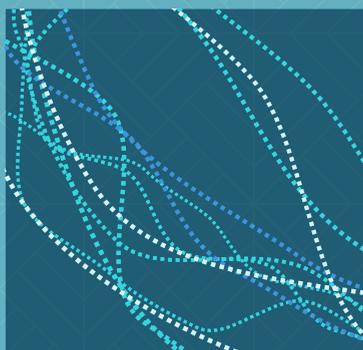


DERECHO PENAL, INTELIGENCIA
ARTIFICIAL Y NEUROCIENCIAS
DIRITTO PENALE, INTELLIGENZA
ARTIFICIALE E NEUROSCIENZE



Jaime Miguel Peris Riera, Antonella Massaro

41

Collana

L'Unità del Diritto



Roma TiE-Press
2023



Università degli Studi Roma Tre
Dipartimento di Giurisprudenza

NELLA STESSA COLLANA

1. P. CARNEVALE (a cura di), *La Costituzione riscritta. Saggi sulla revisione costituzionale in itinere*, 2016
2. E. BATTELLI, B. CORTESE, A. GEMMA, A. MASSARO (a cura di), *Patrimonio culturale: profili giuridici e tecniche di tutela*, 2017
3. R. BENIGNI (a cura di), *Libertà religiosa, diritti umani e globalizzazione*, 2017
4. A. MASSARO (a cura di), *La tutela della salute nei luoghi di detenzione. Un'indagine di diritto penale intorno a carcere, REMS e CPR*, 2017, 2017
5. V. ZENO-ZENCOVICH, *Comparative Legal Systems. A Short Introduction*, 2017 (I ed.) *Comparative Legal Systems. A Short and Illustrated Introduction*, 2019 (II ed.)
6. M. GAMBACCIANI, *Impresa e lavoro: il nuovo bilanciamento nei licenziamenti collettivi*, 2017
7. A. MASSARO, M. SINISI (a cura di), *Trasparenza nella P.A. e norme anticorruzione: dalla prevenzione alla repressione*, 2017
8. A. D. DE SANTIS (a cura di), *I profili processuali della nuova disciplina sulla responsabilità sanitaria*, 2017
9. V. VITI, *La locazione finanziaria tra tipicità legale e sottotipi*, 2018
10. C. CARDIA, R. BENIGNI (a cura di), *50 Anni dalla Populorum Progressio. Paolo VI: il Papa della modernità. Giustizia tra i popoli e l'amore per l'Italia*, 2018
11. G. NUZZO, *L'abuso del diritto di voto nel concordato preventivo. Interessi protetti e regola di correttezza*, 2018
12. G. GRISI, C. SALVI (a cura di), *A proposito del diritto post-moderno. Atti del Seminario di Leonessa, 22-23 settembre 2017*, 2018
13. G. MAESTRI, *L'ordinamento costituzionale italiano alla prova della democrazia paritaria*, 2018
14. G. CONTE, A. FUSARO, A. SOMMA, V. ZENO-ZENCOVICH (a cura di), *Dialoghi con Guido Alpa. Un volume offerto in occasione del suo LXXI compleanno*, 2018
15. E. PODOGGHE, *Comunicazione e "dignità della donna". Uno studio di genere*, 2018
16. G. GRISI (a cura di), *L'abuso del diritto. In ricordo di Davide Messinetti*, 2019
17. S. ANASTASIA, P. GONNELLA (a cura di), *I paradossi del diritto. Saggi in omaggio a Eligio Resta*, 2019
18. S. DEL GATTO, *Poteri pubblici, iniziativa economica e imprese*, 2019
19. R. BENIGNI, B. CORTESE (a cura di), *La "giurisdizione". Una riflessione storico-giuridica*, 2019

20. M. RUOTOLO (a cura di), *La Costituzione...aperta a tutti*, 2019 (I ed.); M. RUOTOLO, M. CAREDDA (a cura di), *La Costituzione...aperta a tutti*, 2020 (II ed.); M. RUOTOLO, M. CAREDDA (a cura di), *La Costituzione...aperta a tutti*, 2021 (III ed.); M. RUOTOLO, M. CAREDDA (a cura di), *La Costituzione...aperta a tutti*, 2022 (IV ed.)
21. N. POSTERARO, M. SINISI (a cura di), *Questioni di fine vita*, 2020
22. G. RESTA (a cura di), *L'Armonia nel diritto. Contributi a una riflessione su diritto e musica*, 2020
23. F. LATTANZI (a cura di), *Genocidio. Conoscere e ricordare per prevenire*, 2020
24. E. TOTI, *Diritto cinese dei contratti e sistema giuridico romanistico tra legge e dottrina*, 2020
25. B. CORTESE, *La tutela in caso di vizio della res empta e della res locata: inadempimento rispondenza ex fide bona*, 2020
26. M.S. BONOMI, *La motivazione dell'atto amministrativo: dalla disciplina generale alle regole speciali*, 2020
27. G. GRISI, F. RASSU (a cura di), *Perspectives nouvelles du droit. Thèmes, méthodes et historiographie en France et en Italie*, 2020
28. F. APERIO BELLA, A. CARBONE, E. ZAMPETTI (a cura di), *Dialoghi di Diritto Amministrativo. Lavori del Laboratorio di Diritto Amministrativo 2019*, 2020
29. S. CALDARELLI, *I vincoli al bilancio dello Stato e delle pubbliche amministrazioni. Sovranità, autonomia e giurisdizione*, 2020
30. A. MASSARO (a cura di), *Connessioni di Diritto Penale*, 2020
31. R. BENIGNI (a cura di), *Diritto e religione in Italia*, 2021
32. E. CALZOLAIO, R. TORINO, L. VAGNI (a cura di), *Liber amicorum Luigi Moccia*, 2021
33. A. CARRATTA (a cura di), *Limiti esterni di giurisdizione e diritto europeo. A proposito di Cass. Sez. Un. n. 19598/2020*, 2021
34. M. CATENACCI, R. RAMPIONI, V.N. D'ASCOLA (a cura di), *Studi in onore di Antonio Fiorella*, 2021
35. F. GRASSI, O. HAGI KASSIM (a cura di), *Vecchie e nuove certezze nel diritto amministrativo. Elementi essenziali e metodo gradualista. Dibattito sugli scritti di Giampaolo Rossi*, 2021
36. B. CORTESE (a cura di), *Il diritto dei beni culturali. Atti del Convegno OGIPAC in memoria di Paolo Giorgio Ferri. Roma 27 maggio 2021*, 2021
37. V. MANNINO, *Governo misto romano e repubblicanesimo americano*, 2022
38. V. BONTEMPI (a cura di), *Lo Stato digitale nel piano di ripresa e resilienza*, 2022
39. E. BATTELLI, M. LOTTINI, G. SPOTO, E.M. INCUTTI, (a cura di), *Nuovi orizzonti sulla tutela degli animali*, 2022
40. F. APERIO BELLA, A. CARBONE, E. ZAMPETTI (a cura di), *Dialoghi di diritto amministrativo. Lavori del laboratorio di diritto Amministrativo 2020-2021*, 2022



Università degli Studi Roma Tre
Dipartimento di Giurisprudenza

Jaime Miguel Peris Riera, Antonella Massaro

**DERECHO PENAL, INTELIGENCIA
ARTIFICIAL Y NEUROCIENCIAS**
**DIRITTO PENALE, INTELLIGENZA
ARTIFICIALE E NEUROSCIENZE**

41

L'Unità del Diritto
Collana del Dipartimento di Giurisprudenza



Roma TrE-Press
2023

La Collana *L'unità del diritto* è stata varata su iniziativa dei docenti del Dipartimento di Giurisprudenza. Con questa Collana si intende condividere e sostenere scientificamente il progetto editoriale di Roma TrE-Press, che si propone di promuovere la cultura giuridica incentivando la ricerca e diffondendo la conoscenza mediante l'uso del formato digitale ad accesso aperto.

Comitato scientifico della Collana:

Paolo Alvazzi Del Frate, Roberto Baratta, Concetta Brescia Morra, Paolo Carnevale, Antonio Carratta, Mauro Catenacci, Alfonso Celotto, Carlo Colapietro, Emanuele Conte, Tommaso Dalla Massara, Carlo Fantappiè, Elena Granaglia, Giuseppe Grisi, Andrea Guaccero, Luca Luparia Donati, Francesco Macario, Luca Marafioti, Enrico Mezzetti, Giulio Napolitano, Giuseppe Palmisano, Annalisa Pessi, Giorgio Pino, Alberto Franco Pozzolo, Giampiero Proia, Giorgio Resta, Francesco Rimoli, Giuseppe Ruffini, Marco Ruotolo, Maria Alessandra Sandulli, Chris Thomale, Giuseppe Tinelli, Luisa Torchia, Mario Trapani, Vincenzo Zeno-Zencovich, Andrea Zoppini.

Collana pubblicata nel rispetto del Codice etico adottato dal Dipartimento di Giurisprudenza dell'Università degli Studi Roma Tre, in data 22 aprile 2020.

Il volume pubblicato è stato sottoposto a previa e positiva valutazione nella modalità di referaggio *double-blind peer review*.

Coordinamento editoriale:

Gruppo di Lavoro *Roma TrE-Press*

Elaborazione grafica della copertina: **MOSQUITO**, mosquitoroma.it

Caratteri tipografici utilizzati:

American Typewriter condensed, Brandon Grottesque, Gotham rounded (copertina e frontespizio)

Adobe Garamond Pro (testo)

Impaginazione e cura editoriale: Colitti-Roma colitti.it

Edizioni: Roma TrE-Press ©

Roma, febbraio 2023

ISBN: 979-12-5977-149-0

<http://romatrepress.uniroma3.it>

Quest'opera è assoggettata alla disciplina *Creative Commons attribution 4.0 International License* (CC BY-NC-ND 4.0) che impone l'attribuzione della paternità dell'opera, proibisce di alterarla, trasformarla o usarla per produrre un'altra opera, e ne esclude l'uso per ricavarne un profitto commerciale.



L'attività della *Roma TrE-Press* è svolta nell'ambito della

Fondazione Roma Tre-Education, piazza della Repubblica 10, 00185 Roma

Collana del Dipartimento di Giurisprudenza

L'unità del diritto

La Collana di studi giuridici promossa dal Dipartimento di Giurisprudenza dell'Università Roma Tre assume un titolo – quello de *L'unità del diritto* – che può apparire particolarmente impegnativo, perché il fenomeno giuridico riflette la complessità delle società che il diritto mira a regolare, si sviluppa intorno ad una molteplicità di articolazioni e sembra pertanto sfuggire ad una definizione in termini di unità. Anche la scienza del diritto, intesa come riflessione intorno al diritto, come forma di conoscenza che assume il diritto ad oggetto diretto e immediato di indagine, sia nella prospettiva teorica sia in quella storico-positiva relativa ad un singolo ordinamento, soffre a trovare una sua dimensione unitaria. La riflessione intorno al diritto, da qualunque punto di partenza si intenda affrontarla, ammette una pluralità di opzioni metodologiche, contempla una molteplicità di giudizi di valore, si caratterizza inevitabilmente per una pluralità di soluzioni interpretative. L'unico, generalissimo, elemento che sembra contraddistinguerla in senso unitario è dato dal suo essere rivolta alla conoscenza del diritto, dal suo carattere conoscitivo dell'esperienza giuridica complessivamente intesa, una unità, potrebbe dirsi, figlia della diversità e varietà delle scelte di metodo e del pluralismo interpretativo, ma pur sempre una unità quanto meno in questo suo nucleo irriducibile. Ed è allora questo il senso da attribuire al titolo della collana che prende l'avvio, ossia quello di dare ospitalità a contributi di studiosi diversi per formazione e interessi, ma che si riconoscono tutti nella comune esigenza di indagare il fenomeno giuridico applicando con rigore il metodo prescelto, nella consapevolezza della condivisione di un patrimonio formativo e culturale idoneo a creare un'adeguata coscienza di sé e sulla cui base costruire l'impegno scientifico del giurista.

In questa prospettiva, la Collana si ripromette di ospitare non solo contributi scientifici di tipo monografico, raccolte di scritti collettanei, atti di convegni e seminari, ma anche materiali didattici che possano proficuamente essere utilizzati nella formazione dei giovani giuristi.

La Collana entra a far parte della struttura di Roma TrE-Press, che, affiancando alla tradizionale pubblicazione in formato cartaceo quella in formato digitale in *open access*, contribuisce ad una nuova e più ampia diffusione del sapere giuridico.

Prof. Antonio Carratta
Direttore del Dipartimento di Giurisprudenza
Università Roma Tre

Índice

Prólogo / Premessa

15

SECCIÓN I

EL DERECHO PENAL 'TRADICIONAL' ANTE LOS RETOS
DE LA INTELIGENCIA ARTIFICIAL Y DE LAS NEUROCIENCIAS

IL DIRITTO PENALE 'TRADIZIONALE' DI FRONTE ALLE SFIDE
DELL'INTELLIGENZA ARTIFICIALE E DELLE NEUROSCIENZE

CAPÍTULO I

Inteligencia artificial y neurociencias: avances del Derecho penal contemporáneo

Jaime Miguel Peris Riera

1. <i>Introducción</i>	19
2. <i>Las bases antropológicas como punto de partida del avance</i>	21
3. <i>Neurociencia e inteligencia artificial en la teoría del delito</i>	22
4. <i>La peligrosidad de la inteligencia artificial en la aplicación de la norma</i>	23
5. <i>Bibliografía</i>	26

CAPÍTULO II

Intelligenza artificiale e neuroscienze: l'eterno ritorno del diritto penale?

Antonella Massaro

1. <i>Diritto penale e nuove tecnologie: rivoluzione copernicana o eterno ritorno?</i>	27
2. <i>Neuroscienze e diritto penale nella dialettica tra normativismo e ontologismo</i>	28
2.1 <i>La natura strutturalmente normativa del diritto penale e della teoria generale del reato</i>	32
2.1.1. <i>Le neuroscienze e la questione del libero arbitrio</i>	33
2.1.2. <i>Neuroscienze, diritto e processo penale</i>	35
2.1.3. <i>Neuroscienze e diritto penitenziario</i>	37
3. <i>Le relazioni tra intelligenza artificiale e diritto penale</i>	38
3.1 <i>a) l'algoritmo delinquente</i>	39
3.2 <i>b) l'algoritmo investigante e/o inquirente</i>	41
3.3 <i>c) l'algoritmo consulente e/o giudicante</i>	42
4. <i>Corsi, ricorsi e deviazioni nella parabola del diritto calcolabile</i>	46
5. <i>Bibliografía</i>	52

SECCIÓN II

LA INTELIGENCIA ARTIFICIAL ENTRE DERECHO Y PROCESO PENAL

L'INTELLIGENZA ARTIFICIALE TRA DIRITTO E PROCESSO PENALE

CAPÍTULO III

Implicaciones de la inteligencia artificial en el ámbito del Derecho Penal

David Lorenzo Morillas Fernández

1. <i>Enfoque y estado de la cuestión</i>	59
2. <i>Situación actual de la responsabilidad penal por un hecho delictivo cometido por inteligencia artificial</i>	73
2.1. <i>¿Puede considerarse a la inteligencia artificial autora única del delito?</i>	74
2.2. <i>¿Quién debe responder penalmente entonces</i>	79
3. <i>Pronunciamientos jurisprudenciales expresos</i>	83
4. <i>Conclusiones</i>	86
5. <i>Bibliografía</i>	89

CAPÍTULO IV

I miti d'oggi: la certezza del diritto fra utopie e distopie

Lorenza Grossi

1. <i>I 'problemi antichi': la dialettica tra disposizione e norma come costante irrinunciabile del diritto</i>	93
2. <i>Il mito della 'certezza del diritto' e il demone dell'interpretazione giudiziaria nell'Illuminismo giuridico</i>	98
3. <i>La (in)certezza del diritto e la crisi della legalità</i>	100
4. <i>L'intelligenza artificiale: il diritto calcolato</i>	107
5. <i>La bellezza del mito: demistificabile, ma non 'demitificabile'</i>	110
6. <i>Bibliografía</i>	113

CAPÍTULO V

Deshumanizando la culpabilidad: los sistemas inanimados en la teoría del delito

Angelo Giraldi

1. <i>Consideraciones introductorias</i>	119
2. <i>Aproximaciones a la vulnerabilidad del concepto de culpabilidad</i>	124
3. <i>La incidencia de los sistemas de inteligencia artificial en la culpabilidad como presupuesto de la punibilidad</i>	128
3.1 <i>El tentativo de reconocer una nueva subjetividad pasiva: la inteligencia artificial como víctima de delitos</i>	131

3.2 <i>Hacia una subjetividad activa instrumental: la inteligencia artificial como autora de delitos</i>	140
3.3 <i>La inteligencia artificial como instrumento de apoyo para el ejercicio de la justicia</i>	148
4. <i>La incidencia de los sistemas de inteligencia artificial en la culpabilidad como criterio para la determinación de la pena en concreto</i>	149
5. <i>Conclusiones</i>	151
6. <i>Bibliografía</i>	152

CAPÍTULO VI

Algoritmos y big data en la responsabilidad penal: el reto de la cibercriminalidad en el Derecho Penal

Jacinto Pérez-Arias

1. <i>Introducción</i>	159
2. <i>La informática y el Derecho Penal: la inteligencia artificial</i>	161
3. <i>Límites normativos</i>	167
4. <i>Problemas procesales ¿vs? problemas penales</i>	171
5. <i>Sujeto activo, resultado y culpabilidad: juicio de probabilidad estadística</i>	175
6. <i>Conclusión</i>	184
7. <i>Bibliografía</i>	186

CAPÍTULO VII

'Algoritmos predictivos' y justicia penal desde una perspectiva italiana y europea

Laura Notaro

1. <i>Introducción</i>	191
2. <i>Algoritmos y 'justicia predictiva'. Experiencias y tendencias recientes</i>	192
2.1 <i>Herramientas de valoración del riesgo de reincidencia en Estados Unidos</i>	193
2.1.1. <i>Las decisiones sobre el 'pre-trial release' y el debate sobre el 'money bail'</i>	194
2.1.2. <i>El 'sentencing' y el caso 'Loomis'</i>	195
2.2 <i>Herramientas de valoración del riesgo de reincidencia en Europa</i>	197
2.3 <i>Herramientas de 'predicción' del resultado decisorio en el ámbito civil. El caso francés</i>	198
3. <i>Riesgos de la utilización de algoritmos 'predictivos' en el sistema de justicia penal</i>	199
3.1 <i>Riesgos de las herramientas de valoración del riesgo de reincidencia</i>	199
3.1.1. <i>Aspectos críticos conectados con la naturaleza estadístico-actuarial de la base científica</i>	200
3.1.2. <i>Aspectos críticos conectados con el carácter 'computacional' de la herramienta</i>	202
3.2 <i>Riesgos de las herramientas de 'predicción' del resultado decisorio</i>	203
4. <i>Límites jurídicos a la utilización de algoritmos 'predictivos' en el sistema de justicia penal</i>	204
4.1 <i>Instrumentos de nivel internacional. La Carta Ética del CEPEJ del Consejo de Europa</i>	205

4.2 Derecho de la Unión Europea: límites a las «decisiones basadas únicamente en el tratamiento automatizado»	206
4.3 Límites del derecho italiano: la prohibición de pericia criminológica	209
5. Conclusiones	210
6. Bibliografía	211

CAPÍTULO VIII

Proceso penal e inteligencia artificial: prevenciones en torno a su futura aplicación en la fase de juicio oral

Rafael Castillo Felipe, Salvador Tomás Tomás

1. La actual imposibilidad técnica y jurídica de un 'juez-robot' como punto de partida necesario	215
2. Cuestiones problemáticas que podría suscitar la aplicación de la IA en la fase de juicio oral	220
2.1 El uso de la IA para la obtención de información sobre los hechos que fundan la acusación y la defensa	220
2.2 El recurso a la IA como prueba para determinar la fiabilidad o credibilidad de los medios de prueba practicados en el juicio oral	223
2.3 El empleo de la IA para decidir sobre la ilicitud de las diligencias de investigación o medios de prueba practicados	225
2.4 El uso de la IA en el juicio de apreciación de la prueba	226
2.5 La aplicación de la IA en la elaboración de la fundamentación jurídica de la sentencia y en la resolución de recursos posteriores frente a esta	232
3. La cuestión de la transparencia del algoritmo empleado en el enjuiciamiento penal	236
4. Bibliografía	241

SECCIÓN III

NEUROCIENCIAS Y RESPONSABILIDAD PENAL

NEUROSCIENZE E RESPONSABILITÀ PENALE

CAPÍTULO IX

Inteligencia artificial y neurociencia.

Algunas reflexiones sobre las aportaciones que pueden hacer al Derecho Penal

José Manuel Palma Herrera

1. Relación entre Derecho penal, inteligencia artificial y neurociencia. Planteamiento de la cuestión	249
---	-----

2. Escenario 1: incidencia de los estudios neurocientíficos en la libertad como fundamento de la culpabilidad	255
3. Escenario 2: creación de herramientas de inteligencia artificial capaces de interactuar con el propio individuo condicionando su comportamiento	263
4. Escenario 3: prevención, enjuiciamiento y tratamiento del delito	265
5. Bibliografía	268

CAPÍTULO X
Imputabilidad y delito
Mario Trapani

1. Imputabilidad y teoría del delito	271
2. El “contenido” de la imputabilidad	276
3. Imputabilidad e inteligencia artificial	280
4. Bibliografía	283

CAPÍTULO XI
Neurociencias y culpabilidad: recapitulando
María Isabel González Tapia

1. Introducción	285
1.1 Presupuestos dogmáticos y epistémicos	285
1.2 Neuroderecho y Derecho Penal	290
2. Culpabilidad y libre albedrío	292
2.1 «El libre albedrío es la biología que no se ha descubierto todavía»	292
2.2 Para el Derecho Penal, ‘la neurociencia lo cambia todo y no cambia nada’	296
3. Imputabilidad, capacidad de control e impulsividad	301
3.1 Genética y agresividad impulsivo-reactiva: la interacción entre el gen MAOA-L y el maltrato severo en la infancia	302
3.2 Biomarcadores del delincuente violento	304
4. Psicopatía: ¿debería integrarse la falta de empatía emocional dentro del concepto de imputabilidad?	308
5. Bibliografía	311
<i>Autores / Autori</i>	315

Prólogo

El intercambio y la colaboración entre la Universidad de Murcia y la Università degli Studi 'Roma Tre', en el ámbito específico del Derecho Penal, se viene desarrollando de forma casi ininterrumpida desde hace más de una década, a partir del Convenio para la obtención de la doble titulación suscrito por ambas Universidades, por primera vez, en 2010. Se trata, sin duda, de una experiencia a través de la cual se ha reforzado esa vocación de internacionalización que es uno de los rasgos más característicos y reconocibles del Departamento de Giurisprudencia de la Universidad 'Roma Tre'.

A estas alturas, muchos estudiantes y doctorandos han cruzado el puente ideal entre Murcia y Roma, enriqueciendo su perspectiva de estudio e investigación mediante la valiosa aportación de la comparación jurídica y, más en general, del diálogo entre dos sistemas jurídicos que, aunque difieran en muchos aspectos, se asientan sobre bases culturales comunes.

Algo que se vio acompañado, a lo largo de los años, por una fructífera interacción entre los profesores de Derecho penal de las dos Facultades de Derecho, que, bajo la dirección del Prof. Jaime Miguel Peris Riera y del Prof. Mario Trapani, se han dedicado a organizar y celebrar encuentros de estudio, seminarios de profundización

Premessa

Lo scambio e la collaborazione tra l'Università di Murcia e l'Università degli Studi 'Roma Tre', sullo specifico versante del diritto penale, continuano pressoché ininterrotti da oltre un decennio, a partire dalla Convenzione per il rilascio del doppio titolo sottoscritta dai due Atenei, per la prima volta, nel 2010. Si tratta, senza dubbio, di un'esperienza attraverso la quale si è rafforzata quella vocazione all'internazionalizzazione che rappresenta una delle cifre più caratteristiche e riconoscibili del Dipartimento di Giurisprudenza dell'Università degli Studi 'Roma Tre'.

Sono ormai tanti gli studenti e i dottorandi di ricerca che hanno attraversato l'ideale ponte tra Murcia e Roma, arricchendo la propria prospettiva di studio e di ricerca attraverso il prezioso apporto della comparazione giuridica e, più in generale, del dialogo tra due sistemi che, sia pur per molti aspetti differenti, poggiano su basi culturali comuni.

A ciò si è aggiunta, nel corso degli anni, una feconda interazione tra i docenti di diritto penale dei due Dipartimenti di Giurisprudenza, che, sotto la guida del Prof. Jaime Miguel Peris Riera e del Prof. Mario Trapani, si sono impegnati nell'organizzazione e nella realizzazione di incontri di studio, seminari di approfondimento, convegni internazionali.

y congresos internacionales.

En concreto, el 5 de mayo de 2022 tuvo lugar en la Universidad de Murcia el Congreso Internacional 'Inteligencia artificial y neurociencias: avances del Derecho Penal contemporáneo', en colaboración con la Universidad de 'Roma Tre' y patrocinado por el Parlamento Europeo. El intenso diálogo iniciado en aquella ocasión convenció a los directores científicos del Congreso para imaginar su continuación ideal, lo que dio lugar a la realización de este volumen.

La inteligencia artificial y la neurociencia, de hecho, representan una valiosa oportunidad para cuestionar el nuevo rumbo que potencialmente imprimen las nuevas tecnologías al Derecho penal, pero, sobre todo, para profundizar en la reflexión, nunca definitiva, sobre aquellos principios y conceptos 'tradicionales' que identifican una constante, por el momento irrenunciable, de la materia penal.

Murcia-Roma, 9 de diciembre de 2022

Jaime Miguel Peris Riera y
Antonella Massaro

Il 5 maggio 2022, in particolare, si è svolto il Congresso internazionale 'Intelligenza artificiale e neuroscienze: l'evoluzione del Diritto penale contemporaneo', organizzato presso l'Università di Murcia in collaborazione con l'Università degli Studi 'Roma Tre' e sotto l'alto patrocinio del Parlamento europeo. L'intenso dialogo avviato in quell'occasione ha convinto i direttori scientifici del Convegno a immaginare una sua prosecuzione ideale, che ha condotto alla realizzazione di questo volume.

L'intelligenza artificiale e le neuroscienze, in effetti, rappresentano una preziosa occasione per interrogarsi sul nuovo corso potenzialmente impresso dalle nuove tecnologie al diritto penale, ma, soprattutto, per approfondire la riflessione, mai definitiva, su quei principi e concetti 'tradizionali' che individuano una costante, al momento irrinunciabile, della materia penale.

Murcia-Roma, 9 dicembre 2022

Jaime Miguel Peris Riera e
Antonella Massaro

SECCIÓN I

EL DERECHO PENAL 'TRADICIONAL' ANTE LOS RETOS
DE LA INTELIGENCIA ARTIFICIAL Y DE LAS NEUROCIENCIAS

IL DIRITTO PENALE 'TRADIZIONALE' DI FRONTE ALLE SFIDE
DELL'INTELLIGENZA ARTIFICIALE E DELLE NEUROSCIENZE

CAPÍTULO I

Inteligencia artificial y neurociencias: avances del Derecho penal contemporáneo

Jaime Miguel Peris Riera

SUMARIO: 1. Introducción – 2. Las bases antropológicas como punto de partida del avance – 3. Neurociencia e inteligencia artificial en la teoría del delito – 4. La peligrosidad de la inteligencia artificial en la aplicación de la norma – 5. Bibliografía.

1. *Introducción*

La inteligencia artificial y las neurociencias han modificado por completo la sociedad en la que vivimos. Nos ofrecen más vida, y sobre todo más calidad de vida. Es un fenómeno imparable, que está abriendo la puerta a avances sociales, culturales, económicos y científicos que evolucionan cada día con progresión geométrica. Avances, todos ellos, que no hace muchas décadas se hubieran considerado *auténtica ciencia-ficción* (D.L. MORILLAS FERNÁNDEZ¹).

Por ello, no puede extrañar que al jurista le apasione, y al mismo tiempo preocupe, este nuevo fenómeno tecnológico social. Se señala, con gran acierto, que *la sociedad del siglo XXI es la sociedad del ocio y del consumo, pero es también cada vez más, una sociedad que se siente amenazada por la aparición de nuevas fuentes de peligro asociadas, paradójicamente, a ese progreso y bienestar* (J.M. PALMA HERRERA²).

No es nuestra misión entrar en lo estrictamente científico técnico de esta materia. Eso compete a los especialistas del ramo quienes, no solo podrán explicar en qué consiste cada ‘avance’, sino también qué riesgos

¹ D.L. MORILLAS FERNÁNDEZ, *Implicaciones de la inteligencia artificial en el ámbito del Derecho Penal*, en esta obra.

² J.M. PALMA HERRERA, *Inteligencia artificial y neurociencia. Algunas reflexiones sobre las aportaciones que pueden hacer al Derecho Penal*, en esta obra.

puede producir cuando interacciona con otros sistemas tecnológicos. Por el contrario, al jurista le debe ocupar el avance científico cuando interacciona con la sociedad, cuando es capaz de generar actos o efectos que tienen relevancia para la existencia social y, por ende, para el Derecho. Como señala D.L. MORILLAS FERNÁNDEZ³, nuestro propósito es que resulte *accesible a juristas en general*.

Pero más allá de estas generalidades, a nosotros nos interesa lo que afecta a una rama muy específica y principal del ordenamiento jurídico: el derecho penal. Son diversos los enfoques que pueden realizarse sobre la inteligencia artificial y la neurociencia en su relación con el derecho penal. La presente obra es un buen ejemplo de ello, como también lo fue el congreso internacional sobre inteligencia artificial y neurociencias (avances del derecho penal contemporáneo) celebrado⁴ en la Facultad de Derecho de la Universidad de Murcia, con participantes no solo especialistas en esta apasionante materia, sino con una visión internacional (no basado exclusivamente en la legislación nacional o local) que aportaron los ponentes de las Universidades de Murcia y Córdoba (de España), y de la Universidad de Roma Tre y la Universidad de Pisa (de Italia).

De las interesantísimas conclusiones que se alcanzaron en aquel congreso, surgió la idea de avanzar en la investigación y plasmarlo en una obra que evidenciara el estado actual de la cuestión, así como los retos futuros del derecho penal. Un derecho penal que, queramos o no, tiene que adaptarse a unos progresos que generan nueva delincuencia (la llamada cibercriminalidad), pero también facilitan nuevas herramientas que nos servirán para medir o evaluar conceptos jurídicos, y que hasta la fecha quedaban a merced de la mente humana y de la formación universitaria en materia jurídica del intérprete. Algo muy parecido a lo que ocurrió, en su día, con las huellas dactilares o el ADN⁵, avances ambos sin los cuales difícilmente podríamos explicar hoy los resultados que se producen en la averiguación del delito y persecución del culpable. De ahí que la *criminología*, y no solo el derecho penal, sea beneficiaria de toda esta nueva tecnología de la que hablamos (en igual sentido, J. PÉREZ-ARIAS⁶).

³ MORILLAS FERNÁNDEZ, *op. ult. cit.*

⁴ Congreso internacional celebrado el 5 de mayo de 2022 y patrocinado por el Parlamento europeo.

⁵ Puede analizarse con detalle en J.M. PERIS RIERA, *La regulación penal de la manipulación genética en España (principios penales fundamentales y tipificación de las genotecnologías)*, Civitas, Madrid, 1995.

⁶ J. PÉREZ-ARIAS, *Algoritmos y big data en la responsabilidad penal: el reto de la cibercriminalidad en el Derecho Penal*, en esta obra.

2. Las bases antropológicas como punto de partida del avance

Pese a todo, somos muchos los que opinamos que las bases antropológicas que puedan avanzar los descubrimientos de la neurociencia no pueden desplazar ni los conceptos centrales del Derecho Penal ni tampoco la actividad autónoma y decisiva del juez. En igual sentido, *il diritto penale non può essere del tutto permeabile alle scienze naturali* (A. MASSARO⁷).

Todo lo más una función de auxilio, auxilio técnico que puede ser también aportado por la inteligencia artificial; pero siempre vertebrada a partir de los principios fundamentales y garantistas del Derecho Penal.

Se trata, en definitiva, de *'situarse', dimensionalmente, en la dicotomía dialéctica entre el normativismo funcionalista y la política-criminal con respecto a la Dogmática, decidiendo si el Derecho es un saber puramente normativo, que se retroalimenta de valores y directrices que proporciona el contexto al que sirve funcionalmente y que genera el propio marco normativo; o se trata de un Derecho que debe mirar también hacia fuera, con un enfoque integrador y crítico, hacia lo que muestran otros saberes, hacia el substrato empírico de lo regulado, sus causas y sus contextos y hacia las consecuencias o impacto de su aplicación práctica* (M.I. GONZÁLEZ TAPIA⁸).

La presente obra tiene a verdaderos especialistas en la materia, que estudian la temática jurídica desde diversas ópticas, bien desde un aspecto sustantivo, de derecho penal estricto, bien desde una óptica de derecho procesal, que es quien, a la postre, debe vertebrar y aplicar este nuevo derecho penal tecnológico.

Desde un punto de vista jurídico penal, la inteligencia artificial y la neurociencia se han convertido en auténticas tropas de avance en la transformación y adaptación del Derecho. No solo por cuanto hace a la cibercriminalidad como fenómeno criminógeno, y que es, posiblemente, el más conocido por cualquier persona; sino por las nuevas concepciones que pueden desarrollarse en torno a conceptos tan jurídicos y tradicionales como lo son, sin duda, la culpabilidad y, dentro de ella, la imputabilidad, por ejemplo. Baste cotejar las investigaciones realizadas por J.M. PALMA HERRERA⁹ y M.I. GONZÁLEZ TAPIA¹⁰ para ver de lo que hablamos.

Desde una perspectiva procesal, se debe ser conscientes del riesgo

⁷ A. MASSARO, *Intelligenza artificiale e neuroscienze: l'eterno ritorno del diritto penale?*, en esta obra.

⁸ M.I. GONZÁLEZ TAPIA, *Neurociencias e imputabilidad: recapitulando*, en esta obra, pp. 1-2.

⁹ PALMA HERRERA, *op. ult. cit.*

¹⁰ GONZÁLEZ TAPIA, *op. ult. cit.*

que puede entrañar cualquier tendencia de corte utilitarista y preñada de pragmatismo, que haga primar un mal entendimiento de la rapidez o agilidad del sistema judicial, sobre las garantías del ciudadano a un proceso seguro y justo. De ahí que se haya afirmado que la inteligencia artificial es *uno de los muchos retos a los que tiene que hacer frente la justicia del siglo XXI* (R. CASTILLO FELIPE y S. TOMÁS TOMÁS¹¹).

3. Neurociencia e inteligencia artificial en la teoría del delito

Con independencia de la perspectiva dogmática que se sustente sobre la imputabilidad (bien como capacidad de culpabilidad, fundamento del reproche, fundamento de la capacidad de motivación o como capacidad de la pena), los avances en neurociencia y en inteligencia artificial pueden ser tenidos en cuenta como elementos que reforzarían –aunque no fueran exclusivamente autónomos y definitivos– la prueba en el proceso penal.

Lo que podría resumirse en la exigencia de una necesaria ‘decisión humana’ en el proceso penal.

Respecto de esta imputabilidad, debemos señalar que la neurociencia está *arrojando luz sobre otras vulnerabilidades inespecíficas que podrían estar en la base de, nada más y nada menos, la delincuencia antisocial o proporcionando un conocimiento más preciso acerca de, por ejemplo, los efectos del maltrato o abuso severo en la infancia* (M.I. GONZÁLEZ TAPIA¹²).

En efecto, resulta sugestiva la pregunta que se plantea J.M. PALMA HERRERA¹³: *¿es capaz la neurociencia, auxiliada por la inteligencia artificial, de destruir la base de la culpabilidad penal?* La responsabilidad penal que pueda nacer de condicionantes basados en inteligencia artificial debería ser individualizada a partir de los criterios generales inherentes a la teoría general del delito. En este contexto, como advierte A. MASSARO¹⁴, *le neuroscienze possono, a monte, interrogarsi sulle sorti di soggetti rispetto ai quali, per esempio, le evidenze empiriche escludono una capacità di intendere e di volere*.

Nadie puede dudar que, por ejemplo, la cibercriminalidad es un

¹¹ R. CASTILLO FELIPE, S. TOMÁS TOMÁS, *Proceso penal e inteligencia artificial: prevenciones en torno a su futura aplicación en la fase de juicio oral*, en esta obra.

¹² GONZÁLEZ TAPIA, *op. ult. cit.*

¹³ PALMA HERRERA. *op. ult. cit.*

¹⁴ MASSARO, *op. ult. cit.*

fenómeno delictivo preocupante en la actualidad, sobre todo por su difícil persecución. El origen de esta criminalidad comportará casi con seguridad las figuras delictivas contenidas a partir de la valoración del peligro abstracto, incluso figuras de delitos obstáculo. Así lo señala con acierto J. PÉREZ-ARIAS cuando afirma que *la solución no pasaría, en todos los casos, por tipificar la cibercriminalidad como derecho penal del resultado, sino como derecho penal de peligro (o de riesgo)*¹⁵, añadiendo más adelante que con el objeto de *impedir que cada intérprete pueda analizar el peligro de una manera diferente (sobre todo partiendo de que estamos ante un peligro estadístico, que permitiría múltiples interpretaciones, a favor y en contra), nos llevaría a configurar la conducta como delito de peligro abstracto, haciendo innecesaria, con ello, la prueba sobre el riesgo hacia el bien jurídico*¹⁶.

Del mismo modo, los avances en términos de inteligencia artificial nos inducen a adelantar la tutela penal hasta confluir en un Derecho penal que se aleja cada vez más del principio del hecho. Si la victimización de los sistemas inanimados resulta todavía cuestionable, no cabe duda acerca de la necesidad, para el Derecho, de plasmar herramientas eficaces en contra de los posibles ataques de las máquinas hacia el ser humano. Sin embargo, como advierte A. GIRALDI¹⁷, aunque el Derecho tenga que prestar atención a los cambios y a las evoluciones de la técnica, *en modo alguno esto implica la equiparación entre artefactos y seres humanos, ni mucho menos. El Derecho, pues, debe reflejar el axioma según el cual las máquinas quedan al servicio del hombre, tanto en el caso de que la máquina sea dirigida por el ser humano, como cuando ella actúe de forma independiente. Si se accede a desvincular la actuación inhumana del control humano, se daría rienda suelta a la derresponsabilización moral de los individuos, que lleva aparejada una inaceptable debilitación de la tutela de los bienes jurídicos.*

4. La peligrosidad de la inteligencia artificial en la aplicación de la norma

Hay multitud de ejemplos que demuestran la ‘peligrosidad’ de que sea la inteligencia artificial, sin consciencia, quien termine aplicando la norma.

¹⁵ PÉREZ-ARIAS, *op. ult. cit.*

¹⁶ *Ivi*, § 5.

¹⁷ A. GIRALDI, *Deshumanizando la culpabilidad: los sistemas inanimados en la teoría del delito*, en esta obra.

Es el llamado ‘juez-robot’¹⁸. Así, basar las medidas de prevención, como medidas de seguridad y sustitutivos penales, en criterios estrictamente probabilísticos automatizados supone una inamisible renuncia a Derechos fundamentales. En cierta medida, *si el funcionamiento de la herramienta se basa en correlaciones estadísticas identificadas por el propio algoritmo a través del análisis de casos anteriores, el resultado predictivo puede conducir a una ‘desindividualización’ de la decisión* (L. NOTARO¹⁹).

Además de ello, piénsese, como correctamente analizan R. CASTILLO FELIPE y S. TOMÁS TOMÁS, en *el supuesto de que el software aportase datos de hecho no aducidos por las partes o máximas de la experiencia no introducidas por los peritos de las que dependa la acreditación de la conducta delictiva*²⁰.

Ser conscientes de los grandes riesgos que entrañaría una aplicación mecánica de los avances en neurociencia y en inteligencia artificial, no debe comportar un abandono de esta fuente de conocimiento; no debe suponer en absoluto un aislamiento por parte del Derecho Penal de los nuevos avances técnicos en estas materias. En palabras de J. PÉREZ-ARIAS²¹, *el derecho penal necesita responder a nuevos retos que, a su vez, pueden exigir de él nuevos principios (si lo que se pretende es idear una respuesta global y satisfactoria al problema)*.

Resulta una obviedad decir que estos avances deben ser tenidos en cuenta para mejorar la técnica procesal y judicial, pero teniendo presente que todo lo técnicamente posible no tiene por qué ser jurídicamente admitido. Por ello, el desembarco de la IA en el proceso sólo debería producirse con un marco legislativo que proporcione seguridad jurídica, con unos órganos de control integrados por personal especialista y no político que estén bien formados, y con unas garantías de transparencia irrenunciables para quienes pretendan diseñar sistemas que auxilien al tribunal en la difícil labor de juzgar. De otro modo, estarán en juego derechos y garantías que, aunque aplicados en un sistema imperfecto –toda creación humana lo es–, hasta ahora estaban suficientemente protegidos (R. CASTILLO FELIPE y S. TOMÁS TOMÁS²²).

Con todo, la utilización de la inteligencia artificial como instrumento de apoyo para el ejercicio de la justicia conlleva el peligro de alterar la estructura del principio *iudex peritus peritorum*, cuyo alcance fundamenta

¹⁸ *Ivi*, § 1.

¹⁹ L. NOTARO, ‘Algoritmos predictivos’ y justicia penal desde una perspectiva italiana y europea, en esta obra.

²⁰ *Ivi*, § 2.1.

²¹ PÉREZ-ARIAS, *op. ult. cit.*

²² CASTILLO FELIPE, TOMÁS TOMÁS, *op. ult. cit.*

la libre valoración de la prueba en el proceso penal. Ante la amenaza de dicha alteración, es oportuno que la inteligencia artificial siga siendo utilizada bajo el control, la supervisión y la responsabilidad del juez persona física (A. GIRALDI²³). De no ser así, *si ritornerebbe ad un diritto oggettivo, preciso, calcolabile, esatto*, cuya máxima exaltación quizás *rischia di sfociare in un'applicazione delle pene fisse, contrarie alla Costituzione* (L. GROSSI²⁴).

Recuérdese también, al respecto, el alcance de la normativa internacional, que exige el respeto de principios ineludibles. Como advierte L. NOTARO²⁵, *una referencia importante dentro del sistema del Consejo de Europa es la 'Carta Ética Europea sobre el uso de la inteligencia artificial en los sistemas judiciales y su entorno', adoptada por la Comisión Europea Para la Eficiencia de la Justicia (CEPEJ) en 2018. La Carta establece cinco principios fundamentales: 1) respeto por los derechos fundamentales; 2) principio de no discriminación; 3) principio de calidad y seguridad de los datos; 4) principio de transparencia, imparcialidad y justicia; 5) principio 'bajo control del usuario' (garantía del control humano).*

En sentido contrario, desde la perspectiva de la cibercriminalidad, tampoco lo tecnológico es un gran aliado de la administración de justicia, tal y como hoy se configura. Adviértase, por ejemplo, el carácter transfronterizo de la criminalidad a través de las redes, que resulta ser un factor negativo que impide avanzar en el esclarecimiento de hechos. En efecto, *los jueces españoles van a encontrar serios inconvenientes en su labor judicial, cuando una investigación se tiene que realizar, en su totalidad (no en algún aspecto concreto, que sí es frecuente) con diligencias para las que se necesitan el completo auxilio y/o cooperación de terceros países. Y, aun cuando eso se produzca, luego sería complicado abrir juicio oral, porque se requiere presencialidad de los acusados* (J. PÉREZ-ARIAS²⁶).

En definitiva, para huir del peligro que conllevan las tendencias maquiavélicas y pragmatistas, el análisis de las cuestiones jurídicas más relevantes, que se desarrollará en los capítulos siguientes, está encaminado a reforzar las garantías del ciudadano ante un Derecho y un proceso penal que corren el riesgo de perder sus caracteres de justicia. En otras palabras, la posibilidad de aprovechar el empleo de los sistemas de inteligencia artificial, integrados por los avances de la neurociencia, nos permite consolidar la observancia de los derechos más básicos. *Ad ogni fatto' nuovo,*

²³ GIRALDI, *op. ult. cit.*

²⁴ L. GROSSI, *I miti d'oggi: la certezza del diritto fra utopie e distopie*, en esta obra.

²⁵ NOTARO, *op. ult. cit.*

²⁶ PÉREZ-ARIAS, *op. ult. cit.*

che si presenta nella serie storica di un dato contesto, può guardarsi come ad una cesura, che si stacca da ciò che la precede e si pone come pietra angolare di qualcosa di diverso, oppure come uno sviluppo necessario di qualcosa che, ad una più attenta analisi, già si era manifestato all'interno della medesima serie (L. GROSSI²⁷).

Los que, en sentido contrario, huyen de esta nueva realidad social y rechazan *tout court* la potencialidad de sus aportaciones, en palabras de A. MASSARO²⁸ quizás están preocupados por su repercusión positiva: *chi ha paura dell'intelligenza artificiale?*

5. Bibliografía

CASTILLO FELIPE R., TOMÁS TOMÁS S., *Proceso penal e inteligencia artificial: prevenciones en torno a su futura aplicación en la fase de juicio oral*, en esta obra.

GIRALDI A., *Deshumanizando la culpabilidad: los sistemas inanimados en la teoría del delito*, en esta obra.

GONZÁLEZ TAPIA M.I., *Neurociencias e imputabilidad: recapitulando*, en esta obra.

GROSSI L., *I miti d'oggi: la certezza del diritto fra utopie e distopie*, en esta obra.

MASSARO A., *Intelligenza artificiale e neuroscienze: l'eterno ritorno del diritto penale?*, en esta obra.

MORILLAS FERNÁNDEZ D.L., *Implicaciones de la inteligencia artificial en el ámbito del Derecho Penal*, en esta obra.

NOTARO L., *'Algoritmos predictivos' y justicia penal desde una perspectiva italiana y europea*, en esta obra.

PALMA HERRERA J.M., *Inteligencia artificial y neurociencia. Algunas reflexiones sobre las aportaciones que pueden hacer al Derecho Penal*, en esta obra.

PÉREZ-ARIAS J., *Algoritmos y big data en la responsabilidad penal: el reto de la cibercriminalidad en el Derecho Penal*, en esta obra.

PERIS RIERA J.M., *La regulación penal de la manipulación genética en España (principios penales fundamentales y tipificación de las genotecnologías)*, Civitas, Madrid, 1995.

²⁷ GROSSI, *op. ult. cit.*

²⁸ MASSARO, *op. ult. cit.*

CAPÍTULO II

Intelligenza artificiale e neuroscienze: l'eterno ritorno del diritto penale?

Antonella Massaro

SOMMARIO: 1. Diritto penale e nuove tecnologie: rivoluzione copernicana o eterno ritorno? – 2. Neuroscienze e diritto penale nella dialettica tra normativismo e ontologismo – 2.1. La natura strutturalmente normativa del diritto penale e della teoria generale del reato – 2.1.1. Le neuroscienze e la questione del libero arbitrio – 2.1.2. Neuroscienze, diritto e processo penale – 2.1.3. Neuroscienze e diritto penitenziario – 3. Le relazioni tra intelligenza artificiale e diritto penale 3.1. a) l'algoritmo delinquente – 3.2. b) l'algoritmo investigante e/o inquirente – 3.3. c) l'algoritmo consulente e/o giudicante – 4. Corsi, ricorsi e deviazioni nella parabola del diritto calcolabile – 5. Bibliografia.

1. Diritto penale e nuove tecnologie: rivoluzione copernicana o eterno ritorno?

La cornice entro cui ricondurre le implicazioni penalistiche tanto dell'intelligenza artificiale quanto delle neuroscienze potrebbe essere rappresentata dall'impatto delle nuove tecnologie sul diritto penale e, più esattamente, dalle 'reazioni' del sistema penale a fronte delle categorie concettuali veicolate dall'evoluzione tecnologica.

I rapporti tra il diritto penale e le nuove tecnologie, a ben vedere, possono descriversi secondo due modelli, che, sia pur semplificati, schematizzano le principali opinioni registratesi negli ultimi decenni.

Secondo un primo schema, che potrebbe definirsi *modello della rivoluzione copernicana*, le nuove tecnologie avrebbero determinato un autentico mutamento di paradigma del diritto penale: si tratterebbe di una trasformazione strutturale e radicale, che imporrebbe non solo una diversa interpretazione delle categorie tradizionali (a partire da quella, fondamentale, di 'responsabilità'), ma anche una più generale rifondazione del sistema.

Una seconda ipotesi ricostruttiva, per contro, è quella sintetizzabile

come *modello dell'eterno ritorno dell'uguale*, prendendo a prestito, sia pur decontestualizzata, una delle più note teorizzazioni filosofiche di Nietzsche. Le nuove tecnologie finirebbero per riproporre, attraverso dei concetti nuovi e linguaggi innovativi, delle questioni tradizionali e, per certi aspetti ataviche, del diritto penale. L'obiettivo, allora, non sarebbe tanto ridescrivere le fondamenta del sistema penale, ma, piuttosto, valorizzare gli strumenti già esistenti e, almeno in certi casi, scarsamente valorizzati.

Proprio questo secondo modello, in effetti, risulta quello preferibile. Non si tratta, beninteso, di minimizzare la svolta epocale impressa, anzitutto sul piano socio-culturale, dall'inarrestabile accelerazione che ha caratterizzato l'evoluzione tecnologica. La questione, tuttavia, può e deve essere contestualizzata entro il più limitato perimetro della riflessione giuridico-penale, magari valorizzandola come occasione per ripercorrere antichi dilemmi e risalenti questioni irrisolte.

2. *Neuroscienze e diritto penale nella dialettica tra normativismo e ontologismo*

Sul versante delle neuroscienze, il complesso rapporto con il diritto è efficacemente sintetizzato dal titolo del noto articolo di Joshua Greene e Jonathan Cohen (*For the law, neuroscience changes nothing and everything*)¹, cui fa eco, nello stesso anno, la riflessione di Stephen Morse (*New neuroscience, old problems: legal implications of brain science*)².

Il concetto di 'neuroscienze' sintetizza un ampio ed eterogeneo gruppo di discipline scientifiche, il cui tratto comune sarebbe la dimostrazione che le connessioni neuronali influiscono sullo svolgimento di tutte le attività umane, non solo di quelle che consistono in semplici movimenti corporei, ma anche di quelle che, come la volizione o le emozioni, sono tradizionalmente considerate un'espressione della mente e che, in quanto tali, non sarebbero suscettibili di indagine sperimentale³. Risulterebbe superata, detto altrimenti, la consolidata dicotomia tra 'mente' e 'cervello', nel senso che 'la mente è il cervello' o, meglio, 'la mente è quel che il

¹ J. GREENE, J. COHEN, *For The Law, Neuroscience Changes Nothing And Everything*, in «The Royal Society», 2004, pp. 1175 ss.

² S. MORSE, *New Neuroscience, Old Problems: Legal Implications of Brain Science*, in «Cerebrum», 2004, pp. 81 ss.

³ C. GRANDI, *Neuroscienze e responsabilità penale. Nuove soluzioni per problemi antichi?*, Giappichelli, Torino, 2016, p. XI, al quale si rinvia, fin da ora, per ampie indicazioni bibliografiche.

cervello fa': le esperienze soggettive, lungi dal restare confinate in una sfera accessibile solo attraverso l'introspezione, sono 'osservabili dall'esterno' e, dunque, suscettibili di valutazione oggettiva⁴.

Da questa chiave di lettura in senso biologico (anche) di attività complesse della persona umana deriva, spesso, un atteggiamento di scetticismo e diffidenza nei confronti delle neuroscienze, specie da parte di chi ritiene che le stesse, legittimando un riduzionismo rigorosamente determinista, segnerebbero, nel campo del diritto, dei pericolosi passi indietro rispetto alle acquisizioni più recenti e condivise⁵.

Indipendentemente dal fatto che queste obiezioni risultino più o meno convincenti, è innegabile che, con specifico riguardo al diritto penale, le neuroscienze ripropongano il complesso dialogo con le scienze naturali o, se si vuole, il dibattito sulla 'autosufficienza' del diritto positivo e sulla impermeabilità dello stesso rispetto ad apporti *lato sensu* esterni.

L'eterno ritorno che parrebbe veicolato dalle neuroscienze, detto altrimenti, è quello sintetizzato dalla dialettica tra normativismo e ontologismo, sia pur intendendo le categorie in questione in un'accezione ampia.

Si pensi, anzitutto, alle oscillazioni registratesi storicamente nella riflessione giuridica tedesca, seguendo quella ideale linea di sviluppo che dal naturalismo di von Liszt conduce, passando per l'ontologismo di Welzel, alla 'reazione normativista' di Roxin e Jakobs⁶.

Il modello positivista proposto da Binding, come noto, si fondava sull'assoluta centralità del diritto positivo e sul rifiuto di ogni possibile ruolo dei giudizi di valore. Sempre nell'ambito del positivismo giuridico, von Liszt, pur mantenendo salda l'esclusione tanto della riflessione filosofica quanto di possibili valutazioni di tipo assiologico, valorizza la realtà empirica come oggetto delle scienze sperimentali: quello di von Liszt, si è osservato, sarebbe un positivismo giuridico con tinte naturaliste⁷.

Con la riflessione neokantiana, specie nella sua versione di 'filosofia dei valori', si assiste a una decisa 'riabilitazione' della dimensione 'extranormativa': le categorie fondamentali della teoria generale del reato (azione, anti-giuridicità, colpevolezza) non possono comprendersi senza un

⁴ O. DI GIOVINE, voce *Neuroscienze*, in «Enciclopedia del diritto», Annali VII, 2014, p. 712.

⁵ Per queste critiche, e per una loro proposta di superamento, DI GIOVINE, voce *Neuroscienze*, cit., pp. 711 ss.

⁶ *Amplius*, S. MIR PUIG, *Límites del normativismo en derecho penal*, in «Revista electrónica de ciencia penal y criminología», 2005, 18:1 ss.; J.C. MIR, *Ontologismo y normativismo en el finalismo de los años cincuenta*, in «Revista de derecho penal y criminología», 2003, pp. 45 ss.

⁷ MIR PUIG, *Límites del normativismo*, cit., p. 18:4.

riferimento al loro significato assiologico, con un'irruzione dei valori che comporta necessariamente un ridimensionamento delle scienze naturali o, comunque, il riferimento a una metodologia differente da quella strettamente sperimentale⁸.

Welzel, proponendo una visione che, forse troppo sbrigativamente, si trova ricondotta all'etichetta dell'ontologismo⁹, pur reagendo con vigore alle distorsioni che deriverebbero dalla teoria causale, ritiene che il diritto penale sia vincolato alle strutture logiche della realtà (*Sachlogische Strukturen*), intese come regole di valore oggettivo che rappresentano un vincolo per il legislatore¹⁰. L'azione, solo per restare a uno dei tratti più celebri del pensiero welzeliano, lungi dall'esaurirsi in una dimensione meccanicistico-causale, può intendersi solo come esercizio di un'azione finalisticamente orientata¹¹.

Tra le reazioni normativiste al preteso ontologismo di Welzel, una delle più celebri è certamente quella di Roxin, che propone un sistema non più impermeabile alla dimensione del sociale e del politico, ma nel quale, al contrario, la politica criminale diviene un contenuto proprio della teoria generale del reato¹².

L'alternativa tra normativismo e ontologismo, in maniera forse più nitida di quanto avviene nello scenario tedesco, emerge anche nell'esperienza giuridica italiana ed è sintetizzata dalla contrapposizione tra Scuola classica e Scuola positiva¹³.

La Scuola classica fondava il proprio sistema sulla centralità del libero arbitrio, da cui deriverebbe anche la libertà di scelta nella commissione di un reato: il risultato è quello di un 'diritto penale della responsabilità',

⁸ V., ancora, MIR PUIG, *Límites del normativismo*, cit., p. 18:4 ss. Per una delimitazione, certamente non scontata, del concetto di 'neokantismo', A. NISCO, *Neokantismo e scienza del diritto penale. Sull'involuzione autoritaria del pensiero penalistico tedesco nel primo Novecento*, Giappichelli, Torino, 2019, pp. 6 ss.

⁹ J. CEREZO MIR, *Ontologismo e normativismo nel finalismo degli anni Cinquanta*, in *Significato e prospettive del finalismo nell'esperienza giuspenalistica*, a cura di S. Moccia, Napoli, Edizioni Scientifiche Italiane, 2002, 71.

¹⁰ H. WELZEL, *Naturalismus und Wertphilosophie in Strafrecht* (1935), in *Abhandlungen zum Strafrecht und zur Rechtsphilosophie*, Berlino-New York, de Gruyter, 1975, 29 ss.

¹¹ Per tutti i necessari riferimenti, L. CORNACCHIA, *Ein unausrottbares Verständnis? L'eredità del finalismo nel dibattito penalistico italiano*, in «Criminalia», 2013, pp. 577 ss.

¹² C. ROXIN, *Politica criminale e sistema del diritto penale*, Guida Editori, Napoli, 1986, pp. 32-33.

¹³ Cfr. M. SBRICCOLI, *Caratteri originari e tratti permanenti del sistema penale italiano (1860-1990)*, in *Storia del diritto penale e della giustizia. Scritti editi e inediti (1972-2007)*, Giuffrè, Milano, 2009, pp. 497 ss.

corredato da una pena intesa essenzialmente in funzione retributiva. Al centro della riflessione viene collocato lo studio del reato, inteso come entità giuridica ricavabile da un'indagine razionale.

La Scuola positiva, per contro, valorizza una visione fondata sul determinismo causale e, quindi, lo studio di tutti quei fattori (sociali, culturali, biologici) che 'conducono' il soggetto alla commissione di un reato: il modello che ne deriva è quello di un 'diritto penale della pericolosità', che, come alternativa alla pena, teorizza l'impiego di misure di sicurezza. Il dibattito scientifico si sposta, dunque, dal reato al delinquente.

Il codice penale del 1930, come ampiamente noto, tenta un 'compromesso storico' tra le due visioni contrapposte, mentre Arturo Rocco teorizza quel metodo tecnico-giuridico che avrebbe dovuto traghettare il diritto penale fuori dalla crisi da cui sembrava strutturalmente afflitto. 'Metodo tecnico-giuridico' significa, prima di tutto, 'centralità del diritto positivo'¹⁴, con l'obiettivo di scrostare il diritto penale dal peso della riflessione filosofica, sociologica, psicologica o antropologica e di restituirlo all'unica dimensione praticabile, da intendersi come una 'dottrina pura del diritto penale'.

Sebbene una lettura ampia dei concetti di normativismo e di ontologismo si esponga al rischio di indebite semplificazioni, l'eterno ritorno del diritto

¹⁴ ART. ROCCO, *Il problema e il metodo della scienza del diritto penale*, in *Opere giuridiche*, Vol. III, Roma, Società editrice del Foro Italiano, 1933, pp. 268 ss.: «Quale, in particolare, la causa prossima di un tale stato di cose? La diagnosi non pare difficile. L'unica scienza classica del diritto penale ignara, prima, dimentica, poi, degli insegnamenti della scuola storica del diritto, aveva preteso di studiare un diritto penale all'infuori del diritto positivo, si era illusa di potere, essa, foggare un diritto penale diverso da quello consacrato nelle leggi positive dello Stato, un diritto penale di carattere assoluto, immutabile, universale, la cui origine andasse rintracciata nella divinità, o nella rivelazione dell'umana coscienza, o nelle leggi di natura o nelle leggi del pensiero e dell'idea. [...] O noi ci sbagliamo, o non c'altro rimedio che questo: rimedio semplicissimo, almeno ad enunciarlo: tenersi fermi, religiosamente e scrupolosamente attaccati allo studio del diritto. [...]. Ma qui vedo già sorgere all'orizzonte una capitale obiezione. Si dirà che una tale distinzione della scienza del diritto penale dalle scienze psicologica, antropologica e sociologica, da un lato, dalla filosofia del diritto e dalla politica, dall'altro, non è scientificamente e praticamente possibile. Si dirà che, in tal modo, si distrugge il diritto penale come scienza, che se ne fa un vuoto, quanto pericoloso, formalismo; che si riduce la scienza ad un puro esercizio scolastico di astrazioni teoriche; che si bandisce un isolamento cellulare fra le varie scienze criminologiche, quanto mai gravido di nefaste conseguenze pratiche per la società civile. Ma non è punto il nostro pensiero a giungere ad una tal conclusione, né essa è punto una conseguenza necessaria delle nostre affermazioni. Ciò che si vuole è soltanto che la scienza del diritto penale conformemente alla sua natura di scienza giuridica speciale limiti l'oggetto delle sue ricerche dirette, allo studio esclusivo del diritto penale e, conformemente ai suoi mezzi, dell'unico diritto penale che esista come dato dell'esperienza, cioè il diritto penale positivo».

penale, come anticipato, sembra risiedere nel dibattito sull'indipendenza e sull'autosufficienza del dato normativo, a volte affermata con forza, altre volte radicalmente messa in discussione. Focalizzando l'attenzione sui rapporti tra diritto penale e scienze naturali (e tralasciando, quindi, il possibile apporto della filosofia e della sociologia), sono almeno tre gli interrogativi che, sia pur in contesti culturali differenti, sembrano ciclicamente riproporsi:

- a) qual è il rapporto tra il sapere giuridico e il sapere scientifico?
- b) il legislatore deve costruire gli istituti giuridici conformemente a quanto affermato dalle scienze naturali?
- c) l'interprete deve attribuire al diritto positivo un significato conforme alle indicazioni ricavabili dalle scienze naturali?

2.1. La natura strutturalmente normativa del diritto penale e della teoria generale del reato

Gli interrogativi ai quali si è fatto riferimento sono indubbiamente complessi e risultano forse refrattari a risposte univoche e perentorie.

È però possibile assumere come punto fermo sufficientemente solido una premessa di carattere generale: *il diritto penale non può essere del tutto permeabile alle scienze naturali.*

La motivazione di questa premessa è evidente, quasi banale. L'ordinamento giuridico è un sistema di norme, che segue delle regole sue proprie e che, in questo modo, persegue l'obiettivo di una coerenza interna. Le regole in questione rispondono a principi e categorie caratteristici del diritto e che non necessariamente coincidono con le conclusioni cui, per esempio, pervengono le scienze naturali.

Anche gli elementi del reato, per quello che qui maggiormente interessa, sono dotati di una consistenza normativa, la quale, a sua volta, non persegue alcuna vocazione universale. La causalità, il dolo, ma anche la capacità di intendere e di volere, detto altrimenti, sono concetti, ricavabili dal diritto positivo e dell'interpretazione sistematica, 'validi' sono nel limitato contesto del diritto penale. La colpa penale, per esempio, è strutturalmente diversa dalla colpa civile e nessuna delle due coincide del tutto con il concetto di colpa derivante dalla teologia o dalla filosofia morale.

Si tratta, allora, di 'regole del gioco' di cui va verificata la validità sul piano interno, secondo, cioè, i vincoli imposti dal settore dell'ordinamento di volta in volta preso in considerazione. Se alcune regole contenute nel codice penale prevedono delle ipotesi di responsabilità oggettiva che,

tuttavia, contrastano con le indicazioni ricavabili dalla Costituzione, quelle regole andranno riscritte per assicurare la coerenza del sistema in cui sono inserite; non sarebbe invece necessario ridefinire il concetto di 'responsabilità penale' al solo scopo di renderlo conforme alle indicazioni ricavabili dalla ricerca scientifica.

2.1.1. *Le neuroscienze e la questione del libero arbitrio*

Un esempio può forse valere a chiarire il senso delle precedenti considerazioni ed è quello che attiene all'autentico motore immobile del diritto penale: il libero arbitrio come fondamento, anche se solo implicito, del giudizio di responsabilità penale.

Se il mistero del libero arbitrio è una questione che 'fa tremar le vene e i polsi' (non solo al giurista, ma) a chiunque si soffermi a riflettere sulla condizione dell'essere umano, non può certo pretendersi che la questione sia risolta dal diritto, per quanto il diritto penale ponga al centro della sua riflessione proprio la condotta umana e le sue conseguenze.

Il diritto penale, questo è il punto, non 'risolve' il dilemma del libero arbitrio, ma 'prende atto' della sola soluzione compatibile con le regole su cui il sistema penale si fonda. La libertà di volere dell'uomo, così come considerato dal diritto (non solo penale) e, più in generale, dell'uomo inserito nel contesto delle relazioni sociali, assume la consistenza di un vero e proprio assioma: un assunto che enuncia una verità evidente a chiunque, non dimostrabile e al tempo stesso indubitabile, che rappresenta la premessa necessaria per la costruzione di un sistema. La libertà del volere, intesa anzitutto come materiale possibilità di agire diversamente¹⁵, è il tassello senza il quale l'intero mosaico del diritto penale sarebbe destinato a sgretolarsi.

Il nodo problematico, allora, non risiede tanto nella dimostrabilità filosofica o scientifica del libero arbitrio. Si tratta, molto più 'modestamente', di prendere atto di un dato così profondamente radicato che neppure il principio dell'onniscienza di Dio e il conseguente determinismo teologico sono riusciti a mettere in discussione¹⁶ e che, come precisato, rappresenta la

¹⁵ Efficace la titolazione di paragrafo scelta da L. FERRAJOLI, *Diritto e ragione. Teoria del garantismo penale*, Bari, Laterza, 2004, p. 496: 'Avrebbe potuto agire diversamente?': il dilemma metafisico tra determinismo e libero arbitrio. Cfr. anche M. PLANCK, *Legge di causalità e libero arbitrio*, in *La conoscenza del mondo fisico*, Torino, Einaudi, 1949, pp. 97 ss.

¹⁶ Evocativi i versi di D. ALIGHIERI, *Purgatorio*, Canto XVI (67-81), ispirati dall'insegnamento tomistico per cui *astra inclinans non necessitant*: «Voi che vivete ogni cagion recate/ pur suso al cielo, pur come se tutto/ movesse seco di necessitate./ Se così fosse, in voi fora distrutto/ libero arbitrio, e non fora giustizia/ per ben letizia, e per male

condicio sine qua non per ipotizzare l'applicazione di ogni tipo di sanzione, a partire, ovviamente, da quella penale¹⁷.

Lo stesso legislatore del codice Rocco, quasi motivato da una sorta di timore reverenziale nei confronti di un concetto tanto potente ed evocativo quale quello di libero arbitrio, sembra apparentemente prenderne le distanze, per poi giungere a 'constatare' la sua irrinunciabilità¹⁸. I compilatori del codice, premesso che la questione filosofica del libero arbitrio, insoluta e forse insolubile, deve rimanere del tutto estranea al diritto penale¹⁹, si rifugiano nella più 'rassicurante' considerazione per cui anche il volere dell'uomo subisce, come ogni fenomeno naturale, la legge di causalità. Al tempo stesso, tuttavia, si rende necessario ammettere

aver lutto./ Lo cielo i vostri movimenti inizia;/ non dico tutti, ma, posto ch'i' 'il dica,/ lume v'è dato a bene e a malizia,/ e libero voler; che, se fatico/ ne le prime battaglie col ciel dura,/ poi vince tutto, se ben si notrica./ A maggior forza e a miglior natura/ liberi soggiacete; e quella cria/ la mente in voi, che 'l ciel non ha in sua cura». Sul libero arbitrio nella filosofia scolastica si rinvia a L. SCARANO, *Libera volontà e libero arbitrio nel diritto penale*, Giuffrè, 1937, pp. 15 ss.

¹⁷ Cfr. FERRAJOLI, *Diritto e ragione*, cit., p. 504: «Se vogliamo, possiamo ben chiamare 'libero arbitrio' l'alternativa *ex ante* tra possibilità di commettere e possibilità di omettere l'azione proibita che forma il presupposto della scelta tra le due cose: a condizione però che questa alternativa sia considerata non già come ontologica, ma come deontologica, non riferita alla struttura ontica del mondo ma a quella deontica delle norme. Un determinista irriducibile potrebbe obiettare che tutto questo conferisce alla colpevolezza, e con essa al libero arbitrio, il carattere di una convenzione. Ma si tratterebbe di una convenzione che ha il suo fondamento nella struttura stessa del diritto quale fenomeno regolativo, e che può essere serenamente accettata almeno finché accettiamo l'esistenza di regole o prescrizioni e non consideriamo insensato il contenuto». V. anche B. PETROCELLI, *La colpevolezza*, CEDAM, 1955, pp. 65 ss.: «A torto o a ragione, qualunque possa essere la verità da un punto di vista filosofico, è su tale principio che, nella realtà della vita, hanno fondamento tutte le relazioni umane [...]. Ora, chi vede il diritto penale, e il diritto in genere, fondato sul principio della libertà del volere non fa che prendere atto di questa realtà; chi vuol fondarlo su una base diversa esce fuori da questa realtà [...]. L'interprete della legge non può prescindere da ciò che è; non può chiudere gli occhi di fronte al principio della libertà del volere, se è tale principio che ispira nel diritto vigente le norme sulla colpevolezza». V. MANZINI, *Trattato di diritto penale italiano*, V ed. aggiornata dai Professori Nuvolone e Pisapia, Vol. I, Utet, 1984, p. 728, nota 9, ritiene che la volontà libera è un fatto universalmente sentito ed ammesso, indipendentemente dalla sua valutazione filosofica e questo è sufficiente per il diritto penale. Per un tentativo di dimostrazione dell'esistenza della libertà del volere, sul piano biologico e su quello psicologico, v. invece H. WELZEL, *Il nuovo volto del sistema penale*, in «Jus», 1952, pp. 53 ss.

¹⁸ PETROCELLI, *La colpevolezza*, cit., pp. 77 ss.

¹⁹ *Lavori preparatori del codice penale e del codice di procedura penale*, Vol. V, Parte I, *Relazione sul Libro I del Progetto*, Tipografia delle Mantellate, Roma, 1929, p. 43.

che la volontà dell'uomo non soggiace in modo fatale ai possibili motivi contrapposti, posto che ciascuno ha la facoltà di discernere e selezionare (*id est*: scegliere) fra i vari motivi²⁰. Lo svelamento del mero espediente letterale è completo quando nella Relazione al Progetto definitivo si riconosce, in maniera esplicita, la capacità di «autodeterminazione» dell'uomo²¹; si getta poi definitivamente la maschera con la considerazione per cui «non può invero concepirsi un diritto penale senza il presupposto di una capacità dell'uomo di auto-determinarsi per motivi coscienti, dal momento che la sanzione penale, come del resto ogni altra sanzione giuridica e morale, mira appunto a contrapporre ai motivi a delinquere il motivo psicologico che deriva dalla minaccia e dall'applicazione ed esecuzione della pena»²².

Le considerazioni in questione non sono certo scontate, anche per effetto dell'irruzione delle neuroscienze su un terreno ormai sufficientemente consolidato. Secondo alcuni esperimenti neuroscientifici, come quelli condotti da Benjamin Libet all'inizio degli anni Ottanta, molte 'decisioni' sarebbero eseguite dal soggetto prima ancora che lo stesso ne sia consapevole. Un programma 'forte' o 'rifondativo' delle neuroscienze, dunque, mirerebbe a disvelare l'illusione nascosta dai concetti di libero arbitrio e di libertà di autodeterminazione²³. Si tratta, a ben vedere, di premesse difficilmente conciliabili con le esigenze della responsabilità penale: se, *de iure condito*, il diritto penale attuale presuppone la libertà di scelta del singolo, è altresì difficile immaginare, in una prospettiva *de iure condendo*, la teorica possibilità di costruire un sistema che faccia a meno del libero arbitrio.

Questo però non significa che, sia pur mettendo da parte un approccio rifondativo, l'apporto delle neuroscienze sia destinato a restare ininfluenza sul versante penalistico.

2.1.2. Neuroscienze, diritto e processo penale

Non è un caso, del resto, che i programmi neuroscientifici 'deboli' o 'moderati' valorizzino la possibile incidenza delle neuroscienze su questioni

²⁰ *Lavori preparatori*, cit., Vol. IV, parte II, pp. 137-138 e Vol. V, Parte I, p. 140.

²¹ Cfr. SCARANO, *Libera volontà e libero arbitrio nel diritto penale*, cit., pp. 160-161, secondo il quale nel codice Rocco si afferma un 'autodeterminismo' che risponde pienamente «al generale sentimento ad al comune modo di intendere la libertà: il libero arbitrio».

²² *Lavori preparatori*, cit., Vol. IV, parte II, 138.

²³ Per tutte le necessarie indicazioni DI GIOVINE, *Ripensare il diritto penale attraverso le (neuro-)scienze?*, Giappichelli, Torino, 2019, pp. 3 ss.; GRANDI, *Neuroscienze e responsabilità penale*, cit., 31 ss.

lato sensu processuali, a partire dalla prova di alcuni elementi del reato (il dolo, la capacità di intendere e di volere) o dalla ‘misurazione’ della pericolosità sociale, per arrivare alle tecniche di *lie detention* (distinguere tra dichiarazioni vere e dichiarazioni false o menzognere) e a quelle di *memory detention* (rilevare tracce della memoria di un certo accadimento)²⁴.

Senza entrare nel dettaglio delle singole questioni, può solo osservarsi, in un’ottica orientata al diritto penale sostanziale, che *il diritto penale*, a differenza, per esempio, di quanto accade per il diritto civile, *non è concepibile al di fuori del processo*. Questo significa anche che *gli elementi del reato ‘esistono’ solo se sono verificabili*. Quello imposto dalla verificabilità processuale, a ben vedere, rappresenta il primo e irrinunciabile ‘vincolo di realtà’ cui il diritto penale sarebbe sottoposto²⁵. Non è un caso, del resto, che secondo la Corte costituzionale italiana il canone della verificabilità empirica rappresenterebbe una necessaria articolazione del più ampio principio di determinatezza²⁶.

Le acquisizioni neuroscientifiche, allora, potrebbero senza dubbio

²⁴ Su punto si rinvia, ancora, a O. DI GIOVINE, *Ripensare il diritto penale*, cit., pp. 17 ss. e GRANDI, *Neuroscienze e responsabilità penale*, cit., pp. 11 ss. e pp. 128 ss. Nella giurisprudenza italiana, i più noti e significativi riconoscimenti delle evidenze neuroscientifiche in materia di imputabilità sono quelli offerti da G.i.p. Como, 20 maggio 2011, su cui M.T. COLLICA, *Il riconoscimento del ruolo delle neuroscienze nel giudizio di imputabilità*, in *Dir. pen. cont.*, 15 febbraio 2012 e App. Trieste, 1° ottobre 2009, in «Rivista penale», 2010, pp. 70 ss., con nota di A. FORZA, *Le neuroscienze entrano nel processo penale*.

²⁵ *Amplius*, sui ‘vincoli di realtà’ del diritto penale, D. PULITANÒ, *Il diritto penale fra vincoli di realtà e sapere scientifico*, in «Rivista italiana di diritto e procedura penale», 3/2006, spec. pp. 798 ss.; R. BARTOLI, *Dubbio e certezza del diritto penale*, in «Quaderno di storia del penale e della giustizia», 2/2020, pp. 227 ss.

²⁶ Corte cost., 4 marzo 1981, n. 96, che ha dichiarato l’illegittimità costituzionale del reato di plagio. Al punto 14 del *Considerato in diritto*, la pronuncia osserva che «La formulazione letterale dell’art. 603 prevede pertanto un’ipotesi non verificabile nella sua effettuazione e nel suo risultato non essendo né individuabili né accertabili le attività che potrebbero concretamente esplicarsi per ridurre una persona in totale stato di soggezione, né come sarebbe oggettivamente qualificabile questo stato, la cui totalità, legislativamente dichiarata, non è mai stata giudizialmente accertata. Presupponendo la natura psichica dell’azione plagiante è chiaro che questa, per raggiungere l’effetto di porre la vittima in stato di totale soggezione, dovrebbe essere esercitata da persona che possiede una vigoria psichica capace di compiere un siffatto risultato. Non esistono però elementi o modalità per potere accertare queste particolari ed eccezionali qualità né è possibile ricorrere ad accertamenti di cui all’art. 314 c.p.p., non essendo ammesse nel nostro ordinamento perizie sulle qualità psichiche indipendenti da cause patologiche. Né è dimostrabile, in base alle attuali conoscenze ed esperienze, che possano esistere esseri capaci di ottenere con soli mezzi psichici l’asservimento totale di una persona».

contribuire a rafforzare quel vincolo di realtà imposto al diritto penale dalla sua dimensione strutturalmente processuale, se non altro nella gestione del complesso capitolo della 'prova scientifica'²⁷, che rappresenta una componente fondamentale di quell'eterno ritorno tratteggiato dalla traiettoria del diritto penale degli ultimi decenni.

2.1.3. *Neuroscienze e diritto penitenziario*

Il sistema penale è composto da tre anime, tutte necessarie e da sole non sufficienti a delinearlo in maniera complessiva: il diritto penale sostanziale, il diritto processuale penale e il diritto penitenziario (inteso come il settore dell'ordinamento rivolto in maniera specifica all'esecuzione penale).

Se rispetto alle prime due anime è stato ampiamente indagato il possibile impatto delle neuroscienze, sembra restare ancora inespresso o, comunque, sottostimato, il loro potenziale applicativo sul versante più strettamente penitenziario²⁸.

Sono almeno tre i 'livelli di intervento' che possono immaginarsi:

a) le neuroscienze possono confermare gli effetti negativi che l'isolamento carcerario, almeno a certe condizioni, produce sulla salute mentale delle persone ristrette;

b) le neuroscienze possono suggerire percorsi di sostegno e/o di cura differenti a seconda del disturbo e/o della patologia da fronteggiare;

c) le neuroscienze possono, a monte, interrogarsi sulle sorti di soggetti rispetto ai quali, per esempio, le evidenze empiriche escludono una capacità di intendere e di volere: posto che il percorso che si apre è quello delle misure di sicurezza detentive e che, attualmente, lo stesso risponde a un modello 'unitario', che non consente di distinguere tra le diverse cause di non imputabilità o tra diversi gradi di pericolosità del soggetto, il compito prioritario dell'indagine neuroscientifica potrebbe essere proprio quello di offrire un supporto per un nuovo sistema penitenziario, capace di muovere dal cervello delle persone per tutelare la dignità delle stesse.

²⁷ Per un generale inquadramento della questione, sia consentito il rinvio ad A. MASSARO, *Determinatezza della norma penale e calcolabilità giuridica*, Edizioni Scientifiche, Napoli, 2020, pp. 392 ss.

²⁸ Cfr., tra le pubblicazioni in lingua A. BASKIN-SOMMERS, *La neuroscienza dovrebbe migliorare il sistema carcerario, non cercare di dimostrare l'innocenza*, in «Diritto penale e uomo», 29 gennaio 2020.

3. Le relazioni tra intelligenza artificiale e diritto penale

Il generico riferimento a strumenti di intelligenza artificiale che prenderebbero il posto dei giudici si rivela in realtà ampiamente insufficiente, posto che dietro l'etichetta della intelligenza artificiale si nascondono fenomeni molto diversi tra loro e non sempre di agevole comprensione per i non addetti ai lavori.

Il concetto di intelligenza artificiale, in via di prima approssimazione, fa riferimento a quelle macchine che sono capaci di pensare o, meglio, di imitare il pensiero umano²⁹. John McCarthy, al quale si attribuisce la paternità dell'espressione *Artificial Intelligence*, fondava i suoi studi sulla premessa per cui ogni manifestazione dell'apprendimento possa essere descritta in maniera così precisa da consentire a una macchina di simularla: il problema dell'intelligenza artificiale, quindi, consisterebbe essenzialmente nell'ottenere che una macchina si comporti in modo 'intelligente'³⁰. Si è anche precisato che l'intelligenza artificiale consiste nella capacità della macchina di fare generalizzazioni appropriate in modo tempestivo e su una base dati limitata³¹, e, potrebbe aggiungersi, in maniera tale per cui la mente umana sia in grado di riconoscere la razionalità e, conseguentemente, l'affidabilità del processo posto in essere.

Pur senza sottovalutare le implicazioni etico-filosofiche sottese all'interrogativo *Can machines think?*, che Alan Turing si poneva nel

²⁹ Sullo sfondo di questa definizione riecheggia, evidentemente, la nota distinzione tra *intelligenza artificiale forte*, secondo la quale le macchine hanno una mente e, quindi, sono davvero intelligenti, e *intelligenza artificiale debole*, secondo cui quella artificiale sarebbe una mera imitazione, non una duplicazione dell'intelligenza umana. Si tratta di una distinzione che, come noto, si fa risalire a J.R. SEARLE, *Minds, Brains and Programs*, in *Behavioral and Brain Sciences*, 1980, p. 417: «Secondo l'IA debole, il pregio principale del calcolatore nello studio della mente sta nel fatto che esso ci fornisce uno strumento potentissimo: ci permette, ad esempio, di formulare e verificare le ipotesi in un modo più preciso e rigoroso. Secondo l'IA forte, invece, il calcolatore non è semplicemente uno strumento per lo studio della mente, ma piuttosto, quando sia programmato opportunamente, è una vera mente; è cioè possibile affermare che i calcolatori, una volta corredati dei programmi giusti, letteralmente capiscono e posseggono altri stati cognitivi. Per l'IA forte, poiché il calcolatore programmato possiede stati cognitivi, i programmi non sono semplici strumenti che ci permettono di verificare le spiegazioni psicologiche: i programmi sono essi stessi quelle spiegazioni». Amplius J. KAPLAN, *Intelligenza artificiale. Guida al futuro prossimo*, Luiss University Press, 2017, pp. 4 ss.

³⁰ J. MC CARTHY, M.L. MINSKY, N. ROCHESTER, C.E. SHANNON, *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, 1955, disponibile in <<http://www-formal.stanford.edu>>.

³¹ KAPLAN, *Intelligenza artificiale*, cit., 2017, p. 19.

1950 nell'ambito del suo *Imitation Game*³², si tratta, a ben vedere, di una questione che può lasciarsi insoluta. Se, invece, al concetto di 'pensiero' si sostituisce quello, più 'oggettivo', di 'ragionamento', potrebbe senz'altro affermarsi che le macchine sono in grado di ragionare e che sono in grado di farlo anche nelle forme del ragionamento giuridico.

Si tratta di affermazioni che appaiono ancora più evidenti a seguito dello sviluppo, nella più ampia cornice dell'intelligenza artificiale, dei sistemi di *machine learning* e di *deep learning*. Il riferimento è a quelle macchine che sono in grado di 'imparare dall'esperienza' e, quindi, di migliorarsi autonomamente: una delle possibili forme di *machine learning* è rappresentata dal *deep learning*, in cui l'apprendimento automatico avviene mediante l'impiego di reti neurali artificiali, percorrendo fino in fondo la via dell'equiparazione tra il cervello umano e quello meccanico. Tra le applicazioni di *deep learning* più 'familiari' possono ricordarsi le traduzioni, il riconoscimento vocale o la *computer vision*.

Anche le possibili intersezioni tra diritto penale e intelligenza artificiale attengono a settori molto diversi tra loro, sebbene convergenti in una più ampia ottica di sistema. A meri fini descrittivo-classificatori, si potrebbe distinguere tra l'*algoritmo delinquente*, l'*algoritmo inquirente* e l'*algoritmo consulente e/o giudicante*.

3.1. a) l'*algoritmo delinquente*

L'espressione 'algoritmo delinquente' sintetizza le questioni relative agli *Artificial Intelligence-Supported Crimes* e, quindi, all'intelligenza artificiale come strumento coinvolto, in maniera diversa, nella commissione di certi reati: si pensi alle diverse tipologie di reati 'informatici', alle frodi finanziarie, al traffico di sostanze stupefacenti, ma anche ai droni pilotati per uccidere e garantire la sicurezza urbana o, ancora, agli incidenti cagionati da automobili con pilota automatico³³.

La prospettiva di indagine su questo versante è (almeno) duplice.

In primo luogo, il ricorso sempre più diffuso all'intelligenza artificiale

³² A.M. TURING, *Computing machinery and intelligence*, in «Mind», 1950, p. 433.

³³ R. BORSARI, *Intelligenza Artificiale e responsabilità penale: prime considerazioni*, in *www.medialaws.eu*, 3/2019, § 2; T. KING, M. AGGARWAL, M. TADDEO, L. FLORIDI, *Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions*, in *Science and Engineering Ethics*, 2019, pp. 1 ss.; F. BASILE, *Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*, in «Diritto penale e Uomo», 29 settembre 2019, pp. 24 ss.

anche ‘a fini criminali’ lascia emergere nuove modalità di offesa con le quali il giudice e il legislatore sono chiamati a confrontarsi, per verificare se le fattispecie di reato tradizionali siano sufficienti a fronteggiare le ‘nuove’ esigenze di tutela o se, per contro, si rendano necessari interventi capaci di tenere nella dovuta considerazione le potenzialità di offesa dell’algoritmo criminale.

In secondo luogo, sembrerebbe che il paradigma tradizionale di responsabilità penale, fondato sul reato commesso dalla persona fisica, si stia preparando a subire l’ennesima scossa di assestamento: dopo le discussioni volte a superare il ‘dogma’ del *societas delinquere non potest*, si tratta di chiarire fino a che punto possa valere il principio per cui *machina delinquere non potest*³⁴, posto che, almeno in certi casi, il modello della ‘responsabilità indiretta dell’uomo’ non parrebbe adattabile alle vicende più complesse, specie quelle che vedono coinvolti meccanismi di *machine learning*³⁵.

Sempre nell’ambito dei ‘soggetti del reato’, poi, potrebbe ipotizzarsi che l’algoritmo risulti coinvolto non come soggetto attivo, ma come potenziale vittima. L’equiparazione macchina-persona, specie sul versante dei sentimenti e delle emozioni, dovrebbe condurre a considerare i *robot* come possibili titolari di interessi giuridici da tutelare: se per l’algoritmo-criminale il parallelismo più immediato è quello con la responsabilità penale degli enti, per l’algoritmo-vittima si richiama spesso la possibile analogia con la tutela degli animali, progressivamente elevati dalla riduttiva qualifica di ‘oggetti’ a quella di ‘esseri senzienti’ e, quindi, di possibili ‘soggetti di diritto’³⁶. Si può immaginare la rilevanza penale di una violenza sessuale commessa a danno di un androide? O il maltrattamento (non il semplice danneggiamento) di un *robot* che svolge assistenza nei confronti di soggetti disabili³⁷?

³⁴ A. CAPPELLINI, *Machina delinquere non potest? Brevi appunti su intelligenza artificiale e responsabilità penale*, in «Discrimen», 27 marzo 2019, pp. 1 ss., il cui sottolinea altresì per la fallacia del parallelo con la responsabilità delle persone giuridiche.

³⁵ BORSARI, *Intelligenza Artificiale e responsabilità penale*, cit., § 4. V. anche F. BASILE, *Intelligenza artificiale e diritto penale*, cit., pp. 29 ss., cui si rinvia per ampi riferimenti bibliografici.

³⁶ Cfr. MASSARO, *I reati ‘contro gli animali’ tra aspirazioni zoocentriche e ineliminabili residui antropocentrici*, in «Cultura e Diritti», 1-2/2018, pp. 79 ss.

³⁷ In forma problematica BASILE, *Intelligenza artificiale e diritto penale*, cit., pp. 32-33.

3.2. b) *l'algoritmo investigante elo inquirente*

Gli impieghi dell'intelligenza artificiale come strumento di supporto all'attività del pubblico ministero e della polizia giudiziaria si riferiscono in primo luogo ai c.d. *software* di polizia predittiva, i quali, rielaborando un numero di dati superiore a quelli che sarebbe in grado di gestire un operatore umano, indirizzano nella individuazione delle zone calde (*hotspots*), intese come quelle in cui è elevato il rischio della commissione di reati, oppure supportano nella attività di *crime linking*, prevedendo la futura commissione di reati da parte di soggetti determinati o da determinare³⁸.

A ciò si aggiungono programmi che sono d'ausilio nella ricostruzione di fatti a partire da indizi: STEVIE, per esempio, costruisce storie coerenti con i dati che gli vengono forniti, mentre ALIBI valuta le spiegazioni alternative del comportamento di un soggetto a fronte di un reato che gli viene contestato³⁹.

Nell'esperienza italiana, particolarmente noto è l'esempio offerto da KEYCRIME, un programma sviluppato dal poliziotto Mario Venturi e attualmente in dotazione alla Questura di Milano. I dati dimostrano che l'impiego di KEYCRIME non solo ha determinato una significativa diminuzione di certi reati (soprattutto rapine), ma avrebbe fornito un aiuto significativo nella risoluzione di reati già commessi⁴⁰. La peculiarità del *software* risiede nel fatto che lo stesso è in grado di processare anche dati che attengono al profilo comportamentale dell'autore, aumentando per esempio il suo grado di *accuracy*, rispetto ad algoritmi analoghi, quando si tratti di prevedere la futura commissione di reati da parte di un determinato soggetto.

I rischi, almeno potenziali, sono piuttosto evidenti. Anzitutto la polizia potrebbe agire sulla base di *biased machines*, intese come macchine programmate sulla base di pregiudizi, soprattutto etnico-razziali, che conducano a intensificare i controlli e la repressione soprattutto in certe zone e a riferimento di determinate categorie di soggetti⁴¹: la macchina

³⁸ BASILE, *Intelligenza artificiale e diritto penale*, cit., pp. 10 ss.;

³⁹ *Amplius* E. NISSAN, *Digital Technologies and Artificial Intelligence's Present and Foreseeable Impact on Lawyering, Judging, Policing and Law Enforcement*, in «AI & Society», 2017, pp. 450 ss.

⁴⁰ R. PELLICIA, *Polizia predittiva: il futuro della prevenzione criminale?*, in *www.cyberlaws.it*, 9 maggio 2019.

⁴¹ Un efficace inquadramento della questione è offerto da R.M. O'DONNELL, *Challenging Racist Predictive Policing Algorithms under the Equal Protection Clause*, in «New York University Law Review», 2019, pp. 544 ss.

sarebbe certamente più veloce ed efficiente dell'operatore umano, ma non, per esempio, meno razzista di coloro che la programmano o (soprattutto) di coloro che sono chiamati a 'gestirne' le indicazioni. Senza contare le questioni relative al trattamento dei dati personali, ma anche all'invasività del controllo della polizia in spazi pubblici⁴².

3.3. c) *l'algoritmo consulente e/o giudicante*

Le ulteriori applicazioni 'giuridiche' dell'intelligenza artificiale sono quelle che attengono a un possibile impiego dell'algoritmo in funzione consulente e/o giudicante, rendendolo uno strumento capace non solo di mettersi al servizio del giudice, ma anche di conferire una più adeguata attuazione al diritto di difesa.

L'intelligenza artificiale, anzitutto, consente di organizzare *banche dati sempre più raffinate*, che non rappresentano un mero deposito statico di dati e informazioni, ma che sono in grado di guidare l'utente nella ricerca dei riferimenti normativi e giurisprudenziali rilevanti o, addirittura, di formulare 'pareri'⁴³.

Gli *algoritmi consulenti*, specie quelli che sfruttano il *machine learning* o addirittura il *deep learning*, rappresentano dunque il primo passo verso l'obiettivo di quella che sempre più spesso è etichettata come 'giustizia predittiva'⁴⁴: l'algoritmo consentirebbe cioè di prevedere, con un grado di attendibilità sempre crescente, l'esito di un processo.

Particolarmente noto, anche per i suoi esiti 'positivi', è l'algoritmo, sviluppato da un gruppo di ricercatori, il quale si è mostrato in grado di 'calcolare' gli esiti delle sentenze della Corte europea dei diritti dell'uomo con un grado di precisione del 79%⁴⁵. I ricercatori precisano che mentre

⁴² J. KREMER, *The End of Freedom in Public Places? Privacy problems arising from surveillance of the European public space*, University of Helsinki, 2017, disponibile in <<https://helda.helsinki.fi>>.

⁴³ Per tutti i necessari riferimenti D. BEN-HARI, Y. FRISCH, A. LAZOVKI, U. ELKAN, D. GREENBAUM, *Artificial Intelligence in the Practice of Law: an Analysis and Proof of Concept Experiment*, in «23 Richmond Journal of Law & Technology», 2017, pp. 31 ss.

⁴⁴ E. RULLI, *Giustizia predittiva, intelligenza artificiale e modelli probabilistici*, in «Analisi giuridica dell'Economia», 2/2018, p. 539 ritiene inappropriato l'impiego dell'aggettivo 'predittiva', nella misura in cui rischia di evocare una fantomatica 'divinazione' relativa agli esiti di un processo.

⁴⁵ N. ALETRAS, D. TSARAPATSANIS, D. PREOTIUC-PIETRO, V. LAMPOS, *Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective*, in «PeerJ Computer Science», 2016, pp. 1 ss.

gli studi precedenti valorizzavano metadati capaci di influenzare i voti dei giudici (gravità del reato, posizioni politiche del giudice, indicazioni provenienti da *amici curiae*), il loro studio si propone di prevedere la decisione del caso basandosi unicamente su informazioni di carattere testuale⁴⁶. L'esito dello studio, inoltre, rivelerebbe quanto determinante sia l'apporto del 'fatto' rispetto al 'diritto', in accordo, ad avviso degli autori, con le indicazioni provenienti dal realismo giuridico⁴⁷.

Tralasciando le implicazioni di carattere 'teorico' che si pretenderebbe di ricavare dallo studio in questione, ciò che importa sottolineare è che l'elevato grado di *accuracy* sia riferito al binomio violazione-non violazione di un certo articolo della CEDU: una previsione, detto altrimenti, che riguarda il dispositivo della pronuncia, non già le sue motivazioni.

L'*algoritmo*, infine, potrebbe essere valutato in funzione *giudicante*: si tratta probabilmente dell'applicazione dell'intelligenza artificiale che maggiormente 'impressiona' l'immaginario collettivo, non solo degli addetti ai lavori.

Anche in questo caso sono molte ed eterogenee le applicazioni dell'intelligenza artificiale che potrebbero venire in considerazione.

Si pensi, per esempio, ai *software* capaci di calcolare il rischio di reiterazione del reato e che, quindi, possono essere utilizzati tanto come base per l'applicazione di misura cautelari o di misure alternative alla detenzione quanto come strumento al quale fare riferimento in sede di commisurazione della pena.

Proprio quest'ultima applicazione ha reso note, ben oltre i confini statunitensi, le prodezze di COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*), che è divenuto l'emblema dei vantaggi, ma soprattutto dei limiti, dell'intelligenza artificiale in funzione giudicante. COMPAS applica algoritmi creati sulla base di numerosi dati personali del reo, ma la società privata che ha elaborato e commercializzato il sistema non ne ha rivelato integralmente il meccanismo di funzionamento. Nel celebre caso *Loomis*, l'imputato contestava l'impiego di COMPAS in sede di commisurazione della pena, tanto per il difetto di trasparenza che ne caratterizzava il funzionamento quanto perché sembrava che l'algoritmo risultasse programmato sulla base di pregiudizi di razza e di genere. La Corte Suprema del Wisconsin ha messo in evidenza i possibili rischi insiti nell'uso di COMPAS e consistenti, essenzialmente, nella mancata trasparenza

⁴⁶ ALETRAS, TSARAPATSANIS, PREOTIUC-PIETRO, LAMPOS, *Predicting judicial decisions of the European Court of Human Rights*, cit., p. 2.

⁴⁷ ALETRAS, TSARAPATSANIS, PREOTIUC-PIETRO, LAMPOS, *Predicting judicial decisions of the European Court of Human Rights*, cit., pp. 11 ss.

dei suoi meccanismi di funzionamento, nell'applicazione di logiche elaborate su base collettiva (in riferimento a gruppi o classi di persone) e non strettamente individuale, nella possibile sovrastima del rischio di commissione di reati a carico di alcune minoranze etniche: nonostante ciò, la Corte ha ritenuto legittima una commisurazione della pena fondata anche, ma non solo, sulle indicazioni fornite da COMPAS⁴⁸. La questione, in vicende come questa, attiene non tanto all'impiego dell'algoritmo, ma alla sua programmazione o, più esattamente, al reclutamento e al controllo dei tecnici chiamati ad elaborare e a gestire l'algoritmo stesso.

Non è un caso che tanto la trasparenza, intesa quale accessibilità, comprensibilità e verificabilità esterna dei processi computazionali utilizzati in sede giudiziaria, quanto il divieto di creare o accentuare, per mezzo degli algoritmi, discriminazioni nei confronti di gruppi o individui, siano tra i principi fondamentali individuati dalla 'Carta etica europea per l'uso dell'intelligenza artificiale nei sistemi di giustizia penale e nei relativi ambienti', elaborata, nell'ambito del Consiglio d'Europa, dalla Commissione europea per l'efficacia della giustizia (CEPEJ)⁴⁹.

L'intelligenza artificiale, poi, potrebbe intervenire in fase di valutazione della prova⁵⁰, con prospettive apprezzabili non solo in tema di prova documentale, ma anche in riferimento alle dichiarazioni rese dalle parti o dai testimoni: si tratterebbe di programmare l'algoritmo in modo che lo stesso possa valorizzare, per esempio, le più convincenti acquisizioni in materia di psicologia della prova o, magari, servirsi delle neuroscienze per distinguere le dichiarazioni menzognere da quelle veritiere⁵¹.

L'algoritmo, infine, potrebbe essere d'ausilio al giudice per testare

⁴⁸ Wisconsin S.C., *State v. Loomis*, 881, N.W.2d 749 (2016). *Amplius* J. NIEVA-FENOLL, *Intelligenza artificiale e processo*, Giappichelli, Torino, 2019, pp. 56 ss.; F. BASILE, *Intelligenza artificiale e diritto penale*, cit., pp. 19 ss.; S. CARRER, *Se l'amicus curiae è un algoritmo: il chiacchierato caso Loomis alla Corte Suprema del Wisconsin*, in «Giurisprudenza Penale Web», 4/2019, pp. 1 ss.

⁴⁹ Gli altri principi sono quelli relativi al rispetto dei diritti fondamentali, alla qualità e sicurezza dei dati, al controllo da parte dell'utente dei risultati prodotti dalla macchina: per una efficace analisi della Carta etica v. S. QUATTROCOLO, *Intelligenza artificiale e giustizia: nella cornice della Carta etica europea, gli spunti per un'urgente discussione tra scienze penali e informatiche*, in «La Legislazione penale», 18 dicembre 2018, pp. 1 ss.

⁵⁰ Fondamentale sul punto L. LUPÀRIA, *Prova giudiziaria e ragionamento artificiale: alcune possibili chiavi di lettura*, in *Il concetto di prova alla luce dell'intelligenza artificiale*, a cura di J. Sallantin, J.J. Szczeciniarz, Giuffrè, Milano, 2005, pp. XIV ss.;

⁵¹ NIEVA-FENOLL, *Intelligenza artificiale e processo*, cit., 70 ss. V. anche A. TRAVERSI, *Intelligenza artificiale applicata alla giustizia: ci sarà un giudice robot?*, in «*Questione Giustizia*», 10 aprile 2019, § 6.

la ragionevole esclusione di ricostruzioni alternative rispetto a quella che condurrebbe ad affermare la responsabilità dell'imputato, anche se l'intelligenza artificiale applicata agli *standard* probatori mostra una strutturale debolezza, trattandosi di strumenti più orientati al futuro che alla ricostruzione di eventi del passato⁵².

Al di là delle questioni relative alla prova, intesa in senso ampio, l'impiego dell'algorithmo giudicante che suscita maggiore interesse nell'ottica di un diritto calcolabile è senza dubbio quello di una possibile applicazione nella vera e propria attività di *sentencing*. L'idea di fondo è quella di una macchina che, lungi dal rappresentare un mero ausilio al giudice, lo sostituisca in maniera pressoché integrale nella redazione della sentenza, almeno nei c.d. casi facili.

L'ostacolo principale che parrebbe opporsi alla integrale redazione di una sentenza da parte di un *robot* sembrerebbe risiedere nel carattere strutturalmente persuasivo dell'argomentazione giuridica, la quale, come dimostrano le critiche radicali al sillogismo di matrice illuministica, non si ridurrebbe alla meccanica e sempre identica applicazione di una fattispecie penale, richiedendo per contro valutazioni che, almeno attualmente, la macchina non è in grado di compiere⁵³.

Qualora, tuttavia, ci si trovasse di fronte a casi facili⁵⁴, anzi facilissimi, tali per cui, svolta l'attività istruttoria l'applicazione della norma al caso concreto si rivelasse 'scontata', potrebbe immaginarsi che la sentenza sia automatizzata, ma a questo punto l'esito più naturale sarebbe quello della scomparsa della motivazione. Se la decisione può essere presa da un algorithmo, il cui funzionamento è noto e sulla cui precisione si ritiene di poter fare affidamento, allora, si è recentemente sostenuto, la motivazione diverrebbe superflua⁵⁵.

L'algorithmo, del resto, non solo è in grado di ragionare come un essere

⁵² NIEVA-FENOLL, *Intelligenza artificiale e processo*, cit., pp. 95 ss.

⁵³ NIEVA-FENOLL, *Intelligenza artificiale e processo*, cit., 105 ss. Sui *computer tools* in grado di fornire un supporto argomentativo v. NISSAN, *Digital Technologies and Artificial Intelligence's Present and Foreseeable Impact*, cit., 448-449. Per un'analisi critica delle applicazioni degli attuali strumenti di intelligenza artificiale alle attività di vero e proprio 'giudizio' C. TATA, *The Application of Judicial Intelligence and 'Rules' to Systems Supporting Discretionary Judicial Decision-Making*, in «Artificial Intelligence and Law», 1998, pp. 203 ss.

⁵⁴ La (possibilità di una) distinzione tra casi facili e casi difficili per delimitare i possibili ambiti applicativi dell'intelligenza artificiale in sede penale è la premessa da cui muove DI GIOVINE, *Il Judge-bot e le sequenze giuridiche in materia penale (intelligenza artificiale e stabilizzazione del precedente)*, in «Cassazione penale», 3/2020, p. 952.

⁵⁵ NIEVA-FENOLL, *Intelligenza artificiale e processo*, cit., p. 108 e p. 134 per le implicazioni in materia di impugnazioni.

umano, ma può essere programmato in modo tale da escludere le euristiche e i *biases* cognitivi che, in maniera pressoché inevitabile, condizionano la condotta umana, anche quella di un giudice⁵⁶.

In conclusione, il giudice-algoritmo sarebbe più veloce, più efficiente, con meno pregiudizi, capace di rendere decisioni prevedibili e, quindi, di apportare un contributo significativo (almeno) a una stabilizzazione della giurisprudenza. In sintesi: chi ha paura dell'intelligenza artificiale?

4. Corsi, ricorsi e deviazioni nella parabola del diritto calcolabile

La metafora del diritto come macchina appartiene a secoli di cultura e politica occidentale⁵⁷, trovandosi di fatto impiegata in contesti e per finalità differenti. Se, da una parte, l'assimilazione del diritto a una macchina vale a evidenziare il profilo della sua artificialità, dall'altra parte, la metafora in questione è legata a filo doppio agli ideali illuministici del diritto calcolabile, fondato su una legislazione certa e su un giudice ridimensionato al ruolo di 'bocca della legge'. A quest'ultima visione si avvicina anche quella del diritto come calcolatore, capace di funzionare secondo un codice binario (lecito/illecito, permesso/vietato) e strutturato secondo criteri di razionalità formale⁵⁸.

Il binomio 'diritto e macchina', tuttavia, rappresenta un'immagine tutt'altro che indiscussa: non si è esitato, anzi, a ritenere che «la metafora del diritto come 'macchina' è fra le più infelici che mai siano state utilizzate»⁵⁹.

Anche Max Weber, quando osservava che il capitalismo moderno necessiterebbe di un diritto che si possa calcolare proprio come si fa con una macchina, prendeva atto del timore diffuso tra gli operatori del diritto

⁵⁶ NIEVA-FENOLL, *Intelligenza artificiale e processo*, cit., pp. 32 ss.; DI GIOVINE, *Il Judge-bot e le sequenze giuridiche in materia penale*, cit., p. 954.

⁵⁷ G. ITZCOVICH, *Il diritto come macchina. Razionalizzazione del diritto e forma giuridica in Max Weber*, in «Materiali per una storia della cultura giuridica», 31/2, 2001, p. 379, anche per i necessari riferimenti.

⁵⁸ F. POGGI, *Il diritto meccanico. La metafora del diritto come macchina e i suoi limiti*, in «Diritto e questioni pubbliche», 9/2009, pp. 395-396.

⁵⁹ G. TUZET, *Sul possibile moto della macchina*, in «Diritto e questioni pubbliche», 9/2009, p. 359. Per il dibattito suscitato dallo scritto di Tuzet v. ITZCOVICH, *Sulla metafora del diritto come macchina*, cit., pp. 379 ss.; M. NARVÁEZ MORA, *Metaforear*, in «Diritto e questioni pubbliche», 9/2009, p. 385; F. POGGI, *Il diritto meccanico*, cit., pp. 395 ss.; la replica dello stesso TUZET, *Il diritto non è una macchina*, in «Diritto e questioni pubbliche», 9/2009, pp. 401 ss.

di essere ridotti a meri automi: il progressivo diffondersi di orientamenti realisti, antidogmatici o che, comunque, propongono una alternativa al positivismo giuridico, deriverebbe anche dalla esigenza dei giudici di affrancarsi dalla immagine della 'distributore automatico del diritto'⁶⁰, di mero ricognitore del diritto positivo nell'ambito di quel processo di burocratizzazione della giurisprudenza che, ancora una volta, risulta funzionale a un diritto che possa ritenersi davvero calcolabile⁶¹.

Le obiezioni a un modello di 'diritto meccanico' vanno oltre le questioni relative alla concreta praticabilità di una interpretazione meramente (o prevalentemente) dichiarativa e attengono, sostanzialmente, al rischio di una poco auspicabile deresponsabilizzazione del giudicante-automa: l'applicazione del diritto richiederebbe ad ogni passo una condotta posta in essere da soggetti capaci di autocontrollo e quindi responsabili⁶² o, se si vuole, di giudici (non inanimati ma) con l'anima, capaci di portare «il grande peso di questa immane responsabilità che è il rendere giustizia»⁶³. Senza contare che gli uomini in carne e ossa sono esseri fallibili, ma anche in grado di imparare dai propri errori o, comunque, di 'svecchiare' soluzioni interpretative o modelli argomentativi non più adeguati alla crescente complessità dei fatti che il diritto è chiamato a regolare⁶⁴.

Proprio nel momento storico in cui, anche negli ordinamenti di *civil law*, si assiste a una sempre più marcata componente giurisprudenziale del diritto penale (non si esita a parlare, in certi casi, di una vera e propria 'giurisprudenza fonte'), l'idea di un diritto meccanico sta tornando a monopolizzare il dibattito teorico e politico. Non si tratta più, tuttavia, delle appassionate discussioni intorno a una metafora più o meno efficace e/o condivisibile, ma di prospettive di indagine volte a sondare la praticabilità di una giustizia 'letteralmente' governata dalle macchine o, meglio, con l'ausilio di intelligenze artificiali e di algoritmi in via di

⁶⁰ M. WEBER, *Economia e società*, IV, Edizioni di Comunità, Roma, 1995, p. 79.

⁶¹ Sulla idea di Weber della macchina come 'spirito represso' e, quindi, di un processo di razionalizzazione già compiutamente dispiegatosi, con conseguente 'ridimensionamento' del ruolo del giudice, ITZCOVICH, *Il diritto come macchina*, cit., pp. 372-373 (spec. nota 30) e pp. 391 ss.

⁶² TUZET, *Sul possibile moto della macchina*, cit., p. 359.

⁶³ P. CALAMANDREI, *Processo e democrazia*, in *Opere giuridiche*, I, Roma TrE-Press, Roma, 2019, p. 650.

⁶⁴ Cfr. WEBER, *Economia e società*, III, Edizioni di Comunità, Roma, 1980, p. 41 e p. 170, dove si osserva come al diritto fondato su costruzioni meramente logico-razionali si obietta di essere 'fuori della vita', posto che le soluzioni dallo stesso indicate si pongono spesso in rapporto irrazionale con le aspettative degli operatori del traffico economico, chiamati a operare in situazioni di crescente complessità dei conflitti di interesse.

progressivo affinamento.

A prima vista sembrerebbe di trovarsi di fronte a un ennesimo ‘corso e ricorso della storia’ che non presenta i caratteri di un’autentica ‘rivoluzione’, ma, piuttosto, quelli di un ritorno alle origini. L’intelligenza artificiale, almeno in alcune sue possibili applicazioni, sembrerebbe riproporre la ‘vecchia e superata’ idea del sillogismo giudiziario, a conferma che certi modi di intendere il ragionamento umano applicato al diritto sono talmente radicati da lasciare emergere ancora qualche spunto dalla ‘preziosa ingenuità’ di Beccaria⁶⁵.

Tra l’idea del giudice-automa come metafora di matrice illuministica e del diritto-automatizzato come reale applicazione di sistemi di intelligenza artificiale di nuova generazione, tuttavia, sembrerebbero ravvisabili non solo punti di contatto evidente, ma anche altrettanto macroscopici aspetti di divergenza.

Il punto di contatto è rappresentato dalla dichiarata reazione a un *caos* giurisprudenziale epidermicamente avvertito come intollerabile.

L’età della codificazione ha rappresentato un momento di svolta o, almeno, di ripartenza, sintetizzata non tanto dall’immissione nel sistema di qualcosa di nuovo, ma nella ordinata razionalizzazione della tradizione giuridica precedente. Nella età della decodificazione la riserva di codice recentemente introdotta all’art. 3-*bis* c.p. suona più come l’impotente e disperato tentativo di indirizzare il legislatore futuro, sia pur attraverso la debole arma dell’autovincolo legislativo, senza alcuna possibilità di riparare ai danni, in parte inevitabili, cagionati dal legislatore passato⁶⁶. Le grida manzoniane hanno lasciato il posto a contrasti che, a distanza di poche settimane, si registrano nella medesima sezione della Corte di cassazione, con le Sezioni unite chiamate a interventi sempre più frequenti e ravvicinati, e sempre meno stabili.

La soluzione che l’Illuminismo giuridico proponeva, al fine di emendare alcuni dei più vistosi ‘difetti della giurisprudenza’, era quella di marginalizzare il giudice, marcando la sua differenza rispetto al legislatore e confinandolo nel ruolo di innocuo burocrate. L’interesse primario degli illuministi non era né quello di sistematizzare il diritto né, tanto meno, quello di comprendere l’inesplicabile modo di ragionare dei giuristi:

⁶⁵ Il riferimento è al titolo di V. VELLUZZI, *La preziosa ingenuità: Beccaria, lo spirito della legge e il sillogismo giudiziale*, in «*Rivista internazionale di filosofia del diritto*», 2014, pp. 687 ss.

⁶⁶ Per una dettagliata disamina critica della riserva di codice introdotta con l’art. 3-*bis* c.p. si rinvia a M. DONINI, *La riserva di codice (art. 3-bis c.p.) tra democrazia normante e principi costituzionali. Apertura di un dibattito*, in «*La Legislazione Penale*», 20 novembre 2018, pp. 1 ss.

volevano più semplicemente abolirli, sostituendoli con la legge e con i burocrati chiamati ad applicarla⁶⁷.

I possibili impieghi dell'intelligenza artificiale, in ambito penalistico ma non solo, sembrerebbero muovere da una premessa apparentemente speculare: l'obiettivo non sarebbe quello di marginalizzare il giudice e il suo operato, ma di offrirgli un prezioso strumento di ausilio che, in ultima analisi, valga a esaltarne il ruolo. Anche la 'Carta etica europea per l'uso dell'intelligenza artificiale nella giustizia' enfatizza il necessario controllo da parte dell'utente sui risultati forniti dalla macchina, posto che l'obiettivo degli strumenti computazionali applicati alla giustizia penale dovrebbe essere quello di accrescere, non già di ridurre, l'autonomia decisionale dell'utente stesso⁶⁸.

Il 'mestiere di giudice' è attualmente soffocato e annichilito da un contenzioso elefantico, che si pone quale avversario temibile della stabilizzazione giurisprudenziale⁶⁹. Se i giudici potessero dedicarsi alle 'cose serie', delegando la sorte di ladri di galline e/o di merendine a macchine veloci, economiche e implacabili, la stabilizzazione della giurisprudenza diverrebbe un obiettivo progressivamente più definito.

I giudici in carne e ossa, del resto, agiscono nella maggior parte dei casi in maniera meccanica: più il carico di lavoro si fa gravoso, più le decisioni diventano standardizzate, più aumenta il ricorso alle 'scorciatoie' delle euristiche⁷⁰. La scrittura di innumerevoli documenti di *routine* rallenta il sistema e aumenta il rischio di errori⁷¹.

Se, allora, non si riesce a deflazionare il processo penale attraverso le vie tradizionali (depenalizzazione e riforma del processo), tanto vale intervenire almeno a valle del fenomeno, lanciando al giudice che annega del mare dei processi una scialuppa di salvataggio che non si stanca, non pretende di 'fare dottrina' con le sue opinioni e (forse anche per questo) sbaglia meno.

Non interessa in questa sede valutare l'effettiva praticabilità degli impieghi dell'intelligenza artificiale cui si è fatto rapido cenno, ma è sufficiente soffermarsi su quello che sembrerebbe il prezzo più alto da pagare in termini di sistema. Il rischio, pressoché unanimemente sottolineato, è quello di una stagnazione della giurisprudenza⁷², di una

⁶⁷ M. BARBERIS, *Cosa resta del sillogismo giudiziale? Riflessioni a partire da Beccaria*, in «Materiali per una storia della cultura giuridica», 1/2015, p. 163.

⁶⁸ QUATTROCOLO, *Intelligenza artificiale e giustizia*, cit., p. 9.

⁶⁹ DI GIOVINE, *Il Judge-bot e le sequenze giuridiche in materia penale*, cit., p. 952.

⁷⁰ NIEVA-FENOLL, *Intelligenza artificiale e processo*, cit., pp. 32-33.

⁷¹ K. BRATING, J.C. LESTER, C.B. CALLAWAY, *Automating Judicial Document Drafting: A Discourse-Based Approach*, in «Artificial Intelligence and Law», 1998, p. 114.

⁷² NIEVA-FENOLL, *Intelligenza artificiale e processo*, cit., p. 21.

ingessatura ermeneutica⁷³, con una conseguente eterogenesi dei fini: l'intelligenza artificiale dovrebbe contribuire all'evoluzione del sistema, non già bloccarlo con 'pastroie intelligenti'.

Accanto a chi esalta il ruolo del giudice assistito dalla macchina, del resto, non si manca di osservare che nel futuro i giudici e le loro Corti risulteranno sempre meno necessari⁷⁴ e che gli avvocati, dal canto loro, diventeranno una razza in via di estinzione⁷⁵.

A ben vedere, le prospettive più interessanti in materia di rapporti tra giudice e macchina sono quelle che non si limitano ad auspicare un affiancamento dell'uomo da parte dell'algoritmo né quelle che insistono sulle sempre più ampie prospettive di un giudice-robot 'autosufficiente', ma quelle secondo cui l'effetto dell'intelligenza artificiale applicata al processo penale possa influire sulla selezione degli operatori del diritto: non solo dei giudici, ma anche degli avvocati.

L'impiego dell'intelligenza artificiale renderà forse necessario un minor numero di giudici, ma quel che è certo è che dovrà trattarsi di professionisti altamente specializzati, capaci di comprendere e mettere in relazione i risultati forniti dalla macchina⁷⁶ e capaci di apportare un reale valore aggiunto nella soluzione di quei 'casi difficili' che alla macchina non possono integralmente delegarsi. Senza contare, vale la pena precisarlo, che, almeno di regola, l'etichetta di 'caso facile' o 'caso difficile' può essere attribuita solo *ex post*, rendendosi quindi necessario un filtro preventivo non sempre così scontato e automatico.

Il miglior antidoto a un possibile irrigidimento del sistema, poi, è la presenza, in sede di impugnazione e soprattutto in Corte di cassazione, di giudici capaci di non accontentarsi sempre e comunque della immutabilità del risultato fornito dalla macchina: «il criterio principale di selezione dei membri degli organi di ultima istanza dovrà essere la loro creatività»⁷⁷.

Analoghe considerazioni, *mutatis mutandis*, valgono per gli avvocati. L'avvocato dovrà conoscere il funzionamento della macchina, senza contare

⁷³ C. PARODI, V. SELLAROLI, *Sistema penale e intelligenza artificiale: molte speranze e qualche equivoco*, in «Diritto penale contemporaneo», 6/2019, p. 51.

⁷⁴ BEN-HARI, FRISCH, LAZOVKI, ELDAN, GREENBAUM, *Artificial Intelligence in the Practice of Law*, cit., p. 35.

⁷⁵ BEN-HARI, FRISCH, LAZOVKI, ELDAN, GREENBAUM, *Artificial Intelligence in the Practice of Law*, cit., p. 37. Per l'impatto dell'intelligenza artificiale sull'esercizio della professione forense v. anche P. MORO, *Intelligenza artificiale e professioni legali. La questione del metodo*, in «Journal of Ethics and Legal Technologies», 2019, pp. 24 ss.

⁷⁶ NIEVA-FENOLL, *Intelligenza artificiale e processo*, cit., p. 21.

⁷⁷ NIEVA-FENOLL, *Intelligenza artificiale e processo*, cit., p. 135.

che gli algoritmi predittivi, come precisato, rappresentano un valido strumento anche per un rafforzamento del diritto di difesa. L'avvocato, però, dovrà essere in grado di elaborare strategie capaci di 'resistere all'algoritmo' quando, per esempio, lo impongano le circostanze del caso concreto o quando la macchina dovesse adagiarsi nella stanca ripetizione di soluzioni anacronistiche e/o sistematicamente irragionevoli.

Si tratta di uno scenario che, prima ancora che proiettarsi verso il futuro, dovrebbe condurre ad aprire gli occhi sul presente, a partire dagli attuali meccanismi di reclutamento dei professionisti del diritto. Si è scritto, per esempio, che le Università dovrebbero tenere il passo con l'intelligenza artificiale e che quindi bisognerebbe insegnare fin da subito ai futuri giuristi come rapportarsi con gli algoritmi. Se, tuttavia, per fare davvero la differenza è necessario valorizzare la creatività di giudici e avvocati, è anzitutto quella creatività che le aule universitarie dovrebbero contribuire a lasciar emergere⁷⁸.

Potrebbe forse ipotizzarsi che l'obsolescenza delle macchine lascerà in breve tempo il posto a quella degli esseri umani e che le radicali trasformazioni già in corso nel mondo del lavoro sconfineranno nell'approdo definitivo secondo cui 'le persone non servono'⁷⁹. Visto che, tuttavia, allo stato le persone servono ancora, sarebbe affrettata la pretesa di modellare gli uomini sulle macchine (dietro l'apparente pretesa di modellare le macchine sugli uomini), alimentando una spirale che renderebbe l'uomo sempre più fungibile: il giurista, per contro, andrebbe formato prima e selezionato poi proprio valorizzando quelle capacità e quelle attitudini che gli consentano di usare la macchina come strumento e di differenziarsi da un algoritmo, per quanto evoluto lo stesso possa risultare.

Le frontiere aperte dall'incalzante e vertiginoso sviluppo degli algoritmi intelligenti paiono rispolverare il sogno di ogni giurista, filosofo o pensatore che abbia inseguito la chimera di un diritto calcolabile: una macchina è in grado di calcolare il diritto, a conferma del fatto che la calcolabilità non si esaurisce né nella sovrastruttura di cui l'ideologia borghese si serve per 'giustificare' il capitalismo né nell'ingenua credenza di chi si rifugia in

⁷⁸ Cfr. BEN-HARI, FRISCH, LAZOVKI, ELKAN, GREENBAUM, *Artificial Intelligence in the Practice of Law*, cit., p. 39: «Law schools will change dramatically, not least because we will need fewer lawyers. Moreover, the nature of legal learning will change to include subjects that are not taught in law schools today—creativity, understanding of statistics, big data analysis, and more».

⁷⁹ Il riferimento è, evidentemente, al titolo di KAPLAN, *Le persone non servono. Lavoro e ricchezza nell'epoca dell'intelligenza artificiale*, Luiss University Press, Roma, 2016. V. anche K. SCHWAB, *La quarta rivoluzione industriale*, Franco Angeli, Milano, 2016.

vetuste mitologie giuridiche per impedire che il diritto si prenda la sua definitiva rivincita sulla legge.

La posta in gioco, tuttavia, è molto alta. Il mutamento giurisprudenziale, se non sconfinava nel *caos* ingestibile di una giurisprudenza che fatica a trovare una sua (sia pur momentanea) stabilizzazione, rappresenta un fenomeno del tutto fisiologico che, anzi, assicura un 'automatico' adattamento dell'ordinamento giuridico ai cambiamenti sociali, culturali e tecnologici. Se, per ipotesi, si raggiungesse il risultato di un 'diritto calcolabile' in senso stretto, la contropartita sarebbe quella di una sostanziale stagnazione del diritto: la giurisprudenza si ripeterebbe pressoché identica a se stessa, con il rischio di condurre alla svalutazione delle peculiarità dei singoli casi e di pervenire a formule standardizzate di gestione delle questioni giuridiche.

Si tratta chiaramente di un rischio inversamente proporzionale al ruolo che si intende riservare alla 'mente umana' rispetto all'attività svolta dall'algoritmo e che quindi tende ad attenuarsi in quei modelli che immaginano un supporto dell'intelligenza artificiale nei settori in cui l'intervento della stessa potrebbe contribuire a ridurre il margine di errore e/o ad assicurare un più elevato grado di oggettività di giudizi che altrimenti resterebbero affidati a un imperscrutabile soggettivismo. A diverse conclusioni, invece, dovrebbe giungersi qualora si immaginasse di affrancare alcuni giudizi dall'intervento del 'giudice uomo', addirittura ipotizzando un superamento dell'obbligo di motivazione e del sistema delle impugnazioni.

A ciò si aggiunga che, qualsiasi sia l'apporto degli algoritmi intelligenti che possa immaginarsi in riferimento al diritto e alla giustizia penale, dovrebbe pur sempre porsi la questione di un necessario controllo dei parametri di programmazione: fino a quando si tratta di strumenti gestiti 'tecnicamente' da soggetti privati, si porrà, in maniera necessaria, un problema di tenuta di alcuni dei principi fondamentali su cui si fonda lo Stato di diritto.

5. Bibliografia

- ALETRAS N., TSARAPATSANIS D., PREOTIUC-PIETRO D., LAMPOS V., *Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective*, in «PeerJ Computer Science», 2016.
- BARBERIS M., *Cosa resta del sillogismo giudiziale? Riflessioni a partire da Beccaria*, in «Materiali per una storia della cultura giuridica», 1/2015.

- BARTOLI R., *Dubbio e certezza del diritto penale*, in «Quaderno di storia del penale e della giustizia», 2/2020.
- BASILE F., *Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*, in «Diritto penale e Uomo», 29 settembre 2019.
- BASKIN-SOMMERS A., *La neuroscienza dovrebbe migliorare il sistema carcerario, non cercare di dimostrare l'innocenza*, in «Diritto penale e uomo», 29 gennaio 2020.
- BEN-HARI D., FRISCH Y., LAZOVKI A., ELKAN U., GREENBAUM D., *Artificial Intelligence in the Practice of Law: an Analysis and Proof of Concept Experiment*, in «23 Richmond Journal of Law & Technology», 2017.
- BORSARI R., *Intelligenza Artificiale e responsabilità penale: prime considerazioni*, in «MediaLaws», 3/2019.
- BRATING K., LESTER J.C., CALLAWAY C.B., *Automating Judicial Document Drafting: A Discourse-Based Approach*, in «Artificial Intelligence and Law», 1998.
- CALAMANDREI P., *Processo e democrazia*, in *Opere giuridiche*, I, Roma TrE-Press, Roma, 2019.
- CAPPELLINI A., *Machina delinquere non potest? Brevi appunti su intelligenza artificiale e responsabilità penale*, in «Discrimen», 27 marzo 2019.
- CARRER S., *Se l'amicus curiae è un algoritmo: il chiacchierato caso Loomis alla Corte Suprema del Wisconsin*, in «Giurisprudenza Penale Web», 4/2019.
- CEREZO MIR J., *Ontologismo e normativismo nel finalismo degli anni Cinquanta*, in *Significato e prospettive del finalismo nell'esperienza giuspenalistica*, a cura di S. Moccia, Napoli, Edizioni Scientifiche Italiane, 2002.
- COLLICA M.T., *Il riconoscimento del ruolo delle neuroscienze nel giudizio di imputabilità*, in «Diritto penale contemporaneo», 15 febbraio 2012.
- CORNACCHIA L., *Ein unausrottbares Verständnis? L'eredità del finalismo nel dibattito penalistico italiano*, in «Criminalia», 2013.
- DI GIOVINE O., *Il Judge-bot e le sequenze giuridiche in materia penale (intelligenza artificiale e stabilizzazione del precedente)*, in «Cassazione penale», 3/2020.
- DI GIOVINE O., *Ripensare il diritto penale attraverso le (neuro-)scienze?*, Giappichelli, Torino, 2019.
- DI GIOVINE O., voce *Neuroscienze*, in «Enciclopedia del diritto», Annali VII, 2014.
- DONINI M., *La riserva di codice (art. 3-bis c.p.) tra democrazia normante e principi costituzionali. Apertura di un dibattito*, in «La Legislazione Penale», 20 novembre 2018, pp. 1 ss.

- FERRAJOLI L., *Diritto e ragione. Teoria del garantismo penale*, Bari, Laterza, 2004.
- FORZA A., *Le neuroscienze entrano nel processo penale*, in «Rivista penale», 2010.
- GRANDI C., *Neuroscienze e responsabilità penale. Nuove soluzioni per problemi antichi?*, Giappichelli, Torino, 2016.
- GREENE J., COHEN J., *For the Law, Neuroscience changes Nothing and Everything*, in «The Royal Society», 2004.
- ITZCOVICH G., *Il diritto come macchina. Razionalizzazione del diritto e forma giuridica in Max Weber*, in «Materiali per una storia della cultura giuridica», 31/2, 2001.
- KAPLAN J., *Intelligenza artificiale. Guida al futuro prossimo*, Luiss University Press, Roma, 2017.
- KAPLAN J., *Le persone non servono. Lavoro e ricchezza nell'epoca dell'intelligenza artificiale*, Luiss University Press, Roma, 2016.
- KING T., AGGARWAL M., TADDEO M., FLORIDI L., *Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions*, in *Science and Engineering Ethics Ethics*, 2019.
- KREMER J., *The End of Freedom in Public Places? Privacy problems arising from surveillance of the European public space*, University of Helsinki, 2017, in <<https://helda.helsinki.fi>>.
- LUPÀRIA L., *Prova giudiziaria e ragionamento artificiale: alcune possibili chiavi di lettura*, in *Il concetto di prova alla luce dell'intelligenza artificiale*, a cura di J. Sallantin, J.J. Szczeciniarz, Giuffrè, Milano, 2005.
- MASSARO A., *Determinatezza della norma penale e calcolabilità giuridica*, Editoriale Scientifica, Napoli, 2020.
- MASSARO A., *I reati 'contro gli animali' tra aspirazioni zoocentriche e ineliminabili residui antropocentrici*, in «Cultura e Diritti», 1-2/2018.
- MC CARTHY J., MINSKY M.L., ROCHESTER N., SHANNON C.E., *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, 1955, in <<http://www-formal.stanford.edu>>.
- MIR J.C., *Ontologismo y normativismo en el finalismo de los años cincuenta*, in «Revista de derecho penal y criminología», 2003.
- MIR PUIG S., *Límites del normativismo en derecho penal*, in «Revista electrónica de ciencia penal y criminología», 2005.
- MORO P., *Intelligenza artificiale e professioni legali. La questione del metodo*, in «Journal of Ethics and Legal Technologies», 2019.
- MORSE S., *New Neuroscience, Old Problems: Legal Implications of Brain Science*, in «Cerebrum», 2004.

- NIEVA-FENOLL J., *Intelligenza artificiale e processo*, Giappichelli, Torino, 2019.
- NISCO A., *Neokantismo e scienza del diritto penale. Sull'involuzione autoritaria del pensiero penalistico tedesco nel primo Novecento*, Giappichelli, Torino, 2019.
- NISSAN E., *Digital Technologies and Artificial Intelligence's Present and Foreseeable Impact on Lawyering, Judging, Policing and Law Enforcement*, in «AI & Society», 2017.
- O'DONNELL R.M., *Challenging Racist Predictive Policing Algorithms under the Equal Protection Clause*, in «New York University Law Review», 2019.
- PELLICCIA R., *Polizia predittiva: il futuro della prevenzione criminale?*, in <<https://www.cyberlaws.it>>, 9 maggio 2019.
- PETROCELLI B., *La colpevolezza*, CEDAM, Padova, 1955.
- PLANCK M., *Legge di causalità e libero arbitrio*, in *La conoscenza del mondo fisico*, Torino, Einaudi, 1949.
- POGGI F., *Il diritto meccanico. La metafora del diritto come macchina e i suoi limiti*, in «Diritto e questioni pubbliche», 9/2009.
- PULITANÒ D., *Il diritto penale fra vincoli di realtà e sapere scientifico*, in «Rivista italiana di diritto e procedura penale», 3/2006.
- QUATTROCOLO S., *Intelligenza artificiale e giustizia: nella cornice della Carta etica europea, gli spunti per un'urgente discussione tra scienze penali e informatiche*, in «La Legislazione penale», 18 dicembre 2018.
- ROCCO ART., *Il problema e il metodo della scienza del diritto penale*, in *Opere giuridiche*, Vol. III, Società editrice del Foro Italiano, Roma, 1933.
- ROXIN C., *Politica criminale e sistema del diritto penale*, Guida Editori, Napoli, 1986.
- RULLI E., *Giustizia predittiva, intelligenza artificiale e modelli probabilistici*, in «Analisi giuridica dell'Economia», 2/2018.
- SBRICCOLI M., *Caratteri originari e tratti permanenti del sistema penale italiano (1860-1990)*, in *Storia del diritto penale e della giustizia. Scritti editi e inediti (1972-2007)*, Giuffrè, Milano, 2009.
- SCARANO L., *Libera volontà e libero arbitrio nel diritto penale*, Giuffrè, Milano, 1937.
- SCHWAB K., *La quarta rivoluzione industriale*, Franco Angeli, Milano, 2016.
- SEARLE J.R., *Minds, Brains and Programs*, in *Behavioral and Brain Sciences*, 1980.
- TATA C., *The Application of Judicial Intelligence and 'Rules' to Systems Supporting Discretionary Judicial Decision-Making*, in «Artificial Intelligence and Law», 1998.

- TURING A.M., *Computing machinery and intelligence*, in «Mind», 1950.
- TUZET G., *Il diritto non è una macchina*, in «Diritto e questioni pubbliche», 9/2009.
- TUZET G., *Sul possibile moto della macchina*, in «Diritto e questioni pubbliche», 9/2009.
- VELLUZZI V., *La preziosa ingenuità: Beccaria, lo spirito della legge e il sillogismo giudiziale*, in «Rivista internazionale di filosofia del diritto», 2014.
- WEBER M., *Economia e società*, III, Edizioni di Comunità, Roma, 1980.
- WELZEL H., *Naturalismus und Wertphilosophie in Strafrecht* (1935), in *Abhandlungen zum Strafrecht und zur Rechtsphilosophie*, Berlino-New York, de Gruyter, 1975, 29 ss.

SECCIÓN II

LA INTELIGENCIA ARTIFICIAL
ENTRE DERECHO Y PROCESO PENAL

L'INTELLIGENZA ARTIFICIALE
TRA DIRITTO E PROCESSO PENALE

CAPÍTULO III

Implicaciones de la inteligencia artificial en el ámbito del Derecho Penal

David Lorenzo Morillas Fernández

SUMARIO: 1. Enfoque y estado de la cuestión – 2. Situación actual de la responsabilidad penal por un hecho delictivo cometido por inteligencia artificial – 2.1. ¿Puede considerarse a la inteligencia artificial autora única del delito? – 2.2. ¿Quién debe responder penalmente entonces? – 3. Pronunciamientos jurisprudenciales expresos – 4. Conclusiones – 5. Bibliografía.

1. Enfoque y estado de la cuestión

Aunque a muchos les pese, la forma más sencilla y, en mi opinión, recomendada de comenzar cualquier trabajo, no debe ser otra que contextualizar el ámbito en la que el autor pretende moverse y eso, en determinadas cuestiones, como la que me ocupa, no es sencillo dada la complejidad y aristas de la temática a tratar, máxime cuando el potencial lector no tiene por qué entender o saber inicialmente el concreto ámbito o las especificidades de la temática a abordar. Dar por supuestos determinados conceptos, contenidos e implicaciones resulta adecuado cuando el destinatario de la información es un público especializado en la materia si bien lo aquí buscado choca diametralmente con semejante afirmación, reuniendo el actual artículo la simple y mera finalidad de resultar accesible a juristas en general, para lo cual se van a ir confrontando y exponiendo ideas desde abajo para ir progresivamente avanzando en intensidad hasta alcanzar un nivel medio.

Así las cosas, la mejor forma de ejemplificar la inteligencia artificial, por mor de que a los más puristas en la materia le resulte una vaguedad o eufemismo, debe partir necesariamente de ejemplos concretos comprensibles para el conjunto de la ciudadanía. En este sentido, me permito traer a colación, dos claros ejemplos conocidos – creo – por todos: i) la serie de televisión de los años ochenta, «El coche fantástico», donde un hombre luchaba contra el

crimen junto con un coche que poseía habilidades y destrezas propias que le permitían realizar funciones de manera completamente independiente y siempre bajo un determinado prisma: proteger la vida humana, siendo capaz de desarrollar conductas ilícitas -básicamente daños- por sí mismo siempre que el fin fuera la protección de la vida¹; o ii) la película «Terminator» – y sus diversas secuelas y precuelas – donde un sistema de inteligencia artificial, llamado Skynet, lidera un ejército de máquinas poseyendo la capacidad para controlar el arsenal militar de los Estados Unidos con independencia de la voluntad humana. En consecuencia, si se tomaran de ejemplo determinadas conductas visionadas en la anterior serie y saga, surgen preguntas relacionadas con la responsabilidad jurídica, ya sea penal o extra-penal, como ¿quién respondería cuándo Michael Knight realiza una persecución y conduce a más de 200 kms/h? Parece claro que el protagonista cometería un delito contra la seguridad vial del artículo 379.1 del Código Penal; pero ¿y si fuera KITT, el coche, quien hace uso de su autonomía para realizar una persecución sin conductor a más de 200 kms/h? ¿a quién atribuir la responsabilidad penal cuando, en alguna de las secuencias más famosas de la serie, KITT realiza una conducción autónoma y, haciendo uso de su turbo, salta y atraviesa una pared de ladrillo causando unos evidentes daños en la propiedad o, con su salto, destroza un enorme espejo que está siendo trasladado en un camión de grandes dimensiones? ¿Cambiaría la calificación de los hechos si Michael pidiera ayuda por su transmisor por ver peligrar su vida o salud y, como quiera que KITT está diseñado para preservar la vida, desarrolla tales comportamientos? ¿Podría aplicarse una causa de justificación a la conducta ilícita de una inteligencia artificial? Del mismo modo, yendo a parámetros más extremos, ¿de quién sería la responsabilidad penal si un humano crea un robot capaz de matar? Parece claro, al igual que en el primer ejemplo anterior, que de la persona, siendo la máquina simplemente el instrumento de ejecución del delito pero ¿y si fuera otra máquina -por ejemplo, Skynet, por seguir la representación- quien, de conformidad con su capacidad de autoaprendizaje, crea un robot (Terminator) que mata a personas (como en las distintas secuelas de la saga)? ¿Tendría relevancia el hecho de que la máquina no hace otra cosa que cumplir con el objetivo para el cual ha sido programada -acabar con la vida de John Connor-?

¹ Conviene llamar la atención en este punto en el hecho de que, al igual que sucede en la realidad, todo invento orientado a la mejora de la sociedad o de su calidad de vida, más pronto que tarde presenta también implicaciones ilícitas y así, en la citada serie, existía otro modelo -KARR- carente del sentido de protección de la vida y antítesis de KITT, por lo que el desarrollo de hechos ilícitos, representado por sentimientos humanos como la ira o la venganza, operaban en él.

Los ejemplos anteriores -y otros muchos- que en la década de los ochenta podían parecer auténtica ciencia-ficción no dejan de ser una realidad completa, con las lógicas limitaciones propios de la ficción en pleno año 2022, donde la sociedad se mueve al ritmo que le marca la tecnología en multitud de ámbitos. Piénsese, a tal efecto, en la propia conducción de vehículos, donde no sólo ya hay sistemas automáticos para aparcar, frenar automáticamente en caso de emergencia, detectar la salida de la calzada y proceder inmediatamente a la protección de los ocupantes tensando los cinturones de seguridad y activando un sistema de absorción de impactos en los asientos con el propósito de proteger la columna vertebral de los usuarios (...) sino igualmente vehículos conducidos autónomamente -vid., por ejemplo, el modelo Tesla, que saca el coche del garaje y lo deja en la calzada, o, yendo más allá, permite la conducción autónoma por vía pública²-, una ingeniería que cada día va a más y recibe el nombre de vehículo autónomo o robótico. Pero es más, ¿qué sucede en el ámbito de la medicina? Los avances de la inteligencia artificial en el diagnóstico y tratamiento de enfermedades son una constante diaria, lo que permite una mejora de la calidad de vida de la población. ¿Y su utilización con fines bélicos? No se trata ya de robots desactivadores de bombas que funcionan con control remoto sino de máquinas con capacidad de lucha autónoma -robots de combate con inteligencia artificial; vehículos no tripulados para transportar armas pesadas, donde la inteligencia artificial valora situaciones de combate y toma decisiones autónomas; el URAN-9, vehículo de combate destinado, entre otras funciones, a destruir otros vehículos blindados- (...).

Con semejantes planteamientos no es de extrañar que, como refirió Pérez Arias, en la era de la ciber-civilización, la sociedad muestre su preocupación por conocer el nivel de protección que el Derecho confiere a víctimas de delitos cometidos a través de un medio tecnológico. Y dentro de esta protección, es aún más lógico que la sociedad se cuestione (y con mayor motivo los intérpretes) si el Derecho Penal, ha sido debidamente actualizado y modernizado hasta el punto de otorgar protección frente a determinados delitos que discurren por medios, y fines, por completo desconocidos, hasta fechas muy recientes³.

² Aun cuando hoy día no está recomendado ya que la activación de esta función exige la supervisión por parte del conductor -exigencia interesante por parte de la compañía que, llegado el caso, pudiera eximirle de responsabilidad penal en caso de atropello, como se planteará más adelante-, no son pocos los vídeos que pueden encontrarse en redes sociales donde los potenciales conductores del vehículo graban la conducción autónoma del Tesla por carretera.

³ J. PÉREZ ARIAS, *Cibercriminalidad: hacia la nueva realidad -virtual- del Derecho Penal*,

Como ha podido observarse, hoy día, la inteligencia artificial juega un papel nuclear en el devenir cotidiano de la ciudadanía con las consiguientes implicaciones jurídicas, entre otras muchas. Ahora bien, si se observa la casuística anterior, no todo responde en términos absolutos a inteligencia artificial ya que, aún dentro de ella, se establecen categorías. En este sentido, Salvadori ha distinguido entre inteligencia artificial y aprendizaje automático como parámetros delimitadores de los llamados agentes artificiales -utilizados para identificar los software- y sistemas de inteligencia artificial -capaces de desarrollar funciones para otro agente, ya sea humano o artificial-, los cuales están apoyando o sustituyendo a las personas en la realización de actividades, siendo en este ámbito donde cabría incluir los ejemplos anteriores vinculados al sector médico-quirúrgico, asistencia a enfermos o ancianos, coches autónomos o algoritmos de justicia predictiva; esto es, no se habla ya de meros instrumentos u objetos sino de verdaderos agentes artificiales capaces de interactuar con la realidad circulante, con otras máquinas inteligentes y seres humanos, siendo capaces de tomar en muy poco tiempo decisiones complejas de indudable relevancia social⁴.

Cuatrecasas Monforte ha clasificado a la inteligencia artificial, entre otros criterios, según el procedimiento en virtud del cual lleva a cabo el aprendizaje, distinguiendo entre aprendizaje supervisado, no supervisado y por refuerzo: i) el supervisado se caracterizaría por la presencia de un programador humano que introduce datos etiquetados y establece distintas variables de entrada en función de las cuales deberá asignarse una etiqueta de salida, mostrando el algoritmo las conclusiones a las que debe llegar; ii) no supervisado, se identificaría por la ausencia de un programador que introduzca datos y parámetros de identificación, siendo el propio algoritmo el que detecta la repetición de patrones y lleva a cabo un auto entrenamiento y aprendizaje, sacando sus propias conclusiones sin indicación externa alguna; y iii) por refuerzo, se basaría en la máxima prueba-error, existiendo una interacción permanente con el sistema, mediante el envío de mensajes notificaciones que le informan de si es correcto o incorrecto lo que va realizando, pudiendo extraer conclusiones tras múltiples valoraciones y determinar qué caminos son los que con más probabilidad llevan al éxito⁵.

La velocidad de los avances tecnológicas es tan creciente e intensa que no es de extrañar la afirmación recogida en la Resolución del Parlamento

en «Revista Electrónica de Doctrina y Jurisprudencia», n. 26, 2021, pp. 175 y 176.

⁴ I. SALVADORI, *Agentes artificiales, opacidad tecnológica y distribución de la responsabilidad penal*, en «Cuadernos de Política Criminal», n. 133, 2021, pp. 138 y 139.

⁵ C. CUATRECASAS MONFORTE, *La Inteligencia Artificial como herramienta de investigación criminal*, La Ley, Madrid, 2022, pp. 32-36.

Europeo, de 3 de mayo de 2022, sobre la inteligencia artificial en la era digital (2020/2266(INI)), la cual habla de que estamos ante una cuarta revolución industrial, caracterizándose ésta, en comparación con las anteriores, iniciadas por la introducción del vapor, la electricidad y luego los ordenadores, porque la cuarta oleada extrae su energía de la abundancia de datos combinada con potentes algoritmos y capacidad informática, expansiva a nivel mundial, con una rápida convergencia y el enorme impacto de los nuevos avances tecnológicos en los países, las economías, las sociedades, las relaciones internacionales y el medio ambiente, lo cual supone un cambio radical de enorme magnitud que repercute diversamente en las distintas partes de la sociedad, en función de sus objetivos, ubicación geográfica o contexto socioeconómico. Ahora bien, semejante transformación digital debe hacerse siempre dentro del pleno respeto de los derechos fundamentales y con el único propósito de que las tecnologías digitales sirvan a la humanidad.

Sea como fuere, conviene tener unas cautelas mínimas a la hora de abordar la problemática en el uso y la regulación de sistemas de inteligencia artificial. En este sentido, la Resolución del Parlamento Europeo, de 6 de octubre de 2021, sobre la inteligencia artificial en el Derecho Penal y su utilización por las autoridades policiales y judiciales en asuntos penales (2020/2016(INI)), ha establecido unos considerandos muy útiles que delimitan los parámetros mínimos a tener en cuenta al respecto y que conviene resaltar:

- i) Las tecnologías digitales en general y la proliferación del tratamiento y el análisis de datos por medio de inteligencia artificial ofrecen enormes riesgos para los derechos fundamentales y las democracias basadas en el Estado de Derecho ya que la inteligencia artificial no debe considerarse un fin en sí misma sino un instrumento al servicio de las personas, con el objetivo primordial de aumentar el bienestar humano, las capacidades humanas y la seguridad.
- ii) No existen programas que puedan igualar a la flexibilidad humana en ámbitos más amplio o tareas que exijan la comprensión del contexto o un análisis crítico. Algunas aplicaciones de inteligencia artificial han alcanzado los niveles de rendimiento de expertos y profesionales humanos – por ejemplo, la tecnología jurídica o ‘legal tech’ –, pudiendo ofrecer resultados a una velocidad muchísimo mayor y una escala más amplia.

- iii) Algunos países hacen uso de aplicaciones o sistemas de inteligencia artificial con fines coercitivos y judiciales debido a la ausencia de regulación y las diferencias normativas. Eso hace que el creciente uso de la inteligencia artificial en el ámbito del Derecho Penal se base en particular en las promesas de que reduciría determinados tipos de delincuencia y daría lugar a decisiones más objetivas. Sin embargo, estas promesas no siempre se cumplen.
- iv) Para que los sistemas de inteligencia artificial resulten fiables, deben diseñarse con el propósito de proteger y beneficiar a todos los miembros de la sociedad, particularmente a las personas vulnerables y marginadas, no debiendo ser, en ningún caso, discriminatorias y tomar decisiones explicables y transparentes, respetando siempre la autonomía humana y los derechos fundamentales.
- v) El uso de aplicaciones de inteligencia artificial ofrece grandes oportunidades para mejorar los métodos de trabajo de las autoridades policiales y judiciales, aumentando la eficacia en la lucha contra determinados tipos de delitos, especialmente los de naturaleza financiera -blanqueo de capitales y financiación del terrorismo-, abusos sexuales, explotación sexual en línea y determinados tipos de ciberdelincuencia. No obstante, también pueden entrañar al mismo tiempo riesgos significativos para los derechos fundamentales de las personas.
- vi) Se hace imprescindible contar con un modelo claro para establecer la responsabilidad penal por los posibles efectos nocivos de los sistemas de inteligencia artificial.

Así las cosas, para concretar el conjunto de considerandos anteriores, resulta fundamental partir de una definición concreta de lo que se entiende por inteligencia artificial, algo que no va a resultar sencillo ya que existen multitud de aristas que dificultan una concreción terminológica unitaria al respecto y así, como refirió Miró Llinares, la premisa inicial resulta por sí compleja debido a las disquisiciones filosóficas no sólo respecto del concepto inteligencia sino porque además según el campo en el que nos movamos los objetivos definitorios serán igualmente modulables⁶. En términos muy parecidos se ha manifestado De la Cuesta Aguado al señalar que el concepto

⁶ F. MIRÓ LLINARES, *Inteligencia artificial y justicia penal: más allá de los resultados lesivos causados por robots*, en «Revista de Derecho Penal y Criminología», n. 20, 2018, pp. 91-92.

de inteligencia, como característica propia del género humano, dista mucho de ser pacífico, tanto en cuanto a su contenido (¿qué es la inteligencia?) como a su extensión (¿solo son inteligentes los humanos?)⁷. Ahora bien, tal y como ha señalado Palma, la cuestión de la inteligencia artificial y la neurociencia se complica en el instante en el que una decisión inteligente puede ser calculada como un algoritmo e imitada por procesos matemáticos⁸.

Sea como fuere, son muchas y muy variadas las definiciones existentes al respecto, presentándose a continuación algunas, de manera meramente orientativa y próximas al ámbito del Derecho Penal:

- El enfoque inicial de la cuestión lo trae perfectamente a colación Valls Prieto cuando refiere que la inteligencia artificial es considerada como la capacidad de la máquina de pensar, de razonar y actuar con inteligencia; esto es, se trataría de un sistema que puede realizar tareas de manera idéntica a como la realizaría un humano en determinados contextos, si bien existen supuestos en los que semejante tarea designada obtiene un rendimiento mejor que en el caso que la llevaran a efecto personas, tal y como puede ser el ejemplo de la búsqueda de información, reconocimiento de tumores en radiografías o jugando al ajedrez. De acuerdo con lo anterior, se podrían diferenciar dos tipos de inteligencia artificial: una, basada en el conocimiento; y otra, en datos. En definitiva, las máquinas tienen la posibilidad de realizar comportamientos automáticos inteligentes en la medida en que pueden realizarlos de forma parecida a los humanos dentro de un contexto concreto en el que tienen que elegir entre diferentes posibilidades y que, en determinadas facetas de la vida, aportan soluciones útiles para las personas. Ahora bien, no conviene olvidar, siendo este extremo especialmente relevante para los aspectos que referiré posteriormente, que la inteligencia artificial es una herramienta creada por las personas, lo cual lleva consigo que, aunque pueda crear cosas -entre ellas más inteligencias artificiales-, siempre va a ser mediante unos patrones impuestos por los humanos, siendo éste el origen de la actividad que desarrolle y, por tanto, la máquina no va a realizar nada que no sea predeterminado por un humano, resultando esta cuestión nuclear al objeto de determinar la responsabilidad penal operable al respecto⁹.

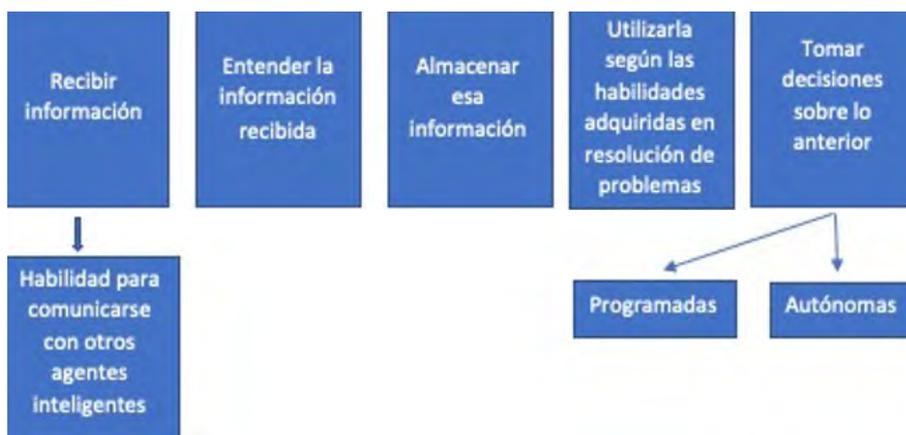
⁷ P. DE LA CUESTA AGUADO, *Inteligencia artificial y Derecho Penal*, en «Revista Penal México», n. 16 y 17, 2019-2020, p. 52.

⁸ M^a F. PALMA, *El Derecho como universo de sentido para la interfaz cerebro e inteligencia artificial*, en «Derecho Penal y comportamiento humano», dirigido por E. Demetrio Crespo, Tirant lo Blanch, Valencia, 2022, p. 176.

⁹ J. VALLS PRIETO, *Sobre la responsabilidad penal por la utilización de sistemas inteligentes*,

Pero claro, como ya se ha apuntado anteriormente, ¿qué cabe entender por inteligencia? Comparto totalmente lo expresado por De la Cuesta Aguado cuando afirma que la inteligencia es una característica propia del ser humano y no existe una significación concreta al respecto, no siendo tampoco éste el momento para profundizar en cuestiones vinculadas a ella, máxime por los múltiples enfoques científicos que podrían operar en la referida cuestión. Es por ello que, siguiendo los postulados manifestados por la citada autora, y con un mero carácter introductorio o ilustrativo a los efectos de este trabajo, voy a identificar la inteligencia de acuerdo con las capacidades que describió: i) capacidad para recibir información; ii) capacidad para entender la información recibida -entendiendo por tal aquella capaz de hacer utilizable por el agente en el contexto y en función de objetivos la información recibida; iii) capacidad para almacenar la información recibida; iv) capacidad para utilizar la información obtenida y procesada mediante las habilidades anteriores en la resolución de problemas; y v) capacidad para tomar decisiones¹⁰.

ESQUEMA DE LAS CAPACIDADES BASE PARA DEFINIR LA INTELIGENCIA¹¹



en «Revista Electrónica de Ciencia Penal y Criminología», 24-27, 2022, pp. 5 y 9.

¹⁰ P. DE LA CUESTA AGUADO, *Inteligencia artificial y Derecho Penal*, cit., p. 52.

¹¹ Elaboración propia sobre los postulados descritos por DE LA CUESTA AGUADO.

Así las cosas, como manifestaron en su día Corvalán y Ciraudo, con la inteligencia artificial se busca que las tecnologías permitan que los sistemas computacionales adquieran autodependencia, reconfiguración autoadaptativa, negociación inteligente, comportamiento de cooperación, supervivencia con intervención humana reducida, entre otros rasgos. La inteligencia artificial no busca crear robots o computadoras inteligentes ya que resulta imposible reproducir el cerebro humano en máquinas, dada su complejidad y la ausencia de datos fiables que permitan determinar a ciencia cierta su funcionamiento integral, lo que se busca de ella es que realice exitosamente ciertas actividades sin necesidad de que estos sistemas se asemejen imiten o copien a un cerebro humano¹².

El error, entiende Cuatrecasas Monforte, radica en realizar una comparación directa entre ambas inteligencias, no siendo el mejor criterio para poder llegar a entender qué es la inteligencia artificial ya que, por un lado, las máquinas pueden realizar tareas que las personas no pueden en absoluto y, por otro, los humanos cuentan con una sensibilidad un sentido común y una capacidad de improvisación que las máquinas, al menos por el momento, no llegan a alcanzar. Así, la comparación entre la inteligencia humana y la artificial implicaría la equiparación de dos ideas con contenido no idéntico que podría inducir a error y/o confusión, por lo que la citada autora considera que la mejor opción es otorgar a la inteligencia artificial una definición autónoma¹³.

Por su parte Valls Prieto, quien ha realizado una muy acertada compilación de diversos documentos de la Unión Europea, refiere las conclusiones a las que llega el grupo de expertos de alto nivel en inteligencia artificial de la Comisión Europea y que, muy sucintamente, puede resumirse, a los efectos que ahora interesa, en lo siguiente: i) la reiterada dificultad para definir lo que es inteligencia, ya sea en humanos o en máquinas, por tratarse de un concepto vago, el cual no debe orientarse hacia la inteligencia en sí misma si no derivarlo al concepto de racionalidad, entendiéndose por ésta la habilidad de elegir la mejor opción para conseguir un determinado objetivo, dados determinados criterios y los recursos disponibles; ii) terminológicamente no es correcto hablar de inteligencia artificial per se sino que prefieren utilizar el término sistema de inteligencia artificial para definir a cualquier componente

¹² J.G. CORVALÁN, D. CIRAUDO, *Inteligencia artificial al servicio de la justicia penal, contravencional y de faltas. 'Prometea' en el ámbito de la ciudad de Buenos Aires*, en *Ciberdelincuencia II*, dirigido por D. Dupuy, Editorial B de f, Buenos Aires, 2018, p. 269.

¹³ CUATRECASAS MONFORTE, *La Inteligencia Artificial como herramienta de investigación criminal*, cit., p. 24.

basado en aquella, ya sea software o hardware¹⁴. Como puede observarse, el primer considerando rehúye de la utilización del término inteligencia en favor de racionalidad, lo cual traduce en una destreza para lograr un fin específico, resultando necesario unos criterios previamente establecidos al respecto, lo cual anularía la absoluta independencia de la máquina, si bien, como se verá posteriormente, pese a ser semejante postulado claro la situación concreta puede complicarse cuando el agente artificial es capaz de ‘aprender’ sobre semejantes parámetros y recurrir a otros criterios basados en el autoaprendizaje.

- El Grupo de Expertos de Alto Nivel en Inteligencia Artificial de la Unión Europea define la inteligencia artificial como aquellos sistemas de software (y posiblemente también de hardware) diseñados por humanos que, dado un objetivo complejo, actúan en la dimensión física o digital percibiendo su entorno mediante la adquisición de datos, interpretándolos de manera estructurada o no estructurada, razonando sobre el conocimiento, o procesando la información, derivada de ellos y decidiendo las mejores acciones a realizar para alcanzar el objetivo dado¹⁵.

Así pues, de la citada definición y sus implicaciones pueden comprobarse dos extremos fundamentales que conviene tener muy presentes a la hora de trasladar al ámbito jurídico: i) el origen del sistema de inteligencia artificial se encuentra embrionariamente ligada al ser humano; y ii) como el propio Grupo reconoce expresamente en el citado documento, la inteligencia artificial integra unos sistemas que muestran un comportamiento inteligente analizando su entorno y emprendiendo acciones -con cierto grado de autonomía- para alcanzar objetivos específicos¹⁶, lo cual, en mi opinión,

¹⁴ VALLS PRIETO, *Sobre la responsabilidad penal por la utilización de sistemas inteligentes*, cit., p. 8.

¹⁵ INDEPENDENT HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE, *A definition of AI: Main capabilities and disciplines*, European Commission, Bruselas, 2019, p. 6.

¹⁶ *Ibidem*, p. 1. Pero es más, incluso en los procesos de aprendizaje automático, como puede ser el caso del aprendizaje profundo, la red neuronal tiene varias capas entre la entrada y la salida que permiten aprender la relación global de entrada-salida en pasos sucesivos, logrando que el enfoque general sea más preciso y con menos necesidad de orientación humana; o el aprendizaje por refuerzo, donde se deja que el sistema de inteligencia artificial tome libremente sus decisiones, a lo largo del tiempo, y en cada decisión se le proporciona una señal de recompensa que le indica si ha sido una buena o una mala decisión con el propósito de que el objetivo del sistema, a lo largo del tiempo, sea maximizar la recompensa positiva recibida (*Ibidem*, p. 4). Sin embargo, tanto en uno como en otro -como en los demás procesos de aprendizaje autónomo- la intervención humana sigue siendo, en mayor o menor medida, necesaria, no pudiendo prescindirse de ella.

vuelve a resultar nuclear ya que la declaración realizada respecto a la ausencia de autonomía absoluta en los sistemas de inteligencia artificial, apuntando nuevamente hacia un cierto grado de autonomía pero sin ser capaz de alcanzar, inicialmente, una independencia total, resulta un postulado esencial al objeto de delimitar el ámbito de la responsabilidad penal derivada de actuaciones ilícitas, máxime cuando un sistema de inteligencia artificial no puede crear nada de la nada; esto es, toda su actuación está basada en datos obtenidos previamente. Miró Llinares ha considerado que la autonomía de una inteligencia artificial no provendría de la capacidad real de adaptar las decisiones a un contexto distinto de aquel para el que ha sido programada; esto es, o bien el aprendizaje por el que se toma la decisión no hubiera devenido de una acción humana, sino de la propia máquina que ha aprendido por sí misma o de otra máquina educadora anterior, o bien su comportamiento no dependiera de ese aprendizaje anterior y de las decisiones atribuidas previamente en la inteligencia artificial de algo ajeno a ello y propio del mismo sistema, siendo, a fecha de hoy, el ser humano quien determina la situación concreta a la que se ve expuesta la inteligencia artificial y su reacción ante una lista cerrada de estímulos¹⁷. Valls Prieto ha sido muy claro al afirmar que, hoy por hoy, no existe un sistema inteligente capaz de reaccionar como un humano, aunque pueda realizar tareas humanas con mayor eficacia¹⁸.

Salvadori ha establecido cuatro tipos de agentes artificiales según el grado/nivel de automatización y autonomía del ente:

- i) De primer nivel, los cuales operan automáticamente estando sujetos al control humano, tal y como podría ser el caso de los coches inteligentes disponibles en el mercado automovilístico actual, en los cuales una persona puede supervisar su funcionamiento e intervenir en cualquier momento para corregir la conducción.
- ii) De segundo nivel, cuyo funcionamiento depende de algoritmos deterministas, pudiendo resolver problemas ya previstos por sus programadores y desarrolladores o reaccionar de manera autónoma ante situaciones preestablecidas, tal y como puede ser el caso, por ejemplo, del software que reconoce señales de tráfico y permite que los vehículos autónomos se detengan ante un semáforo en rojo.

¹⁷ MIRÓ LLINARES, *El sistema penal ante la inteligencia artificial: actitudes, usos y retos*, en *Cibercrimen III*, dirigido por Dupuy D., Corvalán J. G., Editorial B de f, Buenos Aires, 2020, pp. 87 y 88.

¹⁸ VALLS PRIETO, *Sobre la responsabilidad penal por la utilización de sistemas inteligentes*, cit., p. 10.

- iii) De tercer nivel, los cuales incorporan algoritmos de aprendizaje autónomo, capaces de aprender, corregir y mejorar su comportamiento de acuerdo con la experiencia que adquieren. Aquí la intervención humana se mantiene en las fases de aprendizaje, pudiendo el programador hacer un seguimiento del citado proceso de aprendizaje e intervenir en los datos que se recogen y procesan.
- iv) El cuarto nivel o sistema multiagente puede interactuar con otros agentes y objetos adaptando autónomamente su comportamiento en función del entorno en el que opera. El ejemplo más claro se encontraría en los vehículos completamente autónomos, los cuales pueden compartir informaciones que van desde el estado del tráfico, las obras, condiciones meteorológicas, espacios en un parking (...) no sólo con el centro de información de carreteras sino también con otros coches con los que se encuentran conectados, perdiendo, con ello, y de manera progresiva, el hombre el papel de conductor del vehículo y adquiriendo el rol de mero pasajero¹⁹.

- El Parlamento Europeo ha considerado que la denominación inteligencia artificial es un término paraguas que abarca una amplia gama de tecnologías, técnicas y enfoques antiguos y nuevos, siendo más correcto referirse a 'sistemas de inteligencia artificial', entendiéndose cualquier sistema basado en máquinas que a menudo tienen poco más en común que estar guiadas por un conjunto determinado de objetivos definidos por el ser humano, con diversos grados de autonomía en sus acciones, y que realizan predicciones, recomendaciones o toman decisiones basadas en los datos disponibles²⁰.

Una vez más, puede constatarse la intrínseca relación existente entre la persona física y el sistema, no existiendo una independencia absoluta al respecto por parte del ente artificial sino que la intervención del ser humano resulta fundamental en el diseño de la posterior intervención de la máquina, aspecto nuclear, como he venido refiriendo en cada definición, a tener muy presente a la hora de delimitar la responsabilidad penal por el hecho ilícito producido.

- Doctrinalmente existen un sinnúmero de definiciones, si bien, entre todas ellas me permito destacar las llevadas a cabo por: i) Corvalán, quien refiere

¹⁹ SALVADORI, *Agentes artificiales, opacidad tecnológica y distribución de la responsabilidad penal*, cit., pp. 147-149.

²⁰ PARLAMENTO EUROPEO, *Report on artificial intelligence in a digital age*, 2022, p. 12. Recurso electrónico disponible en: <<https://www.europarl.europa.eu>>

que la inteligencia artificial se basa en obtener, por otros métodos artificiales, lo que alcanzamos con la inteligencia humana: el reconocimiento de patrones para alcanzar objetivos o resolver problemas²¹; ii) Kaplan y Haenlein, quienes identifican la inteligencia artificial como la capacidad de un sistema para interpretar correctamente datos externos, aprender de ellos y utilizarlos para alcanzar objetivos y tareas específicas mediante una adaptación flexible²²; o iii) Nilsson, quien la identificó con aquella actividad dedicada a hacer que las máquinas sean inteligentes, y la inteligencia es esa cualidad que permite a una entidad funcionar de manera y con previsión en su entorno²³, añadiendo Cuatrecasas Monforte la capacidad para hacer generalizaciones adecuadas de un modo oportuno, basándose en datos limitados, para después, en su caso, arrojar resultados individualizados lo cual implica la necesidad de observar analizar interpretar y tomar decisiones por sí mismas, de modo que aporten alguna información o solución nueva al usuario²⁴. En consecuencia con lo anterior, se trata de identificar mayoritariamente hoy día a la inteligencia artificial con la característica de racionalidad del agente, elemento clave que, como ya se ha referido, debe concretarse no tanto con la inteligencia sino con la capacidad de desarrollar actuaciones racionales similares o próximas al ser humano y que además va a aportar diferencias sustanciales a la hora de exigir un marco de responsabilidad penal concreto, tal y como referiré posteriormente.

Las formas en las que la inteligencia artificial puede repercutir en la comisión de hechos delictivos resultan muy plurales y diversas, lo cual ha hecho que un amplio sector de la doctrina opte explícitamente por usar el término AI-Crime para referirse a los delitos cometidos a través de inteligencia artificial²⁵. En este sentido, Salvadori ha compilado algunas prácticas que pueden hacer entender la magnitud e implicaciones concretas recurriendo a ejemplos comunes y perfectamente comprensibles

²¹ J.G. CORVALÁN, *Presentación: inteligencia artificial, automatización y predicciones en el Derecho*, en *Cibercrimen III*, cit., p. 14.

²² A. KAPLAN, M. HAENLEIN, *Siri, Siri in my hand: Who's the fairest in the land? On the interpretations, illustrations and implications of artificial intelligence*, en «Business Horizons», n. 62-1, 2018, p. 17.

²³ N. NILSSON, *The Quest for Artificial Intelligence: A History of Ideas and Achievements*, Cambridge University Press, Cambridge, 2010, p.13.

²⁴ CUATRECASAS MONFORTE, *La Inteligencia Artificial como herramienta de investigación criminal*, cit., p. 25.

²⁵ Vid., entre otros, T. KING, N. AGGARWAL, M. TADDEO, L. FLORIDI, *Artificial Intelligence Crime: An Interdisciplinary Analysis of Foresseable Threats and Solutions*, en «Science and Engineering Ethics», n. 26, 2020, p. 2.

al ser habituales en el interactuar diario, siendo los más conocidos los denominados spear phishing²⁶:

- i) Los modernos sistemas de inteligencia artificial pueden estudiar los hábitos de los usuarios online y enviar mensajes de correo electrónico fraudulentos mucho más sofisticados que los que han sido elaborados a través de tradicionales técnicas de social engineering. De este modo, aumentan las posibilidades de inducir a error a los internautas obteniendo de forma ilícita datos personales contenidos en tarjetas de crédito, credenciales para acceder a home banking, datos sanitarios, etc.
- ii) Los sistemas de inteligencia artificial pueden ser programados para acceder ilegalmente a sistemas informáticos, para llevar a cabo ataques distribuidos en la red (los llamados ataques distribuidos de denegación de servicios) o difundir peligrosos malware o ransomware.
- iii) Las denominadas social media bots contribuyen de forma significativa a la proliferación en las redes sociales de discursos de odio y campañas de desinformación, alimentadas de fake news, creadas con el objetivo de manipular la opinión pública, de crear consenso político o de desestabilizar los equilibrios geopolíticos²⁷.

Respecto a la regulación de semejantes prácticas ilícitas, la Unión Europea considera que, como en todo, siempre hay niveles de interacción y, en este sentido, diferencia entre los usos de inteligencia artificial de alto y bajo riesgo, siendo la regulación existente al efecto directamente proporcional a ello; esto es, necesitando las prácticas calificada de ‘alto riesgo’ -debe verificarse su uso concreto y el contexto, la naturaleza, la probabilidad, la gravedad y la irreversibilidad potencial del daño que puede esperarse que se produzca en violación de los derechos fundamentales y de las normas de salud y seguridad establecidas en el Derecho de la Unión- de estrictas

²⁶ Típicos correos electrónicos, aunque también páginas webs o mensajes de texto, falsos, aunque su apariencia pueda parecer real -hay un estudio cibernético previo de la víctima-, en donde se solicita acceso o información en la que el sujeto necesariamente debe abrir sesión en alguna plataforma. En este sentido, da igual la naturaleza de la información asociada, lo importante es la obtención del acceso ya que los usurpadores parten de la premisa de que los usuarios poseen muy pocas claves y la obtenida puede abrir otros accesos; esto es, por utilizar un símil, se trataría de obtener una llave que puede abrir muchas puertas. Esto se complica, aún más, en el caso de inserción de algún tipo de software malicioso en un segundo plano en el equipo de la víctima pudiendo espiar de manera libre.

²⁷ SALVADORI, *Agentes artificiales, opacidad tecnológica y distribución de la responsabilidad penal*, cit., pp. 141-143.

salvaguardias legislativas adicionales, mientras que los casos de uso de 'bajo riesgo' pueden, en muchos casos, requerir requisitos de transparencia simplemente para los usuarios finales y los consumidores²⁸.

2. Situación actual de la responsabilidad penal por un hecho delictivo cometido por inteligencia artificial

La praxis de la cuestión formulada puede resumirse muy sucintamente en la resolución de las siguientes preguntas: ¿puede la inteligencia artificial considerarse autora de un delito y, consecuentemente, ser objeto de pena? Si la respuesta fuera negativa, ¿quién respondería penalmente por semejante ilícito? Es aquí donde, como muy bien ha referido De la Cuesta Aguado, el jurista debe precisar cuándo o con qué grado de inteligencia se supera el umbral de la responsabilidad; es decir, el instante en el que la inteligencia artificial pueda ser considerada responsable de sus decisiones²⁹. Pérez Arias ha manifestado en esta obra que el principal problema con el que se encuentra el Derecho penal aparece en la difícil -si no imposible- conexión causal entre la conducta del autor y el algoritmo en el que finalmente se desencadena la conducta ideada. En consecuencia, el problema penal se puede volver muy peliagudo, porque todo puede partir de un algoritmo que aprende en la red y que no solo depende del autor intelectual (persona física) sino del desarrollo autónomo que generará una vez entre en la red³⁰.

Valls Prieto parte de la premisa de que cuanto más preciso es un análisis realizado por esta tecnología más difícil es conseguir una explicación de por qué la máquina ha tomado esa decisión. Este oscurantismo, que se describe con el concepto de 'black box', hace imposible establecer un nexo causal claro entre la lesión del bien jurídico protegido y el proceso de toma de decisiones por parte del sistema inteligente. Para evitar esta situación se debe realizar una evaluación en dos momentos cruciales: i) antes de empezar a diseñar el artefacto; y ii) antes de su lanzamiento al público en general. Semejante evaluación quedará registrada de tal forma que el juez podrá determinar si se han cumplido con las medidas adecuadas para mitigar los

²⁸ PARLAMENTO EUROPEO, *Report on artificial intelligence in a digital age*, cit., pp. 34 y 35. Recurso electrónico disponible en: <<https://www.europarl.europa.eu>>

²⁹ DE LA CUESTA AGUADO, *Inteligencia artificial y Derecho Penal*, cit., p. 52.

³⁰ PÉREZ ARIAS, *Algoritmos y Big Data en la responsabilidad penal: el reto de la cibercriminalidad en el Derecho penal*, en esta obra, § 4 de su Capítulo.

riesgos conocidos y, en algunos casos, también los desconocidos³¹.

Sea como fuere, Demetrio Crespo considera que los desafíos para la teoría de la acción penal van hoy día más allá debido a la irrupción de la inteligencia artificial, que pone a prueba la capacidad de rendimiento del actual sistema de imputación jurídico penal³².

2.1. *¿Puede considerarse a la inteligencia artificial autora única del delito?*

La respuesta inicialmente debe ser negativa, toda vez que se trata de una máquina que no hace más que obedecer/interactuar sobre los parámetros asignados por su primigenio creador; esto es, lo único que haría sería obedecer las instrucciones asignadas al respecto. Ahora bien, tal y como advierte Salvadori, el funcionamiento de los agentes de software y robots que utilizan tecnologías relacionada con inteligencia artificial y el aprendizaje automático es diferente ya que los sistemas de inteligencia artificial operan sobre la base de algoritmos complejos a través de los cuales adquieren la capacidad de autoaprendizaje, de mejorar su desempeño en el tiempo y de adoptar soluciones dinámicas; esto es, la inteligencia artificial favorece el desarrollo de los agentes artificiales con un nivel de autonomía tal que son capaces de realizar actividades que ni sus programadores, ni desarrolladores ni productores podrían haber previsto o querido. Es más, respecto de los agentes artificiales dotados de una alta autonomía³³, sus actuaciones no pueden ser determinadas ni previstas ex ante por los operadores que lo diseñaron, desarrollaron, probaron o produjeron al depender directamente de los input que adquieren con su experiencia y procesan de manera independiente³⁴.

En este sentido, como ha señalado De la Cuesta Aguado, la clave podría radicar en que la inteligencia fuera capaz de articular decisiones autónomas; esto es, aquéllas tomadas exclusivamente por el propio agente, el cual sería

³¹ VALLS PRIETO, *Sobre la responsabilidad penal por la utilización de sistemas inteligentes*, cit., p. 23.

³² DEMETRIO CRESPO, *El Derecho penal ante el desafío neurotecnológico y algorítmico: reflexiones preliminares*, cit., pp. 24 y 25.

³³ Conviene llamar la atención en este punto ya que se dice expresamente ‘alta autonomía’, no existiendo, por tanto, una autonomía plena, lo cual puede resultar muy relevante a la hora de proceder a la atribución de responsabilidad penal por hechos delictivos; esto es, la ‘alta autonomía’ no conlleva una total independencia del origen sino que sigue manteniéndose una minoritaria dependencia.

³⁴ SALVADORI, *Agentes artificiales, opacidad tecnológica y distribución de la responsabilidad penal*, cit., pp. 143 y 144.

capaz de dirigir su comportamiento o el de otros con su decisión; o, en otras palabras, incluiría la capacidad de actuación sin intervenciones externas, y en particular, humanas, si bien, conviene advertir, que, de momento y dado el estado de la ciencia, por decisiones autónomas debe entenderse autonomía en cuanto a la conducta o comportamiento más adecuado a un fin u objetivo, excluyendo la autonomía de la decisión sobre el fin o el objetivo; esto es, se limita el análisis a problemas en los que los fines u objetivos han sido fijados previamente de forma externa a la gente artificial inteligente y, conforme a ello se implica a una persona (o a varias) en la toma de decisiones, por lo que verdaderamente se le hace responsable a ese sujeto de las consecuencias de los actos del agente artificial. Piénsese en un robot para el control de fronteras que tiene la orden de disparar cada vez que detecta, en una franja de tierra, a personas en movimiento, respondería a la premisa de autónomo, en el sentido en el que pueda actuar sin intervención externa, pero no sería inteligente. Ahora bien, la autonomía relacionada con la inteligencia sería aquella que permite decidir al robot qué debe hacer según las circunstancias, con la posibilidad, incluso, de adoptar comportamientos para los que no ha sido predeterminado en su software. Esta autonomía abriría el paso a la responsabilidad derivada de la libertad de decisión como concepto contrario al de necesidad de decisión. Hoy día existen agentes artificiales inteligentes capaces de adoptar decisiones de comportamiento para la consecución de objetivos en función de fines predeterminados y en atención a circunstancias, lo cual se traduce en decisiones autónomas que hacen que no dependan de autorizaciones o controles de personas³⁵.

Es aquí, precisamente, donde Salvadori focaliza la cuestión en el sentido de considerar a los agentes artificiales capaces de aprender y de autodeterminarse donde puede surgir la problemática alrededor de si el Derecho Penal debe reconocer su autonomía efectiva y plena y equipararlos con sujetos (humanos) penalmente responsables³⁶. Siguiendo esta línea argumental, lejana en el tiempo a fecha de hoy, Romeo Casabona llama la atención sobre el instante futuro en el que la máquina pudiera tener esa intención criminal de manera idéntica a un humano, lo cual implicaría indefectiblemente la necesidad de modificar el actual sistema de responsabilidad penal para adaptarlo a los entes autónomos³⁷, una situación

³⁵ DE LA CUESTA AGUADO, *Inteligencia artificial y Derecho Penal*, cit., pp. 52 y 53.

³⁶ SALVADORI, *Agentes artificiales, opacidad tecnológica y distribución de la responsabilidad penal*, cit., p. 146.

³⁷ C.M. ROMEO CASABONA, *Criminal responsibility of robots and autonomous artificial intelligent systems?*, en «Comunicaciones en propiedad industrial y derecho de la competencia», n. 91, 2020, pp. 170-172.

respecto de la cual, como afirma Miró Llinares, queda aún lejos, toda vez que no existe una inteligencia artificial autónoma como tal, siendo la actual una suma de algoritmos de predicción utilizados para la realización de acciones o recomendaciones para actuar a partir de un conjunto de datos existentes identificación en ellos de patrones y probabilidades, siendo los seres humanos los que determinan completamente el actuar de la máquina, no necesitando ningún tipo de cambio actual en el sistema de atribución de responsabilidad penal, algo que indefectiblemente sucederá en un futuro cuando algún sistema de inteligencia artificial tenga rasgos de autonomía³⁸.

Sea como fuere, Salvadori distingue la situación actual de la hipotética de futuro. Así, mientras ahora los agentes artificiales más avanzados no tienen todavía una conciencia de sí mismos que les permita comprender el valor social negativo de sus acciones y captar la función preceptiva del Derecho penal, quedan igualmente considerables dudas sobre la (presunta) posibilidad de actuar dolosamente, es decir, con 'voluntad consciente' de cometer un hecho constitutivo de delito, o de forma imprudente, como actitud de incumplimiento de la norma cautelar. De forma resumida: un agente artificial (hasta la fecha) no puede ser culpable, como tampoco puede percibir el efecto intimidatorio ni disuasorio de la pena y, menos aún, garantizar la función especial de prevención ya que la sanción no sería adecuada para reducir ni resocializar al agente artificial, quedando únicamente la vía de restablecer los algoritmos que han determinado o influido en su toma de decisión. Ahora bien, de futuro, resulta previsible que los agentes artificiales lleguen a ser completamente autónomos, lo que significaría la posesión de una capacidad propia para actuar y, en algunos aspectos, ser capaces de percibir la realidad que les rodea y de operar de forma 'consciente', lo que devendría en el reconocimiento de la personalidad jurídica a la 'entidad' no humana, algo que tampoco resultaría una novedad, dado que en muchos ordenamientos jurídicos, tanto del common law como del civil law, se reconoce pacíficamente a las entidades y personas jurídicas incluso a efectos penales. A ello debe contribuir, para superar las limitaciones actuales, el hecho de que la existencia de la libertad de la voluntad en la cabeza del hombre no tiene un fundamento ontológico, sino que es el resultado de una construcción social que podría hacerse valer también contra los agentes artificiales completamente autónomos, desde el momento en el que, como las personas, interactúan e intervienen en las relaciones sociales y tienen un poder de conformación de la realidad. En definitiva, concluye

³⁸ F. MIRÓ LLINARES, *El sistema penal ante la inteligencia artificial: actitudes, usos y retos*, cit., pp. 89 y 90.

el autor, la verdadera cuestión subyacente no es si los agentes artificiales están dotados de la libertad de actuar como seres humanos, sino si, en un futuro próximo, podrán ser considerados responsables de la realización de conductas socialmente desviadas³⁹.

En este punto, Dupuy tiene clara la actuación directa desde una perspectiva de responsabilidad civil derivada de los daños ocasionados, no cabiendo realizar una traslación idéntica desde la perspectiva penal debido al referido problema de la culpabilidad, por lo que la autora sugiere dos vías de actuación: aplicando las reglas del Derecho vigente en relación con esta nueva realidad -excluiría una intervención penal- o creando un corpus iuris propio de la robótica, lo que requiere como presupuesto central el reconocimiento de cierta entidad (personalidad jurídica) a los robots como sujetos de Derecho, dando pie al origen del denominado y rechazado Derecho Penal de los robots⁴⁰.

Llevar a término esta última opción requiere, como ya se ha referido, dotar de personalidad a la inteligencia artificial, algo inviable conforme a la normativa española actual ya que, como ha referido Santos González, no pueden ser considerados: i) persona física, ya que chocaría frontalmente con el artículo 30 del Código Civil al requerir unos requisitos biológicos que no pueden darse en un robot; ii) como persona jurídica, sería posible su configuración, requiriendo ser destinatario de derechos y obligaciones, si bien nuevamente vuelven a surgir imposibilidades debido, entre otras circunstancias, a que la persona jurídica se configura sobre la unión de un grupo de personas físicas, algo claramente alejado de los robots, al igual que el hecho de generar daños por sí mismo, algo respecto de lo cual la inteligencia artificial tendría capacidad mientras le persona jurídica no; iii) como animal, tampoco enlazaría con el citado reconocimiento al no existir una base genética o biológica que permita experimentar a la fecha sentimientos naturales; y iv) como cosa, la cual se identifica como algo inanimado, carente de vida, frente a los robots, respecto de los cuales se persigue justamente lo contrario, que realicen conductas inteligentes, tengan autonomía para moverse e interactuar, debiendo considerarse como algo animado⁴¹.

Sin embargo, poco a poco, va cobrando cada vez más peso la idea de crear una personalidad jurídica electrónica, lo cual, según recoge Busato, sería el

³⁹ SALVADORI, *Agentes artificiales, opacidad tecnológica y distribución de la responsabilidad penal*, cit., pp. 151-155.

⁴⁰ DUPUY, *Inteligencia artificial aplicada al Derecho Penal y Procesal Penal*, en *Ciberdelincuencia II*, cit., p. 297.

⁴¹ M^a J. SANTOS GONZÁLEZ, *Regulación legal de la robótica y la inteligencia artificial: retos de futuro*, en «Revista Jurídica de la Universidad de León», n. 4, 2017, pp. 39-43.

producto de dos aspectos distintos: uno, el legislativo, que ha pretendido establecer una dinámica de control jurídico de la evolución científica en el ámbito de la robótica; y, otro, el argumento de equivalencia funcional entre personalidad humana y electrónica, el cual debe tomar como punto de partida, la concepción de acción anclada en la mecánica newtoniana y la expresión de la subjetividad a partir del modelo cartesiano del cogito ergo sum. Este segundo aspecto puede observarse en el hecho de que la evolución de la inteligencia artificial ha llegado a tal punto que pueden observarse en ella características humanas como la comunicación, la internalización del conocimiento, la capacidad de aprendizaje, la toma de decisiones, incluso la creatividad y la espontaneidad, lo que significaría que las máquinas razonan de forma similar a un humano, con lo que merecerían idéntico tratamiento jurídico, concluyendo que los robots dotados de inteligencia artificial son capaces de realizar tanto el actus reus -la capacidad de acción se correspondería con la capacidad de producción de un movimiento físico que produce un resultado- como el mens rea -capacidad de pensar, la cual se vincularía al aprendizaje autónomo- exigidos para la responsabilidad penal; es decir, resultan ser capaces de atender a los requisitos objetivos y subjetivos de la imputación⁴².

Dupuy considera finalmente que los robots, por más autonomía que tengan -drones, vehículos sin conductor, etc.- deben mantener, a fecha de hoy, la configuración y el tratamiento jurídico de las cosas⁴³. Por lo tanto, su utilización por un sujeto para cometer un delito significará que el dominio del hecho siempre le corresponderá al humano y, en base a ello, la imputación se realizará con base en la comisión directa y personal de la acción típica, abriendo un nuevo debate sobre qué persona y en qué condiciones se encontraría detrás de la responsabilidad penal⁴⁴.

⁴² No obstante, el citado autor concluye indicando que no se trata de considerar a la inteligencia artificial como persona incapaz de culpabilidad, sino, más bien, de reconocerlas como incapaces de acción (P. C. BUSATO, *De máquinas y seres vivos*, en *Derecho Penal y comportamiento humano*, dirigido por Demetrio Crespo E., Tirant Lo Blanch, Valencia, 2022, pp. 352-354 y 372).

⁴³ No obstante lo anterior, hay un corriente que viene demandando la creación de una nueva categoría de personalidad y así, como ha referido Santos González, puede surgir la opción de identificar a los robots como 'persona electrónica' o 'robot inteligente artificialmente' (*Ibidem*, pp. 43-46).

⁴⁴ DUPUY, *Inteligencia artificial aplicada al Derecho Penal y Procesal Penal*, cit., pp. 297 y 298.

2.2. ¿Quién debe responder penalmente entonces?

Delimitada la ausencia de responsabilidad directa sobre la inteligencia artificial, en los términos referidos en el apartado anterior, debe concretarse, a título doloso o imprudente, la persona que debe responder por las consecuencias jurídicas derivadas de la actuación de un sistema de inteligencia artificial, debiendo recurrir a criterios que permitan verificar cómo y por qué se ha producido el fallo correspondiente o bien la existencia de una voluntad criminal previamente establecida.

De la Cuesta Aguado ha ejemplificado las distintas hipótesis delictuales posibles respecto de los daños generados por un dron sobre tres modelos de responsabilidad penal:

- a) Responsabilidad personal por acciones o daños u otros resultados típicos, cuando la aeronave actúe bajo el control de una persona.
- b) Responsabilidad personal por acciones o daños u otros resultados típicos cuando la aeronave actúe dirigida por una persona pero en la decisión en concreto hayan intervenido factores ajenos al individuo que dirige o desarrolla el software.
- c) Responsabilidad personal por acciones o daños u otros resultados típicos cuando la aeronave actúe autónomamente, en cuyo caso habría que introducir dos sub-variantes: i) como consecuencia de errores, déficits o deficiencias de cualquier tipo en el software que establece fines, objetivos, límites y condiciones que incidieron en la toma de decisiones; y ii) subsidiariamente cuando no concurriera el supuesto anterior, por la decisión autónoma y derivada del propio aprendizaje de la aeronave⁴⁵.

Valls Prieto ha establecido una metodología de actuación basada en la fase de interacción. En este sentido, considera la existencia de tres niveles de imputación: i) compuesto por desarrolladores y fabricantes del producto que lo diseñan y mandan al mercado; ii) profesionales que utilizan estos sistemas inteligentes para realizar una parte de su trabajo y, por tanto, interacciona con la máquina de forma diferente a cómo lo realizarían los desarrolladores; y iii) los usuarios finales, en muchos casos consumidores del sistema. Como puede comprobarse cada uno de ellos tiene una relación y control diferente respecto al sistema inteligente, lo que, según este autor, deriva en una responsabilidad distinta. No obstante, lo anterior debe ser complementado para determinar si se ha producido un uso inadecuado

⁴⁵ DE LA CUESTA AGUADO, *Inteligencia artificial y Derecho Penal*, cit., p. 54.

e irresponsable de la inteligencia artificial. Para ello, debe igualmente verificarse cómo se ha producido la recolección de datos, su procesamiento y la posterior interpretación de los resultados finales por parte del sistema de inteligencia artificial⁴⁶.

Miró Llinares ha concluido que mientras no puede atribuirse autonomía a las entidades con inteligencia artificial el sistema de la teoría del delito sigue siendo totalmente válido para resolver los diferentes problemas causales y de atribución de responsabilidad, generalmente imprudente, si bien resulta esencial monitorizar la evolución de la inteligencia artificial desde una perspectiva de atribución de responsabilidad para evitar llegar a situaciones en las que el aprendizaje de las máquinas no permita decir que nadie ha tomado una decisión negligente pese a que existan daños⁴⁷.

Ahora bien, yendo paso a paso, puede afirmarse que, en primer lugar, no parece revertir problema jurídico-penal alguno el hecho de que un sujeto utilice consciente y voluntariamente un sistema de inteligencia artificial para desarrollar un comportamiento criminal, respondiendo dolosamente en concepto de autor ya que en estos casos se utilizaría al sistema de inteligencia artificial como herramienta o medio para la comisión del delito -piénsese, por ejemplo, en el uso de un dron para acabar con la vida de alguien-.

Por el contrario, la verdadera problemática subyacerá cuando no existe semejante grado de conocimiento respecto del posible resultado ilícito o simplemente cuando se genera el delito sin ser ésta la voluntad inicial del individuo, abriendo en estos casos la puerta a la imputación a título imprudente.

Sobre este extremo, Valls Prieto hace especial hincapié en la necesidad de determinar cuáles son las normas de cuidado que se tienen que respetar y que son exigibles a los ciudadanos en el uso de sistemas inteligentes para poder delimitar quién ha cometido el delito. Así pues, dentro de los diferentes códigos éticos que existen los principios expuestos por las guías éticas de la Comisión Europea resultan, a juicio de este autor, la fuente más completa, lo cual se traduciría, siguiendo los parámetros de algunos autores⁴⁸, en trasladar el sistema de compliance de las personas jurídicas a los sistemas inteligentes, si bien el modelo y la responsabilidad no podría ser idéntico por los problemas presentados, a saber: los programas de

⁴⁶ VALLS PRIETO, «*Sobre la responsabilidad penal por la utilización de sistemas inteligentes*» cit., pp. 25 y 26.

⁴⁷ MIRÓ LLINARES, *El sistema penal ante la inteligencia artificial: actitudes, usos y retos*, cit., pp. 112 y 113.

⁴⁸ Vid., a tal efecto, ROMEO CASABONA, *Criminal responsibility of robots and autonomous artificial intelligent systems?*, cit., p. 173.

compliance son realizados por personas jurídicas de forma obligatoria y en caso de cumplir les exime de responsabilidad penal. Por contra, el sistema inteligente no puede realizar por sí mismo este cumplimiento normativo, siendo un externo el que tomará la decisión de realizarlo o no. Además, mientras que dentro de la empresa se pueden tomar este tipo de decisiones, en el sistema inteligente no es posible, siendo una persona externa al sistema quién tomará las decisiones. Finalmente, el cumplimiento normativo implica la necesidad de un marco jurídico que regule el incumplimiento, mientras en materia de inteligencia artificial se parte de principios éticos (no normativizados) y, en consecuencia, los sistemas de prevención de riesgos son más amplios que en el caso de compliance⁴⁹.

De la Cuesta Aguado considera fundamental distinguir entre dos situaciones, según la responsabilidad provenga por el producto o por los daños. En el primer caso, piénsese, por ejemplo, en un software que contenga errores, la responsabilidad se compondría de círculos individuales, compartida parcialmente o no, en función de los propios ámbitos de competencia de cada uno de los intervinientes en la fabricación y uso del producto. Esta materia no es objeto ajeno al Derecho Penal y así podrían establecerse distintas responsabilidades individuales según los siguientes supuestos: i) de quien comete el error en su ámbito de competencia que puede implicar errores de distintas personas; ii) por culpa in vigilando; y iii) derivada de la creación o uso de una fuente de peligro. De otro lado, y en relación con la responsabilidad penal por daños, lesiones o, en general, resultado delictivo, aquí el agente artificial podría ser considerado como dependiente de la responsabilidad de una persona, ya sea por errores humanos de consecuencias previsibles o imprevisibles⁵⁰.

Salvadori ha compilado los supuestos en los que podría generarse la aplicación de tipos imprudentes según el rol desempeñado por la persona, siempre que exista una previsibilidad del riesgo⁵¹, destacando las hipótesis siguientes:

⁴⁹ VALLS PRIETO, *Sobre la responsabilidad penal por la utilización de sistemas inteligentes*, cit., pp. 29, 30 y 33.

⁵⁰ DE LA CUESTA AGUADO, *Inteligencia artificial y Derecho Penal*, cit., p. 58.

⁵¹ La previsibilidad del riesgo es el elemento clave para imputar responsabilidad penal por imprudencia, debiendo advertir la posibilidad, como apuntó Hallevy, de que queden fuera de la citada cobertura los casos en los que el resultado no pudiera ser anticipado por el programador o, en última instancia, por el usuario final (G. HALLEVY, *The Criminal Liability of Artificial Intelligence Entities from Science Fiction to Legal Social Control*, en «Akron Intellectual Property Journal», 2016, p. 172).

- i) Los hechos lesivos que tienen lugar debido a un error de programación o funcionamiento del agente artificial podrán ser imputados a título imprudente a quienes lo han programado, desarrollado, producido o probado.
- ii) Quien utiliza un agente artificial podrá ser llamado a responder de las consecuencias lesivas que se deriven de su uso si tiene poder efectivo para prevenir el hecho. Piénsese, como paradigma, en el vehículo semiautomático que, cerca de un cruce, no se detiene en la señal de stop, golpeando fatalmente a un peatón. El fallecimiento del peatón será atribuido al conductor si se demuestra que con la debida diligencia podría (y debería) haber tomado medidas para tomar el control del automóvil y evitar el golpe.
- iii) Cuanto mayor sea el nivel de autonomía del agente artificial, más elevada será la dificultad para imputar responsabilidad penal a personas físicas. Piénsese en un agente de software que actúa como piloto automático en las complejas fases de despegue y aterrizaje de un avión, no proporcionando posibilidad alguna de control manual por parte de los pilotos en caso de emergencia. Si a raíz de un error en el funcionamiento de los sensores que monitorean las condiciones climáticas, la aeronave pierde repentinamente altitud y se estrella contra el suelo provocando la muerte de los pasajeros, no puede hacerse ningún reproche imprudente a los pilotos, si resulta que no podían prever y evitar el accidente. No obstante, habría que hacer un rastreo hasta llegar a los sujetos que diseñaron, programaron y produjeron el software para analizar las complejas fases del vuelo; al fabricante del software, a quien se le podría reprochar a través de la imprudencia, cualquier error cometido en las fases de producción; e incluso al desarrollador o programador por un denominado ‘fallo de programación’ o por no haber previsto la posibilidad de permitir, ante situaciones de emergencia, la intervención manual de los pilotos en situaciones críticas.
- iv) La imputación debe realizarse extensiva en el caso de que verificados problemas de seguridad posteriores -piénsese, por ejemplo, en actuaciones de hackers- no se lancen actualizaciones necesarias para remediar fallos de software o, caso de no ser posible, no se retire del mercado.
- v) La responsabilidad a título imprudente no tiene por qué cerrarse en una única persona ya que si consigue verificarse el error de varios operadores, algunos ordenamientos jurídicos contemplan figuras expresas ante ello, como puede ser el caso del Código Penal italiano y el concurso de condiciones imprudentes independientes o la cooperación imprudente⁵².

⁵² SALVADORI, *Agentes artificiales, opacidad tecnológica y distribución de la responsabilidad penal*, cit., pp. 155-159.

Sobre la casuística anterior, considero muy relevante la hipótesis expuesta relativa al caso de la conducción de un vehículo de manera autónoma⁵³ en el que, fruto de un error del sistema, por el fallo de un sensor o no reconocimiento de una señal -piénsese, en un semáforo en rojo no detectado o simplemente en la salida del garaje del vehículo para posicionarse en la calzada- se genera el atropello de una o varias personas. Inicialmente, parece claro que el ocupante del vehículo automatizado exigirá responsabilidad a la empresa responsable del vehículo, siempre y cuando el mantenimiento y estado de los sensores correspondientes fuera el adecuado, ya que ha adquirido una máquina con una determinada fiabilidad probada científicamente con una serie de test y otros protocolos de seguridad. Sin embargo, las empresas fabricantes y comercializadoras no dejan margen a la improvisación y, para evitar riesgos de este tipo, incluyen cláusulas referentes a que la conducción automática exige la supervisión del conductor quien debe estar en todo momento atento a cualquier incidencia que pueda acontecer para evitar riesgos indeseados o, dicho en otras palabras, hay un deber de cuidado o actuación implícito en el ocupante del vehículo quien, pese a no manejar la máquina, debe estar preparado para tomar el control ante la menor incidencia, pudiendo derivar su inatención de una inobservancia del deber de cuidado que le era exigible con la consiguiente traslación de la responsabilidad penal a su persona.

3. *Pronunciamientos jurisprudenciales expresos*

Pese a que, como ha manifestado Alonso Salgado, aún hoy nuestros Juzgados y Tribunales distan mucho, desde una perspectiva estricta, de ser espacios tecnológicos, máxime cuando siguen instalados todavía en el tránsito transformador del siglo XX al XXI, irrumpe un nuevo desarrollo hijo de la nueva era, que tensiona nuestra tradición legal y nuestra cultura jurídica, que pone a prueba nuestra capacidad de adaptación y que evidencia resistencias, reparos y objeciones. Y ello, en muy buena medida, porque, lejos de lo que inicialmente pudiera parecer, ese nuevo desarrollo al que se

⁵³ Los delitos contra la seguridad vial presentan un amplio espectro de posibilidades e implicaciones tanto incriminadores como exculpatorias, tal y como puede observarse en el caso comentado o en el siguiente epígrafe con el supuesto de la persona con altas dosis de alcohol que aparece inconsciente con su vehículo parado en mitad de la calzada y su posible exención de responsabilidad penal en el hipotético caso de que toda la travesía hubiera sido gestionada la conducción por el sistema de inteligencia artificial.

hace referencia, constituye algo más que un giro meramente cuantitativo⁵⁴.

Así las cosas y pese a que existan un mayor número de resoluciones judiciales vinculadas a los contenidos expuestos y relacionados directamente con la inteligencia artificial, la búsqueda acotada bajo el campo abierto 'inteligencia artificial' y la jurisdicción 'penal' ha dado como resultado un total de doce manifestaciones jurisprudenciales, las cuales se vinculan a circunstancias accesorias, mayoritariamente, del relato de los hechos pero que pueden compilarse en las siguientes categorías:

A) Descripción meramente accesorio, donde se incluye la denominación en el relato de los hechos o los fundamentos jurídicos por cuestiones relacionadas con el objeto del litigio pero sin relevancia jurídica alguna. En este sentido, pueden destacarse, por ejemplo:

a1) El Auto de la Audiencia Provincial de Barcelona 517/2021, de 13 de julio [JUR 2021\377853], refiere el programa piloto llevado a cabo en Cataluña relativo al pronóstico de regresión del recluso en base al cálculo del riesgo de reincidencia mediante el uso de herramientas de inteligencia artificial, como es el caso de la implementación de un algoritmo predictor, si bien no se contempla en la legislación vigente como método científico adecuado de predicción que ayude en la valoración y ponderación del pronóstico.

a2) Llama poderosamente la atención, sobre todo antes de 2013, el hecho de que negocios que incluyen en su denominación la palabra 'Inteligencia Artificial' aparezcan de manera más habitual que otras nomenclaturas en comportamientos delictivos⁵⁵, siendo igualmente llamativo que, cuando una persona se encuentra formada en semejante materia, se haga expresa mención en la resolución judicial, aun cuando no tenga una vinculación relevante o directa con los hechos enjuiciados o el *modus operandi*⁵⁶.

⁵⁴ C. ALONSO SALGADO, *Acerca de la inteligencia artificial en el ámbito penal: especial referencia a la actividad de las fuerzas y cuerpos de seguridad*, en «*Ius et Scientia*», vol. 7, n. 1, 2021, p. 26.

⁵⁵ Vid., a modo de ejemplo, el Auto de la Audiencia Provincial de Madrid 683/2012, de 29 de diciembre [JUR 2012\20403], los querellantes pertenecen a la empresa 'Sistemas de Inteligencia Artificial'; proyecto comercial en materia de inteligencia artificial; Auto de la Audiencia Nacional 22/2005, de 20 de abril [JUR 2008\341211], el menor acudió al negocio denominado 'Central de Inteligencia Artificial SA', dedicado a la venta de equipos de cómputo; Sentencia 1141/2004, de 23 de noviembre [JUR 2005\5995], los acusados prestaban sus servicios profesionales para la entidad 'AIS Aplicaciones de Inteligencia Artificial'.

⁵⁶ Vid., a modo de ejemplo, la Sentencia de la Audiencia Provincial de Madrid 208/2010, de 30 de junio [JUR 2010\32127], donde se destaca del condenado, por un delito de

- B) Uso en los escritos de las partes. Se emplea el término para intentar explicar algunos comportamientos ilícitos. Tal puede ser el caso, por ejemplo, del Auto de la Audiencia Provincial de Burgos 839/2021, de 30 de noviembre, [JUR 2022\60508], donde, en la querella interpuesta, se recoge que «la recopilación de datos junto con el uso de técnicas modernas como la inteligencia artificial ha generado un amplio debate y preocupación respecto a los límites que deben aplicarse, entre los que se encuentra el derecho a la protección de datos de carácter personal».
- C) Autorización de implementación de medidas de vigilancia basadas en inteligencia artificial. El Auto de la Audiencia Provincial de Barcelona 72/2021, de 15 de febrero [ARP 2021\706], invoca el Libro Blanco sobre inteligencia artificial de la Comisión Europea, de 19 de febrero de 2020, con el propósito de verificar lo que se entiende por ‘autenticación’ vinculado al reconocimiento facial en el caso en el que dos sujetos fueron condenados por delito de robo con violencia en la personas en grado de tentativa a las penas de prisión, inhabilitación (...) y prohibición de acceso al centro comercial Mercadona situado en la calle (...) por tiempo de dos años. Ante la citada resolución judicial, la cadena de supermercados pretende, ante la imposibilidad práctica de asegurar el cumplimiento de la pena accesoria, dado que los trabajadores no podían estar pendientes de la gente que entraba en el supermercado y mucho menos de si éstos habían sido efectivamente condenados, lograr la autorización para usar medios electrónicos que permitan detectar la entrada de los penados a través de un circuito cerrado de video grabación, asegurando la proporcionalidad de los medios utilizados al tiempo que la efectividad de la medida, añadiendo los intereses públicos y legítimos de la mercantil de asegurar el cumplimiento de las resoluciones judiciales en las que la misma es la víctima/perjudicada. El funcionamiento del sistema que se pretendía implementar quedaría circunscrito al reconocimiento facial en tiempo real en virtud del cual se detectaría en 0,3 segundos, única y exclusivamente, la entrada de personas con sentencias firmes y medida cautelar de orden de alejamiento en vigor contra Mercadona o contra alguno de sus trabajadoras o trabajadores, procediendo a la eliminación de todo registro de persona que no cumpla con los citados caracteres. Finalmente, la Audiencia rechaza la petición por no resultar proporcional, necesaria ni idónea, además de no respetar las garantías adecuadas en orden a la protección de los derechos y libertades de los interesados,

injurias leves, sus estudios en informática y su especialización en realidad virtual y sistemas de inteligencia artificial; Sentencia del Tribunal Supremo 331/2003, de 5 de marzo [RJ 2003\2814], donde el querellante participaba en un programa de inteligencia artificial.

no ya sólo de los que han sido penados y cuya prohibición de acceso les incumbe, sino del resto de personas que acceden al citado supermercado.

D) La inteligencia artificial en los delitos contra la seguridad vial. La Sentencia de la Audiencia Provincial de Navarra 199/2011, de 15 de diciembre [JUR 2013\320202], recoge el caso de un delito de conducción de vehículos a motor bajo la influencia directa del alcohol, donde unos agentes de la policía municipal se encuentran un vehículo con el motor encendido, las luces puestas pero parado, ocupando uno de los carriles de circulación, estando durmiendo en su interior una persona con el cinturón de seguridad puesto, con un fuerte olor a alcohol⁵⁷ y restos de vómito reciente. Ante las evidencias, el Tribunal verifica que el coche no tuviera un sistema de inteligencia artificial que le permitiera circular sólo por las distintas vías, sin una inteligencia humana que lo guié, lo cual eliminaría la responsabilidad penal del tipo aplicado, cerciorándose, por tanto, de que tuvo que haber una inteligencia humana que lo guiara, siendo sólo posible que fuera el acusado, que en aquel momento tenía no sólo aquella, sino también la voluntad e inteligencia sensiblemente mermadas hasta el punto de que con cinturón de seguridad puesto, motor en marcha y luces encendidas del vehículo, se quedó dormido en el vehículo parado en la citada vía.

La anterior Sentencia recoge la esencia de la problemática que puede derivarse de futuro respecto al uso de los coches autónomos y las connotaciones que puede introducir en los delitos contra la seguridad vial y, más concretamente, los vinculados al artículo 379.2 del Código Penal ya que habrá que verificar, por ejemplo, que hubo conducción humana del vehículo aun cuando el sujeto se encontrara en el asiento del conductor con una ingesta alcohólica o de drogas que incidiera en la conducción.

4. Conclusiones

1. La inteligencia artificial constituye hoy día una realidad creciente e intensa, pudiendo afirmarse que estamos, como especifica la Resolución del Parlamento Europeo, de 3 de mayo de 2022, ante una cuarta revolución industrial ya que su incidencia ha generado un enorme impacto a nivel tecnológico mundial, incidiendo directamente en ámbitos pro-

⁵⁷ Posteriormente, el sujeto dio un resultado de 1,01 mg/l de alcohol.

pios y sensibles de cada país, como puede ser el caso de la economía, la sociedad, las relaciones internacionales o el medio ambiente. Semejante hecho no comporta una carta abierta en virtud de la cual todo tipo de actuaciones queden legitimadas en pro de los ámbitos anteriormente referidos, existiendo un deber de respeto a los derechos fundamentales, lo cual conlleva observar unas cautelas mínimas a la hora de regular su uso, no debiendo olvidar que se trata de un instrumento al servicio de las personas cuyo objetivo primordial ha de ser el bienestar humano y la seguridad.

2. No existe un concepto pacífico respecto a lo que debe entenderse por inteligencia artificial y así son muchas las disquisiciones existentes al respecto, partiendo incluso de la propia definición de inteligencia. Ahora bien, lo que no debe negarse, en pleno año 2022, es la existencia, junto con la inteligencia de los seres humanos, de otros sistemas de inteligencia, en este caso particular artificial, entendiéndose por semejante adjetivación el origen no intrínsecamente humano del citado pensamiento; esto es, la capacidad propia de pensamiento y su ejecución de una máquina, la cual, por un complejo procedimiento de algoritmos, puede realizar tareas similares a las personas, optimizando sus resultados tanto en un prisma temporal como cualitativo, constituyendo ese su principal objetivo fundacional, que no debe ser otro que realizar de manera satisfactoria determinadas actividades de forma precisa, acortando los plazos tradicionales y sin necesidad de una posterior intervención humana.
3. No obstante lo anterior, hay dos elementos que deben tenerse muy presentes desde una perspectiva jurídica: i) el origen del sistema de inteligencia artificial se encuentra embrionariamente ligado al ser humano y; ii) la inteligencia artificial no goza de absoluta autonomía, no siendo capaz de alcanzar una independencia total, al menos a fecha de hoy, lo cual resulta fundamental al objeto de delimitar el ámbito de la responsabilidad penal derivada de actuaciones ilícitas en las que se ve implicada un sistema de inteligencia artificial o, dicho en otras palabras, toda su actuación está basada en datos obtenidos previamente, detrás de los cuales, en algún momento anterior, ha intervenido un ser humano.
4. Como ha demostrado la historia, todo sistema evolutivo que mejora las condiciones de vida del ser humano tiene igualmente connotaciones negativas, las cuales suelen vincularse al ámbito delictivo y, en este caso, se ha acotado el término AI-Crime para referirse a los delitos cometidos a través de inteligencia artificial. Es por ello que los estados se ven obli-

gados a regular y crear instrumentos de control respecto de las prácticas delictivas derivadas del uso ilícito de los sistemas de inteligencia artificial, volviendo a reabrirse debates vinculados a cómo regular los concretos sistemas de responsabilidad penal derivados por la presencia de la inteligencia artificial en la comisión del delito y la clásica pregunta de si son suficientes los mecanismos existentes, bastando únicamente con adaptarlos a la nueva situación criminal o, por el contrario, se requiere de un nuevo sistema normativo que se adecúe a esta realidad.

5. Las preguntas que conviene realizar al respecto son claras: ¿puede considerarse a la inteligencia artificial autora única del delito? Y, en caso de respuesta negativa, ¿quién debería responder penalmente por el acto ilícito producido?

Con relación a la primera, la respuesta debe ser necesariamente negativa ya que una máquina no hace más que obedecer/interactuar sobre los parámetros asignados por su primigenio creador; esto es, por más que quiera, el sistema de inteligencia no va a gozar, al menos hasta la fecha, de una autonomía plena, sí de una alta autonomía pero nunca absoluta, reduciéndose el problema a aquellos casos en los que sus actuaciones no puedan ser determinadas ni previstas *ex ante* por los operadores que la diseñaron, desarrollaron, probaron o produjeron, siempre que dependa directamente de los inputs que adquiera con su experiencia y procese de manera independiente. No obstante, lo que parece claro es que los sistemas de inteligencia artificial carecen de personalidad, capacidad de culpabilidad e intimidación ya que no son capaces de percibir el efecto disuasorio de la pena, lo cual supone un nuevo eslabón a la hora de atribuirles específicamente responsabilidad penal, lo cual no resulta óbice para que algún sector doctrinal empiece a reclamar una regulación específica de los robots, lo que derivaría en un derecho de la robótica y, en el ámbito que ahora me ocupa, un Derecho Penal de los robots, donde alcanzaría una dimensión fundamental un nuevo rol denominado 'personalidad jurídica electrónica', si bien todo esto, repito, no sean más que consideraciones de futuro.

Negada, por tanto, la capacidad jurídica para cometer delitos a los sistemas de inteligencia artificial, cabe plantear la cuestión relativa a ¿quién responde entonces? El sistema de atribución de responsabilidad penal será el propio de las personas físicas, aun cuando un sector doctrinal abogue también por la inclusión de un sistema o régimen similar al de las personas jurídicas. Lo relevante aquí es la determinación de la responsabilidad, a título doloso o imprudente, según el actuar del sujeto o sujetos que, conforme a la relación causal concreta del hecho, sean considerados

responsables del ilícito producido, debiendo retrotraer la intervención de cada uno de los agentes, desde el responsable consumidor o titular del sistema de inteligencia artificial hasta los programadores/trabajadores de la empresa fabricante o comercializadora, quienes, conviene advertir, para evitar riesgos al respecto suelen incluir cláusulas de advertencia tendientes a eliminar una hipotética responsabilidad por la generación de un hecho ilícito en el desarrollo de la actividad concreta de la inteligencia artificial -véase, por ejemplo, en el caso de la conducción automática-.

6. La descripción inteligencia artificial en la jurisprudencia penal española no cuenta con una excesiva representatividad si nos atenemos a su descripción taxativa. En este sentido, conviene llamar la atención en que su aparición expresa puede encuadrarse en alguna de las siguientes categorías: i) como descripción meramente accesoria al supuesto juzgado; ii) breves y escuetas referencias en el escrito de alguna de las partes; iii) como autorización de implementación de medidas de vigilancia basadas en inteligencia artificial; o iv) lo que considero el aspecto más relevante, su implicación en los delitos contra la seguridad vial, donde puede modificar los patrones de actuación, persecución y enjuiciamiento de estos delitos, abriendo un nuevo espectro de posibilidades y actuaciones.

5. Bibliografía

- ALONSO SALGADO C., *Acerca de la inteligencia artificial en el ámbito penal: especial referencia a la actividad de las fuerzas y cuerpos de seguridad*, en «*Ius et Scientia*», vol. 7, n. 1, 2021.
- BUSATO P.C., *De máquinas y seres vivos*, en *Derecho Penal y comportamiento humano*, dirigido por Demetrio Crespo E., Tirant Lo Blanch, Valencia, 2022.
- CORVALÁN J.G., CIRAUDO D., *Inteligencia artificial al servicio de la justicia penal, contravencional y de faltas. 'Prometea' en el ámbito de la ciudad de Buenos Aires*, en *Cibercrimen II*, dirigido por Dupuy D., Editorial B de f, Buenos Aires, 2018.
- CORVALÁN J.G., *Presentación: inteligencia artificial, automatización y predicciones en el Derecho*, en *Cibercrimen III*, dirigido por Dupuy D., Corvalán J.G., Editorial B de f, Buenos Aires, 2020.
- CUATRECASAS MONFORTE C., *La Inteligencia Artificial como herramienta de investigación criminal*, La Ley, Madrid, 2022.

- DE LA CUESTA AGUADO P., *Inteligencia artificial y Derecho Penal*, en «Revista Penal México», n. 16 y 17, 2019-2020.
- DEMETRIO CRESPO E., *El Derecho penal ante el desafío neurotecnológico y algorítmico: reflexiones preliminares*, en *Derecho Penal y comportamiento humano*, dirigido por Demetrio Crespo E., Tirant Lo Blanch, Valencia, 2022.
- DUPUY D., *Inteligencia artificial aplicada al Derecho Penal y Procesal Penal*, en *Ciberdelitos II*, dirigido por Dupuy D., Editorial B de f, Buenos Aires, 2018.
- HALLEVY G., *The Criminal Liability of Artificial Intelligence Entities from Science Fiction to Legal Social Control*, en «Akron Intellectual Property Journal», 2016.
- INDEPENDENT HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE, *A definition of AI: Main capabilities and disciplines*, European Commission, Bruselas, 2019.
- KAPLAN A., HAENLEIN M., *Siri, Siri in my hand: Who's the fairest in the land? On the interpretations, illustrations and implications of artificial intelligence*, en «Business Horizons», n. 62-1, 2018.
- KING T., AGGARWAL N., TADDEO M., FLORIDI L., *Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions*, en «Science and Engineering Ethics», n. 26, 2020.
- MIRÓ LLINARES F., *El sistema penal ante la inteligencia artificial: actitudes, usos y retos*, en *Ciberdelitos III*, dirigido por Dupuy D., Corvalán J. G., Editorial B de f, Buenos Aires, 2020.
- MIRÓ LLINARES F., *Inteligencia artificial y justicia penal: más allá de los resultados lesivos causados por robots*, en «Revista de Derecho Penal y Criminología», n. 20, 2018.
- NILSSON N., *The Quest for Artificial Intelligence: A History of Ideas and Achievements*, Cambridge University Press, Cambridge, 2010.
- PALMA M^a F., *El Derecho como universo de sentido para la interfaz cerebro e inteligencia artificial*, en «Derecho Penal y comportamiento humano», dirigido por Demetrio Crespo E., Valencia, 2022.
- PARLAMENTO EUROPEO, *Report on artificial intelligence in a digital age*, 2022. Recurso electrónico disponible en: <<https://www.europarl.europa.eu>>
- PÉREZ ARIAS J., *Ciberdelincuencia: hacia la nueva realidad -virtual- del Derecho Penal*, en «Revista Electrónica de Doctrina y Jurisprudencia», n. 26, 2021.
- PÉREZ ARIAS J., *Algoritmos y Big Data en la responsabilidad penal: el reto de la ciberdelincuencia en el Derecho penal*, en esta obra.

- ROMEO CASABONA C.M., *Criminal responsibility of robots and autonomous artificial intelligent systems?*, en «Comunicaciones en propiedad industrial y derecho de la competencia», n. 91, 2020.
- SALVADORI I., *Agentes artificiales, opacidad tecnológica y distribución de la responsabilidad penal*, en «Cuadernos de Política Criminal», n. 133, 2021.
- SANTOS GONZÁLEZ M^a J., *Regulación legal de la robótica y la inteligencia artificial: retos de futuro*, en «Revista Jurídica de la Universidad de León», n. 4, 2017.
- VALLS PRIETO J., *Sobre la responsabilidad penal por la utilización de sistemas inteligentes*, en «Revista Electrónica de Ciencia Penal y Criminología», n. 24-27, 2022.

CAPÍTULO IV

I miti d'oggi: la certezza del diritto fra utopie e distopie

Lorenza Grossi

SOMMARIO: 1. I 'problemi antichi': la dialettica tra disposizione e norma come costante irrinunciabile del diritto – 2. Il mito della 'certezza del diritto' e il demone dell'interpretazione giudiziaria nell'Illuminismo giuridico – 3. La (in)certezza del diritto e la crisi della legalità – 4. L'intelligenza artificiale: il diritto calcolato – 5. La bellezza del mito: demistificabile, ma non 'demitificabile' – 6. Bibliografia.

1. I 'problemi antichi': la dialettica tra disposizione e norma come costante irrinunciabile del diritto

L'utilizzo dell'intelligenza artificiale all'interno del processo penale apre nuove sfide (anche) per il diritto penale sostanziale.

La quarta rivoluzione tecnologica sembrerebbe aver preso spazio anche in campi di esclusiva competenza dell'uomo: l'idea di un *software* capace di predire la decisione del giudice o di un giudice-robot non è più solo futuribile, ma si avvicina ad essere il 'prossimo presente'. Necessario diventa, quindi, chiarire quali siano gli effetti che questi strumenti, direttamente o indirettamente, possano apportare all'interno dei nostri sistemi penali.

Una premessa è, però, dovuta: le 'rivoluzioni', è fatto noto, rappresentano delle vere e proprie fratture rispetto al passato e getterebbero le basi per un futuro 'nuovo', nato o ideato con l'illusione di avere poco da spartire con ciò che si pone, logicamente, come suo 'antecedente'.

Ad ogni 'fatto' nuovo, che si presenta nella serie storica di un dato contesto, può guardarsi come ad una cesura, che si stacca da ciò che la precede e si pone come pietra angolare di qualcosa di diverso, oppure come uno sviluppo necessario di qualcosa che, ad una più attenta analisi, già si era manifestato all'interno della medesima serie.

L'impressione, più in generale, è che la storia si presenti, in molti casi, come una concatenazione di utopie che ne illuminano e disegnano il flusso,

destinate a ripresentarsi, sia pur con forme e in tempi diversi.

Circoscrivendo l'oggetto di questa analisi e utilizzando come *fil rouge* la certezza del diritto ed il principio di legalità, non si può prescindere dall'individuazione, sia pur sommaria e forse a tratti approssimativa, delle tappe più rilevanti che hanno segnato l'evoluzione del diritto, avendo cura di chiarire quali siano, da una parte, i 'problemi antichi' e, dall'altra, le soluzioni 'nuove' elaborate e succedutesi nel tempo.

Un'avvertenza metodologica: nelle pagine prevalentemente storiche di questo contributo si utilizzerà, salva qualche eccezione, il «nome» che di un certo istituto è stato poi dato dalla modernità; si ritiene, insomma, convincente quel fortunato insegnamento per cui ciò che è «essenziale non è [...] che si trovi bell'e formulato il *concetto* o posta la *distinzione* o – tanto meno – enunciato il *nome*. L'essenziale è che [...] ci sia la *cosa*»¹.

In via di prima approssimazione, potrebbe dirsi che la legge nasca con l'obiettivo di stabilizzare regole, le quali, una volta affidate allo scrigno delle parole, garantirebbero quell'esigenza di 'certezza' di cui il diritto è presupposto².

¹ E. BETTI, *Diritto romano e dogmatica odierna*, in «Archivio giuridico Filippo Serafini», 1928, p. 58, [i corsivi sono dell'Autore]. Su questi temi della metodologia bettiana A. SCHIAVONE, *Il «nome» e «la cosa»*, in *Quaderni fiorentini*, Giuffrè, Milano, 1978, pp. 293 e ss.

² Emblematica, in questa direzione, l'esperienza giuridica romana a partire dalla 'codificazione' delle XXII tavole, fortemente voluta dai plebei per limitare l'interpretazione, ritenuta talvolta arbitraria, dei *pontifices*, appartenenti alla classe patrizia. Sul punto, *amplius* F. SERRAO, voce *Interpretazione della legge (dir. rom)*, in «Enciclopedia del diritto», XXII, 1972, p. 240; F. DE MARTINO, *Storia della Costituzione romana*, vol. I, Jovene, Napoli, 1951, pp. 302-304, che evidenzia come la codificazione delle XII tavole si sviluppa da pressioni dovute ad una incertezza del diritto, che si risolveva, sul piano applicativo, in forme di disuguaglianza. In questo solco, con tutta evidenza, l'emanazione delle leggi funge da limite al potere del governo, conformandosi, in una sua prima accezione, come 'certezza-eguaglianza', e cioè, autentica garanzia a che la legge venga applicata in maniera uniforme nei confronti di tutti. Non si dimentichi poi che, con riferimento all'epoca antecedente alle XII tavole, è già Pomponio a parlare di *ius 'incertum'* (Pomponio, Libro singolari enchiridii, D.1.2.2.1 e D.1.2.2.4).

Sulla certezza del diritto intesa come 'in sé' del diritto, si veda, per tutti, N. BOBBIO, *La certezza del diritto è un mito?*, in «Rivista Internazionale di Filosofia del Diritto», estratto, 1/1951, Fratelli Bocca – Editori, Editori, Milano-Roma-Firenze, pp. 1 ss.. L'illustre Autore, in particolare, critica apertamente le opinioni espresse da J. Frank, *Law and the Modern Mind*, Cowasd-McLann, VI ed., 1949, il quale, muovendo da una piena coincidenza fra l'attività 'interpretativa' del giudice ed il suo carattere intrinsecamente 'creativo', identifica nella certezza del diritto un'illusione degli individui 'infantili', che non hanno raggiunto l'adeguata maturità, disposti a credere irrazionalmente al 'mito' della certezza del diritto con la stessa tenacia con cui il bambino accetta senza riserve l'autorità, ritenuta onnipotente ed onnisciente del padre, che detta regole valide a garantirgli sicurezza: «[n]on gli viene il

Lo stesso strumento utilizzato, però, tradirebbe l'obiettivo prefissato o comunque lo metterebbe in crisi: la storia del diritto, altrimenti detto, in una certa prospettiva si caratterizzerebbe per quella costante dialettica che esiste fra testo scritto e interpretazione, tra certezza del diritto e evoluzione attraverso l'ermeneutica, tra disposizione e norma³ – dialettica che, con le precisazioni che seguiranno, è destinata ad essere nuovamente coinvolta (o sconvolta) dall'utilizzo di *software* all'interno del diritto penale.

L'idea di veicolare o comunque 'controllare' l'interpretazione sembra potersi rintracciare, sia pur ancora *in nuce*, già in età Augustea, là dove, parallelamente alla progressiva centralità dei senatoconsulti, ora dotati di efficacia normativa, e l'affermarsi di potere normativo diretto da parte del *princeps* (*constitutiones principis*)⁴, l'*interpretatio* va ad assumere con maggiore chiarezza i connotati di una *interpretatio legis* – e non *iuris*, come si era, d'altra parte, imposta durante l'età repubblicana⁵. È sempre durante il principato, poi, che nascono e acquisiscono sempre più rilevanza il *ius respondendi ex auctoritate principis* e l'interpretazione autentica da parte del principe stesso, sì che l'attività giurisprudenziale viene a

sospetto, ad esempio, che la certezza del diritto, anziché essere una illusione di individui viziati da infantilismo inguaribile, sia un *elemento intrinseco del diritto, sì che il diritto o è certo o non è neppure diritto, e che a togliere di mezzo la certezza non si libera l'umanità da una nociva illusione, ma la si priva letteralmente del sussidio e del rimedio del diritto*» (*ibidem*, p. 5, corsivi nostri).

³ Il rinvio è chiaramente a V. CRISAFULLI, voce *Disposizione (e norma)*, in «Enciclopedia del diritto», XII, 1964, pp. 195 ss., intendendo per disposizione l'enunciato linguistico prescrittivo e per norma, invece, il prodotto dell'attività interpretativa che ha ad oggetto la disposizione: detto altrimenti, la regola deducibile una volta che sia stato attribuito 'significato' a lemmi, parole e sintagmi. Sui modelli interpretativi, per tutti, *amplius* A. MASSARO, *Determinatezza della norma penale e calcolabilità giuridica*, Editoriale Scientifica, Napoli, 2020, pp. 189 ss. Sulla dialettica fra disposizione e norma, che si traduce, nella prassi, in un 'rapporto di forza' tra giurisprudenza e legislazione la letteratura è sterminata, si segnalano, tra i tanti, G. GORLA, voce *Giurisprudenza*, in «Enciclopedia del diritto», XIX, 1970, 493 ss.; E. PIOLA CASELLI, voce *Giurisprudenza*, in «Digesto italiano», XII, 1900-1904, pp. 844-847, che già oltre un secolo fa si domandava se potesse immaginarsi una giurisprudenza 'fonte del diritto'; P. CALAMANDREI, *La Cassazione civile. Parte prima*, in *La memoria del diritto*, ristampa in alterata, Roma TrE-Press, Roma, 2019, § 148 ss., con riferimento all'epoca *ante* e *post* rivoluzione francese; da ultimo M. TRAPANI, *Creazione giudiziale della norma penale e suo controllo politico*, in «Archivio penale», 1/2017, 1 ss., che, muovendo da una piena parificazione fra interpretazione e creazione del diritto, ritiene che il giudice *crei* il diritto.

⁴ V. CENDERELLI, voce *Fonti del diritto in diritto romano*, in «Digesto delle discipline privatistiche», 1992, § 10.

⁵ SERRAO, voce *Interpretazione della legge (dir. rom)*, cit., pp. 248-250; A. BURDESE, voce *Interpretazione nel diritto romano*, in «Digesto delle discipline privatistiche», 1993, § 4.

svolgersi sotto il controllo del *princeps*⁶. Sembrerebbe, in generale, che i rapporti fra potere legislativo e potere giudiziario, già in tempi antichi, si caratterizzano nel segno di una soccombenza, logica e politica, del secondo rispetto al primo: è infatti il testo legislativo – ciò che oggi definiremo ‘disposizione’ – un *pruus* logico, sulla cui base (e non oltre quella) germina e si sviluppa l’*interpretatio legis*; d’altra parte la stessa importanza rivestita dalla interpretazione autentica dell’imperatore sarebbe dotata di un grado di autorevolezza maggiore, per ragioni di carattere politico.

La ‘messa a bando’ dell’interpretazione, intesa come fonte creativa del diritto, si avrà, in maniera più netta e dopo un lungo processo⁷, solo con Giustiniano. La ‘creazione’ e l’‘interpretazione’ delle leggi vengono ora a saldarsi in unico binomio, che sintetizza le attività riferibili in via esclusiva al solo imperatore. In questo orizzonte vanno collocati quegli istituti che, all’interno della celebrazione del processo, permettevano forme di ingerenza dell’imperatore, adito tramite ricorso: l’idea di Giustiniano, correttamente, è che il luogo deputato all’interpretazione del diritto è, per l’appunto, il processo⁸. Il potere esclusivo di interpretare le leggi non può, quindi, prescindere dall’individuazione di meccanismi che inseriscano l’imperatore all’interno del medesimo: questi infatti, non solo, si ritaglia un ruolo di ‘chiusura’ nel sistema ‘giudiziario’, rivestendo la carica di ‘giudice supremo’, ma, attraverso l’introduzione di alcuni istituti *ad hoc*, disegna dei percorsi validi ad instaurare forme di dialogo fra giudice ed imperatore tutte le volte in cui dai fatti in causa emergano dubbi di carattere giuridico. In questo solco vanno inseriti la *consultatio ante sententiam* e il ricorso

⁶ BURDESE, voce *Interpretazione nel diritto romano*, cit., § 4; CENDERELLI, voce *Fonti del diritto in diritto romano*, cit., § 9, che chiarisce come il *ius respondendi ex auctoritate principis* non limitava in alcun modo l’esercizio dell’attività di consulenza, che fu sempre libera, ma dotata di maggiore autorevolezza i *responsa* di questo tipo. Si trattava, insomma, di una «sorta di ideale delega della funzione interpretativa del diritto, che il principe rivendicava come propria».

⁷ Si tratta di un percorso che si sviluppa a partire da Costantino, con l’introduzione delle *leges generales*, la cui interpretazione va man mano affermandosi valorizzando soprattutto la *voluntas* o la *mens* del ‘legislatore’, così G.G. ARCHI, *Interpretatio iuris - interpretatio legis interpretatio legum*, in *Studi in onore di Francesco Santoro-Passarelli*, vol. VI, Jovene, Napoli, 1972, p. 39. Si veda anche BURDESE, voce *Interpretazione nel diritto romano*, cit., § 6, che chiarisce come proprio da questo momento storico spetta al solo imperatore l’interpretazione ‘evolutiva’ del *ius* per adeguarlo all’*aequitas*. Sulle *leges generales*, cfr. S. PULIATTI, *Il diritto prima e dopo Costantino*, in *Enc. costantiniana*, 2013, disponibile su <www.treccani.it>.

⁸ Insiste su questo aspetto, in particolare, G. BASSANELLI SOMMARIVA, *L'imperatore unico creatore ed interprete delle leggi e l'autonomia del giudice nel diritto giustiniano*, Giuffrè, Milano, 1983, p. 69.

del giudice all'imperatore su questioni di diritto⁹: la prima dava facoltà al giudice di fare ricorso all'imperatore per questioni giuridicamente dubbie, la seconda, invece, imponeva al giudice di rivolgersi all'imperatore-legislatore nei casi in cui venisse riscontrata una *lacuna*, che, richiedendo ai fini della risoluzione un'attività interpretativa, non poteva in alcun modo prescindere da una decisione dell'unico *interpres legum*¹⁰.

Il problema delle lacune, in effetti, comincia ad apprezzarsi in maniera più significativa proprio in questo momento storico, come diretta conseguenza di un duplice ordine di ragioni, strettamente legate fra loro: da un lato, la pretesa 'completezza'¹¹ dell'ordinamento non potrebbe ammettere 'vuoti' di regolamentazione; dall'altra, l'innovativo divieto di *non liquet*, stabilito per la prima volta nella Nov. 125, imporrebbe di colmare in via ermeneutica la lacuna e di giungere, quindi, ad una decisione¹².

⁹ È discusso se il ricorso del giudice all'imperatore su questioni di diritto assuma una forma di 'dialogo' autonoma. Alcuni Autori, infatti, lo riconducono, in generale alla *consultatio ante sententiam*, così, ad esempio, W. LITEWSKI, *Consultatio ante sententiam*, in «Zeitschrift der Savigny-Stiftung für Rechtsgeschichte (Romanistische Abteilung)», LXXXVI, 1969, pp. 227 ss.; come istituto autonomo si veda BASSANELLI SOMMARIVA, *L'imperatore unico creatore ed interprete delle leggi e l'autonomia del giudice nel diritto giustiniano*, cit., 78-81, cui si rinvia anche per ulteriori indicazioni bibliografiche.

¹⁰ L'imperatore si definisce, infatti, 'conditor' e 'interpres legum' sin da C. 1. 14. 12 del 529, ben prima della più 'famosa' *constitutio* Tanta del 533, § 21, che ribadisce come solo alla *auctoritas Augusta...concessum est leges et condere et interpretari*. Sul tema, *amplius*, BASSANELLI SOMMARIVA, *L'imperatore unico creatore ed interprete delle leggi e l'autonomia del giudice nel diritto giustiniano*, cit., pp. 69 ss., spec. 103; ARCHI, *Interpretatio iuris - interpretatio legis interpretatio legum*, cit., pp. 49 ss. spec. 51, ove l'Autore evidenzia come «Il binomio *conditor et interpres* lungi dall'individuare due attività distinte, come potrebbe pretendere la dommatica moderna, intende piuttosto essere la sintesi di ogni possibile manifestazione creatrice del diritto»; BURDESE, voce *Interpretazione nel diritto romano*, cit., § 6; F. GALLO, «*Princeps*» e «*ius praetorium*», in «Rivista di Diritto Romano», I/2001, p. 18 ss.

¹¹ Sulla completezza del *Corpus iuris civilis* nell'idea di Giustiniano, v. SCHIAVONE, *Ius. L'invenzione del diritto in Occidente*, Einaudi, Torino, 2017, cap. I, § 2. Sembra che il 'dogma della completezza' sia legato a filo doppio con l'idea di 'unitarietà' che connota la codificazione: tutto il diritto, secondo Giustiniano, è ora contenuto nel *Corpus iuris*, senza lasciare altro fuori da sé. Si tratta di premesse che saranno ulteriormente sviluppate, poi, nella tradizione romanistica medioevale; colloca in questa epoca la 'data di nascita' del dogma della completezza N. BOBBIO, *Teoria generale del diritto*, Giappichelli, Torino, 1993, p. 241.

¹² Si veda sul punto A. MORDECHAI RABELLO, *Non liquet. Dal Diritto Moderno al Diritto Romano*, in «Archivio giuridico Filippo Serafini», 185, 1973, pp. 31-55, che individua in quel '*perfecte examinare causam et quod sibi iustum atque legitimum visum sit iudicare*' contenuto nella Nov. 125 del 543 la base per affermare il divieto di *non liquet*. Si tratta, ad ogni modo, di una questione estremamente discussa in dottrina.

Si tratta, in realtà, di questioni, di nuovo, destinate a ripresentarsi, con cadenze simili se pur in tempi diversi, come ‘problema antico’ che necessita di ‘soluzioni nuove’, seguendo un ritmo circolare ed incalzante, destinato a non finire. Ma di ciò meglio in seguito.

2. *Il mito della ‘certezza del diritto’ e il demone dell’interpretazione giudiziaria nell’Illuminismo giuridico*

Il ‘mito’ della certezza del diritto, è fatto arcinoto, si sviluppa in quel florido ambiente culturale che fa da sfondo ‘all’età dei Lumi’: la legge scritta si pone come autentico architrave del ‘moderno’ Stato di diritto, il quale si ammantava di una legalità che è strumento e fine al tempo stesso, e poggia le proprie basi su una netta separazione dei poteri. Alla legge, chiara, generale e astratta, emanata dall’organo rappresentativo della volontà popolare, è affidato il compito di regolare la vita dei consociati, garantendo certezza del precetto e calcolabilità della sanzione: la legge è unica fonte del diritto ed impone, quindi, un feroce ridimensionamento dell’attività interpretativa.

Attorno al ‘mito’ della legge orbitano e disegnano ellitticamente i loro moti ulteriori ‘dogmi’: la certezza del diritto, la coerenza e la completezza dell’ordinamento, che troveranno piena esplicazione nella codificazione.

Al potere giudiziario, definito già «terribile»¹³ o «odioso»¹⁴, in questo ‘nuovo’ sistema toccherebbe una mera funzione dichiarativa: il giudice è *bouche de la loi*¹⁵ o *Subsuntionsautomat*, a lui non è chiesto altro che di compiere un «sillogismo perfetto: la maggiore dev’essere la legge generale,

¹³ CH. L. MONTESQUIEU (DE SECONDAT), *De l’esprit des loix*, t. I, Barillot & Fils, Genève, 1748, Livre XI, chap. VI, p. 247. Per un approfondimento, *amplius*, L. FERRAJOLI, *Diritto e ragione*, 9ª ed., Laterza, Bari, 2008, *passim*, spec. pp. XV, 12, 18, 48-50 (note 16, 20, 23, 24, cui si rimanda per ulteriori indicazioni bibliografiche).

¹⁴ M.J.A.N. CONDORCET (DE CARITAT), *Idées sur le despotisme, à l’usage de ceux qui prononcent ce mot sans l’entendre* (1789), Firmin Didot frères, Paris, 1847, § XII, 155, su cui, di nuovo, FERRAJOLI, *Diritto e ragione*, *passim*, spec. pp. XV, 50 (nota 23).

¹⁵ MONTESQUIEU (DE SECONDAT), *De l’esprit des loix*, cit., Livre XI, chap. VI, è. 256: «Mais le Juges de la Nation ne font, comme nous avons dit, que la bouche qui prononce les paroles de la Loi; des Etres inanimés qui n’en peuvent modérer ni la force ni la rigueur». Sull’interpretazione giudiziaria, si rinvia, ancora, a FERRAJOLI, *Diritto e ragione*, pp. 10 ss., 44-45, 48-50 (note 16 e 24), 94 ss., 143 ss., sul ‘modello di giudice penale’ dell’Illuminismo giuridico, G. AMARELLI, *Il giudice ed il rispetto della legge penale in sede interpretativa. Obsolescenza apparente e attualità permanente del pensiero di Beccaria*, in <<https://www.osservatorioaic.it>>, giugno 2015, pp. 1 ss.

la minore l'azione conforme o no alla legge, la conseguenza la libertà o la pena»¹⁶. Nell'interpretazione, in quest'ordine di idee, si camufferebbe la funzione creativa del giudicante, che verrebbe così a sostituirsi al «legislatore»¹⁷, dando origine a disequaglianze nell'applicazione e quindi, a monte, provocando incertezza del diritto. La legge sembra, in questo contesto, dotata dei caratteri della necessità e della sufficienza: muovendo da un rigido formalismo interpretativo, secondo cui dalla disposizione sarebbe possibile ricavare una sola norma, oggettivazione della volontà univoca del legislatore e deducibile dal significato «proprio» delle parole¹⁸, si sgretola il rapporto fra disposizione e norma (*rectius*: norme), a favore di una solida ed indefettibile centralità della prima. L'introduzione del *référé lefislatif* rappresenta solo l'acme di queste posizioni. Non diversamente dall'esperienza tardoantica, infatti, il giudice aveva, rispettivamente, la facoltà e l'obbligo di rivolgersi al legislatore tutte le volte in cui si fosse trovato innanzi: (i) ad un dubbio di diritto; (ii) ad una lacuna o ad una antinomia, 'accidenti' che imponevano una funzione creativa¹⁹.

Tralasciando la (s)correttezza metodologica di queste impostazioni, basta qui evidenziare come, a ben vedere, la rigida separazione dei poteri veniva comunque superata: se è vero che il giudice non si sostituiva al

¹⁶ C. BECCARIA, *Dei delitti e delle pene*, Nicolo Bettoni, Brescia, 1807, § IV, p. 13, ove l'Autore continua avvertendo come «[q]uando il giudice sia costretto, o voglia fare anche soli due sillogismi, si apre la porta dell'incertezza» [corsivi nostri].

¹⁷ P. VERRI, *Sulla interpretazione delle leggi* (1765), in «Il Caffè», tomo II, foglio XXVIII, consultabile su <<http://illuminismolombardo.it>>: «[s]e il giudice diventa legislatore, la libertà politica è annichilata; il giudice diventa legislatore sì tosto che è lecito interpretar la legge; dunque si proibisca al giudice l'interpretar la legge; dunque si riduca ad esser mero esecutore della legge; dunque eseguisca la legge nel puro e stretto significato delle parole e nella materiale disposizione della lettera. Che il giudice, tosto che la legge è soggetta a interpretarsi più in un senso che in un altro, diventi legislatore, è cosa per sé evidente, basta per esserne convinto il riflettere che *interpretare* vuol dire sostituire se stesso al luogo di chi ha scritto la legge, e indagare cosa il legislatore avrebbe verisimilmente deciso nel tale o tal altro caso su cui non parla chiaramente la legge. *Interpretare* significa far dire al legislatore più di quello che ha detto, e quel *più* è la misura della facoltà legislatrice che si arroga il giudice» [corsivi dell'Autore].

¹⁸ Con riferimento al formalismo (o cognitivismo) interpretativo, R. GUASTINI, *Enunciati interpretativi*, in «Ars interpretandi» 2/1997, 35 ss.; G. ZACCARIA, voce *Intepretazione*, in «Enciclopedia del diritto», Annali V, 2012, p. 695.

¹⁹ Sul *référé lefislatif* facoltativo ed obbligatorio, CALAMANDREI, *La Cassazione civile*, cit., § 150; su questi temi, *amplius*, P. ALVAZZI DEL FRATE, *Giurisprudenza e référé législatif in Francia nel periodo rivoluzionario e napoleonico*, Giappichelli, Torino, 2005, *passim*; ID., *L'interpretazione autentica nel XVIII secolo. Divieto di interpretatio e riferimento al legislatore nell'illuminismo giuridico*, Giappichelli, Torino, 2000, cap. IV, spec. pp. 153 ss.

legislatore, altrettanto non può dirsi per il legislatore, che, di fatto, veniva a sostituirsi al giudice²⁰. Le ingerenze del legislatore, almeno così sembra, godevano di una legittimazione propria ed ultronea, capace anche di superare quei limiti, che pur dovevano essere così cautamente rispettati dall'organo giudiziario. Ad una più attenta analisi, fra l'altro, l'istituto del *référé législatif* si opponeva alla certezza del diritto: l'interpretazione autentica del legislatore, in effetti, veniva applicata retroattivamente, così sacrificando quella che era la base stessa su cui si fondava il sistema 'lego-centrico'²¹. Questo istituto verrà comunque implicitamente abrogato nella sua variante facoltativa dall'art. 4 del Code Napoléon (1804) e definitivamente con la l. 1 aprile 1837.

3. La (in)certeza del diritto e la crisi della legalità

I cardini dell'illuminismo giuridico, che hanno contribuito in maniera preponderante allo sviluppo dei nostri sistemi giuridici, sembrerebbero oggi crollare, restituendosi all'epoca post-moderna come semplici 'mitologie'²² o 'fossili'²³ di un passato remoto. La disposizione ha ormai ceduto il passo alla norma²⁴: dal formalismo interpretativo si è passati, con posizioni mediane²⁵, allo scetticismo. La norma, secondo questa impostazione, non preesisterebbe all'interpretazione, ma sarebbe il mero risultato di una scelta discrezionale di quell'interprete qualificato che è il giudice²⁶, sì da acquistare sempre maggiore rilievo l'idea che la giurisprudenza *sia* fonte del diritto²⁷. Anche il principio di legalità, nell'epoca post-moderna, assume

²⁰ CALAMANDREI, *La Cassazione civile*, cit., § 166, p. 436: «provocando l'emaneazione di una legge in pendenza di lite, trasformava il legislatore in giudice»; ALVAZZI DEL FRATE, *L'interpretazione autentica nel XVIII secolo*, cit., p. 159.

²¹ MORDECHAI RABELLO, *Non liquet. Dal Diritto Moderno al Diritto Romano*, cit., p. 34.

²² Il riferimento è, chiaramente, a P. GROSSI, *Mitologie giuridiche della modernità*, 3ª ed., Giuffrè, Milano, 2007, *passim*.

²³ GROSSI, *Oltre la legalità*, Laterza, Bari, 2020, p. 22: «[i]l vocabolo/concetto di 'legalità' sa di passato, di passato remoto; è un fossile lontano, che rinvia a un momento in cui il diritto fu ridotto a un solo produttore, lo Stato, e a una sola manifestazione/fonte, la legge».

²⁴ Per tutti, GROSSI, *Sull'esperienza giuridica pos-moderna (riflessioni sull'odierno ruolo del notaio)*, in *Quaderni fiorentini*, Giuffrè, Milano, 2018, pp. 334-336.

²⁵ Sulle teorie c.d. miste, *amplius*, GUASTINI, *Enunciati interpretativi*, cit., pp. 38-40.

²⁶ GUASTINI, *Enunciati interpretativi*, cit., pp. 37-38; ZACCARIA, voce *Intepretazione*, in «Enciclopedia del diritto», Annali V, 2012, p. 695.

²⁷ Di recente, v. TRAPANI, *Creazione giudiziale della norma penale e suo controllo politico*,

declinazioni diverse rispetto a quelle che si erano imposte nell'età dei Lumi: se, infatti, i principi di tassatività, determinatezza e irretroattività della legge penale sfavorevole continuano a trovare nella certezza del diritto la loro matrice e *ratio*²⁸, la riserva di legge, nel nuovo tessuto costituzionale, ha assunto altre istanze a proprio fondamento.

Secondo l'impostazione che si ritiene preferibile, la riserva si caratterizza come assoluta: la sola legge del Parlamento e gli atti ad essa equiparati (decreto-legge e decreto legislativo) possono intervenire sulla materia penale. Le ragioni di questa scelta non affondano le radici in una esigenza di certezza del diritto anche perché, a ben vedere, la fonte di rango secondario potrebbe essere, in taluni casi, più chiara, "precisa" e, quindi, comprensibile rispetto a quella emanata dal Parlamento.

Se, dunque, la *ratio* di certezza ben potrebbe essere assolta anche dalle previsioni (più intellegibili) di rango secondario, è sulla *ratio* di garanzia²⁹ che sembra poggiare la riserva di legge prevista dall'art. 25, secondo comma, c.p. L'ordine di ragioni è triplice: (i) il Parlamento è il 'luogo' rappresentativo della volontà popolare; (ii) solo la legge e gli atti ad essa equiparati possono essere sottoposti a forme di controllo da parte del cittadino, attraverso lo strumento di democrazia diretta del *referendum* abrogativo; (iii) sulla legge ed atti ad essa equiparati si svolge il controllo di legittimità costituzionale da parte della Corte.

La riserva di legge, ormai quarant'anni or sono, affronta un tormentato 'momento' di crisi che, andando ben oltre l'eufemismo temporale, scivola

cit., *passim*, spec. pp. 14 ss., che, evidenziando come la particolarità del diritto penale è che esso non viva al di fuori del processo e valorizzando l'idea – pienamente condivisibile – della norma penale reale (M. GALLO, *Diritto penale italiano*, 3^a ed., Giappichelli, 2020, pp. 14-16), ritiene che il giudice (penale) sia in realtà il solo creatore del diritto.

²⁸ TRAPANI, voce *Legge penale*, in «Enciclopedia giuridica Treccani», vol. XXVIII, 1990, p. 3; BRICOLA, *Commento all'art. 25, commi 2 e 3*, in *Commentario della Costituzione*, a cura di G. Branca, Zanichelli, Bologna, 1981, pp. 231-233.

²⁹ TRAPANI, voce *Legge penale*, cit., 3; GALLO, *Diritto penale italiano*, cit., 2020, 34 ss.; BRICOLA, *Commento all'art. 25, commi 2 e 3*, cit., p. 232; G. DELITALA, *Cesare Beccaria e il problema penale*, in «Rivista italiana di diritto e procedura penale», 1964, p. 968: «la sola ragione che giustifichi la scelta del potere legislativo come unico detentore del potere normativo in materia penale, risiede nella rappresentatività di quel potere, nel suo essere espressione non di una stretta oligarchia, ma dell'intero popolo, che attraverso i rappresentanti, si attende che l'esercizio avvenga non già arbitrariamente, ma per il suo bene nel suo interesse»; F. MANTOVANI, *Diritto penale*, 11^a ed., Wolters Kluwer-Cedam, Milano, 2020, p. 52; G. FIANDACA, *Crisi della riserva di legge e disagio della democrazia rappresentativa nell'età del protagonismo giurisdizionale*, in «Criminalia», 2011, p. 82; MASSARO, *Determinatezza della norma penale e calcolabilità giuridica*, cit., p. XIV.

e peggiora verso un buio sempre più pesto³⁰.

A contribuire a questa crisi sono stati tanto fattori endogeni, che hanno permesso, per un motivo o per un altro, una crescente espansione dell'interpretazione giudiziale – sia in funzione 'integrativa' che creativa vera e propria – quanto fattori esogeni. Entro il primo fenomeno si è soliti collocare la crisi del Parlamento, della volontà rappresentativa, con contestuale svuotamento della *ratio* di garanzia della riserva di legge, cui si aggiunge una più generale crisi della fattispecie³¹; nel secondo, invece, assumerebbe centralità la crisi del sistema delle fonti³², ampliato dei trattati europei e internazionali, che ha ormai dovuto abbandonare la solida, perfetta e rassicurante forma della 'piramide', frantumata in un 'arcipelago'³³, il quale sovente assume le forme del 'labirinto'³⁴.

Se il Parlamento è ormai identificato come il primo colpevole di questa crisi, ché ha perso credibilità a furia di leggi malfatte, oscure, indecifrabili³⁵

³⁰ Il primo ad 'accendere i riflettori' sulla crisi della riserva di legge fu BRICOLA, *Legalità e crisi: l'art. 25, commi 2° e 3°, della Costituzione rivisitato alla fine degli anni '70*, in «Questione criminale», 1/1980, pp. 182-183; F. GIUNTA, *Verso un equivalente funzionale della riserva di legge?*, in «Criminalia», 2011, pp. 77-78. Per un chiaro, completo e colorato mosaico in cui oggi si frantuma la crisi, per tutti, MASSARO, *Determinatezza della norma penale e calcolabilità giuridica*, cit., pp. 2-9, cui si rimanda anche per ulteriori indicazioni bibliografiche.

³¹ N. IRTI, *La crisi della fattispecie*, in «Rivista di diritto processuale», 2014, pp. 36-44.

³² Sulla crisi dello Stato che si estende alle fonti, G. PINO, *La gerarchia delle fonti del diritto. Costruzione, decostruzione, ricostruzioni*, in «Ars interpretandi», 1/2011, pp. 19 ss.

³³ ZACCARIA, *Trasformazione e riarticolazione delle fonti del diritto, oggi*, in «Ragion Pratica», 1/2004, p. 116: «più che dell'antica piramide, capace di conferire organicità, compattezza e coerenza al sistema delle fonti, si può forse parlare – con riferimento al panorama attuale – di un complesso e mobile arcipelago di fonti in competizione tra loro, condizionato da elementi di natura fattuale più che normativa, da tentativi di regolazione impressi da punti diversi, anche non giuridici, e perciò caratterizzato, nella sua evoluzione, da grande fluidità, variabilità e instabilità. Il sistema di produzione del diritto diviene notevolmente complesso: dal momento che ciascuna fonte tende fisiologicamente a dilatare il proprio spazio di intervento nel nuovo contesto solo parzialmente regolato, il conseguimento di equilibri pur sempre provvisori, è faticoso e problematico». Non mancano, poi, Autori che si affidano alla metafora della rete, per una sintesi della letteratura sul punto v. MASSARO, *Determinatezza della norma penale e calcolabilità giuridica*, cit., p. 5, nota 25; PINO, *La gerarchia delle fonti del diritto*, cit., p. 21, nota 6.

³⁴ V. MANES, *Il giudice nel labirinto. Profili delle intersezioni tra diritto penale e fonti sovranazionali*, Dike, Roma, 2012, p. 5 ss.

³⁵ G. PECORELLA, *Law in action: per una nuova legalità*, in «Diritto di difesa», 17 novembre 2020, 3-4; ID., *La crisi della legalità come crisi della democrazia rappresentativa*, in «Diritto penale e processo», 7/2018, pp. 945-946.

ed ha delegato, sotto il segno di una perenne urgenza, le principali riforme che si sono succedute di recente³⁶, la crisi della fattispecie è meno intuitiva nelle cause prossime, ma forse più significativa negli effetti.

In questa prospettiva giocherebbe un ruolo principale la stessa formulazione della Costituzione, che, in particolare, sfuggirebbe alla puntuale e nitida fotografia del giudizio ipotetico, ma, collocandosi su un piano assiologico e introducendo valori³⁷, sembrerebbe aver incoraggiato la interpretazione ‘secondo principi’, a scapito della sussunzione sotto il paradigma del ‘se A, allora B’. È proprio su quell’astrazione che caratterizza la costruzione della fattispecie, d’altronde, che i principali caratteri della legge – generalità, astrattezza e quindi ripetibilità – si condenserebbero, innalzandola ad autentico strumento per pensare «l’avvenire»³⁸. Sotto questo profilo, la fattispecie diventa lo strumento principe con cui il legislatore guarda al futuro, *rectius*: dispone per il futuro. È forse questo guardare avanti, che è l’essenza del ‘diritto’, rischia di essere il valore maggiormente sacrificato dall’utilizzo di sistemi di intelligenza artificiale nel diritto penale. Ma di ciò meglio in seguito.

La crisi della fattispecie, come acutamente evidenziato, fra l’altro, si traduce in una centralità della casistica, a discapito del caso: all’applicazione logica viene così a sostituirsi una applicazione statistica³⁹, con la conseguenza che assume sempre maggiore rilevanza il ‘precedente’.

D’altro canto, il panorama sovranazionale e internazionale non solo (ri)disegna il nostro sistema delle fonti, ma concorre a (ri)scrivere i nostri principi.

È proprio dalla Corte di Strasburgo che provengono le indicazioni più innovative sul ‘volto attuale’ della legalità.

³⁶ Basti qui pensare al superamento degli O.P.G. e all’introduzione delle R.E.M.S., avvenuto con d.l. 22 dicembre 2011, n. 211.

³⁷ IRTI, *La crisi della fattispecie*, cit., pp. 38 e 41 sulle norme costituzionali, che l’Illustre Autore definisce come «norme senza fattispecie».

³⁸ *Ibidem*, 36: «[i]l diritto ha la *pretesa di dominare l'imprevedibile*, di attribuire significato agli eventi futuri. Quando i codici civili (e l’art. 11 delle *disposizioni preliminari* nel nostro) ripetono con enfasi solenne: ‘la legge non dispone che per l’avvenire’, essi esprimono la volontà di dominare l’imprevedibile, d’impossessarsi di fatti non ancora accaduti o non ancora compiuti. Il diritto non vuole salvare il passato, ma conquistare il futuro: l’irretroattività è soltanto l’altro lato di questa chiamata verso il domani» [corsivi del testo]; ID., *Per un dialogo sulla calcolabilità giuridica*, in A. Carleo (a cura di), *Calcolabilità giuridica*, Il Mulino, Bologna, 2017, pp. 17-29, spec. 17: «[g]ià la parola *diritto* – se proviene dal latino ‘dirigere’ – indica un volgersi, un andare verso, e dunque si apre ad un dopo» [corsivi dell’Autore].

³⁹ IRTI, *La crisi della fattispecie*, p. 43.

Come noto, con l'obiettivo di permettere una applicazione uniforme dei diritti riconosciuti nella Convenzione, la Corte Edu, superando i particolarismi nazionali, ha elaborato le nozioni autonome di 'legge' e 'materia penale'⁴⁰. Per quel che qui interessa, sembrerebbe che il 'law', nella geometria del pensiero della Corte, si connoti per il suo carattere 'prescrittivo', da intendersi come possibilità di ricavare dal medesimo 'regole di condotta' sufficientemente stabili e capaci di porre il consociato, sia pur con le precisazioni che seguiranno, nella condizione di 'calcolare' le conseguenze giuridico-sanzionatorie derivanti dalle proprie azioni. Senza andare troppo oltre, basterebbe qui ricordare come la legge penale – come un autorevole insegnamento già suggerisce⁴¹ – dovrebbe essere *idonea a funzionare come comando*: sembrerebbe, quindi, che sia proprio per la funzione svolta che la legge si connoti. Detto altrimenti: è possibile parlare di 'legge', nella misura in cui dalla medesima sia possibile ricavare una regola di condotta cogente. Ponendo l'accento sulla natura 'prescrittiva' della regola, è, poi, superato lo scoglio di enucleare un concetto di genere fra sistemi profondamente diversi tra loro – come il *common law* ed il *civil law* – e, a monte, si offre una tutela valida a coinvolgere tanto il diritto scritto quanto quello non scritto – qualsiasi sia il modello giuridico entro cui si colloca lo Stato interno.

Affinché sia rispettata la base legale richiesta dall'art. 7 CEDU⁴², la 'legge', secondo le indicazioni della Corte, deve essere 'accessibile' e 'ragionevolmente prevedibile'⁴³: si tratta di connotati che impongono la definizione dell'oggetto (*cosa* deve essere 'accessibile' e 'ragionevolmente prevedibile') e l'individuazione di un parametro (*per chi* deve essere 'accessibile' e 'ragionevolmente prevedibile')⁴⁴.

Secondo una convincente impostazione, oggetto dell'accessibilità altro non sarebbe che la disposizione: si tratterebbe di indicazioni rivolte in

⁴⁰ *Amplius*, MASSARO, *Appunti di diritto penale europeo*, Giappichelli, Torino, 2020, cap. III.

⁴¹ Da ultimo, M. GALLO, *Stimatissimo Professor Kelsen*, Giappichelli, Torino, 2020, 24: «Tornerò più avanti sulla questione sopra sfiorata, adesso ancora qualche cenno sul comando. La norma ha struttura di giudizio ipotetico; sua funzione: il comando. Riprendo quanto ho già detto e lo completo con la precisazione: la norma tende ad essere un comando».

⁴² L'art. 7 CEDU, rubricato '*Nulla poena sine lege*' dispone: «[n]essuno può essere condannato per una azione o una omissione che, al momento in cui è stata commessa, non costituiva reato secondo il diritto interno o internazionale. Parimenti, non può essere inflitta una pena più grave di quella applicabile al momento in cui il reato è stato commesso».

⁴³ Il caso che 'inaugura' la nozione autonoma di legge penale è Corte EDU, 26 aprile 1979, n. 6538/74, *Sunday Times c. Regno Unito*, § 46: «the word 'law' in the expression 'prescribed by law' covers not only statute but also unwritten law».

⁴⁴ MASSARO, *Appunti di diritto penale europeo*, pp. 97-101.

prima battuta al Legislatore, affinché nella formulazione della fattispecie utilizzi un linguaggio chiaro, preciso e coerente in un'ottica di sistema. Il *deficit* di accessibilità, violando la riserva di legge (intesa quest'ultima come 'regola prescrittiva'), si attergerebbe a vera e propria sanzione nei confronti dello Stato, inadempiente rispetto ai propri doveri di chiarezza nei confronti dei consociati⁴⁵. L'individuazione del parametro risulta, invece, problematico e poco unanime, con opinioni che oscillano per un modello oggettivo o soggettivo. Dalla Corte EDU, con riferimento all'accessibilità, sembrerebbero emergere argomenti a favore di una curvatura prevalentemente soggettiva⁴⁶: il 'vizio' di (in)accessibilità, allora, sulla falsariga di quanto accade con riferimento all'art. 5 c.p., opererebbe come limite alla colpevolezza. L'impressione, tuttavia, è che ai fini di una più corretta e coerente analisi del contenuto, dei limiti e degli effetti (ulteriori, ma non immediati) dell'accessibilità, occorra inevitabilmente muovere dall'oggetto. Se, come anticipato, il *deficit* di accessibilità svuoterebbe di contenuto la riserva di legge, impedendo, in prima battuta, di rinvenire una regola prescrittiva, e, in via consequenziale ma non diretta, al consociato di poter adeguare la propria condotta alla medesima, il parametro non potrebbe che essere rigorosamente oggettivo⁴⁷. Il precetto, cioè, dovrebbe essere inaccessibile per chiunque, con conseguente formula assolutoria 'perché il fatto non è previsto dalla legge come reato': a mancare, in effetti, sarebbe la stessa 'legge'.

Diversamente, la ragionevole prevedibilità avrebbe ad oggetto la norma, e cioè, il prodotto dell'attività interpretativa del giudice: che è tanto

⁴⁵ Si tratta di opinioni espresse, sia pur con specifico riferimento alla scusabilità dell'errore ex art. 5 c.p., da TRAPANI, *Creazione giudiziale della norma penale e suo controllo politico*, cit., p. 55.

⁴⁶ Corte EDU, 20 gennaio 2009, n. 75909/01, *Sud Fondi s.r.l. e altri c. Italia*, § 116; Corte EDU, 25 maggio 1993, n. 14307/88, *Kokkinakis c. Grecia* e le sentenze richiamate in Corte europea dei diritti dell'uomo, *Guida sull'articolo 7 della Convenzione europea dei diritti dell'uomo. Nulla poena sine lege: principio di legalità dei delitti e delle pene*, Aggiornata al 30 aprile 2020, cap. III, disponibile su <<https://www.echr.coe.int>>, cui si rimanda anche per un approfondimento.

⁴⁷ Sembrerebbe muovere in questa direzione, sia pur in maniera esplicita solo con riferimento alla ragionevole prevedibilità, MASSARO, *Appunti di diritto penale europeo*, pp. 97-101; EAD., *Determinatezza della norma penale e calcolabilità giuridica*, cit., p. 153 e ss., spec. p. 167, che, in particolare, identifica nell'art. 5 c.p. un limite della legalità e non della colpevolezza: in questa prospettiva, individuati nell'interprete modello e nella prevedibilità della norma (*rectius*: dell'iter argomentativo attraverso cui si sviluppa la decisione giudiziale), il parametro e l'oggetto del giudizio di inevitabilità ex art. 5 c.p., si ipotizza una applicazione del medesimo nei casi di mutamento giurisprudenziale sfavorevole e di contrasti giurisprudenziali diacronici.

più rilevante, quanto necessaria nel diritto penale, ove la regola non vive al di fuori del processo⁴⁸. In questo modo, la riserva di legge sarebbe rispettata tutte le volte in cui l'interpretazione del giudice risulti, in qualche misura, 'ragionevolmente prevedibile': cioè, detto altrimenti, tutte le volte in cui sia possibile anticipare, in maniera non esatta ma sufficientemente chiara, come una certa condotta verrà valutata dal giudice e quali potranno esserne le conseguenze sul piano sanzionatorio.

Il parametro, secondo l'opzione che sembra più convincente, andrebbe ricercato nell'interprete modello, inteso come «il soggetto in possesso delle conoscenze specialistiche richieste a un operatore giuridico e dunque in grado di interpretare il diritto»⁴⁹: proprio nelle pronunce della Corte che negano una violazione della riserva di legge *sub specie* di ragionevole prevedibilità della norma, il riferimento puntuale e costante è all'eventuale supporto da parte di un esperto del diritto, che sarebbe stato in grado di 'consigliare' il singolo⁵⁰. L'attività interpretativa, in effetti, si esplica su enunciati linguistici connotati da un lessico altamente specialistico, sì che è solo in un orizzonte oggettivo (o oggettivizzato) che potrebbe cogliersi l'irragionevole (im)prevedibilità, a monte, dell'*iter* logico argomentativo del giudicante, e, a valle, del prodotto interpretativo – cioè: la norma.

Nelle pronunce della Corte, fra l'altro, non sfugge come la parte del leone la svolga proprio la ragionevole prevedibilità della norma. Il che, se da un lato ha permesso di 'ammonire' le interpretazioni giurisprudenziali più disinvolte⁵¹, dall'altro ha messo ulteriormente a sistema il progressivo passaggio di centralità dalla disposizione alla norma.

Lo stesso modello decisorio delle Corti europee, in effetti, sembrerebbe avvicinarsi molto più a quello dei sistemi di *common law*, dello *stare decisis*,

⁴⁸ TRAPANI, *Creazione giudiziale della norma penale e suo controllo politico*, cit., *passim*, spec. pp. 14 ss.; M. GALLO, *Diritto penale italiano*, 3^a ed., cit., pp. 14-16.

⁴⁹ MASSARO, *Appunti di diritto penale europeo*, p. 100; EAD., *Determinatezza della norma penale e calcolabilità giuridica*, 138-142.

⁵⁰ MASSARO, *Appunti di diritto penale europeo*, p. 100; EAD., *Determinatezza della norma penale e calcolabilità giuridica*, 138-142; Corte europea dei diritti dell'uomo, *Guida sull'articolo 7 della Convenzione europea dei diritti dell'uomo*, cit., cap. IV.

⁵¹ Si pensi, in Italia, al concorso esterno in associazione mafiosa nel caso Contrada: Corte Edu, 14 aprile 2015, *Contrada c. Italia*, 66655/13. Su cui, MASSARO, *Appunti di diritto penale europeo*, p. 101 ss.; P. GUALTIERI, *Interpretazione, creazione e analogia in materia penale: spunti di riflessione a partire dal caso Contrada*, in «Archivio penale», 1/2018, pp. 1 ss.; O. DI GIOVINE, *Antiformalismo interpretativo: il pollo di Russell e la stabilizzazione del precedente giurisprudenziale*, in «Diritto penale contemporaneo – rivista trimestrale», 2/2015, pp. 16 ss.; M. DONINI, *Il Caso Contrada e la Corte EDU*, in «Rivista italiana di diritto e procedura penale», 1/2016, pp. 346 ss.

rispetto a quelli 'continentali': superando, ulteriormente, anche quelle distanze che sembravano irriducibili proprio per le diversità logiche e ontologiche di ordinamenti fra loro opposti.

In questo caos normativo e interpretativo ecco ora ergersi l'intelligenza artificiale, che, proprio mentre 'il governo delle leggi' vive una patologia senza precedenti e refrattaria alle cure, promette di 'salvare' quell'esigenza di certezza che è presupposto dei nostri sistemi penali.

A fronte dell'espansione sempre maggiore dell'interpretazione soggettiva del giudice si ritornerebbe, con le precisazioni che seguiranno, ad un diritto oggettivo, preciso, calcolabile, esatto, idoneo a mettere ordine a questa Babele che sembra affliggere gli ordinamenti post-moderni. L'idea del giudice come macchina, del ragionamento come operazione sillogistica, anziché risolversi in una metafora sembrerebbe (ancora una volta) trovare una forma nella realtà.

4. *L'intelligenza artificiale: il diritto calcolato*

Come ipotesi di lavoro, si tenterà di chiarire cosa s'intenda per 'intelligenza artificiale', successivamente verrà analizzato il funzionamento dei principali *software* in questo momento in uso nel panorama mondiale e da ultimo si proverà ad approfondire gli effetti che questi sistemi potrebbero apportare al nostro sistema (penale).

Secondo una nota definizione, s'intende per intelligenza artificiale quella disciplina che «studia i fondamenti teorici, le metodologie e le tecniche che consentono di progettare sistemi *hardware* e sistemi di programmi *software* atti a fornire all'elaboratore elettronico prestazioni che, a un osservatore comune, sembrerebbero essere di pertinenza esclusiva dell'intelligenza umana»⁵². Se l'opera definitoria dovesse adempiere alla

⁵² F. AMIGONI, V. SCHIAFFONATI, M. SOMALVICO, voce «Intelligenza artificiale», in «Enciclopedia della Scienza e della Tecnica», 2008, in <<https://www.treccani.it>>, § 1.; cfr. anche M. CRAGLIA (Ed.), A. ANNONI, P. BENZUR, P. BERTOLDI, P. DELIPETREV, G. DE PRATO, C. FEIJOO, E. FERNANDEZ MACIAS, E. GOMEZ, M. IGLESIAS, H. JUNKLEWITZ, M. LÓPEZ COBO, B. MARTENS, S. NASCIMENTO, S. NATIVI, A. POLVORA, I. SANCHEZ, S. TOLAN, I. TUOMI, L. VESNIC ALUJEVIC, *Artificial Intelligence – A European Perspective*, EUR 29425 EN, Publications Office of the European Union, Luxembourg, 2018, secondo cui: «AI is a generic term that refers to any machine or algorithm that is capable of observing its environment, learning, and based on the knowledge and experience gained, taking intelligent action or proposing decisions. There are many different technologies that fall under this broad AI definition. At the moment, ML techniques are the most

sua funzione, a questo punto, il baricentro del discorso non potrebbe che spostarsi sulla definizione di ‘intelligenza umana’, in assenza della quale si rischierebbe di cadere in un vicolo cieco o, peggio, in un circolo vizioso. Ciò che la macchina (*rectius*: il sistema) vuole emulare, è bene chiarirlo, non è l’intelligenza umana *tout court*, ma un tipo di ragionamento (umano) replicabile anche dalla macchina: il ragionamento inferenziale. Un sistema può elaborare una quantità significativa di dati – ed in tempi relativamente brevi – ponendosi come adeguato supporto all’attività umana, specie in un’epoca, come questa, segnata da una mole imponente di dati registrati (c.d. *Big Data*, cui è strettamente collegato il fenomeno di *data mining*, cioè la capacità di un sistema di estrapolare informazioni ‘utili’ da questo magma indefinito di dati).

Il ragionamento attraverso cui il sistema informatico elabora i dati ha un nome: algoritmo. Quest’ultimo non è altro che una procedura ben definita volta alla trasformazione di dati di *input* in dati di *output*⁵³.

Nei sistemi dotati di intelligenza artificiale è richiesto che la macchina sia dotata di un altro algoritmo – c.d. *metalgoritmo*⁵⁴ – capace di costruire da sé nuovi algoritmi, in modo tale da definire un nuovo processo di trasformazione a seconda del problema di volta in volta rappresentato. La macchina, infatti, non è capace di porre domande, ma sa costruire risposte, anche attraverso meccanismi ignoti al suo programmatore. È in questo contesto che va inquadrato il fenomeno del *machine learning*, tale per cui – non diversamente dall’apprendimento umano – la macchina migliora le sue prestazioni a furia dell’esperienza: tanti più sono i dati che la macchina processa, quanto più preciso sarà l’*output*. Una piccola osservazione: gli

widely used»; J. McCARTHY, M.L. MINSKY, N. ROCHESTER, C.E. SHANNON, *A proposal for the Dartmouth summer research project on Artificial Intelligence*, 31 agosto 1955, in <<http://www-formal.stanford.edu>>, ad avviso dei quali: «The artificial intelligence problem is taken to be that of making a machine behave in ways that would be called intelligent if a human were so behaving». Le maggiori definizioni proposte e accettate in ambito internazionale sono riportate da S. SAMOILI, M. LÓPEZ COBO, E. GÓMEZ, G. DE PRATO, F. MARTÍNEZ-PLUMED, B. DELIPETREV, *AI Watch. Defining Artificial Intelligence. Towards an operational definition and taxonomy of artificial intelligence*, EUR 30117 EN, Publications Office of the European Union, Luxembourg, 2020, pp. 17-18. A ben vedere, tutte le definizioni, inclusa quella proposta in questa sede, soffrono di estrema elasticità. Ciò nonostante, appare utile inquadrare, sia pure all’interno di confini così labili, il fenomeno in questione, avendo cura di porre l’accento sul funzionamento e sugli obiettivi che un sistema di intelligenza artificiale si propone.

⁵³ T.H. CORMEN, C.E. LEISERSON, R.L. RIVEST, C. STEIN, *Introduction to Algorithms*, 3ª ed., MIT Press, Cambridge, 2009, p. 5.

⁵⁴ AMIGONI, SCHIAFFONATI, SOMALVICO, voce «Intelligenza artificiale», cit., § 3.

unici poli 'visibili' e 'verificabili' dall'uomo sono i dati di *input* ed i dati di *output*. L'algoritmo, il ragionamento vero e proprio, è precluso all'uomo, il che dispiega i suoi effetti anche sulla prevedibilità del prodotto del ragionamento stesso, essendo plausibile che il sistema operi attraverso procedure non più controllabili dal suo creatore e non sempre verificabili *ex post*⁵⁵. Una macchina così dotata, appare opportuno precisarlo, permette di ottenere dati di *output* in conformità al problema rappresentato e ai dati forniti in *input*, agendo secondo algoritmi dalla stessa costruiti: non è capace di conoscere o approfondire il problema, ma è capace, attraverso un'attività puramente computazionale, di risolverlo.

Una umanizzazione della macchina, quindi, valida a coadiuvare la forza umana in ogni ambito, promettendo maggiori livelli di efficienza.

L'utilizzo di questi sistemi all'interno del processo, come anticipato, sembra carico di promesse: la macchina della giustizia, spesso ingolfata, si dirigerebbe con velocità e precisione. In Cina, ad esempio, è stato istituito il primo Tribunale gestito esclusivamente da *robot*, competente sulle controversie legali relative a questioni digitali e le cui decisioni sono appellabili davanti ad un giudice umano: il Tribunale lavora ininterrottamente, garantendo velocità ed efficienza nel processo e alleggerendo il 'carico' del giudice persona⁵⁶.

In Francia, ancora, è stato creato un *software*, utilizzato soprattutto da avvocati, capace di 'predire' (*rectius*: prevedere) la decisione dell'esito

⁵⁵ P. DOMINGOS, *L'algoritmo definitivo: la macchina che impara da sola e il futuro del nostro mondo*, Bollati Boringhieri, Torino, 2016, pp. 7 ss.; F. PASQUALE, *The black box society: The secret algorithms that control money and information*, Harvard University Press, Cambridge, 2015, pp. 1 ss.; enfatizzano questo aspetto, in dottrina: MANES, *L'oracolo algoritmico e la giustizia penale: al bivio tra tecnologia e tecnocrazia*, in «Discrimen», 15 maggio 2020, 14-15; G. CONTISSA, G. LASAGNI, G. SARTOR, *Quando a decidere in materia penale sono (anche) algoritmi e IA: alla ricerca di un rimedio effettivo*, in «Diritto di internet», 4/2019, pp. 620 e 630; R. BORSARI, *Intelligenza Artificiale e responsabilità penale: prime considerazioni*, in «MediaLaws», 3/2019, pp. 264-265; J. MILLAR, I. KERR, *Delegation, relinquishment, and responsibility: The prospect of expert robots*, in R. Calo, A. M. Froomkin, I. Kerr (a cura di), *Robot Law*, Edward Elgar Publishing, Northampton, 2016, p. 107; C. CASONATO, *Intelligenza artificiale e giustizia: potenzialità e rischi*, «DPCE – online», 3/2020, pp. 3379-3380; A. SIMONCINI, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, in «BioLaw Journal – Rivista di BioDiritto», 1/2019, pp. 78-79; DI GIOVINE, *Il judge-bot e le sequenze giuridiche in materia penale (intelligenza artificiale e stabilizzazione giurisprudenziale)*, in «Cassazione penale» 3/2020, pp. 960-961.

⁵⁶ Si tratta della Beijing internet court (<<http://www.english.bjinternetcourt.gov.cn>>), su altre Corti 'digitalizzate' e, più in generale, sui 'robogiudici', A. CARATTA, *Decisione robotica e valori del processo*, in «Rivista di diritto processuale», 2/2020, pp. 491 ss.

giudiziario⁵⁷: il sistema opera analizzando i precedenti casi emessi dalle Corti così da stabilire, con un certo grado di probabilità, quale sarà, in base agli atti, la decisione del giudice e così ‘supportando’ l’avvocato nella sua strategia difensiva. Semplificando: il *software* è capace di ‘anticipare’ la prossima mossa del giudice.

Ne esce fuori un’idea di diritto calcolabile⁵⁸, che sembra porsi in linea di continuità con quella ‘illuminata’ ed in parte utopica idea di certezza del diritto, che ha contribuito in maniera preponderante a costruire le fondamenta dei nostri sistemi penali.

Il giudice, infatti, sia pur con le dovute precisazioni, di nuovo non sarebbe altro che *bouche de la lois* o *Subsuntionsautomat*: il diritto è calcolabile, l’algoritmo applica la legge che, in sé, diventa prevedibile per tutti⁵⁹. I miti, insomma, ritornano, si solidificano, e anche quando sembrano ormai ‘fossili’ di un passato remoto, resuscitano e si riannodano a quel filo che sembra indistruttibile: in questo caso, la certezza del diritto.

5. *La bellezza del mito: demistificabile, ma non demitificabile*

Non occorre spendere ulteriori parole sui pregi della macchina, già adeguatamente valorizzati: rapidità, efficienza, oggettività, e, ancora, *calcolabilità*.

I limiti, invece, meritano qualche precisazione ulteriore: in questa sede ne sono stati individuati alcuni, ma si tratta di un catalogo aperto, pronto ad essere smentito da una quinta rivoluzione tecnologica, o arricchito, a seconda che si valorizzino le indicazioni dell’attuale diritto positivo – non solo sostanziale, ma anche processuale.

A fini classificatori e limitandosi ad un rapido accenno⁶⁰, meritano

⁵⁷ Rinvenibile al seguente link: <<http://www.predictice.com>>.

⁵⁸ IRTI, *Il tessitore di Goethe (per la decisione robotica)*, in A. Carleo (a cura di), *Decisione robotica*, Il Mulino, Bologna, 2019, pp. 17-22, spec. p. 17, che, muovendo dal concetto di calcolabilità di Max Weber, per cui il capitalismo necessiterebbe di «un diritto che si possa calcolare in modo simile ad una macchina», chiosa constatando che «la similitudine sta cedendo il posto ad una reale identificazione». Il riferimento alla calcolabilità della decisione robotica è inoltre anche in CARATTA, *Decisione robotica e valori del processo*, cit., pp. 491 ss.; MASSARO, *Determinatezza della norma penale e calcolabilità giuridica*, pp. 485 ss.

⁵⁹ *Supra*, § 2.

⁶⁰ Si veda, per un approfondimento, L. NOTARO, *‘Algoritmos predictivos’ y justicia penal desde una perspectiva italiana y europea*, in questo volume.

di essere menzionati gli argini di diritto positivo posti dal legislatore sopranazionale.

Il legislatore europeo, in particolare, ha stabilito all'art. 22 del Regolamento 679/2016 (GDPR) che: «L'interessato ha il diritto di non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato, compresa la profilazione, che produca effetti giuridici che lo riguardano o che incida in modo analogo significativamente sulla sua persona». Non sembra allo stato un freno così significativo, rispetto al quale potrebbe, peraltro, operare il consenso dello stesso interessato, ma è già qualcosa.

In seconda battuta, la Commissione europea per l'efficienza della giustizia (Cepej) ha emanato la Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi⁶¹. Scopo della Carta etica europea è quello di incentivare l'utilizzo 'responsabile' di strumenti di intelligenza artificiale nel processo, enucleando principi che possano definire un modello cui ispirarsi nell'utilizzo, nello sviluppo o nella regolamentazione di questi 'algoritmi intelligenti'. Si tratta di uno strumento di *soft law* attraverso il quale vengono enucleati cinque principi: (i) principio del rispetto dei diritti fondamentali; (ii) principio di non-discriminazione; (iii) principio di qualità e sicurezza; (iv) principio di trasparenza, imparzialità e equità; (v) principio 'del controllo da parte dell'utilizzatore'.

Sembra, invece, maggiormente interessante, in questa sede, esaminare gli effetti che questi *software* produrrebbero all'interno dei nostri sistemi giuridici: l'impressione, più in generale, è che l'analisi condotta 'allo stato degli atti' sui sistemi di intelligenza artificiale possa segnare un significativo mutamento delle basi dei nostri sistemi penali. Si tratterebbe, quindi, di chiarire se ed in che misura ci sia disponibilità a rinunciare alla 'tradizione' come contropartita per un diritto calcolabile e calcolato.

La macchina, come anticipato, è in grado di 'predire' l'esito giudiziario, sempre che questo sia conforme ai *trend* espressi dalle Corti: requisito implicito, quindi, affinché queste macchine operino è che la Corte si ponga in linea di continuità con i precedenti dalla medesima espressi. Messa a bando, di nuovo, l'interpretazione soggettiva del giudice, in qualche misura ritorna ma come precedente 'stabile' o quanto meno stabilizzato – e quindi, sembrerebbe, oggettivizzato. I nostri ordinamenti di *civil law* farebbero bancarotta in favore di un sistema, a questo punto: *ufficialmente*, basato sul vincolo del precedente. Il primo 'prezzo da pagare' sarebbe, quindi, un altro

⁶¹ Consultabile su <<https://www.rm.coe.int>>.

mito: quello del giudice sottoposto alla legge, destinata (ancora una volta) a soccombere in favore della giurisprudenza.

Il diritto oggettivamente oggettivizzato imporrebbe una calcolabilità non solo del precetto, ma anche della sanzione. Dubbi sorgono, al proposito, con riferimento a quegli spazi di discrezionalità che sono riservati e imposti per legge al giudice in sede di applicazione della pena. La finalità rieducativa della pena ed il principio di personalità della responsabilità penale – che, come noto, ha un'estensione maggiore del suo più noto corollario principio di colpevolezza – coinvolgono l'intero *iter* della sanzione: muovendo dalla previsione in astratto, passando per la comminazione in concreto e terminando con la fase dell'esecuzione⁶². Ad esser sacrificato sembrerebbe proprio il secondo segmento di questo percorso, con eventuale 'effetto domino' sul terzo; e la domanda potrebbe suonare così: gli spazi di discrezionalità riservati al giudice possono essere riempiti dalle macchine o il diritto esatto, nella sua massima esaltazione, non rischia di sfociare in un'applicazione delle pene fisse, contrarie alla Costituzione?

L'interpretazione, piaccia o non piaccia, è un'attività irrinunciabile dell'applicazione del diritto. È proprio ciò che permette l'applicazione della fattispecie, che, come già anticipato, è ciò che consente al diritto di guardare al futuro – di guardare 'diritto'. Il rischio, altrimenti detto, sarebbe quello di un diritto destinato a stagnare, a replicare se stesso⁶³. Sembra, spingendosi un po' oltre con la metafora, che al diritto meccanicamente calcolabile tocchi la stessa sorte degli indovini della Divina Commedia: costretti a camminare all'indietro, con la testa capovolta, «perché 'l veder

⁶² L'indissolubile legame fra il primo ed il terzo comma dell'art. 27 Cost. è stato evidenziato dalla Corte costituzionale già nella storica sentenza che ha dichiarato la parziale illegittimità dell'art. 5 Cost., laddove non prevede(va) l'ignoranza scusabile (Corte cost., 24 marzo 1988, n. 364). Si tratta, in ogni caso, di osservazioni ormai consolidate nella giurisprudenza costituzionale, sviluppatesi, a partire da Corte cost. 2 luglio 1990, n. 313, che individua nella finalità rieducativa, strettamente collegata al principio di proporzione, un principio che deve operare in tutte le 'fasi' che attengono alla pena – previsione astratta, concreta applicazione, esecuzione. Per un approfondimento su questi temi, G. LEO, *Automatismi sanzionatori e principi costituzionali, voce per 'Il libro dell'anno Treccani 2014'*, in «Diritto penale contemporaneo», 7 gennaio 2014, pp. 1 e ss., cui si rimanda anche per approfondimenti sulla giurisprudenza più recente.

⁶³ MASSARO, *Determinatezza della norma penale e calcolabilità giuridica*, p. 502.; J. NIEVA-FENOLL, *Intelligenza artificiale e processo*, Giappichelli, Torino, 2019, p. 21; C. PARODI, V. SELLAROLI, *Sistema penale e intelligenza artificiale*, in «Diritto penale contemporaneo», 16 giugno, 2019, p. 51.

dinanzi era lor tolto»⁶⁴.

Ultima resta, a questo punto, una domanda, e cioè: cosa resta del mito della certezza del diritto? Probabilmente un mito: demistificabile, ma mai 'demistificabile', pronto a intrufolarsi negli angoli in cui ce ne sia bisogno e a porsi come valore a cui tendere. E valgano le parole espresse, al proposito, da Marcello Gallo: «siano pure miti. Ché se è bello che si trovino tanti spiriti critici, pronti a spiare, a scoprire e denunciare i miti, i miti duri a morire conservati e insinuati nella vita di oggi, penso per esempio al Barthes delle *Mythologie*, credo anche che dovremmo essere tutti così illuminati da essere capaci di credere ad un mito, sapendo che è un mito»⁶⁵.

6. Bibliografia

- ALIGHIERI D., *Divina Commedia, Inferno*, canto XX.
- ALVAZZI DEL FRATE P., *Giurisprudenza e référé législatif in Francia nel periodo rivoluzionario e napoleonico*, Giappichelli, Torino, 2005.
- ALVAZZI DEL FRATE P., *L'interpretazione autentica nel XVIII secolo. Divieto di interpretatio e 'riferimento al legislatore' nell'illuminismo giuridico*, Giappichelli, Torino, 2000.
- AMARELLI G., *Il giudice ed il rispetto della legge penale in sede interpretativa. Obsolescenza apparente e attualità permanente del pensiero di Beccaria*, in <<https://www.osservatorioaic.it>>, giugno 2015.
- AMIGONI F., V. SCHIAFFONATI V., SOMALVICO M., in «Enciclopedia della Scienza e della Tecnica», 2008, in <<https://www.treccani.it>>.
- ARCHI G. G., *Interpretatio iuris - interpretatio legis interpretatio legum*, in *Studi in onore di Francesco Santoro-Passarelli*, vol. VI, Jovene, Napoli, 1972.
- BASSANELLI SOMMARIVA G., *L'imperatore unico creatore ed interprete delle leggi e l'autonomia del giudice nel diritto giustiniano*, Giuffrè, Milano, 1983.
- BECCARIA C., *Dei delitti e delle pene*, Nicolo Bettoni, Brescia, 1807.
- BETTI E., *Diritto romano e dogmatica odierna*, in «Archivio giuridico Filippo Serafini», 1928.
- BOBBIO N., *La certezza del diritto è un mito?*, in «Rivista Internazionale di Filosofia del Diritto», estratto, 1/1951, Fratelli Bocca – Editori, Milano-Roma-Firenze.

⁶⁴ D. ALIGHIERI, *Divina Commedia, Inferno*, canto XX.

⁶⁵ GALLO, *Diritto penale italiano*, cit., pp. XXIII-XXIV.

- BOBBIO N., *Teoria generale del diritto*, Giappichelli, Torino, 1993.
- BORSARI R., *Intelligenza Artificiale e responsabilità penale: prime considerazioni*, in «MediaLaws», 3/2019.
- BRICOLA F., *Commento all'art. 25, commi 2 e 3*, in *Commentario della Costituzione*, a cura di G. Branca, Zanichelli, Bologna, 1981.
- BRICOLA F., *Legalità e crisi: l'art. 25, commi 2° e 3°, della Costituzione rivisitato alla fine degli anni '70*, in «Questione criminale», 1/1980.
- BURDESE A., voce *Interpretazione nel diritto romano*, in «Digesto delle discipline privatistiche», 1993.
- CALAMANDREI P., *La Cassazione civile. Parte prima*, in *La memoria del diritto*, ristampa in alterata, Roma TrE-Press, Roma, 2019.
- CARATTA A., *Decisione robotica e valori del processo*, in «Rivista di diritto processuale», 2/2020.
- CASONATO C., *Intelligenza artificiale e giustizia: potenzialità e rischi*, «DPCE – online», 3/2020.
- CENDERELLI V., voce *Fonti del diritto in diritto romano*, in «Digesto delle discipline privatistiche», 1992.
- CONDORCET M.J.A.N. (DE CARITAT), *Idées sur le despotisme, à l'usage de ceux qui prononcent ce mot sans l'entendre* (1789), Firmin Didot frères, Paris, 1847.
- CONTISSA G., LASAGNI G., SARTOR G., *Quando a decidere in materia penale sono (anche) algoritmi e IA: alla ricerca di un rimedio effettivo*, in «Diritto di internet», 4/2019.
- CORMEN T.H., LEISERSON C.E., RIVEST R.L., STEIN C., *Introduction to Algorithms*, 3ª ed., MIT Press, Cambridge, 2009.
- CRAGLIA M. (Ed.), ANNONI A., BENCZUR P., BERTOLDI P., DELIPETREV P., DE PRATO G., FEIJOO C., FERNANDEZ MACIAS E., GOMEZ E., IGLESIAS M., JUNKLEWITZ H., LÓPEZ COBO M., MARTENS B., NASCIMENTO S., NATIVI S., POLVORA A., SANCHEZ I., TOLAN S., TUOMI I., VESNIC ALUJEVIC L., *Artificial Intelligence – A European Perspective*, EUR 29425 EN, Publications Office of the European Union, Luxembourg, 2018.
- CRISAFULLI V., voce *Disposizione (e norma)*, in «Enciclopedia del diritto», XII, 1964.
- DE MARTINO F., *Storia della Costituzione romana*, vol. I, Jovene, Napoli, 1951.
- DELITALA G., *Cesare Beccaria e il problema penale*, in «Rivista italiana di diritto e procedura penale», 1964.
- DI GIOVINE O., *Antiformalismo interpretativo: il pollo di Russell e la stabilizzazione del precedente giurisprudenziale*, in «Diritto penale

- contemporaneo – rivista trimestrale», 2/2015.
- DI GIOVINE O., *Il judge-bot e le sequenze giuridiche in materia penale (intelligenza artificiale e stabilizzazione giurisprudenziale)*, in «Cassazione penale» 3/2020.
- DOMINGOS P., *L'algoritmo definitivo: la macchina che impara da sola e il futuro del nostro mondo*, Bollati Boringhieri, Torino, 2016.
- DONINI M., *Il Caso Contrada e la Corte EDU*, in «Rivista italiana di diritto e procedura penale», 1/2016.
- FERRAJOLI L., *Diritto e ragione*, 9^a ed., Laterza, Bari, 2008.
- FIANDACA G., *Crisi della riserva di legge e disagio della democrazia rappresentativa nell'età del protagonismo giurisdizionale*, in «Criminalia», 2011.
- GALLO F., «*Princeps*» e «*ius praetorium*», in «Rivista di Diritto Romano», I/2001.
- GALLO M., *Diritto penale italiano*, 3^a ed., Giappichelli, Torino, 2020.
- GALLO M., *Stimatissimo Professor Kelsen*, Giappichelli, Torino, 2020.
- GIUNTA F., *Verso un equivalente funzionale della riserva di legge?*, in «Criminalia», 2011.
- GORLA G., voce *Giurisprudenza*, in «Enciclopedia del diritto», XIX, 1970.
- GROSSI P., *Mitologie giuridiche della modernità*, 3^a ed., Giuffrè, Milano, 2007.
- GROSSI P., *Oltre la legalità*, Laterza, Bari, 2020.
- GROSSI P., *Sull'esperienza giuridica pos-moderna (riflessioni sull'odierno ruolo del notaio)*, in *Quaderni fiorentini.*, Giuffrè, Milano, 2018.
- GUALTIERI P., *Interpretazione, creazione e analogia in materia penale: spunti di riflessione a partire dal caso Contrada*, in «Archivio penale», 1/2018.
- GUASTINI R., *Enunciati interpretativi*, in «Ars interpretandi», 2/1997.
- IRTI N., *Il tessitore di Goethe (per la decisione robotica)*, in A. Carleo (a cura di), *Decisione robotica*, Il Mulino, Bologna, 2019.
- IRTI N., *La crisi della fattispecie*, «Rivista di diritto processuale», 2014.
- LEO G., *Automatismi sanzionatori e principi costituzionali*, voce per 'Il libro dell'anno Treccani 2014', in «Diritto penale contemporaneo», 7 gennaio 2014.
- LITEWSKI W., *Consultatio ante sententiam*, in «Zeitschrift der Savigny-Stiftung für Rechtsgeschichte (Romanistische Abteilung)», LXXXVI, 1969.
- MANES V., *Il giudice nel labirinto. Profili delle intersezioni tra diritto penale e fonti sovranazionali*, Dike, Roma, 2012.
- MANES V., *L'oracolo algoritmico e la giustizia penale: al bivio tra tecnologia e tecnocrazia*, in «Discrimen», 15 maggio 2020.

- MANTOVANI F., *Diritto penale*, 11^a ed., Wolters Kluwer-Cedam, Milano, 2020.
- MASSARO A., *Appunti di diritto penale europeo*, Giappichelli, Torino, 2020.
- MASSARO A., *Determinatezza della norma penale e calcolabilità giuridica*, Editoriale Scientifica, Napoli, 2020.
- MCCARTHY J., MINSKY M.L., ROCHESTER N., SHANNON C.E., *A proposal for the Dartmouth summer research project on Artificial Intelligence*, 31 agosto 1955, in <<http://www-formal.stanford.edu>>.
- MILLAR J., KERR I., *Delegation, relinquishment, and responsibility: The prospect of expert robots*, in R. Calo, A. M. Froomkin, I. Kerr (a cura di), *Robot Law*, Edward Elgar Publishing, Northampton, 2016.
- MONTESQUIEU CH. L. (DE SECONDAT), *De l'esprit des loix*, t. I, Barillot & Fils, Genève, 1748.
- MORDECHAI RABELLO A., Non liquet. *Dal Diritto Moderno al Diritto Romano*, in «Archivio giuridico Filippo Serafini», 185, 1973.
- N. IRTI, *Per un dialogo sulla calcolabilità giuridica*, in A. Carleo (a cura di), *Calcolabilità giuridica*, Il Mulino, Bologna, 2017.
- NIEVA-FENOLL J., *Intelligenza artificiale e processo*, Giappichelli, Torino, 2019.
- PARODI C., SELLAROLI V., *Sistema penale e intelligenza artificiale*, in «Diritto penale contemporaneo», 16 giugno, 2019.
- PASQUALE F., *The black box society: The secret algorithms that control money and information*, Harvard University Press, Cambridge, 2015.
- PECORELLA G., *La crisi della legalità come crisi della democrazia rappresentativa*, in «Diritto penale e processo», 7/2018.
- PECORELLA G., *Law in action: per una nuova legalità*, in «Diritto di difesa», 17 novembre 2020.
- PINO G., *La gerarchia delle fonti del diritto. Costruzione, decostruzione, ricostruzioni*, in «Ars interpretandi», 1/2011.
- PIOLA CASELLI E., voce *Giurisprudenza*, in «Digesto italiano», XII, 1900-1904.
- PULIATTI S., *Il diritto prima e dopo Costantino*, in *Enc. costantiniana*, 2013, disponibile su <<http://www.treccani.it>>.
- SAMOILI S., LÓPEZ COBO M., GÓMEZ E., DE PRATO G., MARTÍNEZ-PLUMED F., DELIPETREV B., *AI Watch. Defining Artificial Intelligence. Towards an operational definition and taxonomy of artificial intelligence*, EUR 30117 EN, Publications Office of the European Union, Luxembourg, 2020.
- SCHIAVONE A., *Il «nome» e «la cosa»*, in *Quaderni fiorentini*, Giuffrè, Milano, 1978.
- SCHIAVONE A., *Ius. L'invenzione del diritto in Occidente*, Einaudi, Torino, 2017.

- SERRAO F., voce *Interpretazione della legge (dir. rom)*, in «Enciclopedia del diritto», XXII, 1972.
- SERRAO F., voce *Interpretazione della legge (dir. rom)*, in «Enciclopedia del diritto», XXII, 1972.
- SIMONCINI A., *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, in «BioLaw Journal – Rivista di BioDiritto», 1/2019.
- TRAPANI M., *Creazione giudiziale giudiziale della norma penale e suo controllo politico*, in «Archivio penale», 1/2017.
- TRAPANI M., voce *Legge penale*, in «Enciclopedia giuridica Treccani», vol. XXVIII, 1990.
- VERRI P., *Sulla interpretazione delle leggi (1765)*, in «Il Caffè», tomo II, foglio XXVIII, consultabile su <<http://www.illuminismolombardo.it>>.
- ZACCARIA G., *Trasformazione e riarticolazione delle fonti del diritto, oggi*, in «Ragion Pratica», 1/2004.
- ZACCARIA G., voce *Intepretazione*, in «Enciclopedia del diritto», Annali V, 2012.

CAPÍTULO V

Deshumanizando la culpabilidad: los sistemas inanimados en la teoría del delito

Angelo Giraldi

SUMARIO: 1. Consideraciones introductorias – 2. Aproximaciones a la vulnerabilidad del concepto de culpabilidad – 3. La incidencia de los sistemas de inteligencia artificial en la culpabilidad como presupuesto de la punibilidad – 3.1. El tentativo de reconocer una nueva subjetividad pasiva: la inteligencia artificial como víctima de delitos – 3.2. Hacia una subjetividad activa instrumental: la inteligencia artificial como autora de delitos – 3.3. La inteligencia artificial como instrumento de apoyo para el ejercicio de la justicia – 4. La incidencia de los sistemas de inteligencia artificial en la culpabilidad como criterio para la determinación de la pena en concreto – 5. Conclusiones – 6. Bibliografía.

1. *Consideraciones introductorias*

La estructura del Derecho penal contemporáneo se ha plasmado en los siglos a través de los reflejos sociales de la ciencia y de la técnica. No es ninguna novedad que la ciencia jurídica no pueda (o no deba) prescindir de los avances tecnológicos y sociales, a su vez influenciados por los acontecimientos históricos y el desarrollo de los demás ámbitos del conocimiento¹, pues, de hacerlo, se arriesgaría a que la regulación de toda conducta humana –y, a esta altura, también inhumana²– quedase sin

¹ El propio concepto de culpabilidad, así como las teorías de la pena, representan el resultado de reflexiones doctrinales elaboradas en el marco de determinadas realidades históricas y sociales. Al respecto, v. D. SANTAMARIA, voz *Colpevolezza*, en «Enciclopedia del diritto», VII, 1960, pp. 646 ss.; G. MARINI, voz *Colpevolezza*, en «Digesto delle Discipline Penalistiche», II, 1988, pp. 314 ss.; A. VALLINI, *Antiche e nuove tensioni tra colpevolezza e diritto penale artificiale*, Giappichelli, Torino, 2003, pp. 9 ss.; H.H. JESCHECK, *Evolución del concepto jurídico penal de culpabilidad en Alemania y Austria*, en «Revista Electrónica de Ciencia Penal y Criminología», 2003, pp. 2 ss.; J.M. NAVARRETE URIETA, *El nacimiento de la idea de culpabilidad*, en «Aldaba», 18/1992, pp. 31 ss.

² Con las precisiones debidas, que se detallarán *infra*. Hasta la fecha se utiliza el adjetivo 'artificial' en la locución 'inteligencia artificial' para subrayar que, aunque adquieren

efecto alguno. Con ello, sólo se pretende insistir en que una regulación que carece de efectividad refleja un estado anacrónico del ordenamiento jurídico que conlleva riesgos para la seguridad colectiva, objeto último de tutela por el Derecho, especialmente desde una perspectiva penal³. En otras palabras, lo que se pretende es proteger al individuo a través de un «margen de *seguridad cognitiva* frente a la posibilidad de ataques a sus intereses por parte de otros individuos»⁴.

No puede omitirse que las reflexiones que siguen están encaminadas al análisis de cuestiones futuras desde una perspectiva limitada, abordada desde la visión peculiar del Derecho penal. Como es lógico, el análisis ‘integrado’ de las cuestiones afectadas por la inteligencia artificial supone un conocimiento global de los asuntos científicos que, sin duda alguna, exceden del perímetro de estas consideraciones. Así que, sin alguna pretensión de exhaustividad, este escrito persigue a la vez un doble objetivo: por un lado, intentar averiguar el *an* y el *quomodo* de la influencia de los sistemas de inteligencia artificial en la teoría del delito y, en particular, en los elementos que –con fundadas razones de discrepancia en la doctrina⁵–

autonomía en el aprendizaje y en la orientación de los procesos causales, estos sistemas son el fruto de la creación por parte del ser humano.

³ En las palabras de G. JAKOBS, *El principio de culpabilidad*, en «Anuario de Derecho Penal y Ciencias Penales», 3/1992, p. 1074, la finalidad de prevención general positiva de la pena consta de «prevención general, porque pretende dirigirse un efecto en *todos* los ciudadanos; [y se denomina] positiva, porque este efecto no se pretende en que consista en miedo ante la pena, sino en una tranquilización en el sentido de que la norma esta vigente, de que la vigencia de la norma, que se ha visto afectada por el hecho, ha vuelto a ser fortalecida por la pena». Al respecto, v. también C. ROXIN, *Derecho Penal. Parte General. Tomo I. Fundamentos. La estructura de la teoría del delito*, Civitas, Madrid, 1997, p. 97, traducción de D.M. Luzón Peña, M. Díaz y García Conlledo, J. de Vicente Remesal: «[e]n primer lugar, el fin de la conminación penal es de pura prevención general».

⁴ R. ALCÁCER GUIRAO, *Los fines del Derecho Penal. Una aproximación desde la filosofía política*, en «Anuario de Derecho Penal y Ciencias Penales», 1/1998, p. 425: «En palabras de Vehling, el Derecho sirve al fin de “allanar los momentos de inseguridad respecto a la conducta de los otros en los contactos sociales a través de un mínimo de orientación jurídica”». No debe extrañar que la regulación de las conductas humanas, y en particular la represión de algunas de estas a través del Derecho penal, estén (y tiendan a modificarse para estar) ancladas a lo que la razón humana percibe en un determinado momento histórico. Con las palabras *ante litteram* de B. SPINOZA, *Tratado político*, Alianza Editorial, Madrid, 1986, p. 96, traducción de Atilano Domínguez: «no resulta tan inadecuado que los hombres que están habituados a vivir en el Estado llamen pecado a lo que contradice el dictamen de la razón, puesto que los derechos del mejor Estado [...] deben estar fundados en ese dictamen».

⁵ Sobre la evolución del concepto de culpabilidad, v. *amplius* Á. TORIO LÓPEZ, *El concepto individual de culpabilidad*, en «Anuario de Derecho Penal y Ciencias Penales», 2/1985, pp. 285 ss.

conformarían el núcleo de la llamada culpabilidad; por otro, hacer hincapié en la influencia y en los cambios que la introducción masiva de sistemas inanimados autónomos en la realidad social tendrá en el Derecho penal.

Tradicionalmente, cuando se empezó a hablar de inteligencia artificial, en términos muy amplios y cada vez más acogedores, se hizo especial referencia al ámbito de la privacidad y de los datos personales. La masificación de los datos facilitados o introducidos en programas informáticos, bases de datos, aplicaciones y servidores representaba un peligro para la intimidad personal, protegida en España por el propio art. 18 de la Constitución⁶ y, como reflejo, por la Ley penal. No es casualidad que, hasta el día de hoy, los órganos jurisdiccionales se vean implicados en el pronunciamiento de resoluciones que de alguna manera limiten el empleo de la inteligencia artificial⁷.

A título ejemplificativo, cabe mencionar la reciente sentencia de 21 de junio de 2022, dictada por el Tribunal de Justicia de la Unión Europea en al asunto C-817/19. El asunto está relacionado con la Directiva *Passenger Name Record* (PNR)⁸ y su incorporación al ordenamiento interno de Bélgica. Dicha Directiva impone el tratamiento sistemático de un número considerable de datos de los pasajeros aéreos de los vuelos que entran y salen de la Unión Europea, y excepcionalmente también de los vuelos ‘intra-UE’,

⁶ No se prevé análoga expresa previsión normativa en la Constitución italiana, esencialmente por razones históricas. Antes de la promulgación del Convenio Europeo de Derechos Humanos en 1950, la Constitución italiana tutelaba la intimidad personal desde una perspectiva entonces estática, principalmente obstativa, que se desprende de sus arts. 14 («El domicilio es inviolable»), 15 («La libertad y el secreto de la correspondencia y de cualquier otra forma de comunicación son inviolables») y 21 («Todos tienen derecho a manifestar libremente su pensamiento con la palabra, el escrito y cualquier otra forma de difusión»). Con la entrada en vigor del Convenio Europeo de Derechos Humanos, entre los derechos fundamentales mencionados en el art. 2 de la Constitución italiana se asoma el contenido del art. 8 CEDH: «1. Toda persona tiene derecho al respeto de su vida privada y familiar, de su domicilio y de su correspondencia. 2. No podrá haber injerencia de la autoridad pública en el ejercicio de este derecho sino en tanto en cuanto esta injerencia esté prevista por la ley y constituya una medida que, en una sociedad democrática, sea necesaria para la seguridad nacional, la seguridad pública, el bienestar económico del país, la defensa del orden y la prevención de las infracciones penales, la protección de la salud o de la moral, o la protección de los derechos y las libertades de los demás».

⁷ Por razones metodológicas y de espacio, no procede examinar en esta sede el panorama de las resoluciones judiciales que mencionan, directa o indirectamente, la inteligencia artificial. Para un *excursus* de algunas resoluciones, v. D.L. MORILLAS FERNÁNDEZ, *Implicaciones de la inteligencia artificial en el ámbito del Derecho Penal*, en esta obra, § 3.

⁸ Directiva (UE) 2016/681 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativa a la utilización de datos del registro de nombres de los pasajeros (PNR) para la prevención, detección, investigación y enjuiciamiento de los delitos de terrorismo y de la delincuencia grave.

con el fin de luchar contra el terrorismo y demás delitos considerados ‘graves’. No es necesario detenerse sobre el asunto específico, pero resulta curioso mencionar la referencia (quizás anacrónica) que el Tribunal de Justicia hizo a los sistemas de inteligencia artificial.

Con arreglo a la Directiva, la Unidad de Información sobre Pasajeros (UIP) de cada Estado miembro, para realizar la evaluación de los pasajeros antes de su llegada o salida programada del Estado, puede tratar los *passenger name record* exclusivamente con arreglo a criterios predeterminados⁹. Sobre la base de esta disposición, en la resolución del Tribunal de Justicia de la Unión Europea se establece que la evaluación con arreglo a criterios predeterminados impide la utilización de tecnologías de inteligencia artificial en el ámbito de sistemas de *machine learning*, ya que estos serían capaces de modificar el proceso de evaluación sin que intervenga el control humano¹⁰. En particular, sería capaz de alterar los criterios de evaluación que, por expresa previsión normativa, han de ser predeterminados. Con otras palabras, el Tribunal hace hincapié en la opacidad del funcionamiento de los mecanismos de inteligencia artificial, cuya comprensión de cara a averiguar el alcance de un determinado resultado puede resultar hasta imposible¹¹.

⁹ Art. 6, apartado 3, letra b), de la Directiva (UE) 2016/681.

¹⁰ El § 194 de la resolución de 21 de junio de 2022 establece: «Con respecto a los criterios que el UIP puede utilizar para dicha finalidad, hay que destacar ante todo que, según el propio texto del artículo 6, párrafo 3, letra b), de la Directiva PNR, dichos criterios deben ser “preestablecidos”. Como se ha puesto de relieve por parte del abogado general en el párrafo 228 de sus conclusiones, este requisito impide la utilización de tecnologías de inteligencia artificial en el ámbito de sistemas de autoaprendizaje (*machine learning*), que pueden modificar, sin intervención ni control humanos, el proceso de evaluación y, en particular, los criterios de evaluación sobre los cuales se basa el resultado de la aplicación de dicho proceso, además de la ponderación de tales criterios» (traducción de la versión italiana realizada por el autor, al no existir traducción oficial al español).

¹¹ El Tribunal acoge las observaciones formuladas por el Abogado General en el párrafo 228: «En tercer lugar, tanto del tenor del artículo 6, apartado 3, letra b), de la Directiva PNR como del sistema de garantías que acompaña al tratamiento automatizado de los datos PNR establecido por la Directiva PNR se desprende que el funcionamiento de los algoritmos utilizados para realizar el análisis al que se refiere esta disposición debe ser transparente y permitir la trazabilidad del resultado de su aplicación. Evidentemente, este requisito de transparencia no implica que deban hacerse públicos los “perfiles” utilizados. En cambio, requiere que se garantice que pueda identificarse la toma de decisión algorítmica. En efecto, por una parte, el requisito de que los criterios con arreglo a los cuales debe realizarse este análisis sean “predeterminados” excluye la posibilidad de que puedan modificarse sin intervención humana y se opone, en consecuencia, a la utilización de las tecnologías de inteligencia artificial llamadas “machine learning”, las cuales, al mismo tiempo que pueden presentar un grado más elevado de precisión, son difíciles de interpretar, incluso por parte de los operadores que han efectuado el tratamiento automatizado. [...] La transparencia del funcionamiento de los algoritmos utilizados constituye también un requisito necesario para

En cierta medida, y esta es la razón por la cual se ha mencionado dicha resolución, sorprende la censura de la utilización de sistemas inteligentes inanimados en un sector que sólo afecta a la intimidad y la privacidad de los datos de los pasajeros aéreos, al mismo tiempo que las demás Instituciones europeas están demandando (o al menos autorizando) la implementación regulada de estos avances tecnológicos especialmente en el ámbito de la justicia penal.

Cabe recordar, por cierto, que en algunos Países de la Unión Europea ya se encuentran implementados mecanismos de decisión automatizada en procedimientos judiciales. Sin entrar en detalles, en el Reino Unido (entonces parte de la Unión), Francia y Holanda se han implementado modelos de algoritmos en el ámbito de la justicia civil¹², que posiblemente trascienden el alcance de la privacidad relativa a los datos de los pasajeros aéreos. Asimismo, en España se ha introducido la utilización de *VioGén*, una técnica de policía predictiva encaminada a la valoración policial del riesgo¹³, y de *VeriPol*, el polígrafo utilizado por las comisarías para detectar denuncias falsas¹⁴.

De todos modos, la rápida evolución del estado de la técnica ha intervenido en favor de una reconsideración de la inteligencia artificial y de sus efectos, de manera mucho más amplia y dinámica que la exclusivamente ligada a la intimidad personal. Por dicha razón, se ha ido introduciendo prolíficamente un *corpus* europeo de recomendaciones y normas especializado en la materia¹⁵, con el objetivo de fijar algunos límites

que los interesados puedan ejercer sus derechos de reclamación y su derecho a un recurso jurisdiccional efectivo».

¹² Cfr. L. NOTARO, *Intelligenza artificiale e giustizia penale*, en *Intelligenza artificiale e giustizia penale*, dirigido por A. Massaro, Paruzzo Editore, Caltanissetta, 2020, pp. 102 ss.; M. GIALUZ, *Quando la giustizia penale incontra l'intelligenza artificiale: luci e ombre dei risk assessment tools tra Stati Uniti ed Europa*, en «Diritto penale contemporaneo», 29 de mayo de 2019, pp. 10 ss.

¹³ Permítase la referencia a A. GIRALDI, *Algorithms and Big Data Towards a Crime-Preventing Groupware*, en «Roma Tre Law Review», 2/2021, pp. 26 ss.

¹⁴ El polígrafo está siendo utilizado desde 2018. Sobre «la herramienta que más dudas nos genera», v. C. ALONSO SALGADO, *Acerca de la inteligencia artificial en el ámbito penal: especial referencia a la actividad de las fuerzas y cuerpos de seguridad*, en «Ius et Scientia», 1/2021, pp. 32-34.

¹⁵ Cfr. E. GAMERO CASADO, *El enfoque europeo de inteligencia artificial*, en «Revista de Derecho Administrativo», 2021, pp. 268 ss. Más recientemente, con particular referencia al ámbito de la justicia penal, se aprobó la Resolución del Parlamento Europeo, de 6 de octubre de 2021, sobre la inteligencia artificial en el Derecho penal y su utilización por las autoridades policiales y judiciales en asuntos penales. A este respecto, v. MORILLAS FERNÁNDEZ, *Implicaciones de la inteligencia artificial*, cit., § 1. También cabe mencionar, desde una perspectiva más amplia, la Resolución del Parlamento Europeo, de 3 de mayo de 2022, sobre la inteligencia artificial en la era digital, en cuya sección «La IA y el futuro de la democracia» señala algunas criticidades de naturaleza jurídica: «estas tecnologías de IA plantean cuestiones éticas y jurídicas cruciales;

a un mundo –el tecnológico– que de por sí no conocería restricciones.

2. *Aproximaciones a la vulnerabilidad del concepto de culpabilidad*

Como se adelantaba al principio, el alcance de los sistemas de inteligencia artificial se abordará en estas consideraciones desde la perspectiva peculiar del Derecho penal y, en particular, de aquel conjunto de elementos que, según la doctrina que se acoja, se denomina culpabilidad. De ahí, resulta imprescindible efectuar algunas premisas de carácter metodológico.

Los fundamentos de la culpabilidad –que encuentra su acogida en el Derecho penal español¹⁶ y, según la mayoría de la doctrina actual¹⁷, en el

señala que determinadas tecnologías de IA permiten automatizar el tratamiento de la información a una escala sin precedentes, lo que allana el camino para la vigilancia colectiva y la injerencia de manera ilegal y *representa una amenaza para los derechos fundamentales* [énfasis añadido], en especial el derecho a la privacidad y la protección de datos»; además recuerda esta resolución que «muchos regímenes autoritarios utilizan sistemas de IA para controlar, espiar, seguir y clasificar a sus ciudadanos, restringir su libertad de circulación y ejercer vigilancia colectiva; hace hincapié en que cualquier forma de marcaje normativo ciudadano por parte de las autoridades públicas, especialmente en el ámbito de las fuerzas de seguridad del Estado, el control de las fronteras y el poder judicial, así como su uso por parte de empresas privadas o particulares, conduce a la pérdida de autonomía y privacidad y no está en consonancia con los valores europeos; recuerda que tecnologías como la cibervigilancia y el reconocimiento biométrico, que pueden utilizarse para estos fines, están sujetas al Reglamento sobre el control de las exportaciones de la Unión; muestra su profunda preocupación por los casos anteriores de empresas de la Unión que han vendido a regímenes autoritarios de países no pertenecientes a la Unión sistemas biométricos cuyo uso sería ilegal en la Unión y condena estos casos».

¹⁶ L. MORILLAS CUEVA, *Sistema de Derecho Penal. Parte general*, Dykinson, Madrid, 2018, pp. 597 ss.; M. COBO DEL ROSAL, T.S. VIVES ANTÓN, *Derecho penal. Parte general*, Tirant lo Blanch, Valencia, 1999, p. 549; F. MUÑOZ CONDE, M. GARCÍA ARÁN, *Derecho Penal. Parte General*, Tirant lo Blanch, Valencia, 2019, pp. 341 ss. Con diferentes matices, varios autores le atribuyen valor en sí, prescindiendo de su función generalpreventiva. Cfr. S. MIR PUIG, *Función de la pena y teoría del delito en el Estado social y democrático de Derecho*, Bosch, Barcelona, 1982, pp. 91 ss.; M. PÉREZ MANZANO, *Culpabilidad y prevención. Las teorías de la prevención general positiva en la fundamentación de la imputación subjetiva y de la pena*, Universidad Autónoma de Madrid, Madrid, 1990, p. 213: «El argumento conforme al cual la culpabilidad debe ser mantenida porque tiene efectos general integradores, no puede ser esgrimido como fundamental, pues entonces la culpabilidad carece de valor en sí y el criterio determinante son las mudables exigencias sociales. La necesidad de mantener la culpabilidad es consecuencia de un argumento valorativo derivado de los principios del Estado de Derecho».

¹⁷ G. FIANDACA, E. MUSCO, *Diritto penale. Parte generale*, Zanichelli, Bolonia, 2019, pp. 329 ss.; G. MARINUCCI, E. DOLCINI, G.L. GATTA, *Manuale di Diritto Penale. Parte Generale*, Giuffrè, Milán, 2018, pp. 335 ss.

italiano— han ido evolucionando y se orientan hoy hacia la superación del modelo *punitur quia peccatum*. Desde el Derecho penal contemporáneo, constitucionalmente orientado, se ha ido huyendo de las teorías retributivas, en aras de la afflictividad del proceso sancionador y de su teleología reeducativa y resocializadora¹⁸. Con base en estas evoluciones del concepto de culpabilidad, y de su trascendencia en el sistema jurídico-penal vigente, es evidente que la introducción de sujetos inanimados en el entorno social tiene varias repercusiones en los elementos que lo conforman. Dicho de otro modo, desde la perspectiva psicológica o desde la más avanzada concepción normativa, en un mundo dominado (o transformado) por la actuación de sistemas de inteligencia artificial, el Derecho penal puede ver amenazados sus principios rectores.

Con el fin de analizar cuáles pueden ser las implicaciones de un cambio radical que supone la llegada de objetos inanimados en la esfera de los sujetos de derecho, se intentará esbozar un análisis de la culpabilidad penal distinguiendo entre sus dos vertientes. Por un lado, se examinará la influencia de los sistemas de inteligencia artificial en la culpabilidad en su condición de presupuesto de la punibilidad¹⁹, y por otro, se analizará la cara más empírica de la moneda, es decir, la culpabilidad como criterio de determinación de la sanción penal.

Antes de ceñir el enfoque de estos temas, parece oportuno acercarse al concepto de inteligencia artificial y de los correspondientes sistemas. En efecto, pese a que la expresión se halle incorporada al lenguaje común, se hace necesario aclarar su significado con vistas a estas reflexiones. El concepto de inteligencia artificial ha sido definido como el potencial biopsicológico de procesar informaciones para solventar problemas o crear productos que tienen valor en una determinada cultura social²⁰ y, asimismo, como el proceso de llevar una máquina a comportarse de una manera que

¹⁸ Cfr. los artículos 25.2 de la Constitución española («Las penas privativas de libertad y las medidas de seguridad estarán orientadas hacia la reeducación y reinserción social y no podrán consistir en trabajos forzados») y 27, apartado tercero, de la Constitución italiana («Las penas no pueden consistir en tratamientos contrarios al sentido de humanidad y deben tender a la reeducación del condenado»). Se llega a dudar que pueda defenderse la retribución como «desvinculada del fin». Así, W. HASSEMER, *¿Por qué castigar? Razones por las que merece la pena la pena*, Tirant lo Blanch, Valencia, 2016, p. 62, traducción de M. Cancio Meliá y F. Muñoz Conde.

¹⁹ Es decir, la culpabilidad como elemento constitutivo del delito, pese al escenario dogmático que se acoja con respecto a la estructura del delito.

²⁰ H. GARDNER, *Intelligence Reframed. Multiple Intelligences for the 21st Century*, Basic Books, New York, 2000, p. 33: «biopsychological potential to process information [...] to solve problems or create products that are of value in a culture».

se consideraría inteligente si un humano actuara así²¹.

No debe extrañar, pues, que al tratarse de un mundo en continua evolución, ya hayan aparecido distintas etapas en el desarrollo de dichos sistemas. En cambio, lo que sí debería impresionar –y quizás no lo hace por el justificable recelo hacia el Derecho positivo– es el consuetudinario aletargamiento del ordenamiento jurídico, que mientras sigue ocupándose de –y asolando la ley ante– presuntos incrementos de ‘alarma social’, aparta la adecuación de la legislación vigente a los retos tecnológicos que avanzan, y que ya han dejado atrás las primeras etapas de su desarrollo. Evidentemente, la primera generación de los sistemas de inteligencia artificial (*artificial narrow intelligence*, que permitió desde la creación de los asistentes de voz hasta el diseño de los primeros coches autónomos) está ahora al alcance de cualquiera y ha sido ya superada por la *artificial general intelligence*, la segunda generación capaz de razonar, plantear y solventar problemas de forma independiente e incluso inesperada (*machine learning*) por sus diseñadores. Está previsto para el futuro próximo el desarrollo de una tercera generación (*artificial super intelligence*), que algunos llaman la «verdadera inteligencia artificial», formada por sistemas conscientes y capacitados para relegar a un segundo plano al ser humano²², lo que asustaría, desde el punto de vista jurídico, al poder imaginar hasta la posibilidad de que alcance al mismo procedimiento de creación y aprobación de las leyes.

En cualquier caso, el concepto de sistema de inteligencia artificial que se utilizará a lo largo de este escrito incluye todo sistema inanimado, creado por el ser humano y capaz de alimentarse por sí solo, susceptible de causar autónomamente modificaciones del mundo exterior, relevantes para la sociedad y, por consiguiente, para el Derecho. En dicha definición, lo más amplia posible, caben máquinas, robots, algoritmos, y cualquier otro tipo de sujeto inhumano, hasta el robot aspirador de todos los días.

Regresando a las consideraciones sobre la culpabilidad penal, conviene distinguir el alcance de los sistemas de inteligencia artificial según impacten

²¹ La definición «making a machine behave in ways that would be called intelligent if a human were so behaving» se remonta a la Conferencia de Dartmouth de 1955, en la que probablemente tuvo lugar la primera aparición de la locución ‘inteligencia artificial’. Cfr. J. MCCARTHY, M.L. MINSKY, N. ROCHESTER, C.E. SHANNON, *A proposal for the Dartmouth summer research project on artificial intelligence*, 31 de agosto de 1955, disponible en <<http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>> (consultado el 18 de noviembre de 2022).

²² La delimitación de las tres etapas de desarrollo de la inteligencia artificial se encuentra en A. KAPLAN, M. HAENLEIN, *Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence*, en «Business Horizons», 1/2019, p. 16.

en el elemento constitutivo del delito o en el criterio para la determinación de la pena en concreto.

Con respecto al primero de los casos, es decir, a la culpabilidad como presupuesto de la punibilidad, y entonces, como elemento del delito, el análisis estará encaminado a averiguar las implicaciones que conllevan los nuevos sujetos de derecho, dependiendo de que sean autores o víctimas de la comisión de un hecho abstractamente delictivo, o bien instrumentos de apoyo al ejercicio de la justicia. Es posible adelantar desde ahora que, sin perjuicio de las investigaciones desarrolladas en sentido contrario, todavía resulta difícil imaginar una concepción de culpabilidad que tenga carácter unitario, es decir, que permita armonizar en una categoría autónoma el conjunto de los distintos elementos que la concepción normativa moderna deposita en la propia culpabilidad²³. Evidentemente, a la luz de la identidad del efecto que producen las causas tipificadas, lejos de insistir en categorías *praeter legem* como la exigibilidad de una conducta distinta²⁴, parecería correcto descartar la idea de que, en el marco de la misma norma penal, existan causas que eximen de la responsabilidad penal sin llegar a justificar el hecho²⁵.

En segundo lugar, con respecto a la culpabilidad concebida como criterio para la determinación de la pena, es necesario advertir que el análisis tomará en consideración la medición de la pena en concreto, es decir, la que debe imponer el órgano judicial posteriormente a la averiguación de la existencia del hecho delictivo. Con ello, sólo se intenta restringir el examen

²³ M. GALLO, *Il concetto unitario di colpevolezza*, Giuffrè, Milán, 1951, pp. 95-96: «intesa come fatto giuridico, e precisamente come criterio di imputazione, la colpevolezza sfugge ad ogni concettualizzazione che non sia quella meramente estrinseca della specifica funzionalità assoluta dagli elementi dai quali si afferma costituirsi». Para el autor, el contenido de la culpabilidad coincide con el concepto de antijuridicidad, o más exactamente, se enmarca en el seno de esta, constituyendo una forma especialmente cualificada de antijuridicidad. Cfr. *ivi*, p. 144.

²⁴ Sucintamente, «la referencia a la inexigibilidad de una conducta conforme a la norma como fundamento de las denominadas causas de inculpabilidad o exculpación, no casa bien con el hecho de que al declarar el legislador una conducta como antijurídica, sí parece exigirla». Así, T. AGUADO CORREA, *Principio de inexigibilidad de otra conducta en las categorías del delito*, en «Revista Digital de la Maestría en Ciencias Penales de la Universidad de Costa Rica», 3/2011, p. 24.

²⁵ GALLO, *La teoria dell'azione «finalistica» nella più recente dottrina tedesca*, Giuffrè, Milán, 1950, p. 60. Cabe recordar que, en la perspectiva bipartita que apoya el autor, el hecho puede definirse típico al constar de la conducta, el nexo causal y el resultado, además de la ausencia de causas de justificación, que se encajan en el tipo como elementos negativos construidos positivamente.

al efecto último²⁶ del principio de culpabilidad, ya que, hasta la fecha, la determinación de la pena en abstracto le corresponde íntegramente al legislador, que actúa (o debería actuar) atendiendo a criterios jurídicos y a juicios de valor de carácter político-criminal. Se ha aludido antes a la posibilidad de que los sistemas de inteligencia artificial lleguen incluso a intervenir en el *iter legis*. Sin embargo, parece que de momento el control de dicho proceso queda exclusivamente en manos del legislador y no ha habido propuestas de incorporación de mecanismos artificiales en una actividad tan crucial, cuya justificación sigue rigiéndose por la población en aras de la insoslayable reserva de ley penal.

En este sentido, pese a los cambios de los marcos edictales realizados por los distintos mandatos parlamentarios, lo importante es que se respeten los principios rectores en atención al peligro o al daño causado al bien jurídico protegido por la norma penal. Ahora bien, en el ámbito de la determinación de la pena en concreto, el alcance de la inteligencia artificial se estudiará según se indica: por un lado, la inteligencia artificial pondría en peligro la operatividad de la proporcionalidad de la pena, al prever mecanismos que impedirían la explotación de todo segmento del marco edictal; por otro lado, puede darse el caso de la máquina que, al aplicar la pena en lugar del juez persona física, genera una condena injusta o ilícita, cuyo pronunciamiento puede adquirir a su vez relevancia penal.

3. La incidencia de los sistemas de inteligencia artificial en la culpabilidad como presupuesto de la punibilidad

Con respecto al conjunto de elementos que conformaría la categoría autónoma de la culpabilidad, ya se ha adelantado la existencia de discrepancias entre la doctrina contemporánea. En principio, varios han sido los esfuerzos para dotarla de independencia dogmática; tanto es así

²⁶ En realidad, la determinación judicial de la pena no sería el último de los efectos jurídicamente relevantes de la culpabilidad, al jugar sus elementos (especialmente los que están vinculados con las circunstancias personales del reo) otro papel en la fase de post-determinación de la pena. En cierta medida, es fundamental que el reo continúe a percibir justo, equitativo, y entonces proporcionado, el tratamiento sancionador que se le otorgó en su momento. Al respecto, véase la literatura citada en GIRALDI, *Política, discrecionalidad y Derecho en las implicaciones empíricas del principio de proporcionalidad de la pena*, en «Revista Internacional de Doctrina y Jurisprudencia», 22/2020, pp. 23 ss.

que la doctrina predominante, en una perspectiva pluripartita²⁷, incluye en la culpabilidad la imputabilidad, la conciencia de la antijuridicidad de la conducta realizada y la exigibilidad de una actuación distinta por parte del sujeto activo²⁸. Sin embargo, persiste la disyuntiva sobre la inclusión de las formas de culpabilidad (dolo, imprudencia y, según algunos autores, dolo mixto a imprudencia²⁹) en su propio ámbito³⁰ o en el tipo subjetivo

²⁷ La autonomía de la categoría dogmática se afirma con distintos matices desde ambas las teorías tripartitas y cuatripartitas del delito, actualmente mayoritarias.

²⁸ MUÑOZ CONDE, GARCÍA ARÁN, *Derecho Penal*, cit., pp. 341-342; COBO DEL ROSAL, VIVES ANTÓN, *Derecho penal*, cit., p. 567, incluye en la estructura del juicio de culpabilidad la imputabilidad, las formas de culpabilidad (dolo y culpa) y la exigibilidad; MIR PUIG, *Derecho penal. Parte general*, Editorial Reppertor, Barcelona, 2011, p. 139, compartiendo las dudas sobre la ambigüedad del término y las connotaciones moralizantes de la culpabilidad, más bien se refiere a ella incluyendo en la esfera de la imputación personal «tanto la indicada exigencia de infracción personal de la norma primaria, como la necesidad de un sujeto penalmente responsable».

²⁹ La teoría del dolo mixto a culpa fundamentaría la regulación expresa, en el ordenamiento italiano, de la preterintencionalidad y del homicidio preterintencional (arts. 43 y 584 del Código Penal italiano), sobre la base de que las conductas lesivas serían dolosas y la causación del resultado mortal se imputaría a título de imprudencia. *Contra*, la tesis de la unicidad del elemento psicológico de la preterintencionalidad, actualmente dominante en la jurisprudencia italiana, según la cual el elemento subjetivo del delito subsidiario (lesiones) absorbería la previsibilidad del resultado causado (muerte). Cfr., *amplius*, F. BASILE, *L'alternativa tra responsabilità oggettiva e colpa in attività illecita per l'imputazione della conseguenza ulteriore non voluta, alla luce della sentenza Ronci delle Sezioni Unite sull'art. 586 c.p.*, en «Rivista italiana di diritto e procedura penale», 3/2011, pp. 911 ss. En España, tras la derogación de la atenuante de preterintencionalidad y de los delitos cualificados por el resultado, con el fin de erradicar los supuestos de *versari in re illicita*, la mayoría de la jurisprudencia está de acuerdo en aplicar el concurso ideal del art. 77 CP entre las lesiones dolosas y el homicidio imprudente (cfr. recientemente STS 2833/2022, de 7 de julio). Dicha solución concursal ya había sido propuesta por MIR PUIG, *Preterintencionalidad y límites del art. 50 del Código Penal*, en «Revista Jurídica de Cataluña», 1979, p. 99. V. también G. QUINTERO OLIVARES, *Las vicisitudes del dolo y la subsistencia de la preterintencionalidad*, en *Constitución, derechos fundamentales y sistema penal. Semblanzas y estudios con motivo del setenta aniversario del profesor Tomás Salvador Vives Antón*, dirigido por J.C. Carbonell Mateu, J.L. González Cussac, E. Orts Berenguer, Tirant lo Blanch, Valencia, 2009, p. 1573 ss.

³⁰ COBO DEL ROSAL, VIVES ANTÓN, *Derecho penal*, cit., pp. 616-617. Aunque un determinado momento del ánimo pueda considerarse para determinar si concurre el injusto típico y, a la vez, para decidir sobre la reprochabilidad de dicho injusto, «en ningún caso, el dolo o la imprudencia, como tales, puedan pertenecer al injusto típico. [...] El dolo, aunque presuponga ciertas realidades psicológicas, no se identifica, desde luego, con ellas. Y no puede pertenecer al tipo de injusto desde el momento en que representa un modo o forma en que la realización del tipo puede reprocharse a su autor».

(*subjektive Tatbestand*)³¹, o bien en ambos con distintas funciones³².

En cualquier caso, prescindiendo de la validez de las observaciones formuladas acerca de la imposibilidad de construir una concepción unitaria de culpabilidad³³, al entrar en la esfera de las relaciones jurídicas un conjunto de sujetos inanimados, vuelven a adquirir pleno auge las polémicas sobre la cercanía entre culpabilidad y peligrosidad³⁴, y quizás resultará oportuno cuestionar la validez de la estructura actual del delito.

El análisis de la influencia de los sistemas de inteligencia artificial en la culpabilidad, entendida como elemento constitutivo del delito y, por ende, presupuesto de la punibilidad, se desarrollará con arreglo a las distintas formas en que la existencia de estos sujetos puede interactuar con las conductas delictivas.

En primer lugar, se plantearán las cuestiones relativas a la eventual atribución de subjetividad *pasiva* a los sistemas ‘inanimados’, que alcanzarían la condición de potenciales víctimas de un delito, con todos los problemas de índole procesal que tal reconocimiento pueda entrañar. En este sentido, también hay que pensar en la dificultad de determinar quién –y sobre todo cómo– asumiría su representación y defensa en el juicio, teniendo presente

³¹ MUÑOZ CONDE, GARCÍA ARÁN, *Derecho Penal*, cit., p. 249. «Volver a incluir estos elementos en la culpabilidad me parece una complicación innecesaria y una falta antisistemática. Y lo mismo creo, si se considera que el dolo pertenece a la culpabilidad –como hace la teoría tradicional–. Lo que no se puede hacer es querer incluir un mismo hecho en dos categorías sistemáticas distintas, pues entonces ¿para qué sirven las clasificaciones y distinciones sistemáticas?». Así, MUÑOZ CONDE, *Introducción*, en C. ROXIN, *Política criminal y sistema del derecho penal*, Hammurabi, Buenos Aires, 2002, p. 28, traducción de F. Muñoz Conde.

³² Las disyuntivas sobre la ubicación del elemento psicológico del delito ponen de manifiesto la «riqueza de pensamiento jurídico[que] crece en la medida en que aumentan las soluciones que se ofrecen». Así, QUINTERO OLIVARES, *Parte General del Derecho Penal. Adaptada al programa de ingreso en las carreras judicial y fiscal*, Aranzadi, Pamplona, 2015, p. 161. Véanse, al respecto, las posiciones que apoyan la ubicuidad del dolo en MORILLAS CUEVA, *Sistema de Derecho Penal*, cit., pp. 671 ss.: «como factor directivo del comportamiento es el núcleo central del injusto típico de la acción; como resultado final del proceso de formación de la voluntad que pertenece a la culpabilidad es también un integrante de esta». En la misma línea, ya se expresó ROXIN, *Derecho Penal*, cit., p. 228, al afirmar que «la tan discutida cuestión de si el dolo “pertenece” al injusto o a la culpabilidad [...] es para el sistema político criminal una alternativa planteada erróneamente, ya que el dolo puede ser significativo desde ambos puntos de vista en contextos jurídicos respectivamente diversos».

³³ Cfr. *supra*, *sub* nota 23.

³⁴ Piénsese en que el dolo y la imprudencia, en el marco de las teorías que los incorporan como elementos de la culpabilidad, fundamentarían tanto el reproche penal como el juicio de peligrosidad criminal.

que la norma penal «real» consta de segmentos tanto sustantivos como procesales, ambos irrenunciables³⁵.

En segundo lugar, a la inversa, se analizarán las cuestiones más controvertidas que lleva consigo la atribución de subjetividad *activa* a los sistemas de inteligencia artificial. En esta perspectiva, merece la pena cuestionar la oportunidad de reprochar a una máquina, un robot o incluso un algoritmo. De considerarlo oportuno, habrá que analizar las posibles medidas de adaptación del actual sistema penal a efectos del reconocimiento de autoría directa por parte de los sujetos de inteligencia artificial.

Por último, la utilización de los sistemas de inteligencia artificial como *instrumento* de apoyo en el ámbito de la justicia exige el desarrollo de algunas reflexiones sobre la posible afectación del criterio de libre valoración de la prueba y, de cara al futuro, sobre el posible fomento de una revolución (patológica) en la teoría del delito.

3.1. *El tentativo de reconocer una nueva subjetividad pasiva: la inteligencia artificial como víctima de delitos*

El reconocimiento de subjetividad pasiva a los sistemas de inteligencia artificial, entendidos con los matices susodichos, permite el planteamiento de dos distintas cuestiones que merecen ser analizadas por separado. Por un lado, hay que subrayar las dificultades ligadas a la vigencia de un Derecho penal necesariamente lesivo, en el que – pese al marco teórico elaborado por la doctrina del delito como *personales Unrecht*³⁶ – las exigencias de protección de un determinado bien jurídico han de primar sobre el desvalor personal de la acción. Por otro lado, la convergencia entre los conceptos de culpabilidad y peligrosidad hacen resurgir viejas preocupaciones a la hora de examinar los casos en que los sistemas de inteligencia artificial pudieran ser víctimas.

En primer lugar, debe advertirse que la doctrina penalista se ha interrogado extensamente sobre la oportunidad de introducir nuevos tipos

³⁵ GALLO, *Diritto penale italiano. Appunti di parte generale*, vol. I, Giappichelli, Torino, 2021, p. 15: «la norma penale reale, la proposizione cioè che esaurisce il significato che ad una determinata condotta umana è attribuito in quel settore dell'ordinamento definito penale e criminale, consta di segmenti che delineano la fattispecie condizionante e le conseguenze condizionate e di segmenti che dispongono come deve essere accertata detta fattispecie e come vanno pronunciate (irrogate) ed eseguite le conseguenze sanzionatorie».

³⁶ Sobre la dificultad de concebir un injusto no culpable desde la doctrina alemana del injusto personal, v. E.J. LAMPE, *Das personale Unrecht*, Duncker & Humblot, Berlín, 1967, p. 111; H. OTTO, *Personales Unrecht, Schuld und Strafe*, en «Zeitschrift für die gesamte Strafrechtswissenschaft», 3/1975, pp. 539 ss.

delictivos, o bien de modificar los ya existentes, de manera que adquieran relevancia penal los ataques dirigidos a los sistemas de inteligencia artificial³⁷. La cuestión se enmarca, desde una perspectiva más amplia, en la posibilidad de reconocer la titularidad de derechos y deberes a sujetos que, en principio, no son humanos.

A este respecto, ha evolucionado mucho la manera de entender la personalidad jurídica en relación con la inteligencia artificial. En un primer momento, parecía hasta complejo construir máquinas capaces de lo que hoy, en cambio, pudiera ser logro del futuro próximo, si no lo es ya del presente³⁸. El mismo concepto de *machina sapiens*³⁹, anteriormente sólo concebible en el ámbito de la ciencia ficción, parece ahora el objetivo tendencial de la sociedad, que exige su consideración por el ordenamiento jurídico.

Para que las entidades puedan definirse inteligentes, se ha señalado que deberían poseer al menos los siguientes atributos: en primer lugar, la capacidad de comunicar (*communication*), es decir, la posibilidad de que la máquina entienda de algún modo las interacciones que se le dirijan; en segundo lugar, el conocimiento de sí (*internal knowledge*), es decir, la capacidad de comprenderse a sí mismo; después, el conocimiento del mundo exterior (*world knowledge*), que implica la conciencia de los acontecimientos e informaciones externas a sí; la intencionalidad (*intentionality* o *goal-driven behavior*), es decir, la habilidad de realizar una o más acciones para alcanzar un determinado objetivo; por último, la creatividad (*creativity*), o sea la capacidad de buscar una solución alternativa

³⁷ BASILE, *Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*, en «Diritto Penale e Uomo», 10/2019, pp. 32-33.

³⁸ Sobre el reconocimiento de la personalidad jurídica a máquinas y robots, y acerca de la posibilidad de crear una inteligencia artificial en un contexto temporal ya superado, v. L.B. SOLUM, *Legal Personhood for Artificial Intelligences*, en «North Carolina Law Review», 4/1992, pp. 1231 ss. Para formarse una idea sobre la vivacidad del debate ya en ese entonces, v. *ivi*, p. 1237, *sub nota* 24.

³⁹ Existen notables discrepancias sobre este concepto, a la luz de la perenne complejidad de proporcionar una definición válida incluso para la inteligencia, por así decirlo, natural. En este sentido, cfr. el pensamiento de Aristóteles en las palabras de H. VELÁZQUEZ FERNÁNDEZ, *¿Qué tan natural es la inteligencia artificial? Sobre los límites y alcances de la biomimesis computacional*, en «Naturaleza y Libertad. Revista de estudios interdisciplinarios», 12/2019, p. 238: «cuando intentamos imitar lo natural mediante lo artificial, es importante diferenciar las artes que reproducen la naturaleza, de otras que solo intentan perfeccionarla, completarla o modificarla para bien». Desde una perspectiva jurídico-penal, v. G. HALLEVY, *The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control*, en «Akron Intellectual Property Journal», 2/2010, pp. 175 ss.

en el caso de que la acción (intencional) no haya alcanzado su finalidad⁴⁰.

Ahora bien, no es esta la sede para evaluar las evidencias científicas que se han aportado a favor y en contra de la humanización de los sistemas inanimados. Pues, desde un punto de vista ontológico, la afirmación de que los artefactos no tienen –y quizás nunca podrán tener– sentimientos y no sienten emociones reales, es quizás la respuesta de una pregunta erróneamente planteada. En realidad, en un sentido similar al acertadamente brindado por KAPLAN, la pregunta básica de si los sistemas inanimados son humanizables no debería dirigirse a un público indefinido, sino que exigiría una respuesta personalizada sobre la base de los sujetos o las cosas a que cada uno decide extender la «cortesía de su empatía»⁴¹.

Sin embargo, las reflexiones ontológicas mencionadas exceden del objetivo que se propone el presente estudio, que a lo mejor debería ceñirse a la deontología de dicho fenómeno humanizador. En efecto, desde el punto de vista del Derecho penal, no se trata de establecer si se le concede la condición de ‘vida humana’ a un sistema de inteligencia artificial; más bien se trataría de averiguar si la protección de dicha ‘vida’ en el marco de los bienes jurídicos penalmente relevantes está permitida por la *extrema ratio* que trasluce del carácter necesariamente lesivo de la Ley penal.

Aun así, la cuestión formulada escaparía del alcance del concepto de culpabilidad, tanto en su vertiente psicológica como en la concepción normativa más moderna. La incorporación de un determinado bien jurídico entre los que merecen tutela penal va más allá de las problemáticas inherentes al principio de culpabilidad y, junto con todas las implicaciones de carácter político-criminal, se sitúa idealmente en la teoría del bien jurídico y de la legitimación del *ius puniendi* estatal⁴².

Podría objetarse que, para dotar al hecho delictivo del carácter de

⁴⁰ La enucleación de los atributos de las entidades inteligentes se desarrolló por R.C. SCHANK, *What Is AI, Anyway?*, en «AI Magazine», 4/1987, p. 60.

⁴¹ En sentido similar se expresa, con un sugestivo ejemplo, J. KAPLAN, *Artificial Intelligence. What Everyone Needs to Know*, Oxford University Press, Oxford, 2016, pp. 85 ss. «On the one hand, our intuitions lead us to believe that machines, no matter how sophisticated, raise no ethical concerns in their own right. On the other, we can't comfortably exclude certain entities from the community of living things based solely on what materials they are composed of. My personal opinion, not universally shared, is that what's at issue here is little more than a decision we get to make as to whom, or to what, we choose to extend the courtesy of our empathy» (*ivi*, p. 86).

⁴² Sobre estas teorías, v. R. HEFENDEHL, A.V. HIRSCH, W. WOHLERS (eds.), *La teoría del bien jurídico. ¿Fundamento de legitimación del Derecho penal o juego de abalorios dogmático?*, Marcial Pons, Madrid, 2016, *passim*, traducción de R.A. Guirao, M. Martín Lorenzo, Í. Ortiz de Urbina Gimeno.

‘culpable’, la lesión causada al bien jurídico sea subsidiaria frente al desvalor personal de la acción. Esta objeción fue impulsada, en su momento, por una parte de la doctrina alemana y obtuvo particular notoriedad en Italia, debido a la formulación del art. 27, apartado primero, de la Carta constitucional («la responsabilidad penal es *personalis*»). Sin embargo, es suficiente recordar la doctrina ampliamente consolidada, conforme a la cual dicha ofensa no representa una mera *quantité négligeable*, sino más bien un requisito necesario para que un hecho sea punible con arreglo a las Leyes penales⁴³. Desde luego se trata de una afirmación que se debe al posicionamiento, por parte de la Constitución, sobre la teleología de la pena y de las medidas de seguridad⁴⁴.

Ahora bien, por lo que respecta al presente estudio, se puede afirmar con suficiente seguridad que el Derecho penal, en última instancia, está encaminado a promover y proteger los intereses y los valores típicos de los seres humanos. En este sentido, se ha afirmado que el reconocimiento de derechos y obligaciones de entidades inhumanas siempre se lleva a cabo desde una perspectiva antropocéntrica, es decir, con arreglo al objetivo último de la protección de la persona humana⁴⁵. No obstante, mientras es

⁴³ F. BRICOLA, voz *Teoria generale del reato*, en «Novissimo digesto italiano», XIV, 1973, p. 63: «La concezione del reato come “*personales Unrecht*”, nella misura in cui riduce l’offesa al bene giuridico al rango di mera *quantité négligeable*, non si pone in esplicita antitesi con la Costituzione tedesca la quale, a differenza di quella italiana, non prende posizione sul teleologismo delle pene e delle misure di sicurezza e, quindi, non cristallizza, sulla base di quanto si dirà in seguito, la natura *necessariamente* offensiva dell’illecito penale».

⁴⁴ *Ivi*, pp. 62 ss. Cfr. también *supra*, *sub* nota 18. En España, la teleología de las medidas de seguridad se desprende expresamente del art. 27.3 de la Constitución. En Italia, en cambio, al no constar una expresa previsión de esa finalidad en el caso de las medidas de seguridad, sino solo para las penas, se ha debatido largamente sobre su naturaleza, inicialmente considerada de carácter administrativo (véase la propia rúbrica del Código penal italiano, «De las medidas administrativas de seguridad»), y sobre su finalidad, originalmente correspondiente a la mera prevención especial. Sin embargo, hoy la unanimidad de la doctrina está de acuerdo en apreciar su naturaleza de sanciones penales y en dotarlas de los rasgos finales reeducativos, al igual que las penas. Cfr., al respecto, A. CAVALIERE, *Misure di sicurezza e doppio binario. Considerazioni introduttive*, en «Rivista italiana di diritto e procedura penale», 1/2022, pp. 343 ss.; L. FURNARI, *Misure di sicurezza e doppio binario: un declino inarrestabile?*, en «Rivista italiana di diritto e procedura penale», 1993, 569 ss.; E. MUSCO, voz *Misure di sicurezza*, en «Enciclopedia del diritto», I, 1997, § 22; P. NUVOLONE, voz *Misure di prevenzione e misure di sicurezza*, en «Enciclopedia del diritto», XXVI, 1976, § 17.

⁴⁵ Con las palabras de S. RIONDATO, *Robotica e diritto penale (robots, ibridi, chimere e “animali tecnologici”)*, en *Genetics, robotics, law, punishment*, dirigido por D. Provolò, S. Riondato, F. Yenisey, Padova University Press, Padua, 2014, p. 603: «tali estensioni della nozione di “persona” e, più specificatamente, la qualificazione delle entità non-umane

cierto que tal reconocimiento se lleva a cabo desde una visión exquisitamente ligada a la percepción de dichas entidades por parte del ser humano en cada época, plantea problemas la tipificación de algunas conductas como, por ejemplo, los delitos contra los animales domésticos. Desde la perspectiva de la protección del ser humano, la tutela de la flora y de la fauna se habría incorporado al Código Penal de manera instrumental, de cara a proteger el equilibrio medioambiental a modo de delito obstáculo, con el objetivo último de tutelar al ser humano. Sin embargo, la introducción de los delitos contra los animales domésticos plantea alguna duda acerca de esa construcción, si bien podría considerarse que, de todos modos, la tutela a través del Derecho penal se ha otorgado de cara a proteger los sentimientos de las personas hacia aquellos.

Asimismo, se ha afirmado con acierto que no hay razones válidas para obstaculizar el reconocimiento de la subjetividad pasiva a los sistemas de inteligencia artificial, como sujetos dotados de derechos y obligaciones, siempre y cuando dicha atribución sea meramente ficticia y se justifique con la persecución de finalidades ‘humanas’, o al menos controladas por los seres humanos⁴⁶.

En este sentido, antes de acceder al otorgamiento de subjetividad pasiva a los sistemas de inteligencia artificial, merece la pena mencionar dos ejemplos sugestivos⁴⁷ que quizás proporcionen una idea sobre el alcance de los comportamientos de las máquinas en el marco de conductas que, si se dirigiesen hacia seres humanos, constituirían lesión de un bien jurídico penalmente protegido.

Es el caso, primero, de los programas de *doll therapy* o *pet therapy*,

come “soggetti” del reato, siano solitamente finalizzate a promuovere o proteggere gli interessi e i valori tipici dell’essere umano. Infatti, l’attribuzione di diritti, interessi ed obblighi a entità non-umane – siano esse animali o società – avviene sempre in base ai valori e al senso di giustizia del legislatore umano».

⁴⁶ *Ibidem*. Resulta oportuno destacar la discrepancia con la posición del autor, cuando sostiene que también los robots capaces de actuar de forma independiente e imprevisible están programados por los seres humanos, y por ello los propios seres humanos pueden impedir cualquier consecuencia. Hoy en día, la evolución de los sistemas de *machine learning* permite afirmar lo contrario: el aprendizaje autónomo, característico de los sistemas de inteligencia artificial, impide un control humano efectivo sobre las consecuencias de sus actuaciones independientes, constituyendo quizás el problema más nítido de la implementación de estas máquinas en la sociedad. Cfr., al respecto, J. PÉREZ ARIAS, *Algoritmos y big data en la responsabilidad penal: el reto de la cibercriminalidad en el Derecho Penal*, en esta obra, § 5; P.C. ARIAS, *Artificial Intelligence & Machine Learning: a model for a new judicial system?*, en «Revista Internacional Jurídica y Empresarial», 2020, pp. 81 ss.

⁴⁷ Se utilizan los ejemplos brindados por BASILE, *Intelligenza artificiale*, *loc. ult. cit.*

terapias dirigidas a sujetos enfermos (mentales, sobre todo) para que puedan desarrollar sentimientos y emociones, en las cuales las personas implicadas (no necesariamente inimputables) pueden llegar a causar la destrucción de estos robots, su maltrato y eventual lesión; y, en segundo lugar, del fenómeno denominado ‘violación robótica’⁴⁸, con particular referencia a las agresiones sexuales hacia robots que tienen rasgos parecidos a los de sujetos menores.

Los actos de ‘violencia’ perpetrados hacia las máquinas susodichas no constituyen delito con arreglo a la legislación vigente, a no ser que se enmarquen dentro de los tipos delictivos referidos exclusivamente a la protección de bienes materiales (*e.g.*, daños). Ahora bien, recuperando la sugestiva imagen retratada por BASILE⁴⁹, si la agresión sexual ‘robótica’ se llevara a cabo hacia un androide que está desarrollando la labor encomendada, es decir, un robot diseñado para ser recepcionista de un hotel o bien empleado en una tienda de ropa, ¿el hecho de que la conducta quede impune y su autor libre podría despertar un sentimiento de (legítima) inquietud por parte de la colectividad? En otras palabras, reformulando la pregunta con las lentes del jurista, *de iure condendo* ¿dicha conducta podría fundamentar un reproche penal o, al menos, un juicio de peligrosidad del autor?

Resulta claro que, finalmente, la alternativa conlleva distintas consecuencias jurídicas. Si se accede a fundamentar un reproche penal por las conductas llevadas a cabo en contra de un sistema inanimado, habría que reformar el Código penal con la introducción de la subjetividad pasiva en cuestión, que conllevaría la aplicación de una determinada pena. En cambio, si se decide fundamentar un juicio de peligrosidad del autor, las consecuencias en términos de medidas de seguridad o de prevención son distintas según la legislación vigente en cada ordenamiento, habida cuenta de la considerable diferencia que existe entre los diversos sistemas jurídicos positivos⁵⁰.

Con respecto a esta segunda opción, en España el principio de legalidad establecido en el art. 25 de la Constitución se ha interpretado de manera estricta, así que el Código penal de 1995 impide la aplicación de medidas

⁴⁸ Sobre este fenómeno, v. C. TORRAS GENÍS, L. DELICADO MORATALLA, *Blanco y Frio como El Polo Norte y El Polo Sur: Seres humanos y robots*, en «Sociología y tecnología: Revista digital de sociología del sistema tecnocientífico», 1/2022, pp. 263 ss. J. DANAHER, *Robotic Rape and Robotic Child Sexual Abuse: Should They be Criminalised?*, en «Criminal Law and Philosophy», 1/2017, pp. 71 ss.; M.H. MARAS, L.R. SHAPIRO, *Child Sex Dolls and Robots: More Than Just an Uncanny Valley*, en «Journal of Internet Law», 2017, pp. 3 ss.

⁴⁹ BASILE, *Intelligenza artificiale*, cit., p. 33.

⁵⁰ COBO DEL ROSAL, VIVES ANTÓN, *Derecho penal*, cit., p. 985 ss.

de seguridad a sujetos socialmente peligrosos, y restringe su imposición a quienes se enmarquen en el concepto de peligrosidad criminal, es decir, a aquellos individuos que, tras haber cometido algún delito, ostenten el riesgo de volver a violar la Ley penal. Pues, en este contexto, sólo cabría la posibilidad de fundamentar un juicio de peligrosidad si el autor de la conducta en contra del sistema inanimado haya realizado un hecho delictivo. Ahora bien, se podría sostener este fundamento del juicio de peligrosidad sobre la base de que el autor, en la mayoría de los casos, al menos cometería un delito de daños⁵¹, y la medida de seguridad se haría necesaria de cara a prevenir la reiteración futura del delito.

En Italia, en cambio, la problemática relativa a la viabilidad de un juicio de peligrosidad abre las puertas a debates todavía pendientes⁵². En el Código penal italiano, que acoge el sistema dualista, el concepto de peligrosidad está ligado a la peligrosidad social, que surte sus efectos tanto (i) después de la comisión de un delito, como (ii) antes de –y sin– que este se cometa, incluyendo en esta categoría la peligrosidad derivada de la realización de un cuasi-delito⁵³. De apreciarse la peligrosidad del sujeto, en el primer caso se podría aplicar una medida de seguridad, y en el segundo tanto una medida de seguridad⁵⁴ como una medida de prevención⁵⁵.

⁵¹ Por los problemas de causalidad que normalmente plantean, no se mencionan demás delitos que puedan apreciarse de forma colateral a la conducta realizada, como por ejemplo las lesiones psíquicas ocasionadas en los compañeros humanos del robot recepcionista violado.

⁵² Cfr., *amplius*, E. SQUILLACI, 'Ragioni' e 'costi' dell'attuale paradigma di prevenzione, en «Archivio penale web», 2/2020, pp. 1 ss.; A. MENCARELLI, *Le misure di sicurezza detentive al vaglio della Corte europea dei diritti dell'uomo*, en *La tutela della salute nei luoghi di detenzione. Un'indagine di diritto penale intorno a carcere, REMS e CPR*, dirigido por A. Massaro, RomaTrE-Press, Roma, 2017, pp. 333 ss.; A. MASSARO, *Pericolosità sociale e misure di sicurezza detentive nel processo di "definitivo superamento" degli ospedali psichiatrici giudiziari: la lettura della Corte costituzionale con la sentenza n. 186 del 2015*, en «Archivio penale», 2/2015, pp. 1 ss.; M. BERTOLINO, *Declinazioni attuali della pericolosità sociale: pene e misure di sicurezza a confronto*, en «Archivio penale», 2/2014, pp. 459 ss.

⁵³ Se trata de los casos previstos en los arts. 49, último apartado, y 115, último apartado, del Código penal italiano. Respectivamente, por un lado, se puede aplicar una medida de seguridad al sujeto que haya cometido un hecho creyendo erróneamente que este constituya un delito (mientras, en realidad, no lo constituye), y por otro, al sujeto que instigue a otro a cometer un delito, cuando la instigación no se haya acogido.

⁵⁴ De conformidad con el art. 202, apartado segundo, del Código penal italiano, cuando el sujeto no ha cometido un hecho delictivo, sólo se podrá aplicar una medida de seguridad si así lo prevé expresamente la Ley penal. Es el caso de los cuasi-delitos mencionados en la nota anterior, que están previstos específicamente en el Código.

⁵⁵ Con los requisitos y las formas previstas en el Código de las leyes antimafia y de las medidas de prevención (D.Lvo. italiano n. 159/2011, de 6 de septiembre). Para una

Por otro lado, con respecto al primer caso, es decir, a la fundamentación de un reproche penal, el ordenamiento vigente exigiría ser adaptado para incluir, entre los sujetos pasivos de cada delito que se considere relevante, los sistemas de inteligencia artificial. Esto se puede realizar, esencialmente, a través de la introducción de un título *ad hoc*, que incluyera delitos específicos cometidos en contra de dichos sistemas, o bien mediante el estudio de una cláusula general de aplicación automática en conjunción con las normas ya existentes, al hilo de la ‘cláusula de equivalencia’ establecida en el art. 40, apartado segundo, del Código penal italiano⁵⁶, si bien en este caso habría que manejar una pluralidad de desafíos interpretativos que quizás el Derecho penal –y con esta sinécdoque sus intérpretes– tendrían el placer de prevenir.

Pues bien, de considerar que el catálogo de los bienes jurídicos protegidos merece incorporar –aun bajo la máscara del *délit obstacle*– la integridad de las máquinas, aparecen las dudas teleológicas con respecto a la aplicación de la sanción correspondiente, que finalmente repercuten en la exigencia de culpabilidad del autor. De hecho, esta sanción quizás tendría sentido desde la rechazable visión orientada hacia el Derecho penal del enemigo impulsado por JAKOBS, que justificaría el «amplio adelantamiento de la punibilidad, es decir, [el] cambio de la perspectiva del hecho producido por la del hecho que se va a producir»⁵⁷. Sin embargo, para cumplir con las exigencias de prevención especial propiciadas por ROXIN, habría que averiguar «hasta qué punto, en los casos de circunstancias personales irregulares o condicionadas por la situación conviene una sanción penal a una conducta que, en principio, está amenazada con una pena»⁵⁸. Y es evidente que, en palabras del mismo autor, «una educación para la vida legal en la libertad del Estado de Derecho puede difícilmente llevarse a cabo con la privación de todas las libertades»⁵⁹.

imagen de los antiguos supuestos de peligrosidad *ante delictum*, v. F. TAGLIARINI, voz *Pericolosità*, en «Enciclopedia del diritto», XXXIII, 1983, § 12.

⁵⁶ «No impedir un evento que se tiene la obligación de impedir equivale a causarlo». Con esta fórmula, el Código penal italiano amplía la esfera de los delitos incorporando las omisiones impropias.

⁵⁷ JAKOBS, *La autocomprensión de la ciencia del Derecho penal ante los desafíos del presente*, en *La ciencia del Derecho penal ante el nuevo milenio*, dirigido por A. Eser, W. Hassemer, B. Burkhardt, Tirant lo Blanch, Valencia, 2004, p. 58.

⁵⁸ ROXIN, *Política criminal*, cit., p. 59.

⁵⁹ *Ivi*, p. 50. Hay determinadas conductas que «en el caso concreto devengan en un problema individual y deban tenerse en cuenta también dentro de lo posible y de los límites que permita la prevención general las finalidades preventivas especiales, p. ej., la resocialización del delincuente o, por lo menos, su no desocialización». Así,

Con ello, no sería conveniente instrumentalizar la intervención penal anticipándola hasta las conductas dirigidas a los sistemas inanimados, ya que tendría un carácter marcadamente 'educativo', cuyo objetivo difícilmente podría alcanzarse a través de la privación de las libertades esenciales. Más aún, no se puede olvidar que «cuanto mayor sea una intervención en un derecho fundamental, tanto mayor deberá ser la certeza de las premisas que fundamentan la intervención»⁶⁰.

Todavía está por analizar la posibilidad de utilizar la opción menos invasiva, es decir, aplicar una medida de seguridad al autor de las conductas en cuestión que se revele peligroso. En este sentido, de acuerdo con el modelo italiano, además de poder declarar la peligrosidad en su forma específica (habitualidad, profesionalidad y/o tendencia a delinquir⁶¹), también se puede valorar su modalidad genérica o simple⁶², para cuya apreciación es necesario considerar los criterios indicados en el mismo artículo que establece las reglas para medir la pena⁶³, con una formulación que todavía refleja el influjo de algunas exigencias (en parte vetustas) del contenido de la culpabilidad⁶⁴.

Pues bien, además de las complejidades debidas a la diferencia entre los distintos sistemas penales, que implicaría un tratamiento heterogéneo en los varios ordenamientos europeos, cabe insistir en las mismas razones formuladas anteriormente con respecto a la aplicación de la pena. En efecto, los elementos que conforman y de los que se desprende la peligrosidad social de un sujeto se acercan insidiosamente a las circunstancias valoradas en el ámbito de la culpabilidad. Pese a que peligrosidad y culpabilidad tengan

Id., *Culpabilidad y prevención en Derecho penal*, Editorial Reus, Madrid, 1981, p. 29, traducción de F. Muñoz Conde.

⁶⁰ R. ALEXY, *Teoría de la argumentación jurídica. La teoría del discurso racional como teoría de la fundamentación jurídica*, Centro de Estudios Políticos y Constitucionales, Madrid, 1989, p. 370, traducción de M. Atienza e I. Espejo.

⁶¹ Cfr. el art. 109, apartado primero, del Código penal italiano: «la declaración de habitualidad o de profesionalidad en el delito o bien la de tendencia a delinquir conlleva la aplicación de medidas de seguridad».

⁶² Se trata de la peligrosidad social de los sujetos que, tras haber realizado un hecho delictivo o un cuasi-delito (cfr. *supra*, *sub* notas 53 y 54), manifiesten el riesgo de reiteración delictiva. Sobre la diferenciación entre peligrosidad simple y específica, cfr. V. MANZINI, *Trattato di diritto penale italiano*, vol. III, Utet, Torino, 1981, pp. 249 ss.

⁶³ El art. 203, apartado segundo, del Código penal italiano establece que la condición de persona socialmente peligrosa se desprende de las circunstancias indicadas en el art. 133, que a su vez incluye los criterios para la medición de la pena (gravedad del hecho y capacidad para delinquir).

⁶⁴ Cfr. SANTAMARIA, *op. ult. cit.*, pp. 660 ss.

que mantenerse distintas por fundamentar dos consecuencias jurídicas diferentes⁶⁵, es evidente que las circunstancias personales del sujeto, que se valoran –desde una perspectiva distinta, intrafactual o extrafactual– terminan por ser las mismas.

Quizás las exigencias de prevención especial reeducativa, que conlleva la comisión de conductas abstractamente delictivas hacia los sistemas de inteligencia artificial, justificarían la anticipación de la tutela penal en el marco de un Derecho penal de la voluntad (*Willensstrafrecht*) o, tal vez mejor, de la actitud interna (*Gesinnungsstrafrecht*), que sin embargo son incompatibles con el principio del hecho, impuesto por los mandatos constitucionales⁶⁶.

Hoy en día, pues, y hasta que haya cambios relevantes en el ámbito de la ingeniería relacionada con de los sistemas de inteligencia artificial, el principio de subsidiariedad quizás respalda la protección de estos nuevos ‘sujetos pasivos’ a través de otras ramas del Derecho⁶⁷.

3.2. *Hacia una subjetividad activa instrumental: la inteligencia artificial como autora de delitos*

Después de haber analizado las cuestiones relativas a la inteligencia artificial como potencial víctima de un delito, cabe plantear las problemáticas inherentes a la otra cara de la moneda, es decir, a la posibilidad de que los sistemas inanimados puedan encajar en el concepto de autoría a efectos de la Ley penal. En otras palabras, para conseguir el objetivo que pretende alcanzar este escrito no se puede prescindir del análisis sobre la oportunidad de castigar un sistema inanimado y, de así estimarla, habrá que examinar sus implicaciones en la adecuación del ordenamiento jurídico-penal.

En primer lugar, pues, surge la cuestión de si es oportuno ‘reprochar’

⁶⁵ *Ivi*, p. 663 ss.

⁶⁶ BRICOLA, *op. ult. cit.*, pp. 65 ss.; J.M. PERIS RIERA, *Delitos de peligro y sociedad de riesgo: una constante discusión en la dogmática penal de la última década*, en *Estudios penales en homenaje al profesor Cobo del Rosal*, dirigido por J.C. Carbonell Mateu, B. del Rosal Blasco, L. Morillas Cueva, E. Orts Berenguer, M. Quintanar Díez, Dykinson, Madrid, 2005, pp. 687 ss. Con respecto a la proximidad entre el Derecho penal del autor y el Derecho penal del enemigo, v. G. JAKOBS, M. CANCIO MELIÁ, *Derecho penal del enemigo*, Civitas, Madrid, 2003, pp. 100 ss.

⁶⁷ Sobre las cuestiones de legitimidad de la intervención penal, v. M. CORCOY BIDASOLO, *Exigibilidad en el ámbito del conocimiento y control de riesgos: teorización*, en «Revista catalana de seguretat pública», 13/2003, pp. 31 ss; *amplius*, EAD., *Delitos de peligro y protección de bienes jurídico-penales supraindividuales*, Tirant lo Blanch, Valencia, 1999, *passim*.

a un sistema inanimado, en el sentido de que se le pueda conceder subjetividad activa. Sólo después, de considerar adecuada dicha concesión, se podrá analizar la medida en que podría adaptarse la legislación vigente. La doctrina se ha ido interrogando sobre la duda de si *machina delinquere potest*⁶⁸, al igual que ocurrió en su momento con las personas jurídicas⁶⁹. De hecho, la fórmula aquí referida se ha plasmado sobre la base de otro conocido brocardo que, en lugar de la máquina, alude a la sociedad⁷⁰.

Pues bien, ante la duda se puede intentar razonar *a contrario* y plantear la hipótesis de una respuesta negativa. Por razones de lógica (quizás simplista), y sobre todo en virtud del empirismo que informa hasta las quiebras más evidentes del Derecho penal, la habitualidad en concebir formas de autoría estrictamente ligadas a los seres humanos podría llevar a excluir perentoriamente que los sistemas inanimados adquieran subjetividad activa. No obstante, dicha denegación supondría la elusión de exigencias de política criminal que hoy en día están lanzando signos de alarma.

En efecto, no faltarían motivos para obstaculizar el reconocimiento de la subjetividad activa en cuestión. Es suficiente pensar en la 'personalidad' de la responsabilidad penal, mencionada arriba, o en las propias palabras de COBO DEL ROSAL sobre la persona jurídica⁷¹, para asombrarse ante una forma de autoría tan revolucionaria y, al mismo tiempo, temerosa. Sin embargo, al igual que en la historia se han sucedido modificaciones

⁶⁸ Cfr. J.M. MUÑOZ VELA, *Inteligencia Artificial y responsabilidad penal*, en «Derecho Digital e Innovación. Digital Law and Innovation Review», 11/2022, pp. 1 ss.; J. VALLS PRIETO, *Inteligencia artificial, Derechos humanos y bienes jurídicos*, Aranzadi, Pamplona, 2021, *passim*; P.M. DE LA CUESTA AGUADO, *Inteligencia artificial y responsabilidad penal*, en «Revista penal México», 2020, pp. 51 ss.; I. LLEDÓ BENITO, *El impacto de la robótica. La inteligencia artificial y la responsabilidad penal en los robots inteligentes*, en «Foro galego. Revista xurídica», 2020, pp. 173 ss.; D. LIMA, *Could AI Agents Be Held Criminally Liable? Artificial Intelligence and the Challenges for Criminal Law*, en «South Carolina Law Review», 3/2018, pp. 677 ss.

⁶⁹ *Amplius*, J. PÉREZ ARIAS, *Sistema de atribución de responsabilidad penal a las personas jurídicas*, Dykinson, Madrid, 2014, *passim*.

⁷⁰ BASILE, *Intelligenza artificiale*, cit., *sub nota* 100, p. 27.

⁷¹ «Desde otra perspectiva, puede concluirse que, en ninguna parte del vigente Código penal se afirma que se construya una autoría criminal de las personas jurídicas. Porque la autoría ni se puede improvisar ni se puede derivar de una prejuiciosa y errática interpretación, por muy buena y amplia voluntad represiva que conlleven mentalidades defensoras extremadamente de la sociedad o del Estado». Si se lee el Código penal «como debe ser, de forma no locoide e imprejujada, no hay posibilidad de afirmar una *societas* como realizadora, en régimen de autoría estricta o de complicidad, de infracción penal de clase alguna». Así, COBO DEL ROSAL, *Societas delinquere non potest*, en «Anales de derecho», 2012, pp. 8-9.

(incluso) en la teoría del delito, como resultado de la evolución de la realidad social en las distintas épocas⁷², quizás es posible plantear cambios también con respecto a conceptos más estrechos. Pues, las renovadas exigencias sociales necesitan, al menos, ser evaluadas por el Derecho, que en su caso se adaptará con las herramientas que posee.

Resulta incontrovertido que ya estamos en una época en la que el *machine learning* ha salido del control de los diseñadores de los sistemas de inteligencia artificial. Desde luego, si en el futuro se da la posibilidad de supervisar exhaustivamente este proceso de autoaprendizaje de las máquinas, las consideraciones que aquí se están desarrollando tendrían un alcance diferente y, por ello, habrían de reformularse. Sin embargo, en la actualidad, la ciencia parece viajar en sentido inverso. Pues, como ya se ha dicho, la evolución de la tecnología está llevando a la peligrosa aparición de sistemas que, aunque se hayan diseñado para alcanzar determinados fines, escapan al control de los programadores y se atreven no sólo a actuar de forma maquiavélica, de modo que cualquier medio (aun ilegal) se hace adecuado mientras se llegue al objetivo, sino también a buscar otros propósitos que ni se les habían ocurrido a los diseñadores.

Siendo este el escenario, es oportuno que el Derecho preste atención a los cambios repentinos y a las evoluciones de la técnica, ya que tienen efectos considerables en la sociedad. En modo alguno esto implica la equiparación entre artefactos y seres humanos, ni mucho menos. Las máquinas quedan al servicio del hombre, tanto en el caso de que la máquina sea dirigida por el ser humano, como cuando ella actúe de forma independiente. Y para que estén al servicio del hombre, el Derecho positivo debe reflejar este axioma.

De no ser así, es decir, de desvincular la actuación inhumana del control humano, se daría rienda suelta a la derresponsabilización moral de los individuos, que lleva aparejada una inaceptable debilitación de la tutela de los bienes jurídicos⁷³. A modo de ejemplo, cualquier sujeto interesado en cometer una conducta delictiva podría alcanzar su objetivo amparándose en la utilización de una máquina y en su capacidad de autoaprendizaje, habida cuenta de las dificultades probatorias y causales que propiciarían su impunidad.

Resumiendo, para contestar a la pregunta que impulsó estas reflexiones, es oportuno y necesario que el Derecho se plantee la posibilidad de

⁷² Cfr. *supra*, *sub* nota 1.

⁷³ Así advierte BASILE, *Intelligenza artificiale*, cit., p. 29, mencionando, *sub* nota 106, la «deshumanización de las ofensas» descrita por F. CONSULICH, *Il nastro di Mobius. Intelligenza artificiale e imputazione penale nelle nuove forme di abuso del mercato*, en «Banca Borsa Titoli di Credito», 2/2018, p. 204.

‘reprochar’ a los sistemas inanimados de inteligencia artificial, a fin de evitar una quiebra evidente de los bienes jurídicos ya protegidos por el ordenamiento. No resulta difícil imaginar la época en que los accidentes de tráfico se deban a coches conducidos por sí mismos, las pizarras digitales de los centros de educación se conviertan en robots autónomos capaces de lesionar a los discentes, los mismos robots adquieran impulsos sexuales aptos para llevarlos a cometer violaciones, o los algoritmos aprecien el valor del dinero hasta llegar a sustraerlo de unas cuentas bancarias y traspasarlo a otras. Así pues, el Derecho debe reaccionar antes de perder el control sobre la ordenación de una sociedad que paulatinamente está avanzando hacia nuevas fronteras; en otras palabras, antes de que la tenencia de un carné de conducir se convierta en algo tan insólito como, a día de hoy, no tenerlo.

Ahora bien, al haber apreciado la oportunidad (*an*) de atribuir subjetividad activa a los sistemas de inteligencia artificial, procede analizar la manera (*quomodo*) en que dicho reconocimiento se puede llevar a cabo en el marco del ordenamiento penal. Hasta la fecha, la responsabilidad penal gira en torno a la comisión u omisión de conductas realizadas exclusivamente por el ser humano, no sólo desde la perspectiva determinista, ya que también son delictivos algunos hechos realizados sin que subsista la intención del agente. En este sentido, de nada sirve objetar que también las personas jurídicas poseen subjetividad activa en Derecho penal. De hecho, como se dirá en breve, entre la *societas* y la *machina* hay diferencias significativas, empezando por la dirección siempre ‘humana’ de la persona jurídica, que al contrario de los productos de inteligencia artificial, no tiene autonomía cognitiva y decisional propia.

Lo que ahora se plantea, en cambio, es el traspaso de actuaciones penalmente relevantes del ser humano a una pluralidad de máquinas independientes. De algún modo se sale de la esfera de actuación humana para entrar en un escenario donde, por leves que sean, los comportamientos se desarrollan autónomamente por sistemas inanimados, lo que conlleva una alienación de la responsabilidad del individuo⁷⁴. Debe advertirse sin más tardanza que se presupone un contexto en el que las máquinas son capaces de actuar independientemente del control humano; de lo contrario, las consideraciones que figuran a continuación carecerían de sentido.

El análisis se desarrollará siguiendo dos directrices paralelas que sólo se escinden por razones de comodidad expositiva. Por un lado, se intentará examinar las implicaciones del reconocimiento de la subjetividad activa de los sistemas inanimados en la teoría del delito; por otro, se estudiará

⁷⁴ Cfr. C. BAGNOLI, *Teoria della responsabilità*, Il Mulino, Bologna, 2019, p. 77.

el mismo fenómeno desde un punto de vista penológico, pues no es posible imaginar un sistema de normas penales que despliegue sus efectos coercitivos sin hallarse respaldado por la posibilidad de conminar una sanción efectiva.

En el marco de la teoría del delito y, más detenidamente, del contenido de la culpabilidad, aunque se cuestione la idea de un concepto unitario⁷⁵ no cabe duda de la necesidad de integrar los varios elementos establecidos por el Derecho positivo para apreciar la existencia de un delito. Ahora bien, en la medida en que se teorice una autoría mecánica, esta implicaría la acreditación de una culpabilidad ‘deshumana’⁷⁶, o en sentido similar pero distinto ‘deshumanizada’. En ambos los casos, hay que afrontar la dificultad, por un lado, de adaptar conceptos que son propios de las personas físicas y, por otro, de reformularlos según las peculiaridades de los nuevos sujetos activos.

En el primer caso, para que se pueda hablar de una culpabilidad ‘deshumana’ se debería adaptar el régimen penal ordinario para centrarlo en las máquinas, cuya responsabilidad penal tendría que superar los peldaños exigidos por la dogmática (más bien, por el Derecho positivo). En este sentido, para que se pueda apreciar la existencia de un delito, la adaptación del Derecho penal conllevaría la incorporación de conceptos todavía ambiguos, elaborados por las distintas ciencias que operan en el sector de la inteligencia artificial. En otras palabras, el Derecho penal debería aceptar incondicionalmente los fundamentos científicos del *machine learning*, según los cuales las propias máquinas dispondrían de capacidad natural, o bien de capacidad de entender y querer, serían capaces de comprender el alcance de la antijuridicidad de una conducta y, más aún, podrían realizar comportamientos que se encajan en el complejo marco del dolo y de la imprudencia.

Ahora bien, aunque ningún ordenamiento se haya atrevido a acogerla, en la doctrina hubo manifestaciones aisladas de apoyo a dicha postura. Entrando en este posicionamiento, los problemas más evidentes están relacionados con el dolo y la imprudencia, además de la conciencia de la antijuridicidad y su contenido. Efectivamente, ¿es posible imaginar que un algoritmo posea la voluntad consciente de realizar un hecho delictivo, o más aún, sea capaz de desatender una regla cautelar que fundamente un juicio de reprochabilidad por imprudencia? Las opiniones que admiten dichas posibilidades⁷⁷ parten de la configurabilidad en abstracto de una tipicidad mecánica, dejando al margen como mínimo los problemas relativos a la

⁷⁵ Cfr. *supra*, sub nota 23.

⁷⁶ BASILE, *Intelligenza artificiale*, cit., p. 30.

⁷⁷ HALLEVY, *op. ult. cit.*, pp. 171 ss.

estructura de la antijuridicidad en los sistemas jurídicos continentales y las críticas que surgirían desde la perspectiva procesal.

Aunque se trate de un modelo de responsabilidad directa abstractamente configurable, se trataría de condenar por acciones y omisiones naturalísticas –*rectius*, actividades e inactividades– a las que el Derecho penal otorga relevancia de por sí, sin que siquiera existan instrumentos para averiguar el elemento subjetivo subyacente. Se ha hablado de *general intent*⁷⁸, pero no resulta claro cómo este vínculo psicológico entre la conducta y la máquina pueda resultar probado a efectos penales. Además, con respecto a la antijuridicidad, los modelos antropocéntricos de algunas causas de justificación –y lo mismo debe predicarse sobre algunas de las demás eximentes consideradas, por la mayoría, causas de exculpación⁷⁹– podrían suponer, de forma sistemática, la impunidad de la máquina.

Piénsese, a modo de ejemplo y sin entrar en casos más complejos, en el estado de necesidad alegado por (¿quién representa?) el sistema de inteligencia artificial que, para evitar un mal propio (en el que cabría simplemente su alteración o descomposición mecánica), lesione un bien jurídico de un ser humano; o incluso en el sistema inanimado que, diseñado por o al servicio de la autoridad, intente cumplir el deber (legítimo) de realizar la tarea encomendada y para eso se vea obligado a realizar una conducta delictiva. Todo ello sin perjuicio de los problemas de carácter procesal⁸⁰, empezando por la dificultad de encontrar un procurador dispuesto a representar la máquina en el juicio sin poder siquiera comunicar con ella, hasta llegar a enflaquecer la intolerancia del delito «ad ogni articolazione che faccia in qualche modo ricadere sull'imputato l'onere della prova (o il rischio della mancata allegazione) di elementi d'ordine positivo che ne caratterizzano la struttura»⁸¹.

Acoger este modelo de responsabilidad directa se hace, pues, inviable por suponer la creación de un tejido de problemas⁸² –entre ellos los aquí referidos– difícilmente solucionables. Mientras se ‘humanizan’ los sistemas de inteligencia artificial, permanecerían ‘deshumanos’ –y, entonces, sin contenido– algunos de los elementos que se han plasmado en la teoría del delito sobre la base de las exigencias y especificidades del ser humano.

Por otro lado, con un grado de viabilidad más elevado se puede plantear

⁷⁸ *Ivi*, pp. 177 ss.

⁷⁹ Cfr. el concepto de antijuridicidad especialmente cualificada (*supra*, *sub* nota 23).

⁸⁰ Al respecto, v. R. CASTILLO FELIPE, S. TOMÁS TOMÁS, *Proceso penal e inteligencia artificial: prevenciones en torno a su futura aplicación en la fase de juicio oral*, en esta obra.

⁸¹ BRICOLA, *op. ult. cit.*, p. 92.

⁸² Cfr., *amplius*, A. CAPPELLINI, *Machina delinquere non potest? Brevi appunti su intelligenza artificiale e responsabilità penale*, en «Criminalia», 2018, pp. 499 ss.

la elaboración de una culpabilidad ya no deshumana, sino ‘deshumanizada’, en el marco de una dogmática que sabe adaptarse a las renovadas necesidades sociales, que constan de una creciente demanda de tutela frente a posibles ataques de los sistemas de inteligencia artificial lesivos de bienes jurídicos constitucionalmente protegidos.

La elaboración de este modelo de responsabilidad penal de los sistemas de inteligencia artificial supondría un cambio de paradigma, es decir, la creación de una *fictio iuris* capaz de otorgarles personalidad jurídica y aptitud autorial. Se trataría de una responsabilidad autónoma frente a la que seguiría perteneciendo a las personas físicas. De hecho, como ya se advirtió anteriormente, aquí se toman en consideración los comportamientos de los sistemas inanimados que actúan de forma independiente, pues, de actuar conjuntamente con un sujeto humano, la implementación de este modelo no perjudicaría la responsabilidad individual.

La creación de dicha ficción jurídica se vería facilitada por el antecedente significativo de las *societates*, aunque existan diferencias considerables entre las dos figuras. Desde el punto de vista de la regulación aplicable, quizás convendría dotar a los sistemas de inteligencia artificial de personalidad jurídica a través de la inscripción en registros públicos, para que pueda plasmarse coherentemente su peculiar imputabilidad. Del mismo modo, sería oportuno restringir el área de lo penalmente relevante mediante la selección de un *numerus clausus* de hechos punibles, bien a través de una cláusula extensiva genérica, bien con la incorporación de tipos autónomos de aplicación exclusiva a los sujetos activos de nuevo cuño, además de rediseñar el sistema de las eximentes para adaptarlo a exigencias ya no humanas.

Sin embargo, algunas soluciones escogidas para las personas jurídicas no serían adecuadas para regular el fenómeno de la inteligencia artificial. En primer lugar, no procedería restringir la subjetividad activa excluyendo los sistemas inanimados pertenecientes al Estado, ni podría estigmatizarse un catálogo de personas físicas al hilo del primer apartado del actual art. 31 *bis* del Código penal español. En esto radica la principal diferencia entre personas jurídicas y máquinas: estas últimas carecen del sustrato humano que dirija su actuación. Entiéndase: en el caso de que las máquinas cometiesen un delito en virtud de la dirección humana, no cabe duda de que sea la persona física la que debe responder penalmente, con arreglo a la legislación existente; el problema surge cuando la máquina actúa independientemente de las determinaciones (aun dolosas o imprudentes) humanas.

Por ello, también hay que regular de forma distinta el criterio de imputación subjetiva para dichas actuaciones mecánicas. Pues, la tesis predominante en el ámbito de las personas jurídicas acerca de la necesidad

de acudir a la persona física que realmente llevó a cabo la conducta delictiva, encuentra respaldo en la existencia de ese sustrato humano que aquí, en el caso de los sistemas autónomos de inteligencia artificial, no se puede apreciar. Ante la carencia de pruebas científicamente satisfactorias que permitan averiguar un verdadero elemento volitivo de las máquinas, y frente a la imposibilidad de escoger un 'agente modelo' que presente *eaedem condiciones et professiones*, quizás el caso de la inteligencia artificial puede constituir el único supuesto de responsabilidad objetiva que la Constitución esté dispuesta a tolerar. De ser así, la responsabilidad penal de los sistemas inanimados prescindiría de la existencia del vínculo psicológico entre el 'autor' y el hecho, y asimismo de la cognoscibilidad de la Ley penal por parte del sujeto activo.

Por último, después de haber intentado examinar algunas de las implicaciones del reconocimiento de la subjetividad activa de los sistemas inanimados, queda por abordar el mismo fenómeno desde un punto de vista penológico. No hace falta detenerse en la exigencia de que, en relación con todo delito, las penas sean efectivas y proporcionadas al hecho cometido⁸³; de no ser así, la tipificación del delito quedaría sin efecto alguno.

Acogiendo la adaptación del Derecho penal que se ha propuesto para admitir que los sistemas de inteligencia artificial puedan cometer delitos, el debate sobre el fundamento y las funciones de las penas correspondiente requeriría un examen mucho más amplio. Aun así, se intentará desarrollar unas breves consideraciones tanto desde la perspectiva teleológica, con respecto a las tres teorías clásicas sobre la finalidad de la sanción, como desde la perspectiva práctica, pues si la sanción efectiva no tiene la aptitud para lograr el objetivo fijado a nivel abstracto, se hace innecesaria y el ejercicio del *ius puniendi* estatal representaría un mero desgaste. Debe advertirse previamente que, mientras «el fundamento del castigo penal que recibiría la persona jurídica sería la existencia de defectos en su organización, la tenencia de una estructura opaca, la inexistencia de controles sobre sus empleados»⁸⁴, en el caso de las máquinas no puede limitarse a eso.

La idea que debería apoyarse es que la elaboración del modelo de

⁸³ Ya advertía dichas exigencias C. BECCARIA, *Tratado de los delitos y de las penas*, Heliasta, Buenos Aires, 1993, pp. 68 ss.

⁸⁴ A. NIETO MARTÍN, *Problemas fundamentales del cumplimiento normativo en el derecho penal*, en *Compliance y teoría del Derecho penal*, dirigido por L. Kuhlen, J.P. Montiel, Í. Ortiz de Urbina Gimeno, Marcial Pons, Madrid, 2013, pp. 21 ss. Se expresa en contra de la responsabilidad penal de las personas jurídicas L. GRACIA MARTÍN, *Crítica de las modernas construcciones de una mal llamada responsabilidad penal de la persona jurídica*, en «Revista Electrónica de Ciencia Penal y Criminología», 2016, pp. 1 ss.

responsabilidad penal de los sistemas en cuestión, en el que se incluye la aplicación de las sanciones correspondientes, tiende exclusivamente a la protección última del ser humano. Por dicha razón, y dado que en los modernos sistemas constitucionales se ha ido sentando la idea de que la retribución subyacente a la pena sólo tiene sentido como medio preventivo, con las sanciones dirigidas a las máquinas se intentaría alcanzar, por un lado, una finalidad de prevención especial, que está encaminada a prevenir la reiteración delictiva por parte del ‘autor’, y por otro lado, un objetivo de prevención general, porque la vigencia de estas normas, además de tranquilizar a la colectividad, podría tender a evitar que se diseñen sistemas de inteligencia artificial parecidos. Con ello, como consecuencia del acogimiento del *versari in re illicita*, se debería renunciar a la función reeducativa prevista en el régimen general.

A nivel práctico, las sanciones que pueden conminarse a las máquinas no se enmarcan en el catálogo tradicional de las penas. En este sentido, es hasta antieconómico aplicar a una máquina la pena de prisión, así como imposible atacar un patrimonio que no posee. Con las dificultades que conllevan las hipótesis más complejas –*e.g.*, ¿cómo se plantea la destrucción o inutilización de un algoritmo?– se trata de estudiar un nuevo sistema penológico que esté de acuerdo con las finalidades que merecen ser perseguidas. Al efecto, téngase en cuenta que las dos sanciones quizás más fáciles de imaginar (el apagado definitivo y temporal de la máquina) traicionarían la teleología fijada arriba. Mientras el apagado temporal del sistema inanimado no lograría prevenir la reiteración delictiva por el mismo, su destrucción (o inutilización, o apagado definitivo) –que en el fondo refleja los caracteres de la pena capital– sería aceptable al haber renunciado al propósito reeducativo, pero a lo mejor obstaculizaría la posibilidad de mejorar el diseño de las demás máquinas.

3.3. La inteligencia artificial como instrumento de apoyo para el ejercicio de la justicia

Aunque no se puedan abordar en esta sede los aspectos procesales implicados, parece oportuno hacer mención a las consecuencias del uso de la inteligencia artificial como instrumento de apoyo para el ejercicio de la justicia.

Por un lado, hay que hacer hincapié en la posible afectación del modelo de libre valoración de la prueba por parte del juez, que se vería afectado en el caso de intervenir instrumentos de inteligencia artificial,

como los polígrafos, en la detección de la veracidad de las pruebas. Dicha implementación podría intervenir tanto de manera dura, si al juez no se le deja la posibilidad de moldear los resultados elaborados por la máquina, como de manera débil, si la mediación del juez persona física permite mantener en vigor el principio del *iudex peritus peritorum*. Es evidente que, en el marco de una justicia que ha de ser administrada en favor del reo, las máquinas que analizan los hechos, estudian las pruebas e incluso producen resoluciones judiciales, no pueden alcanzar la autonomía que supondría la opción de intervención más dura⁸⁵. Se hace necesario, ahora más que nunca, el control constante del órgano judicial para escapar de los supuestos en que la mecánica se apodera del control sobre las libertades fundamentales de los individuos. Así se reafirman las exigencias de culpabilidad que imponen la huida de Derechos penales desvinculados del hecho y apegados a los procesos emocionales internos del autor.

Por otro lado, aunque parezca ciencia ficción, no se puede descartar *a priori* que alguien llegue a formular una teoría del delito en la que el medio tecnológico de apoyo a la justicia se convierta en auténtico elemento negativo del hecho. Casi actuando como una causa de justificación, el reproche penal pasaría a ser el producto de factores que incluyen entre sí la estimación de la máquina acerca de la producción (o no) de un hecho delictivo. Las mismas consideraciones que se acaban de formular con respecto a la necesidad de un juicio ‘natural’ fundamentan, *a fortiori*, el rechazo de tal deterioro de la teoría del delito.

4. La incidencia de los sistemas de inteligencia artificial en la culpabilidad como criterio para la determinación de la pena en concreto

La segunda cara de la moneda ‘culpabilidad’ tiene que ver con el ámbito de la determinación de la pena en concreto⁸⁶. Desde esta perspectiva, como ya se ha adelantado, por un lado la inteligencia artificial pondría en peligro la operatividad de la proporcionalidad de la pena, al prever mecanismos que podrían favorecer la aplicación de sanciones fijas, y por otro lado, puede ocurrir que la máquina, al aplicar la pena en lugar del juez persona física, genere una condena injusta o arbitraria, o bien sea dictada como consecuencia de influencias sobornadoras externas.

Este último caso, antes de convertir la máquina en la autora de un delito

⁸⁵ Cfr. CASTILLO FELIPE, TOMÁS TOMÁS, *op. ult. cit.*, § 2.4.

⁸⁶ Cfr. las consideraciones críticas efectuadas *supra*, *sub* notas 63 y 64.

de cohecho o prevaricación, no representa sino un ulterior argumento de apoyo a los razonamientos sobre la inoportunidad de que los sistemas de inteligencia artificial sustituyan al juez persona física. En el marco de una intervención débil de las máquinas como instrumentos de apoyo al ejercicio de la justicia, no surgirían estos problemas. El único culpable de haber dictado una resolución injusta o ilícita seguiría siendo el juez, de acuerdo con las leyes que regulan su responsabilidad penal.

En cambio, un ataque más evidente se dirige al principio de proporcionalidad, cuyo fundamento constitucional le impide ser afectado por el uso de máquinas en el ejercicio de la justicia. Estas, quebrantando los matices de la discrecionalidad judicial clásica, no lograrían tomar en consideración todos los segmentos que constituyen un determinado marco edictal, llevando en última instancia a la aplicación de penas fijas, de por sí incompatibles con el ordenamiento jurídico-penal vigente. Para analizar este fenómeno, es suficiente pensar en el funcionamiento de los algoritmos que estarían detrás del cálculo de las penas operado por los sistemas de inteligencia artificial. Cada uno de los factores que intervienen en la averiguación de la verdad judicial conllevaría, en términos de porcentaje o unidad mínima (un día), el aumento o la disminución de pena empezando por el mínimo establecido para el caso específico.

Pese a que un listado exhaustivo de 'factores' (o circunstancias) a los que se debe atribuir un determinado efecto (aumento o rebaja) penológico es materialmente inconcebible, la conminación final de la pena sería el producto de operaciones matemáticas llevadas a cabo por el algoritmo. Para ejemplificar, habría grupos de circunstancias favorables, que llevarían aparejada una modificación (positiva) de la pena, así como grupos de factores desfavorables, con la opuesta repercusión (negativa) en la sanción a aplicar. Por muy detallado que sea el catálogo de las circunstancias 'matemáticas' teorizado e introducido en el algoritmo, siempre existirían segmentos de la horquilla que, aun formando parte de la pena en abstracto, pasarían a ser evanescentes en concreto.

Así pues, se generaría una casuística cerrada (presencia de un factor desfavorable: modificación en x puntos porcentuales superiores a la pena mínima; presencia de dos factores favorables y uno desfavorable: disminución en y puntos porcentuales; y así sucesivamente) que no es sino una manifestación del automatismo del proceso sancionador. En definitiva, acoger este mecanismo perverso significaría instrumentalizar la 'popularidad' de una peligrosa seguridad jurídica celando, al mismo tiempo, el sacrificio de ineludibles garantías constitucionales frente al poder punitivo del Estado.

5. Conclusiones

Las breves reflexiones desarrolladas en estas páginas permiten formular las siguientes conclusiones:

- i) La inteligencia artificial se ha desarrollado hasta llegar a influir no sólo en la realidad social, sino también en las instituciones básicas del Derecho penal. La implementación de máquinas capaces de autodeterminarse y, en cierta medida, de actuar independientemente de la voluntad humana, ha generado la exigencia de replantear la regulación jurídica frente a los posibles ataques sufridos y causados por estos sistemas inanimados.
- ii) Al analizar la influencia de la inteligencia artificial en el Derecho y, más en detalle, en el Derecho penal, es imprescindible tomar como referencia la definición de delito que hace hincapié en los mandatos constitucionales, para evitar el menoscabo de los derechos y las libertades fundamentales, conquistas del Estado de Derecho contemporáneo: «possiamo definire il reato come un fatto previsto in forma tassativa dalla legge, di realizzazione esclusiva dell'agente o in ogni caso al medesimo riconducibile tramite un atteggiamento colpevole (doloso o colposo), idoneo ad offendere un valore costituzionalmente significativo, minacciato con una pena proporzionata anche alla significatività del valore tutelato e strutturalmente caratterizzato dal teleologismo costituzionalmente attribuito alla sanzione penale e, infine, intollerante rispetto ad ogni articolazione probatoria che faccia in qualche modo ricadere sull'imputato l'onere della prova (o il rischio della mancata allegazione) di elementi d'ordine positivo che ne caratterizzano la struttura»⁸⁷.
- iii) La influencia de los sistemas de inteligencia artificial se manifiesta ante todo en la teoría del delito, según el sistema inanimado aparezca como potencial víctima o autor de un hecho delictivo, o bien como instrumento de apoyo para el ejercicio de la justicia.
- iv) Al intentar conceder subjetividad pasiva a los sistemas de inteligencia artificial, surgen las dudas teleológicas con respecto a la aplicación de la sanción correspondiente al autor del hecho. No parece conveniente anticipar la intervención penal hasta las conductas dirigidas a los sistemas inanimados, concretamente no lesivas de bienes jurídicos pertenecientes a los seres humanos. La finalidad educativa que se perseguiría con dicha intervención penal se enmarcaría en un Derecho penal de la voluntad o de la actitud interna, que sin embargo son incompatibles con el princi-

⁸⁷ BRICOLA, *op. et loc. ult. cit.*

pio del hecho, impuesto por los mandatos constitucionales.

- v) En cambio, parece oportuno conceder a los sistemas de inteligencia artificial subjetividad activa, si bien a través de una *fictio iuris* que supone la renuncia a algunas exigencias dogmáticas básicas. Habida cuenta de la creciente autonomía que el *machine learning* les va proporcionando, resulta necesario tutelar al ser humano frente a los ataques realizables por estos sistemas inanimados. La hipótesis de una responsabilidad directa de las máquinas con las herramientas del Derecho actualmente vigente supondría la implementación de una culpabilidad ‘deshumana’, que erigiría su estructura aceptando postulados científicos ajenos al Derecho penal y todavía envueltos por mucha ambigüedad. Por tanto, se propone la creación de un modelo alternativo de responsabilidad penal de los sistemas de inteligencia artificial, basado en una ficción jurídica ‘des-humanizada’ que permita renunciar incluso a las más elementales exigencias del delito ‘humano’, hasta llegar a posibilitar el único supuesto de responsabilidad objetiva que la Constitución esté dispuesta a tolerar.
- vi) La utilización de la inteligencia artificial como instrumento de apoyo para el ejercicio de la justicia conlleva el peligro de alterar la estructura del principio *iudex peritus peritorum*, cuyo alcance fundamenta la libre valoración de la prueba en el proceso penal. Ante la amenaza de dicha alteración, es oportuno que la inteligencia artificial siga siendo utilizada bajo el control, la supervisión y la responsabilidad del juez persona física.
- vii) Con respecto a los efectos de la inteligencia artificial en la culpabilidad entendida como criterio para la determinación de la pena en concreto, se rechaza íntegramente la creación de algoritmos que, de algún modo, puedan quebrantar el contenido del principio de proporcionalidad, constitucionalmente garantizado frente a la arbitrariedad de las decisiones legislativas y judiciales.

6. Bibliografía

- AGUADO CORREA T., *Principio de inexigibilidad de otra conducta en las categorías del delito*, en «Revista Digital de la Maestría en Ciencias Penales de la Universidad de Costa Rica», 3/2011.
- ALCÁCER GUIRAO R., *Los fines del Derecho Penal. Una aproximación desde la filosofía política*, en «Anuario de Derecho Penal y Ciencias Penales», 1/1998.

- ALEX Y R., *Teoría de la argumentación jurídica. La teoría del discurso racional como teoría de la fundamentación jurídica*, Centro de Estudios Políticos y Constitucionales, Madrid, 1989, traducción de M. Atienza e I. Espejo.
- ALONSO SALGADO C., *Acerca de la inteligencia artificial en el ámbito penal: especial referencia a la actividad de las fuerzas y cuerpos de seguridad*, en «Ius et Scientia», 1/2021.
- ARIAS P.C., *Artificial Intelligence & Machine Learning: a model for a new judicial system?*, en «Revista Internacional Jurídica y Empresarial», 2020.
- BAGNOLI C., *Teoria della responsabilità*, Il Mulino, Bolonia, 2019.
- BASILE F., *Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*, en «Diritto Penale e Uomo», 10/2019.
- BASILE F., *L'alternativa tra responsabilità oggettiva e colpa in attività illecita per l'imputazione della conseguenza ulteriore non voluta, alla luce della sentenza Ronci delle Sezioni Unite sull'art. 586 c.p.*, en «Rivista italiana di diritto e procedura penale», 3/2011.
- BECCARIA C., *Tratado de los delitos y de las penas*, Heliasta, Buenos Aires, 1993.
- BERTOLINO M., *Declinazioni attuali della pericolosità sociale: pene e misure di sicurezza a confronto*, en «Archivio penale», 2/2014.
- BRICOLA F., voz *Teoria generale del reato*, en «Novissimo digesto italiano», XIV, 1973.
- CAPPELLINI A., *Machina delinquere non potest? Brevi appunti su intelligenza artificiale e responsabilità penale*, en «Criminalia», 2018.
- CASTILLO FELIPE R., TOMÁS TOMÁS S., *Proceso penal e inteligencia artificial: prevenciones en torno a su futura aplicación en la fase de juicio oral*, en esta obra.
- CAVALIERE A., *Misure di sicurezza e doppio binario. Considerazioni introduttive*, en «Rivista italiana di diritto e procedura penale», 1/2022.
- COBO DEL ROSAL M., *Societas delinquere non potest*, en «Anales de derecho», 2012.
- COBO DEL ROSAL M., VIVES ANTÓN T.S., *Derecho penal. Parte general*, Tirant lo Blanch, Valencia, 1999.
- CONSULICH F., *Il nastro di Mobius. Intelligenza artificiale e imputazione penale nelle nuove forme di abuso del mercato*, en «Banca Borsa Titoli di Credito», 2/2018.
- CORCOY BIDASOLO M., *Exigibilidad en el ámbito del conocimiento y control de riesgos: teorización*, en «Revista catalana de seguretat pública», 13/2003.

- CORCOY BIDASOLO M., *Delitos de peligro y protección de bienes jurídico-penales supraindividuales*, Tirant lo Blanch, Valencia, 1999.
- DANAHER J., *Robotic Rape and Robotic Child Sexual Abuse: Should They be Criminalised?*, en «Criminal Law and Philosophy», 1/2017.
- DE LA CUESTA AGUADO P.M., *Inteligencia artificial y responsabilidad penal*, en «Revista penal México», 2020.
- FIANDACA G., MUSCO E., *Diritto penale. Parte generale*, Zanichelli, Bolonia, 2019.
- FORNARI L., *Misure di sicurezza e doppio binario: un declino inarrestabile?*, en «Rivista italiana di diritto e procedura penale», 1993.
- GALLO M., *Diritto penale italiano. Appunti di parte generale*, vol. I, Giappichelli, Torino, 2021.
- GALLO M., *Il concetto unitario di colpevolezza*, Giuffrè, Milán, 1951.
- GALLO M., *La teoria dell'azione «finalistica» nella più recente dottrina tedesca*, Giuffrè, Milán, 1950.
- GAMERO CASADO E., *El enfoque europeo de inteligencia artificial*, en «Revista de Derecho Administrativo», 2021.
- GARDNER G., *Intelligence Reframed. Multiple Intelligences for the 21st Century*, Basic Books, New York, 2000.
- GIALUZ M., *Quando la giustizia penale incontra l'intelligenza artificiale: luci e ombre dei risk assessment tools tra Stati Uniti ed Europa*, en «Diritto penale contemporaneo», 29 de mayo de 2019.
- GIRALDI A., *Algorithms and Big Data Towards a Crime-Preventing Groupware*, en «Roma Tre Law Review», 2/2021.
- GIRALDI A., *Política, discrecionalidad y Derecho en las implicaciones empíricas del principio de proporcionalidad de la pena*, en «Revista Internacional de Doctrina y Jurisprudencia», 22/2020.
- GRACIA MARTÍN L., *Crítica de las modernas construcciones de una mal llamada responsabilidad penal de la persona jurídica*, en «Revista Electrónica de Ciencia Penal y Criminología», 2016.
- HALLEVY G., *The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control*, en «Akron Intellectual Property Journal», 2/2010.
- HASSEMER W., *¿Por qué castigar? Razones por las que merece la pena la pena*, Tirant lo Blanch, Valencia, 2016, traducción de M. Cancio Meliá y F. Muñoz Conde.
- HEFENDEHL R., HIRSCH A.V., WOHLERS W. (eds.), *La teoría del bien jurídico. ¿Fundamento de legitimación del Derecho penal o juego de abalorios dogmático?*, Marcial Pons, Madrid, 2016, traducción de R.A. Guirao, M. Martín Lorenzo, Í. Ortiz de Urbina Gimeno.

- JAKOBS G., *La autocomprensión de la ciencia del Derecho penal ante los desafíos del presente*, en *La ciencia del Derecho penal ante el nuevo milenio*, dirigido por A. Eser, W. Hassemmer, B. Burkhardt, Tirant lo Blanch, Valencia, 2004.
- JAKOBS G., *El principio de culpabilidad*, en «Anuario de Derecho Penal y Ciencias Penales», 3/1992.
- JAKOBS G., CANCIO MELIÁ M., *Derecho penal del enemigo*, Civitas, Madrid, 2003.
- JESCHECK H.H., *Evolución del concepto jurídico penal de culpabilidad en Alemania y Austria*, en «Revista Electrónica de Ciencia Penal y Criminología», 2003.
- KAPLAN A., HAENLEIN M., *Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence*, en «Business Horizons», 1/2019.
- KAPLAN J., *Artificial Intelligence. What Everyone Needs to Know*, Oxford University Press, Oxford, 2016.
- LAMPE E.J., *Das personale Unrecht*, Duncker & Humblot, Berlín, 1967.
- LIMA D., *Could AI Agents Be Held Criminally Liable? Artificial Intelligence and the Challenges for Criminal Law*, en «South Carolina Law Review», 3/2018.
- LLEDÓ BENITO I., *El impacto de la robótica. La inteligencia artificial y la responsabilidad penal en los robots inteligentes*, en «Foro galego. Revista xurídica», 2020.
- MANZINI V., *Trattato di diritto penale italiano*, vol. III, Utet, Torino, 1981.
- MARAS M.H., SHAPIRO L.R., *Child Sex Dolls and Robots: More Than Just an Uncanny Valley*, en «Journal of Internet Law», 2017.
- MARINI G., voz *Colpevolezza*, en «Digesto delle Discipline Penalistiche», II, 1988.
- MARINUCCI G., DOLCINI E., GATTA G.L., *Manuale di Diritto Penale. Parte Generale*, Giuffrè, Milán, 2018.
- MASSARO A., *Pericolosità sociale e misure di sicurezza detentive nel processo di "definitivo superamento" degli ospedali psichiatrici giudiziari: la lettura della Corte costituzionale con la sentenza n. 186 del 2015*, en «Archivio penale», 2/2015.
- MCCARTHY J., MINSKY M.L., ROCHESTER N., SHANNON C.E., *A proposal for the Dartmouth summer research project on artificial intelligence*, 31 de agosto de 1955, disponible en <<http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>> (consultado el 18 de noviembre de 2022).
- MENCARELLI A., *Le misure di sicurezza detentive al vaglio della Corte europea dei diritti dell'uomo*, en *La tutela della salute nei luoghi di detenzione. Un'indagine di diritto penale intorno a carcere, REMS e CPR*, dirigido por A. Massaro, RomaTrE-Press, Roma, 2017.

- MIR PUIG S., *Derecho penal. Parte general*, Editorial Reppertor, Barcelona, 2011.
- MIR PUIG S., *Función de la pena y teoría del delito en el Estado social y democrático de Derecho*, Bosch, Barcelona, 1982.
- MIR PUIG S., *Preterintencionalidad y límites del art. 50 del Código Penal*, en «Revista Jurídica de Cataluña», 1979.
- MORILLAS CUEVA L., *Sistema de Derecho Penal. Parte general*, Dykinson, Madrid, 2018.
- MORILLAS FERNÁNDEZ D.L., *Implicaciones de la inteligencia artificial en el ámbito del Derecho Penal*, en esta obra.
- MUÑOZ CONDE F., GARCÍA ARÁN M., *Derecho Penal. Parte General*, Tirant lo Blanch, Valencia, 2019.
- MUÑOZ CONDE F., *Introducción*, en C. ROXIN, *Política criminal y sistema del derecho penal*, Hammurabi, Buenos Aires, 2002, traducción de F. Muñoz Conde.
- MUÑOZ VELA J.M., *Inteligencia Artificial y responsabilidad penal*, en «Derecho Digital e Innovación. Digital Law and Innovation Review», 11/2022.
- MUSCO E., voz *Misure di sicurezza*, en «Enciclopedia del diritto», I, 1997.
- NAVARRETE URIETA J.M., *El nacimiento de la idea de culpabilidad*, en «Aldaba», 18/1992.
- NIETO MARTÍN A., *Problemas fundamentales del cumplimiento normativo en el derecho penal*, en *Compliance y teoría del Derecho penal*, dirigido por L. Kuhlen, J.P. Montiel, Í. Ortiz de Urbina Gimeno, Marcial Pons, Madrid, 2013.
- NOTARO L., *Intelligenza artificiale e giustizia penale*, en *Intelligenza artificiale e giustizia penale*, dirigido por A. Massaro, Paruzzo Editore, Caltanissetta, 2020.
- NUVOLONE P., voz *Misure di prevenzione e misure di sicurezza*, en «Enciclopedia del diritto», XXVI, 1976.
- OTTO H., *Personales Unrecht, Schuld und Strafe*, en «Zeitschrift für die gesamte Strafrechtswissenschaft», 3/1975.
- PÉREZ ARIAS J., *Algoritmos y big data en la responsabilidad penal: el reto de la cibercriminalidad en el Derecho Penal*, en esta obra.
- PÉREZ ARIAS J., *Sistema de atribución de responsabilidad penal a las personas jurídicas*, Dykinson, Madrid, 2014.
- PÉREZ MANZANO M., *Culpabilidad y prevención. Las teorías de la prevención general positiva en la fundamentación de la imputación subjetiva y de la pena*, Universidad Autónoma de Madrid, Madrid, 1990.

- PERIS RIERA J.M., *Delitos de peligro y sociedad de riesgo: una constante discusión en la dogmática penal de la última década*, en *Estudios penales en homenaje al profesor Cobo del Rosal*, dirigido por J.C. Carbonell Mateu, B. del Rosal Blasco, L. Morillas Cueva, E. Orts Berenguer, M. Quintanar Díez, Dykinson, Madrid, 2005.
- QUINTERO OLIVARES G., *Parte General del Derecho Penal. Adaptada al programa de ingreso en las carreras judicial y fiscal*, Aranzadi, Pamplona, 2015.
- QUINTERO OLIVARES G., *Las vicisitudes del dolo y la subsistencia de la preterintencionalidad*, en *Constitución, derechos fundamentales y sistema penal. Semblanzas y estudios con motivo del setenta aniversario del profesor Tomás Salvador Vives Antón*, dirigido por J.C. Carbonell Mateu, J.L. González Cussac, E. Orts Berenguer, Tirant lo Blanch, Valencia, 2009.
- RIONDATO S., *Robotica e diritto penale (robots, ibridi, chimere e "animali tecnologici")*, en *Genetics, robotics, law, punishment*, dirigido por D. Provolo, S. Riondato, F. Yenisey, Padova University Press, Padua, 2014.
- ROXIN C., *Culpabilidad y prevención en Derecho penal*, Editorial Reus, Madrid, 1981, traducción de F. Muñoz Conde.
- ROXIN C., *Derecho Penal. Parte General. Tomo I. Fundamentos. La estructura de la teoría del delito*, Civitas, Madrid, 1997, traducción de D.M. Luzón Peña, M. Díaz y García Conlledo, J. de Vicente Remesal.
- SANTAMARIA D., voz *Colpevolezza*, en «Enciclopedia del diritto», VII, 1960.
- SCHANK R.C., *What Is AI, Anyway?*, en «AI Magazine», 4/1987.
- SOLUM L.B., *Legal Personhood for Artificial Intelligences*, en «North Carolina Law Review», 4/1992.
- SPINOZA B., *Tratado político*, Alianza Editorial, Madrid, 1986, traducción de Atilano Domínguez.
- SQUILLACI E., *'Ragioni' e 'costi' dell'attuale paradigma di prevenzione*, en «Archivio penale web», 2/2020.
- TAGLIARINI F., voz *Pericolosità*, en «Enciclopedia del diritto», XXXIII, 1983.
- TORIO LÓPEZ Á., *El concepto individual de culpabilidad*, en «Anuario de Derecho Penal y Ciencias Penales», 2/1985.
- TORRAS GENÍS C., DELICADO MORATALA L., *Blanco y Frío como El Polo Norte y El Polo Sur: Seres humanos y robots*, en «Sociología y tecnociencia: Revista digital de sociología del sistema tecnocientífico», 1/2022.
- VALLINI A., *Antiche e nuove tensioni tra colpevolezza e diritto penale artificiale*, Giappichelli, Torino, 2003.
- VALLS PRIETO J., *Inteligencia artificial, Derechos humanos y bienes jurídicos*, Aranzadi, Pamplona, 2021.

VELÁZQUEZ FERNÁNDEZ H., *¿Qué tan natural es la inteligencia artificial? Sobre los límites y alcances de la biomímesis computacional*, en «Naturaleza y Libertad. Revista de estudios interdisciplinarios», 12/2019.

CAPÍTULO VI

Algoritmos y big data en la responsabilidad penal: el reto de la cibercriminalidad en el Derecho Penal

Jacinto Pérez-Arias

SUMARIO: 1. Introducción – 2. La informática y el Derecho Penal: la inteligencia artificial – 3. Límites normativos – 4. Problemas procesales ¿vs? problemas penales – 5. Sujeto activo, resultado y culpabilidad: juicio de probabilidad estadística – 6. Conclusión – 7. Bibliografía

1. *Introducción*

Es difícil encarar, en un trabajo de esta naturaleza, un problema con tantas aristas como la cibercriminalidad. El derecho penal, anclado en sus principios clásicos, necesita responder a nuevos retos que, a su vez, pueden exigir de él nuevos principios (si lo que se pretende es idear una respuesta global y satisfactoria al problema). Mientras tanto eso llegue, e inspirándonos en G. RADBRUCH¹, debemos ver cómo mejorar el derecho penal, hasta tener algo mejor que el derecho penal.

Ante todo, debemos valorar si la cibercriminalidad constituye un problema jurídico-penal o no y, si lo es, si tiene solución jurídico-penal o no. Y decimos un problema jurídico penal, pues no debemos confundir el problema sustantivo (material) con las dificultades procesales que puedan surgir en la investigación y enjuiciamiento de los delitos derivados de, o con causa en, la cibercriminalidad.

Como señala J.M. PERIS RIERA², ser conscientes de los grandes riesgos que entrañaría una aplicación mecánica de los avances en neurociencia y en inteligencia artificial, no debe comportar un abandono de esta fuente de conocimiento; no debe suponer en absoluto un aislamiento por parte del

¹ G. RADBRUCH, *Einführung in die Rechtswissenschaft*, Quelle & Meyer, Stuttgart, 1929, p. 11.

² J. PERIS RIERA, *Inteligencia Artificial y Neurociencias: Avances Del Derecho Penal Contemporáneo*, en esta obra, p. 3

derecho penal de los nuevos avances técnicos en estas materias. Más bien deben ser tenidos en cuenta para mejorar la técnica procesal y judicial, pero teniendo presente que todo lo técnicamente posible no tiene por qué ser jurídicamente admitido.

La multidisciplinar temática de la ciberdelincuencia afecta, también, a la criminología (la llamada cibercriminología, en término acuñado por JAISHANKAR³), donde se están haciendo avances importantes en la materia, sobre todo, en lo relacionado con el perfil del ciberdelincuente; y ello, aunque la propia ONU concluyera, en su XIII Congreso sobre Prevención del Delito y Justicia Penal de 2015, que no había un perfil estándar del ciberdelincuente⁴. No obstante, y como se ha matizado, esta poco halagüeña intuición nace de la propia naturaleza técnica, pero igualmente dinámica y constantemente innovadora de los denominados ciberdelitos: por un lado, su perpetración requiere de unos conocimientos cualificados en materia informática; por el otro, su versatilidad y el avance en la disponibilidad de las nuevas Tecnologías de la Información y la Comunicación (TIC's), hacen posible que, al igual que los interfaz informáticos necesarios para su comisión, la realización de actos delictivos a través de medios telemáticos se encuentre, cada vez más frecuentemente, al alcance de cualquiera que sepa manejar –aunque sea de forma rudimentaria o a nivel ‘usuario’– un dispositivo. Más aún, es posible que las cualificaciones técnicas poseídas por algunos sujetos sirvan de puerta de entrada a otros con conocimientos mucho menores, que se aprovechan del trabajo ya realizado por los primeros para cometer hechos delictivos a través de medios informáticos (S. CÁMARA ARROYO⁵).

Como se señala por J.M. TAMARIT SUMALLA⁶, la evolución hacia la sociedad de la información ha supuesto una transformación profunda de las relaciones sociales, y uno de los efectos de dicho proceso ha sido la transformación de la delincuencia. Buen ejemplo de ello es la evolución de la delincuencia juvenil, que siempre ha sido un fenómeno particularmente

³ Considerado el fundador de esta rama de la criminología, JAISHANKAR entiende la cibercriminología como una materia multidisciplinar que abarca diversos campos, tales como la Criminología, la Victimología, la Sociología, la Ciencia de Internet y las Ciencias de la computación (así lo recuerda S. CÁMARA ARROYO. *Estudios criminológicos contemporáneos (IX): La Cibercriminología y el perfil del ciberdelincuente*, en «Derecho y Cambio Social», Núm. 60, Abril-junio 2020, pp. 471-472.

⁴ Informe del 13º Congreso de las Naciones Unidas sobre Prevención del Delito y Justicia Penal, A/CONF.222/17. 29 de abril de 2015.

⁵ CÁMARA ARROYO. *Estudios criminológicos contemporáneos*, cit., pp. 482.

⁶ J.M. TAMARIT SUMALLA. *Ciberdelincuencia y cibervictimización*, en «Revista de los Estudios de Derecho y Ciencia Política», n.º 22, Junio, 2016, p. 30.

estudiado por la criminología y que puede funcionar como un punto de observación de las tendencias, de anticipación del futuro y de experimentación de nuevas formas de justicia y de tratamiento de la criminalidad.

En su día señalamos (J. PÉREZ-ARIAS⁷), que el sucesivo incremento de los inventos tecnológicos, y el casi preocupante (ab)uso de las nuevas tecnologías (hasta para las cosas más básicas), ha convertido lo cibernético en un nuevo mundo, con reglas nuevas y riesgos y/o peligros desconocidos. Solo ha de pensarse en los datos personales que cada individuo entrega a la red, y en la posibilidad de uso, por terceros anónimos, de tales datos, incluso sin el conocimiento directo ni indirecto de la víctima. Es evidente que este nuevo mundo (esa nueva modernidad a la que se refería U. BECK⁸) arroja un cúmulo de nuevos riesgos, que merecen, al menos teóricamente -no necesariamente en términos legislativos- un estudio y tratamiento especializado, que tenga en cuenta todas las variables que confluyen en este, cada vez más expandido, fenómeno social de redes (o, dicho de otro modo, en esta nueva sociedad cibernética, que se impone sobre la ya denominada sociedad tradicional o antigua).

2. La informática y el Derecho Penal: la inteligencia artificial

El Derecho Penal no queda extramuros de todo este nuevo mundo⁹. Por ello, se ha afirmado que la revolución social y técnica que implican las nuevas tecnologías traen consigo también efectos en lo que al Derecho penal respecta (R.M. MATA Y MARTÍN¹⁰).

La cibercriminalidad, como fenómeno técnico, exigiría una valoración exhaustiva de todas y cada una de las categorías de la teoría jurídica del delito¹¹, por lo que podemos extraer ya una primera conclusión: estamos ante

⁷ J. PÉREZ-ARIAS, *Cibercriminalidad: Hacia la nueva realidad -virtual- del derecho penal*, en «Revista internacional de doctrina y jurisprudencia», Volumen 26 (diciembre 2021), Universidad de Almería, p. 182.

⁸ U. BECK. *Sociedad de riesgo. Hacia una nueva modernidad*, Paidós, Barcelona-Buenos Aires-México, 2002.

⁹ PÉREZ-ARIAS, *Cibercriminalidad: Hacia la nueva realidad -virtual- del derecho penal*, cit., p. 176.

¹⁰ R.M. MATA Y MARTÍN, *Criminalidad informática: una introducción al cibercrimen*, en «Actualidad Penal», 37, Semana del 6 al 12 Oct. 2003, Ref. XXXVI, tomo 3, p. 935.

¹¹ Como señala PERIS RIERA, la responsabilidad penal que pueda nacer de condicionantes basados en inteligencia artificial debería ser individualizada a partir de los criterios generales inherentes a la teoría general del delito (PERIS RIERA, *Inteligencia Artificial y*

una cuestión compleja¹². Así lo entiende también A. GIRALDI¹³, para quien la inteligencia artificial se ha desarrollado hasta llegar a influir no sólo en la realidad social, sino también en las instituciones básicas del Derecho penal.

Ahora bien, su complejidad no puede confundirse ni con su desconocimiento (todo lo desconocido parece complejo), ni con el incómodo y habitual uso indiscriminado de términos ingleses (siempre de poca ayuda conceptual). Cuando lejos de divulgar se pretende construir una teoría válida, los anglicismos pierden importancia. Solo cuentan las bases teóricas y su respaldo dogmático. Y ahí es donde empieza el verdadero problema. Es hora, pues, de que la doctrina comience a construir el necesario *andamio para las ideas*, en palabras tan expresivas de A. MUÑOZ ALONSO¹⁴.

Se señala por F. MIRÓ LINARES¹⁵, y estamos por completo de acuerdo, que algunos términos generan una expectativa superior a su propia realidad, y cuando se compara lo que expresa el concepto con la concreción del mismo en la práctica, llega la pequeña o gran decepción.

Cada vez que se impone una tendencia, rápidamente surgen las dudas de si existe una adecuada regulación penal y un adecuado tratamiento teórico. Pasó con la lucha por el medioambiente, pasó con la responsabilidad penal de las personas jurídicas y obviamente tenía que pasar con la cibercriminalidad.

A nuestro entender, la cibercriminalidad comparte caracteres de estas dos temáticas penales: Del medio ambiente comparte su peligrosidad y la eterna cuestión de si estamos o no admitiendo conductas generadoras de riesgos no tolerables, lo que nos llevará a la ardua cuestión del peligro relevante; y de las personas jurídicas participa de la singularidad de un sujeto penal altamente discutible (en nuestro caso, la inteligencia artificial) y con clara falta de sintonía con el concepto clásico de autor.

Una vez surge la tendencia, el problema se pone de moda y alguna literatura jurídica o criminológica, y los medios divulgativos empiezan a crear teorías, en la mayoría de los casos sin base firme ni normativa. Estas

Neurociencias: Avances Del Derecho Penal Contemporáneo, en esta misma obra, p. 3).

¹² Como afirma L. HERNÁNDEZ DÍAZ, la primera dificultad a la hora de afrontar el análisis de los delitos informáticos es su conceptualización (L. HERNÁNDEZ DÍAZ, *El delito informático*, en «EGUZKILORE. Cuaderno del Instituto Vasco de Criminología», Número 23. San Sebastián, diciembre 2009, p. 228.

¹³ A. GIRALDI, *Deshumanizando la culpabilidad: los sistemas inanimados en la teoría del delito*, en esta obra.

¹⁴ A. MUÑOZ ALONSO, *Andamios para las ideas*. Colección Aula de ideas, volumen 1, Aula, Madrid. 1952.

¹⁵ F. MIRÓ LINARES, *El sistema penal ante la inteligencia artificial: Actitudes, usos, retos*, en *Cibercrimen III* (Dir. D. DUPUY/ J.G. CORVALÁN), B de f, Buenos Aires, 2020, p. 81.

teorías chocan con la realidad, y terminan generando folios de disertaciones (sin hablar de horas de televisión), sin apenas contar con una regulación cierta, seria y sistemática. Si tres palabras rectificadoras del legislador convierten bibliotecas enteras en basura (J. KICHMANN¹⁶), podemos imaginar qué es lo que ocurre cuando la literatura científica inicia su andadura de razón práctica mucho antes de que el legislador haya mostrado su derrotero político criminal.

Los dos principales problemas, a nuestro entender, son, de un lado, de sujeto (con todo lo que ello comporta desde perspectivas de capacidad consciente) y, de otro, de conducta. Pero adviértase que hablo de sujeto y no de autor, porque para admitir la autoría antes se ha debido depurar si existe o no un sujeto con aptitud para realizar hechos por algún concepto y con capacidad de culpabilidad. No se puede hablar en directo de autor, porque el autor no es una premisa teórica, sino una consecuencia práctica¹⁷. Hablar de autoría, sin tener claro esto, nos llevaría nuevamente a la creación artificial de un sujeto difícil de digerir, como ocurrió, mediante la reforma del código penal español operada a través de la Ley Orgánica 5/2010, con la persona jurídica.

Con todo lo relevante y novedoso que pudiera parece la temática, el problema de la cibercriminalidad no es un absoluto desconocido para el Derecho. Aparte de la normativa europea, que ha sido la gran precursora de toda la regulación española en la materia, a nivel dogmático encontramos a C.M. ROMEO CASABONA¹⁸, que a finales de la década de los 80, del siglo pasado, hablaba ya de la relación entre el poder informático y seguridad jurídica.

Sin embargo, lo informático ha dado paso al ciberespacio, entendido como aquella realidad espacio-virtual, que no tiene una localización física y que abarca los sistemas de información y comunicación contenidos en la Red (S. MORÁN BLANCO¹⁹). Hoy no solo preocupa el acceso ilícito a un terminal, que también, sino la existencia de un espacio virtual dinámico, en el que se genera biografía vital de cada uno, personal y con contenido y trascendencia económica. ¿Somos conocedores de que nuestra vida está expuesta a un espacio sin control y que los ataques pueden venir incluso

¹⁶ Frase que afirmó en una conferencia en la sociedad jurídica de Berlín celebrada en 1847, y publicada al año siguiente.

¹⁷ Sin olvidar su trascendencia teórica.

¹⁸ C.M. ROMEO CASABONA, *Poder informático y seguridad jurídica. la función tutelar del derecho penal ante las nuevas tecnologías de la información*, Fundesco, Madrid, 1988, pp. 42 y ss.

¹⁹ S. MORÁN BLANCO. *La ciberseguridad y el uso de las tecnologías de la información y la comunicación (tic) por el terrorismo*, en «Revista Española de Derecho Internacional», vol. 69/2, julio-diciembre 2017, p. 197.

sin tocar, en el momento de la acción, nuestro ordenador, nuestra tableta o nuestro móvil?: Esa es la otra cara del mundo de los datos (de la Big Data) y de la inteligencia artificial²⁰.

Y todavía existe más riesgo desde que esas aplicaciones -tan útiles²¹- tienen capacidad para activar micrófonos, cámaras y deambulan por nuestros datos más íntimos porque nosotros lo hemos consentido. Esto genera también un problema en la víctima, dado que, desde el punto de vista penal, en este tipo de delitos, la víctima es crucial, en la mayoría de los casos, para la posible comisión del delito. De hecho, lo que permitirá activar la conducta en su contra -insistimos que, en la mayoría de los casos, no en todos- será la 'confianza' que deposita en algo que tiene constantemente en sus propias manos (el teléfono móvil, por ejemplo). Nadie suele pensar que su intimidad está en riesgo. Pero lo relevante y peligroso de ese artefacto, como sabemos, no está entre las manos, está en la red, en esa 'nube'. El peligro, por tanto, no es lo tangible (que es lo que genera falsa confianza), sino lo intangible (esa autopista de datos, que encuentra su entrada y salida en el artefacto que uno tiene entre sus manos).

Como afirma la SAP de Barcelona²², la modalidad ciberdelictiva más habitual es la de los fraudes [engaños] en internet, que se ha convertido, con sus 192.375 casos en el 88,1% de total de los ciberdelitos. Coincidimos con J.R. AGUSTINA²³, cuando afirma que la relación entre ofensor y víctima, mediada por 'máscaras virtuales', facilita al ofensor el recurso a apariencias engañosas, técnicas de camuflaje y de manipulación, y potencian en la víctima una serie de déficits cognitivo-conductuales que incrementan notablemente los riesgos de victimización.

A nivel conceptual, la inteligencia artificial es la habilidad de una máquina de presentar las mismas capacidades que los seres humanos, como el razonamiento, el aprendizaje, la creatividad y la capacidad de

²⁰ Término acuñado por John McCarthy, en el año 1956 (véase para mayor profundidad M. CUMBRERAS, P. LÓPEZ, *¿Es necesario un marco ético para guiar el desarrollo y uso de la inteligencia artificial en las organizaciones?*, en *La robótica y la inteligencia artificial en la nueva era de la revolución industrial 4.0* (Coord. M.J. CRUZ BLANCA/I. LLEDÓ BENITO), Dykinson, Madrid, 2021. p. 450.

²¹ Utilidad no solo en lo personal, pues no debemos olvidar el evidente progreso que se produce en la ciencia forense, gracias a esta tecnología.

²² SAP Barcelona de 9 de julio de 2020 (ECLI:ES:APB:2020:7301), p. 3

²³ J.R. AGUSTINA. *Cibercriminalidad y perspectiva victimológica: un enfoque general explicativo de la cibervictimización*, en «Cuadernos de política criminal», 136, Mayo 2022, p. 159.

planear²⁴. Esta habilidad, como señala R. KURZWEIL²⁵ y recuerdan S. RUSSEL/P. NORVIG²⁶, es el arte de desarrollar máquinas con capacidad para realizar funciones que cuando son realizadas por personas requieren de inteligencia. Sin embargo, así como la inteligencia humana está limitada por el nacimiento, la inteligencia artificial no encuentra más límites que el estado de desarrollo de la ciencia; de tal forma que cuanto más avance el conocimiento sobre el procesamiento informático, más capacidad de desarrollo tendrá la máquina con inteligencia artificial. O, dicho con otras palabras, la inteligencia artificial es una fuente de riesgos que será mayor conforme el sistema inteligente adquiera la capacidad de reproducir la forma de funcionamiento del cerebro humano y goce de mayor autonomía en la toma de decisiones (J.M. PALMA HERRERA²⁷).

Y eso es así, porque la máquina solo requiere el enfoque racional, que implica una combinación de matemáticas e ingeniería²⁸. De hecho, el gran problema que genera actualmente la big data es que estamos en presencia de un número ingente de datos cuya administración y gestión no puede realizarse a través de las tradicionales bases de datos, sino a través de esa nueva tecnología en que consiste la inteligencia artificial²⁹.

No puede extrañar, pues, que la actual sociedad esté por completo digitalizada. La vida actual ha cedido por completo al avance de una técnica que tiene capacidad para abarcarlo prácticamente todo, y es lógico que en el mundo personal y en el mundo empresarial se quieran usar recursos que abaraten costes e incrementan resultados/beneficios.

La desinformación social es la clave que dificulta todo. Una desinformación que se observa en la facilidad de crear (y creer) hechos inexistentes (*fake news*), en la extrema confianza y 'falsa' necesidad de las redes sociales y en el comercio electrónico. Es más, esa desinformación es la gran aliada de la ciberdelincuencia, pues es la que explica el recelo de los ciudadanos frente a las instituciones (con potestad legislativa), cuando éstas quieren regular y poner límites normativos al ciberespacio; unos límites que

²⁴ Definición obtenida de <<https://www.europarl.europa.eu>>

²⁵ R. KURZWEIL. *The Age of Intelligent Machines*, MIT Press, Cambridge, 1990.

²⁶ S. RUSSEL, P. NORVIG, *Inteligencia artificial. Un enfoque moderno*, Pearson, Madrid, 2008, p. 2.

²⁷ J.M. PALMA HERRERA. *Inteligencia artificial y neurociencia. Algunas reflexiones sobre las aportaciones que pueden hacer al Derecho Penal*, en esta obra.

²⁸ RUSSEL, NORVIG. *Inteligencia artificial. Un enfoque moderno*, cit., p. 2.

²⁹ Para mayor profundidad, vid C. MATÉ JIMÉNEZ. *Big data. Un nuevo paradigma de análisis de datos*. *anales de mecánica y electricidad*, en <<https://www.iit.comillas.edu>>, noviembre-diciembre 2014, pp. 10-16.

se perciben por la sociedad como una limitación de derechos y no como una barrera de protección. Y lo más preocupante es que, apenas están saliendo a la luz los primeros riesgos, la sociedad ya considera, paradójicamente, que limitar las redes supondría una intolerable limitación de derechos.

Ningún Estado, por ejemplo, quiere quedar al margen del comercio electrónico, pero esto no garantiza que el riesgo sea soportable en términos jurídicos. Si P. BONFANTE³⁰ o S. PEROZZI³¹ representaban al Estado y a la familia en los mismos términos, pero a diferente escala, quizás sea la hora de entender que la realidad virtual está hecha a escala de la vida real (física), y que nunca concebiríamos la vida social real sin límites o respeto al derecho, o libertad, ajeno. Esta necesidad de la red, esta desinformación y este uso y abuso de lo tecnológico es, lógicamente, aprovechada por los nuevos ciberdelincuentes, porque es en la red donde están hoy en día los datos, la intimidación, el valor económico, los intereses, etc. De ahí que se sostenga que la inteligencia artificial juega un papel nuclear en el devenir cotidiano de la ciudadanía con las consiguientes implicaciones jurídicas, entre otras muchas (D.L. MORILLAS FERNÁNDEZ³²).

Que la ciberactividad es un fenómeno de especial trascendencia e importancia lo evidencia la propia ONU cuando, en año 2015, afirmó que uno de los principales elementos impulsores de la ciberdelincuencia contemporánea y del uso creciente de pruebas digitales es el desarrollo de la conectividad electrónica global. Hoy existen casi 3.000 millones de usuarios de Internet, cerca del 40% de la población mundial³³. No debe extrañar, por ello, que se haya mantenido, a nivel jurisprudencial, que lamentablemente la ciberdelincuencia sigue su auge exponencial³⁴. Como resalta J.M. PALMA HERRERA³⁵, la sociedad del siglo XXI es la sociedad del ocio y del consumo, pero es también cada vez más, una sociedad que se siente amenazada por la aparición de nuevas fuentes de peligro asociadas,

³⁰ P. BONFANTE, *Corso di diritto romano*, Giuffrè, Milano, 1963.

³¹ S. PEROZZI, *Istituzioni di Diritto Romano*, F. Vallardi, Milano, 1947, pp. 311 y ss.

³² D.L. MORILLAS FERNÁNDEZ. *Implicaciones de la inteligencia artificial en el ámbito del Derecho Penal*, en esta obra.

³³ ONU. (A/CONF.222/12) *13º Congreso de las Naciones Unidas sobre Prevención del Delito y Justicia Penal. Seminario 3: El fortalecimiento de las respuestas de prevención del delito y justicia pena frente a las formas de delincuencia en evolución, como la ciberdelincuencia y el tráfico de bienes culturales, incluidas las lecciones aprendidas y la cooperación internacional*, celebrado el 2 de febrero de 2015, p. 7.

³⁴ SAP Barcelona de 9 de julio de 2020 (ECLI:ES:APB:2020:7301), p. 3.

³⁵ J.M. PALMA HERRERA, *Inteligencia artificial y neurociencia. Algunas reflexiones sobre las aportaciones que pueden hacer al Derecho Penal*, en esta obra.

paradójicamente, a ese progreso y bienestar.

La ciberseguridad se está configurando como un derecho de los ciudadanos y un deber de los Estados de garantizar el libre ejercicio de los derechos fundamentales y libertades públicas en la red, promoviendo medios para la seguridad e integridad de las infraestructuras y la información. En el caso de los robots inteligentes existe una interconexión continua e instantánea entre el mundo físico y digital y se deberá velar simultáneamente por la seguridad en ambos entornos. Los robots inteligentes, como sistemas físicos cibernéticos, no están exentos de sufrir ataques y las principales amenazas a las que se enfrentan son el malware, el ciber espionaje, denegaciones de servicio, pérdida de información, spam, phishing o daño físico entre otros (I. LLEDÓ BENITO³⁶).

3. *Límites normativos*

Lo que no está tan claro es qué institución (política) puede poner barreras normativas a la ciberactividad, máxime cuando estos límites son percibidos como restricción de falsos derechos por el electorado del que dependen. Un electorado que, a su vez, y para evidenciar una circularidad en bucle, será manipulado con opiniones, a favor y en contra, que serán generadas a través de esas mismas redes sociales que pretenden ser reguladas. Un ejemplo lo vemos a diario en Twitter; o en las cámaras de seguridad ciudadana: Todos queremos la seguridad que nos brindan, al tiempo que nos consideramos espionados. Es preciso llegar a ese punto de equilibrio aristotélico como medida de la virtud: cualquier solución implica, paradójicamente, una restricción, un nuevo problema.

Han sido varios los intentos de regular aspectos concretos de la ciberdelincuencia, sobre todo de aquella parte que afecta a la seguridad y protección de los datos que es, al fin y al cabo, el origen de toda actividad ciberdelictiva. Como en su día señalamos (J. PÉREZ-ARIAS³⁷), las siguientes normas internacionales resultan de interés:

- 1.- Como primera norma supranacional, dentro del ámbito territorial europeo, debemos partir de la ya clásica Directiva 95/46/CE, de 24 de

³⁶ I. LLEDÓ BENITO, *Ciberseguridad Versus Ciberdelincuencia*, en *El derecho penal, robots, IA y cibercriminalidad: desafíos éticos y jurídicos. ¿Hacia una distopía?*, Dykinson, Madrid, 2022, p. 18.

³⁷ PÉREZ-ARIAS, *Cibercriminalidad: Hacia la nueva realidad -virtual- del derecho penal*, cit., p. 178-179.

octubre de 1995; norma que, actualmente, se encuentra derogada por el Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016. Esta norma inició la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos. Se trata de una norma que, dentro del ámbito que estudiamos, es secundaria e indirecta, ya que su objeto no es, como resulta evidente (dado el año de su aprobación), la realidad cibernética y, sobre todo, el contexto de protección de datos que circulan en la red, tal y como lo conocemos hoy en día³⁸.

Sin embargo, en esta Directiva 95/46/CE se procuró tutelar los datos personales almacenados en bases de datos de numerosos organismos públicos y privados, cuya nueva trascendencia permite su aplicación para un sinnúmero de fines diversos, que van desde la seguridad pública y la defensa del Estado (v.gr. registros de naturaleza penal) hasta la ponderación de riesgos en la concesión de préstamos o servicios (G. EDUARDO ABOSO³⁹).

2.- El llamado Convenio de Budapest (Convenio sobre la ciberdelincuencia) de 23 de noviembre de 2001, cuyo objetivo, tal y como expone su preámbulo, es satisfacer la necesidad de aplicar, con carácter prioritario, una política penal común con objeto de proteger a la sociedad frente a la ciberdelincuencia, en particular mediante la adopción de una legislación adecuada y la mejora de la cooperación internacional. Y todo ello a partir de los profundos cambios provocados por la digitalización, la convergencia y la globalización continuas de redes informáticas. En definitiva, el Convenio sobre Ciberdelincuencia, firmado en Budapest el 23 de noviembre de 2001, supone la respuesta a la necesidad de tener medios eficaces de cooperación para la lucha contra la cibercriminalidad (M.C RAYÓN BALLESTEROS/J.A. GÓMEZ HERNÁNDEZ⁴⁰)

Es de destacar la capacidad de antelación que tuvo este convenio, que ya en el año 2001, cuando las redes aún se encontraban en un momento muy incipiente (respecto, al menos, de consumidores particulares y finales), se percató del peligro de las redes informáticas y la información

³⁸ Puede profundizarse más en A. PALMA ORTIGOSA, *Decisiones automatizadas y protección de datos*, Dykinson, Madrid, 2022.

³⁹ G. EDUARDO ABOSO, *Derecho Penal Cibernético. La cibercriminalidad y el Derecho Penal en la moderna sociedad de la información y la tecnología de la comunicación*, B de f, Buenos Aires, 2017, p. 59.

⁴⁰ M.C. RAYÓN BALLESTEROS, J.A. GÓMEZ HERNÁNDEZ, *Cibercrimen: particularidades en su investigación y enjuiciamiento*, en «Anuario jurídico y económico», Número 47, enero 2014, p. 212.

electrónica cuando éstas fueran utilizadas igualmente para cometer delitos y del riesgo de que las pruebas relativas a dichos delitos fueran almacenadas y transmitidas por las redes (así se expresa el preámbulo del convenio). En aquel año ya se hablaba de la necesidad de que los Estados firmantes (España lo ratificó, no obstante, en el año 2010⁴¹) incorporaran en la regulación sustantiva penales determinados delitos que, poco a poco, y en años posteriores, fueron quedando integrados en el código penal español, a partir de sus diversas reformas⁴². Estas figuras eran los delitos contra la confidencialidad, la integridad y la disponibilidad de los datos y sistemas informáticos (artículos 3 a 6 del convenio); delitos informáticos, propiamente dichos (artículos 7 a 8); delitos relacionados con el contenido (en materia de pornografía infantil, art. 9); delitos relacionados con infracciones de la propiedad intelectual y de los derechos afines (art. 10), así como otras disposiciones en materia de Derecho Penal general.

- 3.- La Decisión Marco 2005/222/JAI del Consejo de 24 de febrero de 2005 relativa a los ataques contra los sistemas de información. La etiología de la norma la plasma su misma exposición de motivos, al indicar que se ha comprobado la existencia de ataques contra los sistemas de información, en particular como consecuencia de la amenaza de la delincuencia organizada, y crece la inquietud ante la posibilidad de ataques terroristas contra sistemas de información que forman parte de las infraestructuras vitales de los Estados miembros. Esto pone en peligro la realización de una sociedad de la información segura y de un espacio de libertad, seguridad y justicia, y por tanto exige una respuesta por parte de la Unión Europea.
- 4.- Lo sorprendente es que el legislador español, en el intento de estar a la 'vanguardia', transpone esa decisión marco en el mismo año que Europa había propuesto una Directiva²¹ para derogarla. No es de extrañar, puesto que la decisión marco era de 2005 y la reforma del código penal

⁴¹ Véase Instrumento de Ratificación del Convenio sobre la Ciberdelincuencia, hecho en Budapest el 23 de noviembre de 2001 (BOE núm. 226, de 17 de septiembre de 2010). Dicho Convenio entró en vigor para España el 1 de octubre de 2010, de conformidad con lo establecido en su artículo 36.4.

⁴² Como afirma N.J. DE LA MATA BARRANCO, las transposiciones europeas se vienen haciendo en nuestro Código Penal desde el respeto a la sistemática tradicional del mismo, en su caso con el habitual recurso a los títulos, capítulos o artículos bis, ter, etc. (N.J. DE LA MATA BARRANCO, *Reflexiones sobre el bien jurídico a proteger en el delito de acceso informático ilícito (art. 97 CP). El concepto de de privacidad informática y la tutela del buen funcionamiento de los sistemas de información y comunicación*, en «Cuadernos de política criminal», segunda época, Núm. 118, mayo 2016, p. 55.

tuvo lugar cinco años después. Afortunadamente para el legislador español, esta Directiva no fue aprobada hasta el año 2013, mediante Directiva 2013/40/UE del Parlamento Europeo y del Consejo de 12 de agosto de 2013, relativa a los ataques contra los sistemas de información y por la que se sustituye la Decisión marco 2005/222/JAI. No será, pues, hasta la reforma del código penal del 2015 cuando el legislador español traspusiera la Directiva 2013/40/UE²². La misión de esta reforma fue la de pretender superar las limitaciones de la regulación vigente para ofrecer respuesta a la delincuencia informática en el sentido de la normativa europea (preámbulo, punto XIII, LO 1/2015).

El trasfondo de toda esta normativa internacional trata de conocer la seguridad de los sistemas de información en sí mismos considerados y, en su caso, la confidencialidad de los datos (y su integridad y disponibilidad frente a conductas de daños) que, por una parte, permiten el correcto funcionamiento de tales sistemas y, por otra, garantizan el funcionamiento de una sociedad cada vez más vinculada a dichos sistemas (N.J. DE LA MATA⁴³). Que la cuestión no es sencilla lo demuestra que, a día de hoy, no exista en el derecho penal español un delito informático en sentido estricto y tampoco un concepto penal de cibercriminalidad o ciberdelito, y ese es el punto preocupante del que partimos. En efecto, el Código Penal español, a fuerza de tratado internacional, ha ido introduciendo, poco a poco, delito a delito (allá donde fuera necesario) determinadas modalidades delictivas, que incluían lo informático como medio comisivo autónomo y específico, pero en ningún caso ha denominado a alguno de ellos delito informático, ni mucho menos ha creado un título autónomo a este fenómeno criminal (J. PÉREZ-ARIAS⁴⁴). La falta de una regulación específica evidencia la complejidad del tratamiento jurídico de este fenómeno criminal. Como señala I. BENÍTEZ ORTUZAR⁴⁵, ni existe ni es fácil que se consensue un concepto unitario perfectamente definido de 'delito informático' que incluya todas las modalidades delictivas que tienen en los datos incorporados en los sistemas informáticos o en los propios

⁴³ MATA BARRANCO. *Reflexiones sobre el bien jurídico a proteger en el delito de acceso informático ilícito (art. 97 CP). El concepto de de privacidad informática y la tutela del buen funcionamiento de los sistemas de información y comunicación*, cit., p. 57.

⁴⁴ PÉREZ-ARIAS, *Cibercriminalidad: Hacia la nueva realidad -virtual- del derecho penal*, cit., p. 176.

⁴⁵ I. BENÍTEZ ORTUZAR, *Ciberdelitos: la implementación en el ordenamiento interno de los acuerdos internacionales en materia de ciberdelincuencia*, en *La robótica y la inteligencia artificial en la nueva era de la revolución industrial 4.0* (Coord. M.J. CRUZ BLANCA, I. LLEDÓ BENITO), Dykinson, Madrid, 2021, p. 81.

sistemas informáticos el objeto o el instrumento del delito.

Con todo, y como se ha señalado por la jurisprudencia⁴⁶, los ciberdelitos -aquellos que ya tienen respaldo normativo- deberían establecer también penas específicas (como la inhabilitación especial) y, sin embargo, ninguna se establece específicamente relacionada con las TIC. De hecho, si prestamos atención a las figuras delictivas más próximas a la cibercriminalidad, observamos que las penas previstas son las tradicionales penas de prisión y multa. En definitiva, hay todavía mucha distancia que recorrer, porque son muchas las carencias e imprecisiones existentes en la normal penal.

4. *Problemas procesales ¿vs? problemas penales*

El debate debe girar en torno a dos rasgos que se han de conciliar, y que resulta difícil que lo hagan: Riesgo y resultado. En ambos casos, nos enfrentamos con inconvenientes de índole procesal y penal. Respecto de los problemas procesales, nos encontramos con dos factores esenciales a tener en cuenta: la desterritorialización y el anonimato. Problemas éstos que, a la postre, también generarán disfunciones en el derecho penal sustantivo, como es el derecho aplicable⁴⁷ o la identificación última del

⁴⁶ Entre otras, puede verse la SAP Barcelona, de 6 de mayo de 2015 (ECLI: ES:APB:2015:2946), p. 5. Como sigue indicando esta Sentencia, el objetivo de la imposición de prohibiciones de uso de las TIC en los EE.UU. no es apoyar la rehabilitación, ni tan solo incapacitar completamente y constantemente al sujeto, sino limitar simplemente las oportunidades delictivas y, sobre todo, introducir una infracción técnica susceptible de ser detectada a medio plazo y que justifique la revocación de la medida y el retorno a la prisión. Por lo contrario, la situación en nuestro país es la opuesta. Como se ha visto, nuestro marco legal continúa orientado hacia un modelo rehabilitador, en el cual las medidas de control cumplen sobre todo una función auxiliar del tratamiento y casi no están reguladas, de manera que sería muy cuestionable un sistema de control del de las condiciones tan agresivo como el norteamericano. Por otra parte, a pesar de algunas mejoras en los últimos años, en la práctica nuestro aparato de ejecución penal en el entorno comunitario es muy débil, a duras penas puede hacer frente a las necesidades más urgentes y no está en condiciones de llevar a cabo un control intensivo.

⁴⁷ En igual sentido, cuando afirma que a efectos procesales, hay que matizar que la conducta delictiva puede tener su origen en uno o varios países y los resultados producirse en otro u otros, incluso puede resultar difícil determinar dónde se ha cometido la acción o por parte de quién. Obviamente esto afecta a la competencia jurisdiccional, a la ley penal aplicable y al procedimiento que se tramitará para su investigación y enjuiciamiento, ya que la regla general tradicional se refiere al lugar de comisión del delito o *locus commissi delicti* (principio de territorialidad) contenido en la Ley Orgánica del Poder

autor. La ciberdelincuencia, pues, representa una amenaza global, técnica, transfronteriza y anónima, y comprensiva de cualquier tipo de actividad ilegal (G. MARTÍNEZ ATIENZA⁴⁸).

Si se quiere llegar a una categorización del cibercrimen es importante asimilar las características que posee ese nuevo fenómeno (J. PÉREZ-ARIAS⁴⁹). Entre todas las características que podrían desarrollarse sobre las tecnologías de la información y la comunicación (TIC), nos concentraremos en tres: 1. la inmediatez de las comunicaciones a distancia, 2. la posibilidad de la realización de acciones masivas (automatizadas o no) y 3. la posibilidad de realizar acciones con un determinado nivel de anonimato (M. TEMPERINI⁵⁰).

En cuanto a la desterritorialización, la red cuenta con un importante rasgo característico: se trata de un espacio sin fronteras. Es la verdadera globalización. Esto hace que muchos de los ataques sean cometidos desde servidores o ubicaciones extranjeras, que impiden, por propia jurisdicción, avanzar (de un modo fructífero) en una investigación policial y/o judicial. Se nos podrá indicar, con acierto, que una correcta interpretación del artículo 23 de la Ley Orgánica del Poder Judicial permitiría investigar un hecho delictivo con estos caracteres transfronterizos. Sin embargo, a nivel práctico (no teórico ni hermenéutico), esa investigación no fructificaría en un resultado apto para su enjuiciamiento. Pensemos que un país extranjero (sobre todo extracomunitario) no facilitará, en muchos casos, datos o identidades de sus propios nacionales, sobre todo cuando estemos hablando de ciertos países, que se han convertido en un auténtico territorio offshore (por referencia analógica al paraíso fiscal) de la ciberdelincuencia. Todos tenemos, al menos, uno o dos países en mente.

Incluso delinquiendo desde España, la ingeniería informática y el sistema de redes permite simular una ubicación falsa gracias al empleo de un proxy encadenado o una VPN. Por tanto, los jueces españoles van a encontrar serios inconvenientes en su labor judicial, cuando una investigación se tiene que realizar, en su totalidad (no en algún aspecto concreto, que sí es frecuente) con diligencias para las que se necesitan el completo auxilio y/o cooperación de terceros países. Y, aun cuando eso se produzca, luego

Judicial. (RAYÓN BALLESTEROS, GÓMEZ HERNÁNDEZ, *Cibercrimen: particularidades en su investigación y enjuiciamiento*, cit., p. 216).

⁴⁸ G. MARTÍNEZ ATIENZA, *El blanqueo de capitales y su sanción administrativa y penal*. «Revista de Derecho vLex», número 202, Marzo 2021, p. 1.

⁴⁹ *Ibidem*, p. 191.

⁵⁰ M. TEMPERINI. *Delitos informáticos y cibercrimen: Alcances, conceptos y características*, en *Cibercrimen y delitos informáticos. Los nuevos tipos penales en la era de internet*, Erreius, Buenos Aires, 2018, p. 52.

sería complicado abrir juicio oral, porque se requiere presencialidad de los acusados. En la actualidad resulta poco frecuente encontrar jurisprudencia que analice la cuestión de la inteligencia artificial en el sentido que aquí analizamos (más allá de tratarla como medio comisivo específico en algunas figuras delictivas), y eso es la prueba evidente de que resulta complicado que un asunto de esta naturaleza llegue a los tribunales. De ahí que, en palabras de D.L. MORILLAS FERNÁNDEZ⁵¹, la descripción de la inteligencia artificial en la jurisprudencia penal española no cuente con una excesiva representatividad si nos atenemos a su descripción taxativa.

El anonimato es el otro gran inconveniente que, desde perspectivas procesales (luego veremos que también de derecho penal sustantivo), nos debe llamar la atención. En efecto, en la red sólo existe un sistema de identificación: la dirección IP (Internet Protocol). Como se tiene afirmado, cuando un terminal se conecta a Internet es necesario que disponga de un número que lo identifique de forma unívoca, que lo diferencie de cualquier otro ordenador en la red y permita localizarlo, de igual forma que las direcciones postales distinguen cada calle y cada casa de cualquier otra en el mundo. Cuando un usuario solicita o envía una información, se genera una carta o paquete en el que se escribe tanto la dirección destino como el remitente (L. SALVADOR CARRASCO⁵²).

Pero imaginemos este protocolo a nivel o escala mundial y con toda la tecnología disponible, y cuyos servicios de anonimización permitirán convertir en papel mojado todos los protocolos de identificación. De ahí que se señale que uno de los mayores escollos en materia de responsabilidad penal en el ámbito de la comisión de los llamados ciberdelitos se vincula con el papel que cumplen las empresas proveedoras de servicio en Internet por los delitos cometidos por terceros (G. EDUARDO ABOSO⁵³). La anonimización, junto a la desterritorialización, son los auténticos obstáculos procesales de difícil superación, ya que no puede haber proceso sin tribunal con jurisdicción (territorial, como presupuesto del proceso⁵⁴) ni sin persona a quien investigar, imputar y finalmente acusar.

Por lo que hace al derecho penal, el principal problema lo tenemos en

⁵¹ MORILLAS FERNÁNDEZ, *Implicaciones de la inteligencia artificial en el ámbito del Derecho Penal*, en esta obra. p. 23.

⁵² L. SALVADOR CARRASCO, *Redes de anonimización en internet: cómo funcionan y cuáles son sus límites*, en «Revista de Instituto Español de Estudios Estratégicos», número 16/12. 2012, p. 3.

⁵³ EDUARDO ABOSO. *Derecho Penal Cibernético. La cibercriminalidad y el Derecho Penal en la moderna sociedad de la información y la tecnología de la comunicación*, cit., p. 425.

⁵⁴ V. GIMENO SENDRA, *Derecho penal procesal*, Civitas-Thomson Reuters, Navarra, 2012, pp. 160-161.

la difícil -si no imposible⁵⁵- conexión causal entre la conducta del autor y el algoritmo en el que finalmente se desencadena la conducta ideada. El problema penal es peliagudo, porque todo se hace partir de un algoritmo que aprende en la red, y que no solo depende del autor intelectual (persona física), sino del desarrollo autónomo que aquella programación generará una vez entre en la red. Es la llamada inteligencia artificial (y que tanta utilidad ofrece cuando es Alexa, Siri o nuestro teléfono quien la tiene). Estamos, como se ha dicho, ante una misma tecnología que facilita la comunicación pero que puede explotarse con fines terroristas y delincuenciales (S. MORÁN BLANCO⁵⁶). Al ser difícil conectar causalmente el resultado con la conducta de un sujeto (porque intervienen muchas concausas), será fácil imaginar que también resultará poco fácil individualizar al sujeto primario que hay detrás de esa inteligencia artificial. De ahí la importancia de delimitar y enfocar correctamente el problema informático en su relación con el derecho penal⁵⁷.

Posiblemente, sea más necesario formular preguntas que improvisar respuestas. Preguntas, cuyas respuestas futuras, pueden generar algo de luz a este fenómeno criminal. Desde el punto de vista penal, ¿estamos tan solo ante un mero medio comisivo más (delito informático) o ante una realidad nueva, con tratamiento jurídico diferenciado (sujetos (v.g. *bots*), bienes jurídicos afectados (v.g. seguridad ciberespacial), conductas, víctimas, finalidades etc.)? Es evidente que aquí hay dos tendencias claramente: Aquella que podríamos tildar de regla general, que considera que el fenómeno informático no es más que un medio comisivo, y que los delitos afectados siguen siendo los tradicionales, sin que el medio comisivo cambie o afecte el tratamiento penal tradicional de la figura; y de otro lado, una nueva tendencia que considera que estamos ante una nueva realidad, esto es, ante nuevos delitos, no tratados específicamente (aunque pudieran encajar de manera alambicada en alguna modalidad concreta) en el código penal. Es el caso, por ejemplo, del secuestro de datos (*ransomware*) o del falso antivirus (*Scareware*). Decidir si estamos ante una tendencia u otra, haría variar no solo el estudio y la sistemática del fenómeno, sino también la técnica del legislador (J. PEREZ ARIAS⁵⁸).

Cuanto más queramos abstraer el fenómeno cibernético, mayor

⁵⁵ Así lo será en numerosos casos.

⁵⁶ MORÁN BLANCO, *La ciberseguridad y el uso de las tecnologías de la información y la comunicación (tic) por el terrorismo*, cit., p. 197.

⁵⁷ Muy en la línea de como se hizo, en su día, con el delito de tráfico de drogas.

⁵⁸ PÉREZ-ARIAS, *Cibercriminalidad: Hacia la nueva realidad -virtual- del derecho penal*, cit., p. 192.

problema habrá para legislar sobre ello. Si, en cambio, lo informático (aparte su naturaleza compleja) se considera, a estos efectos, un mero medio comisivo⁵⁹, obviamente su tratamiento jurídico no pasará de ser un añadido descriptivo al tipo que proceda. Este es el modo en el que, hoy en día, se regulan los delitos relacionados con el ciberespacio en el código penal español.

Pero es que, además, al derecho penal no le puede resultar indiferente el anonimato al que antes hacíamos referencia. En efecto, desde perspectivas penales, no hay posibilidad de delito ni de consecuencias penales sin sujeto. Y es evidente que, en multitud de casos, el anonimato no va a permitir imputar objetivamente la conducta a persona determinada ni determinable. No estamos hablando de la culpabilidad o no del sujeto, sino de algo previo, de la existencia o no de sujeto a quien atribuir un resultado que se presenta como delictivo. Sin lugar a dudas, desde perspectivas prácticas, es el proceso quien debe individualizar el sujeto responsable; pero desde una visión técnica penal y técnica legislativa, se debe analizar si la conducta dañosa es el producto o resultado de una persona o de un conjunto algorítmico que aprende de manera autónoma en el ciberespacio, y que ya no es controlado por ningún sujeto. Sobre esto profundizamos en el siguiente apartado.

5. Sujeto activo, resultado y culpabilidad: juicio de probabilidad estadística

La consecuencia de este aprendizaje autónomo (característico de la inteligencia artificial) es que el resultado final (aquel que culmina con la lesión concreta del bien jurídico) puede no haber sido, ni siquiera, imaginado con precisión por el autor, más que a título de juicio remoto de probabilidad. Por tanto, el autor quiere, pero solo puede esperar, lo que nos lleva a una forma de culpabilidad dolosa híbrida o de tipo mixto, y con esto, a otro problema: La contingente voluntad⁶⁰ haría inexplicable el uso

⁵⁹ La problemática de si lo informático es un medio comisivo o si supone un nuevo tipo de delito también se lo plantean S. GARAT, J. REALE, *La reforma penal en materia de cibercrimen en la República Argentina*. en *Cibercrimen II* (Dir. D. DUPUY), B de f, Buenos Aires, 2018, p. 506.

⁶⁰ Voluntad entendida como finalidad perseguida y buscada de propósito. Es contingente, porque puede estar presente, o puede no estarlo, por lo que siempre quedará el margen de duda de hasta qué punto el comportamiento del sujeto puede imputarse a título de dolo, aunque sea dolo eventual.

directo del delito doloso, si se contempla, incluso como mera hipótesis, el resultado como fruto del puro accidente. En efecto, en la mayoría de las ocasiones, la forma de culpabilidad característica del autor (desde una perspectiva empírica y no teórica) será la del dolo eventual, lo que, si bien permite hablar de la consciencia de la acción, anticipa la fina línea que separa la mera culpa o imprudencia consciente de aquel tipo de dolo.

El problema no es nada novedoso en el derecho penal, aunque es cierto que hoy se incluye un nuevo elemento a la discusión: No se trata solo de estar ante un sujeto que se representa el resultado como posible o probable, sino ante un sujeto que, aunque quisiera tener certeza de que el resultado se va a producir, no tiene capacidad de valorar conscientemente si quiera su causación probable. Es obvio que nos referimos a la ciberdelincuencia más profunda (la que de verdad es preocupante), aquella que no consiste en un mero ataque de terminal a terminal de fácil averiguación causal. No debe olvidarse que, en todos estos casos, la conducta desplegada se produce mediante lenguaje binario sobre datos anónimos al azar, no mediante conductas humanas sobre sujetos conocidos. Es el llamado *aprendizaje no supervisado con capacidad de autoorganización* (J.G. CORVALÁN⁶¹).

No se trata de imputar el resultado a bulto, sino de hacerlo con absoluto respeto a una causalidad adecuada y relevante, que no solo debe existir, sino ser la prueba de la existencia de una conducta clara, causalmente directa y determinante del resultado. Pese a esto, se mantiene que los algoritmos trabajan de forma superficial (no pueden abstraer o valorar en conjunto) y sin basarse en relaciones causales (solo en correlaciones significativas)⁶². No obstante, si la causalidad no es posible definirla con exactitud, o el resultado es obra de la previsibilidad tan solo, no es posible hablar de conducta dolosa en sentido estricto; y solo en el segundo caso sería posible imputarla a título imprudente o, en el mejor de los casos, como dolo eventual, dependiendo del grado de probabilidad con la que se representó el resultado el autor (con lo que ello supone, además, cuando se ha de decidir imputar un delito doloso, en el que el margen de exactitud de ese juicio de probabilidad se representa impreciso siempre). Si atendemos al resultado, en definitiva, se tendría que castigar la conducta como imprudente (de existir en esa figura concreta, ex artículo 12 del código penal español), aun cuando el origen de

⁶¹ J.G. CORVALÁN. *Presentación: Inteligencia artificial. Automatización y predicciones en el Derecho*, en *Ciberdelincuencia III* (Dir. D. DUPUY/J.G. CORVALÁN), B de f, Buenos Aires, 2020, p. 31.

⁶² C. SOUZA DE MENEZES, J.R. AGUSTINA. *Big Data, inteligencia artificial y policía predictiva*, en *Ciberdelincuencia III* (Dir. D. DUPUY/J.G. CORVALÁN), B de f, Buenos Aires, 2020, p. 160.

la acción sea claramente doloso. Imputar un delito doloso, en estos casos, sería una clara interpretación contra reo. Y ello, como decíamos arriba, en el supuesto remoto de que ese resultado concreto pudiéramos imputarlo a un sujeto concreto, que no será muy frecuente.

Y esto es una nueva disfunción sustantiva, pues nadie duda del carácter doloso de la conducta, pero la garantía del derecho penal no debe ceder a la utilidad social que reporta el castigo; castigo que, de producirse en estas condiciones, iría más allá de la culpabilidad⁶³, lo que se encuentra no solo en contra de la norma constitucional⁶⁴ sino de los artículos 5⁶⁵ y 10⁶⁶ del código penal español. Dicho de otro modo, aun pudiendo demostrar y valorar el resultado (con su correspondiente lesión del bien jurídico), como consecuencia causal de la conducta primaria del sujeto, quedaría por definir a título de qué forma de culpabilidad puede serle reprochado el delito. Y no es que se mezclen categorías dogmáticas (la causalidad es propia del tipo de injusto y el dolo⁶⁷ y la imprudencia de la culpabilidad); pero es evidente que cuando todo se difumina (y así ocurre con esta forma de criminalidad anónima), al final todo se mezcla y confunde. Sin hecho criminal imputable objetivamente (lesión del bien jurídico o fin de protección de la norma, incremento del riesgo, y causalidad), difícilmente podremos imputar subjetivamente el resultado a una persona concreta.

En otro caso, se produciría una identificación entre bien jurídico y objeto de la conducta, que conduciría al renacimiento del *versari in re illicita*⁶⁸ (y a la consiguiente abolición, si quiera parcial, del principio de

⁶³ La culpabilidad es el fundamento y el límite de la pena, tal y como recuerda M.I. GONZÁLEZ TAPIA. *Neurociencias e imputabilidad: recapitulando*, en esta obra.

⁶⁴ Como mantuvo la STC 150/1991, «La CE consagra sin duda el principio de culpabilidad como principio estructural básico del Derecho Penal» (STS 150/1991, de 4 de julio de 1991, ECLI: ES:TC:1991:150)

⁶⁵ Artículo 5 del código penal español: No hay pena sin dolo o imprudencia.

⁶⁶ Artículo 10 del código penal español: Son delitos las acciones y omisiones dolosas o imprudentes penadas por la Ley.

⁶⁷ Salvo en aquellos delitos con elemento subjetivo, en cuyo caso el dolo podría estar presente desde el mismo análisis de la conducta.

⁶⁸ En cuanto al *Versari in re illicita* la entendemos en el sentido explicado por M. COBO DEL ROSAL, esto es, en su forma tradicional, y más rigurosa, el *versari* implica la imputación a título de dolo consecuencias fortuitas o respecto de las que solamente existe culpa; Sin embargo, con el transcurso del tiempo, han prevalecido formas atenuadas, en las cuales únicamente se responde a título de dolo por las consecuencias no queridas cuando concurre, al menos culpa respecto de ellas o bien se imputan las consecuencias fortuitas, pero no a título de dolo, sino al más leve de imprudencia (M. COBO DEL ROSAL, T. VIVES ANTÓN, *Derecho Penal. Parte General*, Tirant lo Blanch, Valencia, 1991, p. 489).

legalidad y del principio de culpabilidad), determinándose el delito por la causación del resultado y la culpabilidad por la coincidencia gramatical entre este resultado (que coincidiría con la conducta) y el tipo descrito en la norma (J. PÉREZ-ARIAS⁶⁹).

En efecto, la conducta, a veces, no funciona o no consigue el objetivo donde el autor esperaba, aunque se haya puesto el mecanismo en marcha, lo que nos lleva a la dificultad de vincular a una persona determinada con un hecho concreto y un resultado esperado. En palabras de JAKOBS, el comportamiento de un autor o partícipe se define como causa determinante del curso lesivo⁷⁰. Por ello, al margen de figuras concretas allá donde sea posible (una estafa, injurias graves, pero no a través de bots, que nos llevaría a la misma problemática) la solución no pasaría, en todos los casos, por tipificar la cibercriminalidad como derecho penal del resultado, sino como derecho penal de peligro (o de riesgo), lo que ello ya de por sí comporta.

De ahí, que el legislador, si quiere ofrecer una respuesta jurídico penal eficaz, deba tipificar conductas de peligro hipotético o potencial, donde el objeto de investigación e imputación no sea el resultado, sino la idoneidad de la conducta cibernética para originar un riesgo grave en el bien jurídico protegido (adelantando con ello la barrera de protección, y sin exigir llegar a un resultado de imposible imputación causal, que comportaría la plena desprotección del bien jurídico). Se trataría de un delito de actividad, cuyo merecimiento de pena, en palabras de H. JESCHECK/T. WEIGEND⁷¹, descansa sobre la peligrosidad general de la acción típica para determinados bienes jurídicos.

Por tanto, hay mucho de probabilidad e intención, y muy poco -a veces nada- de conexión directa (objetiva y subjetiva) con un resultado concreto (en términos naturalísticos). Mas aún, en la mayoría de los casos, la conducta final lesiva y el resultado ni siquiera serán conocidos, ni esperados, ni imaginados, con carácter previo por el autor. La peligrosidad del ciberdelincuente estriba en que este lanza un producto (que busca solo y crece solo), y solo tiene que esperar el resultado (si lo hay). Es, en realidad, una caza con red (J. PÉREZ-ARIAS⁷²).

⁶⁹ PÉREZ-ARIAS, *Sistema de atribución de responsabilidad penal a las personas jurídicas*, Dykinson, Madrid, 2014, pp. 161-162.

⁷⁰ G. JAKOBS. *Concurrencia de riesgos: curso lesivo y curso hipotético en el Derecho penal*, en «El Derecho Penal. Lecciones y Ensayos», N° 54, 1990, p. 54.

⁷¹ H. JESCHECK, T. WEIGEND, *Tratado de Derecho Penal. Parte General* (trad. M. OLMEDO CARDENETE). Comares, Granada, 2002, p. 283.

⁷² PÉREZ-ARIAS, *Cibercriminalidad: Hacia la nueva realidad -virtual- del derecho penal*, cit., p. 189.

Esto nos llevará, por otro lado, a la impunidad -por cuestiones procesales- de muchas de las conductas que actualmente abundan, pero sería desproporcionado teorizar al contrario y a contracorriente, esto es, castigando, a través de delitos de resultado, conductas que, en términos causales, no están determinadas con plena precisión y seguridad (lo que ello comporta para la antijuridicidad de la acción y para la culpabilidad del sujeto). En efecto, no sería posible, desde una perspectiva jurídico penal y constitucional, atribuir a un sujeto la comisión de un hecho, cuando no existe prueba real que desvirtúe correctamente la presunción de inocencia sobre su autoría.

Esto resulta pacífico cuando se trata de otra tipología delictiva, y no hay motivo para no aplicar el mismo principio a la cibercriminalidad: Si no existe prueba, el tribunal no puede, en ningún caso, desvirtuar la presunción de inocencia mediante hipótesis contra reo o argumentos alambicados y especulativos. De hecho, y como se ha señalado en reiterada jurisprudencia, es ya una doctrina consolidada que la presunción de inocencia debe entenderse como un derecho a no ser condenado sin pruebas de cargo válidas. Ello implica que en la Sentencia condenatoria deben expresarse las pruebas de cargo que sustentan la declaración de responsabilidad jurídico-penal, las cuales, a su vez, han de proceder de verdaderos actos de prueba obtenidos con todas las garantías que exigen la Ley y la Constitución, y normalmente practicados en el acto del juicio oral (STC 171/2000⁷³, 26

⁷³ ECLI:ES:TC:2000:171, p. 5. Esta interesante sentencia del Tribunal Constitucional español sigue diciendo que este Tribunal ha admitido asimismo que la prueba de cargo puede ser por indicios, cuando el hecho objeto de prueba no es el constitutivo de delito sino otro intermedio que permite llegar a él por inferencia lógica, siempre que se cumplan los siguientes requisitos: a) la prueba indiciaria ha de partir de hechos plenamente probados; y b) los hechos constitutivos de delito han de deducirse de esos hechos completamente probados a través de un proceso mental, razonado y acorde con las reglas del criterio humano, explicitado en la Sentencia. La falta de concordancia con las reglas del criterio humano o, en otros términos, la irrazonabilidad, se puede producir, tanto por falta de lógica o de coherencia en la inferencia, cuando los indicios constatados excluyan el hecho que de ellos se hace derivar o no conduzcan naturalmente a él, cuanto por el carácter excesivamente abierto, débil o indeterminado de la inferencia. El control de dichos requisitos debe ser extremadamente cauteloso, al carecer este Tribunal de la necesaria intermediación de la actividad probatoria, que sólo tiene lugar en presencia del órgano judicial que ha de decidir el proceso y con intervención de las partes. No obstante lo anterior, las especiales características de la prueba indiciaria y la elaboración subjetiva de su valoración por el Juez o Tribunal que haya presenciado la prueba hacen que el Tribunal Constitucional deba exigir en las resoluciones judiciales que se someten a su conocimiento por esta causa una motivación suficiente de la inferencia y del resultado de la valoración, de tal manera que una y otro, en este ámbito de enjuiciamiento, ‘no resulten tan abiertos que quepan tal pluralidad de conclusiones alternativas que ninguna de ellas pueda darse por probada’.

de Junio de 2000).

En estos casos, la conducta no podría ser objeto de imputación objetiva, ya que, a pesar de poder valorar, en el mejor de los casos, el incremento del riesgo con la conducta primaria del agente (aquella que pone en circulación el algoritmo, tras su programación) y la probable y/o previsible lesión del bien jurídico (fin de protección de la norma), no existiría técnicamente, salvo que aceptemos la divagación argumental, la debida e individualizable relación causal entre el hecho primario del agente y el hecho final concreto (no aquel primario) y el resultado lesivo que se termina produciendo.

Ni siquiera podríamos hablar de la llamada causalidad cumulativa, aunque estemos ante varios actos que terminan produciendo -todos ellos juntos- el resultado. Y no lo estamos, porque el aprendizaje del algoritmo autónomo no es, técnicamente, un acto consciente (sobre la causalidad) de la persona natural (del programador), por lo que difícilmente puede considerarse ese acto algorítmico de aprendizaje como una conducta consciente y subsiguiente de la persona. Además de ello, el acto algorítmico de aprendizaje no es por sí solo capaz de producir el resultado sin ese acto primario, ni del acto primario podemos derivar como consecuencia causal directa y necesaria la producción del resultado, sin tener en cuenta aquellos otros sucesos algorítmicos posibles y autónomos posteriores. En palabras de S. MIR PUIG⁷⁴, existiría causalidad cumulativa cuando el resultado fuera causado por dos o más condiciones cada una de las cuales resultó suficiente por sí sola para producirlo. Y, en todo caso, tendríamos que estar ante actos humanos en sentido estricto, esto es, ante manifestaciones conscientes de la voluntad de la persona, cualidad inexistente, por definición, en un acto algorítmico artificial autónomo.

Por tanto, en nuestra opinión, la mayoría de las conductas derivadas de inteligencia artificial escaparían de los parámetros de previsibilidad racional, y entrarían por completo en el terreno de la predicción estadística, hecho este absolutamente distinto. Si por previsibilidad se ha de entender la posibilidad de prever la conducta de otros sujetos (I. LIFANTE VIDAL⁷⁵), la predicción estadística requiere método científico, porque el resultado no es previsible, sin más, sin llevar a cabo un tratamiento y estudio sobre la combinación matemática de datos que explique la causalidad (con su margen de error correspondiente). Por tanto, la conducta escapa de la previsibilidad tradicional, ya que la causación se convierte en una

⁷⁴ S. MIR PUIG. *Derecho Penal. Parte General*, Reppertor, Barcelona, 2012, p. 250.

⁷⁵ I. LIFANTE VIDAL. *La relevancia de la previsibilidad jurídica*, en *DOXA. Cuadernos de Filosofía del Derecho. Edición especial*, Marcial Pons, Madrid, 2017, p. 145.

predicción de difícil intuición a priori. De ahí que consideremos el delito de peligro como única forma no solo de anticipación de la lesión impredecible y dudosa (en términos causales), sino para poder valorar como intencionada una conducta claramente peligrosa y dolosa. En otro caso, se produciría la paradoja de necesitar inteligencia artificial -no parámetros jurídicos- para poder medir, a su vez, la responsabilidad penal de un sujeto que actúa -hecho relevante- a través de una fórmula algorítmica que auto-aprende (hecho no sujeto a control humano). Quizás no tarde mucho en implantarse esto que se ha dado en llamar el '*juez-robot*'⁷⁶. No debemos olvidar que la ciencia forense es una de las grandes beneficiarias de estas nuevas tecnologías. Sería, como señala J. PERIS RIERA⁷⁷, algo muy parecido a lo que ocurrió, en su día, con las huellas dactilares o el ADN, avances ambos sin los cuales difícilmente podríamos explicar hoy los resultados que se producen en la averiguación del delito y persecución del culpable.

Quedaría por valorar si la mera activación del algoritmo permitiría en un momento posterior concretar el peligro para el bien jurídico pues, en ese caso, el delito debería ser de peligro concreto⁷⁸ (convirtiéndose el riesgo concreto en auténtico elemento del tipo). En otro caso, esto es, impedir que cada intérprete pueda analizar el peligro de una manera diferente (sobre todo partiendo de que estamos ante un peligro estadístico, que permitiría múltiples interpretaciones, a favor y en contra) nos llevaría a configurar la conducta como delito de peligro abstracto (de mera actividad), haciendo innecesaria, con ello, la prueba sobre el efectivo riesgo hacia el bien jurídico. No obstante, se ha de matizar que el delito de peligro abstracto no puede ser el modo de solucionar un problema de imputación objetiva, ni tampoco resulta muy claro que este tipo de modalidades tenga en realidad bien jurídico.

Entendemos, y solo sirve para complicar aún más la cuestión, que se debe distinguir, con J. SCHULENBURG⁷⁹, entre bien jurídico y objeto del bien jurídico. Si una conducta lesiona o pone en peligro algún elemento u

⁷⁶ R. CASTILLO FELIPE, S. TOMÁS TOMÁS, *Proceso penal e inteligencia artificial: reflexiones en torno a su futura aplicación en la fase de juicio oral*, en esta obra, § 1.

⁷⁷ PERIS RIERA. *Inteligencia Artificial y neurociencias: Avances del derecho penal contemporáneo*, en esta obra, § 1.

⁷⁸ Si se considerase que la activación del algoritmo permite acreditar el concreto riesgo para el bien jurídico.

⁷⁹ Como considera el autor J. SCHULENBURG, *Relaciones dogmáticas entre bienes jurídicos, estructura del delito e imputación objetiva*, en *La teoría del bien jurídico ¿Fundamento de legitimación del Derecho Penal o juego de abalorios dogmático?* (Dir. R. HEFENDEHL, A.V. HIRSCH, W. WOHLERS), Marcial Pons, Madrid, 2016, pp. 341-354.

objeto del bien jurídico (por tanto, algo tangible o preciso y determinado) no es necesario el delito de peligro abstracto, siendo suficiente con el delito de resultado o de peligro concreto; y al contrario, cuando lo que se pretende proteger no es un elemento u objeto concreto del bien jurídico, ni el bien jurídico en sí considerado puede ser entendido como individual sino como colectivo, es cuando procedería establecer, en este contexto temático, el delito como de peligro abstracto. Por ello, se decía más arriba que el legislador debe decidir, antes que nada, si la ciberseguridad, se ha convertido en un bien jurídico autónomo y colectivo (huyendo de *conceptos ideales e inatacables*, como destaca HEFENDEHL⁸⁰) y, que fundamentará su protección más allá del concreto ataque a un bien jurídico individual o elemento individual concreto de un bien jurídico más colectivo. En este caso, se podrían tipificar tres conductas; una, de resultado pluriofensivo; otra, de peligro concreto y otra de peligro abstracto. Con ello se propiciaría la sanción de la lesión como delito de resultado y la aplicación del delito de peligro (concreto) si no se produce resultado alguno, o con la conducta se ha producido un resultado, pero también riesgo para otros sujetos, en cuyo caso concurriría en concurso ideal⁸¹ con el delito de resultado). El delito de peligro abstracto quedaría para conductas muy iniciales.

Entendemos que la solución pasaría por distinguir entre aquellos ataques a la ciberseguridad (como bien jurídico colectivo y pseudo abstracto), de aquellos otros que atentan contra bienes individuales (como el patrimonio, intimidad, etc.). En este segundo caso, resultaría suficiente con tipificar un delito de resultado o peligro concreto (en términos muy parecidos a como se hace en la actualidad en los casos citados, y siendo el fenómeno tecnológico un medio comisivo específico).

No obstante, y aunque profundizar en ello excedería de un trabajo de esta naturaleza, de quedar reguladas ambas modalidades (resultado y peligro concreto para los casos de bien jurídico individual y peligro abstracto para los de bien jurídico colectivo), se podría interpretar que el delito de peligro abstracto no sería más que una tentativa del delito de resultado (habrá quien sugiera, incluso, que no es más que un acto preparatorio), lo que podría hacer innecesaria su tipificación, por redundante, para evitar problemas

⁸⁰ R. HEFENDEHL, *Las jornadas desde la perspectiva de un partidario del bien jurídico*, en *La teoría del bien jurídico ¿Fundamento de legitimación del Derecho Penal o juego de abalorios dogmático?* (Dir. R. HEFENDEHL, A.V. HIRSCH, W. WOHLERS), Marcial Pons, Madrid. 2016, p. 401.

⁸¹ Si como consecuencia de la conducta solo se ha producido lesión del bien jurídico individual, sin haberse causado riesgo para otros sujetos, el delito de resultado absorbería el delito de peligro concreto, en virtud de lo dispuesto del artículo 8 del código penal español.

hermenéuticos de tipo práctico. En este caso, y teniendo tipificada una modalidad de resultado (o incluso de mera actividad, pero lesivo), sería suficiente con la aplicación de los principios generales del código penal español (concretamente su artículo 16.1⁸², para la tentativa (inacaba, en todo caso, para el delito de mera actividad), o el artículo 18 para los actos preparatorios), para valorar las distintas fases delictivas.

En cualquier caso, alejar al operador de interpretaciones creativas y de moda, impediría la absolución por falta de prueba respecto del riesgo concreto (hecho éste que será habitual), lo que es más propio de los delitos de peligro abstracto. Como señala S. MIR PUIG⁸³ lo que no podría admitirse es que en los delitos de peligro abstracto falte el tipo siempre que se pruebe que a posteriori no resultó peligro concreto, añadiendo que esto contradiría el fundamento político-criminal de los delitos de peligro abstracto, que ha de verse en la conveniencia de no dejar a juicio de cada cual la estimación de la peligrosidad de acciones que con frecuencia son lesivas (peligro estadístico).

Incluso, y si, como decimos, se planteara el legislador un bien jurídico autónomo colectivo y difuso, no extrañaría que se decantara por la fórmula del delito obstáculo, esto es, y como ha relatado la jurisprudencia del Tribunal Supremo, aquel fenómeno de criminalización anticipada mediante el que se castigan conductas en un momento anterior a la lesión del bien jurídico e incluso con anterioridad a que se genere un peligro concreto o abstracto para el bien jurídico. Son por lo tanto delitos que se configuran como auténticos obstáculos que tienen como función impedir que lleguen a producirse los actos delictivos futuros que se tipifican en otros preceptos. En ellos el principio de ofensividad cede ante la necesidad de prevención general y se presentan como tipos penales formales o de mera desobediencia mediante los que se anticipa, si bien en algunos casos de forma muy cuestionable, la barrera defensiva que supone la aplicación de toda norma penal (STS 2315/2017⁸⁴). En realidad, y como señala P. CUESTA PASTOR⁸⁵, la utilización por el legislador del delito obstáculo equivale a la criminalización de los actos preparatorios para cometer ciertos delitos.

⁸² Donde se regula la tentativa del delito.

⁸³ MIR PUIG, *Derecho Penal. Parte General*, p. 241.

⁸⁴ STS 2315/2017, de 8/6/2017 (ECLI: ES:TS:2017:2315), p. 6.

⁸⁵ P. CUESTA PASTOR, *Delitos obstáculo. Tensión entre política criminal y teoría del bien jurídico*, Editorial Comares, Granada, 2002, p. 44.

6. Conclusión

Resulta difícil mantener una conclusión cerrada y definitiva (y no provisional) en una temática tan cambiante como es la ciberdelincuencia. En la actualidad, la solución que entendemos correcta pasaría por aceptar una de estas tres propuestas y que se exponen de manera esquemática:

En primer lugar, seguir entendiendo que lo informático es solo un medio comisivo específico y, por tanto, seguir regulando la cibercriminalidad como mera modalidad agravada dentro de las categorías tradicionales de delito (es lo que ocurre con la estafa, los daños, la intimidación etc.). Sin embargo, debemos tener en cuenta que el uso de la tecnología empieza a generar una realidad que merece ser tipificada de manera específica.

Por ejemplo, si una injuria se comete a través de cualquier aplicación conocida en la red (Twitter, Facebook, Instagram, ...), se produce un *efecto eco* que está buscado de propósito por el autor. Este hecho es un claro elemento subjetivo del tipo que añade un plus de antijuridicidad a la conducta que requiere un tratamiento jurídico diferenciado. El derecho al olvido ni tan siquiera es capaz de eliminar este efecto eco. Es algo más que la mera publicidad⁸⁶, es la imposibilidad de que la conducta cese, porque siempre habrá un rastro visible del atentado al honor.

Como se ha señalado, y hacemos nuestro, los buscadores se encargan de sacar a luz dichos contenidos, una y otra vez, bajo un mismo principio: la libertad de expresión y de circulación de contenidos online aplicable a una internet que, según algunos, se autorregula y no requiere legislación (F. TOMEO⁸⁷). Sin embargo, la experiencia demuestra la necesidad de poner límites a esta actividad. El problema es cómo hacerlo, sin que suponga merma de otros derechos. Se ha de partir del principio general de que no existe en la realidad (y tampoco, por tanto, en la red) un derecho al insulto.

En segundo lugar, crear un título autónomo en el que el legislador trate de manera conjunta el fenómeno de la cibercriminalidad, con la posibilidad incluso de establecer disposiciones comunes. Se trataría de trasladar los tipos actuales que regulan lo informático (que se convertirían así en delitos pluriofensivos) más los nuevos específicos que deben ser tipificados, con el consiguiente problema de sistematización. Esta posibilidad procede si se considerara que la seguridad en el ciberespacio tiene entidad autónoma

⁸⁶ El artículo 211 del código penal español establece que la calumnia y la injuria se reputarán hechas con publicidad cuando se propaguen por medio de la imprenta, la radiodifusión o por cualquier otro medio de eficacia semejante.

⁸⁷ F. TOMEO. *Calumnias e injurias on line*, en *Cibercrimen II* (Dir. D. DUPUY), B de f, Buenos Aires, 2018, p. 111.

y suficiente, como para erigirse en bien jurídico protegido independiente. Esta independencia nos permitiría configurar la ciberseguridad como un bien jurídico colectivo y permitiría establecer, en determinadas condiciones, delitos de peligro abstracto o, incluso, delitos obstáculo.

Entiendo que establecer esta modalidad típica⁸⁸ para resolver problemas de imputación objetiva sería pervertir los principios esenciales del derecho penal, dado que se estaría ocultando, a través de estas nuevas figuras legislativas, un claro y reprochable delito de sospecha como objeto de castigo. Allá donde haya un bien jurídico individual, sin posibilidad de considerarlo colectivo, precisa de la creación de delitos de resultado o, a lo sumo, de peligro concreto (donde el efectivo riesgo se convierte, en un elemento más del delito).

Y, en tercer lugar, establecer una circunstancia modificativa de la responsabilidad (agravante) con el fenómeno cibernético, de tal forma que el libro II no deba, salvo excepciones, regular de manera autónoma y específica cada una de las modalidades delictivas en las que quepa imaginar la cibercriminalidad. El problema que habría que analizar es la proporcionalidad de la pena, porque en muchas ocasiones la mera agravante no satisface la proporcionalidad de la pena en determinados delitos.

Todas las posibilidades se encuentran abiertas y son claramente mutables por otras, dependiendo, en mi opinión, de que el avance de la técnica permita dos cosas básicamente: De un lado, demostrar que el fenómeno cibercriminal es un hecho que no puede adaptarse a figuras tradicionales (salvo en supuestos específicos); y de otro, determinar con exactitud y precisión quién hay detrás del algoritmo, y probar que el derrotero seguido por éste es perfectamente conocido por su autor, al margen de que la víctima sea o no determinada desde un inicio. Esto, que hoy no ocurre, porque no hay medios que lo permitan (al menos medios directos que sean aptos para su prueba en un procedimiento penal) es lo único que permitiría tratar el ciberdelito, en todos sus aspectos, como delitos de resultado (o, en el peor de los casos, de peligro concreto), y no tan solo de las siempre discutibles figuras de peligro abstracto y obstáculo.

Con todo, el estudio caso a caso es imprescindible, y habrá de cuestionarse en cada modalidad delictiva la mejor opción, en función de los datos empíricos contrastados y sus resultados en un procedimiento judicial. Una vez queden resueltas las consideraciones de principios en derecho penal, restará entrar a detalle en la llamada parte especial, donde la solución se deberá hacer depender de otros factores más concretos y perfilados.

⁸⁸ Delitos de peligro abstracto.

En definitiva, estamos ante un nuevo fenómeno en el que, de momento, es mejor formular preguntas que respuestas categóricas, y entender que el vocabulario (los anglicismos) no construye teorías, solo distrae del problema central que sigue sin tener estructura ni solución.

7. Bibliografía

- AGUSTINA J.R., *Cibercriminalidad y perspectiva victimológica: un enfoque general explicativo de la cibervictimización*. «Cuadernos de política criminal». Segunda época. Madrid. Núm. 136, Mayo 2022.
- BECK U., *Sociedad de riesgo. Hacia una nueva modernidad*, Paidós, Barcelona-Buenos Aires-México, 2002.
- BENÍTEZ ORTÚZAR I., *Ciberdelitos. la implementación en el ordenamiento interno de los acuerdos internacionales en materia de ciberdelincuencia*, en *La robótica y la inteligencia artificial en la nueva era de la revolución industrial 4.0* (Coord. M.J. CRUZ BLANCA/I. LLEDÓ BENITO), Dykinson, Madrid, 2021.
- BONFANTE P., *Corso di diritto romano*, Giuffrè, Milano, 1963.
- CÁMARA ARROYO S., *Estudios criminológicos contemporáneos (IX): La Cibercriminología y el perfil del ciberdelincuente*, en «Derecho y Cambio Social», Núm. 60, Abril-junio 2020.
- CASTILLO FELIPE R., TOMÁS TOMÁS S., *Proceso penal e inteligencia artificial. Reflexiones en torno a su futura aplicación en la fase de juicio oral*, en esta obra.
- COBO DEL ROSAL M., VIVES ANTÓN T., *Derecho Penal. Parte General*, Tirant lo Blanch, Valencia, 1991.
- CORVALÁN, J.G., *Presentación: Inteligencia artificial. Automatización y predicciones en el Derecho*, en *Cibercrimen III* (Dir. D. DUPUY/J.G. CORVALÁN), B de f, Buenos Aires, 2020.
- CUESTA PASTOR P., *Delitos obstáculo. Tensión entre política criminal y teoría del bien jurídico*, Editorial Comares, Granada, 2002.
- CUMBRERAS M., LÓPEZ P., *¿Es necesario un marco ético para guiar el desarrollo y uso de la inteligencia artificial en las organizaciones?*, en *La robótica y la inteligencia artificial en la nueva era de la revolución industrial 4.0* (Coord. M.J. CRUZ BLANCA/I. LLEDÓ BENITO), Dykinson, Madrid, 2021.
- DE LA MATA BARRANCO N.J., *Reflexiones sobre el bien jurídico a proteger en el delito de acceso informático ilícito (art. 97 CP). El concepto de de*

- privacidad informática y la tutela del buen funcionamiento de los sistemas de información y comunicación*, en «Cuadernos de política criminal», segunda época, Núm. 118, mayo 2016.
- EDUARDO ABOSO G., *Derecho Penal Cibernético. La cibercriminalidad y el Derecho Penal en la moderna sociedad de la información y la tecnología de la comunicación*, B de f, Buenos Aires, 2017.
- GIRALDI A., *Deshumanizando la culpabilidad: los sistemas inanimados en la teoría del delito*, en esta obra.
- GARAT S., REALE J., *La reforma penal en materia de cibercrimen en la República Argentina*, en *Cibercrimen II* (Dir. D. DUPUY), B de f, Buenos Aires, 2018.
- GONZÁLEZ TAPIA M.I, *Neurociencias e imputabilidad: recapitulando*, en esta obra.
- HEFENDEHL R, *Las jornadas desde la perspectiva de un partidario del bien jurídico*, en *La teoría del bien jurídico ¿Fundamento de legitimación del derecho penal o juego de abalorios dogmático?* (Dir. R. HEFENDEHL, A.V. HIRSCH, W. WOHLERS), Marcial Pons, Madrid, 2016.
- HERNÁNDEZ DÍAZ L., *El delito informático*. «EGUZZILORE, Cuaderno del Instituto Vasco de Criminología», 23, diciembre 2009.
- JAKOBS G., *Concurrencia de riesgos: curso lesivo y curso hipotético en el Derecho penal*, en «El Derecho Penal Lecciones y Ensayos», N° 54, 1990.
- H. JESCHECK, T. WEIGEND, *Tratado de Derecho Penal. Parte General* (trad. M. OLMEDO CARDENETE), Comares, Granada, 2002.
- KURZWEIL R., *The Age of Intelligent Machines*, MIT Press, Cambridge, 1990.
- LINFANTE VIDAL I., *La relevancia de la previsibilidad jurídica*, en «DOXA, Cuadernos de Filosofía del Derecho», edición especial, 2017.
- LLEDÓ BENITO I., *Ciberseguridad Versus Ciberdelincuencia*, en *El derecho penal, robots, IA y cibercriminalidad: desafíos éticos y jurídicos ¿Hacia una distopía?*, Dykinson, Madrid, 2022.
- MATA Y MARTÍN R.M., *Criminalidad informática: una introducción al cibercrimen*. «Actualidad Penal», Núm. 37, Sección Doctrina, Semana del 6 al 12 Oct. 2003, Ref. XXXVI, tomo 3.
- MARTÍNEZ ATIENZA G., *El blanqueo de capitales y su sanción administrativa y penal*. «Revista de Derecho vLex», número 202, Marzo 2021.
- MATÉ JIMÉNEZ C., *Big data. Un nuevo paradigma de análisis de datos*, en «Anales de mecánica y electricidad», noviembre-diciembre 2014.
- MIR PUIG S., *Derecho Penal. Parte General*, Reppertor, Barcelona, 2012.
- MORÁN BLANCO S., *La ciberseguridad y el uso de las tecnologías de la*

- información y la comunicación (tic) por el terrorismo*, en «Revista Española de Derecho Internacional». Sección ESTUDIOS. Vol. 69/2, julio-diciembre 2017.
- MORILLAS FERNÁNDEZ D.L., *Implicaciones de la inteligencia artificial en el ámbito del derecho penal*, en esta obra.
- MUÑOZ ALONSO A., *Andamios para las ideas*. Colección Aula de ideas, volumen 1, Aula, Madrid, 1952.
- PALMA HERRERA, J.M. *Inteligencia artificial y neurociencia. Algunas reflexiones sobre las aportaciones que pueden hacer al Derecho Penal*, en esta obra.
- PALMA ORTIGOSA A., *Decisiones automatizadas y protección de datos*, Dykinson, Madrid, 2022.
- PÉREZ-ARIAS J., *Sistema de atribución de responsabilidad penal a las personas jurídicas*, Madrid, 2014.
- PÉREZ-ARIAS J., *Cibercriminalidad: hacia la nueva realidad -virtual- del derecho penal*, en «Revista Internacional de Doctrina y Jurisprudencia», volumen 26, 2021.
- PERIS RIERA J., *Inteligencia Artificial y Neurociencias: Avances Del Derecho Penal Contemporáneo*, en esta obra.
- PEROZZI S., *Istituzioni di Diritto Romano*, Milano, 1947.
- RADBRUCH G., *Einführung in die Rechtswissenschaft*, Quelle & Meyer, Stuttgart, 1929.
- RAYÓN BALLESTEROS M.C., GÓMEZ HERNÁNDEZ J.A., *Cibercrimen: particularidades en su investigación y enjuiciamiento*, en «Anuario Jurídico y Económico Escurialense», XLVII, 2014.
- ROMEO CASABONA C.M., *Poder informático y seguridad jurídica: la función tutelar del derecho penal ante las nuevas tecnologías de la información*, Fundesco, Madrid, 1988.
- RUSSEL S., NORVIG P., *Inteligencia artificial. Un enfoque moderno*, Pearson, Madrid, 2008.
- SALVADOR CARRASCO L., *Redes de anonimización en internet: cómo funcionan y cuáles son sus límites*, en «Revista de Instituto Español de Estudios Estratégicos», número 16/12. 2012.
- SCHULENBURG J., *Relaciones dogmáticas entre bienes jurídicos, estructura del delito e imputación objetiva*, en *La teoría del bien jurídico ¿Fundamento de legitimación del derecho penal o juego de abalorios dogmático?* (Dir. R. HEFENDEHL/A.V. HIRSCH/W. WOHLERS), Madrid, 2016.
- SOUZA DE MENEZES, C., AGUSTINA J.R., *Big Data, inteligencia artificial y policía predictiva*, en *Cibercrimen III* (Dir. D. DUPUY/J.G. CORVALÁN), B de f, Buenos Aires, 2020.
- TAMARIT SUMALLA J.M., *Ciberdelincuencia y cibervictimización*, en «Revista

- de los Estudios de Derecho y Ciencia Política», n.º 22, Junio, 2016.
- TEMPERINI M., *Delitos informáticos y cibercrimen: Alcances, conceptos y características*, en *Cibercrimen y delitos informáticos. Los nuevos tipos penales en la era de internet*, Erreius, Buenos Aires, 2018.
- TOMEIO, F. *Calumnias e injurias on line*, en *Cibercrimen II* (Dir. D. DUPUY), B de f, Buenos Aires, 2018.

CAPÍTULO VII

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

Laura Notaro

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

1. *Introducción*

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

con un cierto nivel de aproximación, la libertad provisional.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

2.2. Herramientas de valoración del riesgo de reincidencia en Europa

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

²⁰ *Supra*, par. 2.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

learning)³¹.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

³⁸ *Ivi*, par. 137.

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

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

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

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

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

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

³⁹ *Ivi*, par. 138.

⁴⁰ *Ivi*, parr. 141 ss.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

5. Conclusiones

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

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

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

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

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

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

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

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

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

6. Bibliografía

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

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

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

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

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

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

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

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

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

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

CAPÍTULO VIII

*Proceso penal e inteligencia artificial: prevenciones en torno a su futura aplicación en la fase de juicio oral**

Rafael Castillo Felipe, Salvador Tomás Tomás

SUMARIO: 1. La actual imposibilidad técnica y jurídica de un ‘juez-robot’ como punto de partida necesario – 2. Cuestiones problemáticas que podría suscitar la aplicación de la IA en la fase de juicio oral – 2.1. El uso de la IA para la obtención de información sobre los hechos que fundan la acusación y la defensa – 2.2. El recurso a la IA como prueba para determinar la fiabilidad o credibilidad de los medios de prueba practicados en el juicio oral – 2.3. El empleo de la IA para decidir sobre la ilicitud de las diligencias de investigación o medios de prueba practicados – 2.4. El uso de la IA en el juicio de apreciación de la prueba – 2.5. La aplicación de la IA en la elaboración de la fundamentación jurídica de la sentencia y en la resolución de recursos posteriores frente a esta – 3. La cuestión de la transparencia del algoritmo empleado en el enjuiciamiento penal – 4. Bibliografía.

1. La actual imposibilidad técnica y jurídica de un ‘juez-robot’ como punto de partida necesario

La problemática de la inteligencia artificial –en lo sucesivo, IA– aplicada al proceso, tanto penal como civil, es uno de los muchos retos a los que tiene que hacer frente la justicia del siglo XXI. En los últimos tiempos, las aportaciones en este campo son ubérrimas y, en su mayor parte, se mueven entre lo tópico y lo anecdótico. Así, la doctrina ha identificado una serie de problemas comunes que podría ofrecer el despliegue de la tecnología citada en los tribunales, sin embargo, las experiencias de uso reales son todavía muy limitadas. Circunstancia que impone a cualquier estudioso que se acerque al tema una elevada dosis de abstracción con el fin de situarse ante una realidad muy distinta a la que puede contemplarse hoy en el foro.

* El presente trabajo se ha realizado en el marco del proyecto de investigación «Hacia una Justicia eficiente y sostenible: una aproximación multidisciplinar» (UAL2020-SEJ-C1965), financiado por la Consejería de Transformación Económica, Industria, Conocimiento y Universidades, Junta de Andalucía, cofinanciado por el Programa Operativo FEDER 2014-2020.

Partiendo de dicho ejercicio de abstracción o, si se prefiere, de imaginación, el presente trabajo se centrará en la posibilidad de aplicar IA en la fase de juicio oral del proceso penal y, en su caso, en la resolución de los eventuales recursos interpuestos contra las sentencias dictadas en el mismo¹. El ordenamiento procesal penal español será el régimen jurídico que tomaremos en consideración para dibujar y resolver las distintas hipótesis.

Entre los problemas más sugerentes y de mayor enjundia, se encuentra, sin lugar a dudas, el del ‘juez-robot’. Esta expresión tiene un fuerte potencial evocador; en concreto, de un *ius dicere* dispensado de forma automática por una máquina inteligente, o que, como poco, simula que lo es.

Si bien, a pesar de lo atractivo del tópico, consideramos que no es necesario detenerse en él, por cuanto, hoy por hoy, en el ordenamiento procesal español no es viable que un programa informático administre justicia por sí solo. Los obstáculos en este punto son, por una parte, técnicos y, por otra, jurídicos.

El principal obstáculo técnico vendría dado por un factor endógeno de la IA, cual sería las limitaciones de los sistemas expertos para reproducir algunas parcelas del razonamiento humano que intervienen en la decisión judicial (v. gr., piénsese en los obstáculos derivados del procesamiento del lenguaje natural o en la imposibilidad de emular la creatividad humana)².

¹ Las implicaciones del uso de la IA en materia de policía predictiva y en la fase de instrucción requieren un estudio aparte. En cuanto a otras cuestiones, como las transformaciones que provocará la IA en el rol de los operadores jurídicos, o la eficiencia que puede alcanzarse automatizando las labores de gestión procesal –v. gr., evitación de dobles señalamientos–, vid. R. CASTILLO FELIPE, *Proceso civil e inteligencia artificial* en AA.VV., *Proceso civil y nuevas tecnologías*, J. SIGÜENZA LÓPEZ (dir.), Thomson-Aranzadi, Cizur Menor (Navarra), 2021, pp. 259-296.

² Sobre las limitaciones actuales de la IA puede acudir, con carácter general, a las explicaciones de PEARL sobre la escala de la causalidad. Según este matemático, dicha escala consta de tres niveles: el primer peldaño sería la observación, esto es, el aprendizaje por asociación, que sería en el que se encuentra actualmente la tecnología que interesa a nuestro estudio, en tanto en cuanto aprende a la vista de los datos; el segundo estaría constituido por el aprendizaje por intervención, caracterizado por tomar en consideración acciones que no tienen un precedente. Finalmente, el último nivel comprendería el aprendizaje mediante contrafactuales o hipótesis alternativas. Mientras la IA no alcance el tercer peldaño, no podrá funcionar como la mente humana. Estas explicaciones pueden hallarse en J. PEARL; D. MCKENZIE, *The Book of Why. The New Science of Cause and effect*, Penguin Random House UK, Londres, 2019, singularmente en el capítulo 10 –ubicado, en nuestra versión de *ebook* (formato EPUB), en las pp. 673-716–. Una explicación comprensible sobre los límites de los actuales sistemas de *machine learning* y *deep learning* se localiza también en J.J. MURILLO FUENTES, ¿Qué es lo que no funciona en los algoritmos de inteligencia artificial? en AA.VV., *Uso de la información y de los datos personales en los procesos: los cambios en la era*

A ello habría que sumar, asimismo, un inconveniente exógeno a los sistemas expertos, contingente, pero de probable aparición: las dificultades que enfrenta cualquier intento de cambio tecnológico en nuestro sistema judicial. Las experiencias previas con la implantación generalizada del sistema LexNET y la introducción del expediente judicial electrónico evidencian que el tiempo requerido para incorporar nuevas herramientas informáticas es dilatado, los resultados iniciales son pocos y los riesgos de causar disfuncionalidades graves en el desempeño normal de la Justicia muy altos³. Por tanto, los precedentes no invitan a pensar que el aterrizaje de los algoritmos decisorios en el mundo judicial vaya a ser distinto y, sobre todo, que pueda llevarse a cabo de la noche a la mañana⁴.

digital, I. COLOMER HERNÁNDEZ (dir.); M.A. CATALINA BENAVENTE; S. OUBIÑA BARBOLA (coords.), Thomson-Aranzadi, Cizur Menor (Navarra), 2022, pp. 151-171. En doctrina jurídica, puede acudirse a las limitaciones que han remarcado: J. NIEVA FENOLL, *Prueba científica. Cuestiones de futuro: neurociencia e inteligencia artificial* en AA.VV., *La prueba en el proceso*, Atelier, Barcelona, 2018, pp. 473-495, vid., especialmente, pp. 487-495; *Id.*, *Inteligencia artificial y proceso judicial*, Marcial Pons, Madrid, 2018, pp. 79 y ss.; M. FERNÁNDEZ SALMERÓN, *De la reutilización de sentencias al 'Big data' judicial. Aproximación a la metamorfosis experimentada por los modelos de uso de la información en el marco de la actividad jurisdiccional*, en AA.VV., *Modernización digital e innovación en la Administración de Justicia*, F. GÓMEZ MANRESA; M. FERNÁNDEZ SALMERÓN (coords.), Thomson-Aranzadi, Cizur Menor (Navarra), 2019, pp. 63-102; S. BARONA VÍLAR, *Algoritmización del Derecho y de la Justicia. De la inteligencia artificial a la Smart Justice*, Tirant Lo Blanch, Valencia, 2021, pp. 558 y ss., 584-587 y 621; T. ARMENTA DEU, *Derivas de la justicia. Tutela de los derechos y solución de controversias en tiempos de cambio*, Marcial Pons, Madrid, 2021, pp. 302, 308 y 309; y, de nuevo, NIEVA FENOLL, *Inteligencia artificial y proceso judicial: perspectivas tras un alto tecnológico en el camino*, en «Revista General de Derecho Procesal» [en línea], núm. 57, 2022 [consultada por última vez el 9 de noviembre de 2022], disponible en: <<https://acortar.link/DTrjhX>>.

³ Todavía habría que estudiar con calma –labor que escapa al objeto de este estudio– si los escollos que han aparecido durante la puesta en marcha de las plataformas anteriores son de corte económico, técnico u obedecen a una mala planificación que genera fallos imprevistos y respuestas tan improvisadas como erráticas a los mismos. Sobre la implantación del expediente judicial electrónico pueden consultarse las consideraciones de J.L. MEDRANO SÁNCHEZ, *Luces y sombras del expediente judicial digital* en AA.VV., *Aciertos, excesos y carencias en la tramitación del proceso*, J.F. HERRERO PEREZAGUA; J. LÓPEZ SÁNCHEZ (dirs.), Atelier, Barcelona, 2020, pp. 49-85. Por lo demás, los fallos de LexNET son notorios para cualquier profesional que forme parte de la Administración de Justicia o colabore con la misma y no los listaremos aquí. En otro estudio previo hemos dado cuenta de los más graves, vid. CASTILLO FELIPE, *Proceso civil e inteligencia artificial*, cit., p. 281, nota 51.

⁴ Realizamos la afirmación aun a sabiendas de las previsiones de la «Estrategia de Justicia 2030», en cuyo marco se apuesta por la potenciación de la IA como uno de los medios para lograr «eficiencia digital» del «servicio de la justicia» (en relación con este punto, vid. F. BUENO DE MATA, *Hacia un proceso civil eficiente: transformaciones judiciales en un contexto pandémico*, Tirant Lo Blanch, Valencia, 2021, pp. 99 y ss.). Si bien, hay que

Junto a los impedimentos técnicos, encontramos obstáculos de índole jurídica. En concreto, el principio de exclusividad jurisdiccional, instituido por el art. 117.1 de la Constitución Española de 1978 –en lo sucesivo, CE–, en conjunción con sus arts. 23.2 y 103, y el art. 301 de la Ley Orgánica 6/1985, de 1 de julio, del Poder Judicial –a partir de ahora, LOPJ–, impiden que una persona jurídica o una cosa, como sería un programa informático constituido por un conjunto de datos, impartan justicia.

En relación con este punto, cabe anotar que, por el momento, las iniciativas legislativas sobre el uso de la IA vedan el uso de algoritmos decisorios sin supervisión humana en el ámbito del proceso. En este sentido, la propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial, de 21 abril de 2021, cuyo artículo 14 impone la vigilancia humana de los sistemas de IA cuando se apliquen a sectores de alto de riesgo, entre los cuales está incluida la Justicia⁵.

Por su parte, el art. 57 del PLMEDJ contempla la necesidad de validación humana de las propuestas de resolución emitidas por los algoritmos que pudieran asistir a jueces, letrados de la Administración de Justicia y fiscales en el ámbito de sus respectivas competencias⁶.

Por consiguiente, teniendo en cuenta los obstáculos jurídicos mencionados y los textos prelegislativos apuntados, en esta contribución partiremos del estadio previo al del juez autómatas, conformado por la posibilidad de que la IA se utilice para auxiliar al tribunal encargado de dictar la sentencia o incluso a las partes que tratan de destruir o confirmar la presunción de inocencia.

En contra de lo que pudiera parecer, el escenario que constituye nuestro punto de partida sí tendría cabida en el estado actual de la legislación,

tener en cuenta, por un lado, que, transcurridos dos años desde la puesta en marcha de la iniciativa, todavía no se han promulgado las leyes sobre las que aquella descansa. Así sucede, en lo que aquí interesa, con el Proyecto de Ley de Medidas de Eficiencia digital del Servicio Público de la Justicia –en lo sucesivo, PLMEDJ– que se encuentra todavía en tramitación parlamentaria (vid. <<https://acortar.link/5OKML6>>). Por otro lado, en tiempos de eslóganes y consignas, conviene recordar que la publicación de una ley no lleva aparejada una suerte de efecto alquímico en virtud del cual se opere la transformación automática de la realidad sobre la que incide. De este modo, aunque la LMEDJ llegue a ver la luz, se requerirá tiempo para acometer las actuaciones materiales que permitan introducir sistemas de IA de manera generalizada en los órganos judiciales.

⁵ Disponible en: <<https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=celex:52021PC0206>> [consultado por última vez el 9 de noviembre de 2022].

⁶ Vid. Boletín Oficial de las Cortes Generales. Congreso de los Diputados –BOE– [en línea], serie A, núm. 116-1, de 12 de septiembre de 2022 [consultado por última vez el 9 de noviembre de 2022], disponible en: <<https://acortar.link/5OKML6>>.

pues el art. 42 de la Ley 18/2011, reguladora del uso de las tecnologías de la información y la comunicación en el ámbito de la Administración de Justicia –en adelante, LTICAJ– ampara actuaciones judiciales automatizadas realizadas por sistemas expertos⁷.

De este modo, en el caso de que el uso de una IA fuerte en la fase de juicio oral –basada en el llamado ‘*deep learning*’– llegue a ser técnicamente viable, podría aportar las siguientes utilidades⁸:

En primer término, podría suministrar información sobre los hechos que conforman el objeto del proceso penal; sobre hipótesis que ayuden a decidir sobre la licitud o ilicitud de los medios de prueba propuestos por las partes; y sobre otros hechos complementarios que ayuden a determinar la fiabilidad o credibilidad de los medios de prueba, es decir, para practicar prueba sobre la prueba.

En segundo término, la IA podría ayudar al tribunal en la apreciación de la prueba practicada sobre los hechos constitutivos de la pretensión penal, así como sobre los relativos a circunstancias eximentes, atenuantes y agravantes.

En tercer lugar, los sistemas expertos podrían resultar útiles tanto para auxiliar a los jueces y tribunales a elaborar la fundamentación jurídica de la sentencia como para resolver los recursos posteriores frente a aquella.

Cada uno de los tres aspectos enunciados nos enfrenta con cuestiones espinosas distintas y con un problema común, a saber, el de la transparencia del algoritmo que actúa como una suerte de red neuronal de la IA.

⁷ Pese a que la definición de actualización judicial automatizada que contiene el anexo al que remite el art. 3 LTICAJ incluye la posibilidad de que el acto en cuestión se realice sin intervención de una persona física, solo cabe entender que lo único permitido es un uso auxiliar o supervisado de la IA. En otro caso, la previsión resultaría contraria a las disposiciones de la CE mencionadas *supra*, tal y como defendimos en otro lugar, vid. CASTILLO FELIPE, *Proceso civil e inteligencia artificial*, cit., p. 265. También, BUENO DE MATA, *Macrodatos, inteligencia artificial y proceso: luces y sombras*, en «Revista General de Derecho Procesal» [en línea], núm. 51, 2020 [consultado por última vez el 8 de noviembre de 2022], disponible en: <<https://acortar.link/gnhfjV>>.

⁸ Existen otras eventuales funciones en la fase procesal que nos ocupa que no afectarían a la labor de enjuiciamiento, por lo que no nos detendremos en ellas. Valga, a modo de ejemplo, la posibilidad de utilizar IA para asegurar la identidad de las partes que intervienen en una audiencia telemática. Escenario que quedaría comprendido dentro de las funciones de auxilio que los sistemas expertos pueden proporcionar a los letrados de la Administración de Justicia. Sobre las potencialidades de la tecnología de reconocimiento facial y la IA puede acudir, con carácter general, a M.J. PÉREZ ESTRADA, *La inteligencia artificial como prueba científica en el proceso penal español*, en «Revista Brasileira de Direito Processual Penal» [en línea], Vol. 7, núm. 2, pp. 1385-1410 [consultado por última vez el 10 de noviembre de 2022], disponible en: <<https://acortar.link/fehWmK>>.

2. Cuestiones problemáticas que podría suscitar la aplicación de la IA en la fase de juicio oral

2.1. El uso de la IA para la obtención de información sobre los hechos que fundan la acusación y la defensa

En principio, un juez equipado con un sistema de inteligencia artificial podría acudir a este para la averiguación de los hechos objeto de enjuiciamiento. Cabe pensar en dos hipótesis: por un lado, en que el sistema revele al juzgador datos, patrones o conexiones fácticas que habían permanecido ocultos o invisibles para las partes; escenario no imposible si la máquina fuera capaz de nutrirse de datos extraídos de la red acerca del hecho punible objeto de enjuiciamiento, relativos a sus autores, a las víctimas o a terceros relacionados con el proceso⁹.

Por otro lado, podemos representarnos el evento de que el juez use la IA para analizar los datos fácticos aportados por las partes mediante la asociación con los datos generales que han servido para entrenar al algoritmo o incluso a partir de la información de otros procesos fallados por el mismo tribunal –no aludiremos ahora mismo al problema del sesgo y a la imposibilidad del programa de desvincularse de las decisiones adoptadas en procesos previos para salvaguardar la presunción de inocencia, singularmente de aquellos que guarden conexión subjetivo-objetiva con el que está siendo enjuiciado–.

En una y otra hipótesis los problemas que pueden surgir son de índole variada:

En primer término, resulta evidente que al juez o tribunal enjuiciador le estaría vedado considerar extremos fácticos relativos al hecho punible que no han sido incluidos en los escritos de calificación (arts. 650.1º y 729.2

⁹ Ciertamente, en este caso, lo lógico sería que a la altura del juicio oral los datos susceptibles de ser obtenidos a partir de sistemas expertos se hayan incorporado ya a la causa; o dicho de otro modo: asumiendo que el algoritmo pueda servir para descubrir aspectos fácticos y no solo para analizarlos, el momento para emplearlo y provocar dicho descubrimiento es la fase de instrucción y no la de juicio oral. Sobre los problemas que pueden surgir en caso de empleo de una misma IA en ambas fases volveremos *infra*. Por lo demás, este primer supuesto no es ciencia ficción, baste ver las posibilidades de análisis de imágenes, archivos, conversaciones mantenidas en chats, etc., que brindan algunos programas en materia de detección de delitos de pornografía infantil, *child grooming* o trata de personas. Al respecto vid. M. LLORENTE-SÁNCHEZ ARJONA, *La inteligencia artificial como nueva estrategia de prevención en los delitos de violencia sexual*, en AA.VV., *Uso de la información y de los datos personales en los procesos*, cit., pp. 259-281.

del Real Decreto de 14 de septiembre de 1882 por el que se aprueba la Ley de Enjuiciamiento Criminal –en lo sucesivo, LECrim–¹⁰. Por ello, en caso de arrojar la IA un hallazgo de esta clase en el juicio oral, habría que acudir necesariamente al mecanismo del art. 746.6º de la norma procesal penal y decretar la suspensión al tiempo que se acuerda una instrucción suplementaria –suponiendo, claro está, que el descubrimiento tenga la entidad suficiente para alterar sustancialmente el objeto del proceso–.

En segundo término, se ha adelantado que la IA podría servir para analizar los hechos objeto de calificación y los hechos auxiliares, estableciendo las asociaciones causales correspondientes entre los hechos del caso y los datos que nutren al sistema o incluso aplicando máximas de la experiencia a aquellos¹¹. Sucedería aquí algo similar a lo que acontece hoy día con los controvertidos informes de inteligencia, utilizados para deducir conclusiones que no se basan única y exclusivamente en los datos y hechos sobre los que versa el proceso¹².

En tal escenario, el problema se traslada al control del funcionamiento del sistema experto y, en particular, a la información manejada por este. Sin perjuicio de las consideraciones que realizamos en un momento ulterior, valga por ahora señalar que, en el supuesto de que el *software* aportase datos de hecho no aducidos por las partes o máximas de la experiencia no introducidas por los peritos de las que dependa la acreditación de la conducta delictiva, surgirían inconvenientes derivados del alcance con el

¹⁰ Imaginemos que, con ocasión del enjuiciamiento de un delito de pertenencia a grupo criminal (art. 570 ter de la Ley Orgánica 1/1995, de 23 de noviembre, del Código Penal –en adelante, CP–), el sistema brinda información al juez de una serie de imágenes publicadas en redes sociales donde se ve juntos a los presuntos integrantes del grupo.

¹¹ En este sentido, con ocasión del estudio de la denominada ‘prueba sobre la prueba’, F. GASCÓN INCHAUSTI, *El control de la fiabilidad probatoria: prueba sobre la prueba en el proceso penal*, Ediciones de la Revista General de Derecho, Valencia, 1999, p. 164, nota 55, ha señalado que el juzgador podría acudir a determinadas fuentes de prueba que no obren en autos, pero que no constituyan investigación en sentido estricto, tal y como sería la consulta a los peritos adscritos al tribunal. Obsérvese, como se indicará más tarde, que la IA podría llegar a difuminar las fronteras entre el perito y el juez al ser capaz de realizar inferencias en el marco de saberes técnicos especializados.

¹² Vid. en este sentido, BARONA VILAR, *Algoritmización del Derecho*, cit., pp. 504 y ss.; V. MORENO CATENA, *Los datos en el sistema de justicia y la propuesta de Reglamento UE sobre inteligencia artificial*, en AA.VV., *Uso de la información y de los datos personales en los procesos: los cambios en la era digital*, cit., pp. 47-73, vid., especialmente, pp. 65 y ss. Sobre la polémica que ha mantenido la jurisprudencia en torno a la naturaleza de los informes de inteligencia hemos consultado F. ORTEGO PÉREZ, *Consideraciones críticas en torno a la denominada prueba pericial de «inteligencia policial»*, en AA.VV. *Peritaje y prueba pericial*, J. PICÓ I JUNOY (dir.), C. DE MIRANDA VÁZQUEZ (COORD.), Bosch, Barcelona, 2017, pp. 583-589.

que la jurisprudencia de nuestro Tribunal Supremo –en lo sucesivo, TS– ha interpretado los poderes del juez penal. Como es sabido, el art. 729.2º LECrim autoriza al juez a practicar pruebas no propuestas por ninguna de las partes que considere necesarias para la comprobación de los hechos que hayan sido objeto de calificación. El alcance de este precepto ha sido discutido y se ha visto reducido por una interpretación hipertrofiada del principio acusatorio y limitadora del principio de oficialidad, de modo que la prueba de oficio se ha circunscrito mayoritariamente a los casos de ‘prueba sobre la prueba’, o, lo que es lo mismo, a aquellas que persigan determinar la fiabilidad de un medio probatorio que ha sido previamente admitido y practicado por las partes (v. gr., obtención de información y perfil de un testigo para probar su parcialidad a la hora de prestar declaración)¹³. En consecuencia, si se estuviera a la interpretación restrictiva enunciada, que no compartimos, el problema surgiría si el uso del algoritmo conlleva que el juez, por extensión, se valga de información fáctica accesoria o máximas de las experiencias no invocadas por las partes para decidir sobre la condena o absolución; esto es, si se traspasan los límites actuales que impone el precepto *supra* citado y las interpretaciones restrictivas del mismo¹⁴.

La cuestión no es baladí. Si se admite que estos sistemas pueden funcionar recurriendo a fuentes no expresamente incorporadas en la causa para suministrar conclusiones al juez, habrá que convenir que la IA puede ser fuente, pero también medio de prueba y, por ende, la posibilidad de su uso adquiere también relevancia desde la óptica del debate sobre el alcance de los poderes del juez penal¹⁵. En nuestra opinión, sería absurdo señalar que el juez que recurra a un algoritmo que tome en consideración fuentes accesorias y de carácter general, no incorporadas expresamente a la causa,

¹³ Vid. GASCÓN INCHAUSTI, *El control de la fiabilidad probatoria*, cit., pp. 161 y ss.; J. TOMÉ GARCÍA, *Fase decisoria (II). La prueba*, en A. DE LA OLIVA SANTOS, A. et. al., *Derecho Procesal Penal*, 8ª ed., Centro de Estudios Ramón Areces, Madrid, 2008, pp. 507-508; A. DE LA OLIVA SANTOS, *El papel del juez en el proceso civil*, Thomson-Civitas, Cizur Menor (Navarra), 2012, p. 30, nota 7; ARMENTA DEU, *Lecciones de Derecho Procesal penal*, 9ª ed., Marcial Pons, Madrid, 2015, pp. 282-283.

¹⁴ Dejando al margen el hecho de que lo idóneo sería el uso de la IA que sea capaz de suministrar máximas de la experiencia en instrucción, pensemos en la hipótesis de un sujeto acusado por la comisión de un delito de falso testimonio (art. 457 CP) y un juez auxiliado por un sistema que sea capaz de aplicar automáticamente las reglas de la psicología del testimonio. Conforme a la interpretación restrictiva anotada, el juez perdería su imparcialidad al recurrir a la IA, que bien podría ser una herramienta que utilice habitualmente en la toma de declaraciones a los testigos.

¹⁵ El carácter de fuente de prueba de los instrumentos tecnológicos que predecían riesgos lo ha anotado ya BARONA VILAR, *Algoritmización del Derecho*, cit., pp. 502 y 588.

está perdiendo su imparcialidad y asumiendo el rol de parte practicando prueba de cargo o descargo. De ahí que rechazemos la exclusión del empleo de algoritmos en el enjuiciamiento señalando que el tribunal sería a la vez juez y parte. Y es que, si bien se mira, la clave volverá a estar, en este caso, en las cautelas relativas a la utilización de la máquina, singularmente en la posibilidad de conocer y contradecir los resultados arrojados por el sistema a partir del acceso a sus paquetes de datos, modo de entrenamiento y funcionamiento real.

En tercer y último lugar, el recurso a la IA podría generar problemas de ilicitud probatoria derivados de un indebido tratamiento o acceso a los datos manejados por el programa informático (art. 11.1 LOPJ)¹⁶. Ilicitud que, en función del caso, quizás no fuera tal, sobre todo teniendo en cuenta la última jurisprudencia y pronunciamientos constitucionales en esta materia¹⁷. Conforme al margen que proporcionan los mismos, podría plantearse si el indebido tratamiento de datos por parte de un *software* de IA creado por una empresa privada –un particular, a los efectos que nos interesa– genera *per se* ilicitud probatoria o, por el contrario, podría ser objeto de ‘desconexión de antijuridicidad’¹⁸. En particular, si se piensa en que el tratamiento ilícito realizado por la IA no tenga como finalidad específica el enjuiciamiento penal.

2.2. El recurso a la IA como prueba para determinar la fiabilidad o credibilidad de los medios de prueba practicados en el juicio oral

Si seguimos razonando en la línea anotada *supra*, la IA podría ser útil al

¹⁶ Que estos sistemas pueden generar problemas para el derecho a la intimidad, el denominado «derecho al entorno virtual» y el derecho a la protección de datos es algo que ha venido advirtiendo la doctrina, vid. NIEVA FENOLL, *Inteligencia artificial y proceso*, cit., pp. 150-152; BUENO DE MATA, *Macrodatos, inteligencia artificial y proceso*, cit., pp. 4 y ss. del documento en formato pdf; L. ESTÉVEZ MENDOZA, *Algoritmos policiales basados en IA y derechos fundamentales a la luz de HART y VALCRI*, en AA.VV., *Justicia: ¿garantías versus eficiencia?*, F. JIMÉNEZ CONDE, R. BELLIDO PENADÉS (dirs.); P. LLOPIS NADAL, E. DE LUIS GARCÍA (coords.), Tirant Lo Blanch, Valencia, 2019, pp. 663-675, vid. p. 673.

¹⁷ Vid. J.M. ASENCIO MELLADO, *La STC 97/2019, de 16 de julio. Descanse en paz la prueba ilícita*, en «Diario la Ley» [en línea], núm. 9499, 2019 [consultado por última vez el 10 de noviembre de 2022], disponible en: <<https://laleydigital.laleynext.es/Content/Inicio.aspx>>.

¹⁸ El riesgo de rebaja de los estándares de protección es uno de los peligros que comporta la externalización del proceso penal, tal y como ha puesto de relieve GASCÓN INCHAUSTI, *Desafíos para el proceso penal en la era digital: externalización, sumisión pericial e inteligencia artificial*, en AA.VV., *La justicia digital en España y la Unión Europea*, Atelier, Barcelona, 2019, pp. 191-206.

juez para practicar prueba sobre la prueba con el fin de valorar la fiabilidad del resto de medios propuestos (art. 729.3º LECrim)¹⁹. En esta dirección se ha señalado que el sistema informático podría servir para valorar el lenguaje no verbal de los testigos y peritos que prestasen declaración en el juicio oral²⁰.

El tratamiento especial que la ley dispensa al uso de datos de carácter biométrico podría constituir un obstáculo al empleo de sistemas expertos para tal fin. Si bien, la Ley Orgánica 7/2021, de 26 de mayo, permite utilizar estos datos en el marco de un proceso judicial bajo ciertas condiciones²¹ – pensemos, por ejemplo, que, por entenderlo conveniente para su defensa, el acusado muestre su conformidad con el uso de estas herramientas durante su declaración, de manera similar a lo que ha sucedido con la práctica de pruebas neurocientíficas como el *brainfingerprinting*²².

En los casos de uso auxiliar de la IA para determinar la fiabilidad de los

¹⁹ Para un estudio detallado de la prueba sobre la prueba cfr., de nuevo, GASCÓN INCHAUSTI, *El control de la fiabilidad probatoria*, cit., *passim*.

²⁰ Vid. NIEVA FENOLL, *Inteligencia artificial y proceso*, cit., pp. 80 y ss., quien da cuenta del intento llevado a cabo con el programa ADVOKATE; también, BARONA VILAR, *Algoritmización del Derecho*, cit., pp. 591 y ss.; J. BONET NAVARRO, *Valoración de la prueba y resolución mediante inteligencia artificial*, en AA.VV., *Derecho Procesal. Retos y transformaciones*, BUJOSA VADELL, L. (dir.), Atelier, Barcelona, 2021, pp. 315-337, vid. pp. 325-327; V. GUZMÁN FLUJA, *Proceso penal y justicia automatizada*, en «Revista General de Derecho Procesal» [en línea], núm. 53, 2021 [consultado por última vez el 10 de noviembre de 2022], disponible en: <<https://acortar.link/wkVWzh>>; PÉREZ ESTRADA, *La inteligencia artificial como prueba científica en el proceso penal español*, cit., pp. 1398 y ss.; MORENO CATENA, *Los datos en el sistema de justicia y la propuesta de Reglamento UE sobre inteligencia artificial*, cit., p. 66.

²¹ Vid., en lo que aquí interesa, los arts. 13 y 14 de la norma citada. En torno a estos preceptos, resultan de interés las reflexiones de GUZMÁN FLUJA, *Proceso penal y justicia automatizada*, cit., pp. 24 y ss. del documento en formato electrónico. Aunque son anteriores a la transposición de la Directiva 680/2016, relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales por parte de las autoridades competentes para fines de prevención, investigación, detección o enjuiciamiento de infracciones penales o de ejecución de sanciones penales, y a la libre circulación de dichos datos y por la que se deroga la Decisión Marco 2008/977/JAI del Consejo, siguen siendo plenamente pertinentes, ya que los aludidos preceptos de la LO 7/2021, de 26 de mayo, se limitan a transcribir los arts. 10 y 11 de la Directiva.

²² M.L., VILLAMARÍN LÓPEZ, *Neurociencia y detección del engaño en el proceso penal. El uso del escáner cerebral (fMRI) y del brainfingerprinting (P300)*, Marcial Pons, Madrid, 2014, pp. 137 y ss., ha dado cuenta de la utilización de las técnicas reseñadas en nuestros tribunales. Por su parte, NIEVA FENOLL, *Inteligencia artificial y proceso*, cit., pp. 89-90, repara en la cuestión *supra* citada indicando que estos métodos trabajan ya con IA, aunque todavía no revisten un grado de fiabilidad suficiente.

medios de prueba, cabe preguntarse, una vez más, qué es lo que proporcionaría la IA al juez, y la respuesta no puede ser otra que información científica especializada. En efecto, tanto la medición de los fenómenos fisiológicos que se producen durante la declaración de un sujeto como el procesado de dicha información se basa en máximas de la experiencia que aquel desconoce (por ejemplo, de psicología del testimonio)²³. Sin duda, en estos casos, se difuminaría notablemente la línea que tradicionalmente ha separado al juez y al perito y el sistema pasaría a convertirse en una suerte de auxiliar virtual²⁴. De ser esto así podrían surgir inconvenientes para las partes, sobre todo si ellas no pueden interactuar con esta fuente dinámica de procesamiento de la información. A un perito pueden interrogarle e incluso carear al propio con el de la parte contraria: ¿podrán hacer lo mismo con el sistema de IA?

2.3. El empleo de la IA para decidir sobre la ilicitud de las diligencias de investigación o medios de prueba practicados

Como regla general, la aplicación de la IA en la fase de juicio oral servirá para enjuiciar hechos pasados, pues no cabe pensar en un proceso penal cuyo objeto principal sea la represión de conductas futuras, sin perjuicio de que, en ciertos estadios del proceso, resulte necesaria la presciencia –v. gr., para la valoración sobre el riesgo de fuga [arts. 503.3º a) LECrim] o la conducta esperable del condenado si se le concede la libertad condicional (art. 90 CP)–. Con todo, el uso de sistemas expertos podría servir para resolver ciertos enjuiciamientos que requieren la inferencia de hechos futuros en el marco del juicio oral. Pensamos, sobre todo y sin perjuicio de la decisión en sentencia sobre la suspensión de la pena (art. 80 CP) y la responsabilidad civil *ex delicto*, en la resolución de cuestiones sobre ilicitud de la prueba (art. 11.1 LOPJ). En particular, en la teoría del hallazgo inevitable, que permite excluir la ‘conexión de antijuridicidad’ de aquella: habrá que convenir que valorar si la fuente de prueba se hubiera descubierto igual sin la práctica de la diligencia que violó un derecho fundamental requiere de una predicción o cálculo probabilístico en el que la máquina podría ayudar al juez²⁵.

²³ Se detiene en la posible aplicación de la psicología del testimonio a través de la IA, NIEVA FENOLL, *Inteligencia artificial y proceso*, cit., pp. 81 y ss.

²⁴ Extremo que no ha pasado inadvertido a GASCÓN INCHAUSTI, *Desafíos para el proceso penal*, cit., p. 201.

²⁵ Realizamos esta aseveración siendo conscientes de que, para que pudiera convertirse en realidad, sería necesario que la IA hubiera alcanzado el tercer peldaño en la escala de la causalidad, es decir, que fuera capaz de trabajar con contrafactuals o hipótesis alternativas.

2.4. *El uso de la IA en el juicio de apreciación de la prueba*

La tecnología estudiada podría emplearse en el futuro para auxiliar al tribunal en la apreciación de la prueba²⁶. Mientras que el juicio de interpretación, esto es, la determinación de qué información fáctica –con independencia de su verdad o falsedad– se extrae de cada probanza, no creemos que pueda ser arrancado de manos del juez, pues nada habría más peligroso que un juez que dicte sentencia sin ni siquiera haber entendido cuál es la información que arroja una prueba de cargo o de descargo, el juicio de valoración, tendente a fijar cuál es el peso de un medio de prueba a la hora de determinar la certeza positiva o negativa de un hecho, sí que sería susceptible, en parte, de automatización.

En nuestra opinión, estas premisas resultan relevantes para ordenar los factores de la ecuación tribunal-IA. De este modo, tras llevar a cabo el juicio de interpretación, el tribunal habrá de comenzar por valorar qué credibilidad otorga al sistema experto, que funcionaría aquí como una prueba sobre la prueba tendente a fijar la fiabilidad del resto de medios²⁷. Solo si ese juicio fuera positivo merecerá la pena plantearse el resto de

En cuanto al tipo de enjuiciamiento que requiere la citada conexión indica V. LÓPEZ YAGÜES, *La prueba prohibida*, en AA.VV., *Derecho Procesal penal*, ASECIO MELLADO, J. M. (dir.), FUENTES SORIANO, O. (coord.), Tirant Lo Blanch, Valencia, 2021, p. 166, que la existencia o no de conexión de antijuridicidad no constituye un hecho, sino un ‘juicio de experiencia’.

²⁶ Sobre los juicios que componen la apreciación de la prueba y sus diferencias, vid. F. JIMÉNEZ CONDE, *La apreciación de la prueba legal y su impugnación*, Publicaciones del Departamento de Derecho Procesal de la Universidad de Salamanca, Salamanca, 1977, pp. 121 y ss.

²⁷ Para un análisis del régimen de la valoración de la prueba sobre la prueba, vid. GASCÓN INCHAUSTI, *El control de la fiabilidad probatoria*, cit., pp. 43 y ss., donde el autor señala que resulta más realista valorar la prueba sobre la prueba en segundo lugar, aunque reconoce que también sería lógico realizarlo a la inversa. La postura tildada como realista sería pues la contraria a la defendida *supra*. El sustento de este posible modo de valoración derivaría de la posibilidad de evitar la práctica de prueba inútil en los supuestos en que no existan dudas sobre la fiabilidad de los medios de prueba directa practicados. En nuestra opinión, si se piensa, como hemos hecho nosotros, en un escenario en el que el juez tenga a su disposición de modo continuo un sistema de IA para ayudarle, acoger la segunda interpretación propuesta por el brillante procesalista citado restringiría enormemente las posibilidades de uso. Y es que entendemos que si algún día este tipo de artificios tecnológicos llegan a estar disponibles en los tribunales, lo será porque hayan sido testados y exista un convencimiento de que pueden mejorar o ayudar a mejorar el resultado del enjuiciamiento humano con carácter general. Asimismo, cuando el juez no tenga duda inicial sobre la fiabilidad de la prueba practicada no creemos que pueda ser calificado de prueba inútil el recurso a la IA, ya que siempre podría servirle para refrendar la validez de sus propias conclusiones.

problemas que en esta sede estamos tratando.

Así pues, como se ha adelantado anteriormente, en lo que respecta a la valoración individual de la prueba, la doctrina ha puesto de relieve que la IA podría ayudar en la aplicación de los criterios *Daubert*, acuñados por la jurisprudencia estadounidense para decidir sobre la admisibilidad de la prueba pericial y empleados también para su valoración, especialmente, aportando datos externos sobre la técnica utilizada por los peritos, su referendo científico, la habitualidad en su uso y sus márgenes de error²⁸.

Esta labor auxiliar podría darse también respecto de la valoración individual de la prueba documental, de testigos y de declaración del acusado –en este último caso, con especial cuidado en la salvaguarda de los derechos que asisten a este–²⁹. Además, la IA podría brindar ayuda al juez respecto a la valoración conjunta de la prueba y, singularmente, a la hora de fijar la certeza de los hechos mediante el recurso a la prueba de indicios.

A la vista de lo anterior, no nos parece descabellado preguntarnos ¿hasta qué punto usar IA para valorar la prueba no implicaría retornar a un sistema muy semejante al de valoración legal o tasada? En efecto, en estos casos, las premisas que un día informaron el silogismo sobre el descansaban las normas de valoración legal serían sustituidas por el código binario configurado por el programador del sistema experto³⁰; sistema que normalmente se acomodará a directrices de *soft law* o instrucciones técnicas de inexcusable observancia³¹. Estas, al igual que hiciera otrora la premisa mayor de las normas de valoración legal, pueden acabar por secuestrar la libertad enjuiciadora del tribunal.

Somos conscientes de que, mientras rija la regla de libre valoración (art. 741 LECrim), el juez podría apartarse de lo decidido por la IA, pero ello implicará un esfuerzo adicional de motivación que puede no estar dispuesto a hacer o que, simplemente, se presente como imposible al no ser capaz de motivar en contra de un lenguaje y reglas informáticas que

²⁸ Vid. NIEVA FENOLL, *Prueba científica. Cuestiones de futuro: neurociencia e inteligencia artificial*, cit., pp. 493-494; *Id.*, *Inteligencia artificial y proceso*, cit., pp. 96-98.

²⁹ Remitimos en este punto una vez más a la bibliografía reseñada en la nota al pie 20 de este trabajo.

³⁰ En este sentido, BARONA VILAR, *Algoritmización del Derecho*, cit., p. 624, ha anotado que «asistimos a la conversión del algoritmo en ley». Por su parte, NIEVA FENOLL, *Inteligencia artificial*, cit., p. 81, ha advertido que «el uso indiscriminado de la estadística puede acabar propiciando el renacer de algo parecido a las pruebas legales, y ese resultado debería ser evitado en todo caso».

³¹ Vid., en este punto, por todos, las acertadas consideraciones de ARMENTA DEU, *Derivas*, cit., pp. 316 y ss.

desconoce –singularmente, en sistemas con un funcionamiento hartamente complejo³². Surge así el problema, que se adelantó *supra*, de la sumisión judicial al algoritmo transformado en ley. Y es que si actualmente el juez se halla en muchas ocasiones abrumado por las pruebas periciales, sin verdaderas herramientas o criterios para evaluar críticamente la opinión de los expertos, igual fenómeno se reproducirá al acudir a la IA. De hecho, a medida que el recurso al sistema sea mayor y este mejore más difícil puede tornarse para el tribunal alejarse de las conclusiones que arroje la máquina, sobre todo si no se sabe bien cómo funciona³³.

Parte de la doctrina rechaza la conclusión anterior señalando que únicamente se estaría ante la objetivación de decisiones probatorias o fijación de estándares y no ante un sistema similar al de valoración legal³⁴. No podemos aceptar el aserto en toda su amplitud. Convenimos que no existirá valoración tasada –no la puede haber allí donde se reconoce la

³² Las palabras de MURILLO FUENTES, ¿Qué es lo que no funciona en los algoritmos de inteligencia artificial?, cit., p. 164, sobre los sistemas de *deep learning* resultan plenamente pertinentes para ilustrar lo expuesto: «El desarrollo del DL ha permitido grandes avances recientes en un sin fin de tareas. Sin embargo, es conocido que tras la ejecución del algoritmo para entrenar el modelo, no es posible saber realmente en qué se basa el modelo para tomar decisiones y sacar una predicción. Los modelos son muy potentes, pero está documentado el fallo de los mismos ante entradas con determinados cambios que, aparentemente, no deberían de afectar». En la p. 168, el autor ahonda de nuevo en esta idea señalando: «Con un modelo de DL alcanzamos la mejor de las exactitudes posibles en el estado del arte, pero DL se comporta como una caja negra en la que poco se sabe de la relación entre la entrada y la salida ni de por qué la red ha terminado configurada de una determinada forma. Es más (...) es posible alterar levemente la entrada, en una imagen de forma no perceptible, resultando en un cambio en la predicción, sin que se puede explicar bien por qué» (*sic.*). Un ejemplo ilustrativo lo proporciona M. PEIRANO, *El enemigo conoce el sistema*, Penguin Random House, Barcelona, 2019, pp. 139-140, quien da cuenta del experimento realizado por Facebook en el que dos IA negociaron sobre el valor de bienes de uso común –libros, balones, etc.– y acabaron desarrollando un dialecto propio, incomprensible para sus programadores.

³³ Estaríamos ante la paradoja de la prueba científica, vid. G. TUZET, *Filosofia della prova giuridica*, 2ª ed., Giappichelli Editore, Turín, 2016, pp. 206 y ss., cuya valoración depende de un juez que no conoce el saber sobre el que versa la prueba y que puede tener que apreciar esta sin llegar a entender bien aquello que la concreta ciencia le ofrece. Este mismo problema ha sido anotado por L. MUÑOZ SABATÉ, *La prueba de indicios en el proceso judicial*, Wolters Kluwer, Madrid, 2016, p. 86; y, en el específico campo de la prueba neurocientífica, por VILLAMARÍN LÓPEZ, *Neurociencia*, cit., pp. 149-150. En esta paradoja debe buscarse, a nuestro juicio, el origen del fenómeno de la sumisión pericial identificado por la doctrina, vid. GASCÓN INCHAUSTI, *Desafíos para el proceso penal*, cit., pp. 109 y ss.; ARMENTA DEU, *Derivas*, cit., p. 305; BARONA VILAR, *Algoritmización de la justicia*, cit., p. 617, nota 1096.

³⁴ GUZMÁN FLUJA, *Proceso penal y justicia automatizada*, cit., pp. 19-20.

libertad enjuiciadora del tribunal—. También asumimos que es posible que la IA funcione mediante generalizaciones susceptibles de ser consideradas estándares. Ahora bien, no pensamos que cualquier resultado de la IA pueda ser elevado a tal categoría, sobre todo admitiendo la posibilidad de desviación, evolución o involución del *software*.

Sostenemos, por tanto, que los términos probabilísticos en los que se presentan los resultados y la dificultad de contrargumentarlos en un plano que no es el matemático o estadístico crearán, de facto, una vinculación, si se quiere velada, a la decisión automática. A buen seguro esta podrá servir también como excusa para enjugar eventuales responsabilidades en el enjuiciamiento³⁵.

Ahondando un poco más en la cuestión de la fijación de estándares a través de la IA, la doctrina ha estudiado la posibilidad de fijar algorítmicamente cuándo se supera toda duda razonable y, por ende, podría entenderse que se está ante un grado de certeza suficiente para enervar la presunción de inocencia³⁶; en otras palabras, se trataría de que la máquina fijase un porcentaje por encima del cual se excluiría la aplicación de la regla *in dubio pro reo*.

Evidentemente, poder objetivar este punto liberaría al juez penal de una pesada carga. Pero esa liberación sería tan solo aparente, pues el verdadero convencimiento interior del tribunal no consideramos que pueda expresarse porcentualmente y mucho menos la falta de él. Imaginemos que el sistema experto fuera capaz de cuantificar en porcentajes el grado de persuasión del juez necesario para dictar sentencia de condena y este se fija en un 90%. La mera existencia del porcentaje no evitará que el tribunal pueda estar plenamente convencido de la culpabilidad del acusado en un escenario en el que el programa arrojase un porcentaje del 89% o seguir teniendo dudas en la hipótesis de que el valor porcentual sea superior al 90%³⁷. Lo anterior es explicable porque el Derecho y la matemática operan en dos planos diferentes. El estado de duda humano es eminentemente subjetivo y valorativo y arroja matices que desaparecen en la objetivación a la que conduce la segunda. Aun cuando deba buscarse siempre la racionalidad

³⁵ La práctica de fiar una decisión a un algoritmo con funcionamiento opaco para evitar cualquier tipo de responsabilidad en la toma de decisiones se ha localizado ya en otros ámbitos y se denomina *math washing*. Vid. PEIRANO, *El enemigo conoce el sistema*, cit., p. 138.

³⁶ Vid., de nuevo, NIEVA FENOLL, *Inteligencia artificial y proceso*, cit., pp. 105-114; BARONA VILAR, *Algoritmización del Derecho*, cit., pp. 618-619.

³⁷ M. TARUFFO, *Evidence and Truth in Procedural Law*, en AA.VV., *La prueba en el proceso*, cit., pp. 31-47, vid. p. 38, advirtió que resultaban inútiles los intentos de cuantificación de los estándares probatorios, habida cuenta de la imposibilidad de medir con una escala estadística o numérica la culpabilidad del acusado.

de la decisión judicial, esta no tendría que convertirse en una excusa para engrilletar el criterio del juez al de la IA³⁸. Y es que, como se ha anotado con ocasión del análisis de las pruebas probabilísticas y estadísticas, muchas veces los números distraen en torno a la fiabilidad de la prueba y los datos sobre los que esta se construye³⁹. Aserto que es plenamente extrapolable a la cuestión que nos ocupa.

Al margen de lo anterior, la aplicación del sistema experto al enjuiciamiento puede ofrecer otros problemas relacionados con la presunción de inocencia y el derecho al juez no prevenido. Cuando la IA se aplique a una valoración conjunta de la prueba es indispensable preguntarse qué valora el algoritmo, quién lo usa y cuándo se usa. De este modo: ¿hasta qué punto estaría en riesgo el principio del juez no prevenido si el mismo algoritmo empleado en la fase de instrucción se usase más tarde por un juez distinto para dictar sentencia? ¿Y en el caso de que el mismo algoritmo fuera empleado por el Ministerio Fiscal para preparar su acusación y más tarde por el tribunal para juzgar? Si el algoritmo aprende, o simula que lo hace, mediante el procesado de los datos que incorpora o que se le van suministrando, y se le ha cebado con los datos del supuesto de hecho que va a ser objeto de enjuiciamiento en instrucción, el riesgo de que el programa esté ya aleccionado o incluso parta de la conclusión alcanzada en la fase de investigación es totalmente real –v. gr., si se acudió a él para decidir si procedía o no la apertura de juicio oral–. Mientras que el ser humano utiliza varios individuos con potestad jurisdiccional y funciones procesales diversas para respetar la verdad interina que implica la presunción de inocencia, es difícil que un sistema experto pudiera cambiar la conclusión previamente tomada. De ahí el riesgo de contaminación «indirecta o refleja» del órgano de enjuiciamiento con datos de la instrucción previamente manejados por el algoritmo.

Para concluir con las cuestiones relativas a la aplicación de la IA y la valoración de la prueba, podemos preguntarnos qué sucederá con los juicios por jurado. Especialmente teniendo en cuenta que, en 1995, España

³⁸ En relación con este punto, NIEVA FENOLL, *Inteligencia artificial*, cit., p. 115, ha sostenido que el juez será incapaz de motivar por qué la hipótesis de la máquina no le convence, lo cual, en su opinión, supone un mal presagio para la prevalencia del factor humano sobre la IA.

³⁹ S. HAACK, *Proof, Probability, and Statistics: The Dangers of 'delusive exactness'*, en AA.VV., *La prueba en el proceso*, cit., pp. 497-519. También BARONA VILAR, *Algoritmización del Derecho*, cit., p. 610, recuerda que «en los métodos tecnológicos no debe considerarse que concurre una suerte de veracidad implícita con resultados automáticos en la decisión judicial».

optó por implantar un modelo de jurado puro. En principio, la respuesta inmediata impone desterrar cualquier tentativa de aplicación de la IA en estos procesos penales, so pena de pervertir y desnaturalizar la actividad de apreciación de la existencia o inexistencia de los hechos justiciables conforme al sentido común (art. 3 Ley Orgánica 5/1995, de 22 de mayo, del Tribunal del Jurado –en lo sucesivo, LOTJ–)⁴⁰.

Ahora bien, asumiendo un escenario en el que la IA funcionara correctamente y fuera verdaderamente útil, ¿tendría sentido impedir al jurado que se auxiliase de esta herramienta?

Sin duda, el riesgo de seguimiento ciego de los dictados del algoritmo se acentuaría en el caso de los ciudadanos legos que son llamados a participar en la Administración de Justicia, mucho más, si cabe, que respecto de los magistrados profesionales. En este sentido, el hecho de que el canon de motivación del acta del veredicto no sea tan intenso como el de un tribunal profesional no ayudaría precisamente a evitar el inconveniente citado [61.1.d) LOTJ]. Sin embargo, si algún día se colige que el uso de herramientas informáticas para auxiliar a la labor de juzgar permite alcanzar una justicia más fina o depurada, parece que no tendría sentido mantener al jurado privado de esas facilidades. En otro caso, el rechazo de los avances tecnológicos en este campo –insistimos, suficientemente testados– implicaría convertir a este singular tribunal en una antigualla o anacronismo que a ojos de la sociedad podría quedar deslegitimado como forma de participación en la Administración de Justicia⁴¹.

Para solucionar el problema cabría pensar en que se instruyese al jurado sobre el manejo de la herramienta en el trámite de información previa o, incluso, desde una perspectiva más realista, en cambios en el modelo del jurado, hacia un modelo escabinado, si es que entendemos que el juez es el único que debe manejar estos dispositivos por su formación específica.

⁴⁰ La doctrina anota que cuanto mayor es la presencia de la prueba científica en el proceso, menos espacio resta para el sentido común, ya que aquella desplaza con sus reglas generales y muchas veces constatadas las inferencias o conclusiones que pueden alcanzarse a través de este. Vid. TUZET, *Filosofía della prova*, cit., p. 196. Desde esta óptica el propio papel del jurado podría incluso llegar a cuestionarse.

⁴¹ La penetración de la IA en el tribunal del jurado ofrece numerosas aristas que, sin duda, serían dignas de consideración (v.gr., la utilidad de estos sistemas en la fase de selección de los jurados –terreno donde ya existen experiencias– o incluso en la conformación del objeto del veredicto). La extensión máxima de este trabajo nos impide detenernos en ellas.

2.5. *La aplicación de la IA en la elaboración de la fundamentación jurídica de la sentencia y en la resolución de recursos posteriores frente a esta*

Además de asistir al juez en la valoración probatoria, la IA podría utilizarse para ayudar al tribunal en el dictado de la sentencia. De hecho, ya existen programas capaces de construir argumentaciones de manera autónoma⁴². En este caso, el riesgo principal es, sin duda, el sesgo de automatización o, lo que es lo mismo, que el juez siga de manera acrítica y ciega la propuesta de resolución emitida por el sistema experto.

Se podría pensar que esto no sucederá. Empero, existen factores que invitan a colegir lo contrario: por un lado, la archiconocida carga de trabajo que pesa sobre nuestros tribunales; por otro lado, el hecho de que, en España, el sueldo de los jueces se compone de una parte de retribución fija y de otra variable (art. 403.2 LOPJ y Ley 15/2003, de 26 de mayo)⁴³. El monto variable se liga a la productividad, es decir, al número de sentencias dictadas en los tiempos estimados en el Reglamento del Consejo General del Poder Judicial 2/2018⁴⁴. De este modo, la propuesta automatizada de solución podría servir al juez para ahorrar trabajo e incrementar el cobro de incentivos y, en el peor de los casos, para que este omita cualquier intento de razonar sobre el caso. Todo ello disfrazado bajo un falso prisma de eficiencia confundida con rapidez, que es hoy idea recurrente al abordar los males de la Administración de Justicia española⁴⁵.

⁴² Vid., una vez más, NIEVA FENOLL, *Inteligencia artificial*, cit., pp. 28-31 y 115; BARONA VILAR, *Algoritmización del Derecho*, cit., pp. 642 y ss.

⁴³ Sobre la cuestión, aunque anterior a la promulgación del actual Reglamento de 2018, sigue siendo ilustrativo el trabajo de G. DOMÉNECH PASCUAL, *Juzgar a destajo*, Thomson-Civitas, Cizur Menor (Navarra), 2009, *passim*.

⁴⁴ Vid. el Acuerdo de 29 de noviembre de 2018, del Pleno del Consejo General del Poder Judicial, por el que se aprueba el Reglamento 2/2018, para el cumplimiento de la Ley 15/2003, de 26 de mayo, reguladora del régimen retributivo de las Carreras Judicial y Fiscal, en lo relativo a las retribuciones variables por objetivos de los miembros de la Carrera judicial, publicado en el BOE [en línea], núm. 304, de 18 de diciembre de 2018, pp. 124070 a 124122 [consultado por última vez el 14 de noviembre de 2022], disponible en: <<https://www.boe.es/buscar/doc.php?id=BOE-A-2018-17295>>.

⁴⁵ En otro lugar, con ocasión de algunas sugerencias de reforma de la justicia civil, hemos señalado, en clara analogía con lo que sucede con los restaurantes de comida rápida, que la conversión de la rapidez en el objetivo primordial de las reformas procesales –y en la actuación de los tribunales– podría dar lugar a un modelo de ‘fast justice’ o ‘junk justice’, en el que lo cuantiativo desplaza a lo cualitativo. Vid. CASTILLO FELIPE, *Medidas para una justicia más eficiente?: limitación de escritos y sentencias orales*, en AA.VV., *Derecho y pandemia desde una perspectiva global*, M. ESPEJO LERDO DE TEJADA; J. GARCÍA OLIVA; J. MARTÍNEZ CRUZ; J.P. MURGA FERNÁNDEZ (dirs.), A. ANDRADES NAVARRO (coord.), Thomson-Aranzadi, Cizur Menor (Navarra), 2021, pp. 135-150.

El seguimiento judicial acrítico —o supervisión meramente formal— de las propuestas lanzadas por el *software* puede incidir negativamente en el deber de motivación de las sentencias que integra el derecho a la tutela judicial efectiva (arts. 120.3 CE y 24.1 CE)⁴⁶.

En primer término, si el programa selecciona fundamentos jurídicos equivocados o incluye en la motivación ‘retazos’ de sentencias no relacionadas directamente con el caso o que, separados del contexto y supuesto de hecho en el que fueron dictados, puedan inducir a confusión. Evidentemente, el ‘corta y pega’ en las sentencias existe actualmente⁴⁷. Cuando el fundamento indebidamente reproducido no aporta nada a la argumentación y al sentido del fallo, el yerro no debería adquirir una dimensión invalidante de lo resuelto. Ahora bien, si la selección de la norma es desacertada, arbitraria, o si el fragmento de jurisprudencia embutido automáticamente en la sentencia puede variar el sentido de esta o introducir matices relevantes, el problema adquiere otra dimensión, ya que el derecho a la tutela judicial efectiva (art. 24.1 CE) quedaría lesionado⁴⁸.

⁴⁶ GUZMÁN FLUJA, *Proceso penal y justicia automatizada*, cit., p. 35, ha insistido en la necesidad de que la actividad de supervisión humana tenga un alcance material y no meramente formal.

⁴⁷ Lo anotan también NIEVA FENOLL, *Inteligencia artificial*, cit., p. 117; BARONA VILAR, *Algoritmización del Derecho*, cit., p. 627.

⁴⁸ Una síntesis de la jurisprudencia constitucional en esta materia puede verse en la sentencia del Tribunal Constitucional —en adelante, TC—, núm. 63/2021, de 15 de marzo [ECLI:ES:TC:2021:63], fundamento jurídico 3º, dictada, en un expediente administrativo sancionador, con ocasión de la inaplicación de normas derecho transitorio que determinaban la posibilidad o no de practicar notificaciones electrónicas: «la resolución ha de estar suficientemente motivada, es decir, contener los elementos y razones de juicio que permitan conocer cuáles han sido los criterios jurídicos que fundamentan la decisión; y, en segundo lugar, que la motivación debe contener una fundamentación en Derecho, lo que conlleva la garantía de que la decisión no sea consecuencia de una aplicación arbitraria de la legalidad, ni resulte manifiestamente irrazonada o irrazonable, incurso en un error patente o en una evidente contradicción entre los fundamentos jurídicos, o entre éstos y el fallo, ya que, en tal caso, la aplicación de la legalidad sería tan sólo una mera apariencia. Esto es, que no puede considerarse cumplida esta exigencia de fundamentación jurídica con la mera emisión de una declaración de voluntad en un sentido u otro, sino que la decisión adoptada debe ser consecuencia de una exégesis racional del ordenamiento y no fruto de la arbitrariedad. No basta, pues, con una apariencia de motivación; además es preciso que la misma tenga contenido jurídico y no resulte arbitraria, de tal modo que una resolución judicial podrá tacharse de arbitraria cuando, aun constatada la existencia formal de una argumentación, la misma no sea expresión de la administración de justicia, sino simple apariencia de la misma, por ser fruto de un mero voluntarismo judicial o de un proceso deductivo irracional o absurdo». Por lo demás, el propio TC recurre las más de las veces a extractos de sus resoluciones anteriores, hasta el punto de que, en alguna ocasión, ha reinterpretado su propia doctrina a partir de fragmentos individuales y descontextualizados

En segundo lugar, aun cuando el algoritmo acertase en la selección de la norma aplicable, el derecho a la tutela judicial efectiva en su vertiente a obtener una resolución motivada y fundada en derecho (art. 24.1 CE) podría conculcarse si se motivase por remisión al algoritmo. En este caso, el juez no habría exteriorizado las razones que le llevarían a acoger el planteamiento sugerido por la IA. En consecuencia, las partes –y también el resto de la sociedad– desconocerían cómo se ha formado la convicción del tribunal y, sobre todo, si esta obedece a un seguimiento mecánico de las herramientas informáticas. Para ilustrar la cuestión, imagínese una motivación que señalase que la conclusión X se ha alcanzado ‘tras la acertada valoración realizada por el sistema informático de ayuda decisoria’. Sin duda, una motivación de este tenor contravendría la prohibición de emplear motivaciones meramente aparentes o formales establecida por el TC⁴⁹.

Dejando a un lado los problemas concretos de motivación, y entrando en el terreno de la resolución de recursos por parte de tribunales superiores, hay que señalar que el seguimiento a pies juntillas de lo dispuesto por la IA podría anquilosar tanto la jurisprudencia como la denominada ‘jurisprudencia menor’⁵⁰. Téngase en cuenta que la tecnología que nos ocupa funciona sobre la base del tratamiento repetitivo de cientos de supuestos. No obstante, cuando es preciso el plus de creatividad –inexistente en la IA, aunque quizás pueda lograrse en el futuro si se logra avanzar en la escala de la causalidad– para modular la aplicación de las normas o practicar el juicio de subsunción, los sistemas expertos ofrecen limitaciones de importancia. Por ello, consideramos que el uso de esta tecnología para fijar la aplicación

de sentencias anteriores para que estas dijeran lo que no decían. Vid., en este sentido, las consideraciones de ASENCIO MELLADO, *La STC 97/2019, de 16 de julio. Descanse en paz la prueba ilícita*, cit., *passim*, sobre la relectura de la STC núm. 114/1984, de 29 de noviembre [ECLI:ES:TC:1984:114]. Por tanto, si un tribunal compuesto por juristas de reconocido prestigio incurre en este vicio ¿acaso no lo va a sufrir la IA?

⁴⁹ Vid. lo dispuesto, entre otras, en la STC núm. 13/1987, de 5 de febrero [ECLI:ES:TC:1987:13]; o la STC núm. 184/1998, de 28 septiembre [ECLI:ES:TC:1998:184], que otorgó el amparo en un caso en el que la motivación rezaba: «se pone de manifiesto que los hechos declarados probados encuentran su base en la prueba documental y en las admisiones efectuadas por las partes»; y la STC núm. 302/2006, de 23 de octubre [ECLI:ES:TC:2006:302].

⁵⁰ Este riesgo de petrificación judicial ha sido identificado por diversos autores, vid. NIEVA FENOLL, *Inteligencia artificial*, cit., pp. 144 y ss.; S. AMRANI MEKKI, *Garantías frente a eficiencia: ¿es lo racional siempre razonable?*, en AA.VV., *Justicia: ¿garantías versus eficiencia?*, cit., pp. 58-59; M.E. BUAT-MENARD, *La intelligence artificielle au service du pouvoir judiciaire* [en línea], p. 9, [consultado por última vez el 27 de enero de 2021], disponible en: <<https://rm.coe.int/the-so-called-predictive-justice-applications-in-civil-and-criminal-ma/16808e4d85>>, quien refiere que la jurisprudencia podría quedar prisionera del pasado.

del Derecho, sobre todo en sede de recurso de casación, debería excluirse. De este modo, podemos preguntarnos cómo reaccionará el sistema experto cuando tenga que enfrentarse a intrincados concursos de normas (art. 8 CP)⁵¹, delimitar el ámbito de aplicación de ciertos preceptos⁵² o simplemente señalar cómo interactúan nuevas realidades con la legislación existente⁵³.

Para finalizar con la exposición de este apartado, debemos referirnos al riesgo que entrañaría el uso de sistemas expertos para dictar ‘sentencia anticipada’ en orden a presionar al acusado para poner fin al proceso mediante conformidad. Aun cuando pueda ser políticamente incorrecto abordar este asunto, las presiones para que los acusados no se defiendan en juicio y acepten la calificación y pena propuesta por la acusación existen en los tribunales. Cualquiera que haya estado una mañana en un juzgado en el que celebren juicios rápidos, incluso otros procedimientos más enjundiosos, sabrá de lo que hablamos⁵⁴. Un uso indebido de los programas de IA para valorar la prueba o adelantar propuestas de resolución podría convertirse en un elemento más para inducir estas conformidades forzadas, de tal suerte que incluso el inocente acabe accediendo a ellas⁵⁵.

⁵¹ Vid., a modo de ejemplo, el recurso de casación resuelto recientemente por la STS, Sala 2ª, núm. 849/2022, de 27 de octubre (ECLI:ES:TS:2022:3875), en el que el TS entró a conocer del concurso normativo entre el delito de extorsión en grado de tentativa (art. 243 CP) y las amenazas condicionales lucrativas (art. 169.1º CP).

⁵² Cabe pensar aquí, sin ir más lejos, en los problemas que podría ofrecer la IA si tuviera que enfrentarse a la necesidad de concretar la expresión «ejecutar el hecho mediante disfraz», sobre la que el legislador constituye una de las agravantes del art. 22.2º CP. Si se quiere ahondar más en los inconvenientes que venimos analizando, es posible llevar el ejemplo todavía más lejos. Imagínese que el sistema tuviera que decidir sobre la aplicación de este precepto en un escenario pandémico en el que la mascarilla se impone obligatoriamente. Casi con toda seguridad el resultado sería que cualquier ilícito penal cometido por sujeto que observara las normas sanitarias debería sancionarse de manera agravada. Este es el caso extremo, pero cabe representarse otros derivados de las asociaciones del sistema, v. gr., en el caso de que se hubieran dictado diversas sentencias aplicando la agravante de disfraz por uso de pelucas y bigotes falsos, la IA podría llegar a concluir que debe castigarse de manera agravada a cualquier acusado que luzca mostacho por gusto o bisoñé por estar acomplejado con su calvicie, y no por querer dificultar su identificación.

⁵³ Podemos preguntarnos en esta ocasión ¿cómo hubiera resuelto la IA el supuesto que se planteó en la STS, Sala 2ª, núm. 326/2019, de 20 de junio (ECLI:ES:TS:2019:2109), en la que hubo que decidir, prácticamente sin base normativa, sobre la naturaleza del *bitcoin* a la hora de aplicar el art. 110 CP?

⁵⁴ Vid. las notas sobre las ‘conformidades encubiertas’ o *contra legem* que proporciona M. AGUILERA MORALES, *La deriva del ‘principio’ del consenso*, en AA.VV., *Derecho procesal. Retos y transformaciones*, cit., pp. 193-210.

⁵⁵ Vid. J.A. LASCURAÍN SÁNCHEZ; F. GASCÓN INCHAUSTI, *¿Por qué se conforman los inocentes?*, en «InDret» [en línea], núm. 3, 2018 [consultado por última vez el 15 de noviembre de 2022], disponible en: <<https://indret.com/por-que-se-conforman-los-inocentes>>.

El riesgo se acentúa todavía más si se contempla la posibilidad de que el algoritmo, amén de valorar el sustrato fáctico y jurídico de la causa, pudiera trabajar con otros parámetros –por ejemplo, el coste para las arcas públicas de tramitar el litigio o el tiempo medio que consumirá el tribunal en la sustanciación de aquel– que puedan llevar al juez a forzar la conformidad tomando en consideración elementos totalmente ajenos al objeto del proceso.

3. La cuestión de la transparencia del algoritmo empleado en el enjuiciamiento penal

Al principio del capítulo afirmábamos que, en este momento, cualquier trabajo sobre la IA tiene algo de tópico. Y si existe un tópico insoslayable en esta materia es, precisamente, el de la transparencia de los algoritmos que conforman el corazón de los sistemas expertos. La posibilidad de que estos cobijen sesgos compromete las características definitorias de la función jurisdiccional, que son las que, a la postre, generan o debieran generar la confianza en el sistema judicial; singularmente, la independencia, la imparcialidad y el sometimiento a la ley⁵⁶. En idéntico sentido, la aparición de la mácula referida puede poner en jaque derechos fundamentales procesales, aunque justo es decir que tampoco existe consenso sobre cuáles podrían resultar afectados⁵⁷. Quizás haya que estar al carácter público o no

⁵⁶ En nuestra opinión, la característica que sufriría una mayor afección sería la imparcialidad, por cuanto el programa informático y su respuesta falseada eliminarían la neutralidad del tribunal respecto del objeto introduciendo una prevención negativa –el sesgo– hacia el mismo. Empero, el tribunal seguiría conservando su libertad enjuiciadora respecto de otros poderes y tribunales, por lo que no nos parece que se elimine la independencia judicial en su sentido más puro.

⁵⁷ Puede acudirse a los tratamientos que realizan NIEVA FENOLL, *Inteligencia artificial*, cit., pp. 127 y ss.; F. MARTÍN DÍZ, *Inteligencia artificial y proceso: garantías frente a eficiencia en el entorno de los derechos procesales fundamentales*, en AA.VV., *Justicia: ¿garantías versus eficiencia?*, F. JIMÉNEZ CONDE; R. BELLIDO PENADÉS (dirs.); LLOPIS NADAL, DE LUIS GARCÍA (coords.), cit., pp. 815-827; PÉREZ ESTRADA, *El uso de algoritmos en el proceso penal y el derecho a un proceso con todas las garantías*, en AA.VV., *Claves de la justicia penal: feminización, inteligencia artificial, supranacionalidad y seguridad*, S. BARONA VILAR, (ed.), Tirant Lo Blanch, Valencia, 2019, pp. 235-253; A. ZAVRSNIK, *Criminal justice, artificial intelligence systems and human rights*, en «ERA Forum», núm. 20, 2020, pp. 567-583 [consultado por última vez el 15 de noviembre de 2022], disponible en: <<https://link.springer.com/article/10.1007/s12027-020-00602-0>>; G. SCHUMANN BARRAGÁN, *La inteligencia artificial aplicada al proceso penal desde la perspectiva de la UE*, en AA.VV.,

del algoritmo y a las circunstancias de cada caso concreto para dilucidar cuáles son los derechos afectados.

En nuestra opinión, en el caso de sistemas que funcionen de manera transparente pero contengan sesgos, cabe colegir que, al no procurarse un juez imparcial, quedaría comprometido el derecho a un proceso con todas las garantías tipificado en el art. 24.2 CE, en el art. 6 del Convenio Europeo de Derechos Humanos y en el art. 47 de la Carta de Derechos Fundamentales de la Unión Europea –en adelante, CDFUE–.

En cambio, si, además de sesgado, se trata de un algoritmo secreto que no permite conocer los datos que está usando el programa, el origen lícito o ilícito de estos y cuál es el patrón de funcionamiento, la imposibilidad de contradecirlo podría dañar también el derecho de defensa (art. 24.2 CE y 48.2 CDFUE) en relación con el derecho a la tutela judicial efectiva sin indefensión (art. 24.1 CE y 47 CDFUE)⁵⁸.

Así las cosas, la transparencia de las posibles aplicaciones de IA cobra una especial relevancia en aras a la salvaguarda de las garantías básicas del proceso penal. Nos parece que la exigencia de transparencia debe concretarse en un doble plano. En primer lugar, en la necesidad de que las partes puedan conocer con antelación que el juez o tribunal sentenciador actuará asistido por una IA; de otro modo, acusación y defensa no dispondrían de oportunidad real de practicar prueba sobre el funcionamiento de aquella. La cuestión más controvertida estaría en dilucidar cuál es el momento en que ha de ponerse de relieve esta información. Y nos parece que, si el sentido de practicar prueba sobre el funcionamiento del algoritmo y los datos que lo nutren es acreditar posibles fallas que impongan la exclusión del asistente virtual, la información debería comunicarse a las partes antes de la celebración del juicio oral, con el fin de que, en su caso, propongan la correspondiente prueba pericial en sus escritos de calificación. De ahí que el auto de apertura de juicio oral (arts. 633 y 783 LECrim) pudiera ser un buen momento para anunciar que el órgano enjuiciador se apoyará en el sistema experto. Esclarecido esto, todavía faltaría fijar cuál será el momento idóneo para practicar prueba sobre el funcionamiento del sistema. Y es que,

Investigación y proceso penal en el siglo XXI: nuevas tecnologías y protección de datos, S. PEREIRA PUIGVERT; F. ORDOÑEZ PONZ (dirs.), Thomson-Aranzadi, Cizur Menor (Navarra), 2021, pp. 517-539 [consultado por última vez el 15 de noviembre de 2022], disponible en: <<https://acortar.link/BmToHP>>.

⁵⁸ Ciertamente, en la mayor parte de casos, conocer los aspectos reseñados es lo que permitirá determinar si existe o no sesgo. No obstante, podrían darse hipótesis extremas en la que la existencia de este sesgo se evidencie a simple vista a la luz de los resultados que arroja la máquina sin necesidad de ulteriores comprobaciones.

si se considera, como hicimos *supra*, que el uso de un algoritmo sesgado puede llegar a comprometer la imparcialidad del juzgador, cabe preguntarse si podría emplearse el incidente de recusación, o bien aquella prueba ha de practicarse directamente en el juicio oral. Al respecto pensamos que, tanto el trámite de los arts. 52 y ss. LECrim, como las causas de recusación del art. 219 LOPJ, se acomodan mal a esta finalidad. Por ello, lo idóneo sería practicar la prueba sobre el funcionamiento del sistema en el acto de juicio oral.

En lo que atañe al segundo factor en el que se materializa la exigencia de transparencia, puede apuntarse que, conocido por el acusado y por las acusaciones que el juez estará asistido por la IA, es preciso saber cuál es el funcionamiento del mismo y cuáles son los datos de los que se nutre.

El problema, como sucedió en el Caso *Loomis* con la aplicación COMPASS, surge aquí por la posibilidad de que el algoritmo esté salvaguardado por los derechos de propiedad intelectual y los paquetes de datos por el secreto empresarial, aunque todavía no existe consenso pleno en la doctrina mercantilista acerca de cuál es la figura de protección adecuada para cada extremo⁵⁹. Pues bien, ¿cómo puede superarse este obstáculo?

En el marco legislativo actual y fuera del orden penal, el caso Bosco evidencia que los intentos de obtener el código fuente de los programas utilizados por el Estado para determinar la vulnerabilidad de los consumidores a efectos de otorgarles un bono social para el pago de la electricidad han fallado, pues la Administración ha opuesto los arts. 14 y 18 de la Ley 19/2013, de 9 de diciembre, de transparencia⁶⁰. El primero relativo a la existencia de derechos de propiedad intelectual; el segundo a la necesidad de rechazar solicitudes abusivas. Si bien, la sentencia núm. 143/2021, de 30 de diciembre, del Juzgado Central de lo Contencioso-Administrativo núm. 8, da especial importancia también a los problemas de vulnerabilidad del sistema que traería aparejada la revelación del código⁶¹.

⁵⁹ Llama la atención sobre este punto SCHUMANN BARRAGÁN, *La inteligencia artificial aplicada al proceso penal desde la perspectiva de la UE*, cit., p. 19 del documento en formato electrónico.

⁶⁰ La síntesis del caso puede encontrarse en J. DE LA CUEVA, *El derecho a no ser gobernados por algoritmos secretos*, en «El notario del Siglo XXI» [en línea], núm. 87, 2019 [consultado por última vez el 15 de noviembre de 2022], disponible en: <<https://www.elnotario.es/hemeroteca/revista-87/9637-quien-vigila-al-algoritmo>>. Da cuenta del caso, también con cita del anterior trabajo, ARMENTA DEU, *Derivas*, cit., pp. 311-312.

⁶¹ En el momento en que se redacta este trabajo la sentencia está recurrida en apelación. Con posterioridad a la interposición del recurso se ha promulgado la Ley 15/2022, de 12 de julio, integral para la igualdad de trato y no discriminación; norma de vasto ámbito que establece en su art. 23, rubricado «Inteligencia artificial y mecanismos de toma de decisión automatizados», que las administraciones públicas cuidarán que los algoritmos tengan en cuenta criterios de minimización de sesgos, transparencia y

Por suerte, el proyecto de Reglamento sobre IA de la Unión Europea que se dio a conocer en abril de 2021 nos proporciona ya pautas de interés sobre esta materia. Vaya por delante que esta futura norma y el hecho de que adopte la forma de Reglamento deben ser bienvenidos. Como hemos manifestado en otro lugar, en nuestra opinión, el principio de legalidad procesal obliga al legislador a establecer unos principios mínimos a los que deban adaptarse los programadores informáticos que construyan, entrenen e instalen los sistemas de IA, so pena, en otro caso, de que lo que acabe adaptándose a la informática sea la norma y no al revés⁶². Si estos principios vienen dados al legislador nacional desde Europa mejor que mejor.

En este sentido, la propuesta Reglamento establece un modelo de control previo, es decir, impone la auditoría de los algoritmos con anterioridad a su uso efectivo. Del mismo modo, contempla la puesta a disposición de manuales que permitan conocer cómo funciona el sistema y la obligación de revisarlo periódicamente. Cuestión distinta será a quién se atribuya este control. Sería deseable que no fuera al Comité Técnico Estatal de la Administración de Justicia Electrónica (art. 44 LTICAJ), que es a quien el art. 230.6 LOPJ encomienda la validación funcional de los programas que se utilicen en la Administración de Justicia, parece que previo informe del CGPJ⁶³. Baste por ahora decir que ni el Comité Técnico Estatal de la Administración de Justicia ni el CGPJ cuentan con personal técnico especializado que esté en condiciones de auditar debidamente estos programas. De ahí que lo razonable sea la creación de un Instituto de Informática Forense que, de modo similar a lo que sucede con los Institutos de Medicina legal, se encargue de aplicar y testar estos programas⁶⁴. Por supuesto, ello sin perjuicio de las labores de auditoría interna de sus creadores⁶⁵.

El hecho de que el sistema deba funcionar de modo transparente permite

rendición de cuentas. El párrafo segundo de la norma se encarga de dejar sentado que las administraciones públicas priorizarán la transparencia y la capacidad de interpretación de las decisiones adoptadas por los algoritmos.

⁶² Vid. CASTILLO FELIPE, *Proceso civil e inteligencia artificial*, cit., pp. 284-285.

⁶³ En CASTILLO FELIPE, *Proceso civil e inteligencia artificial*, cit., pp. 278-283, se estudia la madeja normativa que ofrecen estos preceptos y el Acuerdo de 15 septiembre de 2005, del Pleno del Consejo General del Poder Judicial, por el que se aprueba el Reglamento 1/2005, de los aspectos accesorios de las actuaciones judiciales –BOE [en línea], núm. 231, de 27 de septiembre de 2005, pp. 31859-31886 [consultado por última vez el 15 de noviembre de 2022], disponible en: <<https://www.boe.es/buscar/doc.php?id=BOE-A-2005-15939>>.

⁶⁴ La idea se ha desarrollado *in extenso* en el trabajo citado en la nota anterior.

⁶⁵ Que exista un modelo mixto o público-privado de auditoría nos parece razonable, sobre todo teniendo en cuenta la posibilidad de que alguno de los actores –no necesariamente privado– persiga falsear los resultados del control para conseguir finalidades espurias.

colegir que, a la vista de la información suministrada por el proveedor, será posible practicar una prueba pericial de parte tendente a cuestionar el funcionamiento del algoritmo. Y es que, cuando hablamos de sistemas inteligentes, capaces de mejorar por sí mismos, el control previo no dota a estos de una suerte de infalibilidad. Por el contrario, únicamente implica que, en el momento de ser examinado, el programa trabaja correctamente, lo que cual no excluye ulteriores fallas.

Desde el prisma de la futura regulación, el inconveniente podría darse si la información suministrada en los manuales y documentación del proveedor es suficiente para que el juez entienda el programa a nivel de usuario pero exigua para que el perito o el organismo al que se encargue la pericia destripe el algoritmo en el juicio oral. En estos casos, debemos preguntarnos si el fabricante podría negarse a proporcionar información adicional al perito aduciendo que el algoritmo está protegido por la propiedad intelectual o por el secreto empresarial.

Teniendo en cuenta que el derecho a un proceso con todas las garantías y el derecho de defensa de las partes podría estar en juego, entendemos que el suministro de información adicional no podría, en principio, denegarse. Ahora bien, tampoco podría procederse como si no existieran los derechos de propiedad intelectual o secretos industriales. Consideramos que la solución pasaría por aplicar analógicamente —o incluso hacer extensiva mediante disposición legal expresa— las previsiones que consagra el art. 15 de la Ley 1/2019, de 20 de febrero, de Secretos Empresariales⁶⁶. Este precepto permite al juez adoptar las medidas necesarias para garantizar la confidencialidad del secreto empresarial en el seno del procedimiento en el que se discute sobre él o cuestiones relativas a las licencias de su explotación. Entre ellas, restringir el número de personas que tienen acceso al secreto empresarial, limitar la publicidad de las vistas en las que vaya a discutirse sobre esta cuestión y eliminar de la resolución judicial que se hará pública las menciones relativas a aquel⁶⁷.

⁶⁶ El artículo citado indica en su apartado 2 que los jueces o tribunales, a instancia de parte o de oficio, podrán adoptar las medidas necesarias para proteger la confidencialidad de la información considerada secreto empresarial en procedimientos que no versen sobre esta materia, pero en los que sea necesario conocer el secreto empresarial para resolver el fondo. Habida cuenta de la amplitud de la norma entendemos que la disposición sería aplicable directamente. Ahora bien, como quiera que no todas las partes del sistema de IA se protegen mediante la técnica del secreto empresarial, la aplicación analógica de la disposición podría resultar pertinente respecto a otras modalidades de protección.

⁶⁷ Otros autores, como SCHUMANN BARRAGÁN, *La inteligencia artificial aplicada al proceso penal desde la perspectiva de la UE*, cit., p. 20, proponen la creación de mecanismos específicos similares a los de los arts. 283 bis a) a k) LEC.

Esta solución valdría también para el caso de que el uso del algoritmo lo proponga una de las partes, por ejemplo, una de las acusaciones.

Para concluir, cabe señalar que el desembarco de la IA en el proceso sólo debería producirse con un marco legislativo que proporcione seguridad jurídica, con unos órganos de control integrados por personal especialista –y no político, que estén bien formados, y con unas garantías de transparencia irrenunciables para quienes pretendan diseñar sistemas que auxilien al tribunal en la difícil labor de juzgar. De otro modo, estarán en juego derechos y garantías que, aunque aplicados en un sistema imperfecto –toda creación humana lo es–, hasta ahora estaban suficientemente protegidos.

4. Bibliografía

- AGUILERA MORALES, M., *La deriva del ‘principio’ del consenso*, en AA.VV., *Derecho Procesal. Retos y transformaciones*, BUJOSA VADELL, L. (dir.), Atelier, Barcelona, 2021, pp. 193-210.
- AMRANI MEKKI, S., *Garantías frente a eficiencia: ¿es lo racional siempre razonable?*, en AA.VV., *Justicia: ¿garantías versus eficiencia?*, JIMÉNEZ CONDE, F., BELLIDO PENADÉS, R. (dirs.); LLOPIS NADAL, P., DE LUIS GARCÍA, E. (coords.), Tirant Lo Blanch, Valencia, 2019, pp. 31-60.
- ARMENTA DEU, T., *Lecciones de Derecho Procesal penal*, 9ª ed., Marcial Pons, Madrid, 2015.
- ARMENTA DEU, T., *Derivas de la justicia. Tutela de los derechos y solución de controversias en tiempos de cambio*, Marcial Pons, Madrid, 2021.
- ASENCIO MELLADO, J. M., *La STC 97/2019, de 16 de julio. Descanse en paz la prueba ilícita*, «Diario la Ley» [en línea], núm. 9499, 2019 [consultado por última vez el 10 de noviembre de 2022], disponible en: <<https://laleydigital.laleynext.es/Content/Inicio.aspx>>.
- BARONA VÍLAR, S., *Algoritmización del Derecho y de la Justicia. De la inteligencia artificial a la Smart Justice*, Tirant Lo Blanch, Valencia, 2021.
- BONET NAVARRO, J., *Valoración de la prueba y resolución mediante inteligencia artificial*, en AA.VV., *Derecho Procesal. Retos y transformaciones*, BUJOSA VADELL, L. (dir.), Atelier, Barcelona, 2021, pp. 315-337.
- BUAT-MENARD, M. E., *La intelligence artificielle au service du pouvoir judiciaire* [en línea], p. 9, [consultado por última vez el 27 de enero de 2021], disponible en: <<https://rm.coe.int/the-so-called-predictive-justice-applications-in-civil-and-criminal-ma/16808e4d85>>.

- BUENO DE MATA, F., *Macrodatos, inteligencia artificial y proceso: luces y sombras*, en «Revista General de Derecho Procesal» [en línea], núm. 51, 2020 [consultado por última vez el 8 de noviembre de 2022], disponible en: <<https://acortar.link/gnhfjV>>.
- BUENO DE MATA, F., *Hacia un proceso civil eficiente: transformaciones judiciales en un contexto pandémico*, Tirant Lo Blanch, Valencia, 2021.
- CASTILLO FELIPE, R., *Proceso civil e inteligencia artificial*, en AA.VV., *Proceso civil y nuevas tecnologías*, SIGÜENZA LÓPEZ, J. (dir.), Thomson-Aranzadi, Cizur Menor (Navarra), 2021, pp. 259-296.
- CASTILLO FELIPE, R., *Medidas para una justicia más ¿eficiente?: limitación de escritos y sentencias orales*, en AA.VV., *Derecho y pandemia desde una perspectiva global*, M. ESPEJO LERDO DE TEJADA; J. GARCÍA OLIVA; J. MARTÍNEZ CRUZ; J.P. MURGA FERNÁNDEZ (dirs.), A. ANDRADES NAVARRO (coord.), Thomson-Aranzadi, Cizur Menor (Navarra), 2021, pp. 135-150.
- DE LA CUEVA, J., *El derecho a no ser gobernados por algoritmos secretos*, en «El notario del Siglo XXI» [en línea], núm. 87, 2019 [consultado por última vez el 15 de noviembre de 2022], disponible en: <<https://www.elnotario.es/hemeroteca/revista-87/9637-quien-vigila-al-algoritmo>>.
- DE LA OLIVA SANTOS, A., *El papel del juez en el proceso civil*, Thomson-Civitas, Cizur Menor (Navarra), 2012.
- DOMÉNECH PASCUAL, G., *Juzgar a destajo*, Thomson-Civitas, Cizur Menor (Navarra), 2009.
- ESTÉVEZ MENDOZA, L., *Algoritmos policiales basados en IA y derechos fundamentales a la luz de HART y VALCRI*, en AA.VV., *Justicia: ¿garantías versus eficiencia?*, JIMÉNEZ CONDE, F., BELLIDO PENADÉS, R. (dirs.); LLOPIS NADAL, P., DE LUIS GARCÍA, E. (coords.), Tirant Lo Blanch, Valencia, 2019, pp. 663-675.
- FERNÁNDEZ SALMERÓN, M., *De la reutilización de sentencias al 'Big data' judicial. Aproximación a la metamorfosis experimentada por los modelos de uso de la información en el marco de la actividad jurisdiccional*, en AA.VV., *Modernización digital e innovación en la Administración de Justicia*, GÓMEZ MANRESA, F.; FERNÁNDEZ SALMERÓN, M. (coords.), Thomson-Aranzadi, Cizur Menor (Navarra), 2019, pp. 63-102.
- GASCÓN INCHAUSTI, F., *El control de la fiabilidad probatoria: prueba sobre la prueba en el proceso penal*, Ediciones de la Revista General de Derecho, Valencia, 1999.
- GASCÓN INCHAUSTI, F., *Desafíos para el proceso penal en la era digital: externalización, sumisión pericial e inteligencia artificial*, en AA.VV., *La justicia digital en España y la Unión Europea*, Atelier, Barcelona, 2019, pp. 191-206.

- GUZMÁN FLUJA, V., *Proceso penal y justicia automatizada*, en «Revista General de Derecho Procesal» [en línea], núm. 53, 2021 [consultado por última vez el 10 de noviembre de 2022], disponible en: <<https://acortar.link/wkWVzh>>.
- HAACK, S., *Proof, Probability, and Statistics: The Dangers of 'delusive exactness'*, en AA.VV., *La prueba en el proceso*, Atelier, Barcelona, 2018, pp. 497-519.
- JIMÉNEZ CONDE, F., *La apreciación de la prueba legal y su impugnación*, Publicaciones del Departamento de Derecho Procesal de la Universidad de Salamanca, Salamanca, 1977.
- LASCURAÍN SÁNCHEZ, J. A.; GASCÓN INCHAUSTI, F., ¿Por qué se conforman los inocentes?, en «InDret» [en línea], núm. 3, 2018 [consultado por última vez el 15 de noviembre de 2022], disponible en: <<https://indret.com/por-que-se-conforman-los-inocentes>>.
- LLORENTE-SÁNCHEZ ARJONA, M., *La inteligencia artificial como nueva estrategia de prevención en los delitos de violencia sexual*, en AA.VV., *Uso de la información y de los datos personales en los procesos*, COLOMER HERNÁNDEZ, I. (dir.), CATALINA BENAVENTE, M. A.; OUBIÑA BARBOLLA, S. (coords.), Thomson-Aranzadi, Cizur Menor (Navarra), 2022, pp. 259-281.
- LÓPEZ YAGÜES, V., *La prueba prohibida*, en AA.VV., *Derecho Procesal penal*, ASENCIO MELLADO, J. M. (dir.), FUENTES SORIANO, O. (coord.), Tirant Lo Blanch, Valencia, 2021, pp. 161-166.
- MARTÍN DIZ, F., *Inteligencia artificial y proceso: garantías frente a eficiencia en el entorno de los derechos procesales fundamentales*, en AA.VV., *Justicia: ¿garantías versus eficiencia?*, JIMÉNEZ CONDE, F., BELLIDO PENADÉS, R. (dirs.); LLOPIS NADAL, P.; DE LUIS GARCÍA, E. (coords.), Tirant Lo Blanch, Valencia, 2019, pp. 815-827.
- MEDRANO SÁNCHEZ, J. L., *Luces y sombras del expediente judicial digital*, en AA.VV., *Aciertos, excesos y carencias en la tramitación del proceso*, HERRERO PEREZAGUA, J. F.; LÓPEZ SÁNCHEZ, J. (dirs.), Atelier, Barcelona, 2020, pp. 49-85.
- MORENO CATENA, V., *Los datos en el sistema de justicia y la propuesta de Reglamento UE sobre inteligencia artificial*, en AA.VV., *Uso de la información y de los datos personales en los procesos: los cambios en la era digital*, COLOMER HERNÁNDEZ, I. (dir.), CATALINA BENAVENTE, M. A.; OUBIÑA BARBOLLA, S. (coords.), Thomson-Aranzadi, Cizur Menor (Navarra), 2022, pp. 47-73.
- MUÑOZ SABATÉ, L., *La prueba de indicios en el proceso judicial*, Wolters Kluwer, Madrid, 2016.
- MURILLO FUENTES, J. J., ¿Qué es lo que no funciona en los algoritmos de inteligencia artificial?, en AA.VV., *Uso de la información y de los datos*

- personales en los procesos: los cambios en la era digital*, COLOMER HERNÁNDEZ, I. (dir.), CATALINA BENAVENTE, M. A.; OUBIÑA BARBOLLA, S. (coords.), Thomson-Aranzadi, Cizur Menor (Navarra), 2022, pp. 151-171.
- NIEVA FENOLL, J., *Prueba científica. Cuestiones de futuro: neurociencia e inteligencia artificial*, en AA.VV., *La prueba en el proceso*, Atelier, Barcelona, 2018, pp. 473-495.
- NIEVA FENOLL, J., *Inteligencia artificial y proceso judicial*, Marcial Pons, Madrid, 2018.
- NIEVA FENOLL, J., *Inteligencia artificial y proceso judicial: perspectivas tras un alto tecnológico en el camino*, en «Revista General de Derecho Procesal» [en línea], núm. 57, 2022 [consultada por última vez el 9 de noviembre de 2022], disponible en: <<https://acortar.link/DTrjhX>>.
- ORTEGO PÉREZ, F., *Consideraciones críticas en torno a la denominada prueba pericial de «inteligencia policial»*, en AA.VV. *Peritaje y prueba pericial*, PICÓ I JUNOY, J. (dir.), DE MIRANDA VÁZQUEZ, C. (coord.), Bosch, Barcelona, 2017, pp. 583-589.
- PEARL, J.; MCKENZIE, D., *The Book of Why. The New Science of Cause and effect*, Penguin Random House UK, Londres, 2019.
- PEIRANO, M., *El enemigo conoce el sistema*, Penguin Random House, Barcelona, 2019.
- PÉREZ ESTRADA, M. J., *La inteligencia artificial como prueba científica en el proceso penal español*, en «Revista Brasileira de Direito Processual Penal» [en línea], Vol. 7, núm. 2, pp. 1385-1410 [consultado por última vez el 10 de noviembre de 2022], disponible en: <<https://acortar.link/fehWmK>>.
- PÉREZ ESTRADA, M. J., *El uso de algoritmos en el proceso penal y el derecho a un proceso con todas las garantías*, en AA.VV., *Claves de la justicia penal: feminización, inteligencia artificial, supranacionalidad y seguridad*, BARONA VILAR, S. (ed.), Tirant Lo Blanch, Valencia, 2019, pp. 235-253.
- SCHUMANN BARRAGÁN, G., *La inteligencia artificial aplicada al proceso penal desde la perspectiva de la UE*, en AA.VV., *Investigación y proceso penal en el siglo XXI: nuevas tecnologías y protección de datos*, PEREIRA PUIGVERT, S.; ORDOÑEZ PONZ, F. (dirs.), Thomson-Aranzadi, Cizur Menor (Navarra), 2021, pp. 517-539 [consultado por última vez el 15 de noviembre de 2022], disponible en: <<https://acortar.link/BmToHP>>.
- TARUFFO, M., *Evidence and Truth in Procedural Law*, en AA.VV., *La prueba en el proceso*, Atelier, Barcelona, 2018, pp. 31-47.
- TUZET, G., *Filosofia della prova giuridica*, 2ª ed., Giappichelli Editore, Turín, 2016.

- TOMÉ GARCÍA, J., *Fase decisoria (II). La prueba*, en DE LA OLIVA SANTOS, A. *et. al.*, *Derecho Procesal Penal*, 8ª ed., Centro de Estudios Ramón Areces, Madrid, 2008, pp. 475-513.
- VILLAMARÍN LÓPEZ, M. L., *Neurociencia y detección del engaño en el proceso penal. El uso del escáner cerebral (fMRI) y del brainfingerprinting (P300)*, Marcial Pons, Madrid, 2014.
- ZAVRSNIK, A., *Criminal justice, artificial intelligence systems and human rights*, en «ERA Forum», núm. 20, 2020, pp. 567-583 [consultado por última vez el 15 de noviembre de 2022], disponible en: <<https://link.springer.com/article/10.1007/s12027-020-00602-0>>.

SECCIÓN III
NEUROCIENCIAS Y RESPONSABILIDAD PENAL
NEUROSCIENZE E RESPONSABILITÀ PENALE

CAPÍTULO IX

Inteligencia artificial y neurociencia. Algunas reflexiones sobre las aportaciones que pueden hacer al Derecho Penal

José Manuel Palma Herrera

SUMARIO: 1. Relación entre Derecho penal, inteligencia artificial y neurociencia. Planteamiento de la cuestión – 2. Escenario 1: incidencia de los estudios neurocientíficos en la libertad como fundamento de la culpabilidad – 3. Escenario 2: creación de herramientas de inteligencia artificial capaces de interactuar con el propio individuo condicionando su comportamiento – 4. Escenario 3: prevención, enjuiciamiento y tratamiento del delito – 5. Bibliografía.

1. Relación entre Derecho penal, inteligencia artificial y neurociencia. Planteamiento de la cuestión

Hasta no hace mucho, saber qué pasaba o cómo funcionaba la mente de determinados delincuentes capaces de las mayores atrocidades era poco menos que impensable. Comportamientos cuya etiología resultaba desconocida acababan sido atribuidos a la maldad intrínseca de algunos seres humanos que, teniendo intacta su capacidad intelectual, se inclinaban por el comportamiento delictivo en el ejercicio de su libertad. Frente a ello, el Derecho Penal, construido sobre presupuestos normativos y a partir del axioma de que el hombre es un ser libre, reaccionaba (de hecho, sigue haciéndolo) con la imposición de una pena fundamentada en un principio de culpabilidad en virtud del cual se reprocha al sujeto haber hecho lo que hizo, cuando pudo haber actuado de otro modo.

Hoy, gracias a los estudios neurológicos y psicológicos llevados a cabo en los últimos 30 años, podríamos encontrarnos en condiciones de conocer cómo funciona el cerebro del ser humano; de conocer cuáles son los procesos biológicos que dirigen su comportamiento.

Podríamos encontrarnos, así, en condiciones de contar con criterios de base científica sobre los que llevar a cabo esa difícil delimitación que en ocasiones tenemos que hacer entre los supuestos de ausencia de acción y

de falta de culpabilidad. Podríamos incluso estar en condiciones de poder demostrar, con esa base científica, si existe o no la libertad del hombre para autodeterminarse, con las consecuencias que su negación acarrearía para el Derecho penal, toda vez que esa libertad ha venido marcando la respuesta penal tanto frente al delincuente imputable como frente al inimputable. Más aún, podríamos estar en condiciones de ser capaces de predecir el comportamiento humano e incluso de determinarlo en la dirección que quisiésemos, con lo que ello podría tener de relevancia tanto en lo que hace a la previsión de medidas de seguridad, como en materia de rehabilitación del penado.

Pero, antes de entrar en lo que estos avances neurocientíficos pueden suponer en el terreno de la culpabilidad, procede que lo hagamos en lo que está suponiendo el desarrollo de la inteligencia artificial en términos generales tanto para el Derecho penal como para la neurociencia.

Desde las últimas décadas del siglo XX vivimos un momento de desarrollo, progreso y bienestar como nunca antes se había conocido en Occidente: la revolución tecnológica a la que asistimos y, sobre todo, que se avecina con la incorporación a la vida cotidiana de sistemas de inteligencia artificial capaces de tomar decisiones autónomas basadas en el aprendizaje y que nos facilitan nuestro día a día; la biotecnología que promete aumentar nuestra esperanza y calidad de vida y mejorar la salud del planeta sustituyendo progresivamente los alimentos de origen animal; la industria del ocio que ofrece nuevas experiencias personales impensables hasta hace unos años, etc. Llevamos años, además, insertos en un proceso de globalización que ha conducido a la supresión de fronteras, la libre circulación de personas y mercancías, la internacionalización de las empresas, o las concentraciones de capital, ampliando considerablemente las posibilidades económicas y de intercambio cultural.

Sin duda, la sociedad del siglo XXI es la sociedad del ocio y del consumo, pero es también cada vez más, una sociedad que se siente amenazada por la aparición de nuevas fuentes de peligro asociadas, paradójicamente, a ese progreso y bienestar.

Se trata de riesgos que la sociedad percibe como estructurales por su capacidad de afectar al conjunto de la sociedad y que provienen de decisiones humanas, múltiples y descentralizadas, que otros toman en el manejo de los avances técnicos y tecnológicos y en el desarrollo de las nuevas formas de actividad productiva o de prestación de servicios en el seno de corporaciones empresariales de carácter internacional¹. Además, estos nuevos riesgos son

¹ J.M. SILVA SÁNCHEZ, *La expansión del Derecho penal. Aspectos de la política criminal en las sociedades postindustriales*, 2ª ed., Civitas, Madrid, 2001, pp. 26-28.

frecuentemente difíciles de evaluar al no conocerse su verdadera potencialidad lesiva ni la de las conductas que los ponen en marcha y por tratarse, por regla general, de riesgos que sólo se manifiestan a largo plazo, como sucede con los avances en inteligencia artificial.

Para mitigar tales peligros, la industria aplica, normalmente con éxito, sus avances científicos y tecnológicos, pero al hacerlo genera nuevas fuentes de peligro para la sociedad que en ocasiones se muestran tan lesivas o más que las que se pretendían evitar incrementando así la sensación de amenaza. Una sensación de amenaza a la que contribuyen las dificultades con las que tropieza el ciudadano para comprender el acelerado cambio tecnológico y para adaptarse a él, así como la desmesurada cobertura mediática que los medios de información llevan a cabo sobre tales cambios².

Aun cuando el recurso al Derecho penal debería ser la última de las soluciones, en ese estado de ‘institucionalización de la inseguridad’³, la sociedad reclama la seguridad que le brinda esta rama del ordenamiento jurídico, instancia en la que deposita toda su confianza, y los poderes públicos, sensibles a esta demanda, le dan a la sociedad más Derecho penal en un proceso expansivo de esta rama del ordenamiento jurídico que pone en tela de juicio algunos de los principios por los que se ha regido durante los dos últimos siglos.

Junto a este proceso expansivo, el Derecho penal vive otro proceso paralelo, que tiende a confundirse, pero que presenta caracteres definitorios específicos. Como el primero, se construye sobre la idea de inseguridad, pero a diferencia de aquel, no se trata ya de la inseguridad generada por las nuevas fuentes de riesgo o las nuevas formas de criminalidad. Ni siquiera de la inseguridad generada por la criminalidad de cuello blanco, sino de la inseguridad generada por la criminalidad clásica en un contexto caracterizado por dos notas definitorias⁴:

- 1) En primer lugar, la desaparición de la actitud de comprensión hacia la criminalidad tradicional, en especial hacia la pequeña delincuencia, una actitud muy difundida en los años 70 y 80, y que se fundaba en una comprensión del delincuente como un ser socialmente desfavorecido y marginado al que la sociedad estaba obligada a prestar ayuda.

² M.J. JIMÉNEZ DÍAZ, *Sociedad del riesgo e intervención penal*, en «Revista Electrónica de Ciencia Penal y Criminología», 16-08, 2014, p. 4.

³ SILVA SÁNCHEZ, *La expansión del Derecho penal. Aspectos de la política criminal en las sociedades postindustriales*, cit., pp. 28-31.

⁴ J.L. DÍEZ RIPOLLÉS, *El nuevo modelo penal de la seguridad ciudadana*, en «Revista Electrónica de Ciencia Penal y Criminología», 06-03, 2004, p. 8.

2) En segundo lugar, una extendida sensación en la sociedad de que las cosas van cada vez peor en temas de prevención de la delincuencia, una sensación avivada por los medios de comunicación, y que se proyecta en una escasa confianza en la capacidad de los poderes públicos para afrontar el problema y en una notable preocupación frente a fenómenos como las migraciones o el fanatismo terrorista. No en vano, los propios especialistas en la materia penal, como pueden ser académicos, abogados y jueces -quizás se salve el Ministerio Público-, son considerados un obstáculo al ideal de justicia que maneja la propia sociedad, pues impiden, con sus teorías o sus decisiones, la aplicación de la ley penal con el rigor que la misma merece. De este modo, el desprestigio de la Justicia (como organización) llega a niveles preocupantes. Todo lo contrario a lo que sucede con las Fuerzas y Cuerpos de Seguridad del Estado, con Guardia Civil y Policía Nacional entre las instituciones mejor valoradas al ser percibidas como garantes de esa ansiada seguridad.

En este estado de cosas, los principios garantistas constituyen también un obstáculo al que, paradójicamente, la sociedad está dispuesta a renunciar revisando los términos del contrato social en condiciones de permitir que el *ius puniendi* estatal cruce fronteras que en otros momentos se habrían considerado infranqueables en un Estado de Derecho. Esa falta de preocupación por las garantías responde a una sociedad dividida en buenos y malos ciudadanos, y en la que el buen ciudadano, cumplidor de las normas, no necesita de unas garantías que no hacen sino favorecer al ciudadano incumplidor, al enemigo de la sociedad y del Derecho.

En este contexto, irrumpe una inteligencia artificial que puede convertirse, en manos del ser humano, en un instrumento para la comisión de hechos delictivos, pero que puede convertirse también, y esto es ya más preocupante, en una fuente de riesgos en sí misma para el ser humano. Una fuente de riesgos que será mayor conforme el sistema inteligente adquiera la capacidad de reproducir la forma de funcionamiento del del cerebro humano y goce de mayor autonomía en la toma de decisiones.

Pero esa misma inteligencia artificial puede ser también, en el contexto descrito, un instrumento idóneo en la labor de lucha contra la criminalidad a través de la predicción del delito que ayude a incrementar la eficacia de las fuerzas policiales y, en consecuencia, a conseguir mayores niveles de esa seguridad que demanda la sociedad; un instrumento que ayude, igualmente al esclarecimiento de hechos delictivos ya cometidos; que ayude a jueces y tribunales en la toma de decisiones relativas a la adopción de medidas cautelares o en la valoración de la prueba; y que ayude, finalmente, en el tratamiento penitenciario con vistas a conseguir mayores tasas de éxito en la

reinserción social.

Si la relación entre la inteligencia artificial y el Derecho penal y Procesal penal está clara, también lo está la relación entre inteligencia artificial y neurociencia, aun cuando desde la comunidad científica se reclama una interconexión aún más estrecha⁵.

Podría decirse que inteligencia artificial y neurociencia se retroalimentan. Como señala algún autor, las máquinas más eficaces para procesar información de forma inteligente no son los procesadores que utilizan algoritmos de aprendizaje automático de última generación, sino los sistemas nerviosos de los seres vivos⁶. Y cuando hablamos de seres vivos, no tenemos que pensar en el hombre como eslabón último, por su complejidad, de la cadena evolutiva.

No debe extrañar, por ello, que la inteligencia artificial se fije en el sistema nervioso y cerebral a través de los conocimientos que ofrece la neurociencia, para avanzar en la creación de sistemas cada vez más complejos. No olvidemos que el objetivo último de la inteligencia artificial es llegar a replicar la forma de actuar del cerebro humano creando máquinas capaces de tomar decisiones de forma autónoma a partir de su interconexión sensorial y cognitiva con el entorno. Así, esa relación, que se remonta al año 1943, cuando WARREN McCULLOCH, un neurófiólogo de la Universidad de Illinois, y WALTER PITTS, un matemático autodidacta, propusieron lo que se considera el primer modelo de red neuronal artificial, se traduce, por ejemplo, en que la inteligencia artificial haya copiado el esquema de interacción sensorial con el mundo exterior propio del sistema nervioso; o en que la inteligencia artificial busque en las estructuras neuronales la necesaria robustez frente a los fallos y la capacidad de autorregulación y autorreparación del sistema nervioso⁷; o se traduce, asimismo, en la búsqueda de formas más eficientes de procesar la información a través de mecanismos de retroalimentación con causalidad circular o de lo que se denomina ‘conectividad en un mundo pequeño’ (muchas conexiones locales y relativamente pocas, pero cruciales, conexiones de largo alcance) como los que utiliza el sistema nervioso⁸.

Y en la otra dirección, la neurociencia aprovecha la capacidad de análisis

⁵ M.F. ZAMBRANO CASTRO, R.S. VARGAS LÓPEZ, E.H. ZAMBRANO FRANCO, K.A. ZAMBRANO FRANCO, *La neurociencia y su relación con la inteligencia artificial*, en «Revista Científica de Investigación actualización del mundo de las Ciencias», vol. 3, n. 3, julio 2018, p. 439.

⁶ P. VARONA, *De la neurociencia a la Inteligencia Artificial y vuelta*, en «Revista Congreso de Ciencia y Tecnología», vol. 10, n. 1, 2015, p. 251.

⁷ VARONA, *De la neurociencia a la Inteligencia Artificial y vuelta*, cit., p. 252.

⁸ ZAMBRANO CASTRO, VARGAS LÓPEZ, ZAMBRANO FRANCO, ZAMBRANO FRANCO, *La neurociencia y su relación con la inteligencia artificial*, cit., p. 434.

de datos, de recreación de escenarios virtuales y formulación de modelos de predicción que permite la inteligencia artificial, que ofrece además grandes posibilidades en lo que se conoce como experimentación en ciclo cerrado permitiendo interactuar con el sistema biológico en escalas de milisegundos midiendo variables que permiten analizar su estado interno (algo que es esencial cuando hablamos de sistemas sólo parcialmente observables, como son los neuronales)⁹, o creando seres como el Hybrot, un pequeño robot controlado por una red de neuronas vivas procedentes del cerebro de una rata creado en el Laboratorio de Neuroingeniería del Instituto de Tecnología de Georgia con el fin de investigar en temas de aprendizaje, memoria y procesamiento de la información con fines terapéuticos, aunque también para estudiar cómo se pueden aplicar las redes neuronales vivas a los sistemas informáticos.

Pues bien, como quiera que la inteligencia artificial de la mano de la investigación neurocientífica puede acabar incidiendo en el comportamiento humano y ofreciendo una valiosísima información sobre la psique del sujeto, los escenarios en los que se movería esa relación inteligencia artificial - neurociencia - Derecho penal serían los siguientes:

- 1º El de la propia discusión de presupuestos básicos de la dogmática penal, como la libertad del ser humano como fundamento de la culpabilidad, con todas las implicaciones que ello tiene sobre la propia concepción del Derecho penal (un Derecho penal de la culpabilidad o de la peligrosidad; de la pena o de la medida de seguridad), o en el de la delimitación entre acción y culpabilidad, partiendo del presupuesto de que existe una zona fronteriza entre ambas de difícil delimitación que hace que cuanto más se amplíen los supuestos de ausencia de acción, más se restrinja la inimputabilidad, y viceversa¹⁰.
- 2º El de la creación de herramientas de inteligencia artificial que interaccionarán con el sistema neurológico del propio individuo pudiendo determinar algunas de sus decisiones: si esas decisiones inducidas por el sistema de inteligencia artificial que interactúa con el sistema neurológico del individuo tienen carácter delictivo, necesariamente habremos de plantearnos en qué medida la capacidad de acción o la culpabilidad del sujeto se encontraban mediatizadas.
- 3º El de la prevención, enjuiciamiento y tratamiento del delito a través de las posibilidades que brindaría un conocimiento más profundo del

⁹ VARONA, *De la neurociencia a la Inteligencia Artificial y vuelta*, cit., p. 254.

¹⁰ M. ALONSO ÁLAMO, *Acción, capacidad de acción y capacidad de culpabilidad: problemas de delimitación*, en «Cuadernos de Política Criminal», n. 131, 2020, pp. 12-13.

cerebro humano unido a la utilización de dispositivos capaces de inducir, reproducir e interpretar determinadas reacciones cerebrales de un sujeto, permitiendo así adelantar la intervención policial, llegar a conclusiones con base científica sobre su participación en un hecho delictivo, o determinar con precisión las posibilidades de reincidencia de ese sujeto.

2. Escenario 1: incidencia de los estudios neurocientíficos en la libertad como fundamento de la culpabilidad

Durante mucho tiempo los partidarios de un Derecho Penal de la culpabilidad por el hecho han refutado la indemostrabilidad del libre albedrío recurriendo, entre otros, a argumentos del orden de que tampoco era demostrable lo contrario, es decir, que el actuar humano estuviese presidido por el determinismo. Hoy, sin embargo, reputados autores sostienen, desde el ámbito de las neurociencias, que éstas habrían desmontado tal axioma, y que la hipótesis de la no existencia de libertad es la que mejor se adapta a los resultados obtenidos experimentalmente.

Estos estudios parten de los experimentos que llevase a cabo Benjamin LIBET en la primera mitad de la década de los ochenta del siglo pasado y que encontraron continuidad en otros autores como HAGGARD¹¹ y EIMER¹² en el Reino Unido o HAYNES¹³ en Alemania. Estos estudios pusieron de manifiesto que en el proceso de toma de decisión y realización de un movimiento corporal había actividad cerebral inconsciente antes de la actividad consciente. Es decir, que antes de que el sujeto tomase de manera consciente la decisión de realizar un movimiento corporal (en los experimentos de LIBET pulsar un botón), su inconsciente había impulsado tal movimiento.

El propio LIBET¹⁴ manifestaría, no obstante, que las conclusiones extraídas

¹¹ P. HAGGARD, *Human volition: towards a neuroscience of will*, en «*Nature reviews. Neuroscience*», vol. 9, diciembre 2008, pp. 934-946.

¹² P. HAGGARD, M. EIMER, *On the relation between brain potentials and the awareness of voluntary movements*, en *Experimental Brain Research* (1999), pp. 128-133, que parecen cuestionar el método seguido por LIBET sugiriendo que deducir la dirección de la causalidad mente-cuerpo sólo sobre la base de la discrepancia temporal resulta complicado por la dificultad de cronometrar con precisión tanto los inicios neuronales como las experiencias subjetivas, llegando a proponer lo que consideran un método más adecuado como es el de las variaciones. Vid. HAGGARD, EIMER, cit., p. 132.

¹³ C.S. SOON, M. BRAAS, H.J. HEINZE, J.D. HAYNES, *Unconscious determinants of free decisions in the human brain*, en «*Nature Neuroscience*», 11, 2008, pp. 543-545.

¹⁴ LIBET, cit., p. 538.

de los análisis experimentales llevados a cabo, ni excluían la responsabilidad individual, ni el libre albedrío; y que aunque el proceso volitivo podía ser iniciado por actividades cerebrales inconscientes, seguía siendo posible el control consciente de los actos voluntarios, de manera que los resultados no debían ser tomados como antagónicos al libre albedrío, sino que sólo afectaban la visión de cómo podría funcionar ese libre albedrío.

En efecto, no parece lo mismo pulsar un botón ante un estímulo visual, o no rozar una alambrada porque sabemos que nos producirá una descarga eléctrica, que tomar una decisión más compleja necesitada de una previa valoración. Por ello, cabe pensar que experimentos como el de LIBET se refieren al funcionamiento de un nivel neuronal muy básico, incapaz de aportar datos significativos respecto a los procesos de decisión del ser humano en contextos normativos¹⁵. Más aún: como señalan algunos autores¹⁶, si ser libre es ser dueño de lo que se hace y organizarlo de acuerdo con una razón, experimentos en los que un sujeto se limita a pulsar un botón u otro sin razones que le lleven a decidirse por uno de ellos, lo único que demostrarían es que cuando no se actúa según un esquema racional, el hombre sí estaría a merced de sus condicionamientos orgánicos.

Pese a este tipo de respuestas y a las dudas del propio LIBET, sus estudios han sido el punto de partida de diversos autores, entre los que destacan ROTH¹⁷, PRINZ¹⁸ y SINGER¹⁹, que, a partir de éste y de otros avances neurocientíficos, vienen cuestionando la libertad en la toma de decisiones y, en algunos casos, la propia responsabilidad penal individual de la forma en la que ha venido siendo tradicionalmente formulada. El planteamiento sería el siguiente: si es el inconsciente quien toma realmente la decisión de llevar a cabo una acción, no podría seguir efectuándose reproche culpabilístico a partir de una libertad que es negada por la evidencia empírica. Y si la libertad es una ficción, el mal podría ser un fenómeno biológico residente en el

¹⁵ M. PÉREZ MANZANO, *El tiempo de la conciencia y la libertad de decisión: bases para una reflexión sobre Neurociencia y responsabilidad penal*, en *Neurociencias y Derecho Penal. Nuevas perspectivas en el ámbito de la culpabilidad y tratamiento jurídico-penal de la peligrosidad*, dirigido por E. Demetrio Crespo, Edisofer, Madrid, 2013, p. 123.

¹⁶ J.M. GIMÉNEZ AMAYA, J.I. MURILLO, *Neurociencia y libertad. Una aproximación interdisciplinar*, en «*Scripta Theologica*», 41, n. 1, 2009, p. 30.

¹⁷ G. ROTH, *Fühlen, Denken, Handeln. Wie das Gehirn unser Verhalten steuert*, Frankfurt am Main, 2001.

¹⁸ W. PRINZ, *Kritik des freien Willens: Bemerkungen über eine soziale Institution*, en «*Psychologische Rundschau*», 55. Jg., Heft 4, 2004, pp. 198-206.

¹⁹ W. SINGER, *Verschaltungen legen uns fest. Wir sollten aufhören, von Freiheit zu sprechen*, en *Hirnforschung und Willensfreiheit: Zur Deutung der neuesten Experimente*, dirigido por C. Geyer, Frankfurt am Main, 2004.

cerebro²⁰ y no una elección del propio sujeto. Un mal que tal vez justificase medidas de intervención penal predelictual a partir de juicios de prognosis basados en la actividad cerebral, como forma de evitar que el 'yo inconsciente' tomase la decisión de cometer un delito.

Como puede verse, con esto lo que se viene a cuestionar son las propias bases del modelo tradicional de Derecho Penal liberal²¹ acentuando, de este modo, la crisis en la que, a mi juicio, se encuentran hoy día algunos de los postulados básicos de este modelo. Una crisis que, desde luego, no invita a mirar con excesivo optimismo el panorama penal y político-criminal de los próximos años.

Se podrá contestar que éste es un planteamiento extremo al que no conducen necesariamente experimentos como los realizados por LIBET²². Es cierto, pero también lo es que cuestionando la libertad del sujeto en la toma de decisiones y el papel que ésta está llamada a desempeñar en el Derecho Penal, se abre una puerta por la que una sociedad alarmada y un gobernante complaciente pueden acabar exigiendo e imponiendo, respectivamente, modelos como el mencionado²³.

Pues bien, ¿es capaz la neurociencia, auxiliada por la inteligencia artificial, de destruir la base de la culpabilidad penal?

²⁰ E. DEMETRIO CRESPO, *Libertad de voluntad, investigación sobre el cerebro y responsabilidad penal*, en «InDret. Revista para el análisis del Derecho», n. 2, 2011, p. 3.

²¹ Vid. un pronóstico de las consecuencias que tendría la negación de la libertad en el Derecho Penal que conocemos, en B. ROMERO FLORES, *Las neurociencias frente a la función de la pena*, en «Anuario de Derecho Penal y Ciencias Penales», vol. 58, 2015, pp. 340-344.

²² DEMETRIO CRESPO, "Compatibilismo humanista": *Una propuesta de conciliación entre Neurociencias y Derecho Penal*, en *Neurociencias y Derecho Penal. Nuevas perspectivas en el ámbito de la culpabilidad y tratamiento jurídico-penal de la peligrosidad*, dirigido por E. Demetrio Crespo, Madrid, 2013, pp. 30-31.

²³ En este sentido, PRINZ, uno de los autores que más modula el alcance de los avances neurocientíficos consciente de la importancia de mantener el libre albedrío como base y garantía del modelo de sociedad, señala que «[...] hablar de libre albedrío refleja discursos e instituciones para el desarrollo político de la voluntad. En la medida en que los agentes de los colectivos sociales se atribuyen mutuamente la autonomía y la responsabilidad de sus decisiones privadas de actuar, también aplicarán estas adscripciones a las decisiones públicas del colectivo mismo. Como consecuencia, reemplazarán el liderazgo autoritario por mecanismos de desarrollo de la voluntad colectiva. En el caso de las democracias occidentales modernas, estos mecanismos están respaldados por la noción de un contrato social entre individuos iguales y autónomos y encarnado en formas democráticas de desarrollo de la voluntad en varios niveles de la sociedad. La idea de democracia se basa en gran medida en la noción de autonomía personal y libre albedrío. Si apreciamos a uno, debemos alabar al otro». PRINZ, *Open Minds: The Social Making of Agency and Intentionality*, Massachusetts, 2012, p. 190.

Para ello, lo primero que deberíamos preguntarnos es si realmente un libre albedrío entendido como rasgo biológico del ser humano (que es el terreno en el que puede moverse la neurociencia), es o debe ser el fundamento de la culpabilidad. En esta pregunta no debemos ver, en ningún caso, una negación o rechazo de los avances científicos. Supone, simplemente, plantearnos si el Derecho en general, y el Derecho Penal en particular, gozan de un estatus propio de científicidad frente a las ciencias empíricas y si existe una metodología jurídica frente a la de las ciencias empíricas que resulte compatible con el conocimiento empíricamente demostrable. Supone, en definitiva, plantearnos si el objeto de estudio del Derecho Penal debe abordarse de forma puramente valorativa, o sobre los conocimientos ofrecidos desde el empirismo.

Resulta evidente que el Derecho no es una ciencia empírica ni por su objeto ni por su método. Ni siquiera el positivismo jurídico consiguió trasponer al terreno de lo jurídico la metodología experimental de las ciencias naturales. Tampoco es una ciencia factual o una ciencia formal al estilo de la lógica o de las matemáticas.

Siendo, no obstante, conocimiento científico, lo cierto es que el carácter de ciencia cultural o valorativa del Derecho Penal condiciona, en lo que hace al método, la interconexión que pueda presentar con otras ciencias, éstas ya sí, de carácter empírico. El Derecho Penal cuenta con un objeto de estudio puramente normativizado, el delito, en el que hoy prácticamente no encontramos referencias ontológicas, y en el que el papel de las ciencias empíricas queda relegado a ofrecer información que será sometida a un proceso de análisis jurídico para extraer conclusiones sobre si concurren o no, en un caso concreto, los conceptos valorativos sobre los que hemos construido ese supraconcepto que es el propio delito. De este modo, el papel de las ciencias empíricas no sería tanto condicionar el propio concepto jurídico penal, como permitir constatar si ese concepto jurídico penal que hemos creado desde el normativismo concurre o no en el caso concreto.

Los datos empíricos ‘brutos’ (siguiendo el acertado término empleado por PÉREZ MANZANO²⁴) nos permiten saber si un vehículo llevaba o no los neumáticos desgastados en un día de lluvia y ello provocó su salida de la vía y el atropello mortal del peatón, un dato que valoraremos para concluir que se generó de manera imprudente un riesgo ilícito que se materializó en un resultado; o que la causa determinante de la muerte fueron los múltiples golpes dados con un objeto contundente en la cabeza de la víctima, dato

²⁴ PÉREZ MANZANO, *Fundamento y fines del Derecho penal. Una revisión a la luz de las aportaciones de la neurociencia*, en «InDret. Revista para el análisis del Derecho», n. 2, 2011, p. 8.

que sometido a un proceso de valoración nos permitirá afirmar que el comportamiento fue doloso; o si un sujeto sufre algún tipo de enfermedad mental que en el momento de cometer el hecho le impedía tener conciencia de la realidad en la que vivía, dato que nos llevará a considerarlo inimputable, etc. Sin embargo, estos datos 'brutos' (y los ofrecidos por la neurociencia lo son) no nos permitirán construir, a través de un método empírico, conceptos como la libertad o la culpabilidad²⁵.

Más aún: las conclusiones a las que parece llegar el debate neurocientífico negando la libertad en la toma de decisiones y dejándolo todo en manos de procesos neurológicos y bioquímicos que el hombre no es capaz de controlar, lo único que acabaría demostrando no es que no exista el libre albedrío. Lo que viene a demostrar es que no existe el libre albedrío como rasgo biológico del ser humano. Es decir, que no existe un concepto ontológico de libertad, por lo que ésta sólo podría ser concebida como categoría normativa. En este sentido, debemos destacar la posición mantenida al respecto por el propio PRINZ²⁶, para quien el libre albedrío es una institución social, hecha por las personas y para las personas, y que al igual que la mayoría de las instituciones sociales, se basa en intuiciones convincentes que se comparten y se comunican entre los individuos²⁷. En este caso, se trata de la intuición, el sentimiento de ser libres, algo que PRINZ considera un hecho empírico que, como cualquier otro, puede explorarse y explicarse utilizando las herramientas de la Psicología, cosa distinta a lo que ocurre con el constructo teórico del libre albedrío.

El problema surge, según este autor, cuando pasamos de ver el libre albedrío como un fenómeno empírico a entenderlo como una construcción teórica, manteniendo que nos sentimos libres (en un sentido empírico) porque somos libres (en un sentido teórico)²⁸ y no al revés. Para PRINZ, no tenemos libre albedrío en el mismo sentido que tenemos órganos como un hígado o un corazón, o habilidades como ver o agarrar. Pero, aunque no tengamos libre albedrío en el sentido de un hecho natural, lo tenemos en el sentido de un artefacto social. El libre albedrío es un artefacto que creamos para nosotros mismos e integramos en el sentimiento de nuestro yo social y mental. Desde esta perspectiva, en ningún caso estamos ante algo ficticio o ilusorio, sino que es algo tan real como los hechos naturales, que gobierna y

²⁵ F. MORALES PRATS, *Precisiones conceptuales en torno a la culpabilidad: convenciones normativas y función individualizadora*, en *El nuevo Código penal: presupuestos y fundamentos. Libro homenaje al Profesor Doctor Don Ángel Torío López*, Granada, 1999, p. 173.

²⁶ PRINZ, *Open Minds: The Social Making of Agency and Intentionality*, cit., pp. 174, ss.

²⁷ PRINZ, cit., p. 188.

²⁸ PRINZ, cit., p. 176.

limita nuestras acciones con la misma lógica que los hechos naturales²⁹.

Siendo, pues, una categoría normativa incluso para quienes, como PRINZ, se pronuncian desde el terreno de la Psicología, hay 'legitimidad científica' desde el Derecho para llenarla de contenido en la dirección que las relaciones de convivencia demanden.

Para estas relaciones de convivencia, la idea de libertad es esencial. No sólo como manifestación de la dignidad humana, sino también como presupuesto de la exigencia de responsabilidad por el incumplimiento de las normas y de los compromisos adquiridos entre particulares, que es base de la propia convivencia. Y es esta perspectiva de las relaciones de convivencia la que nos ha llevado a entender la libertad como la facultad (y no tanto capacidad) que tiene el hombre para obrar de una manera o de otra, o para no obrar, a partir de la capacidad intelectual necesaria para conocer el sentido y el alcance de sus actos u omisiones. Un concepto de libertad normativizado que se presume en todo sujeto por el mero hecho de serlo y como manifestación de la dignidad humana.

¿Tiene esta libertad una base empírica? Los datos 'brutos' que nos llegan desde los estudios neurocientíficos, parecen afirmar que esa libertad no concurre en ningún ser humano al no tener capacidad real de decidir de manera consciente. Sin embargo, esos mismos datos 'brutos' confirman, como vimos antes, que hay algo que sí concurre en el hombre, y es el hecho de sentirse libre, aunque esté equivocado al respecto y sólo sea una ilusión.

HAGGARD³⁰, uno de los continuadores de los estudios de LIBET, señala que la responsabilidad puede depender de los motivos que desencadenaron un proceso neuronal que culminó en una acción, y de si podría haber existido algún tipo de control final que hubiera permitido detener dicha acción. Entiende este autor que ambas decisiones tienen un fuerte elemento normativo, pues, aunque el cerebro de una persona decide las acciones que lleva a

²⁹ PRINZ considera que las intuiciones socialmente compartidas son tan reales como los hechos naturales, algo que obedece a que su estado se deriva del hecho de que las intuiciones que reflejan interpretaciones personales de procesos subpersonales deben, a su vez, ser apoyadas también por procesos subpersonales. En consecuencia, entiende este autor, no hay ninguna razón para considerarlos menos reales o eficaces que la maquinaria subpersonal de toma de decisiones a cuya operación se refieren. Los discursos y las prácticas sociales aseguran que las nociones como el yo, la autoría y el libre albedrío se implementen en las estructuras de conocimiento individuales en términos de intuiciones personales y mecanismos subpersonales. Como resultado, las intuiciones individuales de uno mismo, la autoría de la acción, y libre voluntad tienen precisamente el mismo potencial causal para influir en su comportamiento como lo hacen sus intuiciones sobre el funcionamiento de otras cosas y eventos en el mundo. PRINZ, cit., pp. 188-190.

³⁰ HAGGARD, *Human volition: towards a neuroscience of will*, cit., p. 944.

cabo, la cultura y la educación enseñan a las personas qué motivos resultan aceptables para actuar, cuáles no, y cuándo un control predictivo final debe recomendar una acción de retención. La cultura y la educación, por lo tanto, constituyen, a su juicio, poderosas señales de aprendizaje para los circuitos cognitivo-motores del cerebro. Un enfoque neurocientífico de la responsabilidad puede depender, por tanto, no solo de los procesos neuronales que subyacen a la voluntad, sino también de los sistemas cerebrales que le dan al individuo la capacidad cognitiva general para comprender cómo la sociedad restringe la volición y cómo adaptarse adecuadamente a esas restricciones.

Pues bien, desde el punto y hora en que el hecho de sentirnos libres (aun sin serlo) fuese capaz, como parte de nuestra cultura y nuestra educación, de incidir en nuestra forma de actuar y, por tanto, en nuestra toma -aun inconsciente- de decisiones, permitiendo ese control final, el que fuésemos o no realmente libres pasaría a ser algo secundario a efectos de fundamentar la responsabilidad penal. Lo determinante sería que nos sintiéramos así, y no tanto que lo fuésemos. Respondiendo a esta pregunta, PRINZ sostiene que la intuición de libertad que está presente en el propio individuo sí es capaz de incidir en su comportamiento³¹ introduciendo alteraciones en los algoritmos que subyacen a nuestras decisiones que resultan en una elaboración de los procesos de toma de decisiones³². Es decir, uno de los principales efectos del sentimiento de ser libres sería la capacidad de frenar el sistema de toma de decisiones subpersonal, que está diseñado para ser rápido y eficiente.

Planteadas la cuestión del libre albedrío en estos términos, y teniendo en cuenta que la culpabilidad también aparece como un concepto normativo³³, nada impediría que esa culpabilidad, pese a los avances neurocientíficos, se pudiese seguir construyendo sobre la libertad. No una libertad como rasgo biológico (que los datos 'brutos' proporcionados por las neurociencias nos dicen que no concurre en ningún ser humano), sino como rasgo social del ser humano, que se siente libre (sentimiento que esos mismos datos 'brutos' sí nos confirman que concurre y que es capaz de incidir en nuestra toma de decisiones -aun inconsciente-). Es lo que propone, por ejemplo, BURKHARDT³⁴, para quien lo decisivo es si el sujeto actuó en la creencia de que

³¹ PRINZ, *Open Minds: The Social Making of Agency and Intentionality*, cit., pp. 188-189.

³² PRINZ, cit., p. 189.

³³ Sobre la naturaleza normativa de la culpabilidad ante las dificultades de verificación empírica, vid. DÍEZ RIPOLLÉS, *Los elementos subjetivos del delito. Bases metodológicas*, Valencia, 1990, pp. 91-104.

³⁴ B. BURKHARDT, *La comprensión de la acción desde la perspectiva del agente en el Derecho Penal*, en B. BURKHARDT, K. GÜNTHER, G. JAKOBS, *El problema de la libertad de acción en el Derecho Penal*, Buenos Aires, 2007, p. 45.

tenía alternativas a su comportamiento, toda vez que ese sentimiento sería capaz de influir en su comportamiento, presupuesto suficiente para construir el reproche culpabilístico³⁵.

Es decir, aun pudiendo haber probado que el libre albedrío no existe como realidad biológica, lo que las neurociencias no han conseguido demostrar es que exista un determinismo genético que conduzca causalmente al sujeto en una concreta dirección de manera inexorable y con independencia de las circunstancias que le rodean. Puede que el hombre carezca de libertad en el sentido de que su yo inconsciente se adelante a la hora de tomar las decisiones. Pero ese yo inconsciente decide sobre la base de algoritmos que operan con múltiples factores de tipo social, educativo, situacional, religioso, etc. Entre esos factores figura el propio valor de la libertad y ese sentimiento de ser libre propio del ser humano y fuertemente arraigado en los regímenes democráticos. De este modo, ese sentimiento configuraría el yo inconsciente y coadyuvaría, junto con otros factores, a dirigir su toma de decisiones en un sentido u otro. Una acción libre es, así, una acción ejecutada sabiendo que se podría haber actuado de otra manera³⁶. Interpretado de este modo, lo determinante a efectos de culpabilidad no sería tanto si existe o no la libertad desde un punto de vista biológico, o si el sujeto es capaz de tomar una decisión de modo consciente, o si es su subconsciente el que opera. Lo determinante es que el sujeto se sienta libre; que actúe en la creencia de que tenía alternativas a su comportamiento, toda vez que ello puede condicionar el funcionamiento de su yo inconsciente.

A partir de esta interpretación normativa del libre albedrío, podemos seguir admitiendo el control voluntario de nuestro hacer o no hacer como elemento constitutivo del concepto penal de acción y la capacidad para

³⁵ En contra, DEMETRIO CRESPO, *Libertad de voluntad, investigación sobre el cerebro y responsabilidad penal*, cit., p. 14, para quien el juez debe atenerse a hechos objetivos. DEMETRIO CRESPO, cit., p. 15, siguiendo a MERKEL y a ROTH utiliza, para justificar su posición, el ejemplo del esquizofrénico al que, por mucho que afirme que en el momento de realización del hecho accedió voluntariamente a seguir la voz que le decía que debía matar a alguien, el juez le declarará inimputable. Sin embargo, en el caso del esquizofrénico hay, a mi juicio, una creencia de libertad errónea basada en un conocimiento erróneo: toma una decisión en ejercicio de una libertad inducida a error por su enfermedad mental; es decir, hay una percepción errónea de estar actuando en libertad, que es base suficiente para excluir el reproche culpabilístico por concurrir una causa de inimputabilidad.

³⁶ L. FÄH, S. RAINER, M. KILLIAS, *¿Un nuevo determinismo? La exclusión de las relaciones probabilísticas y de las influencias situacionales en los enfoques neurocientíficos*, en *Derecho penal y Criminología como fundamento de la Política criminal: estudios en homenaje al profesor Alfonso Serrano Gómez*, Dykinson, Madrid, 2006, p. 236.

adecuar nuestro comportamiento a la comprensión de la ilicitud del hecho como elemento constitutivo de la imputabilidad cuestionando, de este modo, el papel que una neurociencia auxiliada por la inteligencia artificial vendría a ser capaz de desarrollar³⁷.

3. *Escenario 2: creación de herramientas de inteligencia artificial capaces de interactuar con el propio individuo condicionando su comportamiento*

NEIL HARBISSEON, en una entrevista narra la siguiente experiencia: «Me conecté a un WiFi público y a través de esta red alguien me mandó una imagen a la cabeza. Yo recibo las imágenes como vibraciones de diferentes colores. Y este fue el único caso en que alguien sin permiso me mandó una imagen en la cabeza. Fue un hackeo físico, pero me gustó. No fue una mala experiencia, pero pudo serlo». NEIL HARBISSEON, que nació pudiendo ver sólo la escala de grises, es la primera persona en el mundo reconocida como cyborg por un gobierno. Tiene una antena implantada en la cabeza que le permite ver y percibir colores invisibles como infrarrojos y ultravioletas, así como recibir imágenes, videos, música o llamadas telefónicas directamente a su cabeza desde aparatos externos.

La interacción con el sistema neurológico del individuo mediante la

³⁷ Con todo, negando el papel de las neurociencias en la construcción de los conceptos de libertad y culpabilidad penal, esa función instrumental al servicio de lo normativo que antes le reconocíamos resulta fundamental en la medida en que arroje luz sobre la concurrencia o no de tales conceptos en el caso concreto y nos permitan una mejor comprensión de nuestro comportamiento ayudándonos al enjuiciamiento de la capacidad de culpabilidad en el caso concreto. Una posición así era defendida por PERIS RIERA ya a mediados de los años noventa, refiriéndose a la incidencia que podían tener sobre el Derecho Penal los avances de los estudios genéticos. En este sentido, PERIS cuestionaba que tales avances pudieran poner en entredicho la propia idea jurídica de culpabilidad, pero sí les reconocía un indudable valor para el enjuiciamiento de la capacidad de culpabilidad en casos concretos. J.M. PERIS RIERA, *La regulación penal de la manipulación genética en España (Principios penales fundamentales y tipificación de las genotecnologías)*, Madrid, 1995, pp. 31-36. En la misma línea, A. URRUELA MORA, *La genética como factor relevante a efectos del juicio de imputabilidad penal. Estado de la cuestión y perspectiva crítica. Particular consideración de la Sentencia de la Corte de Apelación de Trieste (Italia) de 18 de septiembre de 2009*, en «Revista de Derecho y Genoma Humano», n. 32, enero-junio 2010, pp. 178-179.

creación de cyborgs³⁸ o el enhancement neurocognitivo³⁹ no sólo puede suponer peligros para el sujeto portador del elemento tecnológico, sino que podría llegar a ser determinante del actuar de éste en caso de tratarse de auténticos sistemas de inteligencia artificial capaces de tomar decisiones de manera autónoma. En estos casos, dependiendo de en qué medida el sistema llegase a tomar el control del sujeto, podríamos encontrar ante situaciones, no tanto de ausencia de culpabilidad como de ausencia de acción a partir de un concepto típico de acción entendido como comportamiento (activo u omisivo) humano (controlado, en consecuencia, por la voluntad del sujeto) subsumible en un tipo penal. En realidad, cualquier concepto de acción del que se parta toma como presupuesto el control por la voluntad del sujeto que sienta las bases de la exclusión de la tipicidad en lo que son meros movimientos corporales que se producen de forma puramente mecánica, esto es, sin un control consciente por parte del sujeto, así como en aquellos casos en los que la paralización del sujeto escapa, igualmente, a cualquier decisión consciente del individuo. En estos casos, además, si el dispositivo de inteligencia artificial hubiese sido hackeado para provocar el movimiento mecánico o se hubiese previsto o debido prever lo acaecido, nada impediría exigir responsabilidad a un tercero recurriendo a la autoría mediata o incluso al propio sujeto recurriendo a las *actio libera in causa*.

Distinta sería la situación en aquellos casos en los que la interacción del sistema de inteligencia artificial lo que hiciese fuese anular o limitar la capacidad del sujeto para comprender la ilicitud del hecho o para actuar conforme a esa comprensión, situación en la hablaríamos de inimputabilidad o semiimputabilidad que habría que tratar de reconducir, con la actual redacción del Código penal, al art. 20.1 del Código penal, planteándose la cuestión de si estaríamos ante una anomalía o alteración psíquica o ante una situación de trastorno mental transitorio. Partiendo de la enorme dificultad de delimitar la anomalía de la alteración psíquica exigiendo a la primera una base patológica innecesaria en la segunda (como propone algún autor) y de que la doctrina y jurisprudencia vienen decantándose más por una utilización

³⁸ Los cyborgs son seres compuestos por elementos orgánicos y dispositivos cibernéticos buscándose con estos últimos normalmente mejorar las capacidades de la parte orgánica mediante el uso de la tecnología. A.T. RABADÁN, *Horizonte de la Inteligencia Artificial y Neurociencias. Acerca de robots, androides y cyborgs*, en «Jornal Brasileiro de Neurocirurgia», 32, n. 1, 2021, p. 15.

³⁹ Aumento de las facultades neurocognitivas o de los fenómenos sensoriales mediante la cibernética, específicamente a través de la estimulación transcraneal y la intracraneal. RABADÁN, *Horizonte de la Inteligencia Artificial y Neurociencias. Acerca de robots, androides y cyborgs*, cit., p. 16.

indiscriminada de ambos términos, me ofrece serias dudas la posibilidad de afirmar la existencia de una psicosis exógena, creyendo más adecuada la solución del trastorno mental transitorio que permitiría, además, recurrir, una vez más, a las *actio libera in causa* cuando se hubiese previsto o debido prever la comisión del hecho delictivo.

Esta interacción de la máquina con el humano puede, asimismo, provocar en el sujeto erróneas percepciones de la realidad de las cosas susceptibles de integrar situaciones de error de tipo que merecerán el tratamiento previsto para ellas en el art. 14.1 del Código penal. Así, por ejemplo, el dispositivo tecnológico de percepción sensorial que transmite una imagen falsa al cerebro de quien lo lleva implantado (un animal, cuando en realidad se trata de una persona). Estas pueden ser las situaciones más frecuentes en el caso del metaverso como mundo en el que se inserta el sujeto y que, aunque virtual, puede acabar percibiéndose como una auténtica realidad paralela.

4. Escenario 3: prevención, enjuiciamiento y tratamiento del delito

Hasta ahora, el planteamiento que se ha hecho de la inteligencia artificial en este contexto es el de la utilización de sistemas que permiten a los cuerpos policiales ‘anticipar’ la comisión del hecho delictivo adelantándose al delincuente, pero sobre la base de meros cálculos estadísticos a partir de un número elevadísimo de variables y datos que una persona física no tiene capacidad para llevar a cabo, pero que un sistema de inteligencia artificial dotado de unos algoritmos adecuados, sí. Entre esos datos, los derivados de estudios del comportamiento humano. Sin embargo, un conocimiento más profundo del cerebro humano unido a la utilización de dispositivos capaces de leer e interpretar determinadas reacciones cerebrales de un sujeto, permitiría adelantar esa intervención policial preventiva con mucha más precisión de lo que lo permiten hasta ahora las herramientas existentes. De hecho, teniendo en cuenta que existen ya existen mecanismos de análisis facial capaces de identificar, por los gestos de su cara, al individuo que va a sustraer algún artículo en un establecimiento, permitiendo así que ese sujeto sea expulsado del mismo, no cuesta imaginar lo que podría suponer un sistema que en vez de los gestos faciales fuese capaz de detectar e interpretar determinados estímulos cerebrales de la persona. Supondría abrir la mente del sujeto a la intervención policial anticipatoria, pero también abonar el terreno a un cambio en la configuración de las medidas de seguridad que permitiese prescindir de la comisión de un delito previo tal y como exige

el art. 95.1.1^a del Código penal. Y es que si hoy esa comisión de un delito previo viene a ser el medio con el que contamos para acreditar que el sujeto es peligroso, disponiendo de herramientas capaces de acreditar esa peligrosidad con base científico-neurológica, ¿qué necesidad habría ya de esperar a que el sujeto causase el daño que comporta la comisión de un delito para imponerle una medida de seguridad?

En cuanto al proceso penal, las posibilidades que brinda la combinación de neurociencia e inteligencia artificial son, asimismo, enormes, aun cuando resultan difícilmente compatibles con los principios limitadores del ius puniendi y las garantías a las que está sujeto el proceso penal. Así, el conocimiento cerebral de un sujeto podría permitirnos saber qué tipo de reacción provocan en él ciertos estímulos y hasta dónde es capaz de llegar en determinadas circunstancias, de manera que podríamos anticipar su comportamiento y decidir de manera más precisa sobre su prisión provisional. Además, si recreamos mediante inteligencia artificial las condiciones en las que se cometió el delito, podríamos obtener conclusiones rayanas en la certeza sobre su intervención en el delito, siendo capaces, asimismo, de llegar a conclusiones más certeras sobre una posible ausencia de acción o sobre el estado de imputabilidad o inimputabilidad en el que cometió el hecho en cuestión: cómo se determina la resistibilidad o no de la fuerza a propósito de la vis absoluta; o si se mantiene un mínimo de voluntariedad en el actuar impulsivo frente al movimiento reflejo; o si hay capacidad real de oposición en el caso de hipnosis, son preguntas cuya respuesta podría estar más cerca con las consecuencias que ello tendría en la calificación penal de unos hechos. Al mismo tiempo, la neurociencia auxiliada de la inteligencia artificial puede profundizar en el conocimiento de algunas enfermedades, de forma que lo que hasta ahora consideramos excluyente de la culpabilidad podría acabar siendo reconducido a la ausencia de acción⁴⁰. Y, del mismo modo, la difícil delimitación entre la imputabilidad, semiimputabilidad e inimputabilidad podría tener los días contados al poder recrear las condiciones en las que el sujeto cometió el hecho y reproducir el mismo estado mental en el que se encontraba en ese momento, lo que permitiría analizarlo en tiempo real.

Es lo que ocurre, por ejemplo, con las investigaciones neurológicas llevadas a cabo en el campo de las psicopatías, pues aunque existe un alto consenso sobre las características afectivas, interpersonales y comportamentales definitorias de los psicópatas, el diagnóstico de una personalidad psicopática no siempre es fácil, realizándose en ocasiones ese diagnóstico se realiza con

⁴⁰ ALONSO ÁLAMO, *Acción, capacidad de acción y capacidad de culpabilidad: problemas de delimitación*, cit., p. 42.

bastante ligereza⁴¹. Pues bien, los estudios neurocientíficos podrían, con la ayuda de la inteligencia artificial, no sólo facilitar el diagnóstico, sino también arrojar luz sobre cómo funciona realmente el cerebro del psicópata y en qué medida esa personalidad psicopática afecta la capacidad intelectual y/o volitiva del sujeto.

Además, el conocimiento neurológico del individuo y la capacidad de anticipar sus reacciones ante determinados estímulos sería determinante en la fase de ejecución de la pena tanto a la hora de decidir sobre su suspensión y, en su caso, sobre la concesión de la libertad condicional, como a la hora de otorgar permisos, clasificar en un determinado grado penitenciario o, por qué no, a la hora de sustituir medidas de seguridad como la libertad vigilada por mecanismos electrónicos que dirijan el comportamiento del sujeto o, al menos, hagan sonar las alarmas cuando detecten determinadas reacciones cerebrales que puedan considerarse peligrosas por predelictivas. Más aún, ya en el terreno de la privación de libertad, bien a través de la pena o de una medida de seguridad, poder contar con un método de diagnóstico fiable resultaría esencial en el momento de la ejecución de una u otra con vistas a seleccionar aquellos programas de tratamiento que puedan resultar más adecuados atendiendo a la personalidad del sujeto.

Pese a todo, los problemas que la utilización de estas técnicas podría encontrar en el proceso penal son tan grandes que, hoy por hoy, parecen insuperables. ¿Puede obligarse a un sujeto a someterse a este tipo de estudios? ¿No estaríamos atentando contra su intimidad? ¿No estaríamos atentando contra el derecho fundamental a no declarar contra sí mismo y a no confesarse culpable? Circunstancias como estas me permiten aventurar que, aunque la neurociencia y la inteligencia artificial nos ayudarán a conocer con mayor precisión el funcionamiento cerebral, a anticipar la conducta del sujeto e incluso a recrear su actividad cerebral en circunstancias similares a aquellas en las que se cometió el delito, es difícil -confío en ello- que acabemos viendo exámenes neurológicos individualizados de este tipo en el proceso penal. Desde el punto de vista garantista los obstáculos serían prácticamente insalvables.

⁴¹ No pocas veces, basta con que un delincuente resulte frío, distante o falto de empatía, que haya cometido un delito sin motivación aparente, o de una forma especialmente cruel, para calificarlo de manera automática como psicópata.

5. Bibliografía

- ALONSO ÁLAMO M., *Acción, capacidad de acción y capacidad de culpabilidad: problemas de delimitación*, en «Cuadernos de Política Criminal», n. 131, 2020.
- BURKHARDT B., *La comprensión de la acción desde la perspectiva del agente en el Derecho Penal*, en BURKHARDT B., GÜNTHER K., JAKOBS G., *El problema de la libertad de acción en el Derecho Penal*, Buenos Aires, 2007.
- DEMETRIO CRESPO E., *Libertad de voluntad, investigación sobre el cerebro y responsabilidad penal*, en «InDret. Revista para el análisis del Derecho», n. 2, 2011.
- DEMETRIO CRESPO E., “*Compatibilismo humanista*”: *Una propuesta de conciliación entre Neurociencias y Derecho Penal*, en *Neurociencias y Derecho Penal. Nuevas perspectivas en el ámbito de la culpabilidad y tratamiento jurídico-penal de la peligrosidad*, dirigido por Demetrio Crespo E., Madrid, 2013.
- DÍEZ RIPOLLÉS J.L., *Los elementos subjetivos del delito. Bases metodológicas*. Valencia, 1990.
- DÍEZ RIPOLLÉS J.L., *El nuevo modelo penal de la seguridad ciudadana*, en «Revista Electrónica de Ciencia Penal y Criminología», 06-03, 2004.
- FÄH L., RAINER S., KILLIAS M., *¿Un nuevo determinismo? La exclusión de las relaciones probabilísticas y de las influencias situacionales en los enfoques neurocientíficos*, en *Derecho penal y Criminología como fundamento de la Política criminal: estudios en homenaje al profesor Alfonso Serrano Gómez*, Dykinson, Madrid, 2006.
- GIMÉNEZ AMAYA J.M., MURILLO J.I., *Neurociencia y libertad. Una aproximación interdisciplinar*, en «Scripta Theologica», 41, n. 1, 2009.
- HAGGARD P., *Human volition: towards a neuroscience of will*, en «Nature reviews. Neuroscience», vol. 9, diciembre 2008.
- HAGGARD P., EIMER M., *On the relation between brain potentials and the awareness of voluntary movements*, en *Experimental Brain Research*, 1999, 126.
- JIMÉNEZ DÍAZ M.J., *Sociedad del riesgo e intervención penal*, en «Revista Electrónica de Ciencia Penal y Criminología», 16-08, 2014.
- MORALES PRATS F., *Precisiones conceptuales en torno a la culpabilidad: convenciones normativas y función individualizadora*, en *El nuevo Código penal: presupuestos y fundamentos. Libro homenaje al Profesor Doctor Don Ángel Torío López*, Granada, 1999.
- PÉREZ MANZANO M., *Fundamento y fines del Derecho penal. Una revisión a la luz de las aportaciones de la neurociencia*, en «InDret. Revista para el análisis del Derecho», n. 2/2011.

- PÉREZ MANZANO M., *El tiempo de la conciencia y la libertad de decisión: bases para una reflexión sobre Neurociencia y responsabilidad penal*, en *Neurociencias y Derecho Penal. Nuevas perspectivas en el ámbito de la culpabilidad y tratamiento jurídico-penal de la peligrosidad*, dirigido por Demetrio Crespo E., Edisofer, Madrid, 2013.
- PERIS RIERA J.M., *La regulación penal de la manipulación genética en España (Principios penales fundamentales y tipificación de las genotecnologías)*, Madrid, 1995.
- PRINZ W., *Kritik des freien Willens: Bemerkungen über eine soziale Institution*, en «Psychologische Rundschau», 55. Jg., Heft 4, 2004.
- PRINZ W., *Open Minds: The Social Making of Agency and Intentionality*, Massachusetts, 2012.
- RABADÁN A.T., *Horizonte de la Inteligencia Artificial y Neurociencias. Acerca de robots, androides y cyborgs*, en «Jornal Brasileiro de Neurocirurgia», 32 (1), 2021.
- ROMERO FLORES B., *Las neurociencias frente a la función de la pena*, en «Anuario de Derecho Penal y Ciencias Penales», vol. 58, 2015.
- ROTH G., *Fühlen, Denken, Handeln. Wie das Gehirn unser Verhalten steuert*, Frankfurt am Main, 2001.
- SILVA SÁNCHEZ J.M., *La expansión del Derecho penal. Aspectos de la política criminal en las sociedades postindustriales*, 2ª ed., Civitas, Madrid, 2001.
- SINGER W., *Verschaltungen legen uns fest. Wir sollten aufhören, von Freiheit zu sprechen*, en *Hirnforschung und Willensfreiheit: Zur Deutung der neuesten Experimente*, dirigido por Geyer C., Frankfurt am Main, 2004.
- SOON C.S., BRAAS M., HEINZE H.J., HAYNES J.D., *Unconscious determinants of free decisions in the human brain*, en «Nature Neuroscience», 11, 2008.
- URRUELA MORA A., *La genética como factor relevante a efectos del juicio de imputabilidad penal. Estado de la cuestión y perspectiva crítica. Particular consideración de la Sentencia de la Corte de Apelación de Trieste (Italia) de 18 de septiembre de 2009*, en «Revista de Derecho y Genoma Humano», n. 32, enero-junio 2010.
- VARONA P., *De la neurociencia a la Inteligencia Artificial y vuelta*, en «Revista Congreso de Ciencia y Tecnología», vol. 10, n. 1, 2015.
- ZAMBRANO CASTRO M.F., VARGAS LÓPEZ R.S., ZAMBRANO FRANCO E.H., ZAMBRANO FRANCO K.A., *La neurociencia y su relación con la inteligencia artificial*, en «Revista Científica de Investigación actualización del mundo de las Ciencias», vol. 3, n. 3, 2018.

CAPÍTULO X

*Imputabilidad y delito**

Mario Trapani

SUMARIO: 1. Imputabilidad y teoría del delito – 2. El “contenido” de la imputabilidad – 3. Imputabilidad e inteligencia artificial – 4. Bibliografía.

1. *Imputabilidad y teoría del delito*

Ante todo, quiero agradecerles a los amigos y colegas de la Universidad de Murcia, especialmente al Profesor Jaime Peris, así como a mi discípula, la Profesora Antonella Massaro, por haber organizado este interesante congreso.

Quienes me conocen, saben perfectamente que yo no soy políticamente correcto y no sigo el *main stream*, como se conoce hoy en día; así que algunas de las cosas que quiero destacar, pueden no estar conforme a los dogmas del “pensamiento único”, también a los pertenecientes a la ciencia del derecho penal.

En primer lugar, el Derecho es un conjunto de reglas. “Regla de conducta” significa que el destinatario del mandato o de la prohibición puede obedecer o no (es decir, cumplir o no con) su contenido prescriptivo. Este es el presupuesto lógico, no solo del Derecho Penal, sino del Derecho *tout court*. El problema primario en la teoría general de la norma jurídica es establecer si el destinatario de la norma es el ciudadano o el órgano que aplica la consecuencia jurídica, o bien si hay dos diferentes tipos de normas según su destinatario específico. Más particularmente, estando en vigor en la actualidad, según el derecho positivo de todos los ordenamientos de *civil law*, como principio general el de la concreta “dominabilidad” del hecho delictivo cometido por parte de su autor y desarrollando la amenaza de la pena criminal la función de prevención general negativa, de eso se deduce

* El texto de este capítulo corresponde a la ponencia presentada en el Congreso internacional “Inteligencia artificial y neurociencias: avances del Derecho Penal contemporáneo”, celebrado el 5 de mayo de 2022 en la Universidad de Murcia.

que la estructura de la norma penal incriminadora es lógicamente aquella de “imperativo”. Entonces, el derecho penal, en una estricta perspectiva de derecho positivo (y no solo de teoría general del derecho), es un conjunto de imperativos o sea de reglas de conducta que tienen como destinatarios ante todo a los coasociados. Consecuentemente, eso significa que, si la norma penal incriminadora se traduce en una regla de conducta dirigida a seres humanos, su presupuesto lógico es que no se puede hipotetizar cualquier “regla de conducta” si su destinatario no tiene la *posibilidad* de cumplir o incumplir esa regla a través de su “voluntad”, o sea si no tiene “libre albedrío”; entendiendo esa palabra como sinónimo de “libre posibilidad de elección” (en una dimensión moral, entre el “bien” y el “mal”). “Libre albedrío”: expresión que parecía haber desaparecido en la obscuridad del pasado, pero que, gracias a Dios, hoy en día – habiéndose superado en la dogmática penal una visión fundamentalmente materialistico-naturalista – empieza a escucharse otra vez.

Ahora, si hay libre albedrío, la pregunta es: ¿cómo se ubica el libre albedrío frente a la imputabilidad?

Antes de contestar esa pregunta, quería decir algo sobre la posición “sistemática” de la imputabilidad en la teoría del delito. Eso no es un discurso abstracto, tampoco solo teórico, sino que por el contrario tiene consecuencias aplicativas de importancia capital en el Derecho Penal. ¿En qué sentido? Hoy en día la mayoría de los penalistas en España, y en Italia también, siguen, como siempre, a los alemanes en el sentido de que la imputabilidad es un elemento de la “culpabilidad”, culpabilidad que se traduce a su vez en un juicio de “reprochabilidad”, reprochabilidad por algo cometido; reprochabilidad que tiene al final un significado ético y, en cuanto tal, es perfectamente coherente con el pensamiento más profundo alemán, o sea la visión “ética” que está detrás de cualquier “reproche”, no solo moral sino también jurídico. Detengámonos ahora en el concepto de “reprochabilidad”. Una precisión metodológica preliminar: yo soy un dogmático y quiero moverme siempre al interior de la perspectiva del Código Penal, de las leyes penales. En Italia tenemos el código más viejo de Europa, o sea el Código Rocco de 1931, que – en mi opinión – es técnicamente un código casi perfecto. Gracias a Dios, tenemos todavía este código; eso nos permite decir que la idea que la “imputabilidad”, entendida como capacidad natural de entender y de querer, es algo perteneciente a un requisito denominado “culpabilidad”, contrasta de manera clarísima con el “sistema” delineado por el Código Penal italiano. Aunque según la doctrina mayoritaria italiana eso significaría poco o nada, esa construcción tiene, viceversa, consecuencias verdaderamente relevantes en el plano práctico.

En efecto, si la imputabilidad fuese realmente un elemento de la culpabilidad y la culpabilidad, como todos los que siguen la teoría “tripartita” afirman, fuese a su vez un elemento del “delito”, eso significaría que los “no imputables” no podrían cometer “delito”. A esa conclusión debería llegarse siguiendo la concepción tanto de la (antigua) culpabilidad “psicológica” como de la (más moderna) culpabilidad “normativa”. Así, si la imputabilidad fuese simplemente el “presupuesto” del dolo y de la culpa, entendidos como las dos “formas” de relación psíquica con el hecho en las cuales puede manifestarse – y agotarse – la “culpabilidad” (en el sentido de “culpabilidad psicológica”), los “inimputables”, no pudiendo – lógicamente – actuar con dolo o culpa, no podrían consecuentemente cometer hechos normativamente calificables como “delitos”. A la misma conclusión deberían también llegar los autores tripartitos italianos que, aunque siguen la concepción “normativa” de la culpabilidad, ubican todavía sistemáticamente el dolo y la culpa *al interior* del elemento de la culpabilidad (junto con la imputabilidad, la cognoscibilidad de la prohibición y la exigibilidad de un comportamiento diferente), culpabilidad considerada siempre como tercer elemento del delito. Del mismo modo, aunque dolo y culpa – como afirma hoy en día la mayoría de la doctrina tripartita post-welzeliana de molde alemán – están colocados sistemáticamente en la “tipicidad”, como elementos “subjetivos” de ella (hablándose, de manera dogmática y normativamente en absoluto incorrecta, también de dolo o culpa “naturales”), los “inimputables” podrían cometer un “ilícito” pero no un “delito”, faltando de todas formas la “culpabilidad” (aunque entendida en el sentido “normativo”, y siempre colocada sistemáticamente como segundo elemento del “delito”). Es claro a su vez que la distinción conceptual entre “ilícito” y “delito” fue creada precisamente para permitir la aplicación de las medidas de seguridad a los inimputables peligrosos autores de hechos ilícitos, pero sin abandonar la concepción, toda ideológica, del “delito” como “ilícito reprochable” y por tanto “punible”. O sea, un verdadero “prejuicio”.

Así las cosas, al final, ¿cuál es la posición sistemática correcta de la imputabilidad en la teoría del delito? En mi opinión, no debería dudarse que cualquier hecho “ilícito”, siendo desarrollo de un comportamiento humano de acción u omisión, se compone de un elemento objetivo, de un elemento subjetivo y de la falta de causales de justificación. Solo en ese contexto, se separan hoy en día los “tripartitos”, que colocan los elementos del ilícito entre las tres categorías de la “tipicidad objetiva”, de la “tipicidad subjetiva” y de la “antijuridicidad”, de los “bipartitos”, quienes colocan todos los elementos del ilícito en las dos categorías – siempre a entender

normativamente – del “elemento objetivo” y del “elemento subjetivo”, considerando las causales de justificación, como es notorio, “elementos *negativos* del tipo”. Eso significa entonces que la distinción entre “ilícito” y “delito” es simplemente fruto de un prejuicio o, mejor dicho, una inútil superfetación conceptual que constituye solo un extremo homenaje a una larga tradición dogmática que no se quiere abandonar. Convencionalmente, si nos gusta, esa distinción de palabras puede mantenerse, sin embargo, teniendo bien claro que, en una correcta perspectiva dogmática, la “culpabilidad” – entendida como calificación normativa de un hecho típico delictivo *reprochable* a su autor – está *fuera* del “ilícito penal”, o sea del “hecho” – objetivo y subjetivo – constitutivo del “ilícito penal”.

En definitiva, si siempre la doctrina ha calificado como “delito” el “ilícito penal”, no se entiende qué ventaja tendría cambiar la tradición también lingüística. Eso significa entonces que “ilícito penal” y “delito” son la misma cosa; y eso es muy claro en el código penal italiano, según el cual tanto los “imputables” como los “inimputables” pueden cometer el mismo modelo de “delito” (“ilícito penal”) en el sentido “normativo”, comprensivo de los mismos elementos constitutivos.

Las consecuencias dogmáticas y prácticas de lo dicho son muy relevantes. Aunque en ese contexto no puedo detenerme en detalles, es suficiente subrayar como, en mi opinión, de los elementos indicados como constitutivos del concepto de “culpabilidad normativa” (depurada del dolo y de la culpa), la “exigibilidad” de una conducta diferente de la cometida – al menos según el código penal italiano – no tiene algún espacio autónomo al interior de una perspectiva de “reprochabilidad personal” (fuera de los elementos del delito), así como, a su vez, el “error de prohibición inevitable” es relevante también a favor de los inimputables para excluir la aplicabilidad de medidas de seguridad. *Rebus sic stantibus*, eso significa entonces que la “culpabilidad” acabaría *identificándose* con la sola “imputabilidad”, o sea con la capacidad del autor del hecho de motivarse según los imperativos penales si quiere evitar la pena consecuente a la comisión del delito. Pero, si – según un canon metodológico fundamental de la investigación científica – *entia non sunt multiplicanda sine necessitate*, el concepto de “culpabilidad” debería ser definitivamente abandonado en el estudio dogmático de la teoría del delito¹. Consecuentemente, “reprochabilidad” debería ser solo un juicio sintético a cargo del autor del delito que simplemente cumple con todos los requisitos requeridos por la norma incriminadora para ser

¹ Más en detalle, véase mi último trabajo “*Il reato e le sue conseguenze. Punibilità, pena, punizione in un sistema criminale integrale e integrato*”, Roma TrE-Press, Roma, 2022, especialmente el capítulo primero; monografía en *e-book* libremente descargable.

“penado”; en particular, por haber cometido un “delito” siendo “imputable” en el momento de la comisión de su conducta constitutiva. Probablemente, muchas confusiones que todavía ocurren serían evitadas si se considerase, correctamente, que la comisión de un delito no es simplemente el presupuesto de la aplicación de la pena (lo que explica quizá el porqué de la definición del “delito” como “ilícito culpable” y por tanto “punible”), sino también *elemento central y fundamental de otras y diferentes hipótesis de “responsabilidad jurídica”*, como la responsabilidad civil emanada del delito, la responsabilidad por la aplicación de medidas de seguridad y, actualmente en Italia, también la responsabilidad administrativa de las personas jurídicas por la comisión de delitos realizados por la cumbre de la organización en beneficio de la misma persona jurídica².

Todo eso significa entonces, lógicamente antes que dogmática y sistemáticamente, que la “imputabilidad” es un elemento que está *fuera* del “delito”. Fuera del delito significa en primer lugar, como he dicho, que todas las personas tanto “imputables” como “no imputables” pueden cometer un “hecho delictivo”, comprensivo de todos sus elementos constitutivos. Pero también eso significa que la “imputabilidad” es simplemente una característica o – mejor dicho – una “calidad personal” del autor del hecho delictivo que permite al mismo autor ser penado: dogmáticamente es solo “*capacidad de pena*”, nada más y nada menos. Más específicamente, el autor de un “delito” puede ser destinatario de una “pena” o de una “medida de seguridad” según que tenga la “calidad personal” de sujeto “imputable” o “peligroso”.

Ahora, si la imputabilidad es simplemente “capacidad de pena” frente a un sujeto que cometió un delito con todos sus requisitos constitutivos, eso significa simplemente que la imputabilidad está ligada estrictamente a la pena, en particular a la “función” que la pena desarrolla en el sistema penal, o sea al concepto de “prevención” del delito. Concepto clásico de prevención, de entender tradicionalmente como prevención denominada “negativa”, en el sentido que la amenaza de la pena sirve para estimular un sujeto a tener el comportamiento debido. ¿Y cuál es ese sujeto? Solo el sujeto “imputable”. Adicionalmente, como ya he dicho, si la imputabilidad está fuera del delito, fuera del delito está también la posibilidad de conocer el desvalor penal del hecho: en efecto, la denominada *culpa iuris* (como la culpa por el hecho) no se refiere a un individuo particular, sino que puede ser comprobada solo respecto a un “modelo de agente” por definición “impersonal”, de modo que puede referirse indiferentemente a un sujeto imputable o no imputable. A su vez, la exigibilidad de un comportamiento diferente a aquel realizado no

² M. TRAPANI, *Il reato e le sue conseguenze*, cit., *passim*.

tiene ninguna relevancia normativa autónoma, al menos según el derecho positivo italiano, puesto que su razón de ser, y por lo tanto el origen del mismo concepto de “culpabilidad normativa” como “reprochabilidad”, es simplemente permitir a instancias de “justicia” penetrar en el sistema penal y superar lo que habría debido provenir de una aplicación estricta de las normas penales incriminadoras. Pero el problema de la “justicia” es otra cosa. En efecto, “derecho” no es sinónimo de “justicia”: el juez debe aplicar el derecho positivo, no hacer justicia. “Derecho” y “moral” – y la justicia es notoriamente el parámetro de evaluación ética del derecho positivo – tienen que estar rígidamente separados en un derecho penal liberal y garantista. Y si derecho no es justicia, permítanme esta pequeña digresión: mi absoluto rechazo, desde una perspectiva político-criminal “garantista”, del denominado “derecho internacional penal” como instrumento fundado sobre instancias de “verdad” y “justicia” “absolutas” para resolver cuestiones que son, viceversa, exclusivamente políticas en contra de sujetos ya política y militarmente “vencidos”, lo que es algo jurídicamente “sin sentido”.

2. El “contenido” de la imputabilidad

Si la “imputabilidad” es *formalmente* la “capacidad de pena”, es decir la capacidad del autor de un delito de ser penado, la pregunta a la cual debemos ahora contestar es: ¿qué es *sustancialmente* la imputabilidad? Dicho de otra manera: ¿cuál es su contenido específico? En todos los ordenamientos jurídicos, se define la “imputabilidad” – por parte de la doctrina o directamente de la ley penal – como “capacidad de entender y de querer”. Más precisamente, la capacidad de entender y querer es el substrato “naturalístico” que el ordenamiento califica “normativamente” como imputabilidad. Aquí empiezan los problemas. En primer lugar, contrariamente a lo que algunos todavía afirman, no es verdad que solo quien tiene *en abstracto* la capacidad natural de entender y de querer en un determinado momento puede *en concreto* efectivamente representarse y querer algo, lo que significaría *normativamente* actuar con “dolo” (o con “culpa”). Incluso olvidando que, en una correcta perspectiva dogmática y sistemática, como hemos dicho, la imputabilidad no es presupuesto normativo del dolo o de la culpa, en realidad la misma capacidad natural de entender y de querer no tiene nada que ver con los acontecimientos subjetivos-psicológicos en sentido ontológico, o sea representación, voluntad y previsibilidad. Eso porque ya desde la perspectiva naturalística no se

entiende que una cosa son los procesos psíquicos, reales o potenciales, y otra la “*cualidad normal*” de esos procesos. En cuanto “*cualidad*” de los procesos psíquicos, reales o potenciales, eso significa que la capacidad de entender y de querer está *fuera* de esos procesos psíquicos, que a su vez constituyen objeto de su calificación normativa como dolo o culpa. Dicho de otra manera, la capacidad de entender y de querer, ya en el plano naturalístico *no es* “presupuesto lógico” – y por consiguiente necesario – de los procesos psíquicos, reales o potenciales, de representación y voluntad. También un loco, un loco total, un *serial killer* actúa con voluntad y representación no solo naturalísticamente, sino también *jurídicamente* perfectas. Perfectas ¿en qué sentido? Simplemente porque existen en la realidad. Entonces ¿qué es lo que falta en los “incapaces”? Falta simplemente la *cualidad normal* de esa representación o de esa voluntad (según los casos).

A estas alturas, el problema es establecer qué puede entenderse por “*cualidad normal*” de una representación o voluntad.

Una aclaración preliminar: yo no creo que – excepto por exigencias didácticas – pueda distinguirse entre un derecho penal “sustancial”, un derecho penal “procesal” y un derecho penitenciario o de “ejecución” penal: hay solo *una* “norma penal”, porque, a diferencia de lo que sucede en las otras ramas del derecho, la “norma penal incriminadora”, es decir sancionada con una pena criminal, *no existe fuera del proceso*. Proceso que pertenece entonces a la “fisiología” del sistema penal. Siendo así y no pudiéndose condenar a una persona si su responsabilidad penal no ha sido demostrada “más allá de cualquier duda razonable” ¿eso qué significa con referencia, en particular, a la “imputabilidad”? ¿Que en cada proceso penal el fiscal antes y el juez después, si quieren llegar a la condena legítima del autor de un delito, tienen que demostrar cada vez la existencia de su “capacidad de entender y de querer”? Absolutamente no. Ello porque, siendo el derecho, y el derecho penal en particular, una expresión de la razón práctica, como habría dicho Kant, también fiscales y jueces deben tener en cuenta la “normalidad” de las situaciones que tienen que averiguar. Así, cuando un sujeto tiene 18 años, según el código penal italiano, existe una “*presunción*” de imputabilidad. Eso porque en el común de los casos puede decirse que los que tienen 18 años han logrado la plena “madurez”. Sólo cuando hay situaciones particulares a causa de las cuales falta en los mayores de edad la plena capacidad de entender o de querer, podemos decir que no hay imputabilidad. O sea, el sistema penal y el mecanismo mismo de la interacción entre derecho y proceso impone considerar que la “normalidad” en los mayores de edad es la presencia de la capacidad de entender y de querer, presencia que por tanto se presume. Sólo cuando

hay situaciones que pueden integrar causas *taxativamente* previstas por el legislador de “no imputabilidad” tenemos que averiguarlas. Eso significa que, según el código penal italiano, lo que es necesario para la aplicación de la pena es *solo* la calificación *normativa* de “imputable” a cargo del autor del delito; calificación que muchas veces *no coincide* con la presencia de la capacidad *natural* de entender y querer. Lo que es un problema muy grave, también de legitimidad constitucional (problema que no se puede tampoco afrontar en esta sede).

En cuanto a las causas taxativamente previstas de exclusión de la imputabilidad, es necesario aquí subrayar como, desde hace muchos años, los penalistas que en Italia han estudiado específicamente la imputabilidad con referencia a la enfermedad psíquica, para llenar de contenido ese concepto, se han movido según una perspectiva metodológica en mi opinión absurda, que consiste en utilizar textos científicos de médicos, psiquiatras y otros estudiosos de materias similares. Esa manera de enfrentar un problema de naturaleza jurídica sería así mismo ridículo, si, viceversa, los que tienen exclusivamente competencias medicas intentasen utilizar nociones de medicina para llenar de contenido conceptos estrictamente jurídicos (o, incluso peor, utilizar nociones jurídicas para llenar de contenido conceptos estrictamente médicos). Igualmente sería ridículo que yo, que soy jurista y no entiendo nada (o casi) de medicina, intentase aventurarme en conceptos propios de la psiquiatría forense no obstante mi carencia de competencias médicas. Y eso por dos razones fundamentales. En primer lugar, porque los mismos psiquiatras, entre los médicos, son los que menos saben de lo que hablan, es decir del objeto de la materia específica de su competencia. En efecto, ellos mismos lo dicen, afirmando que conocen a lo sumo el 10% de como funciona el cerebro humano y de sus patologías. Y si precisamente los psiquiatras no conocen en forma exacta y en su totalidad el funcionamiento del cerebro, ¿cómo se puede hipotetizar que una rama del derecho, como el derecho penal, deba interesarse en todas las – viejas y nuevas – “patologías” psiquiátricas e incluso psicológicas y ajustarse acriticamente a ellas, cuando los mismos psiquiatras se contradicen a menudo entre ellos (como también ocurre en todas las ramas de la ciencia)? Además, como he dicho antes, yo no soy políticamente correcto: en mi opinión, entre otras cosas, la misma psicología no es una ciencia en sentido propio. Es simplemente (casi) brujería. Más o menos ciencia – en el sentido de “ciencia empírica” – puede ser la psiquiatría, no la psicología. La psicología parte de presupuestos no demostrados, o sea de aserciones apodícticas, de donde extrae lógicamente toda una serie de consideraciones y conclusiones. Tanto es así, que afortunadamente en el código de procedimiento penal italiano

se admite exclusivamente la prueba pericial, no la prueba psicológica, la cual al interior del proceso penal está prohibida (art. 220 inciso 2 c.p.p.); evidentemente porque el mismo legislador es consciente que la psicología no es una verdadera ciencia (empírica). También es evidente el peligro que está detrás de todos estos intentos de ampliar el concepto de “enfermedad mental” para llegar, al final, a la absolución del imputado: fundamentalmente, acoger la idea de que no solo las “viejas” y verdaderas enfermedades mentales puedan excluir la imputabilidad, sino también los simples “disturbios nerviosos”, no mejor identificados. Eso significa destruir la idea misma del derecho penal fundado sobre una “responsabilidad personal” que puede ser excluida solo por la presencia de verdaderas “causas patológicas” de enfermedad mental ya unánimemente reconocidas por la ciencia. De hecho, como es notorio y Popper lo subraya de manera muy clara, la ciencia procede mediante intentos continuos y metodológicamente a través de la “falsabilidad”. Si es así, y yo creo que es así, la conclusión debería ser que el ordenamiento penal presupone que normalmente, sobre la base del *id quod plerumque accidit*, todas las personas mayores de edad son normativamente “imputables” o sea son (potenciales) destinatarios de penas, excepto en algunos casos, que constituyen *excepciones* a la regla y, por consiguiente, tienen que ser interpretados de manera restrictiva. Todo lo dicho explica porque hoy en día es tan difícil hablar del significado *jurídico* de simples “disturbios nerviosos” que sin embargo no llegan a una, reconocida, enfermedad mental, si no – por supuesto – psíquica.

La segunda razón por la cual el concepto jurídico de “imputabilidad”, en particular con referencia a su causa de exclusión representada por la enfermedad mental, no puede deducirse *tout court* de la ciencia médica, está impuesta por el mismo sistema penal correctamente entendido. En efecto, los que piensan que el juez debería siempre “obedecer” en su decisión a las conclusiones del perito psiquiatra, olvidan que tarea de los psiquiatras en la prueba pericial es solo entregar un diagnóstico y/o pronóstico sobre la “salud mental” del imputado; siendo, por el contrario, tarea exclusiva del juez, en calidad de *peritus peritorum*, es decir sin ser vinculado a la evaluación médica, valorar – conforme a las exigencias propias del derecho, sin olvidar lo que es ya unánimemente reconocido por la ciencia médica como enfermedad psíquica – el significado de la prueba pericial para establecer si y eventualmente en qué límites el autor del delito puede ser declarado imputable o no imputable.

3. *Imputabilidad e inteligencia artificial*

Llegamos ahora a unas consideraciones sobre el problema de la “inteligencia artificial”, en particular con relación al recurso a los algoritmos en la decisión del juez. Primera precisión, siempre no políticamente correcta, es que los economistas que usan los algoritmos no hacen verdadera ciencia (al menos en el sentido común de “ciencia exacta”, empíricamente verificable), porque todos nosotros, por experiencia personal, sabemos perfectamente que todo lo que dicen los economistas puntualmente se ve desmentido, puesto que la realidad se va normalmente en dirección opuesta. Eso significa que los algoritmos son meros cálculos de probabilidad, y eso sí sería demasiado peligroso en la perspectiva penal. ¿Porqué un fallo penal basado sobre un mero cálculo de probabilidad, en particular con referencia a la “personalidad” de un ser humano, sería tan peligroso para las “garantías” del imputado? Recuerdo que escuché hace algunos años que alguien puso un ejemplo en una perspectiva “neopositivista”, en el sentido de un “moderno” regreso a las ideas de la vieja Escuela Positiva (no del positivismo dogmático-normativo, entendido como riguroso anclaje al derecho positivo), derivado de un estudio norteamericano. En ese estudio, se descubría que la mayoría de los violadores tienen un particular cromosoma, por lo que, según un simple cálculo de probabilidad, se podría también afirmar que es altamente previsible que quienes poseen ese cromosoma serán en su vida violadores. Aunque admitiendo la “cientificidad” de esa estadística, ¿qué consecuencia se deriva de esa conclusión? ¿Que, siendo la persona que posee ese cromosoma “peligroso para la sociedad” según un “cálculo de probabilidad”, el Estado, en función del nuevo mito actualmente de gran “moda” (también entre los juristas) de la “prevención” de cualquier peligro a toda costa, puede – si no adicionalmente tiene el deber de – matarlo, incluso ya en el vientre de su madre, considerando, en una “comparación de intereses” (criterio tan amado por la “*politically correctness*”) por un lado la “gravedad del peligro” de violaciones y por otra parte que, según el *main stream*, el aborto ya no es normativamente homicidio? Eso tendría un nombre que obviamente todos los demócratas – no los verdaderos democráticos: estoy hablando de “demócratas” entre comillas, que son todo menos democráticos – al menos aparentemente no quieren ni siquiera escuchar: “eugenesia”. Algo propio de un régimen político que se llamaba nazismo. El doctor Mengele, quien tenía totalmente a su disposición miles de personas como material humano para sus experimentos, permitió a la biología y en general a la medicina

progresar al menos treinta años o más en los conocimientos médico-científicos. Desafortunadamente para los “progresistas”, el problema es que no se puede admitir – en primer lugar, “éticamente” – un “progreso” a toda costa, así los resultados útiles para la sociedad que podrían conseguirse fuesen enormes.

Volviendo al cálculo de probabilidad a través de la “inteligencia artificial”, digo inmediatamente que, en mi opinión, el fallo condenatorio del juez por el delito cometido y la consiguiente medición de la pena merecida no puede basarse en algoritmos elaborados por una máquina: solo un juez-ser humano puede *razonablemente* evaluar en su conjunto las pruebas para su decisión, especialmente las que se forman directamente frente a él dentro del contradictorio entre las partes procesales.

Pero lo mismo debe decirse a propósito del cálculo de probabilidad que el juez debe hacer para la afirmación de la “probabilidad” de cometer delitos por parte de una persona, es decir, de su “peligrosidad”. En ese caso, en el cual se intenta averiguar la “probabilidad”, entendida como relevante posibilidad, de algo penalmente negativo que ocurrirá en el futuro (o sea la comisión de delitos), cualquiera sea el sujeto, juez-persona física o máquina, que hace el cálculo de probabilidades, graves problemas se presentan respecto a las “garantías” de la persona imputada. Eso si se recurre tanto a los algoritmos de una inteligencia artificial, como a un juez-persona física, puesto que este último operaría como un brujo frente a una bola de vidrio: de hecho, un juicio de “probabilidad” de futura comisión de delitos no se basa sobre la averiguación de “hechos”, que se fundan sobre “pruebas”, sino solo sobre “prognosis” de comportamientos futuros. Pero es evidente que, razonando así, la probabilidad de verificación de algo en el futuro significa nada más y nada menos que la destrucción del principio fundamental del derecho penal, o sea que la “responsabilidad penal” es “personal”: en el sentido que una persona puede ser tenida como responsable solo por algo que haya sido cometido por él. En realidad, este principio fundamental – correctamente entendido – nos impide aplicar cualquier sanción de naturaleza “penal” a una persona *antes* de la comisión, por su parte, de algo prohibido por la ley. No podemos hacer nada porque de otra forma eso tendría como consecuencia una responsabilidad no personal, sino basada en un mero cálculo de probabilidad, por parte de un algoritmo o de un pronóstico de un juez de ninguna manera “verificable” y por tanto absolutamente arbitrario. Recurso a algoritmos, en particular, que es una manifestación del denominado “transhumanismo” (es decir, al final, de algo “no humano”), última expresión de la ideología de la *new age*.

Sin embargo, quiero hacer algunas precisiones más sobre el concepto

normativo de “peligrosidad”. Todos los penalistas cultos saben que el código penal italiano de 1931 nació de un compromiso entre visiones del derecho penal totalmente diferentes: por un lado, la Escuela Clásica, basada sobre la idea tradicional del derecho penal del “hecho”, imputable objetivamente y subjetivamente a su autor, sancionado con una pena cual retribución por el delito cometido; por otro lado, la Escuela Positiva, que diversamente estaba basada sobre un derecho penal del “autor”, mirando – más que al delito cometido en el pasado – al futuro comportamiento de su autor, o sea a su “peligrosidad social”, afrontable con medidas de seguridad, es decir de defensa preventiva de la sociedad. Así que en nuestro código penal tenemos también, todavía, figuras de “peligrosidad calificada”, como la habitualidad en el delito, la profesionalidad en el mismo, y la reincidencia (en sus tres formas: simple, agravada y reiterada). Siempre bajo el principio de “personalidad” de la responsabilidad penal, creo que un sujeto tiene que ser tenido responsable, en sentido estricto y propio, solo por sus acciones “libres”, es decir libre y autónomamente cometidas, así que un loco no puede ser responsable en el sentido de “punible”, porque ante todo no hay “libre albedrío” (o sea capacidad de entender y de querer). Teniendo en cuenta que el legislador de los años ’30 estaba compuesto por juristas de primera categoría, el código penal establece, precisamente en función de la “personalidad” de la responsabilidad penal, que no puede ser tenido en el mismo plano de evaluación normativa un sujeto que por primera vez, incluso puede ser la única vez en su vida, comete un delito, así sea gravísimo, y otro sujeto que, por el contrario, tiene una historia criminal enorme. Claro que los dos delincuentes, aunque fuesen autores del mismo tipo de delito, no pueden ser penados en concreto de la misma manera. Pero siempre el código de 1931, como consecuencia del compromiso arriba mencionado, prevé también un articulado sistema de peligrosidad social y de medidas de seguridad, que se pone de lado del sistema de las penas. Se trata de la denominada “doble vía”, según la cual pueden ser aplicadas medidas de seguridad a sujetos peligrosos quienes pueden ser no solo inimputables, sino también imputables.

En el sistema originario del código penal, en la mayoría de los casos la “peligrosidad social” estaba *presumida* por la ley, porque era directamente la ley penal la que establecía los presupuestos para la aplicación de una medida de seguridad. Presupuestos que consistían en hechos procesalmente verificables. En 1986, después de una reforma de la peligrosidad social y de las medidas de seguridad en Italia, algunos autores pensaron que todas esas hipótesis de peligrosidad calificada y de peligrosidad presunta, que estaban descritas por la ley, habrían debido entenderse como abrogadas,

porque el juez, como consecuencia de la reforma, tiene ahora el deber de averiguar en cada caso concreto la efectiva peligrosidad del autor del delito. En mi opinión, la disciplina normativa vigente no es así. Todo lo contrario, es todavía correcto considerar que de todas maneras es el legislador quien debe determinar los presupuestos que, conforme a su voluntad política, permiten establecer quienes son los sujetos peligrosos, garantizando así además la “certeza” del derecho también con relación a las medidas de seguridad. En el nuevo sistema, el juez puede intervenir *excepcionalmente* solo para establecer que, no obstante la integración de los presupuestos previstos por la ley, en concreto el sujeto *no es* peligroso. Pero desafortunadamente la idea de la peligrosidad, a declarar por el juez, está por todos lados en el actual sistema de derecho penal considerado en su integralidad, empezando por el sistema procesal: piénsese en la posibilidad de aplicación de una medida de aseguramiento, que puede ser incluso carcelaria, en el curso del proceso en contra de un sujeto calificado como peligroso, o sea cuando es probable, según el juez, que pueda cometer en el futuro *otros* delitos graves o de la misma naturaleza de aquel cometido. Pero, si el proceso se hace para decidir si hay o no comisión de un delito y consiguiente responsabilidad penal, ¿cómo puede tenerse el sujeto en la cárcel simplemente porque se presume que cometerá *otros* delitos si todavía no se sabe si él efectivamente cometió aquel por el cual se está procediendo?

En conclusión, el problema de la prevención del delito tiene que ser rigurosamente ligado, en mi opinión, a la idea de la pena como retribución, entendida como proporción frente al delito cometido. Gravísimo sería, viceversa, la idea de un derecho penal construido en perspectiva de una prevención de delitos que no se sabe si fueron efectivamente cometidos antes de la aplicación de cualquiera medida aflictiva. Ello simplemente porque antes, como ya dijimos, toda decisión “preventiva” estaba sometida a la sola “razón” del juez, por si misma logicamente incontrolable; mientras hoy en día la aplicación de medidas preventivas podría ser remitida a la utilización de algoritmos, u otros mecanismos del mismo sentido. Y eso en mi opinión es muy peligroso.

4. Bibliografía

TRAPANI M., *Il reato e le sue conseguenze. Punibilità, pena, punizione in un sistema criminale integrale e integrato*, Roma TrE-Press, Roma, 2022.

CAPÍTULO XI

*Neurociencias y culpabilidad: recapitulando**

María Isabel González Tapia

SUMARIO: 1. Introducción – 1.1. Presupuestos dogmáticos y epistémicos – 1.2. Neuroderecho y Derecho Penal – 2. Culpabilidad y libre albedrío – 2.1. «El libre albedrío es la biología que no se ha descubierto todavía» – 2.2. Para el Derecho Penal, ‘la neurociencia lo cambia todo y no cambia nada’ – 3. Imputabilidad, capacidad de control e impulsividad – 3.1. Genética y agresividad impulsivo-reactiva: la interacción entre el gen MAOA-L y el maltrato severo en la infancia – 3.2. Biomarcadores del delincuente violento – 4. Psicopatía: ¿debería integrarse la falta de empatía emocional dentro del concepto de imputabilidad? – 5. Bibliografía.

1. *Introducción*

1.1. *Presupuestos dogmáticos y epistémicos*

Estas páginas tienen como objetivo ser una recapitulación de las principales aportaciones de la neurociencia y de la genética del comportamiento, con incidencia en la imputabilidad, en los últimos años. Para abordar esta cuestión, que no es otra que analizar la incidencia que estos saberes científicos pudieran tener sobre la configuración dogmática del Derecho Penal en cualquiera de sus elementos, se requiere un posicionamiento previo acerca de diversos aspectos que orientan necesariamente el enfoque.

En primer lugar, sobre el recurrente y (probablemente) irresoluble problema de lo que Bernardo FEIJOO llama las ‘tormentosas relaciones’ entre la Ciencia empírica y el Derecho Penal¹. Requiere posicionarse acerca

* Este capítulo se realiza en el marco del Proyecto de Investigación Derecho Penal y Comportamiento Humano, ref. RTI2018-097838-B-I00, del Ministerio de Ciencia, Innovación y Universidades, con IP Prof. Dr. E. Demetrio Crespo.

¹ B. FEIJOO SÁNCHEZ, *Derecho Penal y Neurociencias. ¿Una relación tormentosa?*, en «InDret», n. 2, 2011, <<https://raco.cat>>; B. FEIJOO SÁNCHEZ (ed.), *Derecho penal de la culpabilidad y neurociencias*, Civitas, Madrid, 2012, pp. 71 ss.

de sí, y en qué medida, el conocimiento científico puede ser útil y debe informar al Derecho; en nuestro caso, al Derecho Penal. Y para poder contestar si la neurociencia puede aportar al Derecho nuevos enfoques o puede incluso inspirar reformas jurídicas, es preciso conceder que el conocimiento científico, aunque no determine la solución normativa, sí puede, al menos, informarla con sentido, en aquellos campos en los que, como todo lo que tiene que ver con el comportamiento humano, tienen una zona de clara confluencia.

A su vez, es preciso también 'situarse', dimensionalmente, en la dicotomía dialéctica entre el normativismo funcionalista y la política-criminal con respecto a la Dogmática, decidiendo si el Derecho es un saber puramente normativo, que se retroalimenta de valores y directrices que proporciona el contexto al que sirve funcionalmente y que genera el propio marco normativo; o se trata de un Derecho que debe mirar también hacia fuera, con un enfoque integrador y crítico, hacia lo que muestran otros saberes, hacia el substrato empírico de lo regulado, sus causas y sus contextos y hacia las consecuencias o impacto de su aplicación práctica.

Para responder a estos dos primeros posicionamientos previos, se parte aquí de la idea de que, aún con múltiples carencias y ajenos al modelo de las ciencias empíricas, la Dogmática tiene el estatus de ciencia y que por ello debe adoptar una metodología de tal carácter. Como afirmara Enrique GIMBERNAT: «*La ciencia del Derecho lo es, porque – y en tanto en cuanto – los conocimientos que la integran no son subjetivos, arbitrarios, adquiridos en base a prejuicios, sino objetivos, fundamentados, defendibles...*»². También participo de la idea de que el Derecho Penal, como ordenamiento dirigido a salvaguardar las bases fundamentales de la convivencia social en un espacio y en un tiempo determinados, debe ser un *Derecho Penal 'orientado a las consecuencias'*, orientado a la Política Criminal y limitado empíricamente, dentro de un *normativismo moderado* como el propuesto por Claus ROXIN³. Ello significa, ni más ni menos, que la Política Criminal debe orientar las distintas categorías dogmáticas y debe informar el contenido del Derecho Penal; y que ello debe hacerse teniendo en cuenta también el conocimiento que proporcionan otros saberes científicos y que debe evaluarse el impacto real del Derecho Penal para saber si está correctamente orientado a la

² E. GIMBERNAT ORDEIG, *Concepto y método de la ciencia del derecho penal*, Ed. Tecnos, Madrid, 1999, p. 110.

³ C. ROXIN, *Derecho penal: Parte General. Tomo I*, trad. D.M. Luzón Peña, M. Díaz y García Conlledo, J. de Vicente Remesal, vol. I, Thomson Reuters-Civitas, Madrid, 2008; C. ROXIN, *Política criminal y sistema del Derecho penal*, traducción e introducción de F. Muñoz Conde, Bosch, Barcelona, 1972.

consecución de sus finalidades. Obviamente, ni que decir tiene, este planteamiento dogmático moderado se limita también por el garantismo, en este caso de base constitucionalista, que co-fundamenta también Derecho Penal en la persecución de sus fines preventivos⁴.

El tercer posicionamiento que se requiere a la hora de abordar una temática como la propuesta tiene que ver con el papel que se otorga dentro del sistema dogmático a la prevención especial y, sobre todo, en cómo se la entiende, a la hora de explicar el comportamiento humano, tanto en la concreción de la responsabilidad penal, como en el enfoque de las consecuencias jurídicas del delito. Considerar la neurociencia, abiertamente, implica aceptar no 'cerrar los ojos' a que el comportamiento humano está co-determinado por factores neurobiológicos, que en la toma de decisiones influyen factores ambientales pero también biológicos y que ambos deben ser tomados en consideración en todas las categorías dogmáticas orientadas teleológicamente a la prevención (general y especial). Ello conduce a admitir, en definitiva, ese doble contenido a la prevención especial en todos aquellos contextos en los que tiene relevancia dentro del sistema jurídico-penal, desde la prevención, a la propia ejecución de las penas y demás consecuencias jurídicas de la infracción penal. Se participa, pues, del modelo sociológico propuesto por el gran maestro alemán Frank VON LISZT, trasladado a la sensibilidad y a los conocimientos actuales, conforme

⁴ El punto de partida es preguntarse cuál ha de ser la función y los límites que deben asignarse al Derecho Penal español, en un Estado social y democrático de Derecho, que, como el nuestro, propugna en su primer precepto, que los valores superiores son la libertad, la justicia, la igualdad y el pluralismo político. Y de dicho modelo de convivencia se derivarían dos presupuestos centrales. De la acepción social, se deriva la función preventiva y la necesidad de intervención del Estado en los términos previstos en el art. 9.2 CE e implica, como ha señalado QUINTERO OLIVARES, la exigencia de tutelar derechos e intereses personales, pero también sociales, de defensa de la sociedad, presente y futura, frente al delito y su autor. También se derivaría la exigencia de solidaridad con las víctimas, que debe presidir su estatuto jurídico; pero también la obligación de centrarse en lo posible en el delincuente, en las carencias socioeconómicas y biológicas que pudieran ser factores de riesgo para comisión delictiva, para tratar recuperar su capacidad para participar en el programa constitucional. Ahora bien, la defensa de la sociedad frente al delito no puede perseguirse a toda costa. Como compañero inseparable del mismo, que debiera prevalecer incluso frente a él en caso de conflicto, ha de situarse el garantismo, la salvaguarda de los derechos fundamentales del ciudadano y las garantías básicas propias de un Estado de Derecho. Esta idea se deduce claramente de la acepción democrática del Estado y, sobre todo, de un Estado de derecho, que imponen la vigencia del principio de legalidad y el principio rector de que la persona, y no la sociedad como ente global, es el eje central del sistema, de acuerdo con lo expresado en el art. 10 CE. En cuanto a la cita, vid. G. QUINTERO OLIVARES, *Aspectos generales*, en *Derecho Penal Constitucional*, dirigido por G. Quintero Olivares, coordinado por J. Jaria i Manzano, Tirant, Valencia, 2015, p. 27.

al que se reconocería el potencial que puede y debe desplegar también la prevención-especial en el sistema penal y en la que no deben desconocerse las raíces biológicas que también subyacen al comportamiento humano⁵.

De hecho, a mi juicio, como defendiera ya en otro trabajo, este planteamiento (aunque objetivizado y ‘vestido con ropajes de prevención general’) puede detectarse en el Derecho Penal actual, en una tendencia clara hacia un *Derecho Penal más individualizado*, hacia una reconstrucción del Derecho Penal de la peligrosidad⁶. El Derecho Penal moderno, mirado desde una perspectiva material (sin ‘fraudes de etiquetas’), se estructura a través de un sistema ecléctico, una nueva versión de la *vía única* en el que las penas y las medidas de seguridad (junto a las demás consecuencias jurídicas) actúan conjuntamente y en el que la peligrosidad está invadiendo cada vez más el tipo de injusto con *tipos de autor* objetivizados, con penas que se incrementan sustancialmente por esta razón subyacente y cuya ejecución se complementa con auténticas medidas de seguridad, incluyo más allá de la misma, como ocurre con la libertad vigilada o las penas accesorias.

Sin embargo, y creo que es una de las principales críticas que pueden hacerse al Derecho Penal actual, el incremento del rigor punitivo para contrarrestar la peligrosidad criminal se ha fundamentado más en intuiciones, en concepciones ideológicas previas o en legitimaciones populistas, que a partir de estudios científicos que demuestren la mayor peligrosidad de estos sujetos o una mayor necesidad de pena en la tutela de un determinado bien jurídico. Y si ello se hiciera en términos más científicos, si la política criminal se definiera informadamente, habría de contarse también con la Neurociencia y con la Criminología Biosocial, en la medida en que informan de los factores de riesgo biológicos que favorecen la comisión delictiva y pueden ser elementos relevantes para informar con sentido la posible conveniencia de tipos agravados o atenuados fundamentados en la peligrosidad o determinadas figuras delictivas. Y lo mismo cabe decir de la intensificación del recurso al diagnóstico de la peligrosidad en la concreción de la responsabilidad y su generalización en el sistema de las consecuencias jurídicas del delito, puesto que esta dirección político-criminal, en cualquier caso, tendrá mayor eficacia y legitimidad solo cuando (y solo si) la predicción de la peligrosidad

⁵ F. VON LISZT, *La idea de fin en el derecho penal*, Ediciones Olejnik, Santiago de Chile, 2020, *passim*.

⁶ M.I. GONZÁLEZ-TAPIA, *Derecho Penal de la peligrosidad y Neuropredicción: hacia un Derecho Penal Individualizado*, en *Estudios jurídico penales y criminológicos en homenaje al Prof. Dr. Dr. h. c. mult. Lorenzo Morillas Cueva*, coordinado por I.F. Benítez Ortúzar, D.L. Morillas Fernández, J.E. Sáinz-Cantero Caparrós, J.M. Suárez López, vol. 1, Dykinson, Madrid, 2018, pp. 326-51.

mejora sustancialmente su precisión, apoyada en métodos científicos fiables y en un conocimiento criminológico que pueda servir realmente de apoyo a las decisiones de política-criminal. Y ciertamente, hoy por hoy, la fiabilidad del diagnóstico de la peligrosidad está fuertemente cuestionada a nivel científico, y viene apoyándose, según parece, en *factores dinámicos, estáticos y sociobiológicos* estructurales del sujeto, con un claro protagonismo estadístico de los factores sociobiológicos y de comisión delictiva previas en cuanto a la predicción de la reincidencia. Y en este aspecto, sin lugar a dudas, la neurociencia jugará, de la mano de la inteligencia artificial, un papel muy destacado con la denominada *neuropredicción* y con el tratamiento con neurotecnologías o inspiradas en la Neurociencia⁷.

Hechas las precisiones anteriores, en lo que se refiere expresamente a la categoría de la culpabilidad, también sería conveniente formalizar algún otro punto de partida previo.

Así, entre los penalistas parece que la culpabilidad es un elemento en crisis permanente, inmerso en una eterna discusión, respecto de la que no es posible llegar a un acuerdo. Parece que se hubiera convertido, a fuerza de cuestionarla, a fuerza de buscar elementos comunes entre tendencias... en un elemento casi decorativo, despojado (si alguna vez realmente lo tuvo) de valor operacional. Es un elemento al que la Dogmática, a fuerza de de-construirla, ha dejado en versión minimalista. Algunos dicen que su esencia se halla concentrada en el art. 5 del Código Penal, en el «no hay pena sin dolo o imprudencia». Probablemente, pueda decirse que un sujeto es culpable de la realización de un hecho antijurídico cuando, conforme al Derecho Penal (carácter normativo), pueda serle atribuida o imputada responsabilidad por su comisión y dicha relación desencadena una respuesta penal carácter preventivo-sancionadora. Culpabilidad, fuera del reproche moral, es la imputación normativa del hecho antijurídico.

Como categoría dogmática, se ha sostenido también que la culpabilidad es el fundamento y el límite de la pena. Así, sería el elemento del delito conforme al que se responden preguntas (y anuda consecuencias) sobre los destinatarios de la norma penal (¿a quién puede ‘castigarse’ legítimamente?); o sobre ¿cuánto debe ‘castigárseles’ por ello?, según la graduación de la responsabilidad y la exigibilidad... Del mismo modo, como categoría teleológicamente orientada, la culpabilidad no sólo debe estar orientada a los límites de la responsabilidad en base a la idea de proporcionalidad con el hecho realizado, sino también a la finalidad de prevención especial

⁷ GONZÁLEZ-TAPIA, *Peligrosidad de la conducta y del autor en el Derecho Penal actual: implicaciones político-criminales y dogmáticas*, en «Revista de Derecho, Empresa y Sociedad», 13, 2018, pp. 139-57.

y general, en torno a la idea de necesidad de pena. De hecho, ese sería el juego de los fines de la pena, visto en su conjunto. No comparto, como puede apreciarse, la idea purista de separación neta entre ‘culpabilidad y prevención’. Ello, en realidad, ha servido históricamente para *expulsar* formalmente del Derecho Penal de la culpabilidad a la peligrosidad y que, al mismo tiempo, ésta haya *entrado por la puerta de atrás*, camuflada en la tipicidad y gestionada *en los márgenes* del Derecho Penal por parte del Derecho preventivo-sancionador de la ejecución de las penas y de las medidas de seguridad. Se ha dejado sistemáticamente ‘inmaculado y puro’ al Derecho Penal de la culpabilidad, en teoría, pero coexistiendo con el mundo jurídico paralelo de la peligrosidad, con muchas menos garantías, como si fueran o pudieran ser compartimentos estancos... Hipocresía dogmática aún mayor, cuando las cárceles se han convertido en la actualidad, se ha dicho muchas veces, en los ‘*manicomios del siglo XXI*’⁸. En mi opinión, una culpabilidad orientada a las consecuencias implica una mirada global y totalizadora de la reacción penal (de penas, medidas y otras consecuencias...), hacia la decisión en torno al tipo de respuesta penal y al modo de articularla: un ‘Derecho Penal individualizado’, dirigido a la prevención del delito, con la salvaguarda máxima posible de las garantías derivadas de la vigencia de un Estado Democrático y de Derecho.

1.2. *Neuroderecho y Derecho Penal*

Desde otra perspectiva, la reflexión jurídica en torno a la incidencia que la neurociencia podría tener sobre la configuración dogmática del Derecho Penal nos conduce al ámbito del conocido como *Neuroderecho* (*Neurolaw*). Este nuevo campo del saber se ocupa de analizar los nuevos avances científicos en el campo de la neurociencia y de la genética del comportamiento para luego reflexionar acerca de sus implicaciones que tendrían para el Derecho y, muy en particular, para el Derecho Penal. No es que el debate de la relación entre las ciencias empíricas y el derecho sea nuevo, claramente; pero sí que ha cambiado el contexto con la irrupción de la inteligencia artificial y la ciencia empírica que trata de incidir en el Derecho: la neurociencia, que ha despegado de forma espectacular a partir del desarrollo de las técnicas de neuroimagen, en especial, de la resonancia magnética funcional.

⁸ Cfr. el informe hecho por la SECCIÓN INDICAL ESTATAL DE CCOO EN INSTITUCIONES PENITENCIARIAS, *Informe Salud Mental en prisiones*, en <<https://fsc.ccoo.es>> (u.v. 28/07/2017).

Como expresaran Goodenough y Tucker en 2010, el Neuroderecho se ocupa de tres áreas de estudio esenciales⁹. En primer lugar, está creando un contexto normativo bioético aplicable a la neurociencia, lo que incluye las reglas relativas a la investigación neurocientífica y a las posibilidades y límites en cuanto a los tratamientos o intervenciones. En segundo lugar, la neurociencia social se ocupa del conocimiento y del comportamiento del propio Derecho. Esto es, se analiza el proceso de creación, eficacia y aplicación del Derecho y el modo en que es percibido por las personas tanto a nivel individual como a nivel grupal o social. Así, por ejemplo, se estudia cómo pensamos y cómo decidimos acerca del cumplimiento de las normas y cómo se percibe la justicia del castigo en todas sus facetas, informando sobre mejores estrategias preventivo-generales. En tercer lugar, y el que más nos interesa, el Neuroderecho se ocupa de la *cognición* y del *comportamiento* relevante para el Derecho en su aplicación, tanto respecto del Derecho Penal material, como del Derecho Procesal Penal.

Dentro de este tercer pilar del Neuroderecho y con respecto a lo que podría tener mayor incidencia con relación a la imputabilidad, podrían distinguirse, a su vez, otros tres ámbitos fundamentales, que serán desarrollados en cuanto a sus aspectos fundamentales en los epígrafes siguientes. El primero de ellos tiene que ver con el proceso de toma de decisiones del ser humano en lo que se refiere a la libertad de *autodeterminación o libre albedrío* y con respecto al que en los últimos años se ha reeditado el debate acerca de su existencia. El segundo de ellos está vinculado igualmente con la cognición y con el proceso de toma de decisiones, pero, en concreto, con respecto a la toma de decisiones de naturaleza 'moral'. Al fin y al cabo, aunque se concretará posteriormente esta afirmación, la 'decisión' de cometer un delito es una decisión acerca de la infracción de una norma, de la realización de un comportamiento antijurídico, de la puesta en peligro o lesión de bienes jurídicos relevantes ajenos... y, en definitiva, una decisión de carácter y relevancia normativa y, en este sentido, también moral o ética. Como se verá a propósito de la psicopatía, la neurociencia ha favorecido un debate acerca de la *racionalidad subyacente a la imputabilidad y la necesidad o no de incorporar la capacidad emocional*, a través del nuevo paradigma epistemológico de cerebro-mente emocional. En tercer lugar, la neurociencia y la inteligencia artificial están

⁹ O.R. GOODENOUGH, M. TUCKER, *Law and Cognitive Neuroscience*, en «Annual Review of Law and Social Science», 6, n.º 1, 2010, pp. 61-92, en <<https://www.annualreviews.org>> . También, GONZÁLEZ-TAPIA, *Neuroderecho: implicaciones para el Derecho Penal*, en *Locura y Crimen: Psiquiatría y Ley*, editado por A. Medina et al., e-Book: Fundación Española de Psiquiatría y Salud Mental, 2018), pp. 24 ss.

avanzando conjuntamente (lejos aún) en la descripción de la estructura y funcionalidad del cerebro normal o medio y también de los indicadores o *biomarcadores* de las condiciones patológicas, que las identifiquen a través de perfiles genéticos, de volúmenes neurológicos aberrantes o de patrones de actividad anómalos, de su perfil psicológico y conductual vinculado... Y están avanzando también para proporcionar un *modelo neurobiológico explicativo del comportamiento* (también del criminal), que se teoriza *dinámico*, es decir, con una perspectiva temporal hacia atrás en el pasado y hacia delante en la *neuropredicción*; *dimensional*, es decir, graduable, con distintos niveles de afectación o de riesgo...; y *complejo*, integrando otros múltiples factores, con una miríada de datos de toda índole, y atendiendo factores (predisposiciones) inmediatos y remotos hasta las raíces últimas de nuestro comportamiento. Y con respecto a este ámbito, el Neuroderecho reflexiona acerca de las implicaciones que tiene para el Derecho, para el Derecho Penal, poder contar en el futuro con un modelo explicativo de las bases neurobiológicas del comportamiento antisocial.

2. Culpabilidad y libre albedrío

2.1. «El libre albedrío es la biología que no se ha descubierto todavía»

Tanto en las concepciones clásicas de la culpabilidad, construidas sobre la idea de reproche por la infracción de la norma y el merecimiento de pena, como en las concepciones que orbitan alrededor de la idea de motivación y de la necesidad de pena, el estándar de destinatario de la norma, con capacidad de culpabilidad o imputable, es la de un ser humano agente, racional, que toma decisiones y que es motivable por la norma y los valores subyacentes a ella. Así somos realmente o así nos vemos; tanto da. Es decir, en los planteamientos dogmáticos mayoritarios acerca de la culpabilidad, aunque se haya considerado indemostrable, se ha seguido manteniendo el 'libre albedrío' como presupuesto operativo mediato, que se presupone para la atribución de responsabilidad.

En los últimos años, a mi modo de ver, *la neurociencia no ha venido a demostrar el determinismo ni tampoco el libre albedrío en el sentido tradicional y categórico (sí/no) de ambos términos. Lo que sí ha venido es a 'unificar' y a 'matizar' ambos conceptos, introduciendo un enfoque dimensional de los mismos, es decir, graduable.*

No se habla de factores únicos que determinan directamente un tipo de comportamiento. Ello sólo ocurre de forma directa en rarísimas excepciones, p.e., por la afectación al control de los impulsos sexuales debido a un tumor cerebral; la ansiedad, depresión e irritabilidad propias de la enfermedad Huntington, provocada por un único gen¹⁰; la psicopatía adquirida por un daño cerebral traumático como en el celeberrimo caso de Phineas Gage¹¹. De lo que se habla es de *predisposiciones* o vulnerabilidades hacia un determinado comportamiento; (de factores protectores y) de *factores de riesgo múltiples*, que incrementan la probabilidad de que se tome una decisión en un determinado sentido y se realice un determinado tipo de comportamiento en una situación concreta¹². A partir de ahí, lo que se ha venido es a actualizar el contexto científico y a reformular los parámetros tradicionales sobre los que se ha desarrollado el debate jurídico acerca del libre albedrío, el fundamento de la responsabilidad penal y la legitimidad de la pena¹³.

Así, a través de diversos experimentos, Benjamin LIBET y colegas empezaron a plantear que ciertos procesos neuronales inconscientes serían previos y estarían en el origen de la sensación consciente de haber tomado una decisión. En 1983, si bien los trabajos se remontan a los años 60, este autor afirmaba que el cerebro ‘decide’ iniciar o, como mínimo, ‘preparar para iniciar’ (ciertas acciones) antes de que haya ninguna conciencia subjetiva registrada de que dicha decisión ha sido tomada¹⁴. A partir de ahí, aparece la

¹⁰ Vid. por ejemplo, la atenuación en un supuesto de ‘pedofilia adquirida’ en el caso del pediatra Mattiello, Tribunale di Venezia, 8 aprile 2013. También, en caso de un professor en J. BURNS, H.S. RUSSELL, *Right orbitofrontal tumor with pedophilia symptom and constructional apraxia sign*, en «Archives of neurology», vol. 60, 3, 2003, pp. 437-40. P. JENSEN et al., *Crime in Huntington's Disease: A Study of Registered Offences among Patients, Relatives, and Controls*, en «Journal of Neurology, Neurosurgery & Psychiatry», 65, n. 4, 1998, pp. 467-71.

¹¹ A. GARCÍA-MOLINA, *Phineas Gage y el enigma del córtex prefrontal*, en «Neurología», 27, n. 6, 2012, pp. 370-75., entre otros muchos.

¹² Cfr. S. ZEKI, O.R. GOODENOUGH, R.M. SAPOLSKY, *The Frontal Cortex and the Criminal Justice System*, en «Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences», 359, n. 1451, 2004, pp. 1787-96. En dicho trabajo, estos autores señalan una serie de aspectos que confrontan claramente el proceder y necesidades del Derecho y la Ciencia, como el pensamiento categórico frente al in continua o dimensional, la diferencia entre una causa y múltiples causas interactuando.

¹³ Como esta cuestión va a ser objeto de tratamiento más detenido en otras partes de este libro, me voy a centrar únicamente en un ejemplo de posición no compatibilística, es decir, negadora abiertamente del libre albedrío, para simplemente ilustrar el debate que se ha generado.

¹⁴ B. LIBET et al., *Time of conscious intention to act in relation to onset of cerebral activity*

idea de que el cerebro de forma inconsciente estaría, al menos en cierto tipo de decisiones básicas de naturaleza motora, tomando u orientando la decisión antes de que el sujeto tenga siquiera conciencia de ella y, por tanto, antes de que pueda controlarla voluntariamente. Sobre esta base, se ha reavivado la discusión de si el ser humano es libre para autodeterminarse conforme a sentido entre los neurocientíficos y los penalistas de todas las orientaciones posibles: neurodeterministas, no deterministas y posiciones intermedias o compatibilistas¹⁵. *Se ha discutido si el modelo tradicional y popular de 'agente racional' se corresponde con la verdad científica o si lo existe realmente es un modelo de 'agente mecánico' dependiente de la biología inmediata y remota del ser humano, que todavía no se sabe explicar de forma completa.*

En un reciente trabajo, Robert SAPOLSKY¹⁶ nos ilustra magistralmente acerca de esta cuestión desde una posición no compatibilista, es decir, negadora de la existencia del libre albedrío; al menos, de cómo éste se ha entendido tradicionalmente. Explica este autor que a lo largo de la historia enfermedades como la epilepsia, la esquizofrenia o la dislexia fueron malinterpretadas en el pasado y que ello causó un enorme sufrimiento personal y social a quienes las padecieron, hasta que el avance de la ciencia pudo demostrar la biología subyacente a tales trastornos mentales y dar una explicación científica de las mismas, cambiando por completo las concepciones que se tenían sobre ellas y, por consiguiente, también transformando el tratamiento jurídico que habrían de recibir. Y, de la misma forma, nos dice: «lo que hoy llamamos libre albedrío es la biología que no se ha descubierto todavía» (p. 155).

Explica este autor que la mayor parte de los comportamientos delictivos responden al esquema de alguien que tenía, en una situación dada, la alternativa entre un comportamiento correcto (*right*) y otro incorrecto o antijurídico (*wrong*) y tomó la decisión incorrecta, eligiendo el comportamiento antijurídico. Y, afirma, lo que estamos aprendiendo es que la biología tiene mucho que ver en dicha decisión. En concreto, en dicha decisión contraria a los valores normativos tiene una particular incidencia el *córtex prefrontal* del cerebro, que gobierna el control de los impulsos, la planificación a largo plazo, la postposición de la recompensa y la regulación

(*readiness-potential*): *the unconscious initiation of a freely voluntary act*, en «Brain», 106, n. 3, 1983, pp. 623-42.

¹⁵ Cfr. E. DEMETRIO CRESPO, M. MAROTO CALATAYUD (eds.), *Neurociencias y derecho penal: nuevas perspectivas en el ámbito de la culpabilidad y tratamiento jurídico-penal de la peligrosidad*, Edisofer, Madrid, 2013.

¹⁶ P.A. ALCES, R.M. SAPOLSKY, *Nohwere*, en «William & Mary Law Review 1079», 63, n. 4, 2022.

emocional. Si en una situación dada, una persona se siente tentada a hacer algo y no lo hace es, en esencia, porque en esa fracción de segundo que se tarda en tomar una decisión el córtex prefrontal está funcionando correctamente. Sin embargo, daños en el córtex prefrontal pueden producir que un individuo sea capaz conocer el sentido de lo que hace, su carácter incorrecto o antijurídico y, sin embargo, no ser orgánicamente capaz de regular apropiadamente su comportamiento respecto de esa comprensión¹⁷. Al propio tiempo, y en sentido inverso, cuando ya se ha tomado la solución incorrecta, el córtex prefrontal también trabaja para que se realice el comportamiento indebido de la forma más eficaz.

A partir de este planteamiento, cuando se pregunta por qué ha ocurrido un determinado comportamiento o por qué esa persona tomó una determinada decisión, afirma este autor que *la respuesta es inmensamente compleja, porque depende de múltiples factores que interactúan continuamente, desde los sucedidos un segundo antes de la toma de decisión, a factores que tienen su origen millones de años atrás, según nuestra configuración evolutiva*. Así, entre los factores inmediatos que tendrían incidencia relevante, podría señalarse p.e., la ingesta de alcohol o el estrés. En este último supuesto, por ejemplo, si una persona ha estado estresada durante minutos previos, horas o días antes de un determinado evento, tendrá *la amígdala activada* y ello influirá en la evaluación del riesgo percibido (que se percibe mayor) y, por tanto, en la regulación del miedo y la ansiedad que está sintiendo esa persona. Esta situación hará a esa persona más proclive a percibir eventos amenazantes donde cualquier otro vería un evento neutral; y será también más proclive a reaccionar para neutralizar la amenaza que cree que existe. El estrés libera la hormona del *cortisol*, que te hace más proclive a ver una amenaza donde no la hay y a reaccionar para defenderte, y ello podría explicar supuestos en los que policías han disparado sus armas antes de haber siquiera procesado lo que estaban viendo realmente. Cuando las hormonas del estrés son elevadas, esa persona o cualquiera de nosotros sería más proclive a decidir que el teléfono que realmente está sacando un ciudadano de su bolsillo es una pistola y que debe defenderse disparando su arma. Y todo ello es a causa de una biología sobre la que no se tiene control y sobre lo que no somos ni siquiera conscientes.

Además, habría también otros múltiples factores mediatos que afectan al córtex prefrontal, interactuando entre ellos en una complejísima ecuación y de los que no somos siquiera conscientes. Múltiples factores que interactúan

¹⁷ Tambien S. ZEKI, O.R. GOODENOUGH, R.M. SAPOLSKY, *The Frontal Cortex and the Criminal Justice System*, cit., pp. 1790 ss.

para influir en la dirección de la decisión que finalmente toma un sujeto ante un evento que requiera autocontrol y en el que la decisión conforme a derecho, la decisión correcta sea la más difícil de tomar en la situación concreta. Influyen en la capacidad de control y regulación emocional que ejerce el córtex prefrontal, por ejemplo: el desarrollo madurativo del cerebro (hasta los 25 años aproximadamente), las memorias, emociones y pensamientos asentados que el sujeto asocie positiva o negativamente a la situación concreta, la existencia de un estrés postraumático, haber sido víctima o presenciar violencia física, psíquica o sexual o abandono en la infancia, si hay problemas mentales, encarcelamientos o adicciones en la familia cercana, si ha sido crónicamente expuesto a toxinas ambientales como el plomo, si ha nacido pobre, exposición a otros factores nocivos prenatales... Incluso, esas experiencias adversas en la infancia pueden haber causado cambios epigenéticos en la regulación de los genes en su cerebro y que pueden transmitirse a las siguientes generaciones... También incide la herencia genética que se tenga, como, por ejemplo, tener el perfil del MAOA-L y que éste se combine con el maltrato o abuso infantil, pues puede derivar en hombres a una impulsividad por encima de la media; o genes vinculados con la oxitocina que se combinen con abuso en el consumo de alcohol y que pueden incrementar la agresividad reactiva; o la persona de que se trate en el caso concreto, o la música que suena en tal instante... si el autor tiene hambre, si el otro es un extranjero, si está atravesando por dificultades diversas, las fluctuaciones del día o de la estación, sus niveles de testosterona en sangre o azúcar, o de oxitocina... Múltiples factores que pueden en una situación concreta incrementar o decrecer el riesgo de que alguien dispare a otro ante un evento que se percibe como amenazante y frente al que se aprecia la necesidad de defenderse (violencia impulsivo-reactiva).

Múltiples factores, una miríada de combinaciones, diferentes incidencias o efectos estadísticos... una cuestión tan absolutamente compleja como podría ser el universo. Esta es la biología subyacente a nuestro comportamiento y que todavía no conocemos.

2.2. Para el Derecho Penal, 'la neurociencia lo cambia todo y no cambia nada'

El problema del libre albedrío, como fundamento último de la responsabilidad penal, es una cuestión que incide en la columna vertebral de nuestro sistema, en tanto que una opción incompatibilista conduciría a eliminar el elemento intencional y se dinamitaría el sistema desde dentro,

dirigiéndonos hacia un Derecho Penal funcional de la peligrosidad, cuyo fundamento y finalidad habrían de ser: neutralizar dicha peligrosidad del sujeto que delinque o que tiene una propensión a delinquir. En cualquier caso, aparte de las dificultades vinculadas a la definición concreta de ese nuevo Derecho Penal, el problema fundamental seguiría siendo el conjugar este potencial sistema de pura defensa social con los derechos fundamentales del ser humano y con el sistema de garantías materiales y procesales que conforman el modelo socio-liberal de derecho vigente.

Probablemente, en la actualidad ningún neurocientífico, ni tampoco ningún jurista, afirma que en el presente o en el futuro inmediato, la genética o la neurociencia vayan a revolucionar al panorama legal, particularmente el Derecho Penal. Incluso desde posiciones claramente anticompatibilistas, se muestran escépticos ante cambios inmediatos y mucho menos radicales. No parece posible otra posición teniendo en cuenta que, como se ha visto en el apartado anterior, las bases neurobiológicas de nuestro comportamiento es una cuestión increíblemente compleja de la que todavía falta mucho por conocer.

En su conocidísimo artículo, titulado: *For the Law, Neuroscience changes everything and nothing* de 2004¹⁸, Joshua GREENE y Jonathan COHEN negaban categóricamente también la existencia del libre albedrío y apostaban porque en el futuro la Neurociencia tendrá un efecto transformador del Derecho. Vaticinan que lo hará, no sólo minando las actuales premisas sobre las que se asienta nuestro modelo legal de persona, responsabilidad y del Derecho en general, sino de forma mucho más profunda y radical: cambiando las intuiciones morales que la gente tiene sobre la libertad de la voluntad y sobre la responsabilidad. De hecho, entienden que, en verdad, sin ese cambio radical en la percepción social del ser humano, no es posible realmente un cambio en las bases fundamentales del Derecho Penal y todavía no estamos preparados. Hoy por hoy, las personas son entendidas socialmente como *agentes*, es decir, con la autonomía y libertad intrínseca para organizar su comportamiento conscientemente conforme a sentido y en tanto ello subsista, la neurociencia no tendrá potencial alguno de cambio sustancial hacia el Derecho, puesto que se trataría de una imposición científica que carecería de la comprensión y del apoyo social necesario para transformar el Derecho Penal hacia un derecho penal puramente consecuencialista. «The legitimacy of the law itself depends on

¹⁸ J. GREENE, J. COHEN, *For the Law, Neuroscience Changes Nothing and Everything*, en «Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences», 359, n. 1451, 2004, pp. 1775-85. n.º 1451 (29 de noviembre de 2004)

its adequately reflecting the moral intuitions and commitments of society. If neuroscience can change those intuitions, then neuroscience can change the law»¹⁹. La neurociencia cognitiva, a través del descubrimiento de los mecanismos específicos responsables de nuestro comportamiento, podrá ilustrarnos y hacernos comprender que hay algo ‘sospechoso’ en nuestras concepciones ordinarias acerca de la acción humana y de la responsabilidad; y como resultado de ello, que los principios jurídicos vigentes pudieran no ser adecuados a este nuevo entendimiento²⁰. Hasta que dicho cambio en la percepción social del ser humano se consolide, afirman, las posiciones escépticas serían esencialmente correctas. Hoy por hoy, las bases legales de la responsabilidad penal no hacen suposición alguna respecto de las bases neurológicas del comportamiento criminal y ello se corresponde con la percepción social de la libertad del ser humano correspondiente a lo que llaman *folk psychology*.

Con un planteamiento cercano a los anteriores, pero más moderado, GAZZANIGA afirma que la neurociencia actualmente no defiende un determinismo radical. El cerebro es un sistema altamente paralelo y distribuido en el que se producen literalmente millones de decisiones al mismo tiempo. El ser humano, el organismo que procesa en paralelo, es visto socialmente (por los demás seres humanos) marcadamente como un ser automotivado, moralmente coherente, consciente y sujeto capaz de tomar decisiones. Los seres humanos se reconocen entre ellos la ‘assignment of agency’, que vendría a ser como la capacidad para ser reales actores de nuestro comportamiento, con capacidad para actuar de forma voluntaria, consciente. Más aún, incluso se reconoce esta ‘capacidad de actuar’ a los mercados o a las empresas...En cualquier caso, a su juicio, esta discusión no es crucial para determinar dónde ubicamos la responsabilidad. La responsabilidad no está en el cerebro, afirma, sino en el contrato social. La responsabilidad refleja la esperanza que todos tenemos en que cada persona, que responda a la idea de normalidad, seguirá ciertas reglas y será también capaz de motivarse con ellas²¹.

Frente a las anteriores posiciones, una de las voces más autorizadas entre las posiciones escépticas y que ha ejercido una gran influencia en el estado de opinión dominante ‘neuroescéptico’ ha sido Stephen MORSE.

¹⁹ GREENE, COHEN, For the Law, Neuroscience Changes Nothing and Everything, cit., p. 1778.

²⁰ GREENE, COHEN, For the Law, Neuroscience Changes Nothing and Everything, cit., p. 1775.

²¹ M.S. GAZZANIGA, *The Law and Neuroscience*, en «Neuron», 60, n. 3, 2008, pp. 412-15, para la cita, p. 413.

Para este autor, respecto de la neurociencia se habría generado algo así como un *'brain overclaim syndrome'*²². Al igual que ha ocurrido con otras ciencias antes, p.e. la genética, se ha especulado con la posibilidad de que pudieran suponer una auténtica revolución para el Derecho. De la genética se decía, una vez completada la secuenciación del genoma humano, que iba a ser un revulsivo extraordinario para el conocimiento científico. Pero no ha sido así, ni siquiera en el campo natural de la misma como es en la Medicina. Mucho menos, pues, en el campo del Derecho. Miedos tales como que la genética vendría a amenazar nuestra privacidad, acabar con nuestra autonomía o con la heterogeneidad social... la verdad es que, hoy por hoy, no se han cumplido. Y tales miedos con relación a la neurociencia tampoco van a cumplirse, afirma; o no, al menos, en un futuro a corto plazo. Al igual que otras ciencias antes, de las cuales se especulaba que vendrían a 'revolucionar' el Derecho, la Neurociencia no va a implicar una radical transformación de las estructuras jurídicas básicas. La neurociencia, a su juicio, tiene muchas cosas que decir, pero no, al menos en un futuro próximo, la clase de cosas que la gente está esperando que diga, particularmente con relación al Derecho. Por ahora, concluye, puede ofrecer sólo modestas contribuciones que pueden ser útiles para mejorar ciertos aspectos de la práctica de los tribunales, sin necesidad de modificar las bases del Derecho Penal.

Como explicaba también MORSE²³, a propósito de la genética, el Derecho es un campo normativo cuyas nociones de responsabilidad y competencia hacen referencia al comportamiento y al estado mental en el que se encuentra el agente sobre la base de una *'folk-psychology'*. De acuerdo con esta psicología popular, el concepto de persona es el de un agente racional, capaz de autodeterminarse con sentido, de acuerdo a razones y a intenciones. Y la noción de responsabilidad, todos los conceptos implicados en tal juicio de imputación, como la propia idea de Derecho, en atención a su esencial carácter normativo, parten igualmente de esa psicología. El Derecho parte de la necesidad de que exista un comportamiento humano, activo u omisivo, y, en consecuencia, que responda a un acto consciente del sujeto. Después, para dilucidar la posible atribución de responsabilidad, así como el grado de la misma, debe analizarse, en términos generales, la

²² S.J. MORSE, *Brain Overclaim Syndrome and Criminal Responsibility: A Diagnostic Note*. Retrieved from <https://repository.upenn.edu/neuroethics_pubs/28/?utm_source=repository.upenn.edu%2Fneuroethics_pubs%2F28&utm_medium=PDF&utm_campaign=PDFCoverPages>, 2006.

²³ S.J. MORSE, *Genetics and Criminal Responsibility*, en «Trends in Cognitive Sciences», University of Pennsylvania Law School, Public Law Research Paper N. 11-34, 2011.

capacidad del sujeto en el momento de actuar, sus intenciones, la situación motivacional en la que se encontraba... (*mental states*). Este punto de vista puede ser considerado por algunos científicos o filósofos como primitivo o precientífico pero lo cierto es que, en tanto que la ciencia no sea capaz de demostrarnos que es radicalmente incorrecto, el aparato explicatorio básico de la psicología popular va a seguir siendo el pilar básico del Derecho.

En segundo lugar, MORSE apunta también a otra de las dificultades que son tradicionalmente esgrimidas cuando se analizan las relaciones y las implicaciones que pueden establecerse entre las ciencias de la vida o las ciencias sociales y el Derecho. Esto es: los diferentes presupuestos objetivos y metodológicos de los que parten aquellas y el Derecho. Y si la brecha es ya muy significativa entre las ciencias sociales y el Derecho, mucho más lo es cuando se trata de la Neurociencia y el Derecho. Mientras que el criterio de responsabilidad del Derecho es, afirma, esencialmente conductual (actos y estados mentales), el criterio de la neurociencia es mecanicista (estructuras neurológicas y funciones). Desde el punto de vista neurocientífico, afirma este autor, «las neuronas y las redes neuronales no actúan intencionalmente sobre la base de razones. No tienen sentido del pasado, presente y futuro, y no tienen aspiraciones»²⁴. ¿Es posible, entonces, sortear esta dificultad? A juicio de MORSE, dejando aparte aquellos planteamientos que no ven posible ninguna clase de relación, entendiendo que la disparidad epistemológica subyacente conduce a un error categórico radical²⁵, las relaciones entre la Neurociencia y el Derecho deben establecerse de forma interna, como una crítica que debe incorporarse a los conceptos de psicología popular que subyacen en la Ley sin cambiarlos, particularmente en lo referido al libre albedrío y a la noción de responsabilidad. En una relación establecida internamente o como crítica interna al sistema, se acepta la general coherencia y legitimidad del conjunto de doctrinas, prácticas o instituciones de esta. La contribución se formaliza a partir del intento de explicar mejor o de mejorar diferentes cuestiones que puedan no estar funcionando adecuadamente. Una relación interna, en definitiva, no pretende subvertir el sistema básico establecido, sino contribuir a su mejora.

²⁴ MORSE, *Genetics and Criminal Responsibility*, cit., p. 603.

²⁵ Así, p.e. M.R. BENNETT, P.M. HACKER, *Philosophical foundations of neuroscience*, Second Edition, Hoboken, New Jersey, John Wiley & Sons, 2022, *passim*.

3. *Imputabilidad, capacidad de control e impulsividad*

El estudio de la violencia reactiva o impulsividad agresiva es otro de los aspectos centrales respecto de los que la neurociencia está aportando un conocimiento muy relevante para el Derecho Penal.

A primera vista, sin embargo, pudiera pensarse que la Neurociencia no está aportando realmente un conocimiento novedoso²⁶. De sobra es conocido que la impulsividad o falta de control es un factor criminógeno de primer orden y que los déficits en la capacidad de control son una clara vulnerabilidad del sujeto frente a la comisión delictiva funcional o relacionada; tanto es así, que hace tiempo que se han incorporado a la noción de imputabilidad y en las atenuantes correspondientes. Pero si nos fijamos en la jurisprudencia, por ejemplo, la impulsividad prácticamente no es tomada en consideración más allá de los supuestos de enfermedades o trastornos mentales claramente reconocidos y asentados en la literatura psiquiátrica o en el ámbito de las adicciones. Y frente a ello, la neurociencia, además de estar aportando explicaciones científicas cada vez más completas y precisas de estas patologías y aspectos ya conocidos, está en proceso de aportar nuevas evidencias susceptibles de ser utilizadas de forma más fiable en el procedimiento penal, particularmente en la graduación de la imputabilidad²⁷. Así mismo, también está arrojando luz sobre otras vulnerabilidades inespecíficas que podrían estar en la base de, nada más y nada menos, la delincuencia antisocial o proporcionando un conocimiento más preciso acerca de, por ejemplo, los efectos del maltrato o abuso severo en la infancia.

En nuestro ordenamiento, la capacidad de control o capacidad volitiva viene definida como la aptitud que debe tener el sujeto para conformar su comportamiento a lo que prescribe el imperativo que contiene la norma,

²⁶ MORSE ha llegado a afirmar que la neurociencia carece de potencial real en materia probatoria en tanto que aporta únicamente confirmación en supuestos en los que no es realmente necesaria. En los supuestos complejos, por ejemplo, en los supuestos en que una determinada condición alegada no se corresponde con la naturaleza de los hechos realizados, incluso aunque aportara evidencias claras, la información central la seguiría aportando el comportamiento realizado, analizado con métodos psiquiátricos y psicológicos y, ante la duda, debería darse prioridad a lo que muestren los hechos. MORSE, *NeuroEthics NeuroLaw*, en «Faculty Scholarship at Penn Law», 1726. También M.J. VITACCO et al., *The Limited Relevance of Neuroimaging in Insanity Evaluations*, en «Neuroethics», 13, n. 3, 2020, pp. 249-60.

²⁷ C. SCARPAZZA et al., *Translational Application of a Neuro-Scientific Multi-Modal Approach Into Forensic Psychiatric Evaluation: Why and How?*, en «Frontiers in Psychiatry», 12, 2021, pp. 597918 ss.

conforme a lo establecido en el art. 20.1 y 2 del Código Penal. Y debe expresarse, de entrada, que las aportaciones recientes de la neurociencia hacen muy cuestionable que se prescindiera de la valoración de la capacidad volitiva en la prueba de la imputabilidad, como ocurre, por ejemplo, desde el intento de asesinato del presidente Ronald Reagan en 1982, en muchos de los ordenamientos estatales de los Estados Unidos²⁸. Tanto en el Derecho Penal español como en el italiano, sin embargo, la capacidad volitiva se reconoce en el concepto de la imputabilidad y la política subyacente a la misma es, por tanto, el reconocimiento de la afectación en la capacidad de motivación en quienes presentan vulnerabilidades o déficits biológicos en su capacidad de control.

Con respecto a la capacidad volitiva, nos vamos a referir en este apartado a dos ejemplos ilustrativos relacionados. En primer lugar, el mal llamado 'gen guerrero' o MAOA-L interactuado con factores ambientales, en la medida en que ha sido uno de los genes más estudiados hasta el momento y que ha sido alegado en múltiples ocasiones en los Tribunales, particularmente en Estados Unidos; y, en segundo lugar, ya en general, a los *biomarcadores* que están perfilándose con respecto al delincuente violento impulsivo, entre los que se incluye dicho perfil genético.

3.1. *Genética y agresividad impulsivo-reactiva: la interacción entre el gen MAOA-L y el maltrato severo en la infancia*

En la literatura científica existe consenso en que la agresividad es altamente heredable. Se estima que dentro de cualquier comunidad aproximadamente un 50 por ciento de los delitos son cometidos por menos de un 10% de las familias²⁹. Así mismo, que, con respecto al comportamiento antisocial crónico, la genética explicaría alrededor del 50% de las diferencias individuales, siempre mediadas por factores ambientales³⁰.

²⁸ Cfr. E. AHARONI et al., *Can Neurological Evidence Help Courts Assess Criminal Responsibility? Lessons from Law and Neuroscience*, en «Annals of the New York Academy of Sciences», 1124, n. 1, 2008, pp. 145-60.

²⁹ J.C. BARNES, K.M. BEAVER Y B.B. BOUTWELL, *Examining the Genetic Underpinnings to Moffitt's Developmental Taxonomy: A Behavioral Genetic Analysis*, en «Criminology», 49, n. 4, 2011, pp. 923-54. También, D.P. FARRINGTON, G.C. BARNES, S. LAMBERT, *The Concentration of Offending in Families*, en «Legal and Criminological Psychology», 1, n. 1, 1996, pp. 47-63. También, C.J. FERGUSON, *Genetic Contributions to Antisocial Personality and Behavior: A Meta-Analytic Review From an Evolutionary Perspective*, en «The Journal of Social Psychology», 150, n. 2, 2010, pp. 160-180.

³⁰ L.A. BAKER, S. BEZDJIAN, A. RAINE, *Behavioral Genetics: The Science of Antisocial*

Dentro de los genes candidatos a mediar en el comportamiento antisocial está el gen llamado Monoamina oxidasa alfa (MAO-A), regulador de la serotonina y que puede tener diferentes versiones o alelos. Pues bien, parece que si un varón (no mujeres) tiene la versión de baja actividad de este gen, el MAOA-L (de 0 a 2 repeticiones de pares de bases) se tendrán déficits en el sistema serotoninérgico y una mayor propensión hacia el comportamiento antisocial violento, de naturaleza reactiva, en la etapa adulta de la vida. Ahora bien, esta vulnerabilidad solo se materializa si (y solo si) dicho perfil genético ha interactuado con un maltrato/abuso severo en la infancia³¹. En dicha interacción los hombres experimentan una tendencia hacia la agresividad y aparecen, por así decirlo, con una impulsividad patológica, por encima de lo considerado medio o normal. Se trata, pues, de una vulnerabilidad en el sistema regulatorio de la serotonina que, cuando se combina con el maltrato severo desde la infancia se traduce en un endofenotipo caracterizado por: déficits en los circuitos de regulación emocional y de control cognitivo del comportamiento y una personalidad con déficits en el 'andamiaje socio afectivo', caracterizada por una mayor impulsividad y por una mayor emocionalidad negativa: ansiedad, depresión, ira... Finalmente, esta configuración biológica incrementa el riesgo de comportamiento antisocial, con una mayor propensión hacia el comportamiento agresivo en respuesta a un estímulo provocador en hombres³².

A mi juicio, desde un punto de vista dogmático, a partir del conocimiento aportado por la neurociencia, el principal problema que tiene el tratamiento jurídico de la interacción MAOA-L con el maltrato infantil es la falta de consistencia de los resultados científicos en cuanto a la intensidad del impacto de esta condición en la capacidad de control del sujeto. Por tanto, debe excluirse la aplicación de la eximente completa o incompleta derivada del art. 20.1 del Código Penal. A lo sumo, podría aplicarse, como se ha sostenido en un trabajo anterior, una atenuación a través de la circunstancia analógica o a través de la atenuante de arrebató del art. 21.3 del Código Penal y, eventualmente, deja abierta la puerta a la imposición de medidas de seguridad, sobre la base de un riesgo biológico hacia el comportamiento antisocial³³.

Behavior, en «Law and Contemporary Problems», 69, n. 1-2, 2006, pp. 7-46.

³¹ A. CASPI et al., *Role of Genotype in the Cycle of Violence in Maltreated Children*, en «Science», 297, n. 5582, 2002, pp. 851-54, <https://doi.org/10.1126/science.1072290>.

³² A. MEYER-LINDENBERG et al., *Neural Mechanisms of Genetic Risk for Impulsivity and Violence in Humans*, en «Proceedings of the National Academy of Sciences», 103, n. 16, 2006, pp. 6269-74.

³³ M.I. GONZÁLEZ-TAPIA, I. OBSUTH, *'Bad Genes' & Criminal Responsibility*, en

Desde una perspectiva aplicativa han surgido también otros problemas. La tenencia de este perfil genético ha sido alegada en los tribunales, particularmente en Estados Unidos, con un resultado realmente limitado. Nos interesa destacar que, hasta donde alcanzo a conocer fue alegada también en dos casos en Italia. El primero de ellos fue el caso de Abdelmalek Bayout, con la Sentencia de 18 de septiembre de 2009 emitida por la Corte de Apelación de Trieste (Italia) y en la que en un caso de homicidio se apreció una eximente incompleta (por brote psicótico) y una atenuación añadida al reconocer una predisposición genética a la violencia reactivo-impulsiva. También el caso enjuiciado en la Corte de cómo en 2011, donde se condenó a Stefania Albertani a 20 años por el asesinato de su hermana y el intento de homicidio de sus padres, reconociéndosele una atenuación de la pena por la presencia de una alteración cerebral de origen genético. En ambos casos, a mi juicio, se reflejan alguno de los principales problemas aplicativos que presenta la prueba genética y neurocientífica en los Tribunales. En el primer supuesto, planteaba esencialmente el problema de la prueba suficiente acerca del maltrato infantil sufrido por el sujeto y la ausencia de datos suficientes para establecer la graduación del elemento normativo de la imputabilidad, pero, en el segundo, se planteaba directamente la necesaria formación de los profesionales jurídicos en este ámbito, por cuanto ni la naturaleza de los hechos mostraba impulsividad, ni había base alguna para sostener la existencia de maltrato relevante en la infancia, ni la doctrina científica ha reconocido nunca los déficits propios de la interacción MAOA-L x maltrato en mujeres: se ha establecido exclusivamente en hombres³⁴.

3.2. Biomarcadores del delincuente violento

La doctrina neurocientífica ha desplegado una intensa actividad investigadora en lo relativo a las bases neurobiológicas asociadas al comportamiento antisocial de naturaleza violenta o agresiva. Obviamente, dentro de ese comportamiento antisocial, que incumple las normas

«International Journal of Law and Psychiatry», 39, 2015, pp. 60-71.

³⁴ Vid. sobre el *Caso Bayout*, Corte d'assise d'appello di Trieste, 18 settembre 2009, en «Rivista penale», 2010, 70 ss.; E. FERESIN, *Lighter sentence for murderer with 'bad genes'*, en «Nature», 2009. También, sobre el *Caso Albertani*, Tribunale di Como, sent. 20 maggio 2011, Gup Lo Gatto, en «Rivista italiana di medicina legale», 2012, pp. 246 ss., vid. S. BOTTALICO, *Neuroscienze e genetica comportamentale nel processo penale italiano*, en «Rassegna italiana di criminologia», 2013, pp. 70 ss.

establecidas con daño a terceros o a la sociedad en su conjunto, está como subespecie, el comportamiento criminal. Como ya se ha dicho, analizar la etiología del comportamiento humano es una cuestión absolutamente compleja. En lo referido a los potenciales factores *genéticos*, se han señalado una pluralidad de genes candidatos que podrían estar de alguna forma implicados en dicho comportamiento, como el MAOA-L. No obstante, existe un claro consenso en que habría una pluralidad amplia de variaciones genéticas que podrían tener incidencia y que, a día de hoy, todavía falta mucho por saber.

También se ha apuntado hacia el relevante papel que jugarían los *neurotransmisores*, p.e. de la serotonina; las *hormonas*, como la testosterona o sobre todo el cortisol; aspectos *fisiológicos*, como la hiperreactividad (o hipo-) de la amígdala y otras estructuras cerebrales; o factores *estructurales* como volúmenes aberrantes en el córtex prefrontal, el hipocampo, la amígdala, el lóbulo temporal, etc. Así mismo, se han propuesto *otros* biomarcadores como la ‘baja respuesta al miedo’, la desregulación emocional, la impulsividad, la emocionalidad negativa, la afectación del sistema de recompensa/castigo o la baja función ejecutiva...³⁵.

Como puede apreciarse resulta arduo siquiera enumerar todos los factores que han sido apuntados por la doctrina como potenciales biomarcadores del comportamiento violento. Realmente, para el Derecho, ésta sería una cuestión secundaria en su concreción específica, siendo relevante aquí, únicamente, para exponer la complejidad de lo que se está esbozando. En cualquier caso, además, debe entenderse que existe también un absoluto consenso en la incidencia que también ejercen los *factores ambientales* que conforman la experiencia del ser humano. Los factores ambientales a lo largo de la vida, desde las primeras etapas de ésta, co-determinan no solo el comportamiento final exteriorizado del sujeto, sino su propia configuración genética (plasticidad, epigenética), la configuración cerebral subyacente (como se ha visto en el caso en la interacción del MAOA-L con el maltrato) y la expresión de la personalidad resultante. Es un modelo circular y bidireccional que va de los genes al cerebro y a la personalidad y que finalmente resulta en el comportamiento. Y de entre los factores apuntados, se ha dicho, son determinantes las circunstancias adversas que se experimentan en la infancia y en la adolescencia temprana, como la *malnutrición* prenatal e infantil, la exposición prenatal a *metales pesados*, el *maltrato* severo, *abandono* o *abusos* físicos y/o sexuales, la exposición

³⁵ A. RAINE, *The anatomy of violence: the biological roots of crime*, First edition, Pantheon Books, New York, 2013.

a la *violencia* doméstica, *padres* con problemas de alcohol o adicciones o problemas de internalización crónicos, extrema *pobreza*, *institucionalización* (adopción tardía). Por tanto, no estamos hablando de un modelo simple, ingenuo, categórico y determinista, sino un modelo de una extraordinaria complejidad y de naturaleza dinámica, bidireccional y graduable en el que los factores ambientales de naturaleza socioeconómica y familiar juegan una importancia radical³⁶.

Además, en un trabajo reciente pude analizar que en la literatura científica estaba *emergiendo un perfil de delincuente violento, asociado a la violencia reactiva*³⁷. Contar con tal perfil y sus socio-biomarcadores asociados podría tener incidencia en la prueba de los déficits en la capacidad de control, mediante la aportación de biomarcadores esenciales para la prueba del elemento biológico de la imputabilidad. A su vez, de la mano de la capacidad de computación y predicción de la inteligencia artificial, dicho perfil podría estar perfilándose con parámetros dinámicos en términos temporales y graduables, lo que podría constituir en un futuro próximo una sustancial mejora en la prueba del elemento normativo del test de la imputabilidad.

En esencia, dicho perfil vendría a deslindar la psicopatía primaria o de factor 1 (cuya caracterización básica haremos en el siguiente apartado) y a confluir con otras figuras criminológicas y condiciones psiquiátricas análogas de una gran importancia para el Derecho Penal. Así, en los estudios dedicados a analizar los biomarcadores de la psicopatía se ha visto que las principales anomalías detectadas se ubican en la región temporal y en las regiones límbicas y paralímbicas y que éstas pertenecen en su inmensa mayoría exclusivamente a la psicopatía factor 2³⁸, es decir, aquellos sujetos con un persistente comportamiento antisocial, sobre todo de naturaleza agresiva- reactiva³⁹. A su vez, los estudios centrados en la violencia reactiva han destacado biomarcadores relevantes como: *el alelo del*

³⁶ Cfr. también L. MOYA ALBIOL (ed.), *Neurocriminología*, Ediciones Pirámide, Madrid, 2020. D.R. LYON, A. WELSH, *The psychology of criminal and violent behaviour*, Oxford University Press, Don Mills, Ontario, 2017, por todos.

³⁷ Cfr. GONZÁLEZ-TAPIA, *Perfil Neuropsicológico del Delincuente violento: Implicaciones Político-Criminales*, en E. DEMETRIO CRESPO (Dir.) y M. DE LA CUERDA MARTÍN, F. GARCÍA DE LA TORRE GARCÍA (Coords.), *Derecho Penal y comportamiento humano. Avances desde la neurociencia y la inteligencia artificial*, Tirant lo Blanch, 2022, pp. 255 y ss..

³⁸ J. SKEEM et al., *Two Subtypes of Psychopathic Violent Offenders That Parallel Primary and Secondary Variants*, en «Journal of Abnormal Psychology», 116, n. 2, 2007, pp. 395-409.

³⁹ M. JOHANSON et al., *A Systematic Literature Review of Neuroimaging of Psychopathic Traits*, en «Frontiers in Psychiatry», 1, 2020, pp. 1027 ss.

gen MAOA-L de baja actividad, volúmenes alterados en la estructura de la amígdala e hipereactividad de ésta; disfunciones en la 'red predeterminada por defecto'; desinhibición y desregulación emocional, disfunciones en la Teoría de la Mente, así como déficits en la función ejecutiva... En definitiva, una serie de marcadores que implican funcionalmente una hiperreactividad del circuito emocional y una baja respuesta de regulación emocional y de control cognitivo, que coinciden ampliamente con los relacionados con la psicopatía factor 2. Por su parte, desde la perspectiva de la criminología biosocial, dicho perfil convergería también con la idea de que la violencia impulsiva es altamente heredable y de que la exposición a la violencia en la infancia, entre otros factores singularmente adversos, puede ser un factor clave para la transmisión intergeneracional de esta clase violencia. A su vez, confluiría con el perfil de delincuente crónico o *life-course-persistent offender* (LCPO) en la taxonomía propuesta por Terrie MOFFITT⁴⁰.

Analizando lo anterior, la pregunta inmediata es la utilidad o las implicaciones que podría tener el hecho de que la neurociencia pudiera aportarnos realmente un perfil neurobiológico del delincuente violento impulsivo o, en general, el hecho de que pueda aportarnos biomarcadores distintivos de distintas condiciones patológicas. ¿En qué podría mejorar el concepto jurídico y la aplicación de la imputabilidad en los Tribunales? A mi juicio, dada la amplitud con que se formula el elemento biológico en nuestra regulación de la imputabilidad (anomalía o alteración psíquica) y dada la inclusión no sólo de la capacidad cognitiva, sino también de la volitiva y expresamente también la incidencia de las adicciones a las drogas o el alcohol, entiendo que la ampliación del conocimiento neurocientífico en este aspecto no cuestiona el concepto de la imputabilidad. No obstante, sí que puede integrar su contenido con condiciones patológicas que hasta ahora no tienen la consideración de enfermedad mental o trastorno de la personalidad, pero que, como en el caso del MAOA-L x maltrato, suponen una alteración psíquica. Además, al tiempo que la neurociencia puede aportarnos biomarcadores de lo patológico, también se irá conformando paralelamente un 'vector del hombre medio' o 'patrón de normalidad' que hará más fiable y objetiva la prueba pericial al respecto, paliando uno de los principales problemas de la prueba neurocientífica: las pericias contradictorias sobre la base de una misma evidencia probatoria. Igualmente, entiendo que el desarrollo de la neurociencia, de la mano de la inteligencia artificial, podrá mejorar sustancialmente en el futuro la prueba

⁴⁰ T.E. MOFFITT, *Adolescence-Limited and Life-Course-Persistent Antisocial Behavior: A Developmental Taxonomy*, en «Psychological Review», 100, n. 4, 1993, pp. 674-701.

en lo relativo a la graduación de la imputabilidad, esto es, en cuanto a la prueba del elemento normativo, de una forma más fiable, más uniforme y con una proyección temporal.

4. Psicopatía: ¿debería integrarse la falta de empatía emocional dentro del concepto de imputabilidad?

Contemplada la evolución de la literatura científica y criminológica de los últimos años, pareciera como si la psicopatía hubiera pasado de ser un 'fascinante' constructo clínico de contornos difusos, (delimitado sobre la base de la PCL-R, de Robert HARE desde 2003⁴¹, sin reconocimiento explícito en el DSM-V, ni en el CIE 11 todavía), a ser una condición reveladora por antonomasia de la peligrosidad criminal y, al mismo tiempo, el 'arquetipo del mal' y del 'cerebro violento' en la actualidad.

De hecho, probablemente, el primer y más radical problema que presenta la psicopatía a la hora de analizar la corrección o no de su tratamiento jurídico es, precisamente, la circunstancia de que no existe un acuerdo en la literatura científica acerca de su concepto. Analizada dicha doctrina, pueden apreciarse dos variantes claras de psicopatía, absolutamente diversas y que no pueden tratarse de forma unificada. Atendiendo a la crítica y caracterización hecha por SKEEM y colegas en 2007, lo que se agrupa en torno al constructo de la psicopatía tendría dos variantes dispares e irreductibles. En primer lugar, la psicopatía que podríamos denominar resumidamente *Factor 1 o primaria*, que parece reflejar la idea popular que se tiene en torno a la frialdad de carácter, a la crueldad y al encanto superficial que oculta una falta real de empatía emocional y rasgos narcisistas (*callous-unemotional traits*). Esta clase de psicopatía no estaría, como se ha visto en el apartado anterior, conectada a déficits cognitivos ni tampoco de la función ejecutiva; ni tendría raíces genéticas conocidas ni marcadores neurobiológicos claramente definidos. En diversas ocasiones se ha asociado positivamente con el volumen del córtex prefrontal y también con una hiporreactividad de la amígdala que podría explicar su característica frialdad de ánimo y una disminución acusada del miedo. Así mismo, se trataría de una variante asociada claramente a la violencia instrumental, dirigida y motivada por finalidades

⁴¹ R.D. HARE, *Manual for the revised psychopathy checklist*, Multi-Health Systems, Toronto, 2003.

y objetivos. En segundo lugar, podría identificarse la *psicopatía Factor 2 o secundaria (antisocial)*, la cual, como se ha dicho en el apartado anterior, presenta déficits cognitivos y emocionales relevantes, se ha caracterizado tradicionalmente solo por sus rasgos conductuales (paradójicamente no en rasgos psicológicos) y se caracteriza por un persistente comportamiento antisocial, sobre todo de naturaleza agresivo-reactiva.

Aparte de la discusión en torno a su concepto, la psicopatía factor 1 o primaria plantea un problema específico al Derecho Penal con respecto a la imputabilidad y que la neurociencia ha puesto sobre la mesa en los últimos años. En esencia, la psicopatía, caracterizada por la ausencia de empatía emocional de base biológica, parece cuestionar la racionalidad subyacente al juicio de inimputabilidad, que se construye exclusivamente sobre la base de la capacidad cognitiva y de la capacidad de control, pero que no atiende a la falta de capacidad de ser motivado moralmente (a través de los valores normativos). El juicio de imputabilidad, como se sabe, es un juicio sobre la capacidad de motivación del sujeto respecto de la norma penal, del imperativo que incorpora y resulta relevante si el sujeto tiene la capacidad para conocer el sentido de lo que está haciendo (u omitiendo), su carácter antijurídico y si tiene también la capacidad para adecuar su comportamiento a lo prescrito por el Derecho. Pero en la imputabilidad no interesa si el sujeto tiene capacidad o no para ser socializado conforme a valores; si tiene capacidad o no para que le ‘importen’ los valores que pretender salvaguardar el Derecho Penal. ¿Debería, por tanto, incorporarse la capacidad moral o empático-emocional en la racionalidad del concepto de imputabilidad?

La razón de esta pregunta surge a partir del momento en que la neurociencia ha cambiado el paradigma sobre el que se explica el conocimiento y el proceso de toma de decisiones del ser humano. A partir del trascendental libro de Antonio DAMASIO (1994,1999) se puso de manifiesto que el planteamiento tradicional de separación entre mente y cerebro y de ubicar la racionalidad humana únicamente en la primera... simplemente, estaba equivocado⁴². A partir de ahí, la neurociencia ha estado trabajando sobre la base de un nuevo paradigma epistemológico, que podría sintetizarse como cerebro-mente emocional para la racionalidad humana. Decía Antonio DAMASIO que en el célebre caso de Phineas Gages (1848) lo que había quedado dañado era su capacidad emocional, dando como resultado un deterioro del carácter y de ahí un grave deterioro en su capacidad de razonar moralmente y de tomar decisiones adecuadas,

⁴² A.R. DAMASIO, *Descartes' Error: Emotion, Reason and the Human Brain*, Putnam, New York, 1994, *passim*.

por ser incapaz de salvaguardar su propio interés, prever las consecuencias de su comportamiento o actuar conforme a la experiencia anterior. Y esa correlación le sugirió que los sentimientos, las emociones y la regulación biológica guiarían nuestra racionalidad, en una interrelación constante y recíproca. Parafraseando a Descartes, si se me permite, se ha pasado del ‘pienso, luego existo’, al ‘pienso y siento, luego existo’, reconociéndose a la emoción un papel crucial en la toma de decisiones de naturaleza moral. La capacidad emocional, junto a la capacidad cognitiva serían esenciales para la capacidad racional del ser humano⁴³. Y es justo aquella, de la que carece, de serie, el psicópata⁴⁴.

Esta cuestión es, sin duda, muy polémica y compleja⁴⁵. Andrea GLENN y Adrian RAINE, por ejemplo, entienden que los psicópatas podrían tener comprometida su capacidad de agencia, al menos en lo que se refiere a su capacidad de enjuiciamiento moral. Aunque podrían distinguir cognitivamente lo que es correcto y lo que es ilícito, podría faltarles la capacidad emocional suficiente para motivarse hacia el comportamiento moralmente correcto y que ello podría tener repercusiones en su responsabilidad criminal⁴⁶. En un sentido similar, MORSE ha propuesto que en los casos más graves sean considerados inimputables y que se les apliquen medidas de seguridad, como el internamiento voluntario⁴⁷.

En mi opinión, sin embargo, expresada en un trabajo anterior con más detenimiento, hoy por hoy no hay base suficiente que permita replantearse el tratamiento jurídico del psicópata; menos aún del psicópata Factor 1 o primario. Tampoco entiendo que deba incorporarse la capacidad emocional a la estructura de la imputabilidad.

En primer lugar, porque haría falta un consenso suficiente en la literatura científica en cuanto a su concepto, rasgos característicos y en

⁴³ Cfr. A. RAINE, Y. YANG, *Neural foundations of moral reasoning and antisocial behavior*, en «Social Cognitive and Affective Neuroscience», 1, n. 3, 2006, pp. 203–213.

⁴⁴ R.J. BLAIR, Responding to the emotions of others: Dissociating forms of empathy through the study of typical and psychiatric populations, en «Consciousness and Cognition», 14, n. 4, 2005, pp. 698–718.

⁴⁵ Sobre esta cuestión, también M. CANCIO MELIÁ, *Psicopatía y derecho penal: algunas consideraciones introductorias*, en *Neurociencias y derecho penal: nuevas perspectivas en el ámbito de la culpabilidad y tratamiento jurídico penal de la peligrosidad*, dirigido por E. Demetrio Crespo y coordinado por M. Maroto Calatayud, Edisofer, Madrid, 2013, pp. 529-546.

⁴⁶ A.L. GLENN, A. RAINE, *Psychopathy. An introduction to biological findings and their implications*, New York University Press, New York, 2014.

⁴⁷ S.J. MORSE, *Psychopathy and Criminal Responsibility*, en «Neuroethics» 1, 2008, pp. 205–212.

los biomarcadores de este constructo y esa circunstancia no se da en la actualidad. Haría falta, particularmente, una delimitación adecuada con el trastorno de la personalidad antisocial con rasgos psicopáticos. En segundo lugar, porque el Derecho Penal no busca con su estrategia preventiva que los ciudadanos, necesariamente, interioricen los valores subyacentes a la norma, sin perjuicio de reconocer una mayor eficacia de la prevención integradora frente a la meramente intimidatoria. Además, en tercer lugar, porque entiendo que ni la decisión de cometer un delito es una decisión moral, en sentido estricto (es una decisión ilegal o antijurídica), ni el Derecho debe exigir un compromiso interno del sujeto con respecto a los valores de la norma. Simplemente, se debe exigir respeto hacia lo prescrito en la norma, lo que debe lograrse a partir del conocimiento de la prohibición y de la capacidad de control, con independencia de la motivación o compromiso interno del sujeto. No creo que el Derecho deba (ni siquiera que pudiera subsistir como tal) incorporar la capacidad emocional al concepto de imputabilidad y desde ahí a la noción de responsabilidad penal. Baste pensar en lo que ello supondría, por ejemplo, en la delincuencia por adoctrinamiento cultural desde la infancia, como ocurre p.e. en el terrorismo yihadista radical y, en general, en la delincuencia por convicción⁴⁸.

5. Bibliografía

- AHARONI E. et al., *Can Neurological Evidence Help Courts Assess Criminal Responsibility? Lessons from Law and Neuroscience*, en «Annals of the New York Academy of Sciences», 1124, n. 1, 2008, pp. 145-60.
- ALCES P.A., SAPOLSKY R.M., *Nohwere*, en «William & Mary Law Review 1079», 63, n. 4, 2022.
- BAKER L.A., BEZDJIAN S., RAINE A., *Behavioral Genetics: The Science of Antisocial Behavior*, en «Law and Contemporary Problems», 69, n. 1-2, 2006.
- BARNES J.C., BEAVER K.M., BOUTWELL B.B., *Examining the Genetic Underpinnings to Moffitt's Developmental Taxonomy: A Behavioral Genetic Analysis*, en «Criminology», 49, n. 4, 2011.
- BENNETT M.R., HACKER P.M., *Philosophical foundations of neuroscience*, Second Edition, John Wiley & Sons, Hoboken, New Jersey, 2022.

⁴⁸ Cfr. GONZÁLEZ-TAPIA, I. OBSUTH y R. HEEDS, *A New Legal Treatment for Psychopaths? Perplexities for Legal Thinkers*, en «International Journal of Law and Psychiatry», 54, 2017, pp. 46-60.

- BOTTALICO S., *Neuroscienze e genetica comportamentale nel processo penale italiano*, en «Rassegna italiana di criminologia», 2013.
- BURNS J., RUSSELL H.S., *Right orbitofrontal tumor with pedophilia symptom and constructional apraxia sign*, en «Archives of neurology», 60, n. 3, 2003.
- CANCIO MELIÁ M., *Psicopatía y derecho penal: algunas consideraciones introductorias*, en *Neurociencias y derecho penal: nuevas perspectivas en el ámbito de la culpabilidad y tratamiento jurídico penal de la peligrosidad*, dirigido por Demetrio Crespo E. y coordinado por Maroto Calatayud M., Edisofer, Madrid, 2013.
- CASPI A. et al., *Role of Genotype in the Cycle of Violence in Maltreated Children*, en «Science», 297, n. 5582, 2002.
- DAMASIO A.R., *Descartes' Error: Emotion, Reason and the Human Brain*, Putnam, New York, 1994.
- DEMETRIO CRESPO E., MAROTO CALATAYUD M. (eds.), *Neurociencias y derecho penal: nuevas perspectivas en el ámbito de la culpabilidad y tratamiento jurídico-penal de la peligrosidad*, Edisofer, Madrid, 2013.
- FARRINGTON D.P., BARNES G.C., LAMBERT S., *The Concentration of Offending in Families*, en «Legal and Criminological Psychology», 1, n. 1, 1996.
- FERESIN E., *Lighter sentence for murderer with 'bad genes'*, en «Nature», 2009.
- FEIJOO SÁNCHEZ B., *Derecho Penal y Neurociencias. ¿Una relación tormentosa?*, en «InDret», n. 2, 2011.
- FEIJOO SÁNCHEZ B. (ed.), *Derecho penal de la culpabilidad y neurociencias*, Civitas, Madrid, 2012.
- FERGUSON C.J., *Genetic Contributions to Antisocial Personality and Behavior: A Meta-Analytic Review From an Evolutionary Perspective*, en «The Journal of Social Psychology», 150, n. 2, 2010.
- GARCÍA-MOLINA A., *Phineas Gage y el enigma del córtex prefrontal*, en «Neurología», 27, n. 6, 2012.
- GAZZANIGA M.S., *The Law and Neuroscience*, en «Neuron», 60, n. 3, 2008.
- GIMBERNAT ORDEIG E., *Concepto y método de la ciencia del derecho penal*, Ed. Tecnos, Madrid, 1999.
- GLENN A.L., RAINE A., *Psychopathy. An introduction to biological findings and their implications*, New York University Press, New York, 2014.
- GONZÁLEZ-TAPIA M.I., *Derecho Penal de la peligrosidad y Neuropredicción: hacia un 'Derecho Penal Individualizado'*, en *Estudios jurídico penales y criminológicos en homenaje al Prof. Dr. Dr. h. c. mult. Lorenzo Morillas Cueva*, coordinado por Benítez Ortúzar I.F., Morillas Fernández D.L., Sáinz-Cantero Caparrós J.E., Suárez López J.M., vol. 1, Dykinson, Madrid, 2018.
- GONZÁLEZ-TAPIA M.I., *Peligrosidad de la conducta y del autor en el Derecho*

- Penal actual: implicaciones político-criminales y dogmáticas*, en «Revista de Derecho, Empresa y Sociedad» 13, 2018.
- GONZÁLEZ-TAPIA M.I., OBSUTH I., HEEDS R., *A New Legal Treatment for Psychopaths? Perplexities for Legal Thinkers*, en «International Journal of Law and Psychiatry», 54, 2017.
- GONZÁLEZ-TAPIA M.I., OBSUTH I., *'Bad Genes' & Criminal Responsibility*, en «International Journal of Law and Psychiatry», 39, 2015.
- GONZÁLEZ-TAPIA M.I., *Neuroderecho: implicaciones para el Derecho Penal, en Locura y Crimen: Psiquiatría y Ley*, editado por A. Medina et al., e-Book: Fundación Española de Psiquiatría y Salud Mental, 2018.
- GONZÁLEZ-TAPIA M.I., *Perfil Neuropsicológico del Delincuente violento: Implicaciones Político-Criminales*, en E. DEMETRIO CRESPO (Dir.) y M. DE LA CUERDA MARTÍN, F. GARCÍA DE LA TORRE GARCÍA (Coords.), *Derecho Penal y comportamiento humano. Avances desde la neurociencia y la inteligencia artificial*, Tirant lo Blanch, 2022, pp. 255 y ss.
- GOODENOUGH O.R., TUCKER M., *Law and Cognitive Neuroscience*, en «Annual Review of Law and Social Science», 6, n.º 1, 2010.
- GREENE J., COHEN J., *For the Law, Neuroscience Changes Nothing and Everything*, en «Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences», 359, n. 1451, 2004.
- HARE R.D., *Manual for the revised psychopathy checklist*, Multi-Health Systems, Toronto, 2003.
- JENSEN P. et al., *Crime in Huntington's Disease: A Study of Registered Offences among Patients, Relatives, and Controls*, en «Journal of Neurology, Neurosurgery & Psychiatry», 65, n. 4, 1998.
- JOHANSON M. et al., *A Systematic Literature Review of Neuroimaging of Psychopathic Traits*, en «Frontiers in Psychiatry», n. 1, 2020.
- LIBET B. et al., *Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential): the unconscious initiation of a freely voluntary act*, en «Brain», 106, n. 3, 1983.
- LISZT F. v., *La idea de fin en el derecho penal*, Ediciones Olejnik, Santiago de Chile, 2020.
- LYON D.R., WELSH A., *The psychology of criminal and violent behaviour*, Oxford University Press, Don Mills, Ontario, 2017.
- MEYER-LINDENBERG A. et al., *Neural Mechanisms of Genetic Risk for Impulsivity and Violence in Humans*, en «Proceedings of the National Academy of Sciences», 103, n. 16, 2006.
- MOFFITT T.E., *Adolescence-Limited and Life-Course-Persistent Antisocial Behavior: A Developmental Taxonomy*, en «Psychological Review», 100, n. 4, 1993.

- MORSE S.J., *Psychopathy and Criminal Responsibility*, en «Neuroethics», n. 1, 2008.
- MORSE S.J., *Brain Overclaim Syndrome and Criminal Responsibility: A Diagnostic Note*, retrieved from https://Repository.Upenn.Edu/Neuroethics_pubs/28, 2006.
- MORSE S.J., *Genetics and Criminal Responsibility*, en «Trends in Cognitive Sciences», University of Pennsylvania Law School, Public Law Research Paper N. 11-34, 2011.
- MORSE S.J., *NeuroEthics NeuroLaw*, en «Faculty Scholarship at Penn Law», 1726.
- MOYA ALBIOL L. (ed.), *Neurocriminología*, Ediciones Pirámide, Madrid, 2020.
- QUINTERO OLIVARES G., *Aspectos generales*, en *Derecho Penal Constitucional*, dirigido por Quintero Olivares G., coordinado por Jaria i Manzano J., Tirant, Valencia, 2015.
- RAINE A., *The anatomy of violence: the biological roots of crime*, First edition, Pantheon Books, New York.
- RAINE A., YANG Y., *Neural foundations of moral reasoning and antisocial behavior*, en «Social Cognitive and Affective Neuroscience», 1, n. 3, 2006.
- ROXIN C., *Derecho penal: Parte General. Tomo I*, trad. Luzón Peña D.M., Díaz y García Conlledo M., de Vicente Remesal J., vol. I, Thomson Reuters-Civitas, Madrid, 2008.
- ROXIN C., *Política criminal y sistema del Derecho penal*, traducción e introducción de F. Muñoz Conde, Bosch, Barcelona, 1972.
- SCARPAZZA C. et al., *Translational Application of a Neuro-Scientific Multi-Modal Approach Into Forensic Psychiatric Evaluation: Why and How?*, en «Frontiers in Psychiatry», 12, 2021.
- SECCIÓN INDICAL ESTATAL DE CCOO EN INSTITUCIONES PENITENCIARIAS, *Informe Salud Mental en prisiones*, <www2.fsc.ccoo.es> (u.v. 28/07/2017).
- SKEEM J. et al., *Two Subtypes of Psychopathic Violent Offenders That Parallel Primary and Secondary Variants*, en «Journal of Abnormal Psychology», 116, n. 2, 2007.
- VITACCO M.J. et al., *The Limited Relevance of Neuroimaging in Insanity Evaluations*, en «Neuroethics», 13, n. 3, 2020.
- ZEKI S., GOODENOUGH O.R., SAPOLSKY R.M., *The Frontal Cortex and the Criminal Justice System*, en «Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 359», n. 1451, 2004.

Autores

RAFAEL CASTILLO FELIPE, *Profesor Contratado Doctor de Derecho Procesal, Universidad de Murcia*

ANGELO GIRALDI, *Investigador Contratado de Derecho Penal, Universidad de Murcia - Università degli Studi 'Roma Tre'*

MARÍA ISABEL GONZÁLEZ TAPIA, *Profesora Titular de Derecho Penal, Universidad de Córdoba*

LORENZA GROSSI, *Dottoranda di ricerca in Diritto penale, Università degli Studi 'Roma Tre'*

ANTONELLA MASSARO, *Professoressa associata di Diritto penale, Università degli Studi 'Roma Tre'*

DAVID LORENZO MORILLAS FERNÁNDEZ, *Catedrático de Derecho Penal, Universidad de Murcia*

LAURA NOTARO, *Dottoranda di ricerca in Diritto penale, Università di Pisa*

JOSÉ MANUEL PALMA HERRERA, *Catedrático de Derecho Penal, Universidad de Córdoba*

JACINTO PÉREZ-ARIAS, *Profesor Contratado Doctor de Derecho Penal, Universidad de Murcia*

JAIME MIGUEL PERIS RIERA, *Catedrático de Derecho Penal, Universidad de Murcia*

SALVADOR TOMÁS TOMÁS, *Profesor Contratado Doctor de Derecho Procesal, Universidad de Murcia*

MARIO TRAPANI, *Professore ordinario di Diritto penale a r., Università degli Studi 'Roma Tre'*

El impacto de las nuevas tecnologías en el Derecho es cada vez más evidente. El jurista está llamado a afrontar cuestiones nuevas y complejas, al tiempo que verifica la solidez de las categorías e instituciones sobre las que se asienta el edificio jurídico “tradicional”. Por lo que se refiere en particular al Derecho penal, las perspectivas esbozadas tanto por la neurociencia como por la inteligencia artificial han provocado reacciones diferentes. Frente a quienes consideran que se trata de herramientas destinadas a producir sus efectos únicamente en el plano procesal, están quienes prevén o incluso esperan una reescritura de los principios e instituciones fundamentales, empezando por el concepto mismo de responsabilidad. El libro, recorriendo los aspectos más problemáticos por el lado de la neurociencia y la inteligencia artificial, pone de relieve, por un lado, las innegables innovaciones que caracterizan la reflexión penal más reciente y, por otro, se detiene en el resurgimiento, aunque bajo formas peculiares, de lo que, desde hace siglos, representa auténticos nudos irresueltos del sistema penal.

Jaime Miguel Peris Riera es Catedrático de Derecho Penal en la Facultad de Derecho de la Universidad de Murcia (España), donde imparte, entre otras, las asignaturas de Derecho Penal I, Derecho Penal II y Peligrosidad Criminal, y donde es Director del Departamento de Historia Jurídica y de Ciencias Penales y Criminológicas. También dirige el Grupo Investigador en Ciencias Penales y Criminológicas y es autor de numerosas publicaciones sobre el estudio de la parte general y especial del Derecho Penal, la victimología y la biogenética.

Antonella Massaro es Profesora Titular de Derecho Penal en el Departamento de Derecho de la Universidad “Roma Tre” (Italia), donde imparte las asignaturas de Derecho Penal II, Derecho Penal Europeo y *Law and Gender*. Sus temas de investigación incluyen, en particular, el principio de legalidad, el delito culposo por omisión, la eutanasia y los derechos fundamentales de las personas detenidas.