

## Dossier «Revolució 4.0: progrés o precarització?»

Coordinador: Josep Lladós

ELS LÍMITS DE L'AVENÇ TECNOLÒGIC

## El jutge artificial: propera parada?<sup>1</sup>

**David Martínez Zorrilla**

Professor agregat dels Estudis de Dret i Ciència Política (UOC)

**RESUM** Els avenços tecnològics recents, i especialment la intel·ligència artificial, estan afectant de manera molt profunda els més diversos àmbits, entre els quals el jurídic, en les seves més diverses manifestacions. Aquest treball analitza alguns dels més recents desenvolupaments en intel·ligència artificial en el camp de l'aplicació del dret, a fi de destriar fins a quin punt ens estem apropant a la ja no tan utòpica idea del jutge artificial i l'aplicació automatitzada del dret. L'anàlisi s'estructura en tres àmbits diferenciats, encara que complementaris: en primer lloc, s'analitza l'aspecte conceptual (quins elements i activitats implica l'aplicació judicial del dret, a fi de determinar quins processos caldria automatitzar per a poder parlar d'un «jutge artificial»); en segon lloc, es presenten breument algunes tecnologies i aplicacions recents que d'alguna manera incidirien en l'aplicació del dret, a fi de determinar en quin punt ens trobaríem respecte d'un procés d'automatització; i finalment, es reflexiona breument sobre els possibles efectes positius o adversos que comportaria una automatització de l'aplicació del dret.

**PARAULES CLAU** aplicació del dret; intel·ligència artificial; informàtica jurídica; jutge artificial

### *The AI judge: next stop?*

**ABSTRACT** *The recent advances in technology, and particularly artificial intelligence, are having an extremely profound effect on very diverse sectors, including that of the judiciary in its various manifestations. The work presented here analyses some of the more recent developments in artificial intelligence in the field of the application of law, with the aim of discerning the point up to which we are approaching the idea –not so utopian now– of the AI judge and the automated application of law. The analysis is structured in three different areas, which are at the same time complimentary: firstly, there is an analysis of the conceptual aspect (what distinct elements and activities involve the judicial application of law, with the aim of determining what processes would be necessary to automate in order to be able to speak of an “AI judge”); secondly, a number of recent technologies and applications are presented which would influence in some way the appli-*

1. Aquest article és una síntesi i reelaboració parcial de MARTÍNEZ ZORRILLA, D. (2019): «La decisión judicial automatizada: entre la ciencia y la ficción». A: VIAL DUMAS, M. y MARTÍNEZ ZORRILLA, D. (ed.): *Pensando al juez*. Madrid: Marcial Pons, p. 152-173.

*cation of law, with the purpose of determining at what point we would find ourselves with respect to a process of automation; and finally, there will be a brief reflection on the possible positive or negative effects an automation of the application of law would involve.*

**KEYWORDS** *application of law, artificial intelligence, Information Technology Law, AI judge*

## I

Al llarg dels segles, una de les característiques més destacades de l'espècie humana ha estat la de servir-se dels avenços científics i tecnològics per a les més diverses finalitats, ja que ofereixen millores substancials en termes d'eficiència, precisió i estalvi de temps i costos, entre altres aspectes. En aquest sentit, ens permeten desplaçar-nos més ràpidament, construir estructures més grans i sòlides, realitzar càlculs molt complexos de manera més ràpida i precisa o, fins i tot, viure més i millor (gràcies als avenços mèdics), entre altres incomptables exemples.

Vist en perspectiva històrica, no és exagerat afirmar que les tecnologies relacionades amb el que en sentit ampli podem denominar com a «informàtica» serien les més revolucionàries i disruptives de tots els temps, a un nivell equiparable, per exemple, al del domini del foc, la roda o la impremta. I dins d'aquest marc, el que es pot denominar com a «intel·ligència artificial» (o IA, per abreviar) és segurament la seva expressió més avançada, rellevant i amb major projecció de futur en pràcticament tots els àmbits imaginables. Tant per raons d'extensió com perquè no és l'objecte d'aquest treball, no entraré en la complexa i discutida qüestió (de profundes implicacions filosòfiques) de què s'entén exactament per IA<sup>2</sup>, però a ningú se li escapa que actualment podem veure innumbrables mostres d'aplicacions d'IA en els més diversos contextos: des de l'àmbit de la domòtica (en què mitjançant múltiples sensors un dispositiu és capaç de controlar les llums, la climatització, les persianes i tendals, els panys; o els frigorífics 'intel·ligents' que realitzen comandes de subministraments en funció de les existències), al del transport i la mobilitat (des d'ajudes a la conducció, com ara la detecció de senyals de tràfic, el control automàtic de la distància de seguretat, el càlcul de la ruta més ràpida tenint en compte la situació del tràfic en temps real o el reconeixement de vianants per activar si escau la frenada automàtica, fins a la conducció completament autònoma), al de l'assistència a persones malaltes o dependents (robots cuidadors), al del processament massiu de dades per a formular prediccions o, fins i tot, prendre i executar decisions (com, per exemple, on invertir per a obtenir un major rendiment, a partir de l'anàlisi de tendències), i un llarguíssim etcètera.

Encara que l'àmbit del dret, i especialment el de l'Administració de Justícia, mai s'ha caracteritzat precisament per estar a l'avantguarda dels avenços tecnològics, tampoc s'ha mantingut aliè a aquests, i és just reconèixer que en els últims anys s'ha fet un esforç de modernització considerable, com ho mostra, per exemple, el desenvolupament i implementació de LEXNET<sup>3</sup> com a plataforma integral d'administració electrònica de l'Administració de Justícia. No obstant això, quan parlem de l'IA aplicada a l'àmbit judicial, és natural pensar en alguna cosa que va més enllà de la gestió documental o les bases de dades, i ens podríem plantejar, encara que sigui com a mera possibilitat teòrica, l'aplicació automatitzada del dret (presa de decisions legals), arribant fins i tot a la idea d'un «jutge artificial». Es tracta d'alguna cosa totalment fantàsica? Si bé podem dir que en l'actualitat encara estem força lluny de comptar amb processos integrals artificials d'aplicació del dret (excepte, potser, en contextos molt delimitats), els avenços han estat molt significatius en els últims anys (i previsiblement ho seran encara més en el futur), amb resultats sorprenents en no pocs casos, que fan que legítimament ens puguem plantejar la possibilitat que en el futur els processos d'aplicació del dret puguin ser (gairebé) totalment automàtics.

---

2. Sobre aquestes qüestions, vegeu, per exemple, BODEN, M. A. (ed.) (1990): *The Creative Mind: Myths and Mechanisms*. Londres: Routledge.

3. <https://lexnetjusticia.gob.es/inicio> [Data de consulta: 26 de juny de 2019].

Però abans de referir-nos als exemples tecnològics concrets, cal aclarir una qüestió conceptual: en què consisteix o què s'entén exactament per «aplicació del dret»? Quins processos i passos implica? Tots són automatitzables o reproduïbles de manera artificial? D'altra banda, també es planteja una important qüestió ètica o valorativa: fins a quin punt pot ser positiu o desitjable deixar les decisions jurídiques en mans d'algorismes o processos automatitzats?

Tenint en compte l'anterior, aquest treball aborda la qüestió de l'aplicació automatitzada del dret des de diferents dimensions o perspectives d'anàlisi: a) En primer lloc, des d'una perspectiva conceptual, realitzant una primera aproximació a quins serien els diferents processos implicats en l'aplicació del dret per així determinar què s'hauria d'«automatitzar» per a poder parlar d'aplicació automatitzada del dret o decisió judicial automatitzada; b) En segon lloc, des d'una perspectiva descriptiva, fent referència a alguns exemples actuals de desenvolupament de tecnologies que d'alguna manera incideixen en l'aplicació del dret, i reflexionant sobre la seva incidència o rellevància en el procés d'«automatització» de la decisió judicial; i c) Finalment, des d'una perspectiva valorativa, reflexionant sobre la conveniència o la desitjabilitat de deixar en mans d'algorismes automatitzats processos o decisions que fins ara sempre han estat en mans d'éssers humans, i que en major o menor mesura depenen de la seva capacitat de comprensió i de judici.

## II

En què consisteix l'aplicació del dret? En un primer acostament, sembla que ens referim a la resolució de situacions «conforme a dret», és a dir, utilitzant els paràmetres i els criteris establerts pel sistema jurídic, que determinarien, o almenys guiarien, quina és la resposta que s'ha d'oferir a la situació en qüestió sobre la qual s'ha de decidir. Cal tenir en compte que tant els tipus de situacions que d'alguna manera són abordades pel dret, com els tipus de resposta que aquest ofereix, són extremadament variats (pensem, per exemple, des de la imposició d'una sanció per una infracció de tràfic fins a la determinació de la validesa d'un contracte, passant per la protecció d'un dret fonamental, la decisió sobre la responsabilitat per danys i la determinació de la corresponent indemnització, la concessió d'una llicència administrativa o la decisió sobre la incapacitació d'una persona, entre una infinitat d'altres situacions). Aquesta varietat no solament afecta el tipus de situacions que d'alguna manera són abordades pel dret, sinó també les respostes ofertes per aquest: poden ser tant extremadament específiques (com la quantia econòmica de la sanció a imposar per superar el límit màxim de velocitat en 20 km/h en una via interurbana), com àmpliament genèriques i indeterminades (com, per exemple, les mesures cautelars a prendre per un jutge per a assegurar el compliment d'una possible sentència condemnatòria). Per tant, hauria de resultar una definició al més àmplia possible per a poder abastar tota aquesta varietat, al mateix temps que hauria de mantenir els seus elements distintius per no abastar més del necessari i donar així entrada a aspectes estranys al fenomen jurídic. La meua proposta seria entendre l'aplicació del dret com *l'establiment d'una resposta a una determinada situació d'acord amb els paràmetres o criteris establerts pel sistema jurídic*<sup>4</sup>.

Quan un decisor (en general, un funcionari o un jutge) ha de determinar la resposta jurídica al cas que se li planteja, sembla que ha de dur a terme un conjunt d'activitats que resulten quelcom heterogènies, però totes elles necessàries per a poder justificar la seva decisió. Semblaria que, almenys, ha de fer el següent (no necessàriament seguint un ordre cronològic estrict):

- a. La selecció del material jurídic rellevant (recopilació i categorització de la normativa, jurisprudència, i d'altres materials que puguin o s'hagin de tenir en compte per a fonamentar o justificar les decisions jurídiques).

---

4. Aquesta definició procés-producte pateix d'ambigüïtat, encara que de manera intencionada: es refereix tant al fet que el *procés* pel qual es determina una resposta a una situació donada s'ha de regir pels criteris i paràmetres establerts pel sistema jurídic (l'òrgan ha de ser el competent, seguint els procediments i els requisits legalment establerts, etc.), com al fet que el *resultat* o producte (això és, la resposta o decisió) s'ajusti a les exigències jurídiques establertes (és a dir, que sigui, com se sol dir, «conforme a dret»).

- b. La interpretació d'aquest material per a la determinació del seu significat i l'obtenció dels elements del sistema normatiu a utilitzar per a donar una resposta a la situació plantejada.
- c. La determinació dels fets del cas concret (prova), que a vegades també exigeix l'establiment de relacions de causalitat entre els esdeveniments empírics.
- d. La qualificació jurídica dels fets provats (subsumció d'enunciats fàctics en categories jurídiques).
- e. La sistematització del sistema normatiu per a la determinació de la resposta jurídica a cadascuna de les situacions o casos genèrics que regula.
- f. La subsumció del cas concret en el cas genèric corresponent, per així obtenir la reposada jurídica al cas de conformitat amb les normes del sistema.

El que es desprèn, fins i tot, d'una anàlisi molt preliminar és que la decisió judicial comprèn múltiples i molt diverses activitats, que requereixen diferents tècniques i habilitats per part del decisor. Per això, per a poder concebre un model complet o integral de decisió judicial automatitzada, s'hauria de poder comptar amb un sistema integrat de tecnologies i processos capaços de realitzar tasques tan diferents com crear i processar bases de dades, la interpretació del llenguatge natural o el reconeixement d'imatges i sons, entre altres. Per això, semblaria més concebible i factible la creació de tecnologies que realitzin o facilitin la realització de certes parts, tasques o fases del procés de decisió judicial, almenys a curt i mitjà termini.

Ja en una primera aproximació resulta evident que la possibilitat d'automatitzar diferents activitats planteja desafiaments i dificultats molt diferents en funció del tipus de procés que es tracti, i que alguns d'aquests són més fàcilment factibles que uns altres. Per exemple, la selecció i categorització del material jurídic, gràcies a la cada vegada major capacitat de processament d'informació i al *big data*, o la sistematització de sistemes normatius (determinació i assignació de les conseqüències jurídiques dels diferents casos genèrics) no semblen plantejar desafiaments tècnics importants, fins i tot en l'estat actual de la tecnologia o en el del futur immediat. En canvi, altres tasques que normalment vinculem a conceptes i categories tan «humans» com el judici i la comprensió, com podrien ser la interpretació jurídica, la subsumció de fets en categories normatives o la determinació de quan un fet es pot considerar «provat» a partir dels elements probatoris disponibles, semblen difícilment automatitzables o reductibles en un algorisme informàtic. No obstant això, seria un error concloure sense més que aquest tipus de tasques són impossibles d'automatitzar, encara que en el context actual es puguin concebre com alguna cosa molt llunyana o senzillament impossible. Els avenços tecnològics experimentats per la humanitat al llarg de la història en no poques ocasions ens han sorprès mostrant-nos coses que ens semblaven impossibles, i no hi ha cap raó per pensar que no seguirà sent així en el futur. A més, resulta molt rellevant la matèria o el tipus de decisions jurídiques que es tracti. Les possibilitats d'utilitzar satisfactòriament algorismes automatitzats per a la presa de decisions són molt majors quan es tracta d'àmbits molt delimitats i amb criteris molt objectius com, per exemple, el d'infraccions de tràfic o l'àmbit fiscal, mentre que es complica significativament en aquelles altres qüestions amb una major dependència del judici o la valoració subjectiva de qui jutja, com quan cal valorar si hi va haver «mala fe» per part del venedor o si en un robatori hi va haver o no «violència», per usar només dos exemples.

## III

Tot seguit comentaré breument alguns exemples de tecnologies basades en l'IA que d'alguna manera estan relacionades amb l'aplicació del dret. Alguns d'aquests són mers experiments o proves, mentre que en altres casos es tracta de tecnologies plenament operatives<sup>5</sup>.

### i. *Case Cruncher Alpha*<sup>6</sup>

Quatre estudiants de Dret de Cambridge van crear un algorisme d'intel·ligència artificial que, segons afirmaven, era capaç de predir els resultats de decisions judicials o administratives amb un nivell molt alt d'encert, superior al que poden obtenir els agents humans. Per a posar-ho a prova, es va realitzar un experiment en el qual es va proporcionar, a més d'un centenar d'advocats de Londres de bufets prestigiosos, informació bàsica dels fets sobre 775 casos d'assegurances de crèdit (*payment protection insurance*) que ja havien estat decidits per l'autoritat competent (però el resultat de la qual desconeixien), perquè indiquessin quina decisió previsiblement havia pres el *financial Ombudsman* (l'autoritat competent per decidir sobre aquesta matèria). La mateixa informació va ser proporcionada a l'algorisme. Una vegada acabada l'anàlisi, el percentatge d'encert (això és, el nivell de coincidència entre la resposta dels participants i les decisions reals) dels advocats va ser del 66,3%, mentre que l'algorisme *Case Cruncher Alpha* va aconseguir un percentatge del 86,6%. És a dir, l'experiment va demostrar que l'algorisme d'intel·ligència artificial era sensiblement millor que un grup d'advocats competents pel que fa a la predicció del resultat d'un procés judicial o administratiu (d'aquestes característiques, almenys).

Tècnicament, en aquest experiment no es pot parlar pròpiament de *predicció* (ja que que es tractava de casos ja resolts), però això manca d'importància, en el sentit que el mateix algorisme podria ser aplicat a casos no resolts per a proporcionar autèntiques prediccions sobre el seu probable resultat<sup>7</sup>. Però l'aspecte més interessant és el que té a veure precisament amb allò que l'algorisme realitza. Estrictament parlant, no estem davant d'una tecnologia per *decidir* o resoldre casos «conforme a dret», per la qual cosa no hi hauria *aplicació* del dret, sinó per formular hipòtesis sobre com serien resolts. Dit d'una altra manera, l'algorisme no pretén determinar la *resposta correcta* o que procedeix per al cas plantejat d'acord amb els criteris o estàndards jurídics aplicables, sinó formular una proposició sobre quina serà la resposta més probable que dictarà l'òrgan competent, la qual cosa, encara que molt interessant, és alguna cosa conceptualment molt diferent a «aplicar el dret».

No obstant això, és probable que, per a poder realitzar les seves prediccions, l'algorisme d'alguna manera tingui en compte el contingut del dret. Però com que no té coneixement respecte del seu funcionament intern, és difícil valorar en quina mesura el dret és tingut en compte o quina és la seva rellevància en el procés d'elaboració de les prediccions. Tanmateix, considerant que l'objectiu principal és, per dir-ho així, epistèmic (preveure què ocorrerà) i no justificador, és probable que altres aspectes resultin tan o més rellevants, com l'anàlisi històrica de decisions anteriors per a extreure pautes o tendències reiteratives (no necessàriament jurídiques o basades en normes jurídiques) que permetin fonamentar les regularitats en què basar les prediccions. En suma, el que aquesta tecnologia mostraria és la constatació d'alguna cosa ja àmpliament assumida: que les màquines són més eficients que els éssers humans en el processament i anàlisi de grans quantitats d'informació.

---

5. La informació ha estat obtinguda fonamentalment de fonts periodístiques, i atès que els meus coneixements tècnics sobre la intel·ligència artificial són limitats i no he tingut la possibilitat de conèixer o utilitzar directament cap de les tecnologies aquí exposades, no puc garantir que la descripció sigui precisa, malgrat que parteixo de la pressuposició que, en línies generals, les informacions proporcionades per les fonts periodístiques són correctes.

6. Podeu trobar informació sobre aquesta plataforma i l'experiment realitzat a què aquí es fa referència, per exemple, a <http://www.bbc.com/news/technology-41829534> [Data de consulta: 26 de juny de 2019].

7. De fet, aquesta tecnologia ha servit de base per a la creació d'una empresa que proporciona serveis de predicció sobre el resultat de futures o hipotètiques decisions: <http://www.case-crunch.com/> [Data de consulta: 26 de juny de 2019].

Un altre aspecte a considerar és que aquest algorisme no sembla exercir cap funció quant a la determinació dels fets del cas. Pel que sembla deduir-se a partir de la informació disponible, totes les dades relatives al cas, sobre la solució del qual es vol obtenir una predicció, han de ser proporcionades a l'algorisme (per éssers humans, molt probablement). Tampoc queda clar si el software compta amb alguna capacitat que es pugui considerar com una «interpretació» de les fonts normatives, si bé és probable que es limiti a tenir en compte quines són les interpretacions més habituals per part dels tribunals de les diferents disposicions afectades, ja que aquest extrem aporta un gran valor predictiu.

En resum, si tenim en compte el conjunt d'activitats principals vinculades a l'aplicació del dret, tal com s'exposa en l'apartat II, es podria dir que és molt probable que aquesta tecnologia dugui a terme una selecció del material jurídic rellevant i que probablement realitzi tasques de sistematització, encara que és força dubtós que desenvolupi activitats relacionades amb la interpretació jurídica, i sembla que no exerceix en absolut tasques vinculades a la selecció i prova dels fets del cas individual.

## ii. L'algorisme COMPAS dels tribunals de Wisconsin<sup>8</sup>

COMPAS és l'acrònim de *Correctional Offender Management Profiling for Alternative Sanctions*, software utilitzat, entre altres, per l'Administració de Justícia de l'Estat de Wisconsin (EUA) que proporciona una estimació sobre el grau de perillositat i del risc de reincidència del processament, a fi de servir com a criteri per a la determinació de la pena a imposar.

L'algorisme fa ús de l'anàlisi de l'anomenat *big data* per a acumular i analitzar informació de molts casos, i processar-los per a extreure'n pautes o criteris relatius a les circumstàncies o característiques dels subjectes que, suposadament, guarden relació amb la probabilitat de cometre nous actes violents o delictius. Una vegada proporcionada la informació relativa al subjecte individual en qüestió, l'algorisme aplica aquests criteris per a oferir una estimació del seu suposat grau de perillositat, i aquesta informació és tinguda en compte pels tribunals per a la decisió sobre la pena a imposar.

Fa un temps, el sistema COMPAS va tenir una certa repercussió mediàtica per la notícia del cas d'Eric L. Loomis<sup>9</sup>, el qual va ser condemnat a una pena de 6 anys de presó (inusitadament elevada) per fugir amb cotxe de la policia, a causa que l'algorisme va establir un alt grau en la probabilitat de reincidència delictiva. La controvèrsia entorn de COMPAS s'ha centrat fonamentalment en dos punts: a) en la falta de transparència que implica el fet que el *software* és propietat d'una empresa privada (Northpointe Inc.), la qual cosa implica que el codi és secret i no pot ser objecte d'anàlisi o auditoria externa; i b) per les acusacions d'estar esbiaixat i discriminar certs col·lectius com ara els negres<sup>10</sup>, acusacions que l'empresa ha intentat rebatre però les conclusions sobre el caràcter esbiaixat de la qual són gairebé impossibles d'acreditar sense una anàlisi del codi.

Encara que aquestes són qüestions rellevants, en si no afecten el que és la valoració d'un sistema com COMPAS com a exemple d'aplicació de l'IA en la decisió judicial. Més enllà dels possibles problemes de disseny de l'algorisme model, si partim del fet que el sistema legal estableix que cal atendre els criteris de perillositat per a determinar la pena, una tecnologia d'aquest tipus pot resultar més fiable i eficaç que fer dependre la valoració de la perillositat i el risc de reincidència en l'experiència personal del jutge i en el seu propi criteri o estimacions, ja que permet manejar una quantitat d'informació molt major i extreure'n conclusions fonamentades i sense risc de ser afectades pels biaixos o inclinacions que, de manera conscient o inconscient, poden afectar el judici de qui jutja.

8. Vegeu una breu descripció d'aquest a <https://www.theatlantic.com/technology/archive/2018/01/equivant-compas-algorithm/550646/> [Data de consulta: 26 de juny de 2019].

9. Vegeu informació sobre el cas, per exemple, a <https://www.nytimes.com/2017/05/01/us/politics/sent-to-prison-by-a-software-programs-secret-algorithms.html> [Data de consulta: 26 de juny de 2019].

10. Vegeu <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> [Data de consulta: 26 de juny de 2019].



Malgrat tot, cal tenir en compte que estem parlant d'un aspecte molt menor en comparació del conjunt d'activitats i tasques que implica una decisió judicial. Encara que l'algorisme permeti obtenir conclusions precises i objectivables sobre el grau de perillositat i la probabilitat de reincidència de l'acusat, no té cap incidència en pràcticament cap dels àmbits principals de la decisió judicial: ni serveix per seleccionar material normatiu rellevant, ni per a la interpretació jurídica, ni per a la sistematització del dret, ni, estrictament parlant, per a la prova o qualificació jurídica dels fets, a pesar que tingui un impacte en una premissa fàctica que serveix per fonamentar la decisió. Per tant, més que concebre-ho com una aplicació de la intel·ligència judicial a la decisió judicial, s'hauria d'entendre com una eina tecnològica més de suport al decisor, com poden ser les bases de dades jurídiques.

### iii. El *chatbot DoNotPay*<sup>11</sup>

DoNotPay és el nom donat per Joshua Browder, estudiant de Dret de la Universitat d'Stanford, a un algorisme creat per ell de tipus *chatbot*<sup>12</sup>, que en els seus inicis estava enfocat a la impugnació de sancions de tràfic (com ara multes d'aparcament, excés de velocitat, semàfors...) i que és capaç de generar escrits d'impugnació que l'usuari simplement ha d'imprimir i signar. Segons fonts del propi creador, va aconseguir recuperar més de tres milions de dòlars mitjançant els recursos presentats gràcies a la seva tecnologia. Recentment, el seu creador ha treballat en l'ampliació de l'àmbit d'actuació del seu *software*, cobrint altres camps com ara les reclamacions a companyies aèries (cancel·lació de vols, *overbooking*, pèrdua d'equipatge...) o la defensa del consumidor, entre altres. Mitjançant una interfície simple amb un quadre d'entrada de text, l'usuari introdueix la seva petició o el seu problema (per exemple «he rebut una multa d'aparcament») i l'algorisme el va guiant per una sèrie de passos (sol·licitant informació sobre els fets del cas) per acabar amb la generació automàtica d'un escrit d'impugnació (suposadament basat en dret).

Estrictament parlant, aquesta tecnologia no estaria enfocada a la decisió judicial automatitzada, sinó més aviat a allò que es podria entendre com un «advocat virtual» o robot *lawyer*, ja que la seva finalitat és protegir els interessos de la part. Amb tot, sembla evident que per al seu funcionament requereix la realització de certes activitats vinculades a l'aplicació del dret. Així, sembla clar que necessita comptar amb una base de dades de fonts normatives (selecció de material jurídic rellevant) i que ha de ser capaç de realitzar tasques de sistematització per a obtenir les conseqüències jurídiques corresponents als diferents supòsits fàctics, ja que solament d'aquesta manera és possible fonamentar jurídicament la pretensió. És dubtós que l'algorisme dugui a terme una activitat d'interpretació jurídica pròpiament dita, atès que això requereix una complexitat tècnica i de càlcul molt considerable. El més probable és que la pròpia programació contingui el material ja interpretat (vinculant supòsits de fet amb conseqüències jurídiques), sobretot tenint en compte que es tracta d'àmbits la regulació dels quals sol ser força precisa i sense massa marge per a la interpretació. Tampoc sembla que el programa dugui a terme cap tasca vinculada a la prova dels fets, ja que es basa totalment en les dades proporcionades a instància de part.

No obstant això, i a pesar de les seves limitacions tant en les seves funcions com en el seu àmbit, DoNotPay és un bon exemple de com en un futur proper o, fins i tot, en el present la intel·ligència artificial pot resultar molt útil per a decidir sobre qüestions de (relativament) escassa complexitat tècnica des del punt de vista jurídic i d'escassa quantia, agilitant enormement aquests processos.

---

11. Vegeu una descripció general a <https://www.theverge.com/2017/7/12/15960080/chatbot-ai-legal-donotpay-us-uk> [Data de consulta: 26 de juny de 2019].

12. Un *chatbot* és un algorisme d'intel·ligència artificial que permet a un usuari l'entrada de peticions per mitjà del llenguatge natural (oral o escrit) i intenta oferir una resposta adequada. Cada vegada és més gran la seva implantació a les empreses per a les tasques d'atenció al client.

#### iv. La intel·ligència artificial aplicada al llenguatge natural (i jurídic): *Watson/ROSS Intelligence*

La companyia IBM, una de les grans empreses tecnològiques del planeta, fa anys que investiga i desenvolupa en l'àmbit del que es denomina 'computació cognitiva', tecnologia que intenta emular la manera de raonar i interactuar dels éssers humans i en la qual ocupa un lloc central la «comprensió» (sense entrar aquí en debats filosòfics) i la comunicació per mitjà del llenguatge natural. Com a resultat dels seus desenvolupaments va crear Watson, una eina capaç d'«entendre» el llenguatge natural i amb capacitat d'«aprenentatge» a partir de tota la informació que constantment va recopilant i processant. Watson va tenir repercussió mediàtica l'any 2011 per resultar vencedor en el concurs nord-americà de preguntes Jeopardy!, en el qual es va enfrontar als dos concursants humans amb major nombre de victòries fins avui<sup>13</sup>. El més destacable no és tant l'enorme capacitat d'acumulació d'informació i la seva velocitat de processament, sinó el fet que, a mesura que avançava el concurs, Watson va poder entendre correctament, fins i tot, els dobles sentits, els girs lingüístics, les metàfores o la ironia que formen part habitual dels nostres llenguatges naturals, amb la qual cosa es pot dir que compta amb una capacitat notable de comprensió i interpretació lingüística.

Més recentment, una companyia anomenada ROSS Intelligence ha fet ús de la tecnologia de Watson per a aplicar-la a l'àmbit jurídic<sup>14</sup>. La idea principal és usar-la per a processar la ingent quantitat de material jurídic que constantment es genera (legislació, jurisprudència, documents contractuals, etc.) per a comptar no solament amb la màxima i més actualitzada informació, sinó també (i especialment) per a discriminar i seleccionar la més adequada i rellevant per al problema que s'estigui tractant a cada moment, estalviant així grans quantitats de temps i esforç als operadors jurídics. Segons s'afirma, aquesta tecnologia és capaç de processar més de mil milions (*one billion*) de documents legals per segon, redactats en llenguatge natural, i retornar amb molta precisió resultats a les qüestions plantejades (també en llenguatge natural, tal com es formularien a un advocat). A més, com que es tracta d'una tecnologia de *machine learning*, va millorant constantment. Però malgrat aquestes dades sorprenents, aquesta tecnologia no està pensada, almenys de moment, per a substituir els jutges o advocats, sinó més aviat com una eina de suport a la seva activitat, sobretot la de buscar i seleccionar informació jurídica rellevant. Segons el propi CEO de ROSS Intelligence, tasques com ara sospesar la informació, crear arguments jurídics o preparar documents segueixen sent exclusives dels juristes humans.

Aquest últim exemple de tecnologia aplicada a l'àmbit jurídic és especialment interessant perquè guarda una estreta relació amb una de les activitats que *a priori* semblarien més difícils d'automatitzar, com és la interpretació jurídica o la conversió d'expressions redactades en llenguatge natural a una sèrie de paràmetres que permetin una sistematització en termes de correlació de casos amb solucions jurídiques, sense necessitat d'una intervenció humana que determini el significat d'aquests enunciats normatius. Amb tot, com mostra el debat teòric sobre la interpretació jurídica, sembla que aquesta va molt més allà de simplement «entendre» o «captar» el significat clar i unívoc dels enunciats dictats per les autoritats normatives, ja que en molts casos sembla involucrar una presa de decisions o raonaments i arguments que els donen fonament, com en molts dels anomenats «arguments interpretatius» (argument analògic, a fortiori, teleològic, psicològic, sistemàtic, apagògic, etc.) que són habitualment usats en el raonament jurídic. Però, malgrat això, resulten innegables l'abast i la importància dels desenvolupaments tecnològics vinculats a la interpretació del llenguatge natural.

A partir dels exemples exposats en aquest apartat, crec que es pot concloure que, fins avui, tots els avenços tecnològics han de ser concebuts com a eines parcials d'ajuda a la decisió i que encara estem lluny del que es podria considerar com un sistema integral automatitzat de presa de decisions jurídiques (excepte, potser, en àmbits molt concrets i específics en els quals totes les qüestions jurídiques i fàctiques estiguin perfectament delimitades). Alguns aspectes relacionats amb l'aplicació del dret semblen especialment problemàtics, com ara la

13. Vegeu un reportatge periodístic sobre aquesta notícia a <http://www.nytimes.com/2011/02/17/science/17jeopardy-watson.html?pagewanted=all> [Data de consulta: 26 de juny de 2019].

14. Vegeu <https://www.ibm.com/blogs/watson/2016/01/ross-and-watson-tackle-the-law/> [Data de consulta: 26 de juny de 2019].



interpretació jurídica (malgrat els importants avenços en computació cognitiva), la prova dels fets o la qualificació d'aquests en categories jurídiques. Però no és menys cert que els avenços en intel·ligència artificial són ràpids i significatius, i que cal esperar que els seus desenvolupaments encara siguin més grans en un futur, de manera que no es pot descartar en absolut la idea d'un jutge artificial completament autònom.

## IV

Els avenços tecnològics en general i de l'IA en concret no semblen ser una opció, sinó senzillament alguna cosa que està ocorrent i que tindrà un impacte cada vegada més gran. Per això, en lloc de plantejar-nos si s'haurien d'aplicar o no aquests avenços a l'àmbit de la decisió jurídica, potser sigui millor reflexionar sobre els possibles aspectes positius i negatius d'aquests, i també fer propostes sobre cap a on es podrien orientar perquè els resultats siguin més satisfactoris o, almenys, per a evitar possibles perills i problemes.

Com a punt de partida, partiré de la base que la decisió judicial és una activitat primordialment tècnica o instrumental, en el sentit que, sobretot a partir de la Il·lustració i en el context occidental, es vincula més estretament a la idea d'aplicar el *dret* o *decidir conforme a dret*, que a la idea d'impartir *justícia*. Es tracta d'utilitzar les eines, els recursos i els procediments establerts pel sistema jurídic per a donar una resposta que, almenys en principi o teòricament, està predeterminada pel propi sistema. El jutge, en aquest context, es configura com un especialista amb una formació tècnica adequada per a l'obtenció d'aquesta resposta i que, a més, l'ha d'obtenir mitjançant els processos i criteris jurídicament establerts.

Els avenços tecnològics sempre s'han revelat útils per a la realització de tasques instrumentals, millorant aspectes com ara l'eficiència, la precisió i la rapidesa. Per això, almenys *a priori* semblaria que l'activitat d'aplicació del dret es podria veure beneficiada per la implementació d'aquestes noves tecnologies. Ara bé, ja fa força temps que per múltiples raons es va abandonar la concepció formalista que concebia l'activitat judicial o d'aplicació del dret en general com alguna cosa mecànica, en què el decisor seria una mena d'autòmat. Fins i tot les concepcions contemporànies que, com en el cas de Ronald Dworkin<sup>15</sup>, defensen la idea d'una única resposta correcta i l'activitat judicial com a primordialment epistèmica o de descobriment, assumeixen que en la decisió judicial hi intervenen molts aspectes relacionats amb la capacitat de judici i de reflexió, principalment en àmbits com el de la interpretació jurídica. Això planteja alguns problemes i dificultats tant per a la possibilitat efectiva de crear un model automatitzat de decisió judicial, com per a les conseqüències que aquest podria tenir en aquesta activitat. Però ja em referiré a aquestes qüestions més endavant. De moment, assenyalaré alguns aspectes, al meu entendre positius, que *prima facie* comportaria la decisió judicial automatitzada:

*Imparcialitat.* Una primera conseqüència evident és que les decisions estarien lliures dels biaixos que, fins i tot de manera inconscient i reconeixent el gran esforç realitzat per la immensa majoria de jutges per a mantenir la seva imparcialitat, poden afectar els decisors humans, en funció de la seva ideologia, experiència, creences, idees preconcebudes, entorn o context cultural, circumstàncies personals, etc. Un algorisme informàtic és simplement incapaç de comportar-se de manera diferent a com predetermina la seva programació, de manera que es deslliuraria de tots aquests factors. Amb tot, cal tenir en compte un aspecte molt important, que és la possibilitat de *biaixos en la programació de l'algorisme*. Encara que la tecnologia per si mateixa sigui incapaç d'introduir biaixos, aquests poden estar continguts en la programació, ja que els programadors (almenys de moment) són éssers humans. Per això, per a evitar situacions com la controvèrsia relacionada amb l'algorisme COMPAS, seria fonamental el requisit de la transparència. El codi dels algorismes no hauria de ser secret, sinó que hauria de ser públic i susceptible de supervisió i anàlisi per les autoritats corresponents, la qual cosa sens dubte tindria importants repercussions en l'àmbit empresarial, ja que les empreses desenvolupadores de *software* probablement tindrien importants interessos a mantenir el seu codi en secret com a manera d'ocupar una posició avantatjosa enfront de la competència, però en

---

15. Vegeu, per exemple, DWORKIN, R. (1977). *Taking Rights Seriously*. Cambridge (Mass.): Harvard University Press.

aquest punt hauria de prevaler l'interès públic i caldria buscar formes alternatives de fer que aquest àmbit seguís sent atractiu per a les companyies del sector.

*Consistència.* Un altre dels aspectes en què un model de decisió automatitzada tindria un impacte important és el de reduir la variabilitat i la multiplicitat de criteris, interpretacions, etc. en les decisions vinculades a la pròpia pluralitat dels òrgans decisors (jutges i tribunals). Diferents òrgans poden interpretar de manera divergent els mateixos preceptes legals o valorar de manera diferent uns mateixos fets, donant lloc a decisions diferents, fins i tot, en supòsits que comparteixen les mateixes circumstàncies rellevants. L'ús d'algorismes de decisió garantiria una mateixa interpretació i uns mateixos criteris de decisió i, per tant, una major consistència en la «jurisprudència» (si és que es pot utilitzar aquest terme en aquest context), amb conseqüències molt significatives en la previsibilitat de les decisions i la seguretat jurídica. A més, gràcies a l'enorme capacitat d'acumulació i processament de la informació, pràcticament s'eliminaria la possibilitat d'error derivat de no tenir en compte alguna normativa aplicable o algun element jurídicament rellevant.

*Previsibilitat.* Com s'ha comentat en el paràgraf anterior, l'ús d'algorismes garantiria la consistència en les interpretacions i en els criteris de decisió, permetent un alt grau de previsibilitat de la decisió i, amb això, de la seguretat jurídica. En principi, seria possible conèixer amb un alt nivell de precisió les conseqüències jurídiques dels nostres comportaments, fins i tot sense necessitat d'iniciar un procés judicial, de manera que milloraria el grau d'autonomia personal perquè es comptaria amb millors criteris sobre les conseqüències de cada alternativa d'acció, ajudant així a la presa de decisions de l'agent.

D'altra banda, no obstant això, no s'hauria d'oblidar o menysvalorar altres aspectes que plantegen seriosos problemes, tant en la possibilitat efectiva de la decisió judicial automatitzada com en la seva desitjabilitat, fins i tot en el cas de ser tècnicament possible.

- a. Com s'ha posat de manifest anteriorment, sembla que hi ha un ampli consens en la teoria jurídica sobre la inadequació d'un esquema purament «mecanicista» de l'aplicació del dret, i els obstacles perquè un model formalista o mecanicista funcioni tenen a veure primordialment amb certes característiques dels propis sistemes jurídics. Un aspecte important és que el llenguatge jurídic en general, encara que sigui de caràcter tècnic, no deixa de ser un llenguatge natural (en contraposició amb els llenguatges artificials), per la qual cosa està afectat per problemes com ara la vaguetat i l'ambigüitat, que d'alguna manera impliquen un cert grau d'indeterminació i la necessitat d'eleccions o decisions per part de l'interpret. Aquest fet no solament dificulta tècnicament l'elaboració d'un model de decisió automatitzada, atès que els llenguatges de programació algorítmica són artificials, sinó que, en cas d'establir o fixar criteris interpretatius per a enfrontar-se a aquestes dificultats, d'alguna manera «cristal·litzarien» o «solidificarien» certes interpretacions, rebutjant-ne altres igualment legítimes, i, a més, significaria la prevalença del criteri adoptat pels programadors per sobre del d'altres autoritats jurídiques.
- b. En segon lloc, és habitual que, en certs casos, el legislador persegueixi conscientment que hi hagi un cert grau d'indeterminació en la resposta jurídica, o l'atribució d'una certa capacitat de decisió discrecional al decisor. Així ocorre, per exemple, quan els preceptes legals fan referència o incorporen conceptes valoratius o de tipus moral, o conceptes essencialment controvertits, o deleguen directament al jutge la capacitat de decidir les mesures més adequades a prendre (dins d'un marge). De nou, els algorismes podrien incorporar certs paràmetres per a determinar les decisions en aquests casos, però al preu d'imposar una determinada concepció i d'eliminar tota possibilitat de deliberació o argumentació moral.
- c. En tercer lloc, tampoc no es pot descartar la possibilitat d'«atacs» al sistema (*hacking*), com ocorre amb pràcticament tots els sistemes informàtics. El tema de la seguretat informàtica és una carrera contínua entre els qui intenten protegir els sistemes i els qui intenten superar els obstacles i barreres que els protegeixen. Però encara que la seguretat total no existeix, crec que tampoc caldria sobredimensionar el problema o els riscos. Totes les companyies i institucions pateixen atacs constants als seus sistemes informàtics, però la immensa majoria són bloquejats, i en les poques situacions en què els hackers aconseguen burlar les mesures de seguretat, el més habitual és que es reaccioni amb rapidesa i s'aconsegueixin contenir els danys. En certa manera, es podria fer un paral·lelisme amb els casos de corrupció: encara que sempre hi ha casos de jutges corruptes, el més habitual és que siguin molt minoritaris (en

comparació del volum total de jutges i decisors) i que els seus efectes no siguin devastadors. És més, la probabilitat de *hackear* un sistema informàtic de decisió judicial segurament és inferior a la que un jutge sigui corrupte.

D'altra banda, no s'hauria d'oblidar que la irrupció dels avenços tecnològics aplicats a certs àmbits o activitats, algunes vegades, ha provocat o ha contribuït a la pròpia configuració o transformació d'aquestes activitats, i el dret no seria aliè a això<sup>16</sup>. De manera purament especulativa, crec que els canvis podrien prendre dues direccions diferents, encara que hi hauria la possibilitat d'una combinació d'ambdues:

D'una banda, hi ha la possibilitat que la producció normativa es vagi ajustant gradualment a les «necessitats» de l'aplicació automatitzada del dret, en el sentit de fer més senzilla l'adaptació a aquestes tecnologies i que el seu funcionament pugui resultar menys problemàtic. Alguns d'aquests canvis avui ja es poden entreveure, com ocorre amb els anomenats *smart contracts*<sup>17</sup>, que fan ús d'una tecnologia denominada *blockchain* o cadena de blocs, que és pràcticament *inhackeable* per a registrar i fer complir automàticament acords contractuals, la qual cosa pot tenir en el futur un efecte importantíssim, per exemple, en la disminució de la litigiositat associada als contractes, ja que elimina completament la necessitat de prova respecte de l'existència i contingut de la relació contractual.

D'altra banda, una altra manera d'afrontar les dificultats abans assenyalades seria procedint a una cada vegada més marcada diferenciació entre els àmbits o assumptes objecte de decisió artificial i aquells altres que hagin de ser resolts pels éssers humans. D'aquesta manera, les qüestions que puguin ser resoltes mitjançant criteris clars i objectius (probablement la majoria) i que no afectin qüestions de gran transcendència moral (com serien, per exemple, els drets fonamentals, decisions relatives a menors o persones vulnerables, etc.), serien objecte de decisió artificial, mentre que altres en què es consideri important comptar amb la capacitat de judici i deliberació, es reservarien a decisors humans. Fins i tot seria possible un model mixt en el qual la intervenció humana es limités a revisar les decisions prèvies o les propostes de decisió ofertes pels algorismes. En resum, com ha estat la tendència invariable al llarg de la història, es deixaria per a les màquines el que les màquines poden realitzar de manera més eficient i es reservaria als éssers humans allò que les màquina no són capaces de fer o realitzar de manera satisfactòria, i que són de fet les més importants.

---

16. No seria ni molt menys el primer cas. A vegades, una tecnologia ha significat, fins i tot, l'aparició de nous àmbits d'activitat com, per exemple, el cinema o els videojocs com a expressions o manifestacions artístiques. En altres casos ha representat canvis importants en camps que ja existien prèviament, com va ocórrer per exemple amb l'arquitectura gràcies a l'ús de nous materials en l'anomenada 'arquitectura del ferro' del segle XIX, o en l'aparició de nous instruments i tecnologies en la música, que ha incidit de manera destacable, inclús, en el tipus de música que es crea.

17. Vegeu una breu introducció no tècnica dels anomenats *smart contracts* o contractes intel·ligents a [https://retina.elpais.com/retina/2017/12/22/tendencias/1513937575\\_114270.html](https://retina.elpais.com/retina/2017/12/22/tendencias/1513937575_114270.html) [Data de consulta: 26 de juny de 2019].

**Citació recomanada:** MARTÍNEZ ZORRILLA, David. El jutge artificial: propera parada? *Oikonomics* [en línia]. Novembre 2019, n. 12, pp. 1-12. ISSN: 2339-9546. DOI: <https://doi.org/10.7238/o.n12.1914>

---



**David Martínez Zorrilla**

**Professor agregat dels Estudis de Dret i Ciències Polítiques (UOC)**

[dmartinezz@uoc.edu](mailto:dmartinezz@uoc.edu)

David Martínez (Manresa, 1974) és Llicenciat en Dret amb Premi Extraordinari (1997) per la Universitat Autònoma de Barcelona, i Doctor en Dret (2004) per la Universitat Pompeu Fabra. Actualment és professor agregat dels Estudis de Dret i Ciència Política de la Universitat Oberta de Catalunya, i professor associat a l'escola de negocis ESERP, especialitzat en l'àmbit de la teoria general i la filosofia del dret. És autor de diverses monografies i articles acadèmics en publicacions espanyoles i estrangeres, entre les quals destaquen *Conflictes constitucionals, ponderació i indeterminació normativa* (Marcial Pons, 2007); *Metodologia jurídica i argumentació* (Marcial Pons, 2010); «The Structure of Conflicts of Fundamental Legal Rights», a *Law & Philosophy*, vol. 30 (2011); «Constitutional Dilemmas and Balancing. Some Comments on Lorenzo Zucca's Analysis», a *Ràtio Juris*, vol. 24 (2011); «Some Thoughts about the Limits of Alexy's Conception of Principles and Balancing», a Duarte, D. i Silva Sampaio, J. (ed.), *Proportionality in Law. An Analytical Perspective* (Springer, 2018).

Els textos publicats en aquesta revista estan subjectes –llevat que s'indiqui el contrari– a una llicència de Reconeixement 4.0 Internacional de Creative Commons. Podeu copiar-los, distribuir-los, comunicar-los públicament i fer-ne obres derivades sempre que reconegueu els crèdits de les obres (autoria, nom de la revista, institució editora) de la manera especificada pels autors o per la revista. La llicència completa es pot consultar a <https://creativecommons.org/licenses/by/4.0/deed.ca>.

