

Tecnología y discriminación: El Sesgo de Automatización en la praxis jurídica

Joaquín David Rodríguez Álvarez

Profesor Asociado, Departamento de Ciencias Histórico-Jurídicas
Universitat Autónoma de Barcelona

SUMARIO: I. CONTEXTO. II MÁQUINAS DE AMOROSA GRACIA. III. LOS MITOS Y LOS SESGOS. 1. *El primer mito: la agencia moral.* 2. *El segundo mito: la eficiencia.* 3. *El tercer mito, la confiabilidad.* 4. *El sesgo de Automatización.* IV CONCLUSIONES.

IV CONCLUSIONES.

I. CONTEXTO

La praxis jurídica, así como la práctica totalidad de capas de nuestro sistema, ha experimentado una penetración masiva de tecnologías durante las últimas décadas, que van desde las huellas dactilares hasta las pruebas de ADN, pasando ahora por el reconocimiento facial, o las pruebas genéticas, entre otras¹². Una transformación que afecta profundamente la forma en la que comprendemos nuestro contexto³, mientras la corte de justicia se transforma en una reunión científica⁴.

Nos encontramos perdidos en medio de un sistema, donde el juez (y en algunos casos el jurado), está llamado a ser un “experto” en todos los campos, desde la genética hasta la ciencia forense, pasando por el análisis gráfico y ahora la Inteligencia Artificial (en adelante IA), entre la amplia gama de disciplinas que

¹ JASANOFF, S. (2016). *The Ethics of Invention: Technology and the Human Future*. WW Norton & Company.

² JASANOFF, S. (2009). *Science at the bar: Law, science, and technology in America* (1st ed.). Cambridge England etc.: Harvard University Press.

³ RODRÍGUEZ-ÁLVAREZ, J. (2019). Rethinking Paradigms in the Technolo-ecological Transition. *Journal of Law*, 1(2), 169–184.

⁴ ESTEVE-PARDO, J. (2009). *El desconcierto del Leviatán: Política y derecho ante las incertidumbres de la ciencia*. Madrid: Marcial Pons. Ediciones Jurídicas y Sociales.

hoy pueden invocarse en un tribunal. Sin prestar atención al sesgo, que esos conjuntos tecnológicos^{5 6 7 8 9} pueden introducir en la praxis jurídica^{10 11}.

De esta forma, y a pesar de todas las evidencias que indican nuestras limitaciones para tratar de comprender procesos cuyas raíces se sitúan lejos de nuestra realidad (y privilegios) cotidianos¹², insistimos en una manera de lograr un sistema sin prejuicios, donde la ley, basada en principios igualitarios, se pueda aplicar de manera “justa” a todos los miembros de la comunidad sin considerar su color de piel, género, orientación sexual o clase ... Y hoy, esta esperanza está siendo depositada en la Inteligencia Artificial, como nueva solución totémica que puede corregir (o no), nuestra visión parcial del mundo, prometiéndonos análisis sin sesgos, ecuanimidad, y por ende ser mejor juez que nosotros mismos.

Por tanto, podemos decir, que nuestra disciplina ha experimentado un tránsito siguiendo los sueños de Hobbes, de la bruma del derecho eclesiástico a un sistema incrustado en el marco de, no una verdad natural, sino una “verdad tecnológica”^{13 14}. Y hemos sido adoctrinados para creer que la proliferación de tecnologías en los tribunales podrá otorgar “equidad”, entendida como la superación de los sesgos que tradicionalmente ha cristalizado la desigualdad en nuestras sociedades y sistemas jurídicos¹⁵. Aunque nada, nos haga pensar que los

⁵ ANGWIN, J., LARSON, J., MATTU, S., & KRISCHNER, L. (23 Mayo de 2016). Machine Bias., from <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>

⁶ CUMMINGS, M. (2004). Automation Bias in Intelligent Time Critical Decision Support Systems. In *AIAA 1st Intelligent Systems Technical Conference*. Reston, Virigina: American Institute of Aeronautics and Astronautics.

⁷ NOBLE, S. U. (2016). *Algorithms of oppression : how search engines reinforce racism*.

⁸ SKITKA, L. J., MOSIER, K. L., & BURDICK, M. (1999). Does automation bias decision-making? *International Journal of Human Computer Studies*, 51(5), 991–1006.

⁹ STATT, N. (2020, June 10). Amazon bans police from using its facial recognition technology for the next year - The Verge. 11 de Junio de 2020

¹⁰ LEE, C. (2008). The Gay Panic Defense. *U.C. Davis Law Review*, 42.

¹¹ SALERNO, J. M., NAJDOWSKI, C. J., & BOTTOMS, B. L. (2015). Excusing Murder? Conservative Jurors' Acceptance of the Gay-Panic Defense Effects of Race and Socioeconomic Status on Medical Misdiagnosis of Child Abuse View project Miscarriages of Justice Special Issue of Albany Law Review View project. *Psycnet.Apa.Org*, 21(1), 24.

¹² MANDALIOS, J. (2000). Being and Cultural Difference: (Mis)Understanding Otherness in Early Modernity. *Thesis Eleven*, 62(1), 91–108.

¹³ POSTMAN, N. (2011). *Technopoly: The surrender of culture to technology*. Nueva York: Vintage Books.

¹⁴ WIENER, J. L., & MOHR, L. A. (1994). Technopoly: The Surrender of Culture to Technology. *Journal of Public Policy & Marketing*, 13(2), 326–327.

¹⁵ RODRÍGUEZ-ÁLVAREZ, J., & MARTINEZ-QUIRANTE, R. (2019). *Towards a new AI race. The challenge of lethal autonomous weapons systems (Laws) for the United Nations*. Aranzadi.

conjuntos tecnológicos actuales, van a poder producir este escenario, sino todo lo contrario. Toda la evidencia muestra la capacidad de la tecnología para reproducir los sistemas de opresión, agregando un componente acrítico llamado "sesgo de automatización" comprendido como una sobre dependencia, y exceso de confianza en los sistemas de IA orientados al apoyo en las tomas de decisiones.

Nuestro tiempo, se encuentra amenazado por una nueva cosmovisión mediada por la tecnología, donde la nueva verdad parece surgir de las profundidades de la Inteligencia Artificial. Una tecnología que tiene la capacidad de operar como una caja negra heurística, capaz de darnos respuestas manteniendo el "secreto" de su elaboración^{16 17}. Así, hoy en día, parece que la verdad vuelve a ser revelada, no por los textos sagrados que alguna vez inspiraron y "determinaron" nuestra disciplina para generar certezas ante la naturaleza (y ejecutar brujas), sino por nuestra nueva creencia tecnológica donde el *Big Data* y la IA se estructuran como un nuevo Santo Grial, capaz no sólo de predecir, sino también crear, atendiendo a sus características "deterministas"^{18 19 20}.

A día de hoy, es incluso complejo enumerar todas las formas de IA que han proliferado en el ejercicio de nuestra disciplina, muy bien conocido es el caso de COMPASS como ejemplo perfecto de una "caja negra", donde incluso los algoritmos que utiliza son secretos comerciales²¹. Hecho que impide que sean examinados por el público y los afectados, lo que puede entenderse como una clara violación al debido proceso. Un software que, por cierto, no parece mejor que personas seleccionadas de forma aleatoria, además de que parece presentar un sesgo racial a la hora de calcular los índices de reincidencia²².

¹⁶ BARRAT, J. (2013). *Our final invention : artificial intelligence and the end of the human era*. New York: St. Martin's Press.

¹⁷ TVERSKY, A., & KAHNEMAN, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131.

¹⁸ BIMBER, B. (1994). Three faces of technological determinism. In *Does Technolgy drives history?* (pp. 79–100). Cambridge, Massachusetts: Merrit Smith; Leo Marx.

¹⁹ HUGHES, T. (1994). Technological Momentum. In *Does Technology Drive History?: The Dilemma of Technological Determinism* (p. 101). Cambridge, Massachusetts: Mit Press.

²⁰ SMITH, M., & MARX, L. (1994). *Does technology drive history?: The dilemma of technological determinism* (1st ed.). Boston: Mit Pr; Edición.

²¹ ANGWIN, J., LARSON, J., MATTU, S., & KRISCHNER, L. (23 Mayo de 2016). Machine Bias,. from <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>

²² YONG, E. (2018, January 17). A Popular Algorithm Is No Better at Predicting Crimes Than Random People - The Atlantic. Retrieved June 11, 2020, from <https://www.theatlantic.com/technology/archive/2018/01/equivant-compas-algorithm/550646/>

Es necesario notar, que la tecnología, debe ser entendida como un amplificador de la voluntad humana, tal como lo describe Kentaro Toyama ²³que tiene la posibilidad incrustada de generar negatividades externas ^{24 25 26} siendo necesario poner en valor los distintos impactos que puede tener sobre nuestra sociedad. Especialmente en un momento determinado por la penetración masiva de la tecnología, que está transformando la forma en que operamos, adquirimos conocimientos, comprendemos la “otredad” y nos relacionamos con nuestro contexto^{27 28}. Una penetración que tiene toda una gama de impactos en la práctica del derecho, y como el derecho se regula a sí mismo, así como a los propios marcos tecnológicos.

Una penetración tecnológica que opera en un contexto donde podemos observar claramente la existencia de un retraso cultural ^{29 30 31} entre el ámbito tecnocientífico y el jurídico/legislativo. Donde se ponen en marcha un conjunto de metarrelatos para facilitar la penetración tecnológica en nuestra sociedad, aprovechando una serie de debilidades que padecen la mayoría de los sistemas jurídicos occidentales en la actualidad, como la falta de conocimiento tecnológico o la falta de recursos humanos y su traducción en el correcto funcionamiento del sistema (retrasos en el acceso y la sentencia) o “*el hecho de que muchas jurisdicciones dependan de las tasas y multas como ingresos para el sistema de justicia penal y para otros programas*”³².

Una penetración de la tecnología, que viene a su vez retroalimentada por los planificadores públicos, que buscan la solución a los déficits de personal en la

²³ ITU and XPRIZE. (2017). *AI for Good Global Summit*. Geneva. Retrieved from <https://www.itu.int/en/ITU->

²⁴ CAROLAN, M. (2006). Science, expertise, and the democratization of the decision-making process. *Society and Natural Resources*, 19(7), 661–668.

²⁵ FUNTOWICZ, S. O., & RAVETZ, J. R. (1993). The Emergence of Post-Normal Science. *Science, Politics and Morality*, 17, 85–123.

²⁶ FUNTOWICZ, S., & RAVETZ, R. (1992). The good, the true and the post-modern. *Futures*, 24(10), 963–976.

²⁷ LOCKE, T. (1999). Participation, inclusion, exclusion and netactivism: How the Internet invents new forms of democratic activity. In B. D. Barry Hague (Ed.), *Digital democracy: Discourse and decision making in the Information Age* (pp. 211–223). London: Routledge.

²⁸ RUSSELL, S. J. (STUART J., NORVIG, P., & DAVIS, E. (2010). *Artificial intelligence : a modern approach* (3rd Editio). New York: Pearson.

²⁹ BANATHY, B. H. (2000). *Guided Evolution of Society : a Systems View*. Springer US.

³⁰ OGBURN, W. F. (1922). *Social Change With Respect to Nature and Original Change*. Huebsch.

³¹ TOFFLER, A. (1984). *Future shock*. Bantam.

³² HRW. (2012). *The Case against Killer Robots* . Washington.

tecnología. Algo similar a lo ocurrido con la proliferación de “Smart Contracts”³³
³⁴ una nueva tecnología cuyas implicaciones y cumplimiento con los marcos jurídicos es muy cuestionable pero que aceptamos como algo inevitable. Porque básicamente, nos permite reducir el número de partes involucradas así como el procedimiento de supervisión de sus aplicaciones y desarrollo.

Aceptamos, así, una pérdida constante y sistemática humanidad en todos los estratos de la vida, incluso en la guerra como ilustra el caso de los sistemas letales autónomos, cuyo nivel de cumplimiento de las normas del Derecho Internacional Humanitario es también más que cuestionable anticipando una tendencia que podría tener efectos dramáticos en nuestras sociedades, especialmente si esta tendencia se cristaliza en el sistema legal, un sistema que per se debería estar basado en el juicio humano

II. MÁQUINAS DE AMOROSA GRACIA.

“All Watched Over by Machines of Loving Grace” es el título de un poema escrito por Richard Brautigan y publicado por primera vez en 1967 en una colección del mismo nombre. Una representación de un futuro utópico donde nuestra especie finalmente puede encontrar el equilibrio con el resto de las formas de vida de nuestro planeta gracias al papel de la computación moderna. Un poema, relacionado con una escuela ideológica y filosófica, el objetivismo, que floreció en los años 90 bajo la administración Clinton (en el campo económico) gracias al papel del objetivista Alan Greenspan (Presidente de la Reserva Federal de 1987 a 2006), y cuyo impacto en la configuración de la Economía Global actual es clave para comprender la metamorfosis del Estado y sus marcos jurídicos. Porque es precisamente la corriente de la desregulación (abanderada por la administración Regan-Thatcher y sus seguidores) inflamada por la creencia de que el progreso tecnocientífico. Responsable último de la degradación de la idea de democracia y la metamorfosis del Estado en una forma de gobierno privado³⁵ donde el sistema de privilegios que opera, no se erosione sino que se cristalice.

El objetivismo, fue gestado por Ayn Rand, y estaba basado en cuatro postulados: realidad, razón, interés propio y capitalismo³⁶. Una filosofía donde el desarrollo tecnológico jugó un papel fundamental como principal medio para crear una

³³ CHRISTIDIS, K., ACCESS, M. D.-I. (2016). Blockchains and smart contracts for the internet of things. *Ieeexplore.Ieee.Org*.

³⁴ KOSBA, A., MILLER, A., SHI, E., Z. I. symposium (2016). Hawk: The blockchain model of cryptography and privacy-preserving smart contracts. In *IEEE Symposium on Security and Privacy*

³⁵ MBEMBE, A. (2011). *Necropolítica*. Barcelona: Melusina.

³⁶ RAND, A. (2020). *Aynrand.org*.

nueva sociedad.

Un enfoque filosófico, que sigue jugando un papel fundamental en nuestro actual sistema de creencias, con profundas afectaciones sobre nuestro comportamiento, especialmente relacionado con aquellos procesos de toma de decisiones que involucran a la tecnología como fuente y medio de análisis. Como dijo Baudrillard “Vivimos en un mundo donde hay cada vez más información, y cada vez menos significado”³⁷.

Por lo tanto podemos observar, como hoy en día, existe una tendencia de confianza versus la máquina³⁸. Todavía nos encontramos hipnotizados por sus capacidades, su velocidad, su fiabilidad (?) como John Gast mientras pintaba su famosa obra “American Progress”. Estamos “hechizados” por la tecnología de tal manera que no nos permite entender del todo que las máquinas también son falibles³⁹ y cuanta más responsabilidad les demos más dramáticos serán estos fracasos. Y una vez que esto suceda, ¿quién será responsable?

¿Cómo vamos a poder lidiar con errores, con consecuencias fatales que son producto de algoritmos y no de humanos?, ¿cómo vamos a poder proceder en una sociedad donde la toma de decisiones es compartida, cuando no se transfiere o delega directamente a actores no humanos?, como sucede de manera muy visible en el caso de los Sistemas de Armas Autónomas Letales. Pero también puede afectar la concesión de un crédito (credit-scoring), la admisión a una universidad, el cálculo de reincidencia de un acusado o incluso todos los aspectos de la vida como el sistema de *social-scoring* desarrollado por china.

Demasiadas posibilidades y tan pocas respuestas para un fenómeno que está en todas partes, de la Conferencia de Seguridad de Múnich⁴⁰ donde Sophia, la animatrónica, daba lecciones a los líderes mundiales⁴¹ mientras propulsaba la agenda de su creador para facilitar la permeabilidad de nuestro sistema a la IA a

³⁷ BAUDRILLARD, J. (1987). Cool memories 1980-1985. *Collection Débats*.

³⁸ LI, X., HESS, T. J., & VALACICH, J. (2008). Why do we trust new technology? A study of initial trust formation with organizational information systems. *The Journal of Strategic Information Systems*, 17(1), 39–71.

³⁹ O'NEIL, C. (2017). *Weapons of math destruction : how big data increases inequality and threatens democracy*. Broadway Books.

⁴⁰ SHARKEY, N. (2018). Mama Mia It's Sophia: A Show Robot Or Dangerous Platform To Mislead?

⁴¹ WAREHAM, M. (2018). “Want to worry about AI? Then worry about this..” Humans Right Watch. Retrieved April 15, 2019, from <https://www.hrw.org/news/2018/03/07/want-worry-about-ai-then-worry-about>

través de formas como “*Smart Policing*”⁴².

Una narrativa que nos adoctrina para creer que el Big Data y la IA, nos están permitiendo ser libres, incluso de juzgar, sin darnos cuenta de que estamos aceptando retirarnos de algunos aspectos de nuestra existencia que nos resultan incómodos, básicamente porque aún tenemos dudas. Pero hay un sistema capaz, no de erradicarlos, sino de barnizarlos con algoritmos (O’Neil, 2017). Una simulación de inteligencia, que es solo una imitación, nunca una reproducción, básicamente porque las máquinas, no sienten, no tienen empatía u otras cualidades humanas que son básicas para el ejercicio de nuestra disciplina. El juicio humano no puede ser sustituido, ni siquiera guiado por sistemas tecnológicos incapaces de comprender las implicaciones más profundas de la existencia humana.(Conitzer, 2017).

III. LOS MITOS Y LOS SESGOS

“El secreto de toda teoría es que la verdad no existe”

(Baudrillard, 1987)

Históricamente, todas las comunidades humanas se han visto afectadas por un conjunto de “sesgos” que ha redefinido el papel de la “otredad” basado en una multiplicidad de factores donde la experiencia (individual y colectiva) y la ideología/prejuicios han jugado un papel importante en la elaboración de metarrelatos complejos⁴³. Unos metarrelatos que tienden a exagerar las cualidades morales/éticas (incluso físicas e intelectuales) de los miembros de un grupo (el dominante) sobre los “otros” cuyo retrato normalmente incluye factores de “inferioridad”. Un mecanismo que ha sido profundamente estudiado y cuya traducción en nuestra sociedad se puede ver en casi todas partes, también a través de procesos que hemos “normalizado” como es el “perfilado racial”⁴⁴.
⁴⁵cuyo impacto sobre la seguridad pública es muy notorio, con una clara sobrerrepresentación de las minorías en el sistema judicial y penitenciario. Y

⁴² COLDREN, J. R., HUNTOON, A., & MEDARIS, M. (2013). Introducing Smart Policing: Foundations, Principles, and Practice. *Police Quarterly*, 16(3), 275–286.

⁴³ BALIBAR, E., & WALLERSTEIN, I. (1991). *Race, Nation, Class: Ambiguous Identities*.

⁴⁴ WEITZER, R., & TUCH, S. A. (2002). Perceptions of racial profiling: Race, class, and personal experience. *Criminology*, 40(2), 435–456.

⁴⁵ WELCH, K. (2007). Black Criminal Stereotypes and Racial Profiling. *Journal of Contemporary Criminal Justice*, 23(3), 276–288. <https://doi.org/10.1177/1043986207306870>

donde factores como el racismo, el clasismo, o la misoginia juegan un papel importante.

Es necesario señalar, que nosotros como humanos experimentamos una realidad mediatizada a través de la ideología⁴⁶, por lo tanto nuestras acciones, y principales líneas de pensamiento están contaminadas por un conjunto de prejuicios (que todos compartimos, solo cambiando su grado de persistencia). La mayor parte de nosotros (todos nosotros), hemos sido criados en un sistema, donde el racismo, el clasismo, la misoginia, la homofobia, la transfobia (entre otras) opera como una ideología incrustada, por lo que, en mayor o menor grado, forman parte de quienes somos. Incluso afectando a quienes somos parte de estos grupos marginados. ¿Quién no ha sido inducido a aceptar la noción de un “orden natural” que opera en su interior, como lo muestra la dialéctica del amo y el esclavo?⁴⁷

Debemos notar, en este sentido, que los marcos mentales del humanismo tradicional (incrustados en la superestructura del sistema^{48 49} hacen referencia a las ramificaciones de su representación vitrubiana, que tradicionalmente se ha entendido como un postulado universal. Aunque esta imagen arquetípica coincide implícitamente sólo con el varón, blanco, urbanizado, hablante de una lengua estándar, heterosexual, inscrito en la unidad reproductiva básica, ciudadano de pleno derecho de una comunidad política reconocida⁵⁰. Siendo esta quizás, la imagen menos representativa de la realidad contemporánea, y un claro ejemplo de cristalización del sistema de privilegios imperante^{51 52}.

Pero el hecho de que no hayamos podido borrar esos sesgos, no significa que no haya existido un esfuerzo constante para hacerlo, por parte de tantos miembros de la “comunidad legal” y activistas en todo el mundo. Esfuerzos que en algunos casos han sido positivos, y en otros, tienen consecuencias más dudosas: Como en el caso de Amadou Diallo, un inmigrante africano desarmado que recibió 41 disparos por parte de agentes de la Policía de nueva York en la puerta de su

⁴⁶ ZIZEK, S. (1997). The Supposed Subjects of Ideology. *Critical Quarterly*, 39(2), 39–59.

⁴⁷ HEGEL, G. (1966). *Fenomenología del espíritu* (1st ed.). Buenos Aires: Fondo de Cultura Económica.

⁴⁸ BATES, T. (1975). Gramsci and the Theory of Hegemony. *Journal of the History of Ideas*, 32(2), 351–356.

⁴⁹ FEMIA, J. (1987). *Gramsci's political thought: hegemony, consciousness, and the revolutionary process* (4th ed.). London: Clarendon Press.

⁵⁰ DELEUZE, G., & GUATTARI, F. (2006). *Millepiani*. Roma: Castelvecchi.

⁵¹ BRAIDOTI, R. (2015). *Lo Postumano* (1st ed.). Barcelona: Gedisa.

⁵² IRIGARAY, L. (2010). *Speculum. L'altra donna*. Universale económica.

casa⁵³, o el caso de Rodney King^{54 55}, golpeado hasta la muerte por agentes de la policía de Los Angeles. En ambos casos, se presentó una moción de cambio de sede, ya que la defensa entendió que la publicidad negativa sobre sus clientes corrompería al jurado, agregando además, en el caso de Amadou que los vecinos del Bronx suelen albergar “prejuicios” negativos contra policías por lo que se inclinaron a condenarlos, por lo que para tener un “juicio justo” ese cambio era una necesidad.

Por tanto, aquí podemos encontrar un claro ejemplo de uno de esos mecanismos puestos en marcha para evitar sesgos para determinar el sentido del juicio, como es la posibilidad de cambio de sede. Aun cuando podemos incurrir en el error de entender que aquellas comunidades que no han estado expuestas masivamente a los medios y al impacto de un evento, no van a presentar prejuicios, como raciales, o confianza en la policía (por su experiencia específica y construcción de la percepción).

Así en ambos de los casos anteriormente mencionados los agentes de policía fueron declarados no culpables una vez que el juicio se incrustó en las comunidades conservadoras blancas. Un ejemplo que también nos muestra cómo, incluso, los mecanismos establecidos para evitar sesgos negativos sobre los acusados pueden ser mal utilizados para incrustar “sesgos” positivos sobre ellos.

Y aquí es donde, la tecnología, puede estructurarse como promesa ⁵⁶de “equidad”, pero esta afirmación: ¿es verdad? ¿O solo una parte de una narrativa en la que hemos creído acráticamente, poniendo en riesgo no solo a nosotros mismos sino a todo el progreso que hemos realizado en esta área? Uno de los problemas a la hora de afrontar la producción tecnológica y la penetración social es que tendemos a crear narrativas complejas en torno a ella, de tal forma que condicionan y determinan nuestros planteamientos.

A lo largo de la historia han sido no pocos los sistemas tecnológicos que han iluminado metarrelatos complejos que confluyen con formas simbólicas,

⁵³ EMBRICK, D. G. (2015). Two Nations, Revisited: The Lynching of Black and Brown Bodies, Police Brutality, and Racial Control in “Post-Racial” Amerikkka. *Critical Sociology*, 41(6), 835–843.

⁵⁴ FERLDMAN, A. (1994). On cultural anesthesia: From desert storm to Rodney King. *American Ethnologist*, 21(2), 404–418.

⁵⁵ WEITZER, R., & TUCH, S. A. (2002). Perceptions of racial profiling: Race, class, and personal experience. *Criminology*, 40(2), 435–456.

⁵⁶ POSTMAN, N. (2006). *Amusing ourselves to death: Public discourse in the age of show business* (20th ed.). London: Penguin.

generando disociaciones entre realidad y ficción⁵⁷. Así ocurrió con el surgimiento del calendario, asociado a la revolución neolítica como tecnología que permitía la reproducción de ciclos agrícolas y que acabó albergando un pensamiento mágico relacionado con la trazabilidad del tiempo, y su relación con los ciclos astrales. Un tipo de pensamiento que configuró fenómenos de opresión/control a través de la previsión de fenómenos astrales complejos (predicción de eclipses, por ejemplo), entre otros⁵⁸.

Si bien, no es necesario remontarse a pasados remotos para observar la reproductibilidad del fenómeno en sociedades reconocidas como científicas, como es el caso del descubrimiento de la radiación y sus posteriores campos de aplicación. Así, encontramos mitos que han surgido en torno a sus propiedades curativas y la seguridad de su uso en los años 20 y 30 del siglo pasado. Es necesario comprender en qué medida se desarrolló esta mitigación de la seguridad con respecto a la radiactividad, que llevó a la producción de productos de consumo general de todo tipo, incluyendo pastas dentales, chocolates, cremas para la piel, y un largo etcétera, con el único objetivo de fomentando su consentimiento social. Pero estos distan mucho de ser los únicos, y es necesario destacar aquellos de carácter estructural que han tenido profundos efectos en el orden global y en lo humano, mitos que condicionaron las relaciones internacionales durante décadas, como el de la destrucción mutua asegurada, o las que se materializaron en forma de promesas de fuentes de energía limpias, baratas y seguras para el futuro y que tomarían la forma de programas como *Atoms for Peace*.

Así nos encontramos con la estructuración de una serie de metarrelatos en torno a la tecnología que tuvieron la capacidad de acaparar su desarrollo bajo el prisma de la hegemonía cultural, que cabe destacar, no condujo a una liberación futura ni a una producción global de “bien” sino a la perpetuación de dinámicas de opresión y sometimiento cuya materialización se encuentra en todos y cada uno de los conflictos por poderes que se desarrollaron durante la Guerra Fría. período.

Podemos encontrar otros ejemplos donde se repite esta dinámica, como el campo de la biogenética, donde bajo las promesas de cultivos resistentes y el fin del hambre en el mundo, se escondía una agenda orientada a generar marcos tecnológicos que no pivotaran en torno al bien común, sino en torno al interés del sistema de producción capitalista, donde el “gen terminador”, se erige como un

⁵⁷ RODRÍGUEZ, J. (2016). *La civilización ausente: Tecnología y sociedad en la era de la incertidumbre* (1st ed.). Oviedo: Trea.

⁵⁸ ELLUL, J., WILKINSON, J., & MERTON, R. (1964). *The technological society*. Nueva York: Random House.

ejemplo paradigmático, reproduciendo con su mera existencia las dinámicas de opresión y posesión instaladas en el medio rural y agrícola, generando graves amenazas a la soberanía alimentaria (como sucedió en Haití con Monsanto). Otra tecnología con potencial liberador encaminada a la cristalización del orden social y la consagración de un sistema productivo, que sigue fomentando el monocultivo, a pesar de la plena conciencia de sus efectos erosivos sobre el ecosistema.

Y hoy, el ciclo parece repetirse para la IA. En los foros internacionales de primer nivel se vuelven a alcanzar promesas sobre escenarios utópicos, siguiendo el camino del objetivismo. Los escenarios, que se dice, solo se pueden lograr a través de un solo camino: transferencia de datos, transferencia de privacidad y, en última instancia, transferencia de humanidad⁵⁹. Por lo que debemos advertir que el mundo de la Inteligencia Artificial está lleno de sombras que deben ser develadas, sobre todo, teniendo en cuenta que muchos de los datos proporcionados por los usuarios corren el riesgo de ser utilizados en su contra, ya sea por corporaciones privadas o por programas militares, como está pasando con los programas de reconocimiento facial. Siendo necesario explorar los tres principales mitos en torno a la IA, tales como: humanización, eficiencia y confiabilidad.

1. El primer mito: la agencia moral

El primer mito es que, correctamente codificadas, las máquinas pueden adoptar comportamientos ético-morales. Pero claramente, una máquina no es capaz ni de ética ni de moral ni de intuición, básicamente porque no es inteligente “per se” sino solo una simulación de Inteligencia⁶⁰. En todo caso puede ser considerado como un agente moral porque una máquina solo puede aspirar a tener la ética de quien la codificó⁶¹. Lo que significa que será solo una simulación.(Baudrillard, 1994) de la ética del programador, el ingeniero o una combinación de información procedente de la nube. Lo que nos lleva a la pregunta: ¿de quién es la ética que vamos a programar? E incluso si llegamos a este escenario: ¿el sistema evolucionará por sí mismo? ¿O estaremos condenados a una sociedad

⁵⁹ GRIFFIN, A. (2018). Saudi Arabia grants citizenship to a robot for the first time ever | The Independent. 15 de Agosto de 2019, from <https://www.independent.co.uk/life-style/gadgets-and-tech/news/saudi-arabia-robot-sophia-citizenship-android-riyadh-citizen-passport-future-a8021601.html>

⁶⁰ MARTINEZ- QUIRANTE, R., & RODRÍGUEZ-ÁLVAREZ, J. (2020). El lado oscuro de la Inteligencia artificial . IDEES, (48).

⁶¹ RUSSELL, S. J. (STUART J., NORVIG, P., & DAVIS, E. (2010). *Artificial intelligence : a modern approach* (3rd Editio). New York: Pearson.

inmovilizada donde el bien y el mal se cristalizan sobre la base de una construcción algorítmica subjetiva? Y si evolucionara... ¿cuál sería su horizonte?

En resumen, la inteligencia artificial de ninguna manera puede considerarse un agente moral porque es meramente una simulación. No puede, por tanto, comprender, bajo ningún parámetro, algo tan fundamental y fundamental como el valor de una vida humana, ni sentir respeto ni compasión. Características, todas ellas que deben ser inherentes al sistema de justicia.

2. *El segundo Mito, La eficiencia*

El segundo mito o falacia es que la IA puede tomar decisiones de manera más efectiva, equitativa y justa que un ser humano. Nada más lejos de la realidad, en primer lugar, porque la IA emula el sistema ético-ideológico de sus creadores. En otras palabras, reproduce nuestra falta de imparcialidad. Como mostró Cathy O'Neil en su trabajo Armas de destrucción matemática⁶², donde muestra los dramáticos efectos que puede tener sobre la sociedad la creencia acrítica en la infalibilidad de los algoritmos. Un hecho que también se ve reforzado por el trabajo realizado por la American Civil Liberties Union para demostrar que los sistemas de reconocimiento facial tienen una alta tendencia a identificar a sujetos no caucásicos como delincuentes.

En la prueba realizada por ACLU en 2018 utilizando una herramienta de reconocimiento facial, llamada "Rekognition", propiedad de Amazon que está proliferando en los usos policiales y policiales⁶³, el software emparejó incorrectamente a 28 miembros del Congreso, identificándolos como otras personas que habían sido arrestadas por un delito: "*Los miembros del Congreso que fueron emparejados falsamente con la base de datos de fichas policiales que usamos en la prueba incluyen republicanos y demócratas, hombres y mujeres, y legisladores de todas las edades, de todo el país. Las coincidencias falsas fueron desproporcionadamente de personas de color, incluidos seis miembros del Caucus Negro del Congreso, entre ellos la leyenda de los derechos civiles, el representante John Lewis (D-Ga.). Estos resultados demuestran por qué el Congreso debería unirse a la ACLU para pedir una moratoria sobre el uso de la vigilancia facial por parte de las fuerzas del orden*"⁶⁴.

⁶² O'NEIL, C. (2017). *Weapons of math destruction : how big data increases inequality and threatens democracy*. Broadway Books.

⁶³ BRANDON, R. (22 de Mayo de 2018)). Amazon is selling police a real-time facial recognition system - The Verge. (10 de Junio de 2020), de <https://www.theverge.com/2018/5/22/17379968/amazon-rekognition-facial-recognition-surveillance-aclu>

⁶⁴ SNOW, J. (2018, July 26). Amazon's Face Recognition False Matched 28 Members of Congress With Mugshots | American Civil Liberties Union. 10 de Junio de 2020,

Estamos, así, ante una tecnología diseñada en su mayoría por hombres blancos, capaces de transferir los gustos y disgustos de sus creadores. Además, al ser un sistema heurístico, comprender los procesos utilizados por la IA para llegar a una decisión específica es increíblemente desafiante. Y si es imposible, en tantos casos, deconstruir o explicar los procesos detrás de una decisión particular de IA, entonces es irresponsable permitir que infieran en nuestros sistemas de justicia..

3. *El tercer mito, la confiabilidad*

Y finalmente llegamos a la tercera falacia, que sostiene que la inteligencia artificial es más confiable que la inteligencia humana. Este concepto podría aceptarse en circunstancias muy específicas, pero nunca en términos generales. Aquí, cabe destacar el trabajo de la ONG británica Big Brother Watch⁶⁵ se materializó a través de su informe de 2018 "Face off: The lawless growth of facial recognition in UK policing"⁶⁶, que, a través de una investigación basada en información, que involucró más de 50 solicitudes de información, revela como uno de sus hallazgos clave que: "*La gran mayoría de las 'coincidencias' de la policía que utilizan el reconocimiento facial automatizado hasta la fecha han sido inexactas. En promedio, un asombroso 95% de las "coincidencias" identificaron erróneamente a personas inocentes*". (*ibidem* p7). Poniendo serias dudas no solo sobre la precisión/confiabilidad de la tecnología, sino sobre la justificación de su introducción en la aplicación de la ley (especialmente en esta etapa de desarrollo), siendo en este aspecto interesante notar que: "*Las fuerzas policiales han almacenado fotos de todos personas emparejadas incorrectamente por sistemas automatizados de reconocimiento facial, lo que lleva al almacenamiento de fotos biométricas de miles de personas inocentes*". Por lo tanto, podemos observar qué tipo de agendas están en marcha bajo la implementación de estos conjuntos tecnológicos, aún sabiendo que su "utilidad", en relación a su justificación –identificar delincuentes y delincuentes- es en la práctica nula.

Otro ejemplo de esta falta de confiabilidad lo podemos encontrar, en el caso de COMPASS, la tecnología mencionada en la primera sección del presente artículo. Un estudio realizado por ProPublica encontró, después de evaluar las "*puntuaciones de riesgo obtenidas asignadas a más de 7,000 personas arrestadas en el condado de Broward, Florida, en 2013 y 2014 y verificó cuántos fueron acusados de*

de <https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28>

⁶⁵ BBW. (n.d.). Big Brother Watch: Defending Civil Liberties, Protecting Privacy. (4 de Junio de 2020), from <https://bigbrotherwatch.org.uk/>

⁶⁶ CARLO, S., KRUECKEBERG, J., & FERRIS, G. (2018). *Face off: The lawless growth of facial recognition in UK policing*. Retrieved from www.bigbrotherwatch.org.uk

nuevos delitos en los próximos dos años, la misma benchmark utilizado por los creadores del algoritmo. El puntaje demostró ser notablemente poco confiable para pronosticar delitos violentos: solo el 20 por ciento de las personas que se predijo que cometían delitos violentos en realidad lo hicieron". Y "era particularmente probable que la fórmula marcará falsamente a los acusados negros como futuros delincuentes, etiquetándolos erróneamente de esta manera en casi el doble de la tasa de acusados blancos. Los acusados blancos fueron mal etiquetados como de bajo riesgo con más frecuencia que los acusados negros"⁶⁷.

Por lo tanto, aquí tenemos una tecnología, que, si bien las potencialidades de su futuro desarrollo son altas, el estado del arte actual nos muestra que no se corresponde con las expectativas que se han creado en torno a su desempeño, al menos en la actualidad en algunas áreas, siendo directamente imposible en otras como las relacionadas con la agencia moral. Y cómo la creencia acrítica en su humanidad, eficiencia y confiabilidad puede generar enormes problemas para la aplicación de la ley y el sistema de justicia, especialmente cuando se implementa sin un "control humano significativo".

Pero el problema no está solo relacionado con el hecho de que estamos desplegando y usando tecnologías que no cumplen con las expectativas, el problema es que creemos que sí, porque estamos siendo adoctrinados por las narrativas descritas anteriormente, y aquí es donde viene el problema del sesgo de automatización.

4. El sesgo de Automatización

El Sesgo de Automatización es un fenómeno que ha sido profundamente estudiado a lo largo de los últimos años, pero cuyo reconocimiento de impacto en los sistemas de Justicia ha sido sistemáticamente ignorado, permitiendo la penetración de conjuntos de tecnologías, que como se mostró anteriormente, pueden tener efectos dramáticos sobre la justicia.

El sesgo de automatización se puede describir como un fenómeno cognitivo, relacionado con la propensión de los humanos a favorecer las sugerencias de los sistemas automatizados de toma de decisiones e ignorar la información contradictoria hecha sin automatización, incluso si es correcta. Existiendo una larga lista de ejemplos donde se está produciendo este fenómeno, especialmente preocupante en campos como el militar donde incluso tenemos casos de tropas muertas por fuego amigo⁶⁸ debido al sesgo de automatización, lo que representa uno de los mayores desafíos en la carrera por las armas letales autónomas.

⁶⁷ ANGWIN, J., LARSON, J., MATTU, S., & KRISCHNER, L. (23 Mayo de 2016). Machine

⁶⁸ SCHARRÉ, P., & NORTON, W. W. (2018). *Army of None: Autonomous Weapons and the Future of*

El concepto de sesgo de automatización está relacionado con la literatura de psicología social que encontró un sesgo en la interacción humano-humano que mostró que las personas asignan evaluaciones más positivas a las decisiones tomadas por humanos que a un objeto neutral⁶⁹. Y se ha encontrado el mismo tipo de sesgo de positividad para la interacción humano-automatización⁷⁰, donde las decisiones automatizadas se califican más positivamente que neutrales^{71 72}

Este fenómeno se ha convertido en un problema creciente reconocido para la toma de decisiones en varias áreas, como las unidades de cuidados intensivos, las plantas de energía nuclear y las cabinas de los aviones que tienen cada vez más monitores de sistemas computarizados integrados y ayudas para la toma de decisiones para descartar en su mayoría posibles errores humanos.⁷³ Siendo necesario desarrollar herramientas que nos puedan ayudar a corregir este fenómeno. En la mayor parte de los casos analizados en la literatura⁷⁴. Los errores de sesgo de automatización tienden a ocurrir cuando la toma de decisiones depende de computadoras u otras ayudas automatizadas y el ser humano tiene un rol de observador pero es capaz de tomar decisiones.

Entonces aquí tenemos un fenómeno, cuyos efectos pueden ser dramáticos en un sistema de justicia que ha estado expuesto y lo está constantemente, a la penetración de la Automatización. Una penetración que ha sido posible gracias a un proceso de adoctrinamiento que ha facilitado un acercamiento mítico a la IA, muy lejos de las capacidades reales de la tecnología en su estado actual de desarrollo (y muy probable que también en el futuro). Una tecnología que tiene

War Army of None: Autonomous Weapons and the Future of War Arms Control Today. Retrieved from <https://www.armscontrol.org>

⁶⁹ BRUNER, J., & TAGIURI, R. (1954). The perception of People. In *Handbook of Social Psychology: Special fields and applications* (Vol. 2, pp. 634–651). Addison-Wesley.

⁷⁰ MADHAVAN, P., & WIEGMANN, D. A. (2007). Similarities and differences between human-human and human-automation trust: An integrative review. *Theoretical Issues in Ergonomics Science*, 8(4), 277–301.

⁷¹ DZINDOLET, M. T., PETERSON, S. A., POMRANKY, R. A., PIERCE, L. G., & Beck, H. P. (2003). The role of trust in automation reliance. *International Journal of Human Computer Studies*, 58(6), 697–718.

⁷² SKITKA, L. J., MOSIER, K. L., & BURDICK, M. (1999). Does automation bias decision-making? *International Journal of Human Computer Studies*, 51(5), 991–1006. <https://doi.org/10.1006/ijhc.1999.0252>

⁷⁴ CUMMINGS, M. (2004). Automation Bias in Intelligent Time Critical Decision Support Systems. In *AIAA 1st Intelligent Systems Technical Conference*. Reston, Virigina: American Institute of Aeronautics and Astronautics.

la capacidad de reproducir nuestro sesgo, como explica Nobel en su trabajo “Algoritmos de la opresión” (Noble, 2016) donde subraya el papel que puede jugar la raza en la cristalización de los prejuicios “algorítmicos”. Condenándonos a adentrarnos más en la cueva, y en todo caso permitiéndonos corregir el sesgo, ya que era el objetivo principal, o excusa de su despliegue.

IV CONCLUSIONES

La tecnología hoy, tiene la potencialidad de convertirse en fuente de “cristalización de la ética”, debido a las promesas y mitos que envuelven la noción de Autonomía, y su capacidad de hacernos aceptar que es mejor de lo que somos, cuya máxima manifestación es el Sesgo de Automatización. Reforzando el sistema de prejuicios que opera en nuestra sociedad, afectando la percepción de la “otredad” cuestionando el acceso a un juicio justo.

Hemos sido adoctrinados para creer en un conjunto de capacidades que la tecnología no tiene hoy y probablemente nunca tendrá. Permitir que penetre en cada capa de nuestra sociedad, donde el sistema de justicia y el sistema de aplicación de la ley no son una excepción, incluso cuando la tecnología es de dudosa utilidad si se compara con sus expectativas supone un enorme riesgo. Un hecho que no impide una asignación constante de recursos en el ecosistema de IA. Incluso cuando se reconoce el sesgo, lo ha hecho.

La única forma que tenemos de reducir el sesgo en el sistema de justicia es reducir el sesgo en la sociedad, por lo que el sistema legal es solo un reflejo de la sociedad en la que habita. No podemos esperar que la tecnología resuelva por sí sola el problema del sesgo, mucho menos a través de la delegación de procesos críticos, como los que involucran a la aplicación de la ley y el sistema de justicia, por lo que tiene la potencialidad de cristalizar la desigualdad bajo un enfoque algorítmico.

Pero más aún, si aceptamos esta penetración, estaremos renunciando a nuestra responsabilidad. Una responsabilidad que debe entenderse referida al desmantelamiento del mecanismo de opresión y privilegios que opera en nuestra sociedad, y la responsabilidad de ejercer control humano sobre todo el proceso judicial, como sobre todo proceso que pueda tener potenciales implicaciones profundas para un ser humano. ser, como el control de fronteras (“iBorderCtrl,” n.d.) o apuntar a sospechosos entre muchos otros.

Finalmente, aún en el caso de que aceptemos su penetración como algo inevitable, es necesario establecer un conjunto de reglas que subrayen el papel del “control humano significativo” así como la creación de una agencia pública de evaluación de tecnologías, capaz de generar una marco claro de actuación que

pueda orientar a los usuarios, y sensibilizarlos sobre los posibles sesgos para corregirlos.

La aceptación acrítica del análisis que hace la tecnología, tiene hoy, no sólo la capacidad de cristalizar el actual sistema de igualdad a través de algoritmos, sino la potencialidad incrustada de ofrecerles una nueva capa de justificación, cristalizando la discriminación sobre una base algorítmica.

Bibliografía

- ANGWIN, J., LARSON, J., MATTU, S., & KRISCHNER, L. (23 Mayo de 2016). Machine Bias., from <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- BALIBAR, E. (1991). La forma nación: historia e ideología. In *Raza, nación y clase*.
- BALIBAR, E., & WALLERSTEIN, I. (1991). *Race, Nation, Class: Ambiguous Identities*.
- BANATHY, B. H. (2000). *Guided Evolution of Society : a Systems View*. Springer US.
- BARRAT, J. (2013). *Our final invention : artificial intelligence and the end of the human era*. New York: St. Martin's Press.
- BATES, T. (1975). Gramsci and the Theory of Hegemony. *Journal of the History of Ideas*, 32(2), 351–356.
- BAUDRILLARD, J. (1987). Cool memories 1980-1985. *Collection Débats*.
- BAUDRILLARD, J. (1994). *Simulacra and simulation* (1st ed.). Detroit: University of Michigan Press.
- BBW. (n.d.). Big Brother Watch: Defending Civil Liberties, Protecting Privacy. (4 de Junio de 2020), from <https://bigbrotherwatch.org.uk/>
- BHUTA, N. (2016). *Autonomous weapons systems : law, ethics, policy*. Cambridge University Press.
- BIMBER, B. (1994). Three faces of technological determinism. In *Does Technolgy drives history?* (pp. 79–100). Cambridge, Massachusetts: Merrit Smith; Leo Marx.
- BRAIDOTI, R. (2015). *Lo Posthumano* (1st ed.). Barcelona: Gedisa.
- BRANDOM, R. (22 de Mayo de 2018)). Amazon is selling police a real-time facial recognition system - The Verge. (10 de Junio de 2020), de <https://www.theverge.com/2018/5/22/17379968/amazon-rekognition-facial-recognition-surveillance-aclu>
- BRUNER, J., & TAGIURI, R. (1954). The perception of People. In *Handbook of Social Psychology: Special fields and applications* (Vol. 2, pp. 634–651). Addison-Wesley.
- BURRI, T. (2016). The Politics of Robot Autonomy. *European Journal of Risk Regulation*, 7(02), 341–360.
- CARLO, S., KRUECKEBERG, J., & FERRIS, G. (2018). *Face off: The lawless growth of facial recognition in UK policing*. Retrieved from www.bigbrotherwatch.org.uk

- CAROLAN, M. (2006). Science, expertise, and the democratization of the decision-making process. *Society and Natural Resources*, 19(7), 661–668.
- CHRISTIDIS, K., ACCESS, M. D.-I. (2016). Blockchains and smart contracts for the internet of things. *Ieeexplore.Ieee.Org*.
- COLDREN, J. R., HUNTOON, A., & MEDARIS, M. (2013). Introducing Smart Policing: Foundations, Principles, and Practice. *Police Quarterly*, 16(3), 275–286.
- CONITZER, V. (2017). Moral decision Making Framweroks for Artifical Intelligence. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence* (pp. 4831–4835). Durham: Duke University Press.
- CUMMINGS, M. (2004). Automation Bias in Intelligent Time Critical Decision Support Systems. In *AIAA 1st Intelligent Systems Technical Conference*. Reston, Virigina: American Institute of Aeronautics and Astronautics.
- DELEUZE, G., & GUATTARI, F. (2006). *Millepiani*. Roma: Castelvecchi.
- DZINDOLET, M. T., PETERSON, S. A., POMRANKY, R. A., PIERCE, L. G., & Beck, H. P. (2003). The role of trust in automation reliance. *International Journal of Human Computer Studies*, 58(6), 697–718.
- ELLUL, J., WILKINSON, J., & MERTON, R. (1964). *The technological society*. Nueva York: Random House.
- EMBRICK, D. G. (2015). Two Nations, Revisited: The Lynching of Black and Brown Bodies, Police Brutality, and Racial Control in “Post-Racial” Amerikkka. *Critical Sociology*, 41(6), 835–843.
- ESTEVE-PARDO, J. (2009). *El desconcierto del Leviatán: Política y derecho ante las incertidumbres de la ciencia*. Madrid: Marcial Pons. Ediciones Jurídicas y Sociales.
- FEMIA, J. (1987). *Gramsci's political thought: hegemony, consciousness, and the revolutionary process* (4th ed.). London: Claredon Press.
- FERLDMAN, A. (1994). On cultural anesthesia: From desert storm to Rodney King. *American Ethnologist*, 21(2), 404–418.
- FUNTOWICZ, S. O., & RAVETZ, J. R. (1993). The Emergence of Post-Normal Science. *Science, Politics and Morality*, 17, 85–123.
- FUNTOWICZ, S., & RAVETZ, R. (1992). The good, the true and the post-modern. *Futures*, 24(10), 963–976.
- GOODING-WILLIAMS, R. (2013). *Reading Rodney King/reading urban uprising*. New York: Routledge.
- GRIFFIN, A. (2018). Saudi Arabia grants citizenship to a robot for the first time ever | The Independent. Retrieved April 15, 2019, from <https://www.independent.co.uk/lifestyle/gadgets-and-tech/news/saudi-arabia-robot-sophia-citizenship-android-riyadh-citizen-passport-future-a8021601.html>
- HEGEL, G. (1966). *Fenomenología del espíritu* (1st ed.). Buenos Aires: Fondo de Cultura Económica.
- HIRSCHFIELD, P., CRIMINOLOGY, D. S.-T., & 2010, undefined. (2010). Legitimizing police violence Newspaper narratives of deadly force. *Journals.Sagepub.Com*, 14(2), 155–

182. <https://doi.org/10.1177/1362480609351545>
- HRW. (2012). *The Case against Killer Robots*. Washington. Retrieved from <https://www.hrw.org/report/2012/11/19/losing-humanity/case-against-killer-robots>
- HRW. (2018). *US: Criminal Justice System Fuels Poverty Cycle*. Washington.
- HUGHES, T. (1994). Technological Momentum. In *Does Technology Drive History?: The Dilemma of Technological Determinism* (p. 101). Cambridge, Massachusetts: Mit Press.
- IBORDERCTRL. (n.d.). Retrieved April 16, 2019, from <https://www.iborderctrl.eu/>
- IRIGARAY, L. (2010). *Speculum. L'altra donna*. Universale economica.
- ITU and XPRIZE. (2017). *AI for Good Global Summit*. Geneva. Retrieved from https://www.itu.int/en/ITU-T/AI/Documents/Report/AI_for_Good_Global_Summit_Report_2017.pdf
- JASANOFF, S. (2016). *The Ethics of Invention: Technology and the Human Future*. WW Norton & Company.
- JASANOFF, S. (2009). *Science at the bar: Law, science, and technology in America* (1st ed.). Cambridge England etc.: Harvard University Press.
- KOSBA, A., MILLER, A., SHI, E., Z. I. symposium (2016). Hawk: The blockchain model of cryptography and privacy-preserving smart contracts. In *IEEE Symposium on Security and Privacy (SP)*. <https://doi.org/10.1109>
- LEE, C. (2008). The Gay Panic Defense. *U.C. Davis Law Review*, 42.
- LI, X., HESS, T. J., & VALACICH, J. (2008). Why do we trust new technology? A study of initial trust formation with organizational information systems. *The Journal of Strategic Information Systems*, 17(1), 39–71.
- LOCKE, T. (1999). Participation, inclusion, exclusion and netactivism: How the Internet invents new forms of democratic activity. In B. D. Barry Hague (Ed.), *Digital democracy: Discourse and decision making in the Information Age* (pp. 211–223). London: Routledge.
- MADHAVAN, P., & WIEGMANN, D. A. (2007). Similarities and differences between human–human and human–automation trust: An integrative review. *Theoretical Issues in Ergonomics Science*, 8(4), 277–301.
- MANDALIOS, J. (2000). Being and Cultural Difference: (Mis)Understanding Otherness in Early Modernity. *Thesis Eleven*, 62(1), 91–108.
- MARTINEZ- QUIRANTE, R., & RODRÍGUEZ-ÁLVAREZ, J. (2020). El lado oscuro de la Inteligencia artificial . *IDEES*, (48). Retrieved from <https://revistaidees.cat/es/el-lado-oscuro-de-la-inteligencia-artificial/>
- MBEMBE, A. (2011). *Necropolítica*. Barcelona: Melusina.
- NOBLE, S. U. (2016). *Algorithms of oppression : how search engines reinforce racism*.
- O'NEIL, C. (2017). *Weapons of math destruction : how big data increases inequality and threatens democracy*. Broadway Books.
- OGBURN, W. F. (1922). *Social Change With Respect to Nature and Original Change*. Huebsch.
- POSTMAN, N. (2006). *Amusing ourselves to death: Public discourse in the age of show business* (20th ed.). London: Penguin.
- POSTMAN, N. (2011). *Technopoly: The surrender of culture to technology*. Nueva York:

Vintage Books.

RAND, A. (2020). Aynrand.org.

RODRÍGUEZ-ÁLVAREZ, J. (2019). Rethinking Paradigms in the Technolo-ecological Transition. *Journal of Law*, 1(2), 169–184.

RODRÍGUEZ-ÁLVAREZ, J., & MARTINEZ-QUIRANTE, R. (2019). *Towards a new AI race. The challenge of lethal autonomous weapons systems (Laws) for the United Nations*. Aranzadi - Tomson Reuters.

RODRÍGUEZ, J. (2016). *La civilización ausente: Tecnología y sociedad en la era de la incertidumbre* (1st ed.). Oviedo: Trea.

ROFF, H. M. (2014). The Strategic Robot Problem: Lethal Autonomous Weapons in War. *Journal of Military Ethics*, 13(3), 211–227.

RUSSELL, S. J. (STUART J., NORVIG, P., & DAVIS, E. (2010). *Artificial intelligence : a modern approach* (3rd Editio). New York: Pearson.

SALERNO, J. M., NAJDOWSKI, C. J., & BOTTOMS, B. L. (2015). Excusing Murder? Conservative Jurors' Acceptance of the Gay-Panic Defense Effects of Race and Socioeconomic Status on Medical Misdiagnosis of Child Abuse View project Miscarriages of Justice Special Issue of Albany Law Review View project. *Psycnet.Apa.Org*, 21(1), 24. <https://doi.org/10.1037/law0000024>

SCHARRE, P., & NORTON, W. W. (2018). *Army of None: Autonomous Weapons and the Future of War Army of None: Autonomous Weapons and the Future of War Arms Control Today*. Retrieved from <https://www.armscontrol.org>

SHARKEY, N. (2010). Saying 'No!' to Lethal Autonomous Targeting. *Journal of Military Ethics*, 9(4), 369–383.

SHARKEY, N. (2014). Towards a principle for the human supervisory control of robot weapons. *Politica & Società*, III(2/2014), 305–324. <https://doi.org/10.4476/77105>

SHARKEY, N. (2018). Mama Mia It's Sophia: A Show Robot Or Dangerous Platform To Mislead?

SKITKA, L. J., MOSIER, K. L., & BURDICK, M. (1999). Does automation bias decision-making? *International Journal of Human Computer Studies*, 51(5), 991–1006. <https://doi.org/10.1006/ijhc.1999.0252>

SMITH, M., & MARX, L. (1994). *Does technology drive history?: The dilemma of technological determinism* (1st ed.). Boston: Mit Pr; Edición.

SNOW, J. (2018, July 26). Amazon's Face Recognition Falsely Matched 28 Members of Congress With Mugshots | American Civil Liberties Union. 10 de Junio de 2020, de<https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28>

STATT, N. (2018). Amazon told employees it would continue to sell facial recognition software to law enforcement - The Verge. Retrieved April 17, 2019, from <https://www.theverge.com/2018/11/8/18077292/amazon-rekognition-jeff-bezos-andrew-jassy-facial-recognition-ice-rights-violations>

STATT, N. (2020, June 10). Amazon bans police from using its facial recognition

- technology for the next year - The Verge. Retrieved June 11, 2020, from <https://www.theverge.com/2020/6/10/21287101/amazon-rekognition-facial-recognition-police-ban-one-year-ai-racial-bias>
- TOFFLER, A. (1984). *Future shock*. Bantam.
- TVERSKY, A., & KAHNEMAN, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131.
- WALLERSTEIN, I., & BALIBAR, E. (1991). *Raza, nación y clase*. Barcelona: IEPALA.
- WAREHAM, M. (2018). “Want to worry about AI? Then worry about this..” Humans Right Watch. Retrieved April 15, 2019, from <https://www.hrw.org/news/2018/03/07/want-worry-about-ai-then-worry-about>
- WEITZER, R., & TUCH, S. A. (2002). Perceptions of racial profiling: Race, class, and personal experience. *Criminology*, 40(2), 435–456.
- WELCH, K. (2007). Black Criminal Stereotypes and Racial Profiling. *Journal of Contemporary Criminal Justice*, 23(3), 276–288.
- WIENER, J. L., & MOHR, L. A. (1994). Technopoly: The Surrender of Culture to Technology. *Journal of Public Policy & Marketing*, 13(2), 326–327. <https://doi.org/10.1007/BF02948575>
- YAMPOLSKIY, R. V. (2013). *Attempts to Attribute Moral Agency to Intelligent Machines are Misguided*.
- YONG, E. (2018, January 17). A Popular Algorithm Is No Better at Predicting Crimes Than Random People - The Atlantic. Retrieved June 11, 2020, from <https://www.theatlantic.com/technology/archive/2018/01/equivant-compas-algorithm/550646/>
- ZIZEK, S. (1997). The Supposed Subjects of Ideology. *Critical Quarterly*, 39(2), 39–59.