

LA IMPOSIBILIDAD DE UN JUEZ. REALISMO JURÍDICO, INTELIGENCIA ARTIFICIAL Y LA BÚSQUEDA DE UN JUSTO MEDIO¹

THE IMPOSSIBILITY OF A JUDGE. LEGAL REALISM, ARTIFICIAL INTELLIGENCE AND THE SEARCH FOR A MIDDLE GROUND

Dyango Bonsignore Fouquet^{1,a} 

¹ Profesor Ayudante Doctor. Departamento de Derecho Internacional Público y Derecho Penal, Facultad de Derecho, Universidad de Alicante, Carretera San Vicente del Raspeig s/n, San Vicente del Raspeig, Alicante, Carretera San Vicente del Raspeig s/n, San Vicente del Raspeig, Alicante, España

 a.d.bonsignore@ua.es

Resumen

Este artículo contrasta dos líneas teóricas que han tendido a tensionar la labor judicial desde puntos de vista que, tal vez, cabría considerar contrapuestos. Por un lado, se recupera la clásica crítica del “realismo jurídico” que cuestiona la capacidad del juez para decidir con arreglo a las exigencias del ordenamiento jurídico. Aquí, el problema estriba en el carácter excesivamente “humano” del juzgador. Por otro lado, se da voz a la protesta inversa, surgida del debate en torno a las posibilidades de una inteligencia artificial judicial. En este contexto, las opiniones críticas han tendido a plantear que los algoritmos carecen de ciertas cualidades (estructurales y funcionales, pero también “sociológicas”) que imposibilitan la plena sustitución del juzgador humano. La “artificialidad” de la inteligencia es ahora el problema. Se reúnen ambas perspectivas con el propósito de determinar si, entre las presiones por la “abstracción” y la “humanización”, queda espacio alguno para un juez.

Palabras clave: Inteligencia artificial; Derecho penal; IA judicial; realismo jurídico; justicia algorítmica.

Abstract

This paper contrasts two theoretical approaches that have strained the work of the judiciary from an opposing point of view. On the one hand, the classic critique of "legal realism", which challenges the judge's ability to decide in accordance with the requirements of the legal system, is discussed. Here, the problem lies in the excessive "humanity" of the judge. Conversely, the reverse argument arises from the debate about the possibilities of judicial artificial Intelligence. In this context, criticism has been that algorithms lack certain qualities (structural and functional, but also "sociological") that make it impossible to fully replace the human judge. The "artificiality" of intelligence is the concern here. Both perspectives are brought together in order to determine whether, between the pressures for "abstraction" and "humanization", remains any room for a judge.

Keywords: Artificial intelligence; Law; judicial AI; legal realism; algorithmic justice.

¹ Este artículo es fruto de los proyectos de investigación: “Manifestaciones de desigualdad en el actual sistema de justicia penal: examen crítico de las razones de necesidad, oportunidad y peligrosidad para la diferencia (AEQUALITAS)” (RTI2018-096398-B-I00) del Ministerio de Ciencia, Innovación y Universidades, y “Legal Reasoning & Cognitive Science” (2020-1-IT02-KA203-079834) de la Comisión Europea.

1. INTRODUCCIÓN

En el ámbito de las ciencias sociales y jurídicas es corriente la impresión de que una controversia no se encuentra nunca puramente cerrada y que la paz doctrinal llega con el abandono de los combatientes. Sin embargo, ocasionalmente, un cambio en las circunstancias permite el regreso de viejas discusiones al primer plano, tal vez bajo la premisa implícita de que una nueva realidad vaya a hacer posible la victoria definitiva de alguno de los bandos en liza. Algo así sucede con la temática de la que se ocupa este trabajo.

Es bien conocido el extraordinario interés que vienen suscitando la inteligencia artificial (IA) y sus aplicaciones en tiempos recientes. Aunque la IA como campo o ámbito de las ciencias de la computación existe, al menos, desde la década de 1950², el advenimiento de la sociedad digital (con todas sus condiciones de posibilidad) ha propiciado un salto cualitativo en la relevancia social de la inteligencia artificial. Al calor de este seísmo sociotecnológico, de alcance aún incierto, parece haber brotado también el impulso por analizar las posibilidades de la IA desde otros campos menos directamente vinculados con el desarrollo científico-técnico, como la sociología o el Derecho. Naturalmente, hay en esto parte de reconocimiento de un estado de cosas: la inteligencia artificial ya es objeto de un uso extensivo en ámbitos que, tradicionalmente, han concernido a las ciencias “humanas”. Asimismo, el carácter “omnívoros”³ que se atribuye a este tipo de tecnologías de aprendizaje automático conduce a un pronóstico de colonización total: incluso el menos futurista de los observadores parece dispuesto a reconocer que la penetración de la IA en todos y cada uno de los ámbitos de la vida social no es una cuestión de *si*, sino de *cuándo* y de *hasta dónde*. Así las cosas, la naturaleza *disruptiva*⁴ que con frecuencia se atribuye a la inteligencia artificial enfrenta a las sociedades contemporáneas a una situación que, seguramente, quepa considerar genuinamente nueva. Un momento de *crisis*, en sentido propio.

Este es, precisamente, el contexto habilitante para la resurrección de un debate jurídico clásico, revisado a la luz de las nuevas preguntas y respuestas aportadas por la inteligencia artificial. Interesa de manera especial, en este trabajo, el modo en que los avances en inteligencia artificial han propiciado una renovada reflexión sobre la función judicial, azuzada por la hipótesis (verosímil o no) del eventual reemplazo de los jueces humanos por un “Juez IA”⁵. Como veremos, este debate ha enfrascado rápidamente a parte de la literatura en una indagación sobre qué implica “ser” un juez, cuáles son los propósitos irrenunciables que trata de alcanzar esta función social peculiar, y en qué lugar dejan las inteligencias artificiales judiciales (actuales o proyectadas) a sus imperfectos homólogos humanos.

Pero este no solo no es un debate nuevo, sino que tal vez sea uno de los más clásicos, dada la relevancia que las sociedades humanas parecen haber concedido tradicionalmente a la tarea de juzgar. Como no sería posible retrazar extensivamente la historia de esta institución aquí, el propósito de este trabajo es más delimitado. Sugerimos tomar como interlocutor al

² GROSAN, C. & ABRAHAM, A.: *Intelligent Systems. A Modern Approach* Berlin: Springer, 2011, p. 3.

³ MEHOZAY, Y. & FISHER, E.: “The Epistemology of Algorithmic Risk Assessment and the Path towards a Non-Penology Penology”, *Punishment and Society*, vol. 21, nº. 5, 2019, p. 534, <https://doi.org/10.1177/1462474518802336>.

⁴ SOURDIN, T.: *Judges, Technology and Artificial Intelligence. The Artificial Judge* Cheltenham, UK: Edward Elgar Publishing, 2021, p. 2.

⁵ SOURDIN, T.: “Judge v Robot? Artificial Intelligence and Judicial Decision-Making”, *UNSW Law Journal*, vol. 41, nº. 4, 2018, p. 1123; VOLOKH, E.: “Chief Justice Robots,” *Duke Law Journal*, vol. 68, nº. 6, 2019, p. 1134.

“realismo jurídico” (o a una versión condensada de aquel) y su acerada crítica al proceso decisonal seguido por los jueces. A esta corriente de trabajos se quisiera oponer otra, que procede del auge de la discusión sobre IA y Derecho, y que ha criticado la hipótesis del “Juez IA” defendiendo, en su lugar, el irrenunciable elemento humano de la labor de juzgar. En resumen, opondremos al realismo lo que podríamos denominar un “humanismo” jurídico, bajo la premisa de que ambos tienen en común una interrogación fundamental sobre qué constituye un buen juez.

En particular, mientras que desde un extremo parece haberse defendido que los jueces son demasiado humanos para cargar sobre sus hombros las expectativas depositadas por el Derecho (imparcialidad, racionalidad, justicia incluso), desde el otro extremo parece estarse afirmando que la inteligencia artificial, lejos de ser una respuesta a tales problemas, sería demasiado fría y rígida, demasiado *artificial* para satisfacer las expectativas depositadas en la labor judicial. La pregunta que orienta este trabajo estriba, precisamente, en averiguar si subsiste un resquicio entre estas posiciones en el que pueda existir, precisamente, algo parecido a un juez.

El trabajo se estructura como sigue: en primer lugar, se propone abordar de manera condensada una exposición del realismo jurídico, en sus variantes clásicas y contemporáneas. El objetivo no es una descripción exhaustiva de la escuela y sus autores, sino la extracción de algunas de las observaciones y objeciones más importantes planteadas por los “realistas” en relación con el tema a discutir. En segundo lugar, se seguirá idéntico procedimiento para el análisis del argumentario “humanista” buscando destacar lo más interesante de esta defensa del factor humano en la labor judicial. Finalmente, se aborda una reflexión de conjunto, fruto del contraste de ambas posiciones, dirigida a subrayar aquellos aspectos que la controversia analizada pone de manifiesto y las pistas que se deducen para la continuidad del debate.

2. LA JUSTICIA DEL “MUNDO REAL”. REALISMO JURÍDICO Y LOS LÍMITES DEL JUEZ HUMANO

2.1 Pinceladas de iusrealismo

El realismo jurídico ocupa una posición particular en el ámbito de la teoría jurídica. Si prescindimos de matices que en otro contexto serían necesarios, diría que esta posición particular deriva de lo que podría considerarse común a los “realistas”, a saber, una cierta *actitud* en relación con el Derecho y sus operadores. Pese a los riesgos de generalizar demasiado, un aspecto primordial de esta actitud sería tal vez la combinación de pragmatismo y escepticismo⁶: un pragmatismo que entiende que el auténtico Derecho es el que se presenta “en acción” y un escepticismo que cuestiona que la práctica jurídica esté a la altura de los grandes principios y valores que vertebran la teoría jurídica. Aunque seguramente los argumentos del realismo hayan sido exagerados artificialmente en ocasiones⁷, no es difícil de imaginar por qué este tipo de perspectiva pueda ser difícil de asumir por el colectivo de juristas más cercano a la teoría.

⁶ ATIENZA, M.: *Curso de Argumentación Jurídica* Madrid: Trotta, 2013, p. 25–26; CAMPOS ZAMORA, F.J.: “Nociones Fundamentales Del Realismo Jurídico”, *Revista de Ciencias Jurídicas*, nº. 122, 2010, p. 199.

⁷ SCHULTZ, T.: “Arbitral Decision-Making: Legal Realism and Law & Economics”, *Journal of International Dispute Settlement*, vol. 6, nº. 2, 2015, p. 233.

Y es que sobre el Derecho pesa una tradición y unas expectativas que podrían entenderse como incompatibles con la posición adoptada por los realistas. En este sentido, la teoría jurídica se enmarca con frecuencia en continuidad con la Razón ilustrada que sentó las bases de los ordenamientos jurídicos auténticamente modernos. Desde este punto de vista, se deduce una imagen del jurista (y entre ellos, el juez particularmente) como un “teórico aplicado”, una suerte de pequeño *philosophe* cuyo comportamiento derivaría de la combinación de un conocimiento profundo de las leyes y principios jurídicos aplicables, junto con notables competencias deliberativas⁸. Llevada a sus últimas consecuencias, esta representación algo simplificada del ideal del “jurista ilustrado” implicaría también un compromiso profesional con la realización de la ley por encima de cualquier otra prioridad. Tal vez las palabras de la Jueza Sotomayor condensen bien la idea, cuando afirma que “no es el corazón el que conduce a la conclusión de un caso, es la ley”⁹.

Sin duda, el retrato anterior es una exageración a efectos meramente didácticos, aunque no resulta tan extravagante apuntar a que, en ocasiones, se espera que los operadores jurídicos se comporten de manera acorde con su “rol ideal”¹⁰. Dicho esto, esta representación un tanto academicista de las profesiones jurídicas se encuentra en frecuente conflicto con algunas de las principales características que rigen la práctica. No es necesario pensar en profesionales de dudosa integridad para reconocer que, con frecuencia, lidiar con el Derecho puede empujar las prioridades del “jurista ilustrado” a un segundo plano. La dinámica adversarial del proceso, por ejemplo, puede ser desvinculada perfectamente de las funciones de esclarecimiento de los hechos, de la verdad, de la justicia, o de cualquier otro valor trascendental, quedando reducida a un juego de suma cero con ganadores y perdedores¹¹. Visto a la manera de Bourdieu, la práctica jurídica aparece como un juego competitivo en el que la norma es tanto el conjunto de reglas que lo definen como el instrumental del que valerse para alcanzar la victoria¹².

Incluso los jueces, cuya posición se encuentra diseñada, en principio, para garantizar su independencia y neutralidad, pueden tener considerables dificultades para ver en la Ley un marco de referencia completo y exhaustivo que suministra una respuesta determinada para cada caso concreto¹³. Precisamente en este punto es donde se insertan algunas de las críticas

⁸ Similar caracterización hace LAPORTA SAN MIGUEL, F.J.: “La Creación Judicial y El Concepto de Derecho Implícito”, *Revista Jurídica Universidad Autónoma de Madrid*, n. 6, 2002, p. 147. Esta idea es similar a la conocida alusión de Dworkin al “juez Hércules”, si bien no se pretende aquí trazar esa asociación más allá del ejemplo. DWORKIN, R.: “Hard Cases”, *Harvard Law Review*, vol. 88, n.º. 6, 1975, p. 1057–1109, <https://doi.org/10.2307/1340249>.

⁹ “It’s not the heart that compels conclusions in cases, it’s the law”. Citado en WISTRICH, A.J. & RACHLINSKI, J.J.: “Implicit Bias in Judicial Decision Making. How It Affects Judgment and What Judges Can Do About It”, *Enhancing Justice*, American Bar Association, 2017, p. 96. Sin perjuicio de la relevancia pública de Sotomayor en el contexto estadounidense, aquí se la trae a colación por la expresividad de la cita, frecuentemente recuperada por la literatura examinada.

¹⁰ “Puedan o no los jueces tomar decisiones desapasionadas, los políticos y el público esperan, e incluso exigen, que así lo hagan”. Vid. SPAIN BRADLEY, A.: “The Disruptive Neuroscience of Judicial Choice”, *UC Irvine Law Review*, vol. 9, n.º. 1, 2018, p. 3; WISTRICH, A.J., RACHLINSKI, J.J., & GUTHRIE, C.: “Heart versus Head: Do Judges Follow the Law or Follow Their Feelings?”, *Texas Law Review*, vol. 93, 2015, p. 857.

¹¹ LEITER, B.: “Realismo Jurídico Estadounidense”, en FABRA ZAMORA, J.L. & NÚÑEZ VAQUERO, Á. (eds.), *Enciclopedia de Filosofía y Teoría Del Derecho*, Vol. I, México: UNAM, Instituto de Investigaciones Jurídicas, 2015, p. 247. Bastante ilustrativa del asunto es la descripción de la profesión jurídica efectuada en NIETO, A. & FERNÁNDEZ, T.-R.: *El Derecho y El Revés. Diálogo Epistolar Sobre Leyes, Abogados y Jueces*, Ariel, 2006, p. 10–17.

¹² BOURDIEU, P. & WACQUANT, L.: *Una Invitación a La Sociología Reflexiva*, Siglo XXI, 2005, p. 150–52.

más memorables del realismo jurídico, entre cuyos protagonistas se encuentran reputados (e introspectivos) jueces.

Jerome Franck, uno de los autores clásicos asociados al realismo jurídico, criticó fuertemente el retrato clásico de la profesión jurídica y el Derecho como una forma de “fundamentalismo legal” y de “fetichismo normativo”¹⁴. De manera más analítica, aunque igualmente expresiva, Posner diría que:

La “ley” en un contexto judicial es simplemente el material, en el más amplio de los sentidos, con el que los jueces estilizan [*fashion*] sus decisiones. Puesto que el material para la toma de decisiones legalista es incapaz de generar respuestas aceptables para todos los interrogantes legales (...), los jueces por fuerza recurren en ocasiones -bastante frecuentes- a otras fuentes de juicio, incluyendo sus propias opiniones o argumentos políticos, o incluso su idiosincrasia.¹⁵

En este tipo de fragmento se encuentran ya algunos elementos nucleares del realismo y sus sucesores. De hecho, resulta sumamente interesante la representación del proceso decisional del juez que se infiere de la exposición de Posner, y que apunta ya a una distinción fundamental entre la *decisión* en sí misma, y su *justificación*. Sobre este aspecto en particular el autor explica:

El papel del inconsciente en la toma de decisiones del juez queda ocultado por la convención que requiere al juez que explique su decisión a través de una opinión. La opinión judicial puede entenderse mejor como un intento de explicar cómo la decisión (...) incluso alcanzada por intuición, podría haber sido fruto de un razonamiento lógico paso a paso.¹⁶

Obsérvese que, en tales casos, la formulación de una justificación pertinentemente motivada se establece como un filtro sobre la decisión intuitiva, cuya rapidez y espontaneidad vienen ligadas a serias dudas sobre su origen y legitimidad. Se intuye ya una de las objeciones más fuertes dirigidas al legalismo: si el origen de la decisión pertenece al inconsciente, surge la duda de si las piezas con las que se ha formado serían jurídicamente aceptables, en caso de poder ser explicitadas. En cierto sentido, la justificación extiende un manto de legitimidad *ex post*, reconstruyendo a la inversa el proceso racional que conduce de las premisas a las conclusiones, de los hechos al Derecho¹⁷. Aunque se parta de una decisión de origen indeterminado a la que se trata de reconducir al marco normativo vigente, si el argumento del juez convence, la decisión se transforma de premisa en conclusión. Pero todo este mecanismo, que podría parecer razonable a la vista de las limitaciones de la cognición humana, suscita en Posner un comentario inquietante:

[La justificación] es un filtro [*check*] imperfecto, sin embargo, porque el voto sobre cómo ha de decidirse el caso precede a la opinión, y (...) la mayoría de los jueces no tratan la decisión (...) como una hipótesis que probar por medio de una investigación más profunda

¹³ DROBAK, J.N. & NORTH, D.C.: “Understanding Judicial Decision-Making: The Importance of Constraints on Non-Rational Deliberations”, *Journal of Law & Policy*, vol. 865, nº. 1988, 2004, p. 134.

¹⁴ FRANK, J.: *Law and the Modern Mind*, Stevens & Sons, 1949, p. 48 y 163.

¹⁵ POSNER, *How Judges Think*, cit. en KNIGHT, J.: “Are Empiricists Asking the Right Questions about Judicial Decisionmaking?”, *Duke Law Journal*, vol. 58, nº. 7, 2009, p. 1540.

¹⁶ POSNER, *How Judges Think*, cit. en KNIGHT, p. 1542.

¹⁷ DROBAK & NORTH: “Understanding Judicial Decision-Making: The Importance of Constraints on Non-Rational Deliberations”, p. 134.

durante la fase de redacción de la opinión. Esa investigación es principalmente una forma de buscar argumentos y hechos que la sostengan.¹⁸

En estas palabras reside la más importante y problemática de las derivadas del realismo, especialmente para aquellos que no se consideran realistas. La lectura más inocua del asunto implicaría que la mayoría de operadores jurídicos habrán asimilado la norma y sus principios hasta el punto de hacerse inconscientes. El disparo espontáneo del juicio intuitivo no sería *per se* un problema, en la medida en que procedería, de manera significativa al menos, del Derecho interiorizado¹⁹.

Pero esta no es la única interpretación posible del comentario de Posner, que parece querer apuntar a algo bien distinto. Las intuiciones se encuentran especialmente expuestas a la irracionalidad, dada la imposibilidad (por definición) de ofrecer los motivos para una corazonada. Siendo así, resulta difícil descartar de entrada que, en la formación de esa decisión intuitiva, no se hayan visto involucrados motivos “espurios”. A esto se añade que, si tomáramos al pie de la letra lo dicho por Posner, tendríamos que ver cierta “mala fe” en el juez que busca, en el fondo, armarse de razones con las que hacer prosperar una idea preconcebida sobre el resultado jurídico a alcanzar. Y si así fuera, poco queda en la fase de justificación que pudiera contrapesar este efecto, más allá de obligar a un mínimo encaje con el marco jurídico aplicable²⁰. En el mejor de los casos, el juez estaría tan inclinado como el resto de personas a incurrir en *sesgo confirmatorio*²¹, en el peor, encontraríamos un juez “partisano de sí mismo” cuya relación con la legalidad cabría calificar, como mucho, de instrumental.

El problema es, entonces, la relación entre las inclinaciones del decisor y el respeto de la legalidad pues “cuando decimos que las decisiones de un juez son conformes a ‘la ley’ (...) queremos decir que los determinantes de sus decisiones eran cosas que era legal que los jueces tuvieran en cuenta, consciente o inconscientemente”²². Llevado a sus últimas consecuencias, todo esto nos conduce a considerar que no tenemos garantías en cuanto a los criterios que alimentan la intuición, y que solo queda la esperanza de que la labor de justificación sea suficientemente “genuina” como para contrarrestar el efecto de motivaciones espurias (al menos las más claramente contrarias al marco normativo de referencia). Sin embargo, la posibilidad de que un juez haga uso de su posición para *racionalizar*²³, esto es, para ocultar sus auténticos motivos a través del empleo de la discrecionalidad y la elasticidad

¹⁸ POSNER, *How Judges Think*, cit. en KNIGHT: “Are Empiricists Asking the Right Questions about Judicial Decisionmaking?”, p. 1543.

¹⁹ Schultz insiste en que, más allá de posiciones extremadas, es posible entender que el proceso decisional del juez se vea influido por factores legales y extralegales simultáneamente, *vid.* SCHULTZ: “Arbitral Decision-Making: Legal Realism and Law & Economics”, p. 238–39.

²⁰ KNIGHT: “Are Empiricists Asking the Right Questions about Judicial Decisionmaking?”, p. 1548.

²¹ “ Este proceso mental se caracteriza por la tendencia del sujeto a filtrar una información que recibe, de manera que, de forma inconsciente, busca y sobrevalora las pruebas y argumentos que confirman su propia posición inicial, e ignora y no valora las pruebas y argumentos que no respaldan”, MUÑOZ ARANGUREN, A.: “La Influencia de Los Sesgos Cognitivos En Las Decisiones Jurisdiccionales: El Factor Humano. Una Aproximación”, *InDret*, nº. 2, 2011, p. 8.

²² POSNER, *How Judges Think*, cit. en KNIGHT, “Are Empiricists Asking the Right Questions about Judicial Decisionmaking?”, p. 1544.

²³ HABA, E.P.: “Razones Para No Creer En La Actual Teoría (Ilusionista) de La Argumentación. Papel Que Tales Teorizaciones Cumplen Como Más Nueva Ideología de Legitimación Con Respecto, Especialmente, Al Discurso Jurídico Profesional”, *Doxa Cuadernos de Filosofía Del Derecho*, vol. 33, 2010, p. 333; LEITER, “Realismo Jurídico Estadounidense”, p. 242.

interpretativa sigue siendo un riesgo sumamente tangible cuya influencia es preciso reconocer en una parte significativa de las controversias jurídicas.

2.2 Realismo versión 2.0: ciencias cognitivas, sesgos implícitos, intuición y emoción

Hecha una representación sumamente condensada del tipo de crítica realista que interesa a los efectos de este texto, procede actualizar un tanto el debate proyectando la mirada hacia contribuciones más modernas que han abundado en una dirección similar. En este sentido, si los realistas “clásicos” se apoyaban frecuentemente en la psicología para defender sus posturas, en los trabajos contemporáneos este recurso se afianza y expande con el acceso a las más sofisticadas neurociencias y, en general, el campo de las “*cognitive sciences*”. Como el propósito de esta discusión “realista” no es el retrato de la corriente académica en sí, sino la extracción de algunos de los argumentos más llamativos para canalizar el debate, se ha optado por atender a dos de los “grandes temas” que ocupan a la investigación sobre la toma de decisiones humana (y la judicial, en especial). Uno de esos temas es el de los “sesgos inconscientes” (*unconscious biases*), el otro, la neuropsicología de las decisiones²⁴.

A) Sesgos inconscientes. El juez intuitivo

La posición que hemos atribuido al realismo ha tendido a encasillar a sus autores como “escépticos” respecto del Derecho y las profesiones jurídicas. De acuerdo con este retrato, para los realistas los jueces no pueden ser justos, dado que no pueden ser racionales, pues no pueden evitar ser presa de estereotipos, generalizaciones, supersticiones, o los efectos que sobre ellos tengan su estado de cansancio o lo que hayan desayunado²⁵. Más moderadamente, señala Alan Scott Rau que “los decisores que hayan vivido mínimamente en el mundo llegarán invariablemente a los casos con perspectivas, creencias y preconcepciones”²⁶, y en ello reside el temor a la arbitrariedad. Esto explica la intensidad de la reacción contra los realistas: si uno no puede confiar en la motivación enarbolada por un juez, entonces puede que no quepa confiar en nada que pueda ser reconducido a la legalidad o la justicia misma. Curiosamente, que este fuera un diagnóstico algo exagerado no parece hacerlo menos potente.

La relativa minoría en la que las posiciones realistas se desenvuelven en el seno de la academia jurídica no obsta, sin embargo, para encontrar que goza de buena salud entre otras disciplinas de carácter científico. Sin embargo, hasta donde me alcanza, la investigación reciente sobre la toma de decisiones judicial se encuentra lejos de descartar completamente la influencia de las reglas y los precedentes jurídicos. Más matizadamente, el argumento

²⁴ La distinción es en realidad artificial, pues se pueden estudiar los sesgos inconscientes con los métodos y técnicas de las neurociencias. La separación se ha hecho más bien por especificidad metodológica, dado que la mayoría de investigaciones sobre sesgos inconscientes parece hacer uso de los métodos de la psicología tradicional, mientras que el campo neurocientífico representa un subsector especializado, con un abordaje metodológico bien distinto que condiciona tanto el tipo de datos recabados como la interpretación de los mismos. Simplificando, entiéndase la distinción como fundamentalmente expositiva.

²⁵ Esto es, en términos de Schultz, “un meme sobre el realismo jurídico” cuya fama cabe atribuir, no a su defensa por los propios realistas, sino a su empleo como “hombre de paja” por los críticos de aquellos. Vid. SCHULTZ: “Arbitral Decision-Making: Legal Realism and Law & Economics”, p. 233. También, DROBAK & NORTH: “Understanding Judicial Decision-Making: The Importance of Constraints on Non-Rational Deliberations”, p. 136.

²⁶ RAU, A.S.: “Integrity in Private Judging”, *South Texas Law Review*, vol. 38, no. 2, 1997, p. 488.

que parece que se extrae de la literatura es que sería un error tratar de explicar el comportamiento de los jueces *únicamente* sobre la base del marco jurídico preexistente²⁷. Las normas, los valores, los principios... son solo una parte del puzzle y su influencia puede depender de su interacción con otras variables extrajurídicas.

En coherencia con esta perspectiva, se han manejado diversas opciones explicativas. Algunos autores han intentado examinar la influencia de consideraciones relativas a la organización de la actividad judicial (por ejemplo, el efecto que sobre la decisión pueda tener la expectativa de ser contradicho en sede de recurso²⁸). Otros han examinado la influencia de la ideología política y de las dinámicas de grupo en los órganos colegiados²⁹. Finalmente, algunos estudios han tratado de estimar el efecto de variables sociodemográficas más generales³⁰. Resulta significativo que, aunque pervivan posiciones extremadas, no es usual que se descarte la influencia del marco legal. Al contrario, este es generalmente asumido como el marco en que las decisiones son tomadas y que constriñe en mayor o menor medida la libertad de acción del juez³¹.

Con todo, el argumento realista tal vez haya recibido su mayor impulso al calor del debate estadounidense sobre el racismo institucional. En este sentido, la discusión ha promovido múltiples estudios que han tratado de determinar el alcance e intensidad del trato discriminatorio dirigido a las minorías por parte de los órganos judiciales. A los efectos de este trabajo, interesa destacar que los principales avances en esta materia han venido de la mano del estudio de los llamados “sesgos implícitos”, propiciado por el desarrollo del *Implicit Association Test* (IAT)³². Globalmente considerados, estos trabajos apuntan, no al “viejo” racismo explícito y extravagante, sino al sutil espacio de las inclinaciones inconscientes que se deslizan en el pensamiento y el comportamiento de los decisores sin su conocimiento.

Esto supone un grado mayor de sofisticación para la crítica realista, que ya no se presenta como una defensa de la fundamental irracionalidad y mala fe del juzgador. Al contrario, los autores de estos trabajos aceptarían las buenas intenciones y el carácter genuino de la mayoría de los jueces y tribunales. Sin embargo, añadirían que ese no es el problema, pues la característica fundamental de los sesgos implícitos es que producen sus efectos en individuos que jamás apoyarían actitudes o discursos discriminatorios. Frente al paradigma del juez tramposo, aparece el mucho más trágico caso del buen juez que, sin quererlo ni saberlo, discrimina. Que algunos sesgos implícitos, como los motivados por la raza, sean intrínsecamente contrarios al marco jurídico de referencia solo hace la cuestión más problemática.

²⁷ SCHULTZ: “Arbitral Decision-Making: Legal Realism and Law & Economics”, p. 232.

²⁸ Esto en cierto modo sigue siendo reducible a “razonamiento jurídico”, aunque de naturaleza un tanto instrumental.

²⁹ MILES, T.J. & SUNSTEIN, C.R.: “The New Legal Realism”, *University of Chicago Law Review*, vol. 75, nº. 2, 2008, p. 838, <https://doi.org/10.1017/cbo9781107762336>.

³⁰ BREGER, M.L.: “Making the Invisible Visible: Exploring Implicit Bias, Judicial Diversity, and the Bench Trial”, *University of Richmond Law Review*, 53, nº. 4, 2019, p. 1063.

³¹ CROSS, F.B.: “Decisionmaking in the U.S. Circuit Courts of Appeals”, *California Law Review*, vol. 91, nº. 6, 2003, p. 1514, <https://doi.org/10.2307/3481397>; WISTRICH, RACHLINSKI, & GUTHRIE: “Heart versus Head: Do Judges Follow the Law or Follow Their Feelings?”, p. 911.

³² BREGER: “Making the Invisible Visible: Exploring Implicit Bias, Judicial Diversity, and the Bench Trial”, p. 1049.

En este sentido, los resultados de algunos de los trabajos más conocidos en este campo son inquietantes: las personas en general parecen inclinadas albergar sesgos que favorecen a los jóvenes, los ricos y los blancos³³. Los jueces, por su parte, no parecen comportarse de manera diferente al resto de la población en relación con esto³⁴. En particular, los trabajos de Rachlinski y colaboradores mostraron que los jueces se comportaban de manera discriminatoria cuando se encontraban con acusados cuya apariencia y comportamiento cuadraba con sus roles estereotípicos. Es decir, cuanto más se pareciera un acusado afroamericano a la imagen que de él se haría un juez blanco de clase media, más tendía este juez a afincarse en una actitud correspondiente con su grupo. Dicho en términos bruscos, cuanto más “negro” viera el juez al acusado, más “blanco” se volvía él mismo³⁵.

Globalmente consideradas, las investigaciones sugieren que los sesgos implícitos favorecen comportamientos consonantes con ellos, especialmente en contextos poco deliberativos³⁶. Con traintuitivamente, la situación no mejora, sino que empeora, conforme aumenta la experiencia del profesional jurídico. Así, los expertos tienden a confiar más en sus corazonadas, pero también a considerarse libres de sesgos, una combinación que resulta campo abonado para la proliferación de decisiones sesgadas³⁷. Ahora bien, la relación entre sesgos implícitos y comportamiento no es siempre directa o inmediata en la profesión judicial, donde cabe encontrar que incluso jueces con sesgos importantes consiguen adjudicar de manera equidistante³⁸.

Podría plantearse, sin embargo, que forzar a la deliberación mediante la elaboración de justificaciones exhaustivas es el mecanismo por excelencia para poner coto a esta

³³ IRWIN, J.F. & REAL, D.L.: “Unconscious Influences on Judicial Decision-Making: The Illusion of Objectivity”, *McGeorge Law Review*, 42, nº. 1, 2008, p. 4.

³⁴ Con todo, las reflexiones de Rachlinski y colaboradores son más inquietantes por momentos: a diferencia de lo que sucede con otros miembros de la población general con un compromiso igualitario fuerte en términos ideológicos, el compromiso “profesional” de los jueces parece incapaz, por sí solo, de compensar sus sesgos. De hecho, entre jueces caucásicos, los resultados son incluso ligeramente peores que los encontrados en la población general de estadounidenses blancos. *Vid.* RACHLINSKI, J.J. ET AL.: “Does Unconscious Racial Bias Affect Trial Judges?”, *Notre Dame Law Review*, vol. 84, nº. 3, 2009, p. 1221.

³⁵ BREGER: “Making the Invisible Visible: Exploring Implicit Bias, Judicial Diversity, and the Bench Trial”, p. 1056; RACHLINSKI ET AL., “Does Unconscious Racial Bias Affect Trial Judges?”, p. 1209–10. Esta es una observación clásica de la psicología social sobre el prejuicio, según observan AMODIO, D.M. & CIKARA, M.: “The Social Neuroscience of Prejudice”, *Annual Review of Psychology*, vol. 72, 2021, p. 446, <https://doi.org/10.1146/annurev-psych-010419-050928>.

³⁶ IRWIN & REAL: “Unconscious Influences on Judicial Decision-Making: The Illusion of Objectivity”, p. 4; WISTRICH, RACHLINSKI, & GUTHRIE, “Heart versus Head: Do Judges Follow the Law or Follow Their Feelings?”, p. 863.

³⁷ Bennet lo expresa con cierto humor: “En mi reciente estudio empírico nacional, encontré que un 92% de los jueces federales veteranos, el 87% de los no veteranos, el 72% de los magistrados de EE.UU., el 77% de los jueces federales concursales y un 96% de los oficiales de Libertad condicional se situaban a sí mismos en el 25% superior respecto de sus colegas en cuanto a su habilidad para tomar decisiones privadas de sesgo racial. De nuevo, matemáticamente imposible”. Traducción propia, citado en BREGER: “Making the Invisible Visible: Exploring Implicit Bias, Judicial Diversity, and the Bench Trial”, p. 1051, ^{nota 60}.

³⁸ Es posible que los jueces estén especialmente atentos a sesgos problemáticos como el de la raza, lo que les conduciría a comportarse de manera especialmente precavida, algo que se ha mostrado como un buen instrumento para mantener el sesgo bajo control. Cuestión distinta es lo que pueda pasar con otras fuentes de sesgo frente a las que se pueda estar menos concienciado. *Vid.* IRWIN & REAL: “Unconscious Influences on Judicial Decision-Making: The Illusion of Objectivity”, p. 8; WISTRICH & RACHLINSKI, “Implicit Bias in Judicial Decision Making. How It Affects Judgment and What Judges Can Do About It”, p. 101.

impulsividad cognitiva³⁹. Esto no se discute, aunque es cierto que los jueces se han mostrado tan proclives a apoyarse en sus intuiciones como cualquier otra persona (lo que quiere decir, bastante proclives)⁴⁰. Por otro lado, más positivamente, existen indicios de que el influjo de los sesgos implícitos puede ser contrapesado por una actitud vigilante del decisor y mediante el fomento del razonamiento deliberativo, lo cual refuerza la centralidad de la motivación escrita en las decisiones judiciales⁴¹.

La posibilidad de limitar el influjo de sesgos indeseados mediante el adiestramiento de los decisores y el favorecimiento del razonamiento de ritmos “lentos”⁴² ciertamente deja espacio para la posibilidad de un juez bienintencionado y profesional. Esto matiza considerablemente las interpretaciones más extremadas de la crítica realista. Sin embargo, la tranquilidad reposa sobre la premisa de la buena fe⁴³, algo que solo hace más intensa la duda de qué sucede cuando el juzgador alberga sesgos conscientes y voluntarios. En tales casos, cabe esperar que la discrecionalidad siga siendo un arma considerable, y la argumentación jurídica un instrumento sofisticado para doblegar la ley a imagen del decisor⁴⁴.

B) Neurociencia de la decisión judicial. El juez límbico

Como anticipábamos, la expansión extraordinaria acusada por el campo de las neurociencias también ha generado investigaciones con ambiciones similares a las que aquí nos ocupan. Ahora bien, aunque la premisa de partida y el objetivo sean parangonables, flota

³⁹ GUTHRIE, C., RACHLINSKI, J.J., & WISTRICH, A.J.: “Blinking on the Bench: How Judges Decide Cases”, *Cornell Law Review*, vol. 93, nº. 1, 2007, p. 33.

⁴⁰ Algo que se ha medido con el *Cognitive Reflection Test*, mostrando un rendimiento semejante en jueces y otros grupos sociales con similar trayectoria educativa. Vid. IRWIN & REAL: “Unconscious Influences on Judicial Decision-Making: The Illusion of Objectivity”, p. 6.

⁴¹ IRWIN & REAL, 9; WISTRICH & RACHLINSKI: “Implicit Bias in Judicial Decision Making. How It Affects Judgment and What Judges Can Do About It”, p. 117.

⁴² La referencia a la conocida obra de Kahneman resulta prácticamente ociosa, aunque inesquivable a la vista de la estrecha relación temática con el objeto de este trabajo. “Pensar despacio” parece esencial para mantener bajo control a un sistema intuitivo ágil pero proclive a la precipitación. Vid. KAHNEMAN, D.: *Pensar Rápido, Pensar Despacio, Debate, 2012*.

⁴³ Esta buena fe gana importancia en todo contexto en que la labor del juez trasciende la promovida por el formalismo/positivismo. Por este motivo, se la ha vinculado a veces con la tesis de la “única respuesta correcta”, entendiendo esta última (algo que no es pacífico) ante todo como una posición deontológica del jurista. Se trataría, en tal caso, no de un postulado descriptivo, sino de una premisa según la cual, exista o no respuesta correcta, el jurista debe actuar como si tal respuesta existiera, afanándose por hallarla y materializarla. Esto se opone al instrumentalismo que se atribuye a otras profesiones jurídicas, donde “correcto” equivale a “ganador”. No obstante, no termina de tranquilizar la conciencia de aquel que recela de la facilidad con la que diversos horrores se han alzado sobre el convencimiento de la corrección, la verdad, la justicia y otros absolutos. Sobre las distintas visiones de la tesis de la respuesta correcta, vid. ATIENZA, *Curso de Argumentación Jurídica*, p. 552–53. Sobre la buena fe en estos asuntos, v.gr. GILLMAN, H.: “What’s Law Got to Do with It? Judicial Behaviorists Test the ‘Legal Model’ of Judicial Decision Making”, *Law & Social Inquiry*, vol. 26, nº. 02, 2001, p. 488, ^{nota 23}, <https://doi.org/10.1111/j.1747-4469.2001.tb00185.x>.

⁴⁴ Un temor que parece en la base de muchas de las dudas, más o menos vehementes, expresadas frente a la teoría de la argumentación, una vez que esta trasciende un marco restringido y fuertemente apegado a la norma positiva. Vid. HABA: “Razones Para No Creer En La Actual Teoría (Ilusionista) de La Argumentación. Papel Que Tales Teorizaciones Cumplen Como Más Nueva Ideología de Legitimación Con Respecto, Especialmente, Al Discurso Jurídico Profesional”, p. 355–57.

sobre las investigaciones “neuro” una pretensión implícita de profundidad, al ahondar en los confines de la mente⁴⁵. Por asociación de ideas, sus resultados e interpretaciones resucitan un “realismo fuerte” que contrasta con el tono conciliador de las investigaciones del apartado anterior.

Al igual que en los trabajos vistos hasta ahora, las reacciones espontáneas, intuitivas y preconscientes han sido objeto de atención especial. Un ejemplo destacado es el estudio de la neurofisiología del miedo y la evaluación del grado de amenaza inspirada por el acusado. Este es un punto importante de vinculación entre las inclinaciones cognitivas y el funcionamiento del sistema penal, en la medida en que existen decisiones jurídicas específicamente orientadas a valorar el riesgo de un sujeto y adjudicar una respuesta específica en función de esto. Cómo de temible le resulte un individuo al juzgador puede ser una pieza determinante a la hora de decidir su puesta en libertad a través de cualquier mecanismo discrecional (permisos de salida, progresiones de grado, libertad condicional...)⁴⁶. Y esta “temibilidad” depende, no solo del grado de sensibilidad de las correspondientes áreas cerebrales, sino del modo en que esta sensibilidad se manifiesta de manera diferente en unos casos y en otros. Sobre esto, debe añadirse el alineamiento general entre percepción del miedo y significantes culturales del temor (jurídicamente cuestionables) como la ajenidad al grupo, la raza, la extracción social, el estilo de vida o similares⁴⁷. Si algo parecen sugerir la psicología y la neurobiología es que las personas son bastante más hábiles a la hora de interiorizar fuentes de amenaza y reaccionar frente a ellas, que a la hora de realizar evaluaciones certeras sobre el riesgo realmente existente.

Los estudios sobre raza, de nuevo, proporcionan ejemplos interesantes. Algunas investigaciones por medio de resonancia magnética funcional (fMRI) han mostrado, en sujetos experimentales caucásicos, una mayor reacción de temor frente a rasgos faciales afrocentrísticos⁴⁸. Esto no venía alterado en grado alguno por la actitud consciente de los participantes, que afirmaban no albergar sesgos raciales⁴⁹. Si bien, según trabajos recientes, la interpretación de la evidencia disponible es compleja y más indirecta⁵⁰. Algo similar sucede con las reacciones aversivas, que pueden ser también relevantes en la adjudicación penal cuando se trata de pensar en términos de retribución o inocuización. Asimismo, algún trabajo sugiere una degradación de las funciones ejecutivas⁵¹ en sujetos con una elevada activación neurofisiológica frente a afroamericanos. Aplicado al asunto que nos ocupa, esto sugiere que

⁴⁵ SPAIN BRADLEY: “The Disruptive Neuroscience of Judicial Choice”, p. 17.

⁴⁶ PAPIILLON, K.: “The Court’s Brain: Neuroscience and Judicial Decision Making in Criminal Sentencing”, *Court Review*, vol. 48, nº. 1, 2013, p. 53.

⁴⁷ Citando diversa literatura al respecto, *vid.* SPAIN BRADLEY: “The Disruptive Neuroscience of Judicial Choice”, p. 27.

⁴⁸ RONQUILLO, J. ET AL.: “The Effects of Skin Tone on Race-Related Amygdala Activity: An FMRI Investigation”, *Social Cognitive and Affective Neuroscience*, vol. 2, nº. 1, 2007, p. 39–44, <https://doi.org/10.1093/scan/nsl043>. Con comentario al respecto de PAPIILLON: “The Court’s Brain: Neuroscience and Judicial Decision Making in Criminal Sentencing”, p. 51. Nótese, a efectos interpretativos, que el estudio de Ronquillo y colaboradores cuenta con una muestra pequeña de participantes, algo que no resulta del todo inusual en neurociencias.

⁴⁹ WHEELER, M.E. & FISKE, S.T.: “Controlling Racial Prejudice: Social-Cognitive Goals Affect Amygdala and Stereotype Activation”, *Psychological Science*, vol. 16, nº. 1, 2005, p. 56–63, <https://doi.org/10.1111/j.0956-7976.2005.00780.x>.

⁵⁰ AMODIO & CIKARA: “The Social Neuroscience of Prejudice”, p. 448.

⁵¹ RICHESON, J.A. & TRAWALTER, S.: “Why Do Interracial Interactions Impair Executive Function? A Resource Depletion Account”, *Journal of Personality and Social Psychology*, vol. 88, nº. 6, 2005, p. 934–47, <https://doi.org/10.1037/0022-3514.88.6.934>.

algunos decisores podrían tener dificultades para “corregir” sus reacciones espontáneas y mantener la equidistancia.

La empatía también ha sido objeto de estudio en relación con las decisiones judiciales⁵². En este sentido, los resultados son previsibles, y la parte procesal más capaz de suscitar empatía en el juzgador tiene mayores posibilidades de venir favorecida. En términos de castigo, no solo la empatía hacia el acusado puede derivar en un trato más lenitivo, sino que la empatía hacia la víctima puede propiciar una respuesta más punitiva⁵³. A la vista de esto, la sala de justicia se muestra como un lugar donde la retórica y la representación son factores importantes⁵⁴, algo que seguramente no muchos juristas discutirían, pero que abunda en el sentido del escepticismo realista. En este caso, nuevamente, encajar con el “estereotipo” puede ser importante en más de un sentido. Por ejemplo, una víctima que no se muestra tan devastada y vulnerable como cabría esperar puede estar en desventaja a la hora de ser atendida o, incluso, creída. En esta ocasión también los estudios sobre raza complican la historia: algunos autores han observado una pauta que tiende a favorecer la empatía hacia los caucásicos⁵⁵. En coherencia con esto, se ha resaltado la mayor aplicación de la pena de muerte en casos de homicidio en los que el acusado era afroamericano y la víctima caucásica⁵⁶.

Por último, estudios de neuroimagen han mostrado que la capacidad para humanizar a los demás, el percibirlos como personas y reaccionar de conformidad con ello, no es una capacidad autoevidente. Más bien al contrario, se trata de un mecanismo neurofisiológico que se activa de manera selectiva y que, previsiblemente, se muestra hipofuncional ante individuos marginales (como drogadictos o mendigos)⁵⁷. Cabe inferir entonces que la deshumanización de otros ocurre de manera espontánea y adquiere mayores cotas conforme nos desplazamos del individuo prototípico de clase media hacia los diversos subgrupos

⁵² Fuera de las neurociencias, es bien conocido el estudio de: GLYNN, A.N. & SEN, M.: “Identifying Judicial Empathy: Does Having Daughters Cause Judges to Rule for Women’s Issues?”, *American Journal of Political Science*, vol. 59, nº. 1, 2015, p. 37–54, <https://doi.org/10.1111/ajps.12118>. Asimismo, sobre la literatura en materia de empatía y decisiones morales (también judiciales), cabe destacar el reciente trabajo analítico de DECETY, J.: “Why Empathy Is Not a Reliable Source of Information in Moral Decision Making”, *Current Directions in Psychological Science*, vol. 30, nº. 5, 2021, p. 427, <https://doi.org/10.1177/09637214211031943>. Un examen reciente del debate sobre la “empatía judicial” puede encontrarse en CORMACK, W.: “Reassessing the Judicial Empathy Debate: How Empathy Can Distort and Improve Criminal Sentencing”, *Mitchell Hamline Law Review*, vol. 47, 2021.

⁵³ Algo que se ha denominado en ocasiones la “*compassion-hostility paradox*”. Sobre esto, y recopilando bibliografía al respecto, *vid.* JOHNSON, S.L. ET AL.: “When Empathy Bites Back: Cautionary Tales from Neuroscience for Capital Sentencing”, *Fordham Law Review*, vol. 85, nº. 2, 2016, p. 589.

⁵⁴ Algo que, especialmente en delitos graves, puede transformarse en una pugna por la humanización/deshumanización del acusado, especialmente en sistemas procesales más dados a la retórica letrada y a los juicios por jurado. *Vid.* JOHNSON ET AL., p. 574.

⁵⁵ Posiblemente relacionada con la interiorización de respuestas automatizadas basadas en estereotipos culturalmente asentados, así como la (también generalmente comentada) inclinación a empatizar con personas que se perciben como pertenecientes al mismo grupo que el observador (siendo la raza un marcador relevante de “pertenencia”). *Vid.* JOHNSON ET AL., p. 584; PAPILLON, “The Court’s Brain: Neuroscience and Judicial Decision Making in Criminal Sentencing”, p. 55.

⁵⁶ EBERHARDT, J.L. ET AL.: “Looking Deathworthy: Perceived Stereotypicality of Black Defendants Predicts Capital-Sentencing Outcomes”, *Psychological Science*, vol. 17, nº. 5, 2006, p. 383–86, <https://doi.org/10.1111/j.1467-9280.2006.01716.x>.

⁵⁷ HARRIS, L.T. & FISKE, S.T.: “Dehumanizing the Lowest of the Low: Neuroimaging Responses to Extreme Out-Groups”, *Psychological Science*, vol. 17, nº. 10, 2006, p. 847–53, <https://doi.org/10.1111/j.1467-9280.2006.01793.x>.

socialmente desaventajados. Una implicación preocupante de ello parece ser que las personas (y entre ellas, los jueces) parecen tener especial facilidad a la hora de perpetuar las disparidades de clase, raza género y demás, aunque ello entre en contradicción directa con su sistema de valores conscientemente asumido.

2.3 Síntesis. El realismo y los jueces

Llegados a este punto, procede recapitular brevemente antes de pasar al apartado siguiente. Como anticipábamos, en este artículo el realismo jurídico y sus desarrollos se emplean, no como objeto de estudio en sí mismo, sino como plataforma para extraer una serie de observaciones convergentes. En este sentido, ha sido una posición clásica del realismo el afrontar con escepticismo lo que percibían como una idealización del papel del juez. En este sentido, al juez representado como un jurista ilustrado, guiado en exclusiva por el Derecho y sus valores subyacentes, los realistas trataron de oponer la figura de un juez humano, imperfecto, y cuyo proceso decisional no cabía reducir a las variables explícitamente manifestadas en las sentencias. Si la doctrina clásica daba por supuesta la correspondencia entre la decisión y la justificación, dando carta de naturaleza a la primera por lo convincente que fuera la segunda, los realistas quiebran con este consensual estado de cosas. Para el realismo, de entonces y de ahora, la relación entre la decisión y la justificación es problemática, mediata, y existe cierta sospecha de que los extensos argumentos sirvan principalmente de encubrimiento, a la manera de un “blanqueo de capital intelectual”, si se quiere.

Investigaciones que han seguido, directa o indirectamente, la pista a las intuiciones realistas han contribuido a matizar (pero también a consolidar) algunos de sus principales argumentos. En su versión moderada, el proceso decisional de los jueces y tribunales resulta de la composición de múltiples variables interrelacionadas, algunas jurídicas, otras no. De ello se infiere que el marco jurídico actúa como referencia e instrumento simultáneamente, pero no como universo autocontenido del que se deducen naturalmente todas y cada una de las resoluciones sin intervención de variables sociodemográficas, actitudinales y parajurídicas en general. Frente a este mensaje más o menos conciliador se erigen, sin embargo, investigaciones potencialmente más afines a posiciones de “realismo fuerte”, favorecidas por las metodologías neurocientíficas, y que llevan el debate al terreno de las disposiciones emocionales y actitudinales. Sin perjuicio de lo que un análisis sobre el fondo tuviera que decir al respecto, en términos discursivos estos desarrollos refuerzan la idea de las limitaciones de la mente humana, que se imponen al tomador de decisiones con independencia de su voluntad.

Leídos en conjunto, estos y otros trabajos convergen en una misma observación, que es la que querríamos rescatar. Se atisba una línea crítica fundamental que plantea que los jueces y magistrados son, en definitiva, demasiado humanos como para estar a la altura de las expectativas y responsabilidades depositadas sobre sus hombros por parte del ordenamiento jurídico. Esta es una impugnación fuerte que, sin embargo, ha podido ser desestimada hasta la fecha (con más o menos polémica) por la imposibilidad de hacer las cosas de otro modo. Fuera el juez bueno o malo, virtuoso o malicioso, no había más remedio que cargar con ese riesgo en un universo en que solo los seres humanos pueden impartir(se) justicia. Este consenso erigido sobre la resignación iba a venir trastocado con la llegada de la Inteligencia Artificial.

3. JUECES ARTIFICIALES: ¿APOTEOSIS DEL REALISMO O ALIENACIÓN DE LA JUSTICIA?

En los últimos años la conversación en torno a las aplicaciones de la IA ha alcanzado cotas de atención extraordinarias, tanto a nivel político como académico. En este sentido, la relación IA-Derecho ha suscitado múltiples reflexiones de muy distinta naturaleza. A los efectos de este artículo, previsiblemente, resulta de especial interés el debate en torno a la introducción de la IA en el campo de las decisiones expertas y, en particular, en el terreno de las decisiones judiciales. En este ámbito, la producción científica ha tendido a asumir objetivos menos ambiciosos, pensando en las maneras en que la introducción de la IA podría complementar, potenciar y en general acompañar la actividad de la justicia, con el fin de resolver algunos de sus problemas estructurales⁵⁸. Ahora bien, la naturaleza frecuentemente futurista y anticipatoria de las discusiones ha hecho que otra hipótesis haya cobrado atención significativa (tal vez especialmente entre los juristas), a saber, la de la total sustitución de los jueces convencionales por IAs judiciales. El carácter fuertemente especulativo de este debate no lo hace menos pertinente o interesante, especialmente porque permite canalizar eficazmente una discusión sobre aquello que consideramos irrenunciable en la labor de hacer justicia. En el núcleo de esta cuestión reside, diría, una controversia en torno a si la justicia requiere o no de humanidad.

3.1 El Juez IA y la culminación del realismo

Como transición del apartado primero a este, es pertinente dar cuenta de lo que el debate contemporáneo en torno a la inteligencia artificial judicial tiene de “continuidad” implícita del realismo jurídico. La premisa fundamental se intuye con facilidad: si para los realistas el problema era acompasar las imperfecciones del juez humano con las demandas de abstracción de la norma, los algoritmos contemporáneos parecen presentar grandes ventajas a la hora de remediar estos defectos.

El punto de partida, con todo, es menos radical de lo que cabría esperar. Para que pudiera efectuarse una adecuada sustitución del decisor humano por el algorítmico, no es necesario que este último se haga “más humano”. Al contrario, se trata de llegar a un punto en el que el algoritmo pueda desarrollar las tareas que habitualmente ocupan a un juez con un rendimiento igual o mejor. Así lo plantea Christof Winter, por ejemplo: “para que la IA juegue un papel importante en el sistema judicial del futuro, no tiene por qué ser perfecta, basta con que sea mejor que la IH [inteligencia humana]”⁵⁹. Esta posibilidad no es extravagante ni inverosímil, de hecho, resulta hoy por hoy innegable en relación con algunas tareas en las

⁵⁸ De hecho, los expertos y la prudencia coinciden en señalar que esta es la dirección que van a adoptar (ya han adoptado) las incursiones de la inteligencia artificial en materia de justicia y que aquí se encuentran, sin duda, los problemas más inmediatos y urgentes. V.gr. JIN, Y. & HE, H.: “An Artificial-Intelligence-Based Semantic Assist Framework for Judicial Trials”, *Asian Journal of Law and Society*, vol. 7, nº. 3, 2020, p. 532–35, <https://doi.org/10.1017/als.2020.33>; SOURDIN: *Judges, Technology and Artificial Intelligence. The Artificial Judge*, p. 129; SOURDIN: “Judge v Robot? Artificial Intelligence and Judicial Decision-Making”, p. 1123; XU, Z.: “Human Judges in the Era of Artificial Intelligence: Challenges and Opportunities”, *Applied Artificial Intelligence*, vol. 36, nº. 1, 2022, p. 1042, <https://doi.org/10.1080/08839514.2021.2013652>; XU, Z., ZHAO, Y., & DENG, Z.: “The Possibilities and Limits of AI in Chinese Judicial Judgment”, *AI and Society*, 2021, p. 3, <https://doi.org/10.1007/s00146-021-01250-9>.

⁵⁹ WINTER, C.K.: “The Challenges of Artificial Judicial Decision-Making for Liberal Democracy”, en *Judicial Decision-Making: Integrating Empirical and Theoretical Perspectives*, BYSTRANOWSKI, P., JANIK, B., & PRÓCHNICKI, M. (eds.), Springer, 2022, p. 183.

que la capacidad de procesamiento de los sistemas informáticos modernos resulta mucho más adaptada que las capacidades cognitivas humanas⁶⁰.

A este respecto resulta bien interesante un estudio realizado por Kleinberg y colaboradores sobre liberación bajo fianza. Tratando de valorar el rendimiento de un sistema automatizado en comparación con un juez, desarrollaron un algoritmo de aprendizaje automático que mejoraba significativamente el rendimiento de un juez convencional. Y esto, además, con una flexibilidad de la que los decisores humanos no suelen hacer gala: el algoritmo podía optimizarse para minimizar las nuevas detenciones sin por ello alterar las tasas de encarcelamiento o, alternativamente, reducir la tasa de encarcelamiento sin un incremento de las tasas de reincidencia delictiva⁶¹. Todo ello, además, reduciendo al mismo tiempo las disparidades raciales en las decisiones tomadas. En relación con esto, Cass Sunstein defiende que el algoritmo “funciona mucho mejor que los jueces reales (...) en cada una de las dimensiones que importan”⁶². Esta es una postura fuertemente realista que da buena cuenta de las coordenadas del debate.

El mejor rendimiento de los algoritmos respecto de los juzgadores humanos se ha argumentado sobre la base de diversas dimensiones que coinciden con valores jurídicos destacados. Desde la óptica de la igualdad de trato, se ha señalado que hacer uso de sistemas de IA podría dar por zanjada la preocupación por los sesgos cognitivos⁶³, además de permitir una homogeneización de criterios en la aplicación de la ley⁶⁴. Aunque el debate en torno a los sesgos algorítmicos no es un debate menor, se ha planteado que el algoritmo presenta ventajas en términos de transparencia que podrían permitir paliar sus efectos. En efecto, el diseño del algoritmo puede introducir sesgos, como también puede hacerlo el entrenamiento realizado sobre datos sesgados⁶⁵. Sin embargo, gran parte del empuje por desarrollar

⁶⁰ No cabe olvidar, en este sentido, que el fundamento detrás del desarrollo de la más básica de las herramientas informáticas estriba precisamente en la especialización. El propósito nunca ha sido (hasta ahora, al menos) replicar la inteligencia humana, sino complementarla retirándole el peso de tareas que son extremadamente ineficientes para los sistemas cognitivos humanos.

⁶¹ En el primer caso, el modelo pudo producir reducciones de re-arrestos del 14,4 al 24,7%, mientras que en el segundo, pudo bajar las tasas de encarcelamiento del 18,5 hasta el 41,9%, *vid.* KLEINBERG, J. ET AL.: “Human Decisions and Machine Predictions”, *The Quarterly Journal of Economics*, vol. 133, nº. 1, 2018, p. 241, <https://doi.org/10.1093/qje/qjx032>.

⁶² SUNSTEIN, C.R.: “Algorithms, Correcting Biases”, *Social Research*, vol. 86, nº. 2, 2019, p. 500. Desde un punto de vista criminológico, aquí vemos resurgir un argumento central de la criminología positivista clásica. Según Sunstein, los jueces ponen demasiado énfasis en el delito (en sus términos, el “sesgo del delito actual”) en detrimento de otras variables más relevantes para la evaluación del riesgo. Esto, en la jerga decimonónica, significa que el delito y la peligrosidad criminal son diferentes y no tienen por qué estar relacionados. De este modo, los jueces suelen tratar a los delincuentes de alto riesgo como de bajo riesgo, y a los de bajo riesgo como de alto riesgo. A grandes rasgos, los jueces pueden tener razones jurídicas para limitarse al delito o delitos (¿hasta qué punto es legítimo decidir sobre la base de la edad, la raza o el género para hacer predicciones en un contexto penal?), pero cuando se les pone en la tesitura de predecir la delincuencia futura para tomar una decisión, parecen ser tan capaces de atisbar el futuro como cualquier otro ser humano, es decir, no demasiado.

⁶³ SOURDIN: “Judge v Robot? Artificial Intelligence and Judicial Decision-Making”, p. 1128–29; SUNSTEIN, C.R.: “Governing By Algorithm? No Noise and (Potentially) Less Bias”, *Duke Law Journal*, vol. 71, nº. 6, 2022, p. 1188, <https://doi.org/10.2139/ssrn.3925240>; SUNSTEIN: “Algorithms, Correcting Biases”, p. 508; WINTER, “The Challenges of Artificial Judicial Decision-Making for Liberal Democracy”, p. 187.

⁶⁴ SOURDIN: *Judges, Technology and Artificial Intelligence. The Artificial Judge*, p. 133; Xu: “Human Judges in the Era of Artificial Intelligence: Challenges and Opportunities”, p. 1031; XU, ZHAO, & DENG: “The Possibilities and Limits of AI in Chinese Judicial Judgment”, p. 4.

algoritmos “transparentes” o “explicables” estriba, precisamente, en preservar la posibilidad de diseccionar el algoritmo en caso de que su funcionamiento resultase insatisfactorio. De esta manera, se pretende evitar la objeción de que determinadas arquitecturas de aprendizaje automático funcionan como “cajas negras”⁶⁶ algo que, por otro lado, es perfectamente trasladable al juez y sus corazonadas⁶⁷. Recordemos, en este sentido, que las justificaciones no son necesariamente una buena medida de las razones que efectivamente han motivado una decisión. Por su naturaleza, el contenido de una sentencia se construye por medio de piezas predeterminadas sobre la base de ser “contenido jurídicamente aceptable”, con el efecto implícito de silenciar todo aquello que no encaja en tal definición.

La equidad y la transparencia son, naturalmente, dos valores centrales que, por ello, han acumulado considerable atención, pero es cierto que no son las únicas ventajas jurídicamente relevantes. Por ejemplo, no es casual que anteriormente se empleara el ejemplo de un juicio pronóstico, que es una de esas tareas en las que los humanos suelen encontrarse consistentemente en inferioridad respecto de otros instrumentos, incluso aunque no estén basados en ningún tipo de inteligencia artificial⁶⁸. Que los pronósticos supongan una parcela, tal vez menor, de las tareas a realizar por un juez, no le resta relevancia a la vista de las implicaciones que pueden desprenderse, notablemente en la justicia penal. Asimismo, la cuestión de la agilidad del proceso parece tan trivial en la teoría como potencialmente transformadora en la práctica. La universalmente reconocida sobrecarga de la administración de justicia, burocratizada y vetusta hasta la saciedad, podría experimentar una auténtica revolución si pudiera ver sus tareas aligeradas por la informatización y la automatización⁶⁹. En relación con esto, no cabe olvidar que la certeza del castigo y su prontitud son mecanismos

⁶⁵ CITRON, D.K.: “Technological Due Process,” *Washington University Law Review*, 85, nº. 6, 2007, p. 1262; SUNSTEIN: “Governing By Algorithm? No Noise and (Potentially) Less Bias”, p. 1189; WISSER, L.: “Pandora’s Algorithmic Black Box: The Challenges of Using Algorithmic Risk Assessments in Sentencing”, *American Criminal Law Review*, vol. 56, 2019, p. 1819; VARONA, D., LIZAMA-MUE, Y., & SUÁREZ, J.L.: “Machine Learning’s Limitations in Avoiding Automation of Bias”, *AI and Society*, vol. 36, nº. 1, 2021, p. 202, <https://doi.org/10.1007/s00146-020-00996-y>.

⁶⁶ El efecto “black box” se ha convertido en un lugar común del debate sobre algoritmos, de modo que las referencias incluidas son simplemente ejemplificativas. Cabe recordar que el problema de la opacidad no es solo una cuestión de diseño, mientras que el efecto caja negra suele señalar a determinadas arquitecturas de algoritmos que son *inherentemente* opacas (como las SVMs o las redes neuronales artificiales). V.gr. BURRELL, J.: “How the Machine ‘Thinks’: Understanding Opacity in Machine Learning Algorithms”, *Big Data & Society*, vol. 3, no. 1, 2016, p. 3–5, <https://doi.org/10.1177/2053951715622512>; EDWARDS, L. & VEALE, M.: “Enslaving the Algorithm: From a ‘Right to an Explanation’ to a ‘Right to Better Decisions’?”, *IEEE Security and Privacy*, vol. 16, nº. 3, 2018, p. 2–3, <https://doi.org/10.1109/MSP.2018.2701152>; RUDIN, C.: “Please Stop Explaining Black Box Models for High Stakes Decisions”, en *Proceedings of NeurIPS 2018 Workshop on Critiquing and Correcting Trends in Machine Learning*, 2018, p. 1; WASHINGTON, A.L.: “How to Argue with an Algorithm: Lessons from the COMPAS- ProPublica Debate”, *The Colorado Technology Law Journal*, vol. 17, nº. 1 2019, p. 25–26.

⁶⁷ BONEZZI, A., OSTINELLI, M., & MELZNER, J.: “The Human Black-Box: The Illusion of Understanding Human Better Than Algorithmic Decision-Making”, *Journal of Experimental Psychology: General*, vol. February, 2022, p. 1, <https://doi.org/10.1037/xge0001181>; OSWALD, M. ET AL.: “Algorithmic Risk Assessment Policing Models: Lessons from the Durham HART Model and ‘Experimental’ Proportionality”, *Information and Communications Technology Law*, vol. 27, nº. 2, 2018, p. 232, <https://doi.org/10.1080/13600834.2018.1458455>; VOLOKH: “Chief Justice Robots”, p. 1187, nota 154; WINTER: “The Challenges of Artificial Judicial Decision-Making for Liberal Democracy”, p. 192.

⁶⁸ BERK, R.: *Machine Learning Risk Assessments in Criminal Justice Settings*, Springer, 2018, p. 7, <https://doi.org/10.1007/978-3-030-02272-3>; SUNSTEIN: “Governing By Algorithm? No Noise and (Potentially) Less Bias”, p. 1177.

⁶⁹ SOURDIN: *Judges, Technology and Artificial Intelligence. The Artificial Judge*, p. 120; Xu: “Human Judges in the Era of Artificial Intelligence: Challenges and Opportunities”, p. 1032; XU, ZHAO, & DENG: “The Possibilities and Limits of AI in Chinese Judicial Judgment”, p. 4–5. Resulta interesante la lectura de GOMES DE SOUSA, W. ET AL.: “Artificial

comúnmente señalados como más temibles (es decir, eficaces) que un sistema de penas inflacionario⁷⁰. Además, desde la perspectiva de la tutela judicial, recuérdese el adagio que reza “nada se parece tanto a la injusticia como la justicia tardía”⁷¹. En definitiva, hay algo de profundamente “sustantivo” detrás de las cuestiones “operativas”, quizás injustamente desatendidas por no estar tan directamente vinculadas a nociones trascendentales como la verdad, la justicia, la igualdad o la razón. Pero lo trascendente es poco más que un juicio estético si carece de unas mínimas condiciones de posibilidad.

En última instancia, podemos inferir que, para parte de la literatura, la hipótesis del Juez IA cuenta con ventajas importantes, leídas en la mayoría de casos al trasluz de las desventajas atribuidas al juicio humano. Desde este punto de vista, el algoritmo parece capaz de facilitar el cumplimiento de prioridades importantes del sistema de justicia, no solo en términos utilitarios o económicos, pero también deontológicos. Todo ello además, a juzgar por los más explícitos defensores de esta propuesta, al módico coste de sacrificar el romanticismo y la mística con que hemos envuelto la labor judicial.

Interesantemente, esta propuesta lleva implícitos múltiples cambios en la representación de la función judicial. Algunos se infieren de lo visto hasta el momento: un énfasis en el resultado de las decisiones antes que en el mecanismo que las produce (primacía de las tasas de acierto como criterio de calidad), en la seguridad jurídica derivada de la uniformidad de criterios (en lugar de la individualización y la discrecionalidad interpretativa), y tal vez de la agilidad por encima de la exhaustividad. Ahora bien, otras implicaciones parecen menos inmediatas. Por ejemplo, podría discutirse si se está apuntando a una “desdramatización”⁷² del ámbito judicial, en la medida en que la aceleración reduce la necesidad de procedimientos dilatados, lo que resta protagonismo al ritual y sus intervinientes en favor de cierto gerencialismo orientado a una resolución rápida. Esto sería igualmente coherente con desvestir a los operadores jurídicos de la mística de autoridad que los rodea, aproximando

[Intelligence and Speedy Trial in the Judiciary: Myth, Reality or Need? A Case Study in the Brazilian Supreme Court \(STF\)](https://doi.org/10.1016/j.giq.2021.101660), *Government Information Quarterly*, vol. 39, 2022, p. 1–13, <https://doi.org/10.1016/j.giq.2021.101660>.

⁷⁰ Para una revisión reciente, CHALFIN, A. & MCCRARY, J.: “Criminal Deterrence: A Review of the Literature”, *Journal of Economic Literature*, vol. 55, no. 1, 2017, p. 27, <https://doi.org/10.1257/jel.20141147>.

⁷¹ Frase que se atribuye a Séneca y que en inglés tiene un proverbio análogo según el cual “justice delayed is justice denied”. Vid. también la reflexión de SUSSKIND, R.: *Online Courts and the Future of Justice* Oxford: Oxford University Press, 2019, p. 290.

⁷² Este es un aspecto interesante y ambivalente del debate, especialmente en asuntos penales. Por una parte, cabe reconocer que mucho del efecto comunicativo procede de las penurias del proceso, del aire de solemnidad con el que las actuaciones son llevadas a cabo, y de la inclinación general a vincular la legitimidad de la justicia a su abordaje como algo *trascendente*. Sin embargo, con la solemnidad se paga posiblemente un precio. Es bien conocido el precio “kantiano” de la justicia, que obligaría a ejecutar la condena del último individuo sobre la tierra si este fuese culpable de un delito capital. Es igualmente sensato recordar la objeción de fondo planteada por Hulsman y el abolicionismo: al rodear los conflictos penales del peso grandilocuente de la justicia, se facilita la prosperidad de un sistema severo focalizado en responder al mal con otro mal con la esperanza de restablecer el orden ciudadano (y cósmico). Por el contrario, una visión más apegada a las situaciones sociales que dan lugar a conflictos señalados como delitos tendería a destacar las ventajas de una resolución orientada a devolver el conflicto a una escala más próxima a la de los implicados. Desdramatizar puede ser una vía para facilitar el restablecimiento de una situación normalizada para la víctima y el victimario (como sugerirían interpretaciones afines a la justicia restaurativa), tal vez al precio de trivializar la administración de justicia como institución. En un caso como en otro, se trata de decidir sobre las prioridades y el “carácter” que ha de tener la intervención institucional en la conflictividad social. Vid. HULSMAN, L. & BERNAT DE CELIS, J.: *Sistema Penal y Seguridad Ciudadana: Hacia Una Alternativa*, Ariel, 1984, p. 127; SOURDIN: *Judges, Technology and Artificial Intelligence. The Artificial Judge*, p. 106.

el proceso a otras dinámicas de resolución de conflictos de carácter informal o semiformal como la establecida por PayPal para atender a las disputas en el comercio online⁷³.

Finalmente, existe una última transformación interesante que se deduce de una propuesta de esta suerte. Al postular la mayor capacidad de un sistema artificial para atender a necesidades de justicia, igualdad o certeza, se está otorgando un valor positivo a cierto grado de “deshumanización”. Al deconstruir la realidad social en cúmulos de datos, los intervinientes quedan diluidos en una red abstracta de números y correlaciones de las cuales la persona es un subproducto sin especial trascendencia para la decisión. Igualmente, como hemos visto, el decisor artificial es también más abstracto por naturaleza que del decisor humano, pues se mueve en términos de relaciones formales entre variables. En una escapada teórico-política, podría incluso decirse que si los principios de la Ilustración se basaban en una abstracción funcional, la del ciudadano libre, racional e igual a los demás, la IA ofrece empujar esta abstracción al nivel del dato bajo la promesa de un resultado todavía más igualitario. En cierto sentido, pareciera que evaporar la humanidad de la toma de decisiones abre la puerta hacia un perfeccionamiento de la justicia que, inesperadamente, parece casi trascendental, fruto de una penetración más profunda en el tejido de la realidad.

3.2 Combatiendo la “tiranía del algoritmo”. La crítica humanista y la resurrección del juez humano

Si interrumpiéramos aquí el relato sería posible interpretar la llegada de la inteligencia artificial al debate jurídico como la culminación de los esfuerzos iusrealistas. Este diagnóstico sería, sin embargo, precipitado en varios puntos. Por una parte, no está claro que el realismo haya tenido la acogida necesaria entre la comunidad jurídica como para que el empuje de la inteligencia artificial pueda funcionar como “golpe de gracia”. En segundo lugar, resulta dudoso que el enfoque más característico del realismo, de afinidades económicas y psicológicas, sea la única forma de ser “realista”, algo que parece trascendente en el debate que nos ocupa por motivos que seguidamente veremos.

Así las cosas, podríamos sugerir que el enfoque “realista” de la relación entre Derecho e IA era esperable hasta cierto punto. Lo que tal vez no lo fuera tanto es la “resistencia humanista” que puede entreverse, particularmente, en la controversia sobre el Juez IA. Comenzando por lo básico, esta resistencia no debería ser vista como una forma de ludismo jurídico. Aunque posiciones extremadas nunca son descartables, parece más representativo de la opinión mayoritaria una actitud más matizada que funda su crítica, no en un rechazo, sino en el reconocimiento del gran potencial de la IA. Es precisamente por esto que las principales objeciones no se dirigen a la tecnología en sí, sino al exceso de confianza que conduce, en ocasiones, a desatender sus riesgos y limitaciones. Desde este punto de vista, el debate orbita en torno a qué cosas *no pueden* hacer los algoritmos y qué cosas, aun pudiendo, *no deberían* hacer.

En relación con la polémica de la automatización de las decisiones judiciales, esta disyuntiva es importante porque retrata (al menos implícitamente) una protesta que podría leerse como el reverso de la característicamente enarbolada por los realistas. Si, para el realismo, la IA aporta finalmente una respuesta frente a la repetida observación de que los jueces son demasiado “humanos” para estar a la altura de la trascendencia de su misión, para el humanismo la IA es demasiado fría, demasiado *artificial*, como para absorber

⁷³ SOURDIN: *Judges, Technology and Artificial Intelligence. The Artificial Judge*, p. 7.

completamente la función del juez. Con este último argumento, en realidad, se abarcan posturas con diferente grado de intensidad y alcance. En su forma más “restringida”, se sostiene que el modo de funcionar de un algoritmo presenta limitaciones para el tipo de tarea que se le encomendaría en caso de tener que sustituir al juez. No obstante, en su versión más fuerte, lo que se está defendiendo es que hay determinadas funciones intrínsecamente ligadas a la adjudicación que son irrenunciablemente humanas y que no se puede prescindir de ellas sin grave riesgo para la justicia en su conjunto.

Una de las posturas más marcadas que alimentan el argumentario “humanista” deriva de la distinción realizada por Searle entre la inteligencia humana y la artificial, a las que asocia, respectivamente, con la semántica y la sintaxis⁷⁴. De acuerdo con Searle, la IA puede abordar tareas que tradicionalmente han requerido de la inteligencia humana e, incluso, mejorar el *output* obtenido por las personas mediante un procesamiento basado exclusivamente en relaciones formales (sintaxis). La potencia computacional moderna, unida a la gran disponibilidad de datos, permiten que este sistema alcance su máximo rendimiento. Sin embargo, aquello de lo que las inteligencias artificiales carecen es de una mínima aprehensión del significado de las relaciones formales que manejan. La *forma* digital carecería de un *fondo* semántico, que es precisamente el pilar maestro de la inteligencia humana. En idéntico sentido, se ha dicho que la IA es en realidad “inteligencia por *proxy*”⁷⁵, o inteligencia en sentido metafórico: parece inteligencia porque produce resultados que, en humanos, requerirían de inteligencia⁷⁶.

Esta posición de signo “fuerte” es sumamente trascendente en materia de adjudicación, pues obliga a plantear la hipótesis del “Juez IA” por lo que es: una sustitución formal basada en criterios pragmáticos de rendimiento medidos sobre un resultado concreto. Este resultado suele ser la resolución de controversias jurídicas, cuya adecuación se mide en términos de equidad y seguridad jurídica, así como de celeridad y economía. Sin embargo, es legítimo dudar de si puede haber algo parecido a adjudicar sin una *interlocución*, es decir, sin un intercambio entre dos agentes inteligentes (en sentido pleno) cuya comunicación se sustenta, de entrada, en el reconocimiento mutuo como hablantes. A mayor abundamiento, el tipo de adjudicación que suele asociarse a la justicia penal complica las cosas, en la medida en que esta es una rama del Derecho particularmente “semántica” en sus aspiraciones y métodos. Sin perjuicio de la disponibilidad de argumentos utilitarios sobre el castigo y de los múltiples esfuerzos por “purgar” la penalidad de sus tintes morales, no parece que se haya podido prescindir hasta la fecha de ideas poco asépticas de reproche, responsabilidad y culpabilidad. Pensar la hipótesis de un Juez IA nos traslada a escenarios que, por el momento, resultan extraños: una justicia que condena, pero no reprocha, que absuelve, pero no “exculpa” (en términos semánticos, sería algo parecido a “sobreseer”, sin perjuicio de las particularidades procesales), o que, en caso de querer preservar cierto “significado”, debe remitirlo sin

⁷⁴ SEARLE, J.R.: “Minds, Brains, and Programs”, *Behavioural and Brain Sciences*, vol. 3, no. 3, 1980, p. 422.

⁷⁵ SURDEN, H.: “Machine Learning and Law”, *Washington Law Review*, vol. 89, no. 1, 2014, p. 97.

⁷⁶ Conviene recordar que, aunque esta sea una afirmación muy próxima a la definición estricta de inteligencia artificial, no es inusual que la discusión sobre la inteligencia se desplace rápidamente a conjeturas también marcadamente futuristas relacionadas con la superación de la inteligencia humana e, incluso, el advenimiento de entidades artificiales conscientes. Resulta interesante, desde el punto de vista de cómo pensamos los humanos, observar que conforme más esotéricas e insondables se hacen las tecnologías (el *Deep learning* es un ejemplo ya cotidiano de esto) más dispuestos estamos en atribuirles también insondables poderes. Surge en este punto, de manera espontánea, el recuerdo del ya proverbial adagio de Arthur C. Clarke de que “cualquier tecnología lo suficientemente avanzada es indistinguible de la magia”.

intermediarios al mandato del legislador. Sin perjuicio del problema titánico que ello supone en términos de separación de poderes⁷⁷, como mínimo habría que reconocer que el proceso entendido como mecanismo de comunicación quedaría seriamente diluido⁷⁸.

En conexión con esto, aunque con un alcance más restringido, Miller defiende que las explicaciones (y los argumentos jurídicos, añadiríamos) tienen una serie de características distintivas que las hacen específicamente sociales. Son un reflejo de cómo piensan los humanos, de cómo atribuyen causalidad y, significativamente, difieren notablemente de la forma en que una IA abordaría la tarea de “explicar”. Así, cuando las personas señalan un fenómeno específico como “causa” de otro, en realidad suelen seleccionar, entre un cúmulo de factores causales concomitantes, aquel que resulta más apropiado para cargar con el peso de la explicación⁷⁹. Las explicaciones multicausales son relativamente raras en la comunicación social, pues resultan mucho menos eficaces a la hora de transmitir la idea de que se está explicando propiamente (y no enumerando una serie de factores relacionados con el resultado, cuya importancia se desconoce). Explicar una conducta como *imputable* a un agente es un proceso similar de selección de factores causales, en el que se hace primar su voluntad por encima de otros factores concomitantes de diversa naturaleza⁸⁰. Pero este no es el modo en que una IA afronta esta tarea. Para el algoritmo, toda variable relevante es susceptible de entrar en el modelo matemático, y solo saldrá de él si su contribución es intrascendente. Una vez hecho este filtro, en cambio, queda el problema de transformar en explicación un modelo que se expresa en términos de correlaciones. Por un lado, porque correlación no significa causación⁸¹; por otro, porque el efecto de una decisión tomada sobre la base de probabilidades no es el mismo desde el punto de vista de su eficacia discursiva⁸² o de sus necesidades de justificación⁸³. Adicionalmente, no cabe olvidar que las variables

⁷⁷ WINTER: “The Challenges of Artificial Judicial Decision-Making for Liberal Democracy”, p. 197.

⁷⁸ En un caso como este, imaginar que el sujeto condenado (supongamos) se siente interpelado directamente por el legislador, ahora que no hay un juez-interlocutor, es imaginar un diálogo entre el ciudadano y la *volonté générale* que, a pesar de su interesante potencial teórico-político, parece demasiado embellecido por la abstracción como para presentarse en la realidad sin fisuras.

⁷⁹ “Las explicaciones no son solo la presentación de relaciones y causas (atribución causal), sino que son contextuales. Aunque un acontecimiento puede tener muchas causas, a menudo el explicando solo se preocupa por un pequeño subconjunto (relevante para el contexto), el explicador selecciona un subconjunto de este subconjunto (basándose en varios criterios diferentes), y el explicador y el explicando pueden interactuar y discutir sobre esta explicación”. MILLER, T.: “Explanation in Artificial Intelligence: Insights from the Social Sciences”, *Artificial Intelligence*, vol. 267, 2019, p. 3, <https://doi.org/10.1016/j.artint.2018.07.007>.

⁸⁰ Aquí, la argumentación jurídica ha desempeñado tradicionalmente un papel central, al explicar los motivos, legales y racionales que sustentan la decisión, haciendo uso para ello de un tipo de lenguaje que debería ser entendible y aceptable por el público en general y por las partes implicadas en particular. En términos ideales, incluso el condenado debería quedar convencido de los motivos por los que se le condena. En un sentido más moderado, debería al menos poder aceptar que se llegó a su condena por medio del derecho y la razón, de modo que el proceso sea “aceptable” aunque no “aceptado” (por considerar el resultado injusto o errado en algún sentido). Que esta pueda ser una representación ideal de la comunicación propia de una sala de justicia no la hace *per se* prescindible como principio rector e ideal a alcanzar. Este, en su caso, sería un debate distinto relativo a las piezas irrenunciables de un sistema jurídico que, mientras no se plantee explícitamente, no tiene por qué ser presumido en un sentido u otro.

⁸¹ El recordatorio es bien conocido. Baste entonces con una breve referencia al análisis condensado sobre explicación y causalidad de MILLER: “Explanation in Artificial Intelligence: Insights from the Social Sciences”, p. 5.

⁸² No es lo mismo decir que un individuo es culpable, a que “los hechos presentados son compatibles en un 73% con una decisión condenatoria”. Se podría objetar que el hecho de que ambos enunciados no tengan un impacto equivalente deriva de la falta de alfabetización matemática general y reconducir esto, en última instancia, a un

significativas para un modelo algorítmico pueden tener una relevancia limitada para justificar decisiones en asuntos jurídicos, pues lo mismo valen la conducta y la norma, que la presencia de palabras “clave” predictivas a nivel estadístico, pero intrascendentes a nivel explicativo⁸⁴.

Otro aspecto importante que ha suscitado críticas es la inherente falta de flexibilidad de la IA, que se ha discutido al menos de dos maneras: una que podríamos llamar “conceptual” y la otra “temporal”. Desde el punto de vista conceptual, la potencia de la IA se observa particularmente en toda operación o decisión que puede ser sustentada por medio de datos y las relaciones entre ellos. A la inversa, se ha presentado como un punto débil de la inteligencia artificial el tipo de decisión que se basa en conceptos escasamente delimitados, de los que abundan en el lenguaje natural y en el Derecho, fruto de la “textura abierta” que se le atribuye⁸⁵. En efecto, conceptos centrales como los de analogía, ponderación, principios, así como las sutilezas de las interpretaciones no gramaticales de la norma, parecen operaciones tan naturales para el pensamiento humano como difíciles de convertir al universo digital.

Desde el punto de vista “temporal”, por otro lado, la IA basa su rendimiento en inferencias y reglas elaboradas tras el análisis de una cantidad suficientemente importante de datos disponibles. Si regresamos al contexto de las decisiones judiciales, esto quiere decir que la decisión sobre casos nuevos es, en realidad, una predicción de ajuste: se dice “este caso estará correctamente resuelto si se le da la misma respuesta que en este otro subconjunto

remanente de irracionalidad más que habría que exorcizar de la sociedad y del Derecho. Las decisiones judiciales conllevan siempre un margen de error del que, en principio, no se suele hablar demasiado con el fin de no restar eficacia comunicativa a los pronunciamientos judiciales. Ahora bien, parece distinto reconocer un error cuyo origen y entidad resultan desconocidas, que actuar con conocimiento de una determinada tasa de error (especialmente cuando este es en parte fruto de decisiones en el diseño del modelo estadístico), *vid.* BONSIGNORE FOUQUET, D.: “Valoración Del Riesgo y Discrecionalidad Informada Del Juez: A Propósito Del Evidence-Based Sentencing y El Significado Del Error de Predicción”, *Diálogos Jurídicos*, nº 5, 2020, p. 32. Con todo, Sunstein se ha expresado sobre este problema entendiendo que, a pesar de las implicaciones éticas, sigue siendo mejor decidir a conciencia que evadir el problema con la “opacidad” del desconocimiento, *vid.* SUNSTEIN: “Governing By Algorithm? No Noise and (Potentially) Less Bias”, p. 1203.

⁸³ Por ejemplo, aunque los varones jóvenes puedan estar en mayor riesgo de cometer delitos que las mujeres de su rango de edad o que los ancianos, no queda claro en qué modo esta información debería servir para justificar mayor pena para unos, menor pena para otros, o cualquier tipo de consecuencia jurídica en general, más allá de aquellas que tienen por fundamento, precisamente, un juicio probabilístico. MONAHAN, J. & SKEEM, J.L.: “Risk Assessment in Criminal Sentencing”, *Annual Review of Clinical Psychology*, vol. 12, no. 1, 2015, p. 504.

⁸⁴ El ejemplo de los filtros de spam es significativo en este caso. El criterio para detectar un mensaje no deseado frecuentemente no tendrá que ver con el contenido del mensaje, sino con el uso de palabras que aparecen frecuentemente en otros mensajes no deseados, como “amigo”, “urgente”, “premio” o similares. Su adecuación depende del resultado (de que filtren bien), pero no permiten arbitrar un criterio de selección mínimamente relevante en términos de contenido. Trasladar esto a asuntos penales resulta problemático, y podría incluso abrir la puerta a que ciertas palabras se alzaran en auténticos criterios formales de decisión sin relación de contenido ni con los hechos, ni con la norma. Desde este punto de vista (y asumiendo que el procedimiento en casos de justicia algorítmica fuera similar al empleado por los filtros de spam), no queda claro en qué punto quedaría la legitimidad de una sentencia que se basa en una similitud con el precedente medida sobre criterios jurídicamente irrelevantes. Una explicación sumamente didáctica del funcionamiento de los filtros de spam y, en general, una discusión sobre la relevancia de las “razones” de los algoritmos puede encontrarse en BURRELL: “How the Machine ‘Thinks’: Understanding Opacity in Machine Learning Algorithms”, p. 7–9.

⁸⁵ Según expresión exitosa de H.L.A. Hart. Para una discusión, *vid.* BENCH-CAPON, T.J.M. & DUNNE, P.E.: “Argumentation in Artificial Intelligence”, *Artificial Intelligence*, vol. 171, no. 10–15, 2007, p. 633, <https://doi.org/10.1016/j.artint.2007.05.001>; RISSLAND, E.L., ASHLEY, K.D., & LOUI, R.P.: “AI and Law: A Fruitful Synergy”, *Artificial Intelligence*, vol. 150, 2003, p. 5, [https://doi.org/10.1016/S0004-3702\(03\)00122-X](https://doi.org/10.1016/S0004-3702(03)00122-X); SOURDIN, *Judges, Technology and Artificial Intelligence. The Artificial Judge*, p. 229.

de casos con los que muy probablemente está emparentado, a la vista de las características que comparten”. Esto, aunque extraño, puede ser más o menos aceptable como una traslación estadística de un juicio por analogía, que no es muy distinto a cómo actúa (teóricamente) un juez que acude a la jurisprudencia. Sin embargo, en caso de no existir precedentes, se confía en que el Juez tendrá a su disposición otros mecanismos y argumentos para salir del *impasse* y alcanzar una decisión razonable. Una escapatoria semejante no está tan clara en el caso de los algoritmos, cuya referencia constante al pasado para formar sus reglas de decisión los hace primordialmente “retrospectivos”⁸⁶. Esto imprime al ordenamiento jurídico de cierta rigidez, de cierto “conservadurismo” por diseño que, si bien podría argumentarse que está en la naturaleza del tipo de estabilización que se busca por medio del Derecho, se encuentra privado del resquicio concedido a la judicatura para adaptar la norma al periodo histórico en que ha de ser aplicada. En el estado actual de la tecnología, parece que habilidades como la sensibilidad al cambio histórico y cultural, que motivan la innovación judicial y la desviación del precedente⁸⁷, no parecen encajar con el proceso algorítmico.

De manera un tanto paradójica, podemos observar con este ejemplo que algunos de los aspectos que han fundado tradicionalmente la crítica realista y los temores frente a la arbitrariedad judicial, el error y la injusticia, resurgen reconvertidos en virtudes sistémicas⁸⁸. La maleabilidad semántica del lenguaje natural de las leyes es una de las principales virtudes de la maquinaria jurídica cuando se la asocia con una labor judicial cuya finalidad no es solo aplicar la norma, sino lidiar con ella para adaptarla a los tiempos, así como a valores y principios primordiales (digamos, constitucionales) que pueden demandar un cambio interpretativo. Pero este juez que, desviándose del Derecho pasado, refuerza la legitimidad del ordenamiento jurídico presente, no encaja bien con el escepticismo de los realistas. Para ellos, este sería precisamente el espacio de acción de un juez sesgado, de una judicatura inconsistente, y el caldo de cultivo para una arbitrariedad ocultada por la retórica. Curiosamente, pareciera como si el juez perfecto para el realismo fuera el juez *bouche de la loi* en el que, por otra parte, nunca creyeron.

Finalmente, existe un aspecto importante de la controversia que tiene que ver con el papel que desempeña, no ya el juez en sí, sino la institución judicial como institución “encarnada” por el juez en el seno de las comunidades humanas. Aquí encontramos de nuevo una oposición que, frecuentemente, acaba planteándose de manera excesivamente polarizada. Por una parte, como hemos visto, un sector de los realistas favorables a la “hipótesis del Juez IA” se apoya en una visión pragmática de la labor judicial, entendida como mecanismo de resolución de controversias. A esta concepción, sin embargo, el humanismo reacciona en ocasiones con una visión más holística del asunto. Los argumentos son diversos, pero podría decirse que convergen en torno a criticar el reduccionismo derivado de definir la labor del juez en un sentido tan restrictivo⁸⁹. Y esto no solo por influencia de sistemas jurídicos en los que existe un elemento político-electoral en la carrera judicial. Más bien, se suele hacer referencia a la

⁸⁶ KATYAL, S.K.: “Democracy & Distrust in an Era of Artificial Intelligence”, *Daedalus*, vol. 151, nº. 2, 2022, p. 330, https://doi.org/10.1162/DAED_a_01919; SURDEN: “Machine Learning and Law”, p. 106.

⁸⁷ POSNER, R.A.: “What Do Judges and Justices Maximize? (The Same Thing Everybody Else Does)”, *Supreme Court Economic Review*, vol. 3, 1993, p. 18, <https://doi.org/10.1086/scer.3.1147064>; SOURDIN: *Judges, Technology and Artificial Intelligence. The Artificial Judge*, p. 34. Naturalmente, estas son labores “creativas” que se encuentran bastante asumidas en la literatura contemporánea, pero no por ello son pacíficas entre los juristas, v.gr. LAPORTA SAN MIGUEL: “La Creación Judicial y El Concepto de Derecho Implícito”, p. 147.

⁸⁸ Sospecho que esta es una posible lectura (tal vez extremada) del estado de cosas descrito, en relación con la idea de “post-positivismo” jurídico, por ATIENZA: *Curso de Argumentación Jurídica*, p. 28–30.

importancia de que haya una *persona* a cargo del proceso, que atiende a las partes en un sentido tanto profesional como humano, que asume una posición de responsabilidad sobre el resultado del juicio y que, como mínimo, se encuentra obligado a proporcionar los motivos de su decisión a todos los implicados.

En este sentido, se podría incluso criticar al realismo su falta de, precisamente, contacto con un espectro de la realidad que no resulta inmediatamente reconducible al marco económico y psicológico abundante en esta corriente. En este sentido, existe información que pone de manifiesto que la percepción pública de la IA en materia de toma de decisiones es compleja y matizada, pero no necesariamente coherente con el utilitarismo reduccionista⁹⁰. Al contrario, parece que se desprende que, ante cierto tipo de decisiones, normalmente de particular trascendencia (como una sentencia penal) las personas priorizan la obtención de razones por encima del grado de “acierto” del resultado. Algo similar se ha planteado respecto de decisiones como un despido o un rechazo en una oferta de trabajo, aunque por motivos ligeramente distintos: donde la automatización priva de las explicaciones y las razones que daría un decisor humano, se dificulta sensiblemente la posibilidad de adaptación y de mejora para el aspirante. Asimismo, la “seguridad jurídica”, que muchos derivarían de la asignación del mismo resultado a casos iguales, es también valorada de forma matizada, resaltando la necesidad de un sistema en el que el factor humano introduce cierta imprevisibilidad, que es también dinamismo y apertura⁹¹. En el ámbito laboral, esta “disparidad” permite volver a probar suerte en otra entrevista de trabajo, en el ámbito jurisdiccional, da sentido al sistema de recursos. Un Juez IA que maximizara la “igualdad” correría el riesgo de convertirse en determinista, y el determinismo tiene cierta inclinación por alienar a los humanos que, con buen criterio o no, consideran que existe un resquicio de sus vidas sobre el que está a su alcance influir.

Existe información convergente que podría invocarse en apoyo de estas ideas, por ejemplo, el énfasis en la legitimidad institucional que hacen las corrientes de justicia procedimental, así como de justicia restaurativa y resolución alternativa de conflictos. Aunque su relación con la justicia penal haya sido ambivalente, parece que es una premisa central de estos trabajos la importancia de “humanizar” o “desformalizar” la gestión del conflicto⁹². Con independencia de las opiniones que esto suscite, a lo que se apunta en última instancia es

⁸⁹ SOURDIN: *Judges, Technology and Artificial Intelligence. The Artificial Judge*, p. 207–8.

⁹⁰ No se entienda “utilitarismo reduccionista” en un sentido demasiado peyorativo, sino como una forma de caracterizar una concepción de la función judicial guiada por la idea de resolución de conflictos y de la utilidad social derivada de ello. Esta concepción no es necesariamente desacertada en todo contexto, es más, seguramente quepa ensalzar el sentido de las prioridades de muchos autores que escriben desde una posición como esta. Aquí el reduccionismo se refiere a la concepción de la labor judicial y de su lugar en las sociedades humanas, fuertemente condicionada por el enfoque de psicología conductista y de *law and economics* imperante. Como se sugiere en el texto, el “empirismo realista” parece no tener gran cosa que decir sobre las implicaciones de otro tipo de observaciones sobre la realidad, como el lugar sumamente particular que las instituciones judiciales han ocupado en las comunidades humanas. Desde este punto de vista, resulta natural que la concepción del Derecho como un instrumento de humanos para humanos sea considerada como un ejemplo más de la historia del romanticismo: otorgamos a las actividades humanas un valor añadido que no tiene que ver con su utilidad, sino con una proyección general del valor (¿injustificado?) que nos otorgamos.

⁹¹ Las afirmaciones del párrafo relativas a las percepciones sociales remiten, no solo a las opiniones expresadas por la propia academia (como parte integrante, también, del público general), sino de la dirección apuntada en: WOODRUFF, A. ET AL.: “A Cold, Technical Decision-Maker’: Can AI Provide Explainability, Negotiability, and Humanity?”, 2020, p. 1–23; YALCIN, G. ET AL.: “Perceptions of Justice By Algorithms”, *Artificial Intelligence and Law*, 2022, <https://doi.org/10.1007/s10506-022-09312-z>.

a los efectos perjudiciales de la frialdad y el burocratismo. Si esto es así, solo cabe augurar peores resultados de una introducción de la IA Judicial poco sutil. En este sentido, el punto no es que toda decisión deba ser tomada por una persona. Se trata más bien de que, en determinados contextos, existe tal expectativa y esta no puede ser colmada con un “chatbot” convincente (sin engañar a los intervinientes, al menos). Por esto, una distinción entre casos triviales y no triviales⁹³, que puede resultar útil para delimitar un campo para el Juez IA, no debería asentarse sobre criterios exclusivamente jurídicos (complejidad del caso, gravedad de la infracción), sino también en otros relativos a las circunstancias de las partes: su percepción subjetiva de la trascendencia del asunto, su posible situación de vulnerabilidad y otras consideraciones que, en conjunto, hicieran recomendable preservar el “elemento humano”. Me parece que la “resistencia humanista” se dirige, precisamente, a buscar los espacios no triviales de adjudicación, donde no cabe extirpar al juez humano sin grave riesgo para el sistema jurídico.

No resulta extravagante pensar que la función social de “hacer justicia” sea una de las más nucleares para la estabilización de las comunidades humanas. En este sentido, a pesar del papel que la “verdad” desempeña en el Derecho de matriz ilustrada, no parece que la función de adjudicar haya estado indisociablemente ligada a ella en sentido fuerte, o no tanto como a lo que una comunidad pudiera aceptar como *narración adecuada* de los hechos objeto de controversia. Dicho de otro modo, la utilidad y eficacia social de la institución depende de forma muy relevante de los códigos que una sociedad emplea para reconocer las decisiones con eficacia dirimente. Para ello, se sirve de reglas de autoridad, de comportamientos y expresiones ritualizadas, etc. en el fondo, depende de una “práctica” que envuelve la adjudicación y le concede su legitimidad. Por ejemplo, es porque la comunidad considera al juez una autoridad digna de la tarea que escucha su postura y acepta su valor dirimente, aunque discrepe del resultado. En tales casos, la verdad o el acierto pueden cuestionarse, pero el “gesto” adjudicatorio mantiene su eficacia de “cierre”⁹⁴. En este sentido, crear una situación social de esta naturaleza parece requerir de un humano con un determinado capital social

⁹² Naturalmente, aquí se plantea, a efectos discursivos, lo que en realidad es una falsa oposición. La humanidad y la cercanía no están necesariamente reñidas con la preservación de las necesarias formalidades. El punto es, más bien, que un apego excesivo al procedimiento puede ir en detrimento de la dimensión humana (y viceversa), y que este es un factor fuertemente ligado a la percepción de legitimidad institucional. Sobre el particular, y con amplia referencia a la investigación sociológica y psicológica sobre legitimidad, autoridad y justicia (procedimental y distributiva), *vid.* HEGTVEDT, K.A., JOHNSON, C., & WATSON, L.: “Social Dynamics of Legitimacy and Justice”, en *Handbook of Social Justice Theory and Research*, Springer New York, 2016, p. 430, https://doi.org/10.1007/978-1-4939-3216-0_23.

⁹³ Que sospecho que no habría que equiparar a la clásica distinción entre casos “fáciles” y “difíciles”, en la medida en que un caso jurídicamente fácil puede ser no-trivial por muchos motivos, desde el tipo de asunto, sus repercusiones, o el impacto subjetivo que haya tenido en las partes, entre otros. Sobre los casos fáciles y difíciles en el debate que nos ocupa, es clásico referirse a TARUFFO, M.: “Judicial Decisions and Artificial Intelligence”, *Artificial Intelligence and Law*, vol. 6, 1998, p. 318, https://doi.org/10.1007/978-94-015-9010-5_7.

⁹⁴ Esto no viene necesariamente contradicho por las polémicas sociales en torno a las decisiones judiciales, que tienden a impugnar el contenido del Derecho antes que la legitimidad del juzgador. La decisión “vale” en la medida en que se enmarca en las reglas del juego comúnmente aceptadas. Precisamente por esto, la agitación pública no suele proponer la inaplicación de las resoluciones judiciales, sino un cambio en el marco normativo que las determina. Algunas investigaciones que han intentado comparar los efectos de la justicia procedimental a efectos de legitimidad han tendido a señalar la especial importancia del trato dispensado por las autoridades y la adecuación de su comportamiento, con independencia de la resolución dictada. Esto podría deberse a la percepción de que los aspectos procedimentales se encuentran más sometidos al control voluntario del agente de la autoridad que el resultado en sí, sobre el que confluyen muchos más factores de indeterminación. Pero aunque se estuviera hablando, en sentido estricto, de la legitimidad atribuida a los agentes, la literatura suele destacar que se trata de una propiedad

cuyo origen se encuentra en el reconocimiento que la propia comunidad le otorga. Aunque no es inconcebible pensar que una IA pueda quedar investida de mayor legitimidad y autoridad que un juez humano, esto no parece sumamente probable a la vista de que, hasta el momento, el judicial no ha sido un papel desempeñado por cualquier persona, sino por miembros cualificados del colectivo. En este punto, una actitud “realista” demanda escepticismo frente a utopías inminentes que, en su planteamiento, sugieren algo muy parecido a una escatología de la tecnología⁹⁵. Con todo, la religión tal vez tenga más vínculo con la historia del ser humano que el racionalismo materialista, de modo que, sobre este punto, las cartas están seguramente todavía sobre la mesa.

4. REFLEXIONES CONCLUSIVAS: LA IMPOSIBILIDAD DE UN JUEZ

Llegados a este punto, hemos podido ver cómo la irrupción de la inteligencia artificial a empellones ha impulsado una efervescencia de debates y discusiones en el ámbito del Derecho. Entre ellas, la conjetura del “Juez IA” no es seguramente la más precisa, ni sus repercusiones las más inmediatas ni urgentes. No obstante, resulta especialmente interesante porque, a la vista de los niveles de futurismo y especulación que el campo tolera en virtud de su desarrollo aún incipiente, permite canalizar reflexiones que, de otro modo, no tendrían ocasión de presentarse. En particular, lo que comienza frecuentemente como una discusión en torno a las capacidades y limitaciones técnicas de los algoritmos acaba en una reflexión de fondo sobre las implicaciones jurídicas, sociales y éticas de la labor del juez.

Fruto de ello, regresan viejos debates, y aquí nos hemos ocupado de seguir la pista a una impugnación fundamental: la idea de que el ideal del rol judicial resulta incompatible con la realidad a la vista de las limitaciones intrínsecamente humanas del juez. Ya la idea del juez como boca de la ley ofrece pistas tempranas del problema de cuánto poder (y cuánta confianza) otorgar a la institución judicial en el uso de sus facultades discrecionales. Aquí,

que se extiende por “analogía”: la legitimidad inspirada por un operador jurídico redundante sobre la institución en su conjunto. *V.gr.* HEGTVEDT, JOHNSON, & WATSON: “Social Dynamics of Legitimacy and Justice”, p. 433.

⁹⁵ Entiéndase este comentario como una representación conscientemente exagerada. No obstante, el gesto literario tal vez pueda ser llevado hasta sus últimas repercusiones en el espacio más discreto de una nota a pie de página. Es mi impresión particular que, tras un examen panorámico del debate aquí expuesto, un lector con inclinaciones sociológicas o antropológicas podría encontrar detalles bien interesantes desde el punto de vista de su proximidad teórica con debates teológicos en torno a la oposición de la justicia humana *versus* la justicia divina. En primer lugar, del catálogo de imperfecciones del juez humano es fácil extraer la implicación de que ningún mortal se encontraría a la altura de la trascendencia de los valores abstractos que rigen la misión de “hacer justicia”. Siguiendo con la alegoría, resulta tentador ver en la fascinación esperanzada que circunda la “utopía algorítmica” una prolongación de la idea de confeccionar una deidad artificial a cuyos designios encomendarse. Esto resulta menos forzado cuando se recuerda el debate sobre la opacidad y el *Deep learning*: cuando se sostiene la utilidad de un mecanismo decisional cuyo funcionamiento es un “misterio” para sus diseñadores, se externaliza la responsabilidad más allá del juicio y conocimiento humanos. Confiar en lo insondable no es bien distinto de encomendarse a la providencia, si bien la actitud expectante del creyente vendría probablemente reemplazada por la pugna de los teólogos: una discusión sobre cómo y por qué cabe entender que ciertos resultados, de origen insondable, son fruto del buen o mal funcionamiento de la IA. Esto no es muy distinto a ponderar si las calamidades son un castigo merecido, si se trata de un problema de “diseño” o si el problema está en “nosotros”. Finalmente, y desplazando la vista al bando contrario, la defensa abnegada del juez humano que, frente a la uniformidad maquínica, sabría valerse de su “libre albedrío” para tomar la decisión correcta en el momento correcto tiene algo de “Juez redentor”. Un juez por cuyo acto juicioso y noble, no determinado por el estado de cosas previo, materializa una forma de justicia tan trascendente que redime con su gesto al resto de sus imperfectos, sesgados y desviados colegas. Salvar al “juez salvador” o entregarse a la divinidad artificial, actúan como “fantasías” del debate que nos ocupa: por existir en el *imaginario* (como un *fantasma* del debate real) y, tal vez, por expresar un *deseo*, inconfesable, de trascendencia.

no obstante, se ha preferido anclar la discusión al tipo de enfoque y “actitud” adoptados por el realismo jurídico, trazando una línea de continuidad que nace de la problematización de la relación entre decidir y justificar, transita por las limitaciones cognitivas y emotivas que pesan sobre los decisores humanos y desemboca en una esperanzada valoración de las virtudes de la inteligencia artificial judicial. Así retrazada, se hace visible la historia de la institución judicial como ideal inalcanzable, edificada sobre nociones como racionalidad, justicia, equidad, imparcialidad y otros símbolos de trascendencia que encallan contra la imperfección humana. La argumentación jurídica, recurso por excelencia para apaciguar estas incertidumbres, sirvió especialmente en ausencia de alternativas viables, si bien siempre bajo sospecha de dar carta de naturaleza jurídica a motivaciones innobles, inconfesables en sentido pleno. No es de extrañar que la enumeración de las posibilidades auguradas por la inteligencia artificial se haya formulado siempre con apoyo en el diagnóstico poco halagüeño del realismo, entonces. De hecho, el impacto de esta propuesta procede, precisamente, de su vocación de optimizar el sistema de justicia con arreglo a valores propiamente jurídicos: igualdad, prontitud, seguridad jurídica y, como *Gestalt* de todo ello, “justicia”.

Frente al envite de este tipo de argumentos, no obstante, se viene observando un movimiento doctrinal relativamente informe que hemos optado por agrupar en torno a la idea de “resistencia humanista”. Así, más que erigir una oposición radical a los argumentos “realistas” o a la penetración de la IA en la administración de justicia, tratan de poner sobre la mesa aquellos aspectos insuficientemente considerados por un futurismo exagerado. El aspecto nuclear que une estos pronunciamientos heterogéneos, con todo, radica en la defensa de la preservación de un reducto, de extensión variable, para el juez humano. Y ello no por mor del romanticismo, sino por entender que la adjudicación es, en algunos casos, una tarea inevitablemente social: se erige sobre una expectativa de interlocución, de intersubjetividad, que puede versar sobre asuntos fuertemente “semánticos” como, en materia penal, la atribución de responsabilidad y la formulación de reproche. Adicionalmente, la institución judicial ocupa una posición nuclear dentro del organigrama político de los sistemas erigidos sobre la base de una mínima separación de poderes. El hecho de que, en todos estos puntos, se dirija la mirada hacia aspectos sumamente nucleares, tal vez “vitales”, de la labor judicial, concede fuerza a la advertencia de que ciertas transformaciones no son concebibles sin grave riesgo para la sostenibilidad del sistema jurídico-político en su conjunto. Esto no quiere decir que no pueda plantearse la necesidad de una bakuniniana “destrucción creativa”, pero estos no suelen ser los términos del debate energizado por cierto optimismo tecnológico. En este punto, la ascendencia del enfoque realista en trabajos de psicología y economía conductual se enfrenta a la crítica de su reduccionismo y, en última instancia, de falta de realismo, cuando desatiende otras dimensiones propias de la realidad social de donde provienen muchas de las críticas “humanistas”.

Leída conjuntamente, la oposición aquí analizada resulta sumamente interesante en la medida en que se dirige de manera bastante explícita a una discusión sobre los fundamentos mismos de los sistemas jurídico-políticos y la posición que, en ellos, ocupa la adjudicación. Es por ello, además, que resulta en buena medida irresoluble, pues traslada un conflicto que remite a cosmovisiones diferentes sobre el papel del juez. La visión pragmática, que define la labor judicial como instrumento de resolución de conflictos, optimiza la eficiencia del proceso y la eficacia de los rendimientos. Cuantos más casos encuentran su resolución, de la manera más rápida y económica, mejor. La visión humanista, que define la labor judicial, además, como un espacio de comunicación social en el que, por medio de la controversia, se reactivan una y otra vez los elementos fundamentales del Derecho, dirige su atención al “elemento humano”. El debate a propósito de la inteligencia artificial judicial, en este sentido,

aparece como una nueva iteración de otras oposiciones. Clásicamente, el consecuencialismo frente al deontologismo. También, el juez del como *boca de la ley* frente al juez como pequeño filósofo. Sobre estos mimbres, no puede sino plantearse la radical imposibilidad de un juez. Por imposibilidad de materializar los ideales que caracterizan su función, desecharíamos al humano por exceso de humanidad y al algorítmico por excesiva artificialidad.

No obstante, a los efectos de ulteriores discusiones el enfrentamiento aquí descrito cumple una función fundamental. A mi juicio, del grueso de las posturas realistas deberíamos extraer una advertencia sobre los riesgos del juicio humano y sus fallas. Sin perjuicio de detalles, el realismo advierte eficazmente contra el exceso de confianza y sugiere extremar las cautelas en todos aquellos casos en que las inclinaciones cognitivas humanas planteen un riesgo importante para la satisfacción de los valores jurídicos y la tutela de los derechos de las personas. Paralelamente, el humanismo enfatiza la importancia del juicio humano en situaciones de elevada trascendencia, donde la expectativa de interlocución se hace en principio mayor. Consideradas ambas, tal vez cabría inferir que, en la ruta hacia la “algoritmización” de la justicia, el espacio propicio para la experimentación sea el proporcionado por aquellos casos en que el riesgo de sesgo sea significativo y la trascendencia relativamente limitada. En cambio, cuando ambos se presentan con intensidad, es preciso reconocer abiertamente que la toma de posición sobre cuánta “algoritmización” es posible implica un posicionamiento inevitable sobre la arquitectura misma jurídico-política. Como seguramente no puedan optimizarse todos los intereses jurídicamente relevantes al mismo tiempo, el estado “subóptimo” de unos respecto de otros requiere de una justificación expresa.

Tal vez una orientación fructífera para la discusión futura radique en un “realismo no reduccionista”, consciente de que la aplicación del Derecho tiene mucho de juicio por combate, y que el imperio de la razón se busca como producto, pero no se asume como algo dado. Asimismo, semejante realismo tal vez apuntara a que, aunque pueda haber algo de placebo en el irredento romanticismo de las personas, si la justicia ha de ser humana antes que divina, no puede evitar someterse a los términos en que las sociedades humanas le conceden su valor. Tal vez estos términos estén sujetos a cambio, y quizás la justicia algorítmica pueda alcanzar la legitimidad necesaria para hacer prescindible la implicación humana. No obstante, resulta poco claro en qué punto dejaría este desenlace a la conocida narración de la historia cultural de las sociedades occidentales: si con la “muerte de Dios” vino la carga de la responsabilidad de erigir una sociedad por y para sí misma, la utopía algorítmica puede parecer un retorno a los orígenes, fruto del cansancio del “último hombre”.

Bibliografía

- AMODIO, D.M., & M. CIKARA: “The Social Neuroscience of Prejudice”, *Annual Review of Psychology*, vol. 72, 2021, p. 439–69. <https://doi.org/10.1146/annurev-psych-010419-050928>.
- ATIENZA, M.: *Curso de Argumentación Jurídica*, Madrid, Trotta, 2013.
- BENCH-CAPON, T.J.M., & DUNNE P.E.: “Argumentation in Artificial Intelligence”, *Artificial Intelligence*, vol. 171, nº. 10–15, 2007, p. 619–41. <https://doi.org/10.1016/j.artint.2007.05.001>.
- BERK, R.: *Machine Learning Risk Assessments in Criminal Justice Settings*, Springer, 2018. <https://doi.org/10.1007/978-3-030-02272-3>.

- BONEZZI, A., M. OSTINELLI, & MELZNER J.: "The Human Black-Box: The Illusion of Understanding Human Better Than Algorithmic Decision-Making." *Journal of Experimental Psychology: General*, 2022, p. 1–9. <https://doi.org/10.1037/xge0001181>.
- BONSIGNORE FOUQUET, D.: "Valoración Del Riesgo y Discrecionalidad Informada Del Juez: A Propósito Del Evidence-Based Sentencing y El Significado Del Error de Predicción." *Diálogos Jurídicos*, vol. 5, 2020, p. 21–36.
- BOURDIEU, P. & WACQUANT L.: *Una Invitación a La Sociología Reflexiva*, Buenos Aires, Siglo XXI, 2005.
- BREGER, M.L.: "Making the Invisible Visible: Exploring Implicit Bias, Judicial Diversity, and the Bench Trial", *University of Richmond Law Review*, vol. 53, n.º. 4, 2019, p. 1039–84.
- BURRELL, J.: "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms", *Big Data & Society*, vol. 3, n.º. 1, 2016, p. 1–12. <https://doi.org/10.1177/2053951715622512>.
- CAMPOS ZAMORA, F.J.: "Nociones Fundamentales Del Realismo Jurídico", *Revista de Ciencias Jurídicas*, n.º. 122, 2010, p. 191–220.
- CHALFIN, A. & MCCRARY J.: "Criminal Deterrence: A Review of the Literature", *Journal of Economic Literature*, vol. 55, n.º. 1, 2017, p. 5–48. <https://doi.org/10.1257/jel.20141147>.
- CITRON, D.K.: "Technological Due Process", *Washington University Law Review*, vol. 85, n.º. 6, 2007, p. 1249–1313.
- CORMACK, W.: "Reassessing the Judicial Empathy Debate: How Empathy Can Distort and Improve Criminal Sentencing", *Mitchell Hamline Law Review*, vol. 47, 2021.
- CROSS, F.B.: "Decisionmaking in the U.S. Circuit Courts of Appeals", *California Law Review*, vol. 91, n.º. 6, 2003, p. 1457–1515. <https://doi.org/10.2307/3481397>.
- DECETY, J.: "Why Empathy Is Not a Reliable Source of Information in Moral Decision Making", *Current Directions in Psychological Science*, vol. 30, n.º. 5, 2021, p. 425–30. <https://doi.org/10.1177/09637214211031943>.
- DROBAK, J.N., & NORTH D.C.: "Understanding Judicial Decision-Making: The Importance of Constraints on Non-Rational Deliberations", *Journal of Law & Policy*, vol. 865, n.º. 1988, 2004, p. 131–52.
- DWORKIN, R.: "Hard Cases", *Harvard Law Review*, vol. 88, n.º. 6, 1975, p. 1057–1109. <https://doi.org/10.2307/1340249>.
- EBERHARDT, J.L., DAVIES P.G., PURDIE-VAUGHNS V.J., & JOHNSON S.L.: "Looking Deathworthy: Perceived Stereotypicality of Black Defendants Predicts Capital-Sentencing Outcomes", *Psychological Science*, vol. 17, n.º. 5, 2006, p. 383–86. <https://doi.org/10.1111/j.1467-9280.2006.01716.x>.
- EDWARDS, L. & VEALE M.: "Enslaving the Algorithm: From a 'Right to an Explanation' to a 'Right to Better Decisions'?", *IEEE Security and Privacy*, vol. 16, n.º. 3, 2018, p. 46–54. <https://doi.org/10.1109/MSP.2018.2701152>.
- FRANK, J.: *Law and the Modern Mind*, Stevens & Sons, 1949.
- GILLMAN, H.: "What's Law Got to Do with It? Judicial Behaviorists Test the 'Legal Model' of Judicial Decision Making", *Law & Social Inquiry*, vol. 26, n.º. 2, 2001, p. 465–504. <https://doi.org/10.1111/j.1747-4469.2001.tb00185.x>.

- GLYNN, A.N., & SEN M.: "Identifying Judicial Empathy: Does Having Daughters Cause Judges to Rule for Women's Issues?", *American Journal of Political Science*, vol. 59, n.º. 1, 2015, p. 37–54. <https://doi.org/10.1111/ajps.12118>.
- GOMES DE SOUSA, W., ANTUNES FIDELIS R., DE SOUZA BERMEJO, P.H., DA SILVA GONÇALO, A.G. & B. DE SOUZA MELO: "Artificial Intelligence and Speedy Trial in the Judiciary: Myth, Reality or Need? A Case Study in the Brazilian Supreme Court (STF)", *Government Information Quarterly*, vol. 39, 2022, p. 1–13. <https://doi.org/10.1016/j.giq.2021.101660>.
- GROSAN, C. & ABRAHAM, A.: *Intelligent Systems. A Modern Approach*, Springer, 2011.
- GUTHRIE, C., RACHLINSKI, J.J. & WISTRICH, A.J.: "Blinking on the Bench: How Judges Decide Cases", *Cornell Law Review*, vol. 93, n.º. 1, 2007, p. 1–43.
- HABA, E.P.: "Razones Para No Creer En La Actual Teoría (Ilusionista) de La Argumentación. Papel Que Tales Teorizaciones Cumplen Como Más Nueva Ideología de Legitimación Con Respecto, Especialmente, Al Discurso Jurídico Profesional", *Doxa Cuadernos de Filosofía Del Derecho*, vol. 33, 2010, p. 321–60.
- HARRIS, L.T. & FISKE, S.T.: "Dehumanizing the Lowest of the Low: Neuroimaging Responses to Extreme Out-Groups", *Psychological Science*, vol. 17, n.º. 10, 2006, p. 847–53. <https://doi.org/10.1111/j.1467-9280.2006.01793.x>.
- HEGTVEDT, K.A., JOHNSON, C. & WATSON, L.: "Social Dynamics of Legitimacy and Justice", *Handbook of Social Justice Theory and Research*, Springer New York, 2016, p. 425–44. https://doi.org/10.1007/978-1-4939-3216-0_23.
- HULSMAN, L., & BERNAT DE CELIS, J.: *Sistema Penal y Seguridad Ciudadana: Hacia Una Alternativa*, Ariel, 1984.
- IRWIN, J.F. & REAL, D.L.: "Unconscious Influences on Judicial Decision-Making: The Illusion of Objectivity", *McGeorge Law Review*, vol. 42, n.º. 1, 2008, p. 1–18.
- JIN, Y. & HE, H.: "An Artificial-Intelligence-Based Semantic Assist Framework for Judicial Trials", *Asian Journal of Law and Society*, vol. 7, n.º. 3, 2020, p. 531–40. <https://doi.org/10.1017/als.2020.33>.
- JOHNSON, S.L., HRITZ, A.C., ROYER, C.E. & BLUME, J.H.: "When Empathy Bites Back: Cautionary Tales from Neuroscience for Capital Sentencing", *Fordham Law Review*, vol. 85, n.º. 2, 2016, p. 573–98.
- KAHNEMAN, D.: *Pensar Rápido, Pensar Despacio*, Debate, 2012.
- KATYAL, S.K.: "Democracy & Distrust in an Era of Artificial Intelligence", *Daedalus*, vol. 151, n.º. 2, 2022, p. 322–34. https://doi.org/10.1162/DAED_a_01919.
- KLEINBERG, J., LAKKARAJU, H., LESKOVEC, J., LUDWIG, J. & MULLAINATHAN, S.: "Human Decisions and Machine Predictions", *The Quarterly Journal of Economics*, vol. 133, n.º. 1, 2018, p. 237–93. <https://doi.org/10.1093/qje/qjx032>.
- KNIGHT, J.: "Are Empiricists Asking the Right Questions about Judicial Decisionmaking?", *Duke Law Journal*, vol. 58, n.º. 7, 2009, p. 1531–56.
- LAPORTA SAN MIGUEL, F.J.: "La Creación Judicial y El Concepto de Derecho Implícito", *Revista Jurídica Universidad Autónoma de Madrid*, n.º 6, 2002, p. 133–52.

- LEITER, B.: "Realismo Jurídico Estadounidense", en FABRA ZAMORA, J.L. & NÚÑEZ VAQUERO, Á., *Enciclopedia de Filosofía y Teoría Del Derecho*, Vol. I, México, UNAM, Instituto de Investigaciones Jurídicas, 2015, p. 241–76.
- MEHOZAY, Y. & FISHER, E.: "The Epistemology of Algorithmic Risk Assessment and the Path towards a Non-Penology Penology", *Punishment and Society*, vol. 21, n.º. 5, 2019, p. 523–41. <https://doi.org/10.1177/1462474518802336>.
- MILES, T.J. & SUNSTEIN, C.R.: "The New Legal Realism", *University of Chicago Law Review* 75, n.º. 2, 2008, p. 831–52. <https://doi.org/10.1017/cbo9781107762336>.
- MILLER, T.: "Explanation in Artificial Intelligence: Insights from the Social Sciences", *Artificial Intelligence*, vol. 267, 2019, p. 1–38. <https://doi.org/10.1016/j.artint.2018.07.007>.
- MONAHAN, J. & SKEEM, J.L.: "Risk Assessment in Criminal Sentencing", *Annual Review of Clinical Psychology*, vol. 12, n.º. 1, 2015, p. 489–513.
- MUÑOZ ARANGUREN, A.: "La Influencia de Los Sesgos Cognitivos En Las Decisiones Jurisdiccionales: El Factor Humano. Una Aproximación", *InDret*, n.º. 2, 2011, p. 1–39.
- NIETO, A. & FERNÁNDEZ, T.-R.: *El Derecho y El Revés. Diálogo Epistolar Sobre Leyes, Abogados y Jueces*, Ariel, 2006.
- OSWALD, M., GRACE, J., URWIN, S. & BARNES, G.C.: "Algorithmic Risk Assessment Policing Models: Lessons from the Durham HART Model and 'Experimental' Proportionality", *Information and Communications Technology Law*, vol. 27, n.º. 2, 2018, p. 223–50. <https://doi.org/10.1080/13600834.2018.1458455>.
- PAPILLON, K.: "The Court's Brain: Neuroscience and Judicial Decision Making in Criminal Sentencing", *Court Review*, vol. 48, n.º. 1, 2013, p. 48–63.
- POSNER, R.A.: "What Do Judges and Justices Maximize? (The Same Thing Everybody Else Does)", *Supreme Court Economic Review*, vol. 3, 1993, p. 1–41. <https://doi.org/10.1086/scer.3.1147064>.
- RACHLINSKI, J.J., JOHNSON, S.L., WISTRICH, A.J. & GUTHRIE, C.: "Does Unconscious Racial Bias Affect Trial Judges?", *Notre Dame Law Review*, vol. 84, n.º. 3, 2009, p. 1195–1246.
- RAU, A.S.: "Integrity in Private Judging", *South Texas Law Review*, vol. 38, n.º. 2, 1997, p. 485–540.
- RICHENSON, J.A., & TRAWALTER, S.: "Why Do Interracial Interactions Impair Executive Function? A Resource Depletion Account", *Journal of Personality and Social Psychology*, vol. 88, no. 6, 2005, p. 934–47. <https://doi.org/10.1037/0022-3514.88.6.934>.
- RISSLAND, E.L., ASHLEY, K.D. & LOUI, R.P.: "AI and Law: A Fruitful Synergy", *Artificial Intelligence*, vol. 150, 2003, p. 1–15. [https://doi.org/10.1016/S0004-3702\(03\)00122-X](https://doi.org/10.1016/S0004-3702(03)00122-X).
- RONQUILLO, J., DENSON, T.F., LICKEL, B., LU, Z.L., NANDY, A. & MADDUX, K.B.: "The Effects of Skin Tone on Race-Related Amygdala Activity: An fMRI Investigation", *Social Cognitive and Affective Neuroscience*, vol. 2, n.º. 1, 2007, p. 39–44. <https://doi.org/10.1093/scan/nsi043>.
- RUDIN, C.: "Please Stop Explaining Black Box Models for High Stakes Decisions", *Proceedings of NeurIPS 2018 Workshop on Critiquing and Correcting Trends in Machine Learning*, 2018, p. 1–15.
- SCHULTZ, T.: "Arbitral Decision-Making: Legal Realism and Law & Economics", *Journal of International Dispute Settlement* 6, n.º. 2, 2015, p. 231–51.

- SEARLE, J.R.: "Minds, Brains, and Programs", *Behavioural and Brain Sciences*, vol. 3, no. 3, 1980, p. 417–57.
- SOURDIN, T.: "Judge v Robot? Artificial Intelligence and Judicial Decision-Making", *UNSW Law Journal*, vol. 41, n^o. 4, 2018, p. 1114–33.
- SOURDIN, T.: *Judges, Technology and Artificial Intelligence. The Artificial Judge*, Edward Elgar Publishing, 2021.
- SPAIN BRADLEY, A.: "The Disruptive Neuroscience of Judicial Choice", *UC Irvine Law Review*, vol. 9, n^o. 1, 2018, p. 1–52.
- SUNSTEIN, C.R.: "Algorithms, Correcting Biases", *Social Research*, vol. 86, n^o. 2, 2019, p. 499–511.
- SUNSTEIN, C.R.: "Governing By Algorithm? No Noise and (Potentially) Less Bias", *Duke Law Journal*, vol. 71, n^o. 6, 2022, p. 1175–1205. <https://doi.org/10.2139/ssrn.3925240>.
- SURDEN, H.: "Machine Learning and Law", *Washington Law Review*, vol. 89, n^o. 1, 2014, p. 87–116.
- SUSSKIND, R.: *Online Courts and the Future of Justice*, Oxford University Press, 2019.
- TARUFFO, M.: "Judicial Decisions and Artificial Intelligence", *Artificial Intelligence and Law*, vol. 6, 1998, p. 311–24. https://doi.org/10.1007/978-94-015-9010-5_7.
- VARONA, D., LIZAMA-MUE, Y. & SUÁREZ, J.L.: "Machine Learning's Limitations in Avoiding Automation of Bias", *AI and Society*, vol. 36, no. 1, 2021, p. 197–203. <https://doi.org/10.1007/s00146-020-00996-y>.
- VOLOKH, E.: "Chief Justice Robots", *Duke Law Journal*, vol. 68, n^o. 6, 2019, p. 1136–92.
- WASHINGTON, A.L.: "How to Argue with an Algorithm: Lessons from the COMPAS- ProPublica Debate", *The Colorado Technology Law Journal*, vol. 17, n^o. 1, 2019, p. 1–37.
- WHEELER, M.E. & FISKE, S.T.: "Controlling Racial Prejudice: Social-Cognitive Goals Affect Amygdala and Stereotype Activation", *Psychological Science*, vol. 16, n^o. 1, 2005, p. 56–63. <https://doi.org/10.1111/j.0956-7976.2005.00780.x>.
- WINTER, C.K.: "The Challenges of Artificial Judicial Decision-Making for Liberal Democracy", en BYSTRANOWSKI, P., JANIK, B., AND PRÓCHNICKI, M. (eds.), *Judicial Decision-Making: Integrating Empirical and Theoretical Perspectives*, Springer, 2022, p. 179–204.
- WISSER, L.: "Pandora's Algorithmic Black Box: The Challenges of Using Algorithmic Risk Assessments in Sentencing", *American Criminal Law Review*, vol. 56, 2019, p. 1811–32.
- WISTRICH, A.J. & RACHLINSKI, J.J.: "Implicit Bias in Judicial Decision Making. How It Affects Judgment and What Judges Can Do About It", *Enhancing Justice*, American Bar Association, 2017, p. 87–130.
- WISTRICH, A.J., RACHLINSKI, J.J. & GUTHRIE, C.: "Heart versus Head: Do Judges Follow the Law or Follow Their Feelings?", *Texas Law Review*, vol. 93, 2015.
- WOODRUFF, A., ANDERSON, Y.A., ARMSTRONG, K.J., GKIZA, M., JENNINGS J., MOESSNER, C., VIEGAS, F. ET AL.: "A Cold, Technical Decision-Maker': Can AI Provide Explainability, Negotiability, and Humanity?", 2020, 1–23.
- XU, Z.: "Human Judges in the Era of Artificial Intelligence: Challenges and Opportunities", *Applied Artificial Intelligence*, vol. 36, n^o. 1, 2022, p. 1025–45. <https://doi.org/10.1080/08839514.2021.2013652>.

XU, Z., ZHAO, Y. & DENG, Z.: “The Possibilities and Limits of AI in Chinese Judicial Judgment”, *AI and Society*, 2021. <https://doi.org/10.1007/s00146-021-01250-9>.

YALCIN, G., THEMELI, E., STAMHUIS, E., PHILIPSEN, S. & PUNTONI, S.: “Perceptions of Justice By Algorithms”, *Artificial Intelligence and Law*, 2022. <https://doi.org/10.1007/s10506-022-09312-z>.