

CONGRESO “DERECHOS DIGITALES E INTELIGENCIA ARTIFICIAL”

Palacio de la Aljafería, sede de las Cortes de Aragón

Jueves 2 de junio de 2022

Derechos y garantías ante la automatización policial y de la justicia

Pere Simón Castellano

Profesor Titular de Derecho Constitucional

Universidad Internacional de la Rioja UNIR

Secretario de la Sección de Derecho de las Tecnologías de la Información
y la Comunicación (ICAB)



Pere Simón Castellano

Justicia cautelar e inteligencia artificial

LA ALTERNATIVA A LOS ATÁVICOS
HEURÍSTICOS JUDICIALES

PRÓLOGO
VICENTE MAGRO SERVET

JIB
BOSCH EDITOR

MONOGRAFÍAS

La obra aborda el binomio oportunidad desafío que conlleva el empleo de las más modernas tecnologías en los centros penitenciarios. El autor parte de la premisa que entiende la imparable y ambivalente transformación digital, que llegará más temprano que tarde a nuestras prisiones, centrando su atención en los interrogantes relativos a la determinación de qué técnica, para qué fines y, lo más importante, el cómo implementarlas y delimitarlas con pleno respeto a los derechos de los reclusos.

El lector se encuentra ante un trabajo de investigación que explora los usos plausibles de los sistemas de inteligencia artificial en la fase de ejecución penal, que parte de un análisis minucioso de las fortalezas y debilidades de un sistema penitenciario que proyecta aún disonancias fruto de la dicotomía entre teoría y práctica y de la latente pulsión al castigo y retribución de nuestro modelo, ahora dominado por los sistemas actuariales y la necesidad de actuar ex ante delictum, frente a los parcialmente denostados fines constitucionales de inserción y reeducación.

La monografía cuenta con el prólogo, proemio y epílogo de voces contrastadas como las de Esteban Mestre Delgado, Bernardo Feijoo Sánchez y Antonio Andrés Pueyo, que respectivamente, y como reacción a la lectura de esta, entablan un debate con las ideas del autor, elevando en definitiva el nivel de la obra, y sentando las bases para una discusión futura a la que se anima expresamente a participar a los operadores jurídicos y a la doctrina científica más especializada.

PERE SIMÓN CASTELLANO

LA PRISIÓN ALGORÍTMICA

monografías

ALTA CALIDAD EN
INVESTIGACIÓN
JURÍDICA



tirant
lo blanch

+Lectura
GRATIS
en la nube

PERE SIMÓN CASTELLANO

LA PRISIÓN ALGORÍTMICA

(Previa) Qué tecnología y para qué.

- Inteligencia Artificial.

Sistemas expertos; redes neuronales artificiales; robots; sistemas que actúan racionalmente; aprendizaje automático; etc.

Dentro de la IJA encontramos una subcategoría muy específica, que son los sistemas expertos de valoración del riesgo, en adelante, IAJVR.

		Humanos	Racionales
Cognitivo	Sistemas que Piensan como	... lograr que las computadoras "piensen", "máquinas con mente"... (Haugeland 85)	El estudio de las facultades mentales mediante el uso de modelos computacionales (Charniak, McDermot 85)
Conductista	Sistemas que Actúan como	...cómo lograr que las computadoras realicen tareas que, por el momento, los humanos hace mejor (Rich, Knight 91)	La parte de la ciencia de la computación que se ocupa de la automatización de la conducta inteligente (Luger, Stubblefield 93)

Ejemplos del estado del arte

- <https://twitter.com/DaniNovarama/status/1516690248050298882>
- https://twitter.com/_LucasRizzotto/status/1516206578709397504

(Previa) Tecnología en movimiento.

- Cómo funcionan los procesos de desarrollo tecnológico.
- La programación no es estática.
- Deploys, trabajo en dev, equipos de QA...

- Exigen garantías específicas de seguridad de la información.
- Trazabilidad -*audit trails*-, auditabilidad...

Formulando hipótesis

<1>

Los sistemas de IAJVR pueden contribuir a objetivar la toma de decisiones judiciales difíciles vinculadas al cálculo de los niveles residuales de un riesgo determinado (fines policiales y justicia predictiva).

- Prisión preventiva
 - Fuga
 - Reiteración delictiva
 - Destrucción de pruebas
- Ámbito ejecución penal (clasificación penitenciaria, PPR, libertad vigilada, permisos, etc.)



Atajos intuitivos y heurísticos judiciales

- Representatividad: experiencias previas, lo que nos ha marcado.
- Accesibilidad o disponibilidad: más cercano, cómodo, placentero.
- Anclaje o ajuste: contexto, sentidos, marco conceptual.
- Afección: emociones.

Heurísticos judiciales... valoración ambivalente: sesgos que pueden mejorar la equidad medible estadísticamente.

VIEJO CONOCIDO... exigir a los algoritmos que actúen de forma racional cuando nosotros no lo hacemos... exigir más garantías a la firma digital que a la analógica, etc.

Formulando hipótesis

<2>

Los algoritmos perpetúan la desigualdad y discriminan

Versus

La técnica es neutra. El debate regulatorio futuro debería centrarse en los usos y en quién y cómo diseña el algoritmo, así como en la definición de las garantías

Los algoritmos y el fin del mundo

Cómo los algoritmos perpetúan la desigualdad en España

<https://www.lavanguardia.com/tecnologia/20211026/7814332/como-algoritmos-perpetuan-desigualdad-espana.html>

Un algoritmo impreciso condiciona la libertad de los presos

<https://www.lavanguardia.com/vida/20211206/7888727/algoritmo-sirve-denegar-permisos-presos-pese-fallos.html>

Cómo los algoritmos programan la desigualdad

https://retina.elpais.com/retina/2021/11/09/eventos/1636475783_231928.html



Una propuesta más razonable

Algoritmos, desigualdad y otros mitos malintencionados

https://cincodias.elpais.com/cincodias/2021/11/15/legal/1636991256_580299.html

El debate en torno a las garantías jurídicas

Propuesta europea Reglamento IA

- Necesaria identificación y gestión de riesgos a lo largo de toda la vida útil del sistema.
- Alta calidad de las bases de datos utilizadas para nutrir a los algoritmos, evitando así sesgos y discriminaciones.
- Mantenimiento de registros automático de eventos durante la operación del sistema, garantizando la trazabilidad.
- Transparencia, entendida en el sentido de permitir que el usuario interprete los resultados del sistema y los utilice apropiadamente.
- Supervisión humana efectiva durante todo el periodo de utilización del sistema.
- Nivel adecuado de precisión y seguridad cibernética.

El debate en torno a las garantías jurídicas

<3>

En contra de la transparencia algorítmica

A favor de una reinterpretación

En contra...

- Inidónea.
- No aporta nada.
- Falsa sensación de seguridad: puede parecer que es una garantía cuando en realidad no actúa como tal y no aporta nada.

Cuando hablamos de usos de IA los juristas nos referimos habitualmente a la transparencia. Resulta necesario reinterpretar la publicidad de los algoritmos, pues no es un fin en si mismo, y en el estado actual la transparencia únicamente contribuye a crear una falsa sensación de seguridad.

Sobre la transparencia...

Analogía. Sucede como con la publicidad del proceso penal, que no es un fin en sí mismo, sino un vehículo para conseguir un fin mayor, que es la consecución de un juicio justo. Otro debate es el grado de publicidad, puesto que no es lo mismo acceder como público al juicio, que leer las crónicas judiciales en prensa que seguir un juicio por *streaming*.

En realidad, la transparencia del algoritmo (del código) es innecesaria por inidónea, no aporta nada ni consigue ningún fin, y resulta también inaceptable si se defienden las posibilidades de auditabilidad *a posteriori* (no es la vía, por el gap o brecha económica).

Algorithmic Transparency Standard (UK Central Digital and Data Office)

Publicada el 29 de noviembre de 2021. No es preceptivo salvo para sector público.

a) Deben explicar funcionamiento y el problema que trata de resolver.

b) Detalles técnicos:

1. Propietario (responsable u *owner* de proceso, nada relativo a protección de datos)

2. Alcance de la herramienta (diseñada para "X", mantenimiento, arquitectura del sistema, tipo de modelo, uso recurrente en el tiempo)

3. Cómo la herramienta afecta a la toma de decisiones (HIC, HOB o HM) y la participación humana vía *audits* y revisión y mejora continua.

4. Datos personales (descripción de categorías, bases de legitimación, cómo han sido recabados, DPA - acuerdos con corresponsables y encargados-, quién tiene acceso y plazos de conservación)

5. Evaluaciones de impacto (IA ética, IA y desigualdad, IA y proporcionalidad, EIPD).

6. Descripción de riesgos (discriminación, sesgos y daños o perjuicios)

7. Controles previstos y en marcha para mitigar los riesgos identificados.

Algorithmic Transparency Standard (UK Central Digital and Data Office)

No es transparencia del código *stricto sensu*. Es transparencia relativa a las garantías concretas conectadas con usos de la IA considerados de riesgo alto.

Enfoque de riesgo propio de sistemas de Compliance, que va mucho más allá de la mera transparencia.

Garantías interconectadas: "publicidad" vinculada a la explicabilidad "relativa", análisis y definición de riesgos *ex ante*, trabajar en base a controles preventivos, detectivos y correctivos (seguridad de la información, protección de datos y *compliance*), auditabilidad, trazabilidad -rastros participación humana-), etc.

El debate en torno a las garantías jurídicas

<4>

La **explicabilidad** como garantía *relativa*.

Cuándo hay que exigir explicaciones... y la necesidad de combatir las explicaciones limitadas -humanas- que crean una falsa sensación de seguridad.

Determinismo y razonamiento limitado

Los algoritmos de IA son perfectamente explicables, incluso los que evolucionan mediante experiencia y aprendizaje, en un momento dado son perfectamente deterministas y, por tanto, se pueden explicar de manera inequívoca.

Se explican en unos términos lógico-matemáticos que no siguen el esquema simbólico-lingüístico que usamos en el razonamiento consciente los humanos y necesitamos tiempo para descifrar y comprender las razones que nos advierten del “movimiento correcto”.

Ajedrez



Artificial Intelligence. A modern approach

Stuart Russell y Peter Norvig nos hablan de la “falsa sensación de seguridad”.

Qué pasa si sustituimos un árbol de decisión por una red neuronal...

Una red neuronal -no explicable hoy, aunque quizás dentro de unos años- es mucho más potente, rica y efectiva que un árbol de decisión.

Creemos que es más seguro escoger un algoritmo más sencillo, que imita nuestro razonamiento, porque es explicable, pero entramos en una zona de confort que crea una falsa sensación de seguridad y que nos condena a reproducir nuestros sesgos y prejuicios.

Sobre la explicabilidad...

¿De qué nos sirve poder explicar cómo funciona un código algorítmico? ¿Significa eso que éste funciona mejor? Probablemente no, únicamente indica que está limitado a los tradicionales árboles de decisión y a las falibles fórmulas decisorias humanas.



En la terraza del bar



Usos IAJVR

Sistemas expertos (limitados) son muy útiles para fines detectivos y correctivos, como en el caso Amazon, y por ende, terapéuticos.

Para la valoración y cálculo de riesgo están limitados en el diseño si se limitan a funcionar como un árbol de decisión, más propio del razonamiento limitado humano.

¿Y si... vamos más allá, aceptando el uso de lo más seguro?

Ajedrez



El debate en torno a las garantías jurídicas

<5>

Trazabilidad. Auditabilidad. Seguridad de la información.
Enfoque de riesgo y evaluaciones de riesgo -éticos,
discriminación, proporcionalidad, protección de datos,
etc.-.

El debate en torno a las garantías jurídicas

<6>

Las garantías subjetivas en y desde el diseño.

Cómo - *La participación humana.*

Quién *debe participar en el diseño, supervisar la implementación y garantizar la vigencia del modelo.*

El debate en torno a las garantías jurídicas

<7>

El derecho a la segunda oportunidad.

Derecho al recurso, a la defensa, a pedir una segunda opinión.

No todos los sistemas de IA o IAJVR toman decisiones automatizadas con efectos jurídicos.

Debate sobre los usos. Ejemplos: fines detectives, de discovery, entre muchos otros.

El caso de la prisión preventiva.

La predictibilidad del juicio jurisdiccional y las cifras de litigiosidad en España.

Fines policiales.

A modo de conclusión

<1> Sistemas expertos IAJVR deberían ser empleados con finés detectivos y correctivos en sede judicial. A la búsqueda del sesgo desconocido, y dar luz a realidades incómodas. Se trata de sistemas que emulan a los humanos y multiplican sus efectos, facilitando la detección.

<2> La transparencia algorítmica es inidónea y crea una falsa sensación de seguridad. En su lugar y una vez descartadas las decisiones automatizadas en la justicia (HIC), hay que garantizar la transparencia y explicabilidad relativa (en función de los fines y relacionada con extremos concretos), la trazabilidad, el enfoque de riesgo y el trabajo en base a MTO propias de SGSI y RGPD.

<3> ¿Podemos confiar en aquello más potente, rico y eficaz, incluso cuando no lo podemos explicar? Sí, aunque...

Trabajos relacionados

- BUENO DE MATA, Federico. (2020). « Macrodatos, inteligencia artificial y proceso: luces y sombras». Revista General de Derecho Procesal, n. 51.
- COTINO HUESO, L. (2020). «SyRI, ¿a quién sanciono? Garantías frente al uso de inteligencia artificial y decisiones automatizadas en el sector público y la sentencia holandesa de febrero de 2020». La Ley Privacidad, Wolters Kluwer, núm. 4.
- MIRÓ LLINARES, F. (2018). «Inteligencia artificial y justicia penal: más allá de los resultados lesivos causados por robots». Revista de Derecho Penal y Criminología, núm. 20, 2018, págs. 87-130.
- MIRÓ LLINARES, F. (2020). «Policía predictiva: ¿utopía o distopía? Sobre las actitudes hacia el uso de algoritmos de big data para la aplicación de la ley». IDP. Revista de Internet, Derecho y Política, núm. 30 <https://www.raco.cat/index.php/IDP/article/view/373608>
- MARTÍNEZ GARAY, L. (2018). « Peligrosidad, Algoritmos y Due Process: el Caso State V Loomis». Revista de Derecho Penal y Criminología, 3.ª Época, n. 20, pp. 485-502, <https://doi.org/10.5944/rdpc.20.2018.26484>
- NEIRA PENA, A. (2021). « Inteligencia artificial y tutela cautelar. Especial referencia a la prisión provisional». Rev. Bras. de Direito Processual Penal, Porto Alegre, v. 7, n. 3, p. 1897-1933 <https://doi.org/10.22197/rbdpp.v7i3.618>
- NIEVA FENOLL, J. (2018). Inteligencia artificial y proceso judicial. Madrid: Marcial Pons.
- SIMON CASTELLANO, Pere. (2021a). Justicia cautelar e inteligencia artificial. La alternativa a los atávicos heurísticos judiciales. Barcelona: Bosch.
- SIMON CASTELLANO, Pere. (2021b). «Intel·ligència artificial i Administració de justícia: Quo vadis, justitia?». IDP. Revista de Internet, Derecho y Política, Núm. 33, <https://doi.org/10.7238/idp.v0i33.373817>
- VALLS PRIETO, J. (2017). Problemas jurídico penales asociados a las nuevas técnicas de prevención y persecución del crimen mediante inteligencia artificial. Madrid: Dykinson.
- ZARSKY, T. (2016). «The trouble with algorithmic decisions: an analytic road map to examine efficiency and fairness in automated and opaque decision making». Science, Technology & Human Values, vol. 41, núm. 1, págs. 118-132 [en línea] DOI: <https://doi.org/10.1177/0162243915605575>
- ZAVRŠNIK, A. (2019). Algorithmic justice: Algorithms and big data in criminal justice setting. European Journal of criminology, pp. 1-20, pp. 13-14, septiembre de 2019. <https://doi.org/10.1177/1477370819876762>

Bonus track

¿Algoritmos políticamente (in)correctos?

fundación

Manuel Giménez Abad

de Estudios Parlamentarios y del Estado Autonómico



Muchas gracias por vuestra atención.