

CAPÍTULO VII

‘Algoritmos predictivos’ y justicia penal desde una perspectiva italiana y europea

Laura Notaro

SUMARIO: 1. Introducción – 2. Algoritmos y ‘justicia predictiva’. Experiencias y tendencias recientes – 2.1. Herramientas de valoración del riesgo de reincidencia en Estados Unidos – 2.1.1. Las decisiones sobre el ‘*pre-trial release*’ y el debate sobre el ‘*money bail*’ – 2.1.2. El ‘*sentencing*’ y el caso ‘Loomis’ – 2.2. Herramientas de valoración del riesgo de reincidencia en Europa 2.3. Herramientas de ‘predicción’ del resultado decisorio en el ámbito civil. El caso francés – 3. Riesgos de la utilización de algoritmos ‘predictivos’ en el sistema de justicia penal – 3.1. Riesgos de las herramientas de valoración del riesgo de reincidencia – 3.1.1. Aspectos críticos conectados con la naturaleza estadístico-actuarial de la base científica – 3.1.2. Aspectos críticos conectados con el carácter ‘computacional’ de la herramienta – 3.2. Riesgos de las herramientas de ‘predicción’ del resultado decisorio – 4. Límites jurídicos a la utilización de algoritmos ‘predictivos’ en el sistema de justicia penal – 4.1. Instrumentos de nivel internacional. La Carta Ética del CEPEJ del Consejo de Europa – 4.2. Derecho de la Unión Europea: límites a las «decisiones basadas únicamente en el tratamiento automatizado» – 4.3. Límites del derecho italiano: la prohibición de pericia criminológica – 5. Conclusiones – 6. Bibliografía.

1. *Introducción*

El uso de herramientas de Inteligencia Artificial (I.A.) en el ámbito de la justicia penal se ha convertido, incluso en Europa, en un tema cada vez más urgente.

Aunque la experiencia más significativa de aplicación de tecnologías en el juicio penal es todavía la que ofrece el sistema jurídico estadounidense, hoy día es necesario, incluso a nivel europeo, llevar a cabo una reflexión que aborde la admisibilidad de sistemas algorítmicos dentro de la justicia penal. Establecer un debate teórico sobre la naturaleza de estas herramientas y sobre la compatibilidad de su utilización en materia penal, con respecto a los derechos fundamentales y los principios constitucionales de los Estados, permitiría anticipar el desarrollo de la tecnología y garantizar que la difusión

de las herramientas de I.A. en el ámbito de la justicia en general – y sobre todo de la justicia penal – sea preparada por una base cultural sólida y una regulación jurídica clara, adecuada para abordar los retos de la modernidad¹.

En el trabajo que voy a presentar, quiero empezar con una recolección de las principales herramientas predictivas desarrolladas en Estados Unidos y Europa y utilizadas en el sistema de justicia².

En segundo lugar, será posible identificar algunos de los riesgos que se pueden determinar para los principios básicos del derecho y del proceso penal y para los derechos fundamentales.

En tercer lugar, se recordarán algunas disposiciones jurídicas que podrían representar un límite al ingreso de herramientas algorítmicas en el sistema de justicia.

Por último, se intentará ofrecer algunas reflexiones a modo de conclusión sobre la necesidad de un uso responsable de los algoritmos en el sistema de justicia, sobre todo penal.

2. *Algoritmos y ‘justicia predictiva’. Experiencias y tendencias recientes*

En términos resumidos, las tecnologías predictivas que actualmente han sido desarrolladas pueden ser divididas en dos grupos en función del objeto del pronóstico – o sea, de la ‘predicción’ – que prometen hacer.

Por un lado, están los algoritmos que ofrecen una ‘predicción’ del riesgo de reincidencia – es decir, de reiteración del delito – de una persona, sea esa un sospechoso sujeto a investigación penal, un acusado en un procedimiento penal o una persona ya condenada y contra la que se esté ejecutando una pena.

En cuanto a estas herramientas, desarrolladas para ‘predecir’ el riesgo

¹ Señalan la necesidad de que la implementación de políticas públicas para la implementación de herramientas de I.A. sea anticipada por un debate público multidisciplinario, capaz de delinear límites jurídicos al desarrollo de algoritmos que garanticen el respeto de los derechos fundamentales X. RONSIN, V. LAMPOS, V. MAÏTREPIERRE, *In-depth study on the use of AI in judicial systems, notably AI applications processing judicial decisions and data*, en Appendix I, *European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment*, en <<http://www.coe.int>>.

² Queda fuera del análisis que aquí se propone la experiencia de Latinoamérica. Al respecto, cfr. C. RUÍZ LÓPEZ, V. SALAZAR, H. SIERRA OLIVIERI, *Sistemas operados mediante Inteligencia Artificial (IA) y debido proceso penal. Perspectivas de aplicación en Colombia*, en *Disrupción tecnológica, transformación digital y sociedad*, dirigido por J.C. Henao, D. Castaño, tomo III, Bogotá, Universidad Externado de Colombia, 2021, pp. 827 ss.

de cometer nuevos delitos, el punto de observación privilegiado puede identificarse en el ordenamiento de Estados Unidos³, donde los llamados RATs ('*risk assessment tools*') están ya muy extendidos⁴. En Europa, de momento, solo cabe señalar un ejemplo en el Reino Unido.

Por otro lado, deben tenerse en cuenta los algoritmos que ofrecen una predicción del resultado decisorio de un razonamiento judicial. Aunque esta segunda tipología de herramientas todavía no está tan extendida como las de predicción de riesgo y, en todo caso, su uso se limita a la esfera civil. La experiencia más interesante, en este ámbito, es el caso francés.

2.1. Herramientas de valoración del riesgo de reincidencia en Estados Unidos

En el contexto de los Estados Unidos, son esencialmente dos los ámbitos de utilización que más plantean cuestiones de compatibilidad con principios generales y derechos fundamentales.

En primer lugar, las decisiones sobre el llamado '*pre-trial release*', es decir,

³ Cabe señalar, sin embargo, que la cuestión de la difusión de los sistemas computacionales de evaluación de riesgos en la justicia penal estadounidense forma parte de una tendencia más general y antigua de recurrir, en este ámbito, a herramientas predictivas construidas sobre la base de teorías y métodos estadístico-actuariales. Sobre la afirmación de *evidence based practices* (EBP), D. KEHL, P. GUO, S. KESSLER, *Algorithms in the Criminal Justice System, Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing* (Responsive Communities Initiative, Berkman Klein Center for Internet & Society, Harvard Law School), julio 2017, en <<https://cyber.harvard.edu/publications/2017/07/Algorithms>>, p. 7. Destaca algunas cuestiones críticas en este sentido N. SCURICH, *The case against categorical risk estimates*, en «*Behavioral Science Law*», 2018, pp. 1 ss.

⁴ Aunque las herramientas computacionales utilizadas son muy numerosas, las más conocidas y extendidas son 'COMPAS' (*Correctional Offender Management Profiling for Alternative Sanctions*), 'LSI-R' (*Level of Service Inventory Revised*), las dos desarrolladas por empresas comerciales, y 'PSA' (*Public Safety Assessment*), creada por una fundación. Algunos estados, sin embargo, han desarrollado sus propios algoritmos predictivos. Por ejemplo, el *Department of Rehabilitation and Correction* de Ohio desarrolló el algoritmo 'ORAS' (*Ohio Risk Assessment System*) en colaboración con la Universidad de Cincinnati (en este respecto, KEHL, GUO, KESSLER, *Algorithms in the Criminal Justice System*, cit., p. 16). Sobre el algoritmo recientemente adoptado en Pensilvania, como resultado de un trabajo iniciado en 2010 por la *Pennsylvania Commission on Sentencing*, R. HESTER, *Evaluación de riesgos en la sentencia. La experiencia de Pensilvania*, en *Predictive Sentencing: Normative and Empirical Perspectives*, dirigido por J.W. de Keijser, J.V. Roberts, J. Ryberg, Hart Publishing, 2019, pp. 213 ss.; A. DIENER, *Pennsylvania's Proposed, Questionably Constitutional, Risk Assessment Instrument*, en «*Harvard Civil Rights – Civil Liberties Law Review*», 17 de octubre de 2019; A. BASHIR, *Pennsylvania's Misguided Sentencing Risk-Assessment Reform*, en «*The Regulatory Review*», 5 de noviembre de 2019.

con un cierto nivel de aproximación, la libertad provisional.

En segundo lugar, la fase del ‘sentencing’, en la cual se determina la medida de la pena y, de manera más amplia, el tratamiento sancionatorio.

En el ámbito penitenciario y, más en general, de ejecución de penas, también puede encontrarse un uso de algoritmos predictivos del riesgo de reincidencia por el ‘Department of Corrections’, que, sin embargo, parece haber planteado menos problemas, por lo menos en la actualidad del debate.

2.1.1. Las decisiones sobre el ‘pre-trial release’ y el debate sobre el ‘money bail’

Para las decisiones sobre el ‘pre-trial release’, en los Estados Unidos existen numerosas herramientas predictivas del riesgo de reincidencia⁵. Una de las más comunes es PSA (*Public Safety Assessment*)⁶, desarrollada por una organización privada sin ánimo de lucro, la *Arnold Foundation*.

La regulación sobre la utilización de *risk assessment tools* no es uniforme a lo largo de Estados Unidos. Algunos Estados expresamente permiten el uso de algoritmos predictivos para las decisiones en tema de *pre-trial release*, otros lo recomiendan y algunos lo establecen como obligatorio⁷.

⁵ En cuanto a la cuestión del *pre-trial release*, más de 20 diferentes *risk assessment tools* (RATs) se encuentran actualmente en uso en los tribunales de Estados Unidos. Sobre el uso de RATs en las decisiones sobre el *pre-trial release*, A.Z. HUQ, *Equidad racial en la justicia penal algorítmica*, en «Duke Law Journal», 2019, pp. 1043 ss. En la doctrina italiana, M. GIALUZ, *Quando la giustizia penale incontra l’intelligenza artificiale: luci e ombre dei risk assessment tools tra Stati Uniti ed Europa*, en «Diritto penale contemporaneo», 29 de mayo de 2019, p. 4.

⁶ ‘PSA’ es utilizado por cuatro Estados (Arizona, Kentucky, New Jersey e Utah) y en algunas importantes jurisdicciones dentro de Estados Unidos, por ejemplo, Allegheny County (Pittsburgh, Pennsylvania), Cook County (Chicago, Illinois), Harris County (Houston, Texas), Mecklenburg County (Charlotte, North Carolina), Milwaukee County (Wisconsin), San Francisco County (California). Véase, al respecto, la hoja informativa ‘Where is PSA currently used?’ en *advancingpretrial.org*. El *output* del sistema consiste en la atribución de una puntuación entre 1 y 6 para cada uno de los tres ‘riesgos’ que se evalúan: el de no presentación en la audiencia (*Failure To Appear* – FTA), el de nueva detención (*New Criminal Arrest* – NCA) y el de nueva detención para crimines violentos (*New Violent Criminal Arrest* – NVCA). Para los criterios de atribución de la puntuación tras la evaluación de riesgo, véase la hoja informativa ‘How It Works’, en *advancingpretrial.org*. Sobre el algoritmo PSA, en la doctrina italiana, GIALUZ, *Quando la giustizia penale incontra l’intelligenza artificiale*, cit., p. 7; F. BASILE, *Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*, en «Diritto penale e uomo», 29 de septiembre de 2019, p. 18.

⁷ En siete Estados (Alaska, Delaware, Hawái, Indiana, Kentucky, New Jersey, Vermont) la consideración del ‘output’ de estos instrumentos se requiere por la ley, por lo menos en algunos casos. En otros ocho Estados (Colorado, Illinois, Montana, Nueva York, Pensilvania, Rhode Island, Virginia, Virginia Occidental) su adopción está autorizada o recomendada.

Un dato que llama la atención en el marco del estudio de la experiencia estadounidense es la conexión entre el ingreso de *risk assessment tools* en el proceso decisonal sobre el *pre-trial release* y el debate en torno al '*money bail*'⁸. Interesante es el caso del Estado de California, donde en el año 2018 se abolió el *money bail* y se estableció la obligación de utilizar *risk assessment tools*. La reforma fue abrogada tan sólo dos años después por un referéndum, apoyado no solo por la llamada '*bail industry*' – es decir, las empresas de seguros que tenían interés al restablecimiento del *bail* – sino también por organizaciones y redes de activistas que se oponían al *bail*, pero al mismo tiempo señalaban los riesgos del empleo de algoritmos predictivos y, sobre todo, los efectos discriminatorios y la inaccesibilidad al funcionamiento de estas herramientas⁹.

2.1.2. El 'sentencing' y el caso 'Loomis'

El debate estadounidense se ha desarrollado también en torno al segundo ámbito de aplicación de algoritmos predictivos que se ha mencionado al principio: la fase del *sentencing*, en la cual el juez va a establecer el tratamiento sancionatorio que debe imponerse al condenado¹⁰.

Además, se encuentran algunas intervenciones legislativas a nivel estatal (en Idaho, Nueva York y California) que han introducido una regulación del uso de herramientas predictivas en las decisiones sobre el *pre-trial release*: requisitos de imparcialidad y no discriminación, garantías de transparencia y accesibilidad del sistema y de los datos que se ponen en la base del algoritmo, así como la provisión de validación periódica de la herramienta predictiva. En 2019, se señalaban el *House Bill 118* en Idaho, el *Senate Bill 1509* en el Estado de Nueva York y el *Senate Bill 36* en California. En este respecto, A. WIDGERY, *The Statutory Framework of Pretrial Release*, en <<http://www.ncsl.org>>, 8 de noviembre de 2020, p. 7.

⁸ Sobre el '*money bail*', en la doctrina italiana, V. TONDI, *Il Bail. La libertà su cauzione negli ordinamenti anglosassoni*, Cedam-Wolters Kluwer, Padua, 2016.

⁹ GIALUZ, *Quando la giustizia penale incontra l'intelligenza artificiale*, cit., p. 8; J.L. KOEPKE, D.G. ROBINSON, *Danger Ahead: Risk Assessment and the Future of Bail Reform*, en «Washington Law Review», 4/2018, p. 1725.

¹⁰ En la fase del *sentencing*, se utilizan más de 60 diferentes *risk assessment tools*. En este respecto, HUQ, *Racial Equity in Algorithmic Criminal Justice*, cit., p. 1075; A.M. BARRY-JESTER, B. CASSELMAN, D. GOLDSTEIN, *Should Prison Sentences Be Based on Crimes That Haven't Been Committed Yet?*, en <<http://fivethirtyeight.com>>, 4 de agosto de 2015. De manera más general, sobre el uso de algoritmos en el *sentencing*, KEHL, GUO, KESSLER, *Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing*, cit., pp. 13 ss. El uso de herramientas predictivas, sin embargo, está previsto como obligatorio en la legislación de algunos Estados. Requieren el uso de herramientas predictivas en la fase del *sentencing*, por ejemplo, las legislaciones de Arizona, Kentucky, Ohio y Pensilvania. Los dos últimos estados han desarrollado incluso su propio *risk assessment tool*. En este respecto, KEHL, GUO, KESSLER, *Algorithms in the Criminal Justice System*, cit., p. 16; HUQ, *Racial Equity in Algorithmic Criminal Justice*, cit., p. 1075.

Sin duda la herramienta más conocida entre las que se utilizan en este contexto es COMPAS (acrónimo que significa '*Correctional Offender Management Profiling for Alternative Sanctions*')¹¹. Al respecto cabe destacar dos elementos. En primer lugar, como sugiere el nombre, se trata de un algoritmo inicialmente desarrollado para ser utilizado no tanto en el *sentencing*, sino en el contexto de la ejecución penal. En segundo lugar, COMPAS ha sido desarrollado por una empresa privada y su funcionamiento está protegido por secreto empresarial.

Bien se conoce el caso '*Loomis*' de 2016, en el cual el Tribunal Supremo de Wisconsin tuvo que abordar la compatibilidad del uso de COMPAS en la fase del *sentencing* con los derechos del acusado¹².

Los aspectos destacados por *Loomis* se referían a la ausencia de una valoración individual en la decisión sobre el tratamiento sancionatorio y, a nivel procesal, a la inaccesibilidad al funcionamiento del algoritmo, que estaba protegido por el secreto empresarial y cuya validez, de hecho, no podía ser discutida por la defensa. El Tribunal Supremo de Wisconsin, aun confirmando la posibilidad de utilizar COMPAS en el *sentencing*, afirmó que era necesario circunscribir su uso a través de límites y precauciones.

La primera precaución indicada por el Tribunal Supremo de Wisconsin es la información al juez sobre la naturaleza y las características de la herramienta COMPAS. En primer lugar, se trata de un algoritmo protegido por secreto empresarial, que se ha invocado para evitar la divulgación de informaciones sobre su funcionamiento. En segundo lugar, la puntuación que se obtiene tras la evaluación de riesgo se basa en datos que se refieren a grupos de alto riesgo y no puede leerse como una evaluación de riesgo del individuo acusado. Además, se han expresado dudas sobre la clasificación desproporcionada de personas pertenecientes a comunidades minoritarias como individuos con un mayor riesgo de reincidencia. Por último, como se comentaba antes, el algoritmo no se había desarrollado para el *sentencing*, sino para ser utilizado por el '*Department of Corrections*' en las decisiones sobre tratamiento penitenciario, supervisión y libertad condicional.

¹¹ Es una herramienta predictiva computacional desarrollada en 1998 por la empresa 'Northpointe' (hoy 'Equivant'). La base de datos en la que se fundamentaba el funcionamiento del sistema consta de 300.000 casos, examinados por COMPAS entre enero de 2004 y noviembre de 2005. La entrada del algoritmo está representada por la información recogida del expediente del demandado y las respuestas proporcionadas por el sujeto a una serie de preguntas. El producto devuelto por el sistema consiste en un «*risk assessment*» y un «*needs assessment*». La sección de *risk assessment* asigna una puntuación en una escala de 1 a 10 para cada uno de los tres riesgos de reincidencia medidos: *pre-trial recidivism risk*, *general recidivism risk*, y *violent recidivism risk*.

¹² *State v. Loomis*, 881 N.W.2d 749, 753 (Wis. 2016).

A la luz de las perplejidades destacadas por el Tribunal Supremo, COMPAS no puede utilizarse nunca para decidir si a una persona debe aplicársele la pena de prisión o cuanto severa debe ser su sentencia. Sin embargo, COMPAS sí que puede ser utilizado – pero siempre como factor solo relevante y no determinante – para decidir sobre la aplicación de penas alternativas a la prisión, para imponer obligaciones en el marco de la libertad condicional y para evaluar si a la persona puede aplicársele la supervisión en comunidad.

Estos límites son acompañados por la obligación del juez de tomar en cuenta otros factores que apoyen la decisión. Es decir, que una sentencia no puede fundamentarse exclusivamente en una valoración de riesgo llevada a cabo por un algoritmo.

2.2. Herramientas de valoración del riesgo de reincidencia en Europa

Como ejemplo de utilización de tecnologías para la valoración de riesgo en el sistema de justicia penal en Europa, de momento, destaca la experiencia de Reino Unido, donde se ha experimentado el algoritmo denominado HART (*Harm Assessment Risk Tool*)¹³, desarrollado en el marco de una colaboración entre la Universidad de Cambridge y la Comisaría de Durham.

Esta herramienta clasifica los sospechosos en tres grupos, según su nivel de riesgo de reincidencia en los siguientes dos años (riesgo elevado, moderado o bajo) y se ha utilizado por la policía de Durham para seleccionar las personas con riesgo moderado de reincidencia que pueden acceder a un programa de diversión llamado 'Checkpoint'.

Uno de los temas más discutidos en la sociedad civil, sobre todo a la luz de los análisis realizados por organizaciones sin ánimo de lucro para la protección de los derechos fundamentales, se refiere a uno de los dos códigos postales incluidos entre los factores predictivos que constituyen el 'input' de HART. Se trata del 'Mosaic code', una herramienta 'geodemográfica' desarrollada y vendida por la empresa de marketing *Experian*¹⁴ que perfila a los adultos en el Reino Unido poniéndoles en 66 categorías. El algoritmo

¹³ M. OSWALD, J. GRACE, S. URWIN, G.C. BARNES, *Algorithmic risk assessment policing models: lessons from the Durham HART model and "Experimental" proportionality*, en «Information and Communications Technology Law», 2/2018, pp. 223 ss. En la doctrina italiana, GIALUZ, *Quando la giustizia penale incontra l'intelligenza artificiale*, cit., p. 10.

¹⁴ Mas información sobre el producto se puede consultar en <<http://www.experian.co.uk/assets/marketing-services/brochures/mosaic-ps-brochure.pdf>>.

trabaja sobre una base de 850 millones de datos, que incluye el origen étnico y la composición familiar, los datos relacionados con el empleo, la salud, el consumo, los resultados escolares y los datos de navegación en Internet¹⁵.

2.3. Herramientas de 'predicción' del resultado decisorio en el ámbito civil. El caso francés

Como se mencionó previamente, las herramientas de valoración de riesgo no representan la única experiencia de uso de algoritmos en el marco del sistema de justicia. Efectivamente, cabe tomar en cuenta los instrumentos algorítmicos que permiten 'predecir' el resultado decisorio, aunque su uso está todavía limitado a la materia civil.

En el caso de Francia, herramientas algorítmicas de este tipo, sobre todo como apoyo para los abogados¹⁶, se han desarrollado desde 2016, cuando se publicaron en Internet todos los pronunciamientos judiciales dictados en el territorio francés.

La posibilidad de utilizar esta enorme base de datos de acceso abierto ha permitido incluso una 'perfilación' de magistrados, como en el caso 'Supra Legem', en el cual se utilizó un algoritmo para comparar las decisiones de distintos jueces en materia de asilo político y demostrar los prejuicios de algunos de estos¹⁷. A partir de este caso surgió un amplio debate que en 2019 llevó a una intervención del legislador: para abordar las perplejidades planteadas por el poder judicial sobre los riesgos para la independencia de

¹⁵ Cfr. S. CARLO, *Big Brother Watch's written evidence on algorithms in the justice system*, en <<http://bigbrotherwatch.org.uk>>, pp. 1 ss.; BIG BROTHER WATCH TEAM, *A closer look at Experian Big Data and Artificial Intelligence in Durham Police*, en <<http://bigbrotherwatch.org.uk>>, 6 de abril de 2018; BIG BROTHER WATCH TEAM, *Police uses Experian marketing data for AI custody decisions*, en <<http://bigbrotherwatch.org.uk>>, 6 de abril 2018.

¹⁶ Pueden señalarse como mínimo cuatro aplicaciones de este tipo, desarrolladas por *Predictice*, *Case Law Analytics*, *Doctrine.fr*, *Tyr Legal*. Véase C. SZWARC, *La justice predictive: une autre justice?*, en «*Le Mag des Avocats*», 34, 9/2017, pp. 5 ss., disponible en el siguiente enlace: <<https://www.anased.fr/publications/mag34/files/assets/basic-html/page-5.html#>>. El Ministerio de Justicia francés ha comenzado a utilizar uno de estos algoritmos (*Predictice*) en los tribunales de apelación de Rennes y Douai. En este respecto, C. CASTELLI, D. PIANA, *Giustizia predittiva. La qualità della giustizia in due tempi*, en «*Questione giustizia*», 4/2018, p. 156.

¹⁷ El estudio fue realizado por el abogado Michaël Benesty y el ingeniero informático Anthony Sypniewski. Al respecto, M. LANGFORD, M.R. MADSEN, *France Criminalises Research on Judges*, en <<http://verfassungsblog.de>>, 22 de junio de 2019; M. BENESTY, *The Judge Statistical Data Ban. My Story*, en <<http://www.artificiallawyer.com>>, 7 de junio de 2019.

los jueces¹⁸ se introdujo la prohibición de publicar los datos de identificación de los magistrados y se establecieron sanciones penales para todas las formas de uso de estos datos para el análisis, la comparación y la elaboración de perfiles de los jueces¹⁹.

3. Riesgos de la utilización de algoritmos 'predictivos' en el sistema de justicia penal

Después del breve resumen de las herramientas algorítmicas utilizadas en el sistema de justicia que se acaba de presentar, cabe examinar los problemas que se han planteado al respecto. Se trata de considerar los riesgos para los principios del derecho y del proceso penal y para los derechos fundamentales. Al conducir este análisis debe recordarse, una vez más, la distinción fundamental entre algoritmos que ofrecen una valoración del riesgo de reincidencia y los que prometen 'predecir' el resultado decisorio²⁰.

3.1. Riesgos de las herramientas de valoración del riesgo de reincidencia

Empezando con el primer grupo – el de herramientas de valoración de riesgo – es útil poner de relieve una distinción recién sugerida en el marco de la doctrina²¹.

Por un lado, existen problemas y perplejidades que parecen relacionarse con la propia naturaleza estadístico-actuarial de la ciencia que sirve de base para la herramienta y el carácter estadístico de las correlaciones identificadas por el algoritmo²².

¹⁸ B. GALGANI, *Considerazioni sui "precedenti" dell'imputato e del giudice al cospetto dell'IA nel processo penale*, en «Sistema penale», 4/2020, pp. 87 ss.

¹⁹ V. art. 33, al. 3, de la *loi* n. 2019-222 de 23 marzo de 2019: «*Les données d'identité des magistrats et des membres du greffe ne peuvent faire l'objet d'une réutilisation ayant pour objet ou pour effet d'évaluer, d'analyser, de comparer ou de prédire leurs pratiques professionnelles réelles ou supposées. La violation de cette interdiction est punie des peines prévues aux articles 226-18, 226-24 et 226-31 du code pénal, sans préjudice des mesures et sanctions prévues par la loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés*».

²⁰ *Supra*, par. 2.

²¹ S. QUATTROCOLO, *Equo processo penale e sfide della società algoritmica*, en «BioLaw Journal – Rivista di BioDiritto», 1/2019, pp. 144 ss.; EAD., *Questioni nuove e soluzioni antiche? Consolidati paradigmi normativi vs rischi e paure della giustizia digitale "predittiva"*, en «Cassazione penale», 4/2019, pp. 1478 ss.

²² Por ejemplo, el sistema predictivo 'SAVRY' (*Structured Assessment of Violence Risk in*

Por otro lado, están algunos aspectos críticos directamente conectados con el carácter computacional e informatizado de la herramienta.

3.1.1. *Aspectos críticos conectados con la naturaleza estadístico-actuarial de la base científica*

En primer lugar, si el funcionamiento de la herramienta se basa en correlaciones estadísticas identificadas por el propio algoritmo a través del análisis de casos anteriores²³, el resultado predictivo puede conducir a una ‘desindividualización’ de la decisión: por ejemplo, la atribución a un sujeto de un nivel de riesgo elevado no derivaría en un juicio realizado con referencia al individuo específico, sino representaría una extensión a esta persona de datos estadísticos relativos a sujetos con las mismas características.

En el contexto de la aplicación de la pena, el problema descrito podría llegar a vulnerar principios como los de culpabilidad y personalidad de la responsabilidad penal, que requieren una evaluación individualizada no solo en el juicio sobre la responsabilidad por el delito, sino también en la determinación del tratamiento sancionatorio. En las decisiones sobre libertad provisional, el riesgo de una ‘desindividualización’ de la decisión podría ir en contra de las garantías para la protección de la libertad personal: la limitación de este derecho puede admitirse sólo como *extrema ratio*, en presencia de ciertas exigencias de prevención, que deben existir efectivamente, con referencia a la persona en concreto a la que se le aplica la medida coercitiva.

En segundo lugar, el problema de la ‘desindividualización’ de la decisión parece vinculado incluso al riesgo de efectos discriminatorios contra comunidades minoritarias, que ha sido objeto de debate en Estados Unidos y también entre los observadores europeos²⁴.

Youth), utilizado en la justicia juvenil de algunos Estados, es una herramienta con *estructura algorítmica*, pero sin carácter digital o computacional. En realidad, no es un *software*, sino un manual de usuario acompañado de tarjetas de evaluación. Sobre este tema, G.M. VINCENT, J. CHAPMAN, N.E. COOK, *Risk-Needs Assessment in Juvenile Justice: Predictive Validity of the SAVRY, Racial Differences, and the Contribution of Needs Factors*, en «Criminal Justice & Behavior», 2011, pp. 47 ss.; QUATTROCOLO, *Questioni nuove e soluzioni antiche? Consolidati paradigmi normativi vs rischi e paure della giustizia digitale “predittiva”*, cit., pp. 1478 ss.

²³ L. MALDONATO, *Algoritmi predittivi e discrezionalità del giudice: una nuova sfida per la giustizia penale*, en «Diritto penale contemporaneo», 2/2019, p. 411; L. D’AGOSTINO, *Gli algoritmi predittivi per la commisurazione della pena. A proposito dell’esperienza statunitense nel c.d. evidence-based sentencing*, en «Diritto penale contemporaneo», 2/2019, p. 356; GIALUZ, *Quando la giustizia penale incontra l’intelligenza artificiale*, cit., p. 21.

²⁴ MALDONATO, *Algoritmi predittivi e discrezionalità del giudice*, cit., 407; D’AGOSTINO, *Gli algoritmi predittivi per la commisurazione della pena*, cit., p. 364; GIALUZ, *Quando la*

Si el algoritmo 'se entrena' sobre una base de datos formada por precedentes históricos puede reproducir la discriminación implícita en las condiciones de marginalidad social de ciertos grupos, incluso cuando los datos étnicos – o, en general, los datos relativos a la pertenencia a un determinado grupo – no están incluidos entre la información introducida inicialmente en el sistema. El funcionamiento de las herramientas algorítmicas más avanzadas, de hecho, no se basa solo en los datos incluidos durante la programación, sino también en las conexiones identificadas por el propio algoritmo calculando la frecuencia de ciertos elementos. Por lo tanto, es posible que, al elaborar el conjunto de datos a su disposición, el sistema reconozca conexiones entre elementos que en realidad no están conectados a nivel etiológico²⁵ y, por consiguiente, produzca resultados discriminatorios²⁶. Por lo tanto, existe un problema de compatibilidad con el principio de igualdad, tanto en lo que respecta a la fase 'provisional', como en el momento de la determinación del tratamiento sancionador.

En tercer lugar, si el algoritmo está condicionado por prejuicios contra ciertos grupos, puede ocurrir que la pena para un condenado sea más severa debido a su pertenencia a un determinado colectivo y esto podría recordar una lógica de 'derecho penal del autor' o de determinismo penal²⁷.

Por último, cabe mencionar el riesgo de contaminación del juicio sobre la responsabilidad²⁸ debido al hecho de que en los sistemas europeos continentales la decisión sobre el tratamiento sancionador se toma en la misma audiencia en que se decide sobre la atribución de responsabilidad penal²⁹. Conocer la puntuación obtenida tras la evaluación de riesgo del acusado podría condicionar la decisión del juez sobre el hecho.

giustizia penale incontra l'intelligenza artificiale, cit., p. 21.

²⁵ MALDONATO, *Algoritmos predictivos y discreción del juez*, cit., p. 407.

²⁶ Véase RONSIN, LAMPOS, MAÏTREPIERRE, *In-depth study on the use of AI in judicial systems, notably AI applications processing judicial decisions and data*, cit., par. 56 ss., especialmente parr. 66, 71.

²⁷ En este sentido, V. MANES, *L'oracolo algoritmico e la giustizia penale*, en «Discrimen», 15 de mayo de 2020, p. 17; GIALUZ, *Quando la giustizia penale incontra l'intelligenza artificiale*, cit., p. 21.

²⁸ En este respecto, habla de «traslazione del giudizio» O. DI GIOVINE, *Il judge-bot e le sequenze giuridiche in materia penale*, en «Cassazione penale», 3/2020, p. 959. V. también MALDONATO, *Algoritmi predittivi e discrezionalità del giudice*, cit., p. 411; D'AGOSTINO, *Gli algoritmi predittivi per la commisurazione della pena*, cit., p. 367; A. GARAPON, J. LASSÈGUE, *Justice digitale. Révolution graphique et rupture anthropologique*, Puf, 2018, p. 279.

²⁹ Sobre la estructura bifásica que caracteriza el procedimiento penal en los ordenamientos jurídicos de *common law* y la autonomía de la fase del *sentencing*, D'AGOSTINO, *Gli algoritmi predittivi per la commisurazione della pena*, cit., p. 367.

3.1.2. Aspectos críticos conectados con el carácter 'computacional' de la herramienta

Además de los problemas mencionados, deben tenerse en cuenta los aspectos críticos que parecen conectados con el propio carácter 'computacional' de las herramientas predictivas de riesgo.

Un primer aspecto crítico es la ausencia de transparencia, debida al hecho que el sistema puede estar protegido por secreto empresarial y no permitirle el acceso a la defensa del acusado³⁰. La imposibilidad de acceder al sistema por parte de la persona destinataria de la decisión judicial contrasta de manera clara con el derecho al debido proceso y con el derecho de defensa: el juicio se desarrollaría en vulneración de los principios de contradicción e igualdad de armas, ya que la defensa no podría cuestionar la validez de la herramienta predictiva y su resultado, careciendo de medios adecuados para argumentar en contra de la parte motiva de la decisión que se refiere al 'output' del *software* predictivo.

Este aspecto podría remediarse mediante obligaciones de publicar los códigos fuente y toda la información relevante para la accesibilidad del algoritmo, a través de procedimientos competitivos para la selección del proveedor de servicios predictivos o prohibiendo de manera radical el uso de herramientas 'privadas' y reservando a las administraciones públicas el desarrollo de *softwares* para la utilización en el sistema penal.

Sin embargo, algunos problemas persistirían. En cualquier caso, existiría el riesgo de una anulación sustancial del principio de igualdad de armas y del derecho de defensa frente a la elevada especialización requerida para el estudio del algoritmo a efectos de cuestionar su validez y contradecir su resultado. El problema no es nuevo en absoluto ya que cuestiones críticas similares caracterizan cualquier juicio en el que se debata sobre una prueba científica y que requiera, por lo tanto, la participación de profesionales expertos. En el caso de los algoritmos predictivos, sin embargo, el problema podría verse agravado por la amplitud del campo de aplicación potencial de estas herramientas.

Además, no parece bastante claro si el acceso al código fuente puede ser suficiente para permitir a la defensa un conocimiento completo de los mecanismos de funcionamiento de la herramienta, especialmente en el caso de algoritmos capaces de reelaborar los datos introducidos durante la programación desarrollando formas de 'autoaprendizaje' (*machine-*

³⁰ GIALUZ, *Quando la giustizia penale incontra l'intelligenza artificiale*, cit., p. 5 examina algunos de los principales algoritmos empleados en Estados Unidos, distinguiendo entre algoritmos desarrollados por instituciones públicas, algoritmos «comerciales» y algoritmos *sin fin de lucro*.

learning)³¹.

En segundo lugar, además de un riesgo de manipulación³², cabe mencionar el llamado '*automation bias*'. Esto es un sesgo cognitivo que parece conducir a una confianza excesiva e injustificada en el resultado producido por la herramienta predictiva. Incluso si se atribuye a los algoritmos una función puramente consultiva o auxiliar podría observarse una tendencia de los jueces a elegir la solución conforme al resultado del sistema computacional, en lugar de la decisión que se hubiera tomado sin su utilización³³. Se podría generar, de hecho, una delegación sustancial a la herramienta algorítmica de la parte del procedimiento decisorio que se refiere a la valoración del riesgo de reincidencia, lo cual acabaría determinando una 'desresponsabilización' del juez.

3.2. Riesgos de las herramientas de '*predicción*' del resultado decisorio

Con referencia a los algoritmos predictivos del resultado decisorio como los que se han desarrollado en Francia, cabe distinguir entre el uso por particulares y el uso por el juez.

En el primer caso, el mayor riesgo es para la independencia del poder judicial. Sin embargo, la cuestión no es fácil, porque también la regulación introducida por el legislador francés – que prohíbe cualquier uso de datos identificativos de los jueces – podría plantear algunas perplejidades con referencia al control democrático del ejercicio de la función judicial y la libertad de pensamiento e información.

En cuanto al uso de algoritmos predictivos del resultado decisorio por el juez, los riesgos que deben señalarse son el sesgo cognitivo denominado '*automation bias*' y la consiguiente 'desresponsabilización' de los jueces³⁴, que tenderían a 'ratificar' el resultado ofrecido por la máquina.

En segundo lugar, cabe destacarse el riesgo de una 'estandarización' de las decisiones judiciales³⁵.

³¹ Cfr. MALDONATO, *Predictive algorithms and discretion of the judge*, cit., p. 408, habla de «inaccesibilidad ontológica del mecanismo de funcionamiento». V. también J.P. DAVIS, *Law without mind. AI, Ethics and Jurisprudence*, en «University of San Francisco Law research Paper», 2018, p. 6.

³² Cfr. DI GIOVINE, *Il judge-bot e le sequenze giuridiche in materia penale*, cit., p. 957.

³³ G. UBERTIS, *Intelligenza artificiale, giustizia penale, controllo umano significativo*, en «Sistema penale», 11 de noviembre 2020, p. 4; P. COMOGLIO, *Prefazione*, en J. NIEVA-FENOLL, *Intelligenza artificiale e processo*, trad. it. P. Comoglio, Giappichelli, Torino, 2019, X.

³⁴ BASILE, *Intelligenza artificiale e diritto penale*, cit., p. 22.

³⁵ UBERTIS, *Intelligenza artificiale, giustizia penale, controllo umano significativo*, cit., p. 13.

El problema no radica en la ‘estandarización’ de las decisiones en sí: efectivamente, la uniformidad de las opiniones jurisprudenciales es el objetivo de la actividad de tribunales superiores. La cuestión parece residir, más bien, en la forma en que se selecciona la jurisprudencia ‘dominante’: incluso en este caso, el algoritmo funciona sobre la base de correlaciones identificadas a partir de la recurrencia de ciertos grupos de palabras y, por lo tanto, no se fundamenta en la ponderación de la validez de los argumentos jurídicos contenidos en las sentencias.

4. Límites jurídicos a la utilización de algoritmos ‘predictivos’ en el sistema de justicia penal

Los aspectos problemáticos destacados en relación con la entrada de los sistemas algorítmicos en el sector de la justicia penal – como se ha visto en el resumen que se acaba de presentar – parecen mostrar perfiles de contraste con respecto a los principios clave del derecho y del juicio penal.

En cuanto a los *softwares* para evaluar el riesgo de reincidencia deben ponerse de relieve las cuestiones relacionadas con el principio de *personalidad de la responsabilidad penal*, el principio de *igualdad* y los principios de materialidad y lesividad que caracterizan el modelo de *derecho penal del hecho*; bajo el perfil procesal, se han examinado los problemas conectados con el *derecho de defensa* y el de *debido proceso*, que parecería vulnerado por la falta de transparencia de los sistemas informáticos, en particular los de propiedad privada y protegidos por secreto empresarial.

En cuanto a las aplicaciones que ofrecen una previsión del resultado decisorio, es especialmente su uso por los tribunales lo que plantea el problema de la ‘estandarización’ de las decisiones, la cual, sin adecuada claridad en cuanto al funcionamiento de la herramienta algorítmica, parecería vulnerar el principio de *legalidad* y el principio de *independencia del poder judicial*, según el cual el juez está sujeto únicamente a la ley³⁶.

Si bien muchas son las cuestiones críticas, es necesario en todo caso examinar las herramientas jurídicas que actualmente permiten tomar precauciones adecuadas, a nivel internacional, en el contexto del derecho eurounitario y del ordenamiento nacional, con respecto a la posible entrada de algoritmos predictivos en el sistema de justicia penal.

La ‘Carta Ética Europea sobre el uso de la inteligencia artificial en

³⁶ Véase el artículo 110 de la Constitución italiana y el artículo 117 de la Constitución española.

los sistemas judiciales y su entorno', adoptada por la CEPEJ (Comisión Europea Para la Eficiencia de la Justicia) del Consejo de Europa, aunque no es vinculante, sin duda tiene su relevancia en términos de interpretación del derecho vigente y como punto de partida y orientación para el legislador.

En el contexto del Derecho de la Unión Europea, la regulación sobre la protección de datos personales adquiere especial importancia en el discurso sobre los límites al ingreso de herramientas algorítmicas en el sistema penal.

Por último, con referencia al ordenamiento jurídico italiano, el art. 220, par. 2 del Código italiano de Enjuiciamiento Criminal podría constituir un límite al uso de *softwares* predictivos del riesgo de reincidencia.

4.1. *Instrumentos de nivel internacional. La Carta Ética de la CEPEJ del Consejo de Europa*

A nivel internacional, aunque se trata de una herramienta de *soft-law* (o 'derecho indicativo'), una referencia importante dentro del sistema del Consejo de Europa es la 'Carta Ética Europea sobre el uso de la inteligencia artificial en los sistemas judiciales y su entorno', adoptada por la Comisión Europea Para la Eficiencia de la Justicia (CEPEJ) en 2018.

La Carta establece cinco principios fundamentales: 1) respeto por los derechos fundamentales; 2) principio de no discriminación; 3) principio de calidad y seguridad de los datos; 4) principio de transparencia, imparcialidad y justicia; 5) principio 'bajo control del usuario' (garantía del control humano).

El estudio publicado en el primer apéndice de la Carta pone de relieve el tema de los mecanismos de funcionamiento de las herramientas de inteligencia artificial, destacando que las actualmente disponibles – enmarcándose en la noción de IA débil o moderada – no reproducen el modelo de razonamiento humano, sino que elaboran resultados estadísticos sobre la base de una cantidad muy elevada de datos, y presentan, por lo tanto, el riesgo de identificar correlaciones en realidad inexistentes³⁷. Con referencia a su uso en materia penal, además de destacar los problemas relacionados con los efectos discriminatorios³⁸, el estudio subraya la importancia de la accesibilidad a los algoritmos para la efectividad del derecho de defensa y

³⁷ RONSIN, LAMPOS, MAÎTREPPIERRE, *In-depth study on the use of AI in judicial systems, notably AI applications processing judicial decisions and data*, cit., par. 56 ss., especialmente par. 66, 71.

³⁸ *Ivi*, par. 137.

las garantías del debido proceso³⁹. También se pone especial atención en la cuestión de la protección de datos personales: en particular, se destaca la necesidad de un estudio previo de los riesgos del tratamiento de datos personales⁴⁰ con el fin de adoptar las medidas adecuadas para evitar o minimizar su impacto sobre los derechos fundamentales; además, se afirma el llamado «principio de lealtad» en el tratamiento, que prohíbe utilizar datos para fines distintos de los iniciales.

El segundo apéndice de la Carta divide las áreas de uso de la inteligencia artificial en asuntos de justicia en cuatro niveles de precaución: 1) «usos para ser adelantados»; 2) «posibles usos que requieren considerables precauciones metodológicas»; 3) «usos a considerar luego de estudios científicos adicionales»; 4) «usos a tener en cuenta con las reservas más extremas».

En lo que respecta al sector de la justicia penal, el primer grupo de posibles usos de la inteligencia artificial, que se consideran libres de riesgos, incluye el uso de herramientas avanzadas de búsqueda jurisprudencial. La tercera categoría de usos, que sólo podrá admitirse después de nuevas investigaciones científicas, incluye las herramientas para anticipar el resultado decisorio y la elaboración de datos jurisprudenciales destinados a perfilar a los jueces. Por último, entre los usos para los cuales se requiere el máximo nivel de precaución se encuentran, en primer lugar, la valoración del riesgo de reincidencia efectuada mediante la perfilación de personas y, en segundo lugar, los enfoques que atribuyen eficacia vinculante a la «masa» de precedentes seleccionada a través de una medición puramente cuantitativa de la jurisprudencia.

4.2. Derecho de la Unión Europea: límites a las «decisiones basadas únicamente en el tratamiento automatizado»

A nivel de derecho de la Unión Europea, una herramienta jurídica fundamental es el artículo 11 de la Directiva 680/2016 sobre la protección de datos personales, que establece la prohibición de decisiones basadas únicamente en el tratamiento automatizado de datos personales y que parece una sólida barrera frente a los usos más problemáticos de algoritmos predictivos en el sistema penal.

En el contexto del Derecho de la Unión Europea, la regulación adoptada en 2016 sobre la protección de datos personales (Reglamento UE 679/2016 y Directiva 2016/680) contiene algunas disposiciones de gran interés para

³⁹ *Ivi*, par. 138.

⁴⁰ *Ivi*, parr. 141 ss.

reflexionar sobre la posibilidad de utilizar algoritmos en el sistema judicial: con referencia al sector penal, cabe considerar el artículo 11 de la Directiva 2016/680⁴¹, que se refiere específicamente al tratamiento de datos por parte de las «Autoridades competentes fines de prevención, investigación, detección o enjuiciamiento de infracciones penales o de ejecución de sanciones penales».

El apartado primero de dicha disposición obliga a los Estados miembros a establecer la prohibición de decisiones basadas únicamente en un tratamiento automatizado, que produzcan efectos jurídicos negativos o afecten significativamente al interesado de modo similar⁴². Decisiones de este tipo solo podrán permitirse mediante disposiciones específicas de la Unión o del Estado miembro al que está sujeto el responsable del

⁴¹ Artículo 11 («Mecanismo de decisión individual automatizado»): «1. Los Estados miembros dispondrán la prohibición de las decisiones basadas únicamente en un tratamiento automatizado, incluida la elaboración de perfiles, que produzcan efectos jurídicos negativos para el interesado o le afecten significativamente, salvo que estén autorizadas por el Derecho de la Unión o del Estado miembro a la que esté sujeto el responsable del tratamiento y que establezca medidas adecuadas para salvaguardar los derechos y libertades del interesado, al menos el derecho a obtener la intervención humana por parte del responsable del tratamiento. 2. Las decisiones a que se refiere el apartado 1 del presente artículo no se basarán en las categorías especiales de datos personales contempladas en el artículo 10, salvo que se hayan tomado las medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado. 3. La elaboración de perfiles que dé lugar a una discriminación de las personas físicas basándose en las categorías especiales de datos personales establecidas en el artículo 10 quedará prohibida, de conformidad con el Derecho de la Unión».

⁴² Se prevé la prohibición de un contenido similar, con un alcance más general, en el art. 22 del Reglamento UE 2016/679 («Decisiones individuales automatizadas, incluida la elaboración de perfiles»): «1. Todo interesado tendrá derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles, que produzca efectos jurídicos en él o le afecte significativamente de modo similar. 2. El apartado 1 no se aplicará si la decisión: a) es necesaria para la celebración o la ejecución de un contrato entre el interesado y un responsable del tratamiento; b) está autorizada por el Derecho de la Unión o de los Estados miembros que se aplique al responsable del tratamiento y que establezca asimismo medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado, o c) se basa en el consentimiento explícito del interesado. 3. En los casos a que se refiere el apartado 2, letras a) y c), el responsable del tratamiento adoptará las medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado, como mínimo el derecho a obtener intervención humana por parte del responsable, a expresar su punto de vista y a impugnar la decisión. 4. Las decisiones a que se refiere el apartado 2 no se basarán en las categorías especiales de datos personales contempladas en el artículo 9, apartado 1, salvo que se aplique el artículo 9, apartado 2, letra a) o g), y se hayan tomado medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado». Cabe recordar que el art. 15 de la Directiva 95/46/CE ya preveía la prohibición de las decisiones basadas únicamente en un tratamiento automatizado.

tratamiento y siempre que se prevean garantías adecuadas para los derechos del interesado y, como mínimo, su derecho a obtener la intervención humana y a impugnar la decisión.

El apartado segundo prohíbe en general las decisiones basadas en el tratamiento automatizado de datos «relativos a condenas e infracciones penales o a medidas de seguridad conexas» – previstos en el artículo 10 de la Directiva – y las permite únicamente cuando existan medidas adecuadas para proteger al sujeto interesado.

Por último, el tercer apartado introduce una prohibición absoluta de las actividades de perfílación que produzcan efectos discriminatorios contra personas físicas, basándose en datos del tipo mencionado en el artículo 10.

Para determinar los límites del ámbito de aplicación de la prohibición a que se refiere el artículo 11 de la Directiva es decisiva la interpretación del concepto de «decisión basada únicamente en el tratamiento automatizado»⁴³.

La lectura ‘mínima’ es la que limita el objeto de la prohibición a las decisiones *enteramente* confiadas a una máquina; se permitiría, sin embargo, utilizar la tecnología predictiva como mero apoyo para la decisión, siempre que la intervención del operador humano no se limite a la mera ratificación de la ‘decisión’ de la máquina⁴⁴.

La formulación normativa, sin embargo, también podría permitir atribuir un significado más preciso a la prohibición de decisiones basadas en

⁴³ GIALUZ, *Quando la giustizia penale incontra l'intelligenza artificiale*, cit., p. 16.

⁴⁴ Véase M. BRKAN, *Do algorithms Rule the World? Algorithmic Decision-Making in the Framework of the GDPR and Beyond*, en «Electronic Journal», 2017, p. 10: «*In order for the decision not to be based solely on automated processing, the human judgment needs to be such as to verify the machine-generated decision and the human should assess the substance of the decision and not be involved merely as another (empty) procedural step. In other words, in order to escape the prohibition from Article 22 GDPR or Article 11 of the Directive on Data Protection in Criminal Matters, the human has to use the machine only as decision support, whereas the final decision is taken by the human*». En el mismo sentido, v. *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679 (17/IT; WP 251 rev.01)*, p. 23, en la versión revisada adoptada el 6 de febrero de 2018 por el Grupo de Trabajo sobre Protección de Datos del Artículo 29, establecido por el art. 29 de la Directiva 95/46/CE: «*El responsable del tratamiento no puede obviar las disposiciones del artículo 22 inventándose una participación humana. Por ejemplo, si alguien aplica de forma rutinaria perfiles generados automáticamente a personas sin que ello tenga influencia real alguna en el resultado, esto seguiría siendo una decisión basada únicamente en el tratamiento automatizado. Para ser considerada como participación humana, el responsable del tratamiento debe garantizar que cualquier supervisión de la decisión sea significativa, en vez de ser únicamente un gesto simbólico. Debe llevarse a cabo por parte de una persona autorizada y competente para modificar la decisión. Como parte del análisis, debe tener en cuenta todos los datos pertinentes*».

el tratamiento automatizado: según lo que sugiere una parte de la doctrina, se trataría de una regla de evaluación de la prueba que asigne al 'output' de la máquina el valor de un mero indicio, que en todo caso tendrá que corroborarse con otros elementos probatorios⁴⁵.

4.3. Límites del derecho italiano: la prohibición de pericia criminológica

En el derecho procesal italiano se encuentra otra disposición que parece representar un límite considerable al ingreso de tecnologías predictivas del riesgo de reincidencia como las que ya se utilizan a lo largo de Estados Unidos: se trata del segundo apartado del artículo 220 del «*Codice di procedura penale*» italiano (c.p.p.), que no permite el uso de conocimientos criminológicos y psicológicos relacionados con la personalidad del sujeto. Más precisamente, la disposición prohíbe, «*salvo lo dispuesto a efectos de la ejecución de la pena o de la medida de seguridad*», los «*informes periciales para establecer la habitualidad o profesionalidad en el delito, la tendencia a la delincuencia, el carácter y personalidad del imputado y en general las cualidades psíquicas independientes de causas patológicas*»⁴⁶.

La norma establecida por el legislador italiano de 1988, que mantuvo la regulación adoptada por el código anterior⁴⁷, fue motivada por dos

⁴⁵ GIALUZ, *Quando la giustizia penale incontra l'intelligenza artificiale*, cit., p. 16; G. MALGIERI, G. COMANDÈ, *Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation*, en «International Data Privacy Law», vol. 7, 1 de noviembre de 2017, p. 14; MANES, *L'oracolo algoritmico e la giustizia penale*, cit., p. 20.

⁴⁶ Sobre la prohibición de pericia psicológica y criminológica en el ordenamiento jurídico italiano, F. ERAMO, *Il divieto di perizie psicologiche nel processo penale: una nuova conferma dalla Cassazione*, en «Diritto penale e processo», 7/2007, pp. 931 ss.; P. MOSCARINI, *La perizia psicologica e il 'giusto processo'*, en «Diritto penale e processo», 8/2006, pp. 929 ss.; P. MARTUCCI, *Il contributo del criminologo nel processo penale: un problema ancora aperto*, en «Diritto penale e processo», 6/2004, pp. 744 ss.; I. GIANNINI, *Il dibattito sull'ammissibilità della perizia e della consulenza criminologica nel processo penale*, en «Rassegna penitenziaria e criminologica», 3/2003, pp. 87 ss.; A. SAPONARO, *L'esame della personalità del reo nel processo penale: evoluzione, involuzione, modelli alternativi, prospettive*, Cacucci, Bari, 2000, pp. 104 ss.; P. RIVELLO, *Perito e perizia*, en *Digesto delle discipline penalistiche*, vol. IX, pp. 474 ss.; D. BIELLI, *Periti e consulenti nel nuovo processo penale*, en «Giustizia penale», 2/1991, pp. 65 ss.; E. AMODIO, *Perizia e consulenza tecnica nel quadro probatorio del nuovo processo penale*, en «Cassazione penale», 1989, pp. 170 ss.; G. TRANCHINA, *Il divieto di perizia psicologica sull'imputato: una limitazione sicuramente anticostituzionale*, en «Rivista italiana di diritto e procedura penale», 1971, pp. 1325 ss.

⁴⁷ La formulación del artículo 314, el apartado 2 del «Codice di procedura penale» italiano de 1930 era esencialmente idéntica a la de hoy: «*Non sono ammesse perizie per stabilire l'abitudine o la professionalità nel reato, la tendenza a delinquere, il carattere e la*

preocupaciones fundamentales relacionadas con la necesidad de reforzar las garantías del acusado. En primer lugar, la prohibición de pericia criminológica y psicológica parece inspirarse en la presunción de inocencia, la cual podría verse perjudicada por un informe pericial sobre la personalidad del acusado introducido durante el juicio sobre el hecho. Además, en el segundo apartado del art. 220 c.p.p., cabe reconocer una preocupación por la libertad moral del acusado⁴⁸.

Para determinar si el uso de *softwares* de valoración de riesgos está comprendido en la prohibición establecida en el artículo 220 del Código italiano es necesario aclarar si el uso de estas herramientas puede asimilarse a la evaluación prohibida por la norma.

Una parte de la doctrina se ha expresado en sentido favorable a dicha asimilación: según esta opinión, herramientas de evaluación de riesgos como las utilizadas en Estados Unidos representarían una evaluación técnico-científica similar a la pericia ya que en la fase de programación del algoritmo se utilizarían modelos y teorías de la criminología y de la psicología⁴⁹.

5. Conclusiones

Para concluir – y sin poder considerar analíticamente las posibles soluciones para cada problema mencionado – un punto de partida probablemente un poco obvio pero fundamental para abordar el tema de la utilización de algoritmos predictivos en la justicia penal es garantizar una formación adecuada del personal judicial.

Es necesario aclarar que el mecanismo de funcionamiento de estas herramientas – por lo menos hoy en día – no reproduce el razonamiento humano, sino que funciona según diferentes mecanismos, sobre una base

personalità dell'imputato, e in genere le qualità psichiche indipendenti da cause patologiche».

⁴⁸ La misma atención por la libertad moral constituye la *ratio* del art. 188 c.p.p. italiano: «No podrán utilizarse métodos o técnicas capaces de influir en la libertad de autodeterminación o alterar la capacidad de recordar y evaluar los hechos, ni siquiera con el consentimiento de la persona interesada».

⁴⁹ En este sentido, QUATTROCOLO, *Questioni nuove e soluzioni antiche?*, cit., p. 1762. Llega a la misma conclusión MALDONATO, *Algoritmi predittivi e discrezionalità del giudice*, cit., p. 412, autora que destaca el hecho de que en el 'input' del *software* están incluidos datos relacionados con la personalidad del sujeto, incluso los que se extraen de sus respuestas al interrogatorio. Por otro lado, GIALUZ, *Quando la giustizia penale incontra l'intelligenza artificiale*, cit., p. 20, afirma que la asimilación entre el uso de herramientas de evaluación de riesgos y la pericia criminológica prohibida no es obvia.

principalmente estadística. Entonces se necesita que el usuario – ya sea el juez u otro profesional legal – sea consciente de los límites del resultado predictivo y, por lo tanto, pueda abordar críticamente la solución ofrecida por el algoritmo.

De hecho, la necesidad de un mayor conocimiento de la naturaleza del resultado predictivo del algoritmo se ha puesto de relieve tanto por el Tribunal Supremo de Wisconsin en Estados Unidos como por las instituciones de la Unión Europea.

Ya se ha visto que, entre las precauciones indicadas por los jueces estadounidenses, un espacio importante está ocupado por la inserción de informaciones específicas sobre el algoritmo COMPAS y sus límites.

En la misma dirección parece moverse, a nivel de la Unión europea, tanto la propuesta de Reglamento para establecer «normas armonizadas en materia de inteligencia artificial» de abril de 2021⁵⁰ – que está actualmente sometida en primera lectura al Parlamento europeo – como la Resolución del Parlamento europeo sobre la inteligencia artificial en el Derecho penal⁵¹, adoptada en octubre de 2021.

6. Bibliografía

AMODIO E., *Perizia e consulenza tecnica nel quadro probatorio del nuovo processo penale*, en «Cassazione penale», 1989, pp. 170 ss.

BARRY-JESTER A.M., CASSELMAN B., GOLDSTEIN D., *Should Prison Sentences Be Based on Crimes That Haven't Been Committed Yet?*, en <<http://fivethirtyeight.com>>, 4 de agosto de 2015.

BASHIR A., *Pennsylvania's Misguided Sentencing Risk-Assessment Reform*, en «The Regulatory Review», 5 de noviembre de 2019.

BASILE F., *Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*, en «Diritto penale e uomo», 29 de septiembre de 2019, pp. 1 ss.

BENESTY M., *The Judge Statistical Data Ban. My Story*, en <<http://www.artificiallawyer.com>>, 7 de junio de 2019.

⁵⁰ Véase la «Propuesta de Reglamento del Parlamento Europeo y del consejo por el que se establecen normas armonizadas en materia de inteligencia artificial (ley de inteligencia artificial) y se modifican determinados actos legislativos de la unión» (COM/2021/206 final) del 21 de abril de 2021, en *eur-lex.europa.eu*.

⁵¹ Véase la «Resolución del Parlamento Europeo, de 6 de octubre de 2021, sobre la inteligencia artificial en el Derecho penal y su utilización por las autoridades policiales y judiciales en asuntos penales (2020/2016(INI)) en <<http://www.europarl.europa.eu>>.

- BIELLI D., *Periti e consulenti nel nuovo processo penale*, en «Giustizia penale», 2/1991, pp. 65 ss.
- BIG BROTHER WATCH TEAM, *A closer look at Experian Big Data and Artificial Intelligence in Durham Police*, en <<http://bigbrotherwatch.org.uk>>, 6 de abril de 2018.
- BIG BROTHER WATCH TEAM, *Police uses Experian marketing data for AI custody decisions*, en <<http://bigbrotherwatch.org.uk>>, 6 de abril 2018.
- BRKAN M., *Do algoritms Rule the World? Algorithmic Decision-Making in the Framework of the GDPR and Beyond*, en «Electronic Journal», 2017, p. 1 ss.
- CARLO S., *Big Brother Watch's written evidence on algorithms in the justice system*, en <<http://bigbrotherwatch.org.uk>>, pp. 1 ss.
- CASTELLI C., PIANA D., *Giustizia predittiva. La qualità della giustizia in due tempi*, en «Questione giustizia», 4/2018, pp. 153 ss.
- COMOGGIO P., *Prefazione*, en NIEVA-FENOLL J., *Intelligenza artificiale e processo*, trad. it. Comoglio P., Giappichelli, Torino, 2019, pp. I ss.
- D'AGOSTINO L., *Gli algoritmi predittivi per la commisurazione della pena. A proposito dell'esperienza statunitense nel c.d. evidence-based sentencing*, en «Diritto penale contemporaneo», 2/2019, pp. 354 ss.
- DAVIS J.P., *Law without Mind. AI, Ethics and Jurisprudence*, en «University of San Francisco Law research Paper», 2018, p. 6.
- DI GIOVINE O., *Il judge-bot e le sequenze giuridiche in materia penale*, en «Cassazione penale», 3/2020, pp. 951 ss.
- DIENER A., *Pennsylvania's Proposed, Questionably Constitutional, Risk Assessment Instrument*, en «Harvard Civil Rights – Civil Liberties Law Review», 17 de octubre de 2019.
- ERAMO F., *Il divieto di perizie psicologiche nel processo penale: una nuova conferma dalla Cassazione*, en «Diritto penale e processo», 7/2007, pp. 931 ss.
- GALGANI B., *Considerazioni sui "precedenti" dell'imputato e del giudice al cospetto dell'IA nel processo penale*, en «Sistema penale», 4/2020, pp. 87 ss.
- GARAPON A., LASSÈGUE J., *Justice digitale. Révolution graphique et rupture anthropologique*, Puf, 2018, p. 279.
- GIALUZ M., *Quando la giustizia penale incontra l'intelligenza artificiale: luci e ombre dei risk assessment tools tra Stati Uniti ed Europa*, en «Diritto penale contemporaneo», 29 de mayo de 2019, pp. 1 ss.
- GIANNINI I., *Il dibattito sull'ammissibilità della perizia e della consulenza criminologica nel processo penale*, en «Rassegna penitenziaria e criminologica», 3/2003, pp. 87 ss.

- HESTER R., *Evaluación de riesgos en la sentencia. La experiencia de Pensilvania*, en *Predictive Sentencing: Normative and Empirical Perspectives*, dirigido por de Keijser J.W., Roberts J.V., Ryberg J., Hart Publishing, 2019, pp. 213 ss.
- HUQ A.Z., *Equidad racial en la justicia penal algorítmica*, en «Duke Law Journal», 2019, pp. 1043 ss.
- KEHL D., GUO P., KESSLER S., *Algorithms in the Criminal Justice System, Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing* (Responsive Communities Initiative, Berkman Klein Center for Internet & Society, Harvard Law School), julio 2017, en <<https://cyber.harvard.edu/publications/2017/07/Algorithms>>, pp. 1 ss.
- KOEPKE J.L., ROBINSON D.G., *Danger Ahead: Risk Assessment and the Future of Bail Reform*, en «Washington Law Review», 4/2018, pp. 1725 ss.
- LANGFORD M., MADSEN M.R., *France Criminalises Research on Judges*, en <<http://verfassungsblog.de>>, 22 de junio de 2019.
- MALDONATO L., *Algoritmi predittivi e discrezionalità del giudice: una nuova sfida per la giustizia penale*, en «Diritto penale contemporaneo», 2/2019, pp. 401 ss.
- MALGIERI G., COMANDÈ G., *Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation*, en «International Data Privacy Law», vol. 7, 1 de noviembre de 2017, p. 14.
- MANES V., *L'oracolo algoritmico e la giustizia penale*, en «Discrimen», 15 de mayo de 2020, pp. 1 ss.
- MARTUCCI P., *Il contributo del criminologo nel processo penale: un problema ancora aperto*, en «Diritto penale e processo», 6/2004, pp. 744 ss.
- MOSCARINI P., *La perizia psicologica e il 'giusto processo'*, en «Diritto penale e processo», 8/2006, pp. 929 ss.
- OSWALD M., GRACE J., URWIN S., BARNES G.C., *Algorithmic risk assessment policing models: lessons from the Durham HART model and 'Experimental' proportionality*, en «Information and Communications Technology Law», 2/2018, pp. 223 ss.
- QUATTROCOLO S., *Equo processo penale e sfide della società algoritmica*, en «BioLaw Journal – Rivista di BioDiritto», 1/2019, pp. 135 ss.
- QUATTROCOLO S., *Questioni nuove e soluzioni antiche? Consolidati paradigmi normativi vs rischi e paure della giustizia digitale "predittiva"*, en «Cassazione penale», 4/2019, pp. 1478 ss.
- RIVELLO P., *Perito e perizia*, en *Digesto delle discipline penalistiche*, vol. IX,

- pp. 474 ss.
- RON SIN X., LAMPOS V., MAÎTREP IERRE V., *In-depth study on the use of AI in judicial systems, notably AI applications processing judicial decisions and data*, en *Appendix I, European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment*, en < <http://www.coe.int>>.
- RUÍZ LÓPEZ C., SALAZAR V., SIERRA OLIVIERI H., *Sistemas operados mediante Inteligencia Artificial (IA) y debido proceso penal. Perspectivas de aplicación en Colombia*, en *Disrupción tecnológica, transformación digital y sociedad*, dirigido por Henao J.C., Castaño D., tomo III, Bogotá, Universidad Externado de Colombia, 2021, pp. 827 ss.
- SAPONARO A., *L'esame della personalità del reo nel processo penale: evoluzione, involuzione, modelli alternativi, prospettive*, Cacucci, Bari, 2000.
- SCURICH N., *The case against categorical risk estimates*, en «Behavioral Science Law», 2018, pp. 1 ss.
- SZWARC C., *La justice predictive: une autre justice?*, en «Le Mag des Avocats», 34, 9/2017, pp. 5 ss.
- TONDI V., *Il Bail. La libertà su cauzione negli ordinamenti anglosassoni*, Cedam-Wolters Kluwer, Padua, 2016.
- TRANCHINA G., *Il divieto di perizia psicologica sull'imputato: una limitazione sicuramente anticostituzionale*, en «Rivista italiana di diritto e procedura penale», 1971, pp. 1325 ss.
- UBERTIS G., *Intelligenza artificiale, giustizia penale, controllo umano significativo*, en «Sistema penale», 11 de noviembre 2020, pp. 1 ss.
- VINCENT G. M., CHAPMAN J., COOK N. E., *Risk-Needs Assessment in Juvenile Justice: Predictive Validity of the SAVRY, Racial Differences, and the Contribution of Needs Factors*, en «Criminal Justice & Behavior», 2011, pp. 47 ss.
- WIDGERY A., *The Statutory Framework of Pretrial Release*, en <<http://www.ncsl.org>>, 8 de noviembre de 2020, p. 1 ss.