionar la transparencia que tanto se necesita se considera entonces un factor crucial.

Desde nuestro primer informe, los medios de comunicación y los activistas de la sociedad civil han decidido ser una fuerza impulsora para fomentar la responsabilidad en lo

5 Como se detalla en el capítulo de la UE.

6 Cfr. el capítulo de la UE, y en particular la sección sobre el 'Libro Blanco sobre la Inteligencia Artificial' de la Comisión de la UE

Un hallazgo sorprendente en este informe es que, si bien el cambio ocurrió rápidamente con respecto a la implementación de los sistemas de ADM, no ocurrió lo mismo en lo que respecta a la transparencia de estos sistemas.

que respecta a los sistemas de ADM. En Suecia, por ejemplo, los periodistas lograron forzar la publicación del código detrás del sistema de Trelleborg que toma decisiones totalmente automatizadas relacionadas con aplicaciones de subsidios sociales. En Berlín, el proyecto piloto de reconocimiento facial de la estación de tren de Südkreuz no consiguió impulsar una sucesiva implementación en el resto de Alemania. Esto fue gracias a la fuerte oposición de los activistas, tan fuerte que lograron influir en las posiciones de los partidos y, en última instancia, en la agenda política del gobierno.

Los activistas griegos de Homo Digitalis demostraron que ningún viajero auténtico participó en las pruebas piloto de un sistema llamado ‚iBorderCtrl', un proyecto financiado por la UE que tenía como objetivo utilizar la ADM para patrullar fronteras, revelando así que las capacidades de muchos de estos sistemas eran exageradas con frecuencia. Mientras tanto, en Dinamarca suspendieron un sistema de elaboración de perfiles para la detección temprana de riesgos asociados con familias y niños vulnerables (el llamado "modelo Gladsaxe") gracias al trabajo de académicos, periodistas y la Autoridad de Protección de Datos (APD).

Las APD también desempeñaron un papel importante en otros países. En Francia, la autoridad nacional de privacidad declaró que tanto un proyecto de vigilancia de sonido como uno de reconocimiento facial eran ilegales en las escuelas secundarias. En Portugal, la APD se negó a aprobar el lanzamiento de sistemas de videovigilancia por parte de la policía en los municipios de Leiria y Portimão, ya que fue considerado desproporcionado y que equivalía a "un seguimiento sistemático a gran escala de las personas y sus hábitos y comportamientos, así como la identificación de personas a partir de información conectada a sus características físicas". Y, en los Países Bajos, la APD holandesa pidió más transparencia en los algoritmos predictivos utilizados por las agencias del gobierno.

Por último, algunos países han solicitado asesoramiento a un defensor del pueblo. En Dinamarca, este tipo de consulta ayudó a desarrollar estrategias y directrices éticas para el uso de sistemas de ADM en el sector público. En Finlandia, el defensor del pueblo parlamentario consideró ilegales los cálculos fiscales automatizados.

Y, sin embargo, dado el continuo lanzamiento de tales sistemas en toda Europa, uno se pregunta: ¿este nivel de supervisión es suficiente? Cuando el defensor del pueblo polaco cuestionó la legalidad del sistema de detección de sonrisas utilizado en un banco (ya mencionado anteriormente), la decisión no impidió el lanzamiento de una versión de prueba en la ciudad de Sopot, ni impidió que varias empresas mostraran interés en adoptar ese sistema.

/ Falta de inspección, aplicación de las normas, competencia y explicaciones adecuadas

El activismo es más que nada un esfuerzo reactivo. La mayoría de las veces, los activistas sólo pueden reaccionar a un sistema de ADM si está siendo probado o si ya ha sido

implementado. Para cuando los ciudadanos consiguen organizar una respuesta, es posible que sus derechos ya hayan sido violados innecesariamente. Esto puede suceder incluso con las protecciones que deberían otorgar, en la mayoría de los casos, la legislación de la UE y de los Estados miembros. Esta es la razón por la que las medidas proactivas para garantizar los derechos – antes de que se lleven a cabo las versiones de prueba y sus lanzamientos – son tan importantes.

Y, sin embargo, incluso en países donde existe una legislación adecuada, la aplicación de las normas simplemente no está ocurriendo. En España, por ejemplo, la "Actuación administrativa automatizada" está legalmente codificada, define requisitos específicos en términos de control de calidad y supervisión, junto con la auditoría del sistema de información y su código fuente. España también tiene una ley de Libertad de Información. Sin embargo, incluso con estas leyes, sólo en raras ocasiones, escribe nuestro investigador, los organismos públicos divulgan información detallada sobre los sistemas de ADM que utilizan. De manera similar, en Francia, existe una ley de 2016 que exige la transparencia algorítmica, pero nuevamente existe en vano.

Incluso el hecho de llevar un algoritmo a los tribunales, de acuerdo con las disposiciones específicas de una ley de transparencia algorítmica, puede no ser suficiente para hacer cumplir y proteger los derechos de los ciudadanos. Como muestra el caso del algoritmo Parcoursoup para clasificar a los solicitantes universitarios en Francia[7], se pueden crear excepciones a voluntad para proteger a una administración de la asunción de la responsabilidad.

Esto es especialmente preocupante cuando se combina con la falta endémica de habilidades y competencias en torno a los sistemas de ADM en el sector público, algo que lamentan muchos investigadores. ¿Cómo podrían los funcionarios públicos explicar o proveer transparencia de cualquier tipo en torno a sistemas que no comprenden?

Recientemente, algunos países intentaron afrontar este problema. Estonia, por ejemplo, estableció un centro de competencia sobre sistemas de ADM para estudiar mejor cómo se podrían utilizar en el desarrollo de servicios públicos y, en lo específico, con el fin de contribuir a las operaciones del Ministerio de Asuntos Económicos y Comunicaciones y la Cancillería de Estado para la creación de un gobierno electrónico. Suiza también pidió la creación de una "red de competencias" dentro del marco más amplio de la estrategia nacional "Suiza digital".

Y, sin embargo, la falta de alfabetización digital es un problema bien conocido que afecta a una gran proporción de la población en varios países europeos. Además, es difícil pedir que se respeten derechos que no sabes que tienes. Protestas en el Reino Unido y en otros lugares, junto con escándalos importantes basados en sistemas de ADM[8], ciertamente han ayudado a concienciar sobre los riesgos y las oportunidades de automatizar la sociedad. Pero, aunque se trate de una tendencia en aumento, esta conciencia se encuentra todavía en sus primeras etapas en muchos países.

Los resultados de nuestra investigación son claros: si bien los sistemas de ADM ya afectan a todo tipo de actividades y juicios, todavía son implementados principalmente sin ningún debate democrático serio. Además, el hecho de que los mecanismos de imposición y supervisión – si es que existen – se activen después de que la ADM haya sido implementado es la norma, más que la excepción.

Incluso los objetivos de estos sistemas no suelen ser justificados ni explicados a las poblaciones que afectan, sin mencionar los beneficios que se supone que deberían obtener. Basta pensar en el servicio proactivo "AuroraAI" en Finlandia: se supone que debería identificar automáticamente "eventos de la vida", como informan nuestros investigadores finlandeses, y en la mente de quien lo propone debería funcionar como "una niñera" que ayuda a los ciudadanos a satisfacer necesidades en los servicios públicos que puedan surgir junto con ciertas circunstancias de la vida, por ejemplo, mudarse a un nuevo lugar, cambios en las relaciones familiares, etc... pero según nuestros investigadores terminan "dando pequeños empujones", lo que significa que en lugar de empoderar a las personas el sistema podría terminar haciendo todo lo contrario, sugiriendo determinadas decisiones o limitando las opciones de un individuo, a causa del modo en el que fue diseñado y de su arquitectura.

Entonces es más importante si cabe saber qué es lo que se está "optimizando" en materia de servicios públicos: "¿se maximiza el uso del servicio, se minimizan los costos o se mejora el bienestar de los ciudadanos?", preguntan los investigadores. "¿En qué conjunto de criterios están basadas estas decisiones y quién las elige?". El mero hecho de que

7 Cfr. el capítulo sobre Francia

8 Piense en la debacle del algoritmo "Buona Scuola" en Italia, cfr. el capítulo sobre Italia.

no tengamos una respuesta a estas preguntas fundamentales dice mucho sobre el grado de participación y transparencia permitido en el desarrollo, incluso para un sistema de ADM con potencial tan invasivo.

/ La trampa tecno-solucionista

Existe una justificación ideológica generalizada para todo esto. Se llama "solucionismo tecnológico" y todavía afecta gravemente la forma en que se desarrollan muchos de los sistemas de ADM que estudiamos. Aunque el término haya sido condenado durante mucho tiempo como una ideología defectuosa que concibe cada problema social como un "error en el programa" que necesita un "arreglo" a través de la tecnología[9], esta retórica todavía se adopta ampliamente, tanto en los medios de comunicación como en los círculos políticos, para justificar la adopción acrítica de tecnologías automatizadas en la vida pública.

Cuando se promueven como "soluciones", los sistemas de ADM viran inmediatamente hacia el territorio descrito en la Tercera Ley de Arthur C. Clarke: magia. Y es difícil, sino imposible, reglamentar la magia, y más aún dar transparencia y explicaciones en torno a ella. Uno puede ver la mano que se mete dentro del sombrero, y como resultado aparece un conejito, pero el proceso es y *debería permanecer* una "caja negra".

Muchos investigadores involucrados en el proyecto ‚Automating Society' denunciaron esto como el fallo fundamental en el razonamiento detrás de muchos de los sistemas de ADM que describen. Esto también implica, como se muestra en el capítulo sobre Alemania, que la mayoría de las críticas a tales sistemas terminan siendo narradas como un rechazo total a la "innovación", retratando a los defensores de los derechos digitales como "neoluditas". Esto no sólo ignora la realidad histórica del movimiento ludita, que se ocupó de políticas laborales y no de tecnologías *en sí mismas*, pero también, y más importante, amenaza la efectividad de los mecanismos hipotéticos de supervisión e imposición.

En un momento en el que la industria de la "Inteligencia Artificial" está presenciando el surgimiento de grupos de presión "dinámicos", sobre todo en el Reino Unido, esto podría resultar en un "lavado de directrices éticas" y otras respuestas políticas que son ineficaces y estructuralmente inadecuadas para afrontar las consecuencias en los derechos humanos por los sistemas de ADM. Este punto de vista equivale en última instancia a la suposición de que los seres humanos deberíamos adaptarnos a los sistemas de ADM, mucho más que de que los sistemas de ADM deberían adaptarse a las sociedades democráticas.

Para contrarrestar esta narrativa, no debemos abstenernos de cuestiones fundamentales: preguntarse si los sistemas de ADM pueden ser compatibles con la democracia y si pueden ser implementados en beneficio de la sociedad en general, y no solo para algunas partes de ella. Podría ser el caso, en ese sentido, que determinadas actividades humanas – por ejemplo, las relacionadas con la prestación de servicios sociales – no deberían estar sujetas a la automatización, o que ciertas tecnologías – como el reconocimiento facial en vivo en los espacios públicos – no deberían ser promovidas como una búsqueda sin fin para el "liderazgo de la IA", sino que pudieran ser prohibidas por completo.

Aún más importante, debemos rechazar todo marco ideológico que nos impida plantear tales preguntas. Al contrario: lo que necesitamos ver ahora es que las políticas reales cambien para permitir un mayor escrutinio de estos sistemas. En la siguiente sección enumeraremos las exigencias clave que surgen de nuestros hallazgos. Esperamos que sean ampliamente discutidas y finalmente implementadas.

Solo a través de un debate democrático informado, inclusivo y basado en pruebas podemos encontrar el equilibrio adecuado entre los beneficios que los sistemas de ADM pueden proporcionar – y logran – otorgar (en términos de velocidad, eficiencia, equidad, mejor prevención y acceso a los servicios públicos) y los desafíos que plantean a los derechos de todos nosotros.

9 Véase Evgeny Morozov (2014), Para salvar todo, haga clic aquí. La locura del solucionismo tecnológico, Public Affairs, https://www.publicaffairsbooks.com/titles/evgeny-morozov/to-save-everything-click-here/9781610393706/

Recomendaciones de políticas públicas

A la luz de los hallazgos detallados en el informe Automating Society (edición 2020), recomendamos el siguiente conjunto de intervenciones políticas a los responsables políticos del parlamento de la UE y los parlamentos de los Estados miembros, la Comisión de la UE, los gobiernos nacionales, los investigadores, las organizaciones de la sociedad civil (organizaciones de abogados, fundaciones, sindicatos, etc...) y el sector privado (empresas y asociaciones empresariales). Las recomendaciones tienen como objetivo contribuir a garantizar que tanto los sistemas de ADM que están siendo implementado como los que están a punto de serlo en toda Europa sean efectivamente coherentes con los derechos humanos y la democracia:

1. Incrementar la transparencia de los sistemas de ADM

Sin la capacidad de saber con precisión cómo, por qué y con qué fin se implementan los sistemas de ADM, todos los demás esfuerzos para la reconciliación entre dichos sistemas y los derechos fundamentales están condenados al fracaso.

/ Establecer registros públicos para los sistemas de ADM utilizados en el sector público.

Por lo tanto, solicitamos que se promulgue una legislación válida para toda la UE que obligue a los Estados miembros a establecer registros públicos de los sistemas de ADM utilizados por el sector público.

Deben contener una obligación legal para que los responsables del sistema de ADM divulguen y documenten el objetivo del sistema, una explicación del modelo (lógica involucrada) e información sobre quién desarrolló tal sistema. Esta información debe estar disponible de una manera fácilmente legible y accesible, incluyendo datos digitales estructurados basados en un protocolo estandarizado.

Las autoridades públicas tienen la responsabilidad específica de hacer que las características operativas de los sistemas de ADM implementados por la administración pública sean transparentes. Así lo enfatizó un reciente recurso contencioso-administrativo en España, donde se afirma que "cuando el código fuente de un programa informático es ley, [...], el ciudadano tiene tanto derecho a inspeccionar su funcionamiento como lo tiene con respecto a cualquier otra norma jurídica". Si se ratifica, el fallo podría convertirse en un precedente en Europa.

Mientras que las declaraciones de divulgación de los sistemas de ADM deben ser obligatorios en todos los ámbitos del sector público, los mismos requisitos de transparencia también deben aplicarse a los sistemas de ADM por parte de entidades privadas cuando el sistema en cuestión (AI / ADM) tengan un impacto significativo en un individuo, en un grupo específico, o en la ciudadanía.

/ Introducir marcos normativos de acceso a datos legalmente vinculantes para apoyar y permitir la investigación en nombre del interés público.

Aumentar la transparencia no solo requiere revelar información sobre el objetivo, la lógica y el creador de un sistema, sino también la capacidad de analizar y examinar a fondo la entrada de datos y los resultados de un sistema. También requiere que los datos de entrenamiento y los resultados que derivan de ellos sean accesibles a investigadores independientes, periodistas y organizaciones de la sociedad civil para investigaciones de interés público.

Es por eso que sugerimos la introducción de marcos normativos de acceso a datos robustos y legalmente vinculantes, que estén enfocados explícitamente en apoyar y permitir la investigación de interés público en nombre del pleno respeto de la protección de datos y leyes de privacidad.

Teniendo en cuenta las mejores prácticas existentes en el ámbito nacional y de la UE, dichos marcos estratificados deberían incluir sistemas de sanciones, controles y balances, así como revisiones periódicas. Como ya han ilustrado las asociaciones privadas de intercambio de datos, hay preocupaciones legítimas con respecto a la privacidad del usuario y la posible anonimización de ciertos tipos de datos.

Los responsables políticos deben aprender de las infraestructuras que intercambian datos sobre la salud para facilitar el acceso privilegiado a ciertos tipos de datos más granulares, mientras que se aseguran de que los datos personales estén adecuadamente protegidos (por ejemplo, a través de entornos operativos seguros).

Un marco operativo de rendición de cuentas eficaz requerirá de un acceso transparente a los datos de la plataforma, este es un requisito para que numerosos procesos de auditoría sean efectivos.

2. Crear un marco de rendición de cuentas significativo para los sistemas de ADM

Como han demostrado las investigaciones de España y Francia, aun cuando la ley exige la transparencia de un sistema de ADM y/o tal información ha sido divulgada, no siempre se da lugar una rendición de cuentas. Se necesitan ulteriores etapas para garantizar que las leyes y los requisitos sean realmente aplicables.

/ Desarrollar y establecer enfoques para la auditoría efectiva de sistemas algorítmicos.

Para asegurar que la transparencia sea significativa, necesitamos complementar la primera etapa: establecer un registro público con procesos que controlen de manera efectiva los sistemas algorítmicos.

El término "auditing" (auditoría) se utiliza ampliamente, pero no existe un entendimiento común de la definición. Para nosotros, "auditoría", en este contexto, retoma la definición ISO, o sea un "proceso sistemático, independiente y documentado, mediante el cual podemos obtener evidencia – registros, declaraciones, información verificable - que permite verificar el cumplimiento de los requisitos solicitados por una determinada norma".[10]aún no tenemos respuestas satisfactorias a las preguntas complejas[11] planteadas por la auditoría de sistemas algorítmicos; sin embargo, nuestros hallazgos indican claramente la necesidad de encontrar respuestas en un amplio proceso de participación de las partes interesadas y mediante una investigación exhaustiva y dedicada.

Se deben desarrollar tanto criterios como procesos apropiados de auditoría, siguiendo un enfoque de múltiples partes interesadas que tome en consideración el efecto desproporcionado que los sistemas de ADM tienen en los grupos vulnerables y que solicite su participación.

Por lo tanto, pedimos a los responsables políticos que den inicio a dichos procesos junto con las partes interesadas para aclarar las preguntas delineadas, y que proporcionen fuentes de financiación destinadas a permitir la participación de las partes que hasta ahora no han estado representadas adecuadamente.

Además, exigimos la provisión de recursos adecuados para apoyar/financiar proyectos de investigación sobre el desarrollo de modelos con el fin de auditar de manera efectiva los sistemas algorítmicos.

10 https://www.iso.org/obp/ui/#iso:std:iso:19011:ed-3:v1:en

11 Al pensar en modelos potenciales de auditoría algorítmica, surgen varias preguntas. 1) ¿Quién/qué (servicios/plataformas/productos) debería ser auditado? ¿Cómo personalizar los sistemas de auditoría según el tipo de plataforma/tipo de servicio? 2) ¿Cuándo una institución pública debe de realizar una auditoría (en el ámbito de la UE, nacional, local) y cuándo puede ser realizada por entidades/expertos privados (empresas, sociedad civil, investigadores)? 3) ¿Cómo aclarar la distinción entre evaluar el impacto ex ante (es decir, en la fase de diseño) y ex post (es decir, después de su puesta en marcha) y sus respectivos desafíos? 4) ¿Cómo evaluar las ventajas y desventajas de las diferentes virtudes y vicios de la auditabilidad? (por ejemplo: la sencillez, generalidad, relevancia, precisión, flexibilidad, interpretabilidad, privacidad o eficacia de un procedimiento de auditoría pueden estar en tensión). 5) ¿Qué información debe estar disponible para que una auditoría sea efectiva y confiable (por ejemplo, código fuente, datos de capacitación, documentación)? ¿Quien inspecciona necesita tener acceso físico a los sistemas durante la operación para auditar de manera efectiva? 6) ¿Qué obligación de presentar pruebas es necesaria y proporcionada por los vendedores/proveedores de servicios? 7) ¿Cómo podemos garantizar que la auditoría sea posible? ¿Deben tenerse en cuenta los requisitos de auditoría en el diseño de sistemas algorítmicos ("auditables por construcción")? 8) Reglas para la publicidad: Cuando una auditoría es negativa y los problemas no se resuelven, ¿cuál debe ser el comportamiento del auditor, de qué manera se puede divulgar que ocurrió un fallo? 9) ¿Quién audita a los auditores? ¿Cómo nos aseguramos que los auditores serán responsables?

/ Apoyar a las organizaciones de la sociedad civil como organismos de control de los sistemas de ADM

Nuestros hallazgos indican claramente que el trabajo de las organizaciones de la sociedad civil es crucial para cuestionar eficazmente los sistemas de ADM opacos. A través de la investigación y la promoción y, a menudo, cooperando con el mundo académico y con los periodistas, estas organizaciones intervinieron reiteradamente en los debates políticos sobre estos sistemas en los últimos años, y en varios casos consiguieron que el interés público y los derechos fundamentales fueran debidamente considerados tanto antes como después de la implementación de dichos sistemas en muchos países europeos.

Por lo tanto, los actores de la sociedad civil deberían ser apoyados como organismos de control de la "sociedad de la automatización". Como tales, son un componente integral de cualquier marco de rendición de cuentas eficaz para los sistemas de ADM.

/ Prohibir el reconocimiento facial que // pueda corresponder a vigilancia masiva

No todos los sistemas de ADM son igualmente peligrosos, y una reglamentación con enfoque basado en el riesgo, como la de Alemania y UE, lo refleja correctamente. Pero a fin de proporcionar una rendición de cuentas viable para los sistemas que se identifican como de riesgo, se deben implementar mecanismos de supervisión y cumplimiento efectivos. Esto es aún más importante para aquellos sistemas considerados de "alto riesgo" a la hora de infringir los derechos de los usuarios.

Un ejemplo crucial que surgió de nuestros hallazgos es el reconocimiento facial. Se ha demostrado que los sistemas de ADM que se basan en tecnologías biométricas, incluido el reconocimiento facial, representan una amenaza particularmente grave para el interés público y los derechos fundamentales, ya que abren el paso a la vigilancia masiva indiscriminada, y especialmente porque a pesar de todo son implementados extensamente y de un modo opaco.

Exigimos que los usos públicos del reconocimiento facial que podrían corresponder a vigilancia masiva se prohíban a nivel de la UE de manera decisiva y urgente hasta nuevo aviso.

Estas tecnologías ya pueden ser consideradas ilegales en la UE, al menos para ciertos usos, sobre todo si son implementadas sin el "consentimiento específico" de los sujetos escaneados. Esta interpretación legal ha sido sugerida por las autoridades belgas, que emitieron una multa histórica por la implementación de un sistema de reconocimiento facial en el país.

3. Mejorar la alfabetización algorítmica y fortalecer el debate público sobre los sistemas de ADM

Una mayor transparencia de los sistemas de ADM solo puede ser realmente útil si quienes se enfrentan a ellos – los reguladores, el gobierno y los organismos de la industria –, pueden afrontar estos sistemas y su impacto de manera responsable y prudente. Además, las personas afectadas deben poder comprender dónde, por qué y cómo han sido implementados tales sistemas. Es por eso que necesitamos mejorar la alfabetización algorítmica en todos los ámbitos, con la participación de partes interesadas importantes, y también con la ciudadanía, y reforzar debates públicos más diversos sobre los sistemas de ADM y su impacto en la sociedad.

/ Establecer centros independientes de expertos en la ADM

Junto con nuestra demanda de auditoría algorítmica e investigación de apoyo, pedimos también la formación de centros independientes de expertos en la ADM en el ámbito nacional para monitorizar, evaluar, realizar investigaciones, informar y brindar asesoramiento al gobierno y la industria (en coordinación con los reguladores, la sociedad civil y el mundo académico) sobre las implicaciones sociales y de derechos humanos que conlleva el uso de sistemas de ADM. El papel general de estos centros sería crear un marco de rendición de cuentas significativo y aumentar la capacidad alrededor del tema.

Los centros nacionales de especialización deberían involucrar a las organizaciones de la sociedad civil, a los grupos de partes interesadas y a los organismos de aplicación existentes, como las APD y los organismos nacionales de derechos humanos, para beneficiar todos los aspectos del ecosistema y poder generar confianza, transparencia y cooperación entre todos los actores.

Como órganos estatutarios independientes, los centros de especialización tendrían un papel central en la coordinación del desarrollo de políticas y estrategias nacionales relacionadas con los protocolos de ADM, y en ayudar a desarrollar la capacidad (competencia / habilidades) de los reguladores existentes, el gobierno y los organismos de la industria para responder al aumento uso de sistemas de ADM.

Estos centros no deberían tener poderes legislativos, pero sí proporcionar conocimientos fundamentales sobre cómo proteger los derechos humanos individuales y prevenir daños colectivos y sociales. Deberían, por ejemplo, apoyar a las pequeñas y medianas empresas (PYME) en el cumplimiento de sus obligaciones en virtud de la diligencia debida en materia de derechos humanos, incluida la realización de evaluaciones de derechos humanos o evaluaciones de impacto algorítmicas, y mediante la inscripción de sistemas de ADM en el registro público mencionado anteriormente.

/ Promover un debate democrático inclusivo y diverso en torno a los sistemas de ADM.

Además de fortalecer las capacidades y competencias de quienes implementan sistemas de ADM, también es vital promover la alfabetización algorítmica en la ciudadanía a través de un debate más amplio y de diferentes programas.

Nuestros hallazgos sugieren que los sistemas de ADM no solo no permanecen transparentes para la ciudadanía en general cuando están en uso, sino que incluso la decisión de implementar o no un sistema de ADM en primer lugar generalmente se toma sin el conocimiento o la participación del público.

Por lo tanto, es necesario y urgente incluir al (interés) público en la toma de decisiones sobre los sistemas de ADM desde el principio.

De manera más general, necesitamos un debate público diferente a propósito del impacto de los sistemas de ADM. Tenemos que ir más allá de los grupos de expertos y hacer que el tema sea más accesible para el público en general. Eso significa hablar un idioma que no sea el tecno-judicial para atraer al público y despertar su interés.

Para hacerlo, también deben implementarse programas detallados, para construir y promover la alfabetización digital. Si nuestro objetivo es mejorar un debate público informado y crear autonomía digital para los ciudadanos en Europa, tenemos que comenzar por construir y promover la alfabetización digital, con un enfoque particular en las consecuencias sociales, éticas y políticas de la adopción de sistemas de ADM.

Preparando el escenario para el futuro de los sistemas de ADM en Europa

A medida que los sistemas automatizados de toma de decisiones ocupan un lugar central en la distribución de derechos y servicios en Europa, las instituciones de la región reconocen cada vez más su papel en la vida pública, tanto en términos de oportunidades como de desafíos.

Por Kristina Penner y Fabio Chiusi

Desde nuestro primer informe en enero de 2019 – y a pesar de que la UE todavía este sumida en un debate más amplio sobre la inteligencia artificial "confiable" –, muchos organismos, desde el Parlamento de la UE hasta el Consejo de Europa, han publicado documentos destinados a situar la UE y Europa en una trayectoria que las llevará a lidiar con la ADM en los años o incluso décadas por venir.

En el verano de 2019, la recién elegida presidenta de la Comisión, Ursula von der Leyen, autoproclamada "optimista tecnológica", se comprometió a presentar "una legislación para una estrategia europea coordinada sobre las implicaciones humanas y éticas de la inteligencia artificial" y a "reglamentar la inteligencia artificial (IA)" dentro de los 100 días posteriores a la asunción del cargo. En lugar de ello, en febrero de 2020, la Comisión Europea publicó un 'Libro blanco' sobre la IA que contiene "ideas y acciones" – un paquete de estrategias que tiene como objetivo informar a los ciudadanos y sentar las bases para futuras acciones legislativas. También defiende la "soberanía tecnológica" europea: citando las palabras de la misma Von der Leyen, lo que se traduce en "la capacidad que debe tener Europa para tomar sus propias decisiones, basándose en sus propios valores, respetando sus propias reglas", y debería "ayudar a que todos seamos optimistas tecnológicos".

Un segundo esfuerzo fundamental que afecta a los sistemas de ADM en Europa es la Ley de Servicios Digitales (DSA), anunciada en la 'Agenda para Europa' de Von der Leyen y que debería reemplazar la Directiva sobre el Comercio Electrónico que lleva en vigor desde el 2000. Su objetivo es "actualizar nuestras reglas de responsabilidad y seguridad para plataformas, servicios y productos digitales, y completar nuestro mercado único digital", lo que conduce a debates fundamentales sobre el papel de los sistemas de ADM en las políticas de moderación de contenido, la responsabilidad de los intermediarios y la libertad de expresión en general.[12].

Se puede encontrar una estrategia explícita sobre los sistemas de ADM en una Resolución aprobada por el Comité de Mercado Interior y Protección del Consumidor del Parlamento de la UE, y en una Recomendación "sobre los impactos de los sistemas algorítmicos en los derechos humanos" por el Comité de Ministros del Consejo de Europa.

El Consejo de Europa (CoE), en particular, ha ido desempeñando un papel cada vez más importante en el debate político sobre la IA a lo largo del último año, y aunque el impacto real de los esfuerzos aún está por verse, se puede afirmar que sirve como "guardián" de los derechos humanos. Esto es aun más evidente en la Recomendación, 'Unboxing Artificial Intelligence: 10 steps to protect Human Rights', por la comisionada de Derechos Humanos del CoE, Dunja Mijatović, y en el trabajo del Comité Ad Hoc sobre Inteligencia Artificial (CAHAI) establecido en septiembre de 2019.

12 Se pueden encontrar los comentarios y las recomendaciones detalladas sobre los sistemas de ADM contextualizados a la DSA en los resultados del proyecto 'Governing Platforms' de AlgorithmWatch.

MUCHOS OBSERVADORES NOTAN UNA TENSIÓN FUNDAMENTAL ENTRE LAS EMPRESAS Y LOS IMPERATIVOS DE DERECHO EN LA FORMA EN QUE LAS INSTITUCIONES DE LA UE, Y ESPECIALMENTE LA COMISIÓN, ENMARCAN SUS REFLEXIONES Y PROPUESTAS SOBRE IA Y SISTEMAS DE ADM.

Muchos observadores notan una tensión fundamental entre las empresas y los imperativos de derecho en la forma en que las instituciones de la UE, y especialmente la Comisión, enmarcan sus reflexiones y propuestas sobre IA y sistemas de ADM. Por un lado, Europa quiere "aumentar el uso y la demanda de datos (así como productos y servicios habilitados para ello) en todo el Mercado Único"; convirtiéndose así en un "líder" en las aplicaciones comerciales de la IA e impulsando la competitividad de las empresas de la UE frente a la creciente presión de los rivales, Estados Unidos y China. Esto es aún más importante para los sistemas de ADM, ya que se supone que, a través de esta economía de "datos ágiles", la UE "puede convertirse en un modelo de liderazgo para una sociedad que gracias a los datos puede tomar mejores decisiones en las empresas y en el sector público" . Como dice el Libro Blanco sobre IA, "Los datos son el elemento vital del desarrollo económico".

Mientras que, por otro lado, el procesamiento automático de datos sobre la salud, el trabajo y el bienestar de un ciudadano puede generar decisiones con resultados discriminatorios e injustos. La UE lidia con "lado oscuro" de los algoritmos en los procesos de toma de decisiones a través de una serie de principios de su "caja de herramientas". En el caso de los sistemas con alto riesgo, las reglas deben garantizar que los procesos automatizados de toma de decisiones sean compatibles con los derechos humanos y también con controles y contrapesos democráticos. Este es un enfoque único que las instituciones de la UE etiquetan como "centrado en el ser humano", y fundamentalmente opuesto a los métodos aplicados en los EE. UU. (guiados por el beneficio económico) y China (guiados por la seguridad nacional y la vigilancia masiva).

Sin embargo, han surgido dudas sobre si Europa podrá alcanzar ambos objetivos al mismo tiempo. El reconocimiento facial es un buen ejemplo: a pesar de que, como muestra este informe, ahora tenemos muchas pruebas de implementación de sistemas no controlada y opaca en la mayoría de los países miembros, la Comisión Europea no ha actuado con rapidez y decisión para proteger los derechos de los ciudadanos europeos. Como revelaron los borradores que se han filtrado del Libro Blanco de la IA[13], la UE estaba a punto de prohibir la "identificación biométrica remota" en lugares públicos, antes de eludir la cuestión en el último minuto y promover en su lugar un "amplio debate" sobre el tema.

Mientras tanto, el controvertido uso de sistemas de ADM que controlan las fronteras, incluso con reconocimiento facial, todavía están siendo impuestos en los proyectos financiados por la UE.

Políticas y debates políticos

/ El paquete de estrategia de datos europea y el Libro Blanco sobre IA

Si bien la legislación integral prometida "para tener una estrategia europea coordinada sobre las implicaciones humanas y éticas de la inteligencia artificial" (anunciada en la "Agenda para Europa" de Von der Leyen) no ha sido presentada dentro de sus "primeros 100 días en el cargo", la Comisión European publicó una serie de documentos que proporcionan un conjunto de principios e ideas para conformar tal legislación.

> "QUEREMOS INCENTIVAR A NUESTRAS EMPRESAS, A NUESTROS INVESTIGADORES, A LOS INNOVADORES, A LOS EMPRENDEDORES, PARA QUE DESARROLLEN INTELIGENCIA ARTIFICIAL. Y QUEREMOS ANIMAR A NUESTROS CIUDADANOS A QUE SE SIENTAN SEGUROS PARA UTILIZARLA. TENEMOS QUE DAR RIENDA SUELTA A ESTE POTENCIAL."
> URSULA VON DER LEYEN

El 19 de febrero de 2020, se publicaron conjuntamente una "Estrategia Europea de Datos" y un "Libro Blanco sobre inteligencia artificial", estableciendo los principios fundamentales del enfoque estratégico de la UE hacia la IA (incluidos los sistemas de ADM, aunque no son mencionados explícitamente). Estos principios incluyen, poner "a las personas primero" ("tecnología que funciona para las personas"), neutralidad tecnológica (ninguna tecnología es buena o mala en sí misma, algo que se determinará únicamente por su uso) y, por supuesto, la soberanía y el optimismo. Usando las palabras de Von der Leyen: "Queremos incentivar a nuestras empresas, a nuestros investigadores,

13 https://www.politico.eu/article/eu-considers-temporary-ban-on-facial-recognition-in-public-spaces/

a los innovadores, a los emprendedores, para que desarrollen Inteligencia Artificial. Y queremos animar a nuestros ciudadanos a que se sientan seguros para utilizarla. Tenemos que dar rienda suelta a este potencial".

La idea de fondo es que las nuevas tecnologías no deberían traer nuevos valores. El "nuevo mundo digital" imaginado por la administración de Von der Leyen debería proteger plenamente los derechos humanos y civiles. La "excelencia" y la "confianza", destacadas en el mismo título del Libro Blanco, son vistas como los pilares gemelos sobre los que puede y debe apoyarse un modelo europeo de IA, diferenciándolo de las estrategias tanto la de Estados Unidos como la de China.

Sin embargo, esta ambición carece de detalles en el Libro Blanco. Por ejemplo, el Libro Blanco establece una estrategia de riesgo para la reglamentación de la IA, en la que la regulación es proporcional al impacto de los sistemas de "IA" en la vida de los ciudadanos. "Para casos de alto riesgo, como en salud, vigilancia o transporte", dice, "los sistemas de IA deben ser transparentes, rastreables y garantizar la supervisión humana". Las pruebas y las certificaciones de los algoritmos adoptados también se incluyen entre las garantías que deberían establecerse y deberían generalizarse, también para "cosméticos, coches o juguetes". Mientras que los "sistemas de menor riesgo" solo tienen que seguir sistema de etiquetado voluntario: "Los operadores económicos en cuestión recibirían una etiqueta de calidad para sus aplicaciones de IA".

Pero los críticos notaron que la definición misma de "riesgo" en el documento es circular y demasiado vaga, lo que permite que varios sistemas de ADM con gran capacidad de impacto sean pasados por alto en el marco propuesto[14].

Comentarios[15] recolectados por una consulta pública, entre febrero y junio de 2020, recalcan lo controvertida que es esta idea. El 42,5% de los encuestados estuvo de acuerdo en que los "requisitos obligatorios" deberían limitarse a las "aplicaciones de IA de alto riesgo", mientras que el 30,6% dudaba de tal limitación.

Además, no existe una descripción de un mecanismo claro para el cumplimiento de tales requisitos. Y tampoco hay ninguna descripción de un proceso para avanzar hacia ellos.

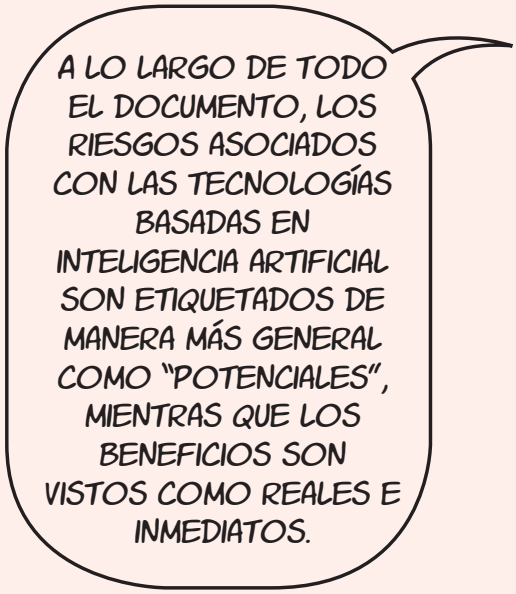
Las consecuencias son evidentes para las tecnologías biométricas y, en particular, para el reconocimiento facial. Sobre esta cuestión, el Libro Blanco propuso una distinción entre la "autenticación" biométrica, que es vista como controvertida (por ejemplo, el reconocimiento facial para desbloquear un teléfono móvil), y la "identificación" biométrica remota (como los sistemas implementados en plazas públicas para identificar a los manifestantes), lo que podría suscitar problemas en materia de derechos humanos y privacidad.

Solo los casos de la última categoría serían problemáticos con el sistema propuesto por la UE. Las Preguntas Frecuentes en el Libro Blanco sostienen que: "esta es la forma más invasiva de reconocimiento facial y, en principio, está prohibida en la UE", a menos que exista un "interés público sustancial" para que sea implementada.

El documento explicativo afirma que "permitir el reconocimiento facial es actualmente la excepción", pero se puede mantener que los hallazgos de este informe contradicen esa afirmación: el reconocimiento facial parece estar convirtiéndose rápidamente en la norma. Un boceto filtrado del Libro Blanco parecía reconocer la urgencia del problema, al incluir la idea de una moratoria de tres a cinco años sobre los usos (en vivo) del reconocimiento facial en lugares públicos, hasta que – si se da el caso – se encuentre una

14 "Dos ejemplos singulares: VioGén, un sistema de ADM para pronosticar la violencia de género, y Ghostwriter, una aplicación para detectar el fraude de exámenes, muy probablemente serían pasados de alto por la reglamentación, aunque conllevan tremendos riesgos" (https://algorithmwatch.org/en/response-european-commission-ai-consultation/)

15 "En total, se recibieron 1215 contribuciones, de las cuales 352 fueron en nombre de una empresa u organizaciones/asociaciones empresariales, 406 de ciudadanos (92% ciudadanos de la UE), 152 en nombre de instituciones académicas/de investigación y 73 de parte de autoridades públicas. Las voces de la sociedad civil estuvieron representadas por 160 encuestados (entre los cuales 9 organizaciones de consumidores, 129 organizaciones no gubernamentales y 22 sindicatos). 72 encuestados contribuyeron como "otros"". Los comentarios llegaron "de todas las partes del mundo", incluidos países como "India, China, Japón, Siria, Irak, Brasil, México, Canadá, Estados Unidos y el Reino Unido". (del Consultation's Summary Report, enlazado aquí: https://ec.europa.eu/digital-single-market/en/news/white-paper-artificial-intelligence-public-consultation-towards-european-approach-excellence)



forma de reconciliación con controles y equilibrios democráticos.

Justo antes de la publicación oficial del Libro Blanco, hasta la Comisaria de la UE, Margrethe Vestager, pidió una "pausa" en este tipo de usos.

Sin embargo, inmediatamente después de la solicitud de Vestager los funcionarios de la Comisión sostuvieron que esta "pausa" no evitaría que los gobiernos nacionales utilicen el reconocimiento facial de acuerdo con las reglas existentes. En última instancia, el borrador final del Documento descartó cualquier mención de una moratoria y pidió "un amplio debate europeo sobre las circunstancias específicas, si las hubiera, que podrían justificar" su uso para fines de identificación biométrica en vivo. Al respecto, el Libro Blanco incluye: justificación, proporcionalidad, la existencia de garantías democráticas y el respeto de los derechos humanos.

A lo largo de todo el documento, los riesgos asociados con las tecnologías basadas en inteligencia artificial son etiquetados de manera más general como "potenciales", mientras que los beneficios son vistos como reales e inmediatos. Esto llevó a muchos actores[16] en la comunidad de derechos humanos a afirmar que la narrativa general del Libro Blanco sugiere una preocupante inversión de las prioridades de la UE, anteponiendo la competitividad global a la protección de los derechos fundamentales.

Sin embargo, en los documentos sí se plantean algunas cuestiones fundacionales. Por ejemplo, la interoperabilidad de tales soluciones y la creación de una red de centros de investigación enfocados en aplicaciones de IA que apunten a la "excelencia" y al desarrollo de competencias.

El objetivo es "atraer más de 20.000 millones de euros de inversiones anuales en la UE destinadas a la IA durante la próxima década".

Un cierto determinismo tecnológico parece afectar también al Libro Blanco. "Es esencial", dice, "que las administraciones públicas, los hospitales, los servicios públicos y de transporte, los supervisores financieros y otras áreas de interés público comiencen rápidamente a implementar productos y servicios que se basen en la IA para sus propias actividades. Las áreas de salud y transporte tendrán un enfoque específico, es donde la tecnología ya está madura para su implementación a gran escala".

Sin embargo, queda por ver si sugerir una implementación apresurada de soluciones de ADM en todas las esferas de la actividad humana es compatible con los esfuerzos de la Comisión Europea para afrontar los desafíos estructurales provocados por los sistemas de ADM a los derechos y la justicia.

/ Resolución del Parlamento de la UE sobre los sistemas de ADM y la protección del consumidor

Una Resolución, aprobada por el Parlamento de la UE en febrero de 2020, afrontó más específicamente los sistemas de ADM en el contexto de la protección del consumidor. La Resolución señaló correctamente que "sistemas complejos de algoritmos y los procesos automatizados de toma de decisiones están siendo creados con rapidez", y que "las oportunidades y los desafíos que entrañan estas técnicas son numerosos y repercuten sobre prácticamente todos los sectores del mercado interior". El texto también destaca la necesidad de "llevar a cabo un estudio del marco jurídico vigente de la Unión", para evaluar si "es capaz de dar respuesta al surgimiento de la IA y la toma de decisiones automatizada".

16 Entre ellos: Access Now (https://www.accessnow.org/cms/assets/uploads/2020/05/EU-white-paper-consultation_AccessNow_May2020.pdf), AI Now (https://ainowinstitute.org/ai-now-comments-to-eu-whitepaper-on-ai.pdf), EDRI (https://edri.org/can-the-eu-make-ai-trustworthy-no-but-they-can-make-it-just/) — y AlgorithmWatch (https://algorithmwatch.org/en/response-european-commission-ai-consultation/).

SI LA "INTELIGENCIA ARTIFICIAL" ES DE HECHO UNA REVOLUCIÓN QUE REQUIERE UN PAQUETE LEGISLATIVO ESPECÍFICO, QUE SUPUESTAMENTE LLEGARÁ DURANTE EL PRIMER TRIMESTRE DE 2021, LOS REPRESENTANTES ELECTOS QUIEREN TENER VOZ AL RESPECTO.

En un llamado a un "planteamiento común de la Unión en cuanto a la obtención de procesos automatizados de toma de decisiones", la Resolución detalla varias condiciones que cualquier sistema de este tipo debería poseer para seguir siendo consistente con los valores europeos. Los consumidores deben estar "debidamente informados" sobre cómo los algoritmos afectan a sus vidas, y deben tener la manera de contactar con un ser humano con poder de decisión y para verificar y corregir las decisiones del sistema si es necesario. También deberían ser informados, "cuando los precios de los bienes o servicios se hayan personalizado basándose en la toma de decisiones automatizada y la elaboración de perfiles del comportamiento de los consumidores".

Al recordar a la Comisión de la UE que se necesita un planteamiento basado en el riesgo cuidadosamente elaborado, la Resolución señala que las garantías deben tener en cuenta que los sistemas de ADM "pueden evolucionar y actuar de manera distinta a la prevista en el momento inicial de comercialización", y que La responsabilidad no siempre es fácil de atribuir cuando se produce un daño como resultado de la implementación de un sistema de ADM.

La resolución recuerda el artículo 22 del RGPD cuando nota que un sujeto humano debe estar siempre informado cuando "estén en juego intereses públicos legítimos", y siempre debe ser el responsable último de las decisiones en "servicios profesionales como los de carácter médico, jurídico o contable, así como en el sector bancario". En particular, una evaluación de riesgos "adecuada" debe preceder a cualquier automatización de los servicios profesionales.

Finalmente, la Resolución enumera los requisitos detallados de calidad y transparencia en la gobernanza de datos: entre ellos, "la importancia de utilizar únicamente conjuntos de datos no sesgados y de calidad, con el fin de mejorar los resultados de los sistemas algorítmicos y de reforzar la confianza y la aceptación de los consumidores"; utilizando "algoritmos explicables e imparciales"; y la necesidad de una "estructuras de revisión" que permita a los consumidores afectados "solicitar la revisión y rectificación humanas de aquellas decisiones automatizadas que sean definitivas y permanentes".

/ Aprovechar al máximo el "derecho de iniciativa" del Parlamento de la UE

En su discurso inaugural, Von der Leyen expresó claramente su apoyo al "derecho de iniciativa" del Parlamento Europeo. "Cuando esta Cámara, actuando en nombre de la mayoría de sus miembros, adopte Resoluciones solicitando a la Comisión que presente propuestas legislativas", dijo, "me comprometo a responder con un acto legislativo en el pleno respeto de los principios de proporcionalidad, subsidiariedad y mejor legislación ".

Si la "Inteligencia Artificial" es de hecho una revolución que requiere un paquete legislativo específico, que supuestamente llegará durante el primer trimestre de 2021, los representantes electos quieren tener voz al respecto. Esto, junto con la intención declarada de Von der Leyen de potenciar sus capacidades legislativas, podría incluso resultar en lo que Politico ha etiquetado como un "momento parla-

mentario", con comisiones parlamentarias dedicándose a redactar varios informes diferentes en consecuencia.

Cada informe investiga aspectos específicos de la automatización en políticas públicas que, aunque están destinados a dar forma a la próxima legislación sobre "IA", también son relevantes para los sistemas de ADM.

Por ejemplo, a través de su "Marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas", la Comisión de Asuntos Jurídicos pide la creación de una "Agencia Europea de Inteligencia Artificial" y, al mismo tiempo, de una red de autoridades nacionales de control en cada Estado miembro para garantizar que las decisiones éticas que implican la automatización sean y sigan siendo éticas.

En "Derechos de propiedad intelectual para el desarrollo de las tecnologías relativas a la inteligencia artificial", la misma Comisión expone su visión del futuro de la propiedad intelectual y la automatización. Por un lado, el proyecto de informe establece que "métodos matemáticos están excluidos de la patentabilidad, a no ser que constituyan invenciones de carácter técnico", mientras que al mismo tiempo afirma que, en lo que respecta a la transparencia algorítmica, "la ingeniería inversa constituye una excepción a la norma de secretos comerciales".

El informe llega a plantearse cómo proteger "las creaciones técnicas y artísticas generadas por la IA, con el fin de fomentar esta forma de creación", imaginando que "determinadas obras generadas por la IA pueden asimilarse a obras intelectuales y que, por lo tanto, podrían estar protegidas por derechos de autor".

Por último, en un tercer documento ("Inteligencia artificial y responsabilidad civil"), la Comisión detalla un "Planteamiento de gestión de riesgos" para la responsabilidad civil de las tecnologías de IA. Según el documento, "la parte que es más capaz de controlar y gestionar un riesgo relacionado con la tecnología es estrictamente responsable, como el único punto de partida para un litigio legal".

Principios importantes sobre el uso de los procesos de ADM en el sistema de justicia penal se pueden encontrar en el informe sobre "La inteligencia artificial en el Derecho penal y su utilización por las autoridades policiales y judiciales en asuntos penales" de la Comisión de Libertades Civiles, Justicia y Asuntos de Interior. Después de una lista detallada de los usos actuales y reales de la "IA" (estos son, en realidad, sistemas de ADM) por parte de las fuerzas policiales[17], el Comité "considera necesario crear un régimen claro y justo para establecer la responsabilidad jurídica de las posibles consecuencias adversas derivadas de estas tecnologías digitales avanzadas".

Luego pasa a detallar algunas de sus características: que no haya decisiones completamente automatizadas[18], resultados algorítmicos que sean "inteligibles para los usuarios", una "evaluación obligatoria del impacto sobre los derechos fundamentales (...) de cualquier sistema de IA en el ámbito policial o judicial" antes de su implementación o adopción, además de "auditorías periódicas obligatorias de todos los sistemas de IA utilizados por las autoridades policiales y judiciales para poner a prueba y evaluar los sistemas algorítmicos una vez que estén en funcionamiento".

En el informe también se pide una moratoria sobre las tecnologías de reconocimiento facial para la aplicación de la ley, "hasta que las normas técnicas puedan considerarse plenamente acordes con los derechos fundamentales, los resultados obtenidos no sean discriminatorios y exista confianza por parte de los ciudadanos con respecto a la necesidad y proporcionalidad de la implementación de estas tecnologías".

El objetivo es lograr impulsar finalmente la transparencia general de dichos sistemas, y aconsejar a los Estados miembros que proporcionen una "comprensión global" de los sistemas de inteligencia artificial adoptados por las fuerzas del orden y el poder judicial, y - junto con la idea de un "registro público" - que detallen el "tipo de herramienta utilizada, tipos de delito a los que se aplica y empresas cuyas herramientas se utilizan".

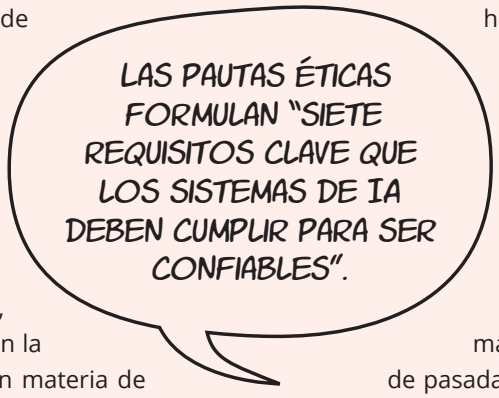
La Comisión de Cultura y Educación y la Comisión de Política Industrial también se encontraban trabajando en sus

17 En p. 5, el informe establece: "Las aplicaciones de IA utilizadas por las autoridades policiales incluyen, por ejemplo, las tecnologías de reconocimiento facial, el reconocimiento automático de matrículas, la identificación por voz, el reconocimiento del habla, las tecnologías de lectura de labios, la vigilancia auditiva (es decir, algoritmos de detección de disparos), la investigación y el análisis autónomos de bases de datos identificadas, la predicción (actuación policial predictiva y análisis de puntos críticos de delincuencia), los instrumentos de detección del comportamiento, las herramientas autónomas para detectar fraudes financieros y la financiación del terrorismo, la vigilancia de las redes sociales (rastreo [scraping] y recopilación de datos para detectar conexiones), los captadores de identidad internacional de abonados móviles (IMSI) y los sistemas automatizados de vigilancia que incorporan diferentes capacidades de detección (como la detección del latido cardíaco y las cámaras térmicas)".

18 "En contextos jurídicos y de aplicación de la ley, la decisión final siempre debe ser tomada por un ser humano" (p. 6)

propios informes a fecha de redacción de este documento.

Todas estas iniciativas llevaron a la creación de una Comisión Especial sobre "Inteligencia Artificial en la Era Digital" (AIDA) el 18 de junio de 2020. Compuesta por 33 miembros, y con una duración inicial de 12 meses, "analizará el impacto futuro" de la IA en la economía de la UE, y "en particular en materia de competencias, empleo, tecnología financiera, educación, salud, transporte, turismo, agricultura, medio ambiente, defensa, industria, energía y administración electrónica".



/ Grupo de expertos de alto nivel sobre IA y Alianza sobre IA

En 2018, el Grupo de alto nivel de expertos (HLEG) sobre IA, una comisión formada por 52 expertos, fue establecido por la Comisión Europea para apoyar la implementación de la Estrategia Europea sobre la IA, para identificar principios que deben ser observados con el fin de obtener una "IA confiable", y, en tanto que junta directiva de la Alianza sobre IA, con el objetivo de crear una plataforma abierta de múltiples partes interesadas (que consta de más de 4.000 miembros en el momento de redactar este documento) que proporcione un punto de partida más amplio para el trabajo del Grupo de expertos de alto nivel en IA.

Después de la publicación del primer boceto de las Directrices de ética de la IA para una IA fiable en diciembre de 2018, seguidas por los comentarios de más de 500 colaboradores, fue publicada una nueva versión en abril de 2019. Presenta un "enfoque centrado en el ser humano" para lograr una IA legal, ética y sólida a lo largo de todo el ciclo de vida del sistema. Sin embargo, sigue tratándose de un marco voluntario sin recomendaciones concretas y aplicables sobre puesta en marcha, implementación y cumplimiento.

Organizaciones de la sociedad civil, de protección del consumidor y derechos comentaron y pidieron la traducción de las directrices en derechos tangibles para las personas[19]. Por ejemplo, Access Now, una organización sin ánimo de lucro que trabaja sobre los derechos digitales, y miembro del HLEG, instó a la CE a aclarar cómo las diferentes partes interesadas pueden probar, aplicar, mejorar, respaldar y hacer cumplir la "IA confiable", reconociendo al mismo tiempo la necesidad de determinar los límites de Europa

En un artículo de opinión, otros dos miembros del HLEG afirmaron que el grupo había "trabajado durante un año y medio, solo para que sus propuestas detalladas fueran en su mayoría ignoradas o mencionadas solo de pasada" por la Comisión Europea que redactó la versión final.[20] También afirmaron que, debido a que el grupo inicialmente tenía la tarea de identificar los riesgos y las "líneas rojas" para la IA, los miembros del grupo señalaron los sistemas autónomos de armas, el puntaje a los ciudadanos y la identificación automatizada de personas mediante el uso del reconocimiento facial como implementaciones de la IA que deberían evitarse. Sin embargo, representantes de la industria, que dominan la comisión[21], lograron eliminar estos principios antes de que se publicara el borrador.

Este desequilibrio para resaltar el potencial de los sistemas de ADM en comparación con los riesgos también se puede observar en su segunda entrega. En las "Recomendaciones de política e inversión para una inteligencia artificial confiable en Europa" del HLEG, publicadas en junio de 2019, hay 33 recomendaciones destinadas a "orientar la IA confiable hacia la sostenibilidad, el crecimiento y la competitividad, así como a la inclusión, y que al mismo tiempo empodere, beneficie y proteja a los seres humanos". El documento es por lo general un llamado a impulsar la adopción y el aumento de la IA en el sector público y privado mediante la inversión en herramientas y aplicaciones "para ayudar a los grupos demográficos vulnerables" y "no dejar a nadie atrás".

Sin embargo, y a pesar de todas las críticas legítimas, ambas directrices siguen expresando preocupaciones y exigencias críticas en relación con los sistemas automatizados de toma de decisiones. Por ejemplo, las pautas éticas formulan "siete requisitos clave que los sistemas de IA deben cumplir para ser confiables". Estas pautas son una guía para la implementación práctica de cada requisito: agencia humana y

19 Por ejemplo, la Organización Europea de Protección al Consumidor (BEUC): insertar enlace a la referencia

20 Mark Coeckelbergh y Thomas Metzinger: Europe needs more guts when it comes to AI ethics, https://background.tagesspiegel.de/digitalisierung/europe-needs-more-guts-when-it-comes-to-ai-ethics

21 El grupo estaba compuesto por 24 representantes empresariales, 17 académicos, 5 organizaciones de la sociedad civil y otros 6 miembros, como la Agencia de los Derechos Fundamentales de la Unión Europea.

supervisión, solidez técnica y seguridad, privacidad y gestión de datos, transparencia, diversidad, no discriminación e imparcialidad; bienestar social y ambiental, y responsabilidad.

Las directrices también proporcionan una guía piloto concreta, llamada "Lista de evaluación de IA confiable", que tiene como objetivo la puesta en marcha de esos principios de alto nivel. El objetivo es que sea adoptado a la hora de "desarrollar, implementar o utilizar sistemas de IA" y adaptarlo "en caso de un uso específico en el que se esté aplicando el sistema".

La lista incluye muchos problemas que están asociados con el riesgo de infringir los derechos humanos a través de los sistemas de ADM. Entre ellos están la falta de agencia y supervisión humana, solidez técnica y problemas de seguridad, la incapacidad de evitar prejuicios injustos o proporcionar acceso equitativo y universal a tales sistemas, y la falta de acceso significativo a los datos que se introducen en dichos sistemas.

Contextualmente, la lista piloto incluida en las pautas proporciona preguntas útiles para ayudar a quienes implementan sistemas de ADM. Por ejemplo, pide una "evaluación del impacto sobre los derechos fundamentales en aquellos casos de usos en los que puedan producirse efectos potencialmente negativos para los derechos fundamentales". También pregunta si se han implementado "mecanismos de control y supervisión específicos" en los casos de sistemas "autónomos o con capacidad de autoaprendizaje", y si existen procesos "para asegurar la calidad y la integridad de sus datos".

Los comentarios detallados también se refieren a cuestiones fundamentales para los sistemas de ADM, como su transparencia y comprensibilidad. Las preguntas incluyen, "¿en qué medida son comprensibles las decisiones y, por tanto, el resultado producido por el sistema de IA?" y "¿en qué medida la decisión del sistema influye en los procesos de adopción de decisiones de la organización?" Estas preguntas son muy relevantes para evaluar los riesgos que plantea la implementación de dichos sistemas.

Para evitar sesgos y resultados discriminatorios, las directrices apuntan a "procesos de supervisión que permitan analizar y abordar el propósito, las restricciones, los requisitos y las decisiones del sistema de un modo claro y transparente", al tiempo que exigen la participación de las partes interesadas en todo el proceso de implementación de sistemas de IA.

Sumado a eso, las recomendaciones de política e inversión prevén la determinación de límites a través de un "diálogo institucionalizado sobre política de IA con los actores afectados", incluidos expertos de la sociedad civil. Además, instan a "prohibir la evaluación de individuos a gran escala permitida por la inteligencia artificial como se define en las Directrices de ética, y a establecer reglas muy claras y estrictas para la vigilancia con fines de seguridad nacional y otros fines que aleguen ser públicos o de interés nacional". Esta prohibición incluiría tecnologías de identificación biométrica y elaboración de perfiles.

En otro punto relevante para los sistemas automatizados de toma de decisiones, el documento también señala que "definir claramente si, cuándo y cómo la IA se puede usar (...) será crucial para hacer realidad una IA fiable", y advierte de que "cualquier forma de evaluación de los ciudadanos puede dar lugar a una pérdida de autonomía [de estos] y poner en peligro el principio de no discriminación", y por lo tanto "solamente debe utilizarse si existe una justificación clara y bajo medidas proporcionadas y justas". También destaca que "la transparencia no puede impedir la discriminación ni garantizar la imparcialidad". Esto significa que debe ofrecerse la posibilidad de optar por no participar en un mecanismo de evaluación, idealmente sin ningún detrimento para el ciudadano individual.

Por un lado, el documento reconoce que "pese a que aportan beneficios sustanciales a las personas y a la sociedad, los sistemas de IA también entrañan determinados riesgos y pueden tener efectos negativos, algunos de los cuales pueden resultar difíciles de prever, identificar o medir (por ejemplo, sobre la democracia, el estado de Derecho y la justicia distributiva, o sobre la propia mente humana)". Por otro lado, sin embargo, el grupo afirma que "una reglamentación innecesariamente preceptiva debería ser evitada".

En julio de 2020, el grupo HLEG de IA también presentó su versión definitiva de la Lista de Evaluación para una Inteligencia Artificial Fiable (ALTAI), creada después de un proceso piloto con aportaciones de más de 350 partes interesadas (entre empresas, personas e instituciones).

La lista, que es completamente voluntaria y carece de implicaciones normativas, tiene como objetivo poner en marcha los siete requisitos establecidos en las Directrices éticas de AI HLEG. La intención es proporcionar, a quienes quieran implementar soluciones de inteligencia artificial que sean compatibles con los valores de la UE – por ejemplo, diseñadores y desarrolladores de sistemas de inteligencia artifi-

cial, científicos de datos, oficiales o especialistas de adquisiciones y oficiales jurídicos/de cumplimiento –, un conjunto de herramientas de autoevaluación.

/ Consejo de Europa: cómo salvaguardar los derechos humanos en la ADM

Además del Comité Ad Hoc sobre Inteligencia Artificial (CAHAI), creado en septiembre de 2019, el Comité de Ministros del Consejo de Europa[22] ha publicado un marco sustancial y persuasivo.

Concebido como un instrumento normativo, su "Recomendación a los Estados miembros sobre el impacto de los sistemas algorítmicos en los derechos humanos" describe[23] "desafíos importantes" que surgen con la manifestación de nuestra "creciente dependencia" de dichos sistemas, y que son relevantes "para las sociedades democráticas y el estado de derecho".

El marco, que fue sometido a un período de consulta pública con observaciones detalladas de las organizaciones de la sociedad civil, va más allá del Libro Blanco de la Comisión Europea en cuanto a garantizar los valores y los derechos humanos.

La Recomendación analiza minuciosamente los efectos y las configuraciones en desarrollo de los sistemas algorítmicos (Apéndice A) centrándose en todas las etapas del proceso que intervienen en la elaboración de un algoritmo, es decir, la adquisición, el diseño, el desarrollo y e implementación en curso.

Aunque en general sigue el enfoque de una 'IA centrada en el ser humano' de las directrices del HLEG, la Recomendación describe "obligaciones de los Estados" como procesables (Apéndice B), así como las responsabilidades de los actores del sector privado (Apéndice C). Además, la Recomendación agrega principios como "autodeterminación informativa"[24], enumera sugerencias detalladas para mecanismos de rendición de cuentas y recursos eficaces, y exige evaluaciones del impacto sobre los derechos humanos.

Aunque el documento reconoce claramente el "potencial significativo de las tecnologías digitales para afrontar los desafíos sociales, así como para la innovación y el desarrollo económico socialmente beneficiosos", también insta a la precaución. Su objetivo es asegurar que estos tipos de sistemas no perpetúen deliberada o accidentalmente "los desequilibrios raciales, de género y otros desequilibrios sociales y laborales que aún no se han eliminado de nuestras sociedades".

Por el contrario, los sistemas algorítmicos deben usarse de manera proactiva y sensible para abordar estos desequilibrios y prestar "atención a las necesidades y voces de los grupos más vulnerables".

Sin embargo, lo más significativo es que la Recomendación identifica el riesgo potencialmente mayor para los derechos humanos cuando los Estados miembros utilizan sistemas algorítmicos con el fin de prestar servicios públicos y aplicar sus políticas. Dado que es imposible para un individuo optar por no participar, al menos sin afrontar las consecuencias negativas de hacerlo, se necesitan precauciones y garantías para el uso de sistemas de ADM en el gobierno y la administración.

La Recomendación también aborda los conflictos y desafíos que surgen de las asociaciones público-privadas ("ni claramente públicas ni claramente privadas") en una amplia gama de usos.

Las recomendaciones para los gobiernos de los Estados miembros incluyen el abandono de procesos y la negación a utilizar sistemas de ADM, si "el control y la supervisión humanos se vuelven impracticables" o se ponen en riesgo los derechos humanos; implementar sistemas de ADM si

22 El CoE es a la vez "un organismo gubernamental donde los enfoques nacionales de los problemas europeos se debaten a la par y un foro para encontrar respuestas colectivas a estos desafíos". Su trabajo incluye "los aspectos políticos de la integración europea, salvaguardar las instituciones democráticas y el estado de derecho y proteger los derechos humanos; en otras palabras, todos los problemas que requieren soluciones paneuropeas coordinadas". Si bien las recomendaciones a los gobiernos de los miembros no son vinculantes, en algunos casos el Comité puede solicitar a los gobiernos de los miembros que se le informe de las medidas tomadas por ellos con respecto a tales recomendaciones (Art. 15b Estatuto). Las relaciones entre el Consejo de Europa y la Unión Europea se establecen en el (1) Compendio de textos que rigen las relaciones entre el Consejo de Europa y la Unión Europea y en el (2) Memorando de entendimiento entre el Consejo de Europa y la Unión Europea.

23 Bajo la supervisión del Comité Director de los Medios de Comunicación y la Sociedad de la Información (CDMSI) y elaborado por el Committee of Experts on Human Rights Dimensions of Automated Data Processing and Different Forms of Artificial Intelligence (MSI-AUT)

24 "Los Estados deben garantizar que todos los diseños, desarrollos e implementación en curso de sistemas algorítmicos permitan una vía para que las personas estén informadas con anticipación sobre el procesamiento de datos relacionado (incluidos sus propósitos y posibles resultados) y para controlar sus datos, incluso mediante la interoperabilidad", dice la Sección 2.1 del Apéndice B.

y sólo si se puede garantizar la transparencia, la rendición de cuentas, la legalidad y la protección de los derechos humanos "en todas las etapas del proceso". Además, el seguimiento y la evaluación de estos sistemas deben ser "constantes", "inclusivos y transparentes", y estar compuestas de un diálogo con todas las partes interesadas pertinentes, así como de un análisis del impacto ambiental y otras posibles externalidades sobre "poblaciones y entornos".

En el Apéndice A, el CoE también define algoritmos de "alto riesgo" para que otros organismos se inspiren en esos ejemplos. Más específicamente, la Recomendación establece que "el término 'alto riesgo' sea aplicado cuando se hace referencia al uso de sistemas algorítmicos en procesos o decisiones que pueden producir graves consecuencias para las personas o en situaciones donde la falta de alternativas genera una probabilidad particularmente alta de violación de los derechos humanos, inclusive mediante la introducción o amplificación de la injusticia distributiva".

El documento, que no requirió la unanimidad de los miembros para su adopción, no es vinculante.

/ Regulación de contenidos terroristas en línea

Después de un largo período en el que el progreso fue muy lento, el reglamento para la prevención de la difusión de contenidos terroristas en línea cobró impulso en 2020. Si la regulación adoptada aún incluye herramientas automatizadas y proactivas para reconocer y eliminar contenido en línea, estas probablemente caerían bajo el Artículo 22 del RGPD.

Como dice el Supervisor Europeo de Protección de Datos (SEPD): "dado que las herramientas automatizadas, según lo previsto en la propuesta, podrían conducir no solo a la eliminación y retención de contenido (y sus datos) relacionados con la persona que subió el video, sino también, en última instancia, a investigaciones penales sobre él o ella, estas herramientas afectarían significativamente a esta persona, incidiendo en su derecho a la libertad de expresión y planteando importantes riesgos para sus derechos y libertades", y por lo tanto corresponden al art. 22(2).

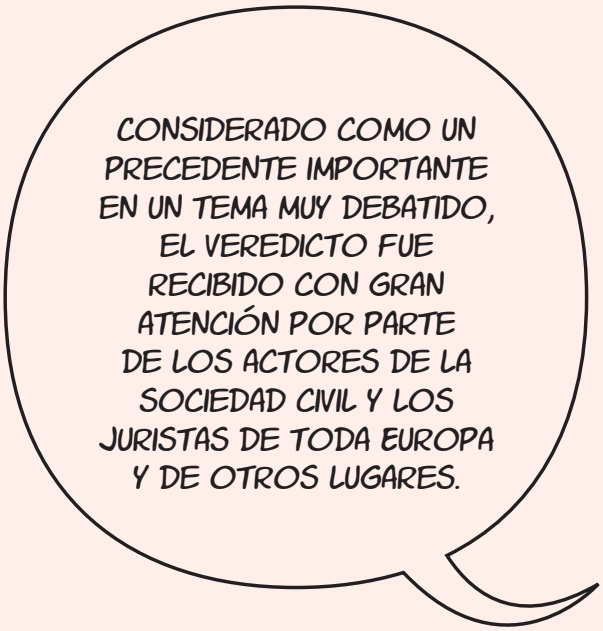
También, y de manera crucial, el uso de estas herramientas requeriría garantías más concretas en comparación con las que prevé actualmente la Comisión. Como explica el grupo de defensa European Digital Rights (EDRi): "la propuesta de Reglamento sobre contenido terrorista necesita una reforma sustancial para estar a la altura de los valores de la Unión y garantizar los derechos y libertades fundamentales de sus ciudadanos".

Una temprana corriente de fuertes críticas a la propuesta inicial por parte de grupos de la sociedad civil, las comisiones del Parlamento Europeo (PE), incluidos los dictámenes y análisis de la Agencia de los Derechos Fundamentales de la Unión Europea (FRA), EDRi, y también un informe conjunto crítico de tres relatores especiales de la ONU, destacaron las amenazas al derecho a la libertad de expresión e información, al derecho a la libertad y al pluralismo de los medios de comunicación, a la libertad para realizar negocios, y a los derechos a la privacidad y protección de datos personales.

Los aspectos críticos incluyen una definición insuficiente que sepa definir qué es contenido terrorista, el objetivo de la regulación (en la actualidad, esta incluye contenido para educación y propósitos periodísticos), las anteriormente mencionadas "medidas proactivas", la falta de supervisión judicial efectiva, las obligaciones de notificación insuficientes para los organismos encargados de hacer cumplir la ley, y la falta de garantías para "casos en los que hay motivos razonables para creer que los derechos fundamentales se ven afectados" (EDRi 2019) .

El SEPD destaca que tales "salvaguardias adecuadas" deben incluir el derecho a obtener la intervención humana y el derecho a una explicación de la decisión adoptada por medios automatizados (EDRi 2019).

Aunque las garantías sugeridas o exigidas se abrieron paso en el borrador de informe del PE sobre la propuesta, aún está por verse quién puede resistir más tiempo antes de la votación final. Durante los diálogos tripartitos a puerta cerrada entre el PE, la nueva CE y el Consejo de la Unión (que comenzaron en octubre de 2019), solo serán posibles pequeños cambios, según un documento filtrado.



Supervisión y regulación

/ Primeras decisiones sobre el cumplimiento de los sistemas de ADM con el RGPD

"Aunque no hubo un gran debate sobre el reconocimiento facial durante las negociaciones sobre el RGPD y la aplicación de la directiva de protección de datos, la legislación fue diseñada para que pudiera adaptarse con el tiempo a medida que evolucionaban las tecnologías. [...] Ahora es el momento de que la UE, ya que discute sobre la ética de la IA y la necesidad de una reglamentación, determine si – y si en algún momento – se podrá o no permitir la tecnología de reconocimiento facial en una sociedad democrática. Si la respuesta es sí, solo entonces pasamos a las preguntas sobre cómo implementar las garantías y la responsabilidad". -SEPD, Wojciech Wiewiórowski

"El procesamiento de reconocimiento facial es un mecanismo biométrico especialmente invasivo, que conlleva importantes riesgos de invasión de la privacidad o las libertades civiles para las personas afectadas" - (CNIL 2019)

Desde el último informe de Automating Society, hemos visto los primeros casos de multas y decisiones relacionadas con incumplimientos de la normativa emitida por las Autoridades nacionales de Protección de Datos (DPA) basadas en en el RGPD. Los siguientes estudios de caso, sin embargo, muestran los límites del RGPD en la práctica cuando se trata del Artículo 22 relacionados con los sistemas de ADM y de cómo está dejando que los reguladores de privacidad realicen evaluaciones caso por caso.

En Suecia, se descubrió que un proyecto de prueba de reconocimiento facial, realizado en la clase de una escuela durante un período limitado de tiempo, violaba varias obligaciones del Reglamento de protección de datos (esp. RGPD art. 2 (14), art. 9 (2)). (European Data Protection Board 2019)

Un caso similar está en suspensión después de que la Comisión Nacional de la Información y las Libertades (CNI) de Francia expresara su preocupación cuando dos escuelas secundarias planeaban introducir tecnología de reconocimiento facial en sociedad con la empresa de tecnología estadounidense Cisco. La opinión no es vinculante y la demanda presentada está en curso.[25].

No se requiere autorización p por los reguladores de datos para realizar tales ensayos, ya que generalmente se considera que el consentimiento de los usuarios es suficiente para procesar datos biométricos. Y, sin embargo, en el caso sueco no lo fue. Esto se debió a los desequilibrios de poder entre el responsable del tratamiento de datos y los sujetos de los datos. En su lugar, se consideró necesaria una evaluación de impacto adecuada y una consulta previa con la APD.

El Supervisor Europeo de Protección de Datos (SEPD) confirmó esa consideración:

"El consentimiento debería ser explícito, así como otorgado libremente, informado y específico. Sin embargo, es indudable que una persona no pueda optar por no participar, y menos aún optar por participar, cuando necesitan acceder a espacios públicos que están cubiertos por la vigilancia de reconocimiento facial. [...] Por último, el cumplimiento de la tecnología respecto a principios como la minimización de datos y la obligación de protección de datos por diseño es altamente dudoso. La tecnología de reconocimiento facial nunca ha sido completamente precisa, y esto tiene serias consecuencias para las personas que son identificadas falsamente, ya sea como delincuentes o no. [...] Sin embargo, sería un error centrarse sólo en cuestiones de privacidad. Esta es fundamentalmente una cuestión ética para una sociedad democrática ". (SEPD 2019)

25 Véase el capítulo de Francia y (Kayalki 2019)

Access Now comentó:

"A medida que se desarrollan más proyectos de reconocimiento facial, ya vemos que el RGPD proporciona garantías útiles para los derechos humanos que se pueden hacer cumplir contra la recopilación y el uso ilegal de datos sensibles como la biometría. Pero la exageración irresponsable y, a menudo, infundada en torno a la eficiencia de tales tecnologías, así como el interés económico subyacente podría llevar a intentos por parte de los gobiernos (central y local) y las empresas privadas a eludir la ley ".

/ Reconocimiento facial automatizado en uso por la policía de Gales del Sur declarado ilegal

En el transcurso de 2020, El Reino Unido fue testigo de la primera aplicación de alto perfil de la Directiva[26] sobre el uso de tecnologías de reconocimiento facial en espacios públicos por parte de la policía. Considerado como un precedente importante en un tema muy debatido, el veredicto fue recibido con gran atención por parte de los actores de la sociedad civil y los juristas de toda Europa y de otros lugares.[27].

El caso fue llevado a la corte por Ed Bridges, un hombre de 37 años de Cardiff, quien afirmó que su rostro fue escaneado sin su consentimiento tanto durante las compras navideñas de 2017 como en una protesta pacífica contra las armas un año después.

El tribunal inicialmente confirmó el uso de tecnología de reconocimiento facial automatizado ("AFR") de la Policía de Gales del Sur, declarándolo lícito y proporcionado. Pero la decisión fue apelada por Liberty, un grupo de derechos civiles, y el Tribunal de Apelación de Inglaterra y Gales decidió anular el rechazo del Tribunal Superior y lo dictaminó ilegal el 11 de agosto de 2020.

Al fallar contra la Policía de Gales del Sur por tres de los cinco motivos, el Tribunal de Apelación encontró "deficiencias fundamentales" en el marco normativo existente en torno al uso de AFR, afirmando que su implementación no cumplía con el principio de "proporcionalidad" y que, además, no se había realizado una Evaluación de Impacto de Protección de Datos (EIPD) adecuada, por lo que no se habían dado varios pasos muy importantes.

Sin embargo, el tribunal no dictaminó que el sistema estuviera produciendo resultados discriminatorios, ya fuera por motivos de sexo o raza, ya que la policía de Gales del Sur no había reunido pruebas suficientes para emitir un juicio al respecto.[28] Aunque el tribunal sí consideró que valía la pena agregar un comentario notable: "Esperamos que, dado que el AFR es una tecnología novedosa y controvertida, todas las fuerzas policiales que tengan la intención de usarla en el futuro deseen asegurarse de que todas las medidas razonables que puedan ser hechas se hayan hecho para asegurarse de que el software utilizado no tenga un sesgo racial o de género".

Tras el fallo, Liberty pidió que la Policía de Gales del Sur y otras fuerzas policiales retiren el uso de tecnologías de reconocimiento facial.

ADM en la práctica: gestión y vigilancia de fronteras

Mientras la Comisión Europea y sus partes interesadas debatían si reglamentar o prohibir las tecnologías de reconocimiento facial, ya se estaba experimentando ampliamente con dichos sistemas en toda Europa.

Esta sección destaca un vínculo crucial y a menudo pasado por alto entre la biometría y los sistemas de gestión de fronteras de la UE, mostrando claramente cómo las tecnologías que pueden producir resultados discriminatorios podrían ser aplicadas alos individuos – por ejemplo, los migrantes –, que ya son quienes más sufren discriminación.

26 La Law Enforcement Directive, en vigor desde mayo de 2018, "se ocupa del procesamiento de datos personales por parte de los controladores de datos para 'fines de aplicación de la ley' - que quedan fuera del alcance del RGPD". https://www.dataprotection.ie/en/organisations/law-enforcement-directive

27 La decisión fue tomada el 4 de septiembre de 2019 por el Tribunal Superior de Cardiff en el caso Bridges v. the South Wales Police (Hight Court of Justice 2019)

28 La policía afirmó que no tenía acceso a la composición demográfica del conjunto de datos de entrenamiento para el algoritmo adoptado, "Neoface". El Tribunal señala que "el hecho es, sin embargo, que la SWP nunca ha tratado de asegurarse, ya sea directamente o mediante una verificación independiente, de que el programa de software en este caso no tuviera un sesgo inaceptable por motivos de raza o sexo".

/ Reconocimiento facial y uso de datos biométricos en las políticas y prácticas de la UE

Durante el último año, el reconocimiento facial y otros tipos de tecnologías de identificación biométrica han atraído mucha atención por parte de los gobiernos, la UE, la sociedad civil y las organizaciones de derechos, especialmente en lo que respecta a la aplicación de la ley y la gestión de fronteras.

Durante 2019, la CE encargó a un consorcio de agencias públicas que "realizara una cartografía de la situación actual del reconocimiento facial en las investigaciones penales en todos los Estados miembros de la UE", para avanzar "hacia el posible intercambio de datos faciales". Encargaron a la consultora Deloitte la realización de un estudio de viabilidad sobre la ampliación del sistema Prüm de imágenes faciales. Prüm es un sistema a gran escala de la UE que conecta bases de datos de ADN, huellas dactilares y registro de vehículos para permitir búsquedas en común. La preocupación es que una base de datos paneuropea de rostros podría usarse para una vigilancia generalizada, injustificada o ilegal.

/ Sistemas de gestión de fronteras sin fronteras

Como informamos en la edición anterior de Automating Society, la implementación de un sistema de gestión de fronteras inteligente e interoperable a nivel global en la UE, propuesto inicialmente por la Comisión en 2013, está en camino. Aunque los nuevos sistemas que se han anunciado (EES,ETIAS[29],ECRIS-TCN[30]) solo comenzarán a operar en 2022, la regulación del Sistema de Entrada / Salida (EES) ya ha introducido imágenes faciales como tipos de identificadores biométricos y también introdujo el uso de tecnología de reconocimiento facial con fines de verificación por primera vez en la legislación de la UE.[31]

La Agencia Europea de Derechos Fundamentales (FRA) confirmó estos cambios: "se espera que el procesamiento de imágenes faciales se introduzca de manera más sistemática en los sistemas informáticos a gran escala en el ámbito de la UE utilizados con fines de asilo, migración y seguridad [...] una vez que se cumplan los requisitos legales y se completen los pasos técnicos".

Según Ana Maria Ruginis Andrei, de la Agencia de la Unión Europea para la Gestión Operativa de Sistemas Informáticos de Gran Magnitud en el Espacio de Libertad, Seguridad y Justicia (eu-LISA), esta nueva arquitectura de interoperabilidad ampliada fue "ensamblada para forjar el motor perfecto para luchar con éxito contra las amenazas a la seguridad interna, controlar eficazmente la migración y superar los puntos ciegos en la gestión de identidades". En la práctica, esto significa "conservar las huellas dactilares, las imágenes faciales y otros datos personales de hasta 300 millones de ciudadanos de fuera de la UE, fusionando datos de cinco sistemas separados". (Campbell 2020)

/ ETIAS: controles automáticos de seguridad fronteriza

El Sistema Europeo de Información y Autorización de Viajes (ETIAS), que todavía no estaba en funcionamiento en el momento de redactar este documento, utilizará diferentes bases de datos para automatizar el control de seguridad digital de los viajeros que no pertenecen a la UE (que no necesitan una visa o "exención de de visado") antes de que lleguen a Europa.

Este sistema recopilará y analizará datos para la "verificación avanzada de posibles riesgos de seguridad o migración irregular" (ETIAS 2020). El objetivo es "facilitar los controles fronterizos; evitar la burocracia y los retrasos de los viajeros a la hora de presentarse en las fronteras; garantizar una evaluación de riesgos coordinada y armonizada de los individuos con nacionalidades de terceros países" (ETIAS 2020).

Ann-Charlotte Nygård, directora de la unidad de Asistencia Técnica y Desarrollo de Capacidades de la FRA, ve dos ries-

29 ETIAS (EU Travel Information and Authorization System o Sistema Europeo de Información y Autorización de Viajes), es el nuevo sistema de "exención de visado" para la gestión de fronteras de la UE desarrollado por eu-LISA. "La información presentada durante la solicitud será procesada automáticamente en las bases de datos de la UE existentes (Eurodac, SIS y VIS), los futuros sistemas EES y ECRIS-TCN y las bases de datos pertinentes de Interpol. Esto permitirá la verificación anticipada de posibles riesgos para la seguridad, la migración irregular y la salud pública". (ETIAS 2019)

30 El Sistema Europeo de Información de Antecedentes Penales - Nacionales de Terceros Países (ECRIS-TCN), que será desarrollado por eu-LISA, será un sistema centralizado de éxito/no éxito para complementar la base de datos de antecedentes penales de la UE (ECRIS) ya existente sobre ciudadanos de fuera de la UE pero condenados dentro de la Unión Europea.

31 EES entrará en funcionamiento en el primer trimestre de 2022, ETIAS lo seguirá a finales de 2022 –, se proponen ser "revolucionarios en el área europea de Justicia y Asuntos de Interior (JAI)".

gos específicos con respecto a ETIAS: "primero, el uso de datos que podría conducir a la discriminación no intencional de ciertos grupos, por ejemplo, si un solicitante es de una etnia particular con alto riesgo de inmigración; el segundo se refiere a un riesgo de seguridad evaluado sobre la base de condenas pasadas en el país de origen. Algunas de estas condenas anteriores podrían ser consideradas poco razonables por los europeos, como las condenas a las personas LGBT en ciertos países. Para evitar esto, [...] los algoritmos deben ser auditados para asegurar que no discriminen, y este tipo de auditoría debería involucrar a expertos de diferentes áreas interdisciplinarias" (Nygård 2019).

/ iBorderCtrl: reconocimiento facial y evaluación de riesgo en las fronteras

iBorderCtrl fue un proyecto que involucró a agencias de seguridad de Hungría, Letonia y Grecia, que aspiraban a "permitir un control fronterizo más rápido y exhaustivo para individuos con nacionalidades de terceros países que cruzan las fronteras terrestres de los Estados miembros de la UE". iBorderCtrl usó tecnología de reconocimiento facial, un detector de mentiras y un sistema de evaluación para avisar a un policía fronterizo humano si consideraba a alguien peligroso o si consideraba que su derecho de entrada era cuestionable.

El proyecto iBorderCtrl finalizó en agosto de 2019 y los resultados, para cualquier posible implementación del sistema en toda la UE, fueron contradictorios.

Aunque "tendrá que ser definido hasta qué punto se utilizará el sistema o partes de él", la página de "Resultados" del proyecto ve "la posibilidad de integrar las funcionalidades similares del nuevo sistema ETIAS y ampliar las capacidades usadas en el procedimiento de cruce de la frontera hasta en los lugares donde se encuentren los viajeros (autobús, automóvil, tren, etc...)".

Sin embargo, los módulos a los que se refiere no se especificaron y las herramientas relacionadas con los sistemas de ADM que se probaron no fueron evaluadas públicamente.

Al mismo tiempo, la página de Preguntas más frecuentes del proyecto confirmó que el sistema que se probó no se considera "actualmente apto para su implementación en la frontera (...) debido a su naturaleza como prototipo y a la infraestructura tecnológica en el ámbito de la UE". Esto significa que "un mayor desarrollo y una integración en los sistemas de la UE existentes serían necesarios si parte de las autoridades fronterizas quisieran usarlo".

En particular, si bien el Consorcio iBorderCtrl pudo mostrar, en principio, la funcionalidad de su tecnología para los controles fronterizos, también está claro que las limitaciones éticas, legales y sociales deben abordarse antes de cualquier lanzamiento real.

/ Proyectos relacionados Horizon2020

Varios proyectos de seguimiento se dedicaron a probar y desarrollar nuevos sistemas y tecnologías para la Gestión y Vigilancia de Fronteras como parte del progra-

LA UE APORTA 8.199.387,75 DE EUROS AL PROYECTO PARA DESARROLLAR "MEJORES MÉTODOS DE VIGILANCIA DE FRONTERAS" PARA CONTRARRESTAR LA MIGRACIÓN IRREGULAR.

ma Horizon2020. Se encuentran en una lista en el sitio web de CORDIS de la Comisión Europea, que proporciona información sobre todas las actividades de investigación apoyadas por la UE relacionadas con el programa. El sitio muestra que 38 proyectos ya se están ejecutando actualmente bajo el programa/tema de la Unión Europea "H2020-EU.3.7.3. - Strengthen security through border management" (Reforzar la seguridad a través de la gestión de fronteras). El programa marco - "Secure societies - Protecting freedom and security of Europe and its citizens" (Sociedades seguras - Proteger la libertad y la seguridad de Europa y sus ciudadanos), cuenta con un presupuesto total de casi 1.700 millones de euros y financia 350 proyectos - afirma que aborda "la inseguridad, ya sea por delitos, violencia, terrorismo, desastres naturales o provocados por el hombre, ataques cibernéticos o abusos a la privacidad, y otras formas de desórdenes sociales y económicos que afectan cada vez más a los ciudadanos" mediante proyectos que desarrollan principalmente nuevos sistemas tecnológicos basado en IA y en la ADM.

Algunos proyectos ya han finalizado y/o sus aplicaciones ya están en uso – por ejemplo, FastPass, ABC4EU, MOBILEPASS y EFFISEC –, todos los cuales se interesaron por los requisitos para el "Control Fronterizo Automatizado (ABC) integrado e interoperable", sistemas de identificación y puertas de embarque "inteligentes" en diferentes pasos fronterizos.

TRESSPASS es un proyecto en curso que comenzó en junio de 2018 y finalizará en noviembre de 2021. La UE aporta casi ocho millones de euros al proyecto, y los coordinadores de iBorderCRL (así como los de FLYSEC y XP-DITE) tienen como objetivo "aprovechar los resultados y conceptos implementados y probados" por iBorderCRL y "expandirlos para una solución de seguridad multimodal basada en el riesgo de cruce de fronteras dentro de un sólido marco legal y ético". (Horizon2020 2019)

El proyecto tiene el objetivo declarado de convertir los controles de seguridad en los pasos fronterizos: de la vieja y obsoleta metodología "basada en reglas" a una nueva estrategia "basada en riesgos". Esto incluye la aplicación de tecnologías biométricas y de detección, un sistema de gestión basado en riesgos y modelos relevantes para saber evaluar la identidad, las posesiones, la capacidad y la intención. Su objetivo es permitir las comprobaciones a través de "enlaces a sistemas heredados y bases de datos externas como VIS/SIS/PNR" y está ya recopilando información de todas las fuentes de datos anteriores con fines de seguridad.

Otro proyecto piloto, FOLDOUT, comenzó en septiembre de 2018 y finalizará en febrero de 2022. La UE aporta 8.199.387,75 de euros al proyecto para desarrollar "mejores métodos de vigilancia de fronteras" para contrarrestar la migración irregular concentrándose en "detectar personas a través de follaje denso en climas extremos" [...] mediante la combinación de "varios sensores y tecnologías y fusionándolas inteligentemente en una plataforma de detección inteligente, eficaz y robusta" para sugerir diferentes escenarios con lo cuales poder reaccionar. Se están realizando proyectos piloto en Bulgaria, con modelos de demostración en Grecia, Finlandia y la Guayana Francesa.

MIRROR, o "Migration-Related Risks caused by misconceptions of Opportunities and Requirement" (Riesgos relacionados con la migración causados por conceptos erróneos de oportunidades y requisitos), comenzó en junio de 2019 y finalizará en mayo de 2022. La UE aporta algo más de cinco millones de euros al proyecto, que tiene como objetivo "comprender cómo se percibe Europa en el exterior, detectar discrepancias entre su imagen y la realidad, detectar casos de manipulación de los medios y desarrollar sus habilidades para poder contrarrestar tales conceptos erróneos y las amenazas a la seguridad resultantes de ellos". Basado en "análisis de amenazas específicas de percepción, el proyecto MIRROR combinará estudios empíricos con métodos de análisis automatizado de texto, multimedia y redes sociales para varios tipos de medios (incluidas las redes sociales)" para desarrollar "tecnología y conocimientos prácticos, [...] validados a fondo con agencias fronterizas y políticos, por ejemplo, a través de programas piloto."

Otros proyectos que ya están cerrados, pero que reciben una mención, son Trusted Biometrics under Spoofing Attacks (TABULA RASA), que empezó en noviembre de 2010 y finalizó en abril de 2014. Analizó "las debilidades del software de procesos de identificación biométrica buscando sus vulnerabilidades respecto a la suplantación de identidad, lo que disminuye la eficiencia de los dispositivos biométricos". Otro proyecto, Bodega, que comenzó en junio de 2015 y finalizó en octubre de 2018, analizó cómo utilizar la "experiencia en factores humanos" cuando se trata de "introducir sistemas de control de fronteras más inteligentes como puertas de embarque automáticas y sistemas de autoservicio basados en biometría".

Referencias

ACCESS NOW (2019). *Comments on the draft recommendation of the Committee of Ministers to Member States on the human rights impacts of algorithmic systems.* <https://www.accessnow.org/cms/assets/uploads/2019/10/Submission-on-CoE-recommendation-on-the-human-rights-impacts-of-algorithmic-systems-21.pdf>

ALGORITHMWATCH (2020). *Our response to the European Commission's consultation on AI* <https://algorithmwatch.org/en/response-european-commission-ai-consultation/>

CAMPBELL, ZACH/JONES, CHRIS (2020). *Leaked Reports Show EU Police Are Planning a Pan-European Network of Facial Recognition Databases.* <https://theintercept.com/2020/02/21/eu-facial-recognition-database/>

CNIL (2019). *French privacy regulator finds facial recognition gates in schools illegal.* <https://www.biometricupdate.com/201910/french-privacy-regulator-finds-facial-recognition-gates-in-schools-illegal>

COECKELBERGH, MARK / METZINGER, THOMAS(2020). *Europe needs more guts when it comes to AI ethics.* <https://background.tagesspiegel.de/digitalisierung/europe-needs-more-guts-when-it-comes-to-ai-ethics>

COMMITTEE OF MINISTERS (2020). *Recommendation CM/Rec(2020)1 of the Committee of Ministers to Member States on the human rights impacts of algorithmic systems.* <https://search.coe.int/cm/pages/result_details.aspx?objectid=09000016809e1154>

COMMISSIONER FOR HUMAN RIGHTS (2020). *Unboxing artificial intelligence: 10 steps to protect human rights.* <https://www.coe.int/en/web/commissioner/-/unboxing-artificial-intelligence-10-steps-to-protect-human-rights>

COMMITTEE ON LEGAL AFFAIRS (2020). *Draft Report: With recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies.* <https://www.europarl.europa.eu/doceo/document/JURI-PR-650508_EN.pdf>

COMMITTEE ON LEGAL AFFAIRS (2020). *Artificial Intelligence and Civil Liability.* <https://www.europarl.europa.eu/RegData/etudes/STUD/2020/621926/IPOL_STU(2020)621926_EN.pdf>

COMMITTEE ON LEGAL AFFAIRS (2020). *Draft Report: On intellectual property rights for the development of artificial intelligence technologies.* <https://www.europarl.europa.eu/doceo/document/JURI-PR-650527_EN.pdf>

COMMITTEE ON CIVIL LIBERTIES, JUSTICE AND HOME AFFAIRS (2020). *Draft Report: On artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters.* https://www.europarl.europa.eu/doceo/document/LIBE-PR-652625_EN.pdf

DELCKER, JANOSCH(2020). *Decoded: Drawing the battle lines — Ghost work — Parliament's moment.* <https://www.politico.eu/newsletter/ai-decoded/politico-ai-decoded-drawing-the-battle-lines-ghost-work-parliaments-moment/?utm_source=POLITICO.EU&utm_campaign=5a7d137f82-EMAIL_CAMPAIGN_2020_09_09_08_59&utm_medium=email&utm_term=0_10959edeb5-5a7d137f82-190607820>

DATA PROTECTION COMMISSION(2020). *Law enforcement directive.* <https://www.dataprotection.ie/en/organisations/law-enforcement-directive>

EDRi (2019). *FRA and EDPS: Terrorist Content Regulation requires improvement for fundamental rights.* <https://edri.org/our-work/fra-edps-terrorist-content-regulation-fundamental-rights-terreg/>

GDPR (Art 22). *Automated individual decision-making, including profiling* <https://gdpr-info.eu/art-22-gdpr/>

EUROPEAN COMMISSION (2018). *White paper: On Artificial Intelligence - A European approach to excellence and trust.* <https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf>

EUROPEAN COMMISSION (2020). *A European data strategy.* https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en

EUROPEAN COMMISSION (2020). *Shaping Europe's digital future – Questions and Answers.* https://ec.europa.eu/commission/presscorner/detail/en/qanda_20_264

EUROPEAN COMMISSION (2020). *White Paper on Artificial Intelligence: Public consultation towards a European approach for excellence and trust.* <https://ec.europa.eu/digital-single-market/en/news/white-paper-artificial-intelligence-public-consultation-towards-european-approach-excellence>

EUROPEAN COMMISSION (2018). *Security Union: A European Travel Information and Authorisation System - Questions & Answers.* <https://ec.europa.eu/commission/presscorner/detail/en/MEMO_18_4362>

EUROPEAN DATA PROTECTION BOARD (2019). *Facial recognition in school renders Sweden's first GDPR fine* <https://edpb.europa.eu/news/national-news/2019/facial-recognition-school-renders-swedens-first-gdpr-fine_en>

EUROPEAN PARLIAMENT (2020). *Artificial intelligence: EU must ensure a fair and safe use for consumers.* <https://www.europarl.europa.eu/news/en/press-room/20200120IPR70622/artificial-intelligence-eu-must-ensure-a-fair-and-safe-use-for-consumers>

EUROPEAN PARLIAMENT (2020). *On automated decision-making processes: ensuring consumer protection and free movement of goods and services* <https://www.europarl.europa.eu/doceo/document/B-9-2020-0094_EN.pdf>

EUROPEAN DATA PROTECTION SUPERVISOR (2019). *Facial recognition: A solution in search of a problem?* <https://edps.europa.eu/press-publications/press-news/blog/facial-recognition-solution-search-problem_de>

ETIAS (2020). *European Travel Information and Authorisation System (ETIAS).* <https://ec.europa.eu/home-affairs/what-we-do/policies/borders-and-visas/smart-borders/etias_en>

ETIAS (2019). *European Travel Information and Authorisation System (ETIAS)* <https://www.eulisa.europa.eu/Publications/Information%20Material/Leaflet%20ETIAS.pdf>

HORIZON2020 (2019). *robusT Risk basEd Screening and alert System for PASSengers and luggage* <https://cordis.europa.eu/project/id/787120/reporting>

HIGH COURT OF JUSTICE (2019). *Bridges v. the South Wales Police*. <https://www.judiciary.uk/wp-content/uploads/2019/09/bridges-swp-judgment-Final03-09-19-1.pdf>

HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE (2020). *Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment.* <https://ec.europa.eu/digital-single-market/en/news/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment>

HUNTON ANDREW KURTH (2020). *UK Court of Appeal Finds Automated Facial Recognition Technology Unlawful in Bridges v South Wales Police* <https://www.huntonprivacyblog.com/2020/08/12/uk-court-of-appeal-finds-automated-facial-recognition-technology-unlawful-in-bridges-v-south-wales-police/>

KAYALKI, LAURA (2019). *French privacy watchdog says facial recognition trial in high schools is illegal* <https://www.politico.eu/article/french-privacy-watchdog-says-facial-recognition-trial-in-high-schools-is-illegal-privacy/>

KAYSER-BRIL, NICOLAS (2020). *EU Commission publishes white paper on AI regulation 20 days before schedule, forgets regulation* <https://algorithmwatch.org/en/story/ai-white-paper/>

LEYEN, URSULA VON DER (2019). *A Union that strives for more - My agenda for Europe* <https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission_en.pdf>

LEYEN, URSULA VON DER (2020). *Paving the road to a technologically sovereign Europe* <https://delano.lu/d/detail/news/paving-road-technologically-sovereign-europe/209497>

LEYEN, URSULA VON DER (2020). *Shaping Europe's digital future.* https://twitter.com/eu_commission/status/1230216379002970112?s=11

LEYEN, URSULA VON DER (2019). *Opening Statement in the European Parliament Plenary Session by Ursula von der Leyen, Candidate for President of the European Commission.* https://ec.europa.eu/commission/presscorner/detail/en/SPEECH_19_4230

NYGÅRD, (2019). *The New Information Architecture as a Driver for Efficiency and Effectiveness in Internal Security.* https://www.eulisa.europa.eu/Publications/Reports/eu-LISA%20Annual%20Conference%20Report%202019.pdf

SABBAGH, DAN (2020). *This article is more than 1 month old South Wales police lose landmark facial recognition case.* <https://www.theguardian.com/technology/2020/aug/11/south-wales-police-lose-landmark-facial-recognition-case>

SOUTH WALES POLICE(2020). *Automated Facial Recognition.* <https://afr.south-wales.police.uk/>

VALERO, JORGE (2020). *Vestager: Facial recognition tech breaches EU data protection rules.* <https://www.euractiv.com/section/digital/news/vestager-facial-recognition-tech-breaches-eu-data-protection-rules/>

ESPAÑA

ARTÍCULO

PÁGINA 36

INVESTIGACIÓN

PÁGINA 42

Encontrarás más información en el artículo a partir de la página siguiente.

En España, el algoritmo VioGén intenta pronosticar la violencia de género

Como parte de un programa para frenar los feminicidios, España desarrolló VioGén, un algoritmo que evalúa el riesgo que afrontan las víctimas de violencia de género. Es un proyecto que aún necesita más desarrollo.

Por Michele Catanzaro

La madrugada del 24 de febrero de 2018, Itziar P., psicóloga residente en Castellón, acudió a una comisaría para denunciar las amenazas de su marido, Ricardo C.

En las grabaciones de sonido que había hecho con su teléfono móvil se podía escuchar al marido diciendo: "Terminaremos todos muertos y yo en la cárcel" y "Te quitaré lo que más amas".

Según Itziar P., Ricardo C. también había roto en pedazos el carro de su hija menor (Martina, de 2 años) y abofeteado a la mayor (Nerea, de 6), cuando ambas estaban bajo su custodia.

El agente de policía le hizo a Itziar P. una serie de preguntas y pasó las respuestas a VioGén, un software que ayuda a la policía a estimar el riesgo de reincidencia en la violencia de género. El agente redactó entonces un informe en el que se consideró que ese riesgo era bajo.

/ Fallo crítico

En los días siguientes, tanto Itziar P. como Ricardo C. fueron llamados a declarar ante el tribunal. Ella pidió que a él se le prohibiera visitar a sus hijas, pero el juez denegó la solicitud basándose, entre otras razones, en la evaluación de bajo riesgo hecha por la policía.

Siete meses después, el 25 de septiembre de 2018, Nerea y Martina estaban durmiendo en la casa de Ricardo C. En las primeras horas de la mañana, Ricardo C. las mató "con ensañamiento" y se tiró por una ventana.

La historia de Itziar P. fue estremecedora. ¿Por qué fue el caso considerado de bajo riesgo? VioGén había fallado en su objetivo de ayudar al personal policial a evaluar el riesgo de nuevas agresiones y, por lo tanto, a decidir un adecuado nivel de protección. Desde que se empezó a utilizar el software por primera vez en 2007, ha habido una serie de casos de "bajo riesgo" que han terminado en homicidios de mujeres o niños.

/ Mejor que nada

El programa es, de lejos, el más sofisticado de su clase en el mundo. Tiene índices de rendimiento razonables. Nadie cree que las cosas estarían mejor sin él – excepto la extrema derecha, que sostiene afirmaciones falsas, como que el software ayuda a las mujeres a denunciar a hombres inocentes.

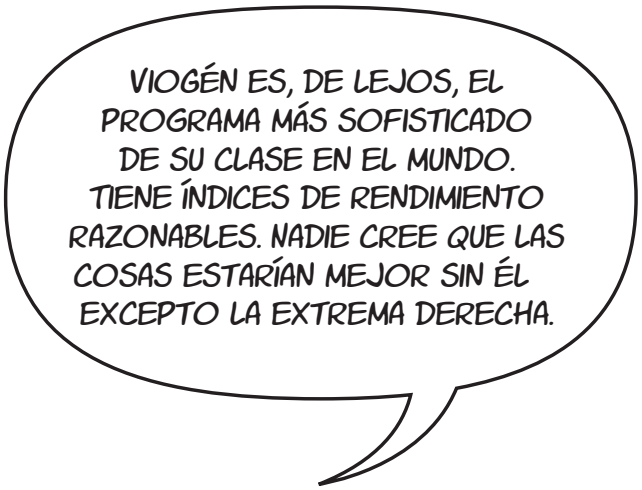
Pero hay voces críticas que señalan algunos defectos. Pocos miembros del personal policial han sido formados en cómo tratar la violencia de género, mientras que otros confían ciegamente en el resultado del software. Además, puede que el programa esté subestimando el riesgo sistemáticamente. Algunas organizaciones de víctimas creen que no tiene sentido obtener una puntuación de riesgo baja. Denunciar a la policía es de por sí una situación de alto riesgo, dicen, porque los abusadores la perciben como un desafío.

En enero de 2020, VioGén ya había evaluado 600.000 casos. Aproximadamente 61.000 de ellos tenían el estado de "activo", lo que significa que estaban bajo seguimiento de la policía (el sistema está diseñado para realizar un control periódico de las mujeres hasta que se las considere seguras).

/ Denunciar una agresión

Cuando una mujer va a denunciar una agresión por parte de su pareja, inicia un proceso que dura al menos un par de horas. Primero, el agente de policía revisa un formulario online junto con ella. El policía marca cada uno de los elementos del formulario VPR (acrónimo de "Valoración Policial de Riesgo") como "presente" o "no presente". En la última versión publicada del formulario (el VPR4.0) hay 39 elementos. Los agentes también pueden basarse en las bases de datos policiales, los testigos y las pruebas materiales.

Las preguntas giran en torno a la gravedad de las agresiones anteriores (por ejemplo, si alguna vez se usaron armas), las características del agresor (celoso, matón, abusador sexual, desempleado, drogadicto, etc.), la vulnerabilidad de la víctima (embarazada, extranjera, dependiente económicamente, etc.), y factores agravantes (como agresiones por parte de otros hombres).



Las respuestas son introducidas automáticamente en una fórmula matemática que calcula un resultado y mide el riesgo de que el agresor repita acciones violentas. Este enfoque cuantitativo es diferente al que utiliza DAS-H, el equivalente británico de VioGén. Este último es básicamente una lista de parámetros a evaluar que ayuda a los agentes a hacerse una idea de la situación.

/ Manteniendo el resultado

En teoría, los agentes pueden aumentar el resultado manualmente si constatan que existe un riesgo mayor. Pero un estudio de 2014 reveló que en el 95% de los casos los policías se atuvieron al resultado automático.

La fórmula utilizada en VioGén es un "algoritmo simple", según Juan José López Ossorio, psicólogo que ha estado a cargo de VioGén desde sus primeras etapas, según señaló por escrito a AlgorithmWatch. El algoritmo da más peso a los elementos que, según estudios empíricos, están más relacionados con la reincidencia, escribió López Ossorio. Se negó a revelar la fórmula exacta.

Una vez que se establece el resultado de un caso, el agente decide sobre un paquete de medidas de protección asociadas a ese nivel de riesgo. En los casos con resultados más bajos, los agentes controlarán discretamente a la mujer de vez en cuando. Con resultados más altos, la policía le dará a la víctima un botón de alarma, rastreará los movimientos del agresor o vigilará su casa. El agente también envía el formulario y la evaluación del riesgo a los fiscales y jueces que tratarán el caso de la mujer.

Tras el primer informe, la policía volverá a reunirse con la mujer para rellenar un segundo formulario, con el fin de evaluar si la situación ha empeorado o mejorado. Esto ocurre periódicamente, con mayor o menor frecuencia según el nivel de riesgo. La policía detiene el seguimiento solo si el caso judicial no sigue adelante y el nivel de riesgo cae por debajo del promedio base.

VioGén es uno de los resultados producidos por una ley pionera en materia de violencia de género que España aprobó en 2004, diez años antes de que el Consejo de Europa adoptara un marco legal común alrededor del tema, el Convenio de Estambul. Hoy en día, el software es utilizado por las principales fuerzas policiales españolas (Policía Nacional y Guardia Civil) y por cientos de cuerpos policiales locales (excepto en Cataluña y en el País Vasco, que cuentan con cuerpos policiales independientes).

"COMPARADO CON EL RESTO Y CONSIDERANDO LAS LIMITACIONES EXISTENTES, VIOGÉN SE ENCUENTRA ENTRE LOS MEJORES SOFTWARE DISPONIBLES",

JUANJO MEDINA, CRIMINÓLOGO.

/ El mejor sistema disponible

VioGén es hasta ahora el mejor dispositivo disponible para proteger la vida de las mujeres, según Ángeles Carmona, presidenta del Observatorio contra la Violencia Doméstica y de Género del Consejo General del Poder Judicial de España (CGPJ).

Carmona recuerda un caso que vio en un juzgado de Sevilla, donde un agresor tenía un alto riesgo de reincidencia, según VioGén. Al acusado le aplicaron una pulsera de control. Un día, la policía vio que la señal de la pulsera avanzaba rápidamente hacia la casa de la víctima. Llegaron justo a tiempo para evitar que él la sofocara con una almohada.

Es imposible saber cuántas vidas se han salvado gracias a VioGén, según Antonio Pueyo, catedrático de psicología de la Universidad de Barcelona, que ha asesorado a VioGén desde el los inicios del proyecto.

Sin embargo, un estudio de 2017 realizado por López Ossorio y su equipo intentó medir lo bien o mal que funciona el protocolo. Encontraron que el Área bajo la curva (AUC) de VioGén, una medida de rendimiento ampliamente utilizada para modelos predictivos, se situaba entre 0,658 y 0,8. Un AUC de 0,5 equivale al lanzamiento de una moneda y un AUC de 1 significa que el modelo nunca falla. Las pruebas de detección del cáncer se consideran buenas cuando su AUC está entre 0,7 y 0,9. En otras palabras, VioGén funciona.

"Comparado con el resto y considerando las limitaciones existentes, VioGén se encuentra entre los mejores software disponibles", dice Juanjo Medina, profesor de Criminología cuantitativa en la Universidad de Manchester, que ha comparado instrumentos de evaluación del riesgo de violencia doméstica.

España es el único lugar donde se puede realizar un seguimiento de las víctimas a través de diferentes regiones. Cerca de 30.000 policías y otros agentes de todo el país tuvieron acceso a VioGén en 2018.

Sin embargo, los casos que se han escapado al algoritmo de VioGén han generado preocupación. El último ocurrió en febrero de 2020, cuando una mujer de 36 años y madre de dos hijos fue degollada por su expareja, quien luego arrojó su cuerpo en un contenedor de basura en la localidad de Moraira. Los dos habían sido registrados en el sistema VioGén después de que la policía denunciara al hombre por agredirla, pero el caso quedó inactivo después de que un juez lo absolviera.

/ Falsos negativos

En 2014, el diario El Mundo publicó un documento filtrado del Consejo General del Poder Judicial que mostraba que 14 de las 15 mujeres que fueron asesinadas ese año después de haber denunciado a su agresor tenían un riesgo bajo o sin especificar (clasificación utilizada para cualquier persona que denuncia una amenaza a la policía).

Algunos críticos dicen que el bajo riesgo ni siquiera debería ser una opción. Denunciar es de por sí un momento de máximo riesgo para una mujer, según Carme Vidal Estruel, portavoz de Tamaia, una asociación que ayuda a las víctimas en Barcelona. Vidal Estruel dice que la situación es similar a divorciarse o quedarse embarazada, son momentos en los que el agresor se da cuenta de que está perdiendo el control de la víctima.

Otra crítica muy común es que pocos agentes entre los que deberían evaluar el resultado del software, reciben suficiente formación en temas de género. Algunos puntos tratados por VioGén suelen producir vergüenza, como los relacionados con la violencia sexual, humillaciones o mensajes íntimos en teléfonos móviles.

Los agentes deben hacer preguntas circulares (en lugar de preguntas directas y contundentes) y evitar transmitir la sensación de que el objeto de la investigación es la mujer, según Chelo Álvarez, presidente de Alanna, asociación de exvíctimas en Valencia. Ángeles Carmona, del CGPJ, recuerda una mujer que denunció a su marido por robarle las llaves del coche. Estaba tan asustada que no pudo decir nada más. Al día siguiente, el hombre la mató.

Pocos agentes son conscientes de estos matices y detalles. En 2017, había un total de 654 agentes en toda España pertenecientes a los Equipos Mujer-Menor (EMUME) de la Guardia Civil. Eso es mucho menos de uno por estación de policía.

/ Requisitos ignorados

Es una situación muy diferente de la requerida por la ley de 2004 que creó VioGén. Según la normal, los casos deben ser tratados por un equipo interdisciplinario que incluya psicólogos, trabajadores sociales y médicos forenses.

Tal equipo debe indagar en los aspectos psicológicos que el formulario VioGén no cubre. Además, debe realizar una valoración forense del agresor. El sistema actual equivale a decidir cómo de peligrosa es una persona sin siquiera hablar con ella, señalan los críticos. Varios equipos fueron creados después de la aprobación de la ley en 2004, pero el proceso se vio reducido tras la crisis financiera de 2008.

Pueyo, el profesor de psicología, reconoce algunas de las críticas pero considera que VioGén debe ser juzgado por su capacidad para predecir nuevos asaltos, no homicidios, porque estos eventos son excepcionales. La probabilidad de que una mujer muera después de denunciar es de una entre diez mil, según López Ossorio.

Sin embargo, el Convenio de Estambul exige precisamente reducir el riesgo de muerte. Y no solo de las mujeres sino también de sus hijos. Pasar por alto el riesgo para los niños es otra de las críticas que recibe VioGén.

La convención entró en vigor en España en 2014, pero los formularios de VioGén no se modificaron hasta que ocurrió el caso de Itziar P. en 2018, según su abogado.

/ VioGén 5.0

En marzo de 2019 se implementó un nuevo protocolo, el quinto gran cambio por el que ha pasado VioGén desde que se empezó a utilizar en 2007. Ahora, el programa identifica casos "de especial relevancia", en los que la peligrosidad es alta, y casos "con menores en riesgo".

Esto se logra través de un "doble procedimiento de evaluación" del nuevo formulario VPR (VPR5,0-H), explica López Ossorio. Se realizan dos cálculos en paralelo: uno relacionado con la reincidencia y otro relacionado con el asalto mortal.

Dependiendo del resultado de este último (llamado "escala H"), el resultado de riesgo puede aumentar automáticamente. Además, el caso puede ser señalado a los fiscales y jueces como "de especial relevancia".

López Ossorio se negó a revelar cómo se construyó la escala H, pero escribió que se basó en un estudio que su grupo realizó durante cuatro años para encontrar qué factores estaban específicamente relacionados con los casos que terminan en homicidios.

El nuevo protocolo parece haber provocado un cambio importante en las puntuaciones de riesgo de VioGén. Pasando de VPR4.0 a VPR5,0-H, el número de casos de riesgo extremo aumentó y los de alto riesgo casi se duplicaron, según López Ossorio.

Como dice la presidenta de la asociación de ex víctimas de Valencia, Chelo Álvarez: "Las cosas están mejorando, pero deberían hacerlo más rápido porque nos están matando".

EL NUEVO PROTOCOLO PARECE HABER PROVOCADO UN CAMBIO IMPORTANTE EN LAS PUNTUACIONES DE RIESGO DE VIOGÉN.

INVESTIGACIÓN

Por José Miguel Calatayud

Contextualización

Los organismos de la administración pública en los ámbitos local, regional y nacional en España utilizan sistemas de ADM (algoritmos para la toma de decisiones automatizada) desde hace años, y si bien los sistemas de ADM totalmente autónomos suelen ser bastante raros — según la información disponible públicamente — los protocolos en los cuales se emplean mecanismos de ADM en el proceso de decisión parecen ser bastante comunes.

Las "actuaciones administrativas automatizadas" están regulamentadas por ley, según la cual antes de implementar cualquier tipo de acción los organismos públicos deberían establecer qué autoridad competente definirá sus funciones y, de ser necesario, se encargarán de controlar el código fuente. Además, la ley de transparencia obliga a los organismos públicos a ser proactivamente transparentes y a otorgar a los ciudadanos acceso a cualquier información que la administración pública pueda llegar a tener.

Sin embargo, en la práctica los organismos públicos divulgan por sí mismos muy poca información de modo sobre los sistemas de ADM que utilizan, y también son reacios a hacerlo cuando los ciudadanos u organizaciones lo solicitan. Todo esto podría llegar a cambiar dependiendo del resultado de un caso judicial todavía abierto: una fundación sin fines de lucro solicitó al gobierno que se publique el código fuente de un sistema de ADM de la misma manera en la que se publican los textos legales de dominio público (hablamos de este caso a continuación en este mismo capítulo).

En los últimos años, ha habido universidades que han publicado investigaciones académicas sobre el uso de ADM por parte de la administración pública, que varios organismo de control también está tratando de vigilar. Pero estos sistemas de ADM son raramente mencionados en los principales medios de comunicación o en los discursos políticos, y parece haber una falta general de concienciación en la opinión pública sobre el uso de sistemas de ADM por los organismos públicos y sobre sus consecuencias en la vida pública y privada.

Empresas españolas, grandes corporaciones internacionales, empresas emergentes y universidades han estado desarrollando y proporcionando a la administración pública sistemas de ADM.

Como ya informó el capítulo del año pasado y tal y como actualiza este , los sistemas de ADM en España son comunes en los campos de vigilancia de seguridad y actividad policial predictiva, en el sector de la salud, y en el análisis de contenido de las redes sociales; y también han ido incrementando su presencia en la asignación de ayudas económicas y sociales, y la automatización de diferentes trámites administrativos.

El hecho de que, a fecha de redacción de este capítulo [11 de mayo de 2020] España no ha tenido un gobierno central estable durante más de cuatro años (desde finales de 2015), y que la administración pública está bastante descentralizada (regiones y pueblos tienen un alto grado de autonomía para implementar sus propias políticas), significa que puede haber diferencias normativas entre regiones, y que los pueblos y ciudades han tenido casi vía libre para implementar procesos de ADM en el ámbito local, particularmente en el marco de las tecnologías de las llamadas "Ciudades Inteligentes ("Smart City").

Catálogo de nuevos casos de ADM

/ Distribución de ayudas económicas

En noviembre de 2017, la Secretaría de Estado de Energía publicó un software llamado BOSCO para las empresas

proveedoras de electricidad. El objetivo de BOSCO era determinar si las personas tenían o no derecho a recibir ayuda financiera para poder pagar sus facturas de electricidad. El razonamiento detrás de este ADM era doble. En primer lugar, tratar de facilitar el proceso a los solicitantes de la ayuda (aunque, a juzgar por la gran cantidad de quejas que recibió el sistema, no resultó ser así) y, en segundo lugar:,igualmente facilitar y optimizar el proceso para las empresas proveedoras de electricidad.

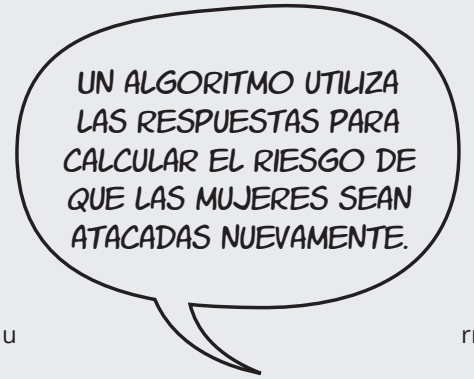
Tras recibir una gran cantidad de información y quejas sobre que el software no funcionaba correctamente, Civio, un medio de investigación sin ánimo de lucro y *lobby* ciudadano con sede en Madrid, descubrió que BOSCO negaba sistemáticamente la ayuda a algunos solicitantes que cumplían los requisitos para ser elegidos. Civio solicitó al gobierno el código fuente de BOSCO para tratar de entender por qué estaban ocurriendo esos errores. La solicitud pasó por tres ministerios diferentes antes de terminar en el Consejo de Transparencia y Buen Gobierno, que rechazó revelar el código diciendo que hacerlo violaría los derechos de autor (a pesar de que el software había sido desarrollado por la propia administración pública).

En julio de 2019, Civio presentó un recurso contencioso-administrativo afirmando que el código fuente de cualquier sistema de ADM utilizado por la administración pública debería ser público, de la misma manera en que lo son las fuentes de derecho (Civio, 2019). El caso, que en el momento de redactar este documento [11 de mayo de 2020] está todavía en curso, podría terminar en el Tribunal Supremo y podría sentar un precedente legal.

/ Evaluación de riesgos en casos de violencia doméstica

Probablemente, el sistema de ADM que más escrutinio público ha recibido en España es el protocolo VioGén, que contiene un algoritmo que evalúa el riesgo de que personas que han sufrido violencia doméstica vayan a ser atacadas nuevamente por su pareja o expareja.

VioGén fue presentado en 2007 tras la aprobación de una ley de 2004 sobre violencia de género que exigía un sistema integrado para proteger a las mujeres. Desde entonces, cada vez que una mujer presenta una denuncia por violencia doméstica el agente de policía u oficial que la atienda debe hacerle una serie de preguntas que provienen de un formulario estandarizado. Un algoritmo utiliza las respuestas para calcular el riesgo de que las mujeres sean atacadas nuevamente. Estos van desde ningún riesgo observado a riesgo bajo, medio, alto o extremo. Si más adelante un oficial a cargo del caso cree que es necesaria una nueva evaluación, VioGén incluye una segunda serie de preguntas y un formulario diferente, que se pueden utilizar para realizar un seguimiento del caso y para que el algoritmo actualice el nivel de riesgo. La idea era que el protocolo VioGén ayude a policías de toda España a realizar evaluaciones consistentes y estandarizadas de los riesgos asociados a la violencia doméstica, para que todos los casos que sean denunciados se beneficien de una respuesta más estructurada por parte de las autoridades, incluyendo un seguimiento y revaluaciones cuando sea necesario.



Una vez que se ha seguido el protocolo, el oficial puede anular la respuesta del algoritmo y decidir otorgar al caso un mayor nivel de riesgo. Cada nivel implica diferentes medidas de protección obligatoria, y solo el nivel de riesgo "extremo" implicaría protección permanente. Aun así, según los documentos oficiales sobre VioGén, la policía puede decidir otorgar medidas de protección adicionales a las establecidas por el protocolo.

El conjunto de preguntas y los algoritmos han sido desarrollados en colaboración entre policías y expertos académicos. Desde sus primeras implementaciones ya han sido actualizados cinco veces, a medida que oficiales y expertos en violencia doméstica observaban cómo funcionaba el protocolo en su aplicación.

Hay una razonable cantidad de información pública disponible sobre el protocolo VioGén en su conjunto (si bien no sobre el algoritmo en sí), y un libro, publicado en septiembre de 2018 por el Ministerio del Interior, contiene un relato bastante franco (para los estándares habituales de información disponible sobre sistemas de ADM) y muy informativo de la historia, el diseño y la implementación de este protocolo (Ministerio del Interior de España, 2018). Sin embargo, el libro no revela el código de los algoritmos o los métodos internos usados para calibrar y medir el peso dado a los diferentes factores y a sus interrelaciones.

La sensibilidad que conlleva el tema de la violencia doméstica y el hecho de que se disponga de una buena cantidad de información sobre el protocolo en su conjunto (si no sobre el algoritmo en sí) han contribuido a que VioGén sea bastante visible en los medios de comunicación (Precedo, 2016). Y ese es también el caso cuando se dan a conocer cifras oficiales sobre víctimas de violencia doméstica y estas incluyen casos que VioGén no consideró de alto o extremo riesgo y que sin embargo terminaron con el asesinato de la mujer (Álvarez, 2014).

Por otro lado, una revisión minuciosa de las cifras oficiales disponibles muestra que estas han sido publicadas de manera inconsistente, a veces contabilizando de forma agregada las evaluaciones iniciales y las de seguimiento de diferentes casos (por ejemplo, algunos informes describen el número total de evaluaciones como el número total de casos individuales, sin tener en cuenta que un mismo caso puede contar con varias evaluaciones), y en ocasiones habían transcurrido varios años entre una valoración inicial de bajo riesgo y un nuevo ataque del agresor, que tras haber cumplido su pena de prisión esta vez mataba a la mujer. Todo esto dificulta extraer conclusiones coherentes sobre las posibles correlaciones entre los casos evaluados com de bajo riesgo y aquellos en los que el agresor acabó matando a la mujer.

Al hablar de VioGén, la mayoría de los expertos en maltrato y violencia de género citados por la prensa se quejan de que los agentes de policía que siguen el protocolo a menudo no han recibido la formación adecuada para tratar los casos de abuso doméstico (Precedo, 2016). Además, el libro que fue publicado por el Ministerio del Interior cita un estudio de 2014 en donde se indica que los agentes que siguieron el protocolo no cambiaron el nivel de riesgo sugerido por el algoritmo en el 95% de los casos. Estos parecen respaldar la idea de que el código fuente del algoritmo debería ser publicado para que pueda ser evaluado adecuadamente.

/ Detección de fraude en la contratación pública

En octubre de 2018, el parlamento regional valenciano aprobó una ley que incluía el uso de un sistema de ADM para detectar posibles casos de fraude en la contratación pública (Europa Press, 2018). La mayoría parlamentaria en el gobierno en ese momento estaba formada por una coalición de partidos progresistas que llegaron al poder en el 2015. Este nuevo gobierno llegó tras 20 años de gobierno del conservador Partido Popular, que al final se vio envuelto en una serie de escándalos de corrupción conectados, entre otras cosas, con la contratación pública.

El sistema de ADM, conocido como Saler, fue financiado por el gobierno valenciano y la organización de derechos humanos FIBGAR, y desarrollado por la Universidad Politécnica de Valencia.

Saler coteja información de distintas bases de datos de la Administración pública valenciana y del Boletín Oficial del Registro Mercantil (BORME). Su finalidad es encontrar señales de alerta cuando detecta un comportamiento sospechoso (por ejemplo, una gran cantidad de pequeños contratos otorgados a la misma empresa). Ante una situación sospechosa, Saler transmite automáticamente la información a la autoridad competente: la Agencia Antifraude, la Fiscalía o el Tribunal de Cuentas.

La idea, aún por demostrar, es que Saler se volverá más eficaz en encontrar señales de alerta relevantes a medida que tenga acceso a más bases de datos y aprenda al ser utilizado por funcionarios públicos. A sus creadores les gustaría exportarlo a otras administraciones regionales de España, y también a la administración central nacional; pero el hecho de que cada región utiliza diferentes índices y bases de datos conllevaría que por el momento este proceso fuera bastante complicado de llevar a cabo.

La presentación pública de este sistema de ADM por parte de las autoridades valencianas obtuvo una mayor cobertura de prensa que la habitual sobre estos temas, probablemente porque se trataba de un sistema para combatir la corrupción y porque su acrónimo original era Satán (Cid, 2018).

/ Videovigilancia y (cuasi) reconocimiento facial

En noviembre de 2019, Ifema, un consorcio de propiedad pública, empezó a instalar cámaras de vigilancia con reconocimiento facial en el centro de conferencias que gestiona en Madrid (Peinado, 2019). Ifema ya había especificado en la información de contratación pública de marzo de 2019 que buscaba "mejoras técnicas en las licencias de reconocimiento facial".

Cabe señalar que las cámaras aún no estaban en uso cuando se llevó a cabo la Conferencia de las Naciones Unidas sobre Cambio Climático COP25, celebrada en Ifema entre el

2 y el 13 de diciembre de 2019, según un representante del consorcio citado por la prensa (Peinado, 2019) .

A finales de 2018, la ciudad de Marbella, un polo turístico de la costa sur de España, empezó a utilizar un sistema de videovigilancia que, al parecer, contaba la más alta definición de España, desarrollado por la firma estadounidense Avigilon (Pérez Colomé, 2019). Según informaciones en prensa, solo la frontera entre el enclave español de Ceuta y Marruecos y algunos estadios de fútbol utilizan cámaras similares. Oficialmente, el sistema no tiene funciones de reconocimiento facial, pero el software utilizado por estas cámaras usa funciones como "búsqueda de apariencia" y "análisis facial". Por lo que parece - según la información oficial de Avigilon - esto permite que el sistema busque personas por su apariencia identificando una serie de rasgos faciales únicos, la ropa, la edad, el sexo y el color del cabello.

Un representante de la empresa citado por la prensa dijo que su software de reconocimiento facial no estaba siendo utilizado por las autoridades públicas en España, pero agregó que la empresa sí había instalado software de reconocimiento facial en sistemas de vigilancia privados en España (Pérez Colomé, 2019).

Tanto en el caso de Ifema como en el de Marbella, los sistemas han sido justificados porque dicen haber aumentado la seguridad y al mismo tiempo la eficiencia de la vigilancia. Pero en ambos casos no está claro cómo funciona exactamente el software de reconocimiento y qué tipo de controles u otras medidas se han podido establecer para proteger la información biométrica de las personas y sus datos personales.

/ Vigilancia policial predictiva

En 2016, la policía local de Rivas-Vaciamadrid, una localidad de 86.000 habitantes de la Comunidad de Madrid, puso en marcha una prueba piloto de Pred-Crime, un software desarrollado por la empresa española EuroCop (Europa Press, 2015). Pred-Crime analiza datos de antecedentes históricos para predecir dónde y cuándo es más probable que se cometan ciertos tipos de delitos , como infracciones de tránsito y robos.

Supuestamente, el plan era implementar completamente este software a lo largo de 2016, pero después de probarlo durante nueve meses el municipio decidió no seguir usándolo. "Necesita más tiempo de desarrollo para ser adecuadamente eficiente", dijo un representante del municipio a la prensa de mayo de 2019 (García, 2019).

> "NECESITA MÁS TIEMPO DE DESARROLLO PARA SER ADECUADAMENTE EFICIENTE", UN REPRESENTANTE MUNICIPAL SOBRE PRED-CRIME.

En su sitio web, EuroCop señala que entre sus clientes hay decenas de municipios de toda España, pero no precisa si están usando su software predictivo u otras herramientas no predictivas que la compañía también comercializa.

En otro aparente caso de vigilancia predictiva, la policía española habría estado utilizando un software que analiza información disponible sobre la víctima de un asesinato y el contexto del crimen, y basándose en ella produce un perfil probable del asesino. Según información publicada por la prensa, entre 2018 y 2019 agentes de la Secretaría de Estado de Seguridad habrían colaborado con la Policía utilizando dicho software en al menos cinco investigaciones (Pérez Colomé, 2019b). No se conocen más detalles sobre este caso.

Según el mismo artículo de prensa, a mediados de 2019 el Ministerio de Medio Ambiente firmó un acuerdo con la Universitat Autònoma de Barcelona para desarrollar un software que pudiera predecir el perfil más probable de pirómanos.

/ Análisis automatizado de los contenidos de las redes sociales

En enero de 2018, la Generalitat de Catalunya puso en marcha un proyecto piloto para intentar medir el impacto en la

ciudadanía de sus iniciativas STEMcat, un plan destinado a fomentar las vocaciones científicas y tecnológicas entre los jóvenes.

Durante un mes y medio, el gobierno catalán utilizó Citibeats. Este software de análisis de texto utilizó algoritmos de aprendizaje automático para recopilar y analizar alrededor de 12.000 tuits que hablaban sobre las disciplinas STEM en Cataluña. La comunicación pública del gobierno catalán sostuvo que, a través de este instrumento, una de las ideas que habían obtenido era que las mujeres eran más receptivas a mensajes sobre ciencias naturales que sobre tecnología. Las autoridades utilizaron esa y otra información para "optimizar su estrategia y proponer nuevas iniciativas" con el fin de estimular el interés en las disciplinas STEM en los jóvenes. El proyecto formó parte de SmartCAT, la estrategia del gobierno catalán para convertirse en una *región inteligente* (como señaló anteriormente este capítulo, los gobiernos regionales en España tienen un alto grado de autonomía para elaborar sus normas). El director de SmartCAT dijo que el software les había permitido "evaluar de una manera más objetiva el impacto de las iniciativas (gubernamentales)" para lograr que la gente se interese por la ciencia y la tecnología.

Citibeats, desarrollado por Social Coin, una empresa emergente radicada en Barcelona, también fue utilizado en diciembre de 2017 por el ayuntamiento de Barcelona para recopilar las actitudes de las personas hacia el transporte público y la movilidad en la ciudad, mediante el análisis de alrededor de 30.000 comentarios de más de 15.000 personas. En un estudio de caso de este proyecto, Citibeats habló de "ciudadanos como sensores".

En los dos casos – y aunque las autoridades elogian el software por su capacidad de recopilar y analizar miles de comentarios en línea (algo que llevaría mucho más tiempo y dinero hacer con los métodos de encuesta tradicionales) –, no está claro qué tan representativas son las muestras y cuán válidas podrían ser las conclusiones y no parece haber información sobre cómo esos análisis influyeron en las políticas públicas.

Desde marzo de 2019, el gobierno regional de Navarra también ha estado utilizando Citibeats para detectar discursos de incitación al odio en línea, mediante el análisis de textos publicados en Facebook, Twitter e Instagram.

Al igual que en otros casos que tratan con datos personales, no está claro cómo funciona el software y tampoco qué mecanismos de supervisión puede tener la autoridad pública al utilizarlo.

/ Evaluación automatizada de ofertas de contratación pública mediante licitación electrónica

La actual ley de contratación pública en España, aprobada en noviembre de 2017 para adherirse a la normativa de la UE, permite a las autoridades públicas conceder licitaciones digitales utilizando dispositivos electrónicos que clasifiquen las diferentes ofertas "a través de métodos de evaluación automatizados".

Esta tipología de licitación digital sólo podrá utilizarse cuando "los requisitos del contrato a adjudicar se puedan establecer de manera precisa" y los "servicios públicos que son su objeto no sean de carácter intelectual, como servicios de arquitectura, ingeniería y consultoría". Los contratos relacionados con la calidad de los alimentos tampoco pueden ser adjudicados mediante licitación digital.

Según la ley, antes de utilizar este método, una autoridad pública debe dar a conocer el "dispositivo electrónico" que va a ser utilizado en el proceso de licitación.

Antes de llevar a cabo la licitación digital, la autoridad pública debe realizar una evaluación completa de todas las ofertas disponibles y luego enviar una invitación electrónica simultánea a todos los postores con derecho a participar. En dicha invitación, la autoridad debe incluir "la fórmula matemática" que se utilizará en la clasificación automática de las ofertas elegibles.

/ Asistencia automatizada para la declaración de impuestos

Desde julio de 2017, la Agencia Tributaria española ha estado utilizando el software Watson de IBM para brindar asistencia automatizada en un aspecto particular de la declaración del IVA que afecta sobre todo a grandes empresas. Tanto IBM como la Autoridad Tributaria destacaron públicamente que el software puede funcionar las 24 horas del día y los 7 días de la semana, por lo que así liberaba a los funcionarios de tener que lidiar con una gran cantidad de correos electrónicos sobre la declaración del IVA. Según la comunicación pública de ambas entidades, entre julio de 2017 y febrero de 2018 el número de correos electrónicos que los funcionarios recibían sobre cuestiones referentes al IVA disminuyó de 900 a 165 por semana. Y, según esas comunicaciones, el asisten-

te automático pasó de recibir 200 preguntas por semana, cuando se lanzó en julio de 2017, a alrededor de 2.000 en noviembre de 2017 (Computing, 2018).

/ Ciudades inteligentes: de la 'Basura Inteligente' al 'Turismo Inteligente'

A medida que las ciudades españolas aspiran a utilizar macrodatos y procesos automatizados para convertirse en ciudades "inteligentes" (Smart Cities), toda una serie de sistemas de ADM han ido siendo implementados en el ámbito municipal. Estos sistemas suelen ayudar la toma de decisiones en vez de funcionar de forma totalmente autónoma.

Un ejemplo es la plataforma Smart Waste (Basura Inteligente), que recopila datos de los sensores instalados en contenedores y camiones, y también de las redes sociales, las encuestas, el censo y de imágenes por satélite. Todos estos datos combinados ayudan a las autoridades locales a decidir qué servicios se necesitarán, cuándo y dónde.

Smart Waste fue desarrollada por The Circular Lab (del centro de innovación de Ecoembes), una organización sin ánimo de lucro encargada de recolectar envases de plástico, latas, cartones, papel y cartón para su reciclaje. Minsait, una división de Indra – una multinacional española consultora de tecnología de transporte, defensa y seguridad -, también ayudó a desarrollar la plataforma.

Durante 2018, el ayuntamiento de Logroño y los gobiernos de La Rioja y Cantabria usaron por primera vez la plataforma como piloto, y hoy está disponible para las autoridades locales y regionales de toda España.

Como ya es habitual en casi todos los casos en España, no hay información disponible sobre cómo funciona el software, qué tipo de resultados produce y qué decisiones o cambios han adoptado las distintas autoridades al usarlo.

Otro ejemplo de uso de sistemas de ADM por parte de autoridades locales es el de Barcelona y la basílica de la Sagrada Familia, una de las atracciones turísticas de la ciudad. Ambos se asociaron con la empresa privada Bismart para desarrollar una aplicación móvil llamada Smart Destination (Destino Inteligente). Esta aplicación analiza los datos públicos disponibles sobre las personas (incluidas sus publicaciones en las redes sociales y la cantidad de personas que hacen cola para ingresar a una atracción turística), los vehículos y el tráfico, el clima, la ocupación de bares y hoteles, etc... y también las preferencias dadas por el usuario para generar automáticamente planes y rutas turísticas personalizadas.

Una vez más, no se sabe exactamente cómo funciona el sistema y tampoco cómo genera los resultados que pone a disposición del público.

Políticas públicas, normativa y debate público en torno a los sistemas de ADM

/ La ley que regula la administración pública

En España, el uso de los procesos de ADM por parte de la administración pública se rige por la ley 40/2015. La

COMO YA ES HABITUAL EN CASI TODOS LOS CASOS EN ESPA A, NO HAY INFORMACIÓN DISPONIBLE SOBRE CÓMO FUNCIONA EL SOFTWARE.

ley define "Actuación administrativa automatizada" como "cualquier acto o actuación realizada íntegramente a través de medios electrónicos por una Administración Pública en el marco de un procedimiento administrativo y en la que no haya intervenido de forma directa un empleado público". Y el texto también establece que antes de tomar cualquier actuación administrativa automatizada se debe establecer cuáles de las autoridades competentes serán las encargadas de definir "las especificaciones, programación, mantenimiento, supervisión y control de calidad y, en su caso, auditoría del sistema de información y de su código fuente". También debe conocerse qué organismo será responsable en caso de impugnación legal de la acción automatizada.

/ Transparencia y acceso a la información

La ley de 2013 que regula el acceso de los ciudadanos a la información pública exige que los organismos públicos sean proactivamente transparentes y que otorguen a los ciudadanos acceso a cualquier contenido y documentos que la administración pública pueda llegar a tener.

Sin embargo, en la práctica los organismos públicos apenas divulgan información sobre los sistemas de ADM que utilizan. Cuando ciudadanos u organizaciones solicitan información al respecto, la administración pública suele ser reacia a otorgarla, lo que significa que el uso de los procesos de ADM ocurre en un contexto opaco. El caso judicial en curso iniciado por Civio, comentado anteriormente en este informe, ilustra este punto.

Desde 2014, la administración pública central ha estado elaborando Planes de Gobierno Abierto, donde se trazan los compromisos hechos para avanzar hacia un gobierno y una administración transparentes. El actual, el III Plan de Gobierno Abierto 2017-19, no incluye acciones relacionadas explícitamente con los sistemas de ADM. El IV Plan de Gobierno Abierto 2019-21 se encuentra, en el momento de redactar este informe [11 de mayo de 2020], en pleno desarrollo por un grupo de trabajo interministerial.

/ Ley de protección de datos

Por lo que respecta a la protección de datos personales, la ley española 3/2018 (que siguió a la aprobación del RGPD) otorga a los ciudadanos el derecho a dar su consentimiento para la recopilación y el procesamiento de sus datos por parte de las autoridades. En principio, los ciudadanos también tienen derecho a saber cuáles son los organismos que almacenan sus datos y qué pretenden hacer con ellos. El texto explícitamente menciona que la ley concierne "cualquier tratamiento total o parcialmente automatizado de datos personales". También dice que cuando se utilizan datos personales "para la elaboración de perfiles" las personas afectadas "el afectado deberá ser informado de su derecho a oponerse a la adopción de decisiones individuales automatizadas que produzcan efectos jurídicos sobre él o le afecten significativamente de modo similar".

El 16 de junio de 2020, el gobierno español anunció que había puesto en marcha el proceso de elaboración de una Carta de Derechos Digitales constituyendo a un grupo de expertos que tendrá el rol de asesor. El objetivo de la Carta es el de añadir nuevos 'derechos digitales' específicos a los ya incluidos en la legislación española. Esta nueva Carta debería incluir derechos "relacionados con la protección de colectivos vulnerables, las nuevas relaciones laborales o el impacto de nuevas tecnologías como la inteligencia artificial", entre otros temas, según informa el comunicado de prensa del gobierno. El anuncio sostenía que en una etapa posterior el proceso sería participativo, abriéndose al público, y que al final el gobierno redactaría la Carta teniendo en cuenta las aportaciones de los expertos y de la ciudadanía. En el momento de redactar este documento, no había información sobre el calendario del proceso ni ningún otro detalle sobre el alcance y el contenido de la Carta.

/ Toma de decisiones políticas y normativa

En materia de toma de decisiones políticas, la Secretaría de Estado de Avance Digital, dependiente del Ministerio de Asuntos Económicos, tiene a su cargo la coordinación de las diferentes estrategias, planes y acciones para la "conectividad y transformación digital de España", e incluye un apartado centrado en la inteligencia artificial.

Solo uno de los 11 planes impulsados por esa oficina, el dedicado a los "territorios inteligentes", que se publicó en diciembre de 2017, menciona el uso de sistemas de ADM. El plan afirma que la calidad de vida de la gente aumentará cuando la automatización permita al gobierno y a la administración pública identificar proactivamente los problemas individuales de las personas y proponer soluciones a estos. En cuanto a la recolección de datos, el plan establece que "deben definirse los mecanismos de recogida y tratamiento de datos, limitando el número y tipos de datos a recoger, controlando sus usos, facilitando el acceso a los mismos y a

la lógica de los algoritmos con los que se fundan decisiones ulteriores".

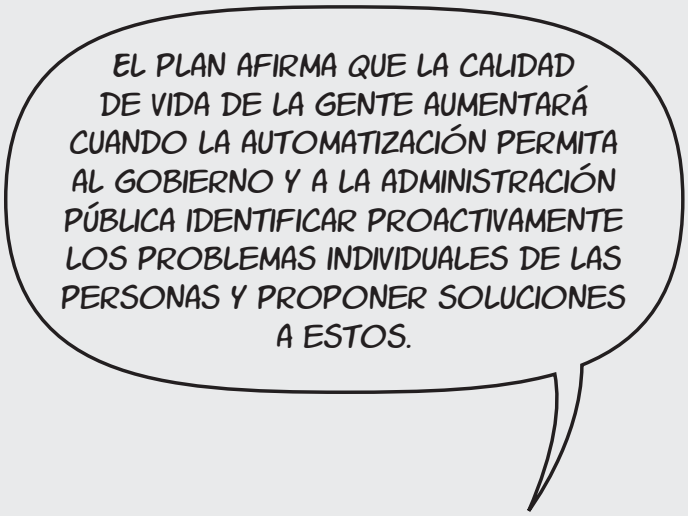
El organismo encargado de "implementar las acciones encomendadas en planes de la Agenda Digital para España" es Red.es, una entidad pública que depende de la Secretaría de Estado para el Avance Digital. Red.es incluye un Observatorio Nacional de las Telecomunicaciones y de la Sociedad de la Información que, en el momento de redactar este informe [11 de mayo de 2020], no parece haber trabajado explícitamente sobre sistemas de ADM.

España es un Estado descentralizado y cada una de las 17 regiones tiene un alto grado de autonomía sobre sus administraciones públicas, las cuales pueden ser diferentes entre sí. En la Administración General del Estado el máximo órgano de gobierno en materia de uso de las tecnologías de la información es la Comisión de Estrategia TIC, en la que todos los ministerios tienen representación y que asume un papel de intermediario entre el poder ejecutivo, encarnado por el Consejo de Ministros, y la propia administración pública. El decreto de 2014 que regula esta Comisión establece que la transformación digital debe significar "no sólo la automatización de los servicios [ofrecidos por la administración pública], sino su rediseño integral", pero no entra en más detalles con respecto a los sistemas de ADM.

/ Gobierno

/ El retrasado Plan Nacional de IA y la Estrategia Nacional de IA de España

En noviembre de 2017, el Ministerio de Industria y Comercio presentó un Grupo de Sabios sobre Inteligencia Artificial y Big Data, formado por personas del mundo académico y del sector privado (en representación de entidades como BBVA - el segundo banco más grande de España -, Telefónica, y el Instituto Vodafone, que la compañía británica describe como su laboratorio de ideas europeo) y le encargó la redacción de un informe oficial (Libro Blanco) sobre IA cuya publicación estaba prevista para julio de 2018. Se esperaba que ese libro blanco sirviera de base para la elaboración de un código ético para la utilización de datos en la administración pública. Este iba a ser el primer paso hacia un Plan Nacional de Inteligencia Artificial Español, pero en el momento de redactar este informe [11 de mayo de 2020] no se sabe nada al respecto. (Descargo de responsabilidad: cuando fue creado, este grupo de expertos incluía entre sus miembros a Lorena Jaume Palasí, cofundadora de AlgorithmWatch quien, en ese momento, todavía estaba involucrada con esta organización).



Como preludio de ese plan, en marzo de 2019 el Ministerio de Ciencia, Innovación y Universidades publicó una Estrategia Nacional de Inteligencia Artificial. Fue difundido poco antes de las elecciones generales de abril de 2019, y fue criticado como vago y apresurado por la prensa especializada (Merino, 2019). De hecho, la Estrategia parece una mera descripción general de las posibilidades y peligros de la IA para el Estado y la administración pública.

El documento aprueba el uso de sistemas de IA por parte de la administración pública, y asume que aumentará la eficiencia (por ejemplo, avanzando hacia la interoperabilidad entre organismos públicos y generando "procedimientos administrativos automatizados") y ahorrará al Estado enormes cantidades de dinero ("miles de millones de euros" solo en el sector de la salud, dice el documento haciendo referencia a un estudio de junio de 2017 de la consultora global PricewaterhouseCoopers) (PwC, 2017). La Estrategia también dice que las amenazas a la seguridad basadas en inteligencia artificial requieren soluciones basadas en inteligencia artificial, y defiende que haya transparencia en "algoritmos y modelos" y quiere ver un "uso honesto de la tecnología", sin describir ningún plan o norma precisos para explicar cómo avanzar en esa dirección.

El 18 de noviembre de 2019, en el acto de clausura del II Congreso Internacional de IA, el ministro de Ciencia, Pedro Duque, dijo que esperaba que el Plan Nacional de IA de España estuviera listo en las siguientes semanas (Fiter, 2019). Sin embargo, en el momento de redactar este informe [11 de mayo de 2020], no ha habido ninguna otra noticia al respecto.

En general, España no ha tenido un gobierno estable desde 2015 y, de hecho, el 10 de noviembre de 2019 España celebró sus cuartas elecciones generales en cuatro años (El País, 2019). Además, en el momento de redactar este informe [11 de mayo de 2020], un gobierno minoritario establecido el 13 de enero tenía dificultades para contener uno de los peores brotes de covid-19 en el mundo. Esos factores probablemente han contribuido a esos retrasos y a la falta de un liderazgo político más decisivo y dedicado en el desarrollo de estrategias y planes sobre el uso de sistemas de IA por parte de la administración pública.

/ Sociedad civil y comunidad académica

/ Investigación y debate académico

En los últimos años, ha habido un animado debate en el mundo académico sobre el uso de los sistemas de ADM en la administración pública (así como en el sector privado), y hay varios investigadores académicos que estudian específicamente la ADM en la administración pública. Sin embargo, pocos de estos debates parecen llegar a la opinión pública y a la ciudadanía en general.

En abril de 2019, la Red de Especialistas en Derecho de las Tecnologías señaló la "llamativa falta de transparencia algorítmica y la ausencia de una adecuada percepción por las Administraciones Públicas sobre la necesidad de aprobación de un marco jurídico específico. Únicamente se observa una cierta preocupación respecto del cumplimiento en materia de protección de datos, que se percibe como un límite". El documento parece ser un buen resumen de la situación en España y pide a las autoridades "impulsar nuevos mecanismos para garantizar estos principios y derechos [de las personas] por defecto y en el diseño". Las conclusiones también señalan el tema de las ciudades inteligentes como un área que necesita atención urgente, ya que los municipios están experimentando libremente con sistemas de ADM sin un debate o regulaciones adecuadas. (Cotino et al., 2019).

Otros destacados grupos de investigación académicos que trabajan en este campo son Idertec, de la Universidad de Murcia y el Grupo de Investigación en Ciencia Web y Computación Social de la Universitat Pompeu Fabra de Barcelona, que cuenta con un área de investigación dedicada a la justicia algorítmica y la transparencia.

/ Organismos de control sin ánimo de lucro

En los últimos años, varias organizaciones privadas han estado realizando investigaciones y propuestas pública que incluyen cuestiones sobre los sistemas de ADM. En Barcelona está Eticas Foundation, una organización sin ánimo de lucro asociada a la consultora Eticas Consulting, que investiga temas que incluyen discriminación algorítmica, derecho y práctica, migración y biometría, vigilancia y seguridad. Su fundadora y directora, Gemma Galdón, ha publicado varios artículos de opinión sobre estas temáticas y suele ser citada en la prensa generalista. Pero a fecha de redacción de este informe, y por lo que se aprecia en su sitio web, la fundación no parece haber estado muy activa recientemente.

En septiembre de 2019, una nueva asociación llamada OdiseIA se presentó en Madrid, describiéndose como un "Observatorio del impacto social y ético de la inteligencia artificial". Los fundadoras de OdiseIA son 10 personas de la comunidad académica, la administración pública y el sector privado, por lo general reconocidas en los círculos de expertos sobre cuestiones de ADM en España. Sin embargo en abril de 2020, OdiseIA aún no había comenzado a publicar sus propias investigaciones, recomendaciones u otro tipo de contenido.

/ Los sistemas de ADM en la prensa

Por lo que respecta al público en general y al discurso político y de los medios de comunicación, parece haber una carencia general de concienciación sobre el uso de los sistemas de ADM en la administración pública, las oportunidades, los desafíos y los riesgos que plantean dichos sistemas, y las implicaciones para la vida pública e individual. este es un punto enfatizado y lamentado por todos los expertos consultados.

Si bien la prensa especializada ha tenido de un modo más consistente un enfoque crítico con respecto a los sistemas de ADM en la administración pública, hasta hace poco la mayor parte de la cobertura sobre los sistemas de ADM por parte de la prensa generalista parecía basarse en poco que en los materiales de relaciones públicas difundidos por las empresas que desarrollan el software y por los organismos públicos que lo usaban. Sin embargo, y tal vez debido en parte a la mala imagen que la palabra "algoritmo" ha recibido en relación con escándalos conectados con Facebook y Google, recientemente la prensa generalista ha estado

adoptando también un enfoque más crítico. Dicho esto, la mayoría de los artículos en la prensa aún se ocupan de sistemas de ADM particulares que ganan visibilidad por algún motivo (como VioGén, discutido anteriormente, que trata el delicado tema de la violencia doméstica) en vez de tratar el uso general de los sistemas de ADM por parte de la administración pública, su marco normativo y sus implicaciones para la vida pública e individual. A menudo, cuando los medios de comunicación hablan de la falta de transparencia y responsabilidad con respecto a los sistemas de ADM en la administración pública, lo hacen analizando cuestiones individuales, caso por caso, en vez de tratar el problema desde un punto de vista estructural.

Conclusiones clave

Aunque no mencionan explícitamente a los sistemas de ADM, las leyes actuales que rigen la administración pública en España (que sí incluyen una definición de "actuaciones administrativas automatizadas") y el acceso de la ciudadanía a la información pública presentan una buena oportunidad para exigir rendición de cuentas a las autoridades y pedir que sean responsables en el uso de los procesos de ADM en la administración pública.

Los marcos normativos actuales serán puestos a prueba por el caso judicial abierto en el que Civio solicitó al gobierno que publique el código fuente de un sistema de ADM de la misma manera en que los textos legales son de dominio público.

También existen buenas prácticas, como la forma proactiva en la que las autoridades han ido publicando información sobre el protocolo VioGén, que podría servir de ejemplo a seguir para otras autoridades. En general, una mayor transparencia sería bienvenida – incluidas las formas de auditar el funcionamiento de los propios sistemas de ADM – y también se necesita un mayor esfuerzo de explicación por parte de las autoridades de su uso de protocolos de ADM. El número de investigadores académicos activos en la investigación de los sistemas de ADM en la administración pública (y también en el sector privado) y la cobertura cada vez más completa del tema por parte de la prensa generalista constituyen también una oportunidad para que la ciudadanía sea más consciente de cómo los sistemas de ADM se utilizan para prestar servicios públicos y las consecuencias que conllevan en la vida pública e individual.

Por otro lado, la normalización del uso de sistemas de ADM por parte de la administración pública sin un debate público adecuado o sin establecer mecanismos de supervisión claros – especialmente a medida que los sistemas de ADM de bajo perfil se extienden en el ámbito local – aumenta el riesgo de crear un entorno administrativo opaco, en el que los ciudadanos pueden terminar careciendo de los medios para responsabilizar a las autoridades sobre dichos sistemas.

PARECE HABER UNA CARENCIA GENERAL DE CONCIENCIACIÓN SOBRE EL USO DE LOS SISTEMAS DE ADM EN LA ADMINISTRACIÓN PÚBLICA, Y LAS OPORTUNIDADES, LOS DESAFÍOS Y LOS RIESGOS QUE PLANTEAN DICHOS SISTEMAS.

Bibliografía

ÁLVAREZ, RAFAEL J. (2014). *Las asesinadas que denunciaron se valoraron como 'riesgo bajo o nulo'.* <https://www.elmundo.es/espana/2014/12/09/54861553ca4741734b8b457e.html>[Consulta: 11 de Enero de 2020]

CID, GUILLERMO (2018). *Ingenieros valencianos crean 'Satan', un 'software' para cazar corruptos: así funciona.* <https://www.elcon dencial.com/tecnologia/2018-10-22/algoritmo-anticorrupcion-valencia- satan_1632428/> [Consulta: 11 de Enero de 2020]

CIVIO (2019). *Being ruled through secret source code or algorithms should never be allowed in a democratic country under the rule of law.* <https://civio.es/novedades/2019/07/12/being-ruled- through-secret-source-code-or-algorithms-should-never-be-allowed-in-a-social-and-democratic-state-under- the-rule-of-law> [Consulta: 11 de Enero de 2020]

COMPUTING (2018). *IBM Watson ficha por la Agencia Tributaria para la gestión del IVA.* <https://www.computing.es/analytics/noticias/1103993046201/ibm-watson-cha-agencia-tributaria-gestion- deliva.1.html> [Consulta: 11 de Enero de 2020]

COTINO, LORENZO ET AL. (2019). *Conclusiones del I Seminario Internacional Derecho Administrativo e Inteligencia Artificial.* <https://www.dropbox.com/s/5px5jkvauiz06vu/CONCLUSIONES_DAIAv nal.pdf> [Consulta: 11 de Enero de 2020]

EL PAÍS (2019). *Socialists win repeat Spanish election, Vox becomes third-biggest force in Congress*. <https://elpais.com/elpais/2019/11/10/inenglish/1573407794_574125.html> [Consulta: 11 de Enero de 2020]

EUROPA PRESS (2015). *La Policía Local usará un 'software' que facilita la prevención de delitos*. <https://www.elmundo.es/tecnologia/2015/12/10/566991ee268e3ee63c8b4634.html> [Consulta: 11 de Enero de 2020]

EUROPA PRESS (2018). *Les Corts aprueban la creación de 'Satan', un sistema de alertas contra la corrupción.* <https://www.europapress.es/comunitat-valenciana/noticia-les-corts-aprueban-creacion-satan- sistema-alertas-luchar-contra-corrupcion-20181017150930.html> [Consulta: 11 de Enero de 2020]

FITER, MIGUEL (2019). *El Gobierno reformará las universidades para hacerlas más exibles.* <https://www.elindependiente.com/politica/2019/11/18/el-gobierno-reformara-las-universidades-para- hacerlas-mas- exibles/> [Consulta: 11 de Enero de 2020]

GARCÍA, TER (2019). *Minority Report en las policías europeas: llegan los sistemas de predicción de delitos.* <https://www.elsaltodiario.com/tecnologia/minority-report-policias-europa-sistemas-algoritmos- prediccion-delitos>[Consulta: 11 de Enero de 2020]

MERINO, MARCOS (2019). *La 'Estrategia para la Inteligencia Artificial en I+D' del Gobierno, poco más que un teaser de la Futura Estrategia Nacional de IA*. <https://www.xataka.com/inteligencia-arti cial/estrategia-para-ia-i-d-i-presentada-gobierno-poco-que-teaser-futura-estrategia-nacional-ia> [Consulta: 11 de Enero de 2020]

PEINADO, FERNANDO (2019). *Las cámaras que leen la cara se extienden por Madrid.* <https://elpais.com/ccaa/2019/11/26/madrid/1574801864_377093.html> [Consulta: 11 de Enero de 2020]

PÉREZ COLOMÉ, JORDI (2019). *Marbella, el mayor laboratorio de videovigilancia de España*. <https://elpais.com/tecnologia/2019/11/21/actualidad/1574348695_231540.html> [Consulta: 11 de Enero de 2020]

PÉREZ COLOMÉ, JORDI (2019b). *Así trabaja el equipo de Interior que predice el perfil de los asesinos.* <https://elpais.com/tecnologia/2019/06/26/actualidad/1561500733_085160.html> [Consulta: 11 de Enero de 2020]

PRECEDO, JOSÉ (2016). *Las primeras medidas de protección a una mujer maltratada las decide un algoritmo informático*. <https://www.eldiario.es/sociedad/primeras-proteccion-maltratada-algoritmo- informatico_0_561344191.html> [Consulta: 11 de Enero de 2020]

PWC (2017). *Sherlock in Health. How artificial intelligence may improve quality and efficiency, whilst reducing healthcare costs in Europe*. <https://www.pwc.de/de/gesundheitswesen-und- pharma/studie-sherlock-in-health.pdf> [Consulta: 11 de Enero de 2020]

SPANISH MINISTRY OF THE INTERIOR (2018). *La valoración policial del riesgo de violencia contra la mujer en España.* <http://www.interior.gob.es/documents/642012/8791743/Libro+Violencia+de+G%C3%A9nero/19523de8- df2b-45f8-80c0-59e3614a9bef> [Consulta: 11 de Enero de 2020]

EQUIPO

/ Beate Autering

Diseñadora gráfica y maquetista

Beate Autering es una diseñadora gráfica freelance. Se graduó en Diseño y dirige el estudio beworx. Crea diseños, gráficos e ilustraciones y también proporciona servicios de edición y postproducción de imágenes. Entre sus clientes se encuentran iRights, mdsCreative, Agentur Sehstern, UNESCO World Heritage, y visitBerlín.

/ José Miguel Calatayud

Editor de la edición en español del informe y autor del capítulo de investigación sobre España.

José Miguel Calatayud es un periodista español freelance. Establecido en Barcelona entre el 2016 y el 2019, José se trasladó a Berlín en el 2020. Actualmente, cubre Europa y se centra en el periodismo de investigación, y está particularmente interesado en la democracia y los derechos humanos. Desde 2019, también ha trabajado con Arena for Journalism in Europe, una fundación sin ánimo de lucro dedicada a la promoción del periodismo colaborativo transfronterizo. En 2017, obtuvo una Open Society Fellowship para llevar a cabo investigaciones periodísticas sobre el activismo ciudadano político en Europa. Entre 2012 y 2014, estuvo establecido en Estambul y cubrió Turquía y la región circundante para el diario El País. Anteriormente, desde 2009esttuvo establecido en el este de África, donde fue corresponsal de la Agencia Efe, y, a partir de 2011, de El País. Su trabajo también ha aparecido en Foreign Policy, Al Jazeera, New Statesman, The Independent, Agence France-Presse, Radio France Internationale y Deutsche Welle, entre otros medios.

/ Michele Catanzaro

Autor del artículo periodístico sobre España

Michele Catanzaro es un periodista freelance establecido en Barcelona. Tiene un doctorado en Física, ha escrito para Nature, El Periódico de Catalunya, Süddeutsche Zeitung y otros medios, y es coautor del libro "Networks: A Very Short Introduction" (2012) y el documental "Fast Track Injustice: The Óscar Sánchez Case" (2014), que recibió el Premio Ninfa de Oro en 2015. Su trabajo también ha sido reconocido por otros premios: Premio Internacional de Periodismo Rey de España, BBVA Innovadata, Valors, Escritor Científico Europeo del Año 2016, Prismas y Colombine. Ha recibido subvencioes del Journalism Fund, Journalism Grants y Climate Investigation Grants. Tiene experiencia en docencia, exposiciones, televisión y eventos, y coordina el proyecto PerCientEx sobre la excelencia en el periodismo científico en España y América Latina. También fue periodista residente en el Instituto de Estudios Teóricos de Heidelberg.

/ Fabio Chiusi

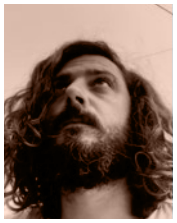
Editor del informe, de la introducción y del capítulo sobre la UE

Photo: Julia Bornkessel

Fabio Chiusi trabaja en AlgorithmWatch como coeditor y director de proyecto de la edición 2020 del informe "Automating Society". Después de una década dedicado al periodismo tecnológico, comenzó a trabajar como asesor e investigador asistente en datos y política (Tactical Tech) y en la IA en periodismo (Polis LSE). Coordinó el informe "Persuasori Social" sobre la regulación de las campañas políticas en las redes sociales para el Proyecto PuntoZero, y trabajó como funcionario de política tecnológica en la Cámara de Diputados del Parlamento italiano durante la legislación vigente. Fabio es miembro del Nexa Center for Internet & Society de Turín y profesor adjunto de la Universidad de San Marino, donde imparte clases de Periodismo y nuevos medios de comunicación, así como de Edición y medios digitales. Es autor de varios ensayos sobre tecnología y sociedad, el último de los cuales es "Io non sono qui. Visioni e inquietudini da un futuro presente" (DeA Planeta, 2018), que actualmente está siendo traducido al polaco y al chino. También escribe como reportero de política tecnológica para el blog colectivo ValigiaBlu.

/ Samuel Daveti

Dibujante de cómics



Samuel Daveti es miembro fundador de la Asociación Cultural Double Shot. Es autor de la novela gráfica en francés, Akron Le guerrier (Soleil, 2009), y editor del volumen antológico Fascia Protetta (Double Shot, 2009). En 2011, se convirtió en miembro fundador del colectivo de cómics autoeditados, Mammaiuto. También escribió Un Lungo Cammino (Mammaiuto, 2014; Shockdom, 2017), que va a ser convertido en una película para la compañía Brandon Box. En 2018, escribió I Tre Cani, con dibujos de Laura Camelli, que ganó el Premio Micheluzzi en el Napoli Comicon 2018 y el premio Boscarato al mejor webcomic en el Festival del Cómic de Treviso.

/ Sarah Fischer

Editora del informe

Sarah Fischer es directora del proyecto "Ética de los algoritmos" de la Bertelsmann Stiftung, donde es la principal responsable de los estudios científicos. Anteriormente, trabajó como becaria posdoctoral en el programa de postgrado "Trust and Communication in a Digitized World" (Confianza y comunicación en un mundo digitalizado) en la Universidad de Münster, donde se centró en el tema de la confianza en los motores de búsqueda. En el mismo grupo de formación de investigación, obtuvo su doctorado con una tesis sobre la confianza en los servicios de salud en internet. Estudió Ciencias de la comunicación en la Universidad Friedrich Schiller de Jena, y es coautora de los artículos "Where Machines can err. Sources of error and responsibilities in processes of algorithmic decision making" y "What Germany knows and believes about algorithms".

/ Leonard Haas

Edición adicional

Leonard Haas trabaja como asistente de investigación en AlgorithmWatch. Entre otras cosas, fue responsable de la concepción, implementación y mantenimiento del Inventario Global de Guías Éticas de la IA. Es estudiante de máster en el campo de las ciencias sociales en la Universidad Humboldt de Berlín y tiene dos títulos de la Universidad de Leipzig, en Humanidades digitales y Ciencias políticas. Sus investigaciones se centran en la automatización del trabajo y el gobierno. Además, está interesado en las políticas de datos de interés público y las luchas laborales en la industria tecnológica.

/ Lorenzo Palloni

Dibujante de cómics

Lorenzo Palloni es dibujante, autor de varias novelas gráficas y cómics web, escritor premiado, y uno de los fundadores del colectivo de dibujantes de cómics Mammaiuto. Actualmente, está trabajando en libros que se van a publicar en Francia e Italia. Lorenzo es también profesor de Escritura de guiones y Narración de cuentos en La Scuola Internazionale di Comics de Reggio Emilia.

/ Kristina Penner

Autora del capítulo de la UE

Photo: Julia Bornkessel

Kristina Penner es asesora ejecutiva en AlgorithmWatch. Sus intereses de investigación incluyen la ADM para sistemas de bienestar social, evaluación y clasificación social, y los impactos sociales que derivan de los sistemas de ADM, así como la sostenibilidad de las nuevas tecnologías desde un punto de vista holístico. Su análisis del sistema de gestión de fronteras de la UE se basa en su experiencia previa en la investigación y el asesoramiento sobre la ley de asilo. También tiene experiencia en proyectos sobre la utilización de los medios de comunicación por la sociedad civil y el periodismo de paz, así como la participación de los interesados en los procesos de paz en Filipinas. Tiene un Máster en Estudios internacionales / Investigaciones sobre la paz y los conflictos de la Universidad Goethe de Frankfurt.

/ Alessio Ravazzani

Dibujante de cómic

Alessio Ravazzani es un diseñador gráfico editorial, caricaturista e ilustrador que colabora con las más prestigiosas editoriales de cómics y novelas gráficas de Italia. Es autor en el colectivo Mammaiuto, del que es miembro desde su fundación.

/ Friederike Reinhold

Edición adicional de la introducción y recomendaciones de normativas

Photo: Julia Bornkessel

En su calidad de asesora superior en materia de normativas, Friederike Reinhold se encarga de impulsar las iniciativas de normativas y propuestas de cambio de AlgorithmWatch. Antes de unirse a AlgorithmWatch, trabajó como Asesora de Normativa Humanitaria en el Ministerio Federal de Relaciones Exteriores de Alemania, con el Consejo Noruego de Refugiados (NRC) en Irán, con Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) en Afganistán, y en el Centro de Ciencias Sociales de Berlín WZB.

/ Matthias Spielkamp

Editor del informe

Photo: Julia Bornkessel

Matthias Spielkamp es cofundador y director ejecutivo de AlgorithmWatch. Ha declarado como experto ante varios comités del Parlamento alemán sobre la IA y la automatización. Matthias es miembro del consejo de administración de la sección alemana de Reporteros sin Fronteras y de los consejos de asesoría de Stiftung Warentest y de la Red de Denunciantes. Ha sido miembro de ZEIT Stiftung, Stiftung Mercator, y el Consejo Americano sobre Alemania. Matthias fundó la revista online mobilsicher.de, que informa sobre la seguridad de los dispositivos móviles, con una audiencia de más de 170.000 lectores mensuales. Ha escrito y editado libros sobre periodismo digital y gobierno de Internet y fue nombrado por Silicon Republic uno de los 15 arquitectos que están construyendo el futuro basado en datos. Tiene un máster en Periodismo por la Universidad de Colorado en Boulder y en Filosofía de la Universidad Libre de Berlín.

/ Marc Thümmler

Coordinador de publicaciones

Photo: Julia Bornkessel

Marc Thümmler está a cargo de las relaciones públicas y la divulgación en AlgorithmWatch. Tiene un máster en Estudios de medios de comunicación, ha trabajado como productor y editor en una compañía cinematográfica y ha gestionado proyectos para el Deutsche Kinemathek y la organización de la sociedad civil Gesicht Zeigen. Además de sus tareas principales en AlgorithmWatch, Marc ha participado en la campaña de micromecenazgo y colaboración abierta distribuida OpenSCHUFA, y coordinó el primer número del informe de la Sociedad de Automatización, publicado en 2019.

ORGANIZACIONES

/ AlgorithmWatch

AlgorithmWatch es una organización sin ánimo de lucro dedicada a la investigación y a la propuesta de recomendaciones, y comprometida con la vigilancia y el análisis de los sistemas automatizados o algorítmicos de toma de decisiones (ADM, en inglés) y de su impacto en la sociedad. Si bien el uso prudente de sistemas de ADM puede beneficiar tanto al individuo como a la sociedad, también puede comportar grandes riesgos. Con el fin de proteger la autonomía humana y los derechos fundamentales y de maximizar el bien común, consideramos crucial que los sistemas de ADM estén sometidos al control democrático. El funcionamiento de sistemas de ADM que afecten significativamente los derechos individuales y colectivos no sólo debe ser público de un modo accesible y claro, sino que además los individuos deben poder entender cómo el sistema toma las decisiones e impugnarlas cuando sea necesario. Es por ello que trabajamos para permitir a los ciudadanos una mejor comprensión de los sistemas de ADM, y que desarrollamos medios para conseguir un control democrático de estos procesos, con una combinación de tecnologías, regulación e instituciones de supervisión adecuadas. De esta forma, tratamos de contribuir a sociedad más justa e inclusiva, y de maximizar el beneficio posible de los sistemas de ADM para todo el conjunto de la sociedad.

https://algorithmwatch.org/en/

ALGORITHM WATCH

/ Bertelsmann Stiftung

La Fundación Bertelsmann trabaja para fomentar la inclusión social para todas las personas. Está comprometida a promover este objetivo mediante programas dedicados a mejorar la educación, desarrollar la democracia, hacer progresar la sociedad, fomentar la salud, revitalizar la cultura y reforzar las economías. A través de sus actividades, la Fundación pretende animar a los ciudadanos a que contribuyan al bien común. Fundada en 1977 por Reinhard Mohn, esta fundación sin ánimo de lucro posee la mayoría de las acciones de Bertelsmann SE & Co. KGaA. La Fundación Bertelsmann es una fundación privada e independiente de partidos políticos. Con su proyecto “Ética de los algoritmos”, la Fundación Bertelsmann examina las consecuencias de los sistemas algorítmicos de toma de decisiones en la sociedad, con el fin de asegurar que esos sistemas se usan al servicio de la sociedad. El objetivo es ayudar a dar forma y desarrollar sistemas algorítmicos que promuevan una mayor inclusión social. Esto implica comprometerse con lo que es mejor para la sociedad más que con lo que es posible técnicamente, para que así las decisiones que dependen de las máquinas puedan servir a la humanidad de la mejor forma posible.

https://www.bertelsmann-stiftung.de/en

BertelsmannStiftung