

## REFLEXIONES SOBRE LA INTELIGENCIA ARTIFICIAL APLICADA AL DERECHO Y EL DERECHO DE LA INTELIGENCIA ARTIFICIAL: ¿VAMOS HACIA EL MUNDO DE *BLACK MIRROR*?

### *REFLECTIONS ON ARTIFICIAL INTELLIGENCE APPLIED TO LAW AND THE LAW OF ARTIFICIAL INTELLIGENCE: ARE WE HEADING TOWARDS THE WORLD OF BLACK MIRROR?*

ENRIQUE CÁCERES NIETO<sup>1\*</sup>

---

**RESUMEN:** El objetivo de este trabajo es proporcionar una panorámica sobre los dos aspectos de la relación entre la inteligencia artificial (IA) y el derecho: la inteligencia artificial aplicada al derecho (IAD) y el derecho de la inteligencia artificial. Consta de tres grandes bloques: 1) una breve historia de la inteligencia artificial; 2) Los temas de la IAD y la exposición de algunos desarrollos en correspondientes tanto al enfoque *top down*, como al *bottom up* que permiten mostrar porque la IAD debe considerarse una nueva área de la Teoría General del Derecho. El tercer bloque versa sobre los peligros de la IA, con énfasis especial en los riesgos para la democracia, la paz y la libertad. El trabajo concluye con una breve consideración sobre ética, derecho e IA y concluye con una breve reflexión sobre la necesidad de reinventar el derecho para afrontar los retos de la IA.

**PALABRAS CLAVE:** Inteligencia Artificial y Derecho; Teoría General del Derecho; Filosofía del Derecho Constructivismo Jurídico

---

<sup>1</sup> \* Investigador de Tiempo Completo en el Instituto de Investigaciones Jurídicas y asociado en el Centro de Ciencias de la Complejidad, ambos de la UNAM. Investigador nacional nivel III. Correo electrónico: [encacer@hotmail.com](mailto:encacer@hotmail.com). ORCID: <https://orcid.org/0000-0001-7196-7970>

**ABSTRACT:** This paper aims to provide an overview of the double relationship between artificial intelligence and law: artificial intelligence applied to law and the law of artificial intelligence. It consists of three main blocks: 1) a brief history of artificial intelligence; 2) the issues of AI and the exposition of some developments in the top-down and bottom-up approaches that show why AI should be considered a new area of the General Theory of Law. The third section deals with the dangers of AI, with special emphasis on the risks to democracy, peace and freedom. The paper concludes with a brief consideration of ethics, law and AI and ends with a short reflection on the need to reinvent law to meet the challenges of AI.

**KEYWORDS:** Artificial Intelligence and Law; General Legal Theory; Philosophy of Law; Legal Constructivism

## I. INTELIGENCIA ARTIFICIAL (IA) HOY

El advenimiento masivo de la inteligencia artificial hace que esté presente en todos los terrenos: vehículos autónomos, robots asistenciales, la determinación de la responsabilidad por consecuencias jurídicas de acciones realizadas por robots, diagnóstico médico realizado por Watson, el sistema desarrollado por IBM, el surgimiento de Ross (un abogado robot contratado por despachos jurídicos), androides conferencistas con nacionalidad jurídica, etcétera. Estas son solo algunas de las turbulencias que han colocado al sistema social y al jurídico como un sistema alejado del estado de equilibrio.

Sin duda, la inteligencia artificial hace nuestra vida cotidiana más fácil a través de teléfonos inteligentes que aprenden las expresiones o frases que empleamos de manera frecuente para comunicarnos y sugerirnos cual utilizar o sustituir; facilita la seguridad de acceso a nuestros dispositivos a través nuestros datos biométricos;

llegar a nuestro destino gracias a aplicaciones como *Google Maps* que incluso nos indica las rutas más rápidas; modelar la propagación de enfermedades; identificar las zonas con el mayor índice de criminalidad y un largo etcétera.

Sin embargo, como suele suceder en el ámbito científico y tecnológico la inteligencia artificial también puede traer efectos colaterales indeseables. Con la llegada de la minería de datos se incrementó la capacidad de manejar volúmenes de información en formas inimaginables a través de la combinación de información en la red, independientemente de su formato, es decir sin importar si son datos numéricos, alfanuméricos, imágenes o sonidos para realizar clasificaciones predictivas susceptibles de ser empleadas para manipular nuestros procesos mentales y nuestro comportamiento, sin que seamos conscientes de ello; esto ha mostrado ser un riesgo para la democracia, como ha sucedido en el caso del sistema de *Cambridge Analytics* (Kaiser, 2019); es una constante fuente de riesgo para el derecho a la privacidad, e incluso para la seguridad nacional.

Como es fácil observar, la revolución tecnológica de la inteligencia artificial implica también una revolución social que, por tanto, no puede quedar al margen del derecho.

El objetivo de este trabajo es brindar una panorámica muy general sobre las dos caras de la relación entre inteligencia artificial y derecho: la inteligencia artificial aplicada al derecho y el derecho de la inteligencia artificial.

El trabajo se puede dividir en tres grandes bloques de ideas: consideraciones generales e introductorias acerca de la inteligencia artificial, incluyendo una breve referencia histórica; la inteligencia artificial aplicada al derecho y finalmente el derecho de la inteligencia artificial. Su finalidad es la de proporcionar un primer andamio para quienes estén interesados en continuar investigando sobre esta novedosa y fascinante relación inter y transdisciplinaria.

## II. INTRODUCCIÓN: BREVE HISTORIA DE LA IA

Por la expresión ‘Inteligencia artificial’ podemos entender la rama de la ciencia computacional dedicada al desarrollo de programas cuyos productos finales, de ser atribuibles a un humano, presupondrían procesos mentales inteligentes.

Por novedosa que parezca, la inteligencia artificial es el producto de conocimientos acumulados a lo largo de siglos, empezando por la lógica Aristotélica, en el siglo III A.C. (Belda, 2017) el surgimiento del concepto de base de datos en el siglo V A.C por parte de Isidoro de Sevilla; la construcción de la *Ars Magna*, una especie de máquina pensante que mediante diales palancas y manivelas podía realizar demostraciones lógicas con sujetos y predicados de teorías teológicas, inventada por Ramón Lulio en el siglo VIII; la primer calculadora creada por el matemático inglés Charles Babbage en el siglo XVIII, inspirado en el uso de tarjetas perforadas inventadas por Joseph Marie Charles para programar una tejedora; la determinación de las bases de la lógica de primer orden por parte de George Boole y Augustus Morgan en contraposición a la lógica clásica aristotélica en el siglo XIX (Belda, 2017, p. 224); el nacimiento de los conceptos fundamentales de la inteligencia artificial por parte de Alan Turing en el siglo XX, quien además hizo aportaciones sumamente importantes en teoría de la computación, la arquitectura de las computadoras, el criptoanálisis y la formulación del famoso *test* que lleva su nombre, conforme al cual si una máquina se comporta en todos los aspectos como un ente inteligente y pasa desapercibida por un humano, quien considera que sus operaciones han sido realizadas por un humano, entonces debe ser considerada inteligente (Belda, 2017). En 1956, dos años después de la muerte de Turín, tuvo lugar un paradigmático encuentro en la universidad de Dartmouth College en New Hampshire (Belda, 2017) conocido como la reunión de Dartmouth. Organizada por John McCarthy, congregó a 11 de las mentes más brillantes de la época, entre las

que se encontraban el matemático Marvin Minsky, creador de la expresión ‘inteligencia artificial’; el psicólogo Frank Ros, pionero de las redes neuronales; Herbert Simon, premio nobel de economía; Claude Shannon considerado el padre de la teoría de la información y Alain Newell, experto en informática y psicología cognitiva. Como resultado de esa reunión, que duró dos meses, se consolidó la columna vertebral de la inteligencia artificial contemporánea, con dos enfoques diferentes y contrapuestos: el enfoque simbólico y el enfoque conexionista, también conocidos como enfoque *top down* y *bottom up*, respectivamente. La visión *top down*, encabezada por Newell y Simon, consideraba que la cognición era un fenómeno de alto nivel, independiente del soporte en el que ocurra. En este sentido, el cerebro no es más que un soporte que puede reemplazarse por tecnologías de silicio. Para este grupo la inteligencia artificial se debe avocar a desarrollar sistemas artificiales capaces realizar el mismo tipo de operaciones simbólicas que el cerebro humano, sin pretender reproducir su estructura.

El grupo que defendía el enfoque *bottom up*, encabezado por Rosenblatt, sostenía lo contrario, es decir, que la estructura del cerebro es una condición necesaria para que pueda tener lugar la emergencia de procesos cognitivos y, por tanto, es imprescindible crear un *hardware* que emule la biología cerebral. Este enfoque, congruente con la visión de Turín, es la base del modelo conexionista en el que se fundamentan los sistemas *Deep* y *Machine Learning* característicos de inteligencia artificial contemporánea.

El auge inicial de ambos enfoques se vio paralizado durante el período de 1970 a 1985 debido a diversos problemas en su desarrollo, que incidieron en la falta de apoyos financieros. Sin embargo, poco a poco fueron surgiendo nuevos avances tanto teóricos como tecnológicos que produjeron un renovado interés por la Inteligencia Artificial con los consiguientes financiamientos.

Con la finalidad de romper la asociación a los intentos infructuosos de la primera etapa y el fracaso de un sistema llamado *General Solving Problem* pensado para resolver cualquier problema que se le planteara, independientemente del dominio, los partidarios del enfoque *top down* comenzaron a desarrollar sistemas expertos: programas basados en la elicitación y representación del conocimiento de uno o más expertos, con el fin de que la computadora emule sus procesos de razonamiento y toma de decisión. Uno de esos sistemas con resultados importantes fue Mycin, apto para ayudar al diagnóstico y tratamiento médico de enfermedades gastrointestinales a un médico no especialista (Belda, 2017).

El enfoque conexionista también siguió avanzando hasta llegar al desarrollo de las redes neuronales, que son programas basados en la arquitectura y dinámica de las redes neuronales humanas.

Sin embargo, tras sólo dos años después de este apogeo la inteligencia artificial entró en un nuevo período de glaciación debido a nuevos problemas tanto económicos como técnicos y, sobre todo, falta de satisfacción de las expectativas tan elevadas que se habían generado alrededor de ella.

Finalmente la IA despegó para alcanzar los vuelos que tiene actualmente debido a nuevos avances científicos y tecnológicos que incluyen el surgimiento de la nube (Belda, 2017), el *Big data* y la minería de datos, que hacen posible el manejo de gigantescos números de datos, independientemente de su formato a diferencia de las bases de datos relacionales tradicionales y a una velocidad de procesamiento nunca antes vista, todo lo cual hizo posible que irrumpiera en todos los aspectos de nuestra vida, para bien y...para mal.

### **III. LA INTELIGENCIA ARTIFICIAL APLICADA AL DERECHO Y EL DERECHO DE LA INTELIGENCIA ARTIFICIAL**

De la misma manera que sucede con otras áreas trans e interdisciplinariedad en el derecho, su relación con la inteligencia artificial presenta dos caras: una en la que el derecho es el dominio al cual se aplican conocimientos de otra disciplina, como ocurre con la informática jurídica o las neurociencias aplicadas a la investigación cognitiva de los operadores jurídicos. Otra en la que el derecho regula la actividad de la otra disciplina, como ocurre con derecho de la informática y del derecho de las neurociencias. En este orden de ideas, por tanto, se debe distinguir entre la inteligencia artificial aplicada al derecho, por una parte, y el derecho de la inteligencia artificial, por la otra, el cual, a su vez, debe distinguirse de la ética de la inteligencia artificial. De abordar estos temas, nos ocuparemos a continuación.

### **IV. LA INTELIGENCIA ARTIFICIAL APLICADA AL DERECHO (IAD)**

Puede definirse como una trans e interdisciplina dedicada al desarrollo de programas cuyos productos finales que, de ser atribuidos a un humano, supondrían procesamiento inteligente por parte de un operador jurídico. Dado su carácter trans e interdisciplinario, la comunidad epistémica encargada de su estudio incluye a especialistas de las más diversas disciplinas: filósofos, científicos de datos, matemáticos, lingüistas, físicos, lógicos, especialistas en ciencias de la computación, etcétera.

La diversidad de temas de investigación de la IAD se puede dividir en dos grandes bloques: el que comprende un carácter meramente teórico (como ocurre con las investigaciones sobre lógica y derecho) y otro de carácter práctico. Un ejemplo de esa pluralidad de temas es el hecho de que se haya aplicado en los siguientes escenarios:

Apoyo a los abogados en el razonamiento jurídico; la redacción de documentos; negociación; producción y gestión de la legislación; análisis de políticas; gestión del flujo de trabajo; seguimiento de la aplicación de la ley; análisis de las pruebas; gestión de los casos; investigación policial y forense; búsqueda y evaluación de pruebas; gestión y recuperación de información; producción de conocimiento jurídico a través de la elaboración de ontologías; enseñanza del derecho a través de casos; diagramas de argumentación jurídica; modelado de interacciones jurídicas de agentes autónomos e instituciones digitales; gestión del cambio organizativo al introducir sistemas de conocimiento jurídico.

Sería imposible dar cuenta en este trabajo de esta diversidad de proyectos y menos aún describir sus bases teóricas, metodológicas y tecnológicas. En lugar de eso expondré brevemente algunos proyectos con el mayor detalle posible dentro de los límites de esta investigación.

## **V. INTELIGENCIA ARTIFICIAL APLICADA AL DERECHO DESDE EL ENFOQUE SIMBÓLICO O *TOP DOWN*: EL CASO DEL SISTEMA EXPERTIUS I**

La historia de la inteligencia artificial aplicada al derecho (IAD) corre paralela a la de la inteligencia artificial general. Cuenta con desarrollos realizados desde el enfoque simbólico *top down* mediante la realización de sistemas expertos, así como de inteligencia artificial *bottom up*.

Un aspecto importante, que suele pasar desapercibido, es que todo proyecto de IAD requiere, como condición necesaria, la definición de un problema relevante para el derecho y el desarrollo de una teoría jurídica computacional. Es decir, una teoría que genere marcos conceptuales y metodológicos susceptibles de servir de base para un desarrollo computacional. En este sentido, la IAD puede considerarse una novedosa área de la teoría general del derecho.

Expertus I es un prototipo de sistema experto basado en el paradigma simbólico o *top down*. Ha sido desarrollado en el Instituto de Investigaciones Jurídicas y el Centro de Ciencia Aplicada y Desarrollo Tecnológico (ambos de la UNAM), bajo el auspicio del Consejo Nacional de Ciencia y Tecnología y en colaboración del Tribunal Superior de Justicia del estado de Tabasco. Su finalidad es ayudar a la toma de decisiones y a la homogenización del conocimiento colectivo de la comunidad judicial, en el dominio del juicio especial de alimentos (Cáceres, 2008) la interfaz de entrada al sistema se muestra en la ilustración 1.

Planteamiento del problema: parte de una clasificación posible de problemas que distingue entre aquellos: 1) convergentes tipo 1, cuyas características son: a) cuentan con una solución única y b) el estado de solución está predeterminado, es decir, es conocido antes de emprender la búsqueda de su solución, por ejemplo, armar un rompecabezas; 2) convergentes tipo 2, con las siguientes características: a) pueden tener una solución única y b) el estado de solución no es determinado inicialmente, pero es determinable. Por ejemplo, los cálculos matemáticos; y, 3) Problemas divergentes cuyas propiedades son: a) la existencia múltiples soluciones plausibles para un mismo problema y b) su solución es razonablemente determinable. A esta categoría corresponden las decisiones judiciales, objeto del sistema.

En congruencia con lo anterior, para los efectos de Expertus I, el poder judicial es considerado una mente colectiva paraconsistente, basada en el principio de individualización cognoscitiva,

conforme al cual, cada uno de los operadores judiciales posee sus propios modelos mentales y, por tanto, puede decidir de manera diferente a como lo harían el resto sus colegas. La paraconsistencia tiene lugar a nivel individual cuando un mismo juzgador puede decidir de manera distinta casos de la misma clase en momentos distintos de su vida, y de manera colectiva cuando esa diversidad de soluciones sucede a nivel institucional pudiendo dar lugar respuestas contradictorias para el mismo tipo de problema. Tomando un término proveniente de la cibernética, la paraconsistencia es fuente de entropía, entendida como el grado de incertidumbre en el comportamiento del sistema.

Considerar a la actividad judicial como el resultado de una mente colectiva reconoce que en la práctica distintos jueces pueden enfrentar problemas diferentes a los que han tenido que resolver sus colegas i.e., que son fuente de conocimiento heurístico que debe ser compartido con el resto de los miembros de la comunidad judicial con el fin de optimizar esfuerzos y “neguentropizar” al sistema.

Los objetivos de Expertus I son: ayudar a los jueces a tomar decisiones en materia de pensión alimenticia y servir como medio para compartir las estrategias de decisión entre los miembros de la comunidad judicial, propiciando trabajo de cognición grupal coherente y estable.

La caracterización académica de Expertus I es la siguiente:

Epistémicamente parte de los supuestos del constructivismo jurídico: metodológicamente tuvo que resolver los dos problemas característicos de este tipo de sistemas: a) la elicitación del conocimiento experto y b) su representación. La solución a dichos problemas se obtuvo mediante dos estrategias: una intra -sistémica consistente el uso de técnicas de análisis de tareas cognitivas, entrevistas a expertos y *shadowing self*; y una extrasistémica consistente en el análisis de cerca de 400 expedientes.

La teoría jurídica computacional consistió en la creación de tablas de representación de oposición dialógica y derrotante (TODD) de los argumentos entre las partes contendientes, incluyendo el proceso de asignación de pesos epistémicos a las pruebas atómicas, así como el de ponderación probatoria. Gracias al uso de las TODD se llevó a cabo la representación del conocimiento de cada uno de los jueces que participaron en el proyecto (análisis intra sistémico) de manera homogénea con el análisis de los expedientes (análisis extra sistémico), esto puede apreciarse en la ilustración 2.

El diseño del sistema consta de los siguientes módulos: un módulo inferencial, encargado de realizar las operaciones lógicas necesarias para obtener la solución del caso; un módulo tutorial que provee al juez de la información necesaria para responder las preguntas del sistema (ver ilustración 3); y, finalmente, un módulo financiero, cuyo objetivo es contribuir a la determinación del monto de la pensión alimenticia, que resultó ser especialmente útil en la fijación de la pensión provisional que debe tomarse al inicio del proceso, cuando aún no se tiene conocimiento del estado financiero de las partes, como indica la ilustración 4.

A pesar de que el desarrollo del prototipo tuvo lugar en el juicio especial de alimentos, tanto el desarrollo teórico como el metodológico son aplicables a cualquier área del derecho.

**ENTIDADES PARTICIPANTES**

EXPERTIUS es un sistema de apoyo a la toma de decisiones judiciales en el dominio del Juicio Especial de Alimentos. Ha sido desarrollado por el Departamento de Inteligencia Artificial Aplicada al Derecho (IA y D) del Instituto de Investigaciones Jurídicas de la UNAM, con el apoyo del CONACYT y la importante colaboración del Tribunal Superior de Justicia del Estado de Tabasco. El conocimiento experto elicitado de los funcionarios judiciales adscritos a este órgano jurisdiccional, en conjunción con las bases teóricas y metodológicas diseñadas ex profeso para este proyecto, constituye uno de los aspectos esenciales del sistema.

EXPERTIUS está conformado por tres módulos básicos: El módulo "Tutorial", que proporciona asesoría para que el juzgador esté en mejores condiciones de resolver los distintos problemas que en la práctica se presentan, vinculados con las múltiples tareas cognitivas que caracterizan a cada fase y sub-fase del juicio especial de alimentos; El módulo "Inferencia", que proporciona asesoría relativa a la resolución de los problemas que plantea la pederación probatoria; Y el módulo "Financiero", que proporciona asesoría relativa a la determinación de los montos correspondientes a las pensiones provisional y definitiva.

EXPERTIUS es un sistema único en su género, y es el resultado de una investigación de vanguardia a nivel internacional. El marco teórico y metodológico que lo soporta ha sido enriquecido y validado en los diversos foros mundiales organizados por la comunidad de la IA y D, cuyos miembros más sobresalientes, como Kevin Ashley (Estados Unidos), Bernhard Schafer (Ecuador), John Zelezniak (Australia), Thomas Gordon (Alemania), Radboud Winkels y Tom Van Engers (Holanda), han contribuido de manera importante.

**LIGAS RELACIONADAS**

**Rationale**

**Código Hipertextual**

**Módulo de Código Hipertextual**

**Base de Casos**

**CONTACTÉNDOS**

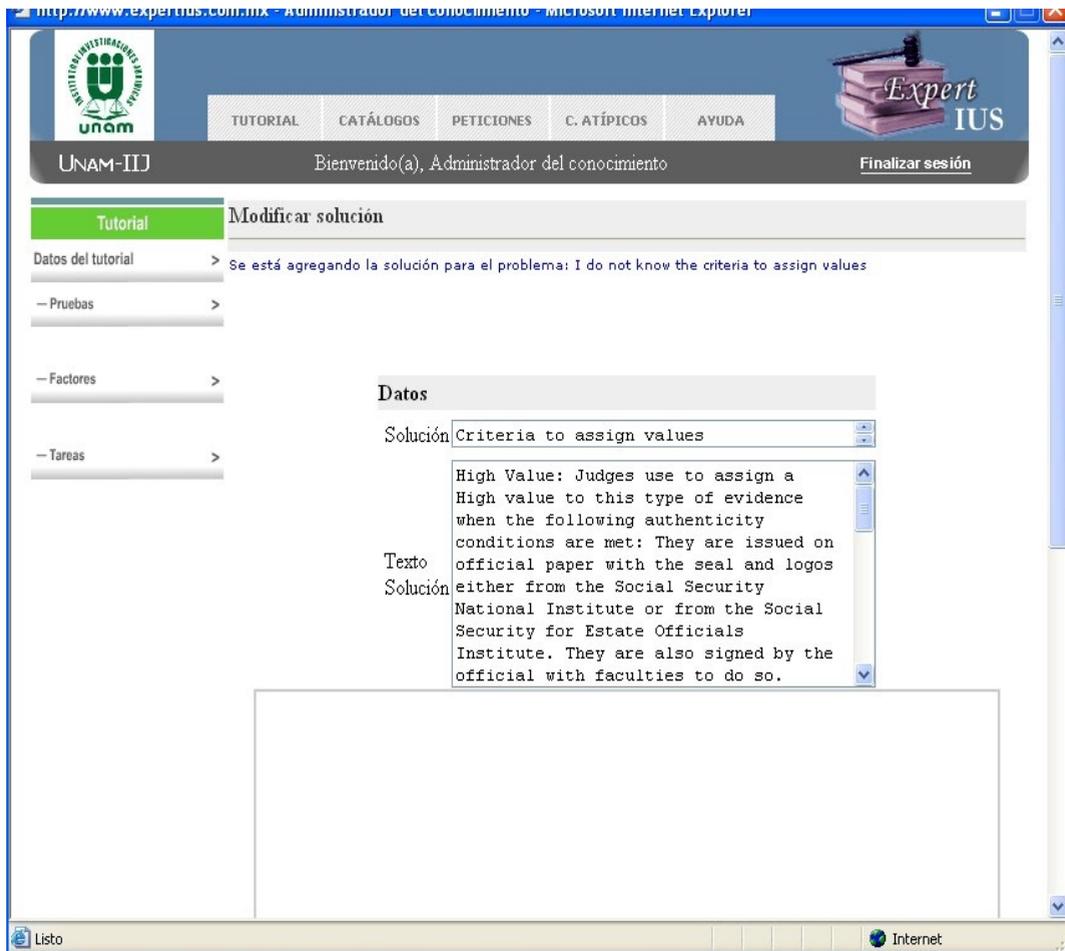
e-mail

IIJ - Unam © 2007 | Conacyt  
Términos de uso | Políticas de Privacidad | Contáctenos

***Ilustración 1. Pantalla principal de Expertius I.***

T	EAH Actor	Sujeto	EAH Demandado	T
PP 1		La firma en el acta de matrimonio, en efecto corresponde a la del demandado. The signature on the marriage certificate in effect corresponds to the signature of the respondent	The signature on the marriage certificate that the plaintiff has produced is not mine <b>Valor: Derrotado</b>	PP 1
MP 1	The testimony of an expert in graphology <b>Valor: Alto</b> <b>Factor:</b> Valor probatorio <b>Valor: Alto</b>			
PP 2		The respondent is my husband <b>Valor: Alto</b>		
MP 2	An official marriage certificate that states that the plaintiff and the respondent are married <b>Valor: Alto</b> <b>Factor:</b> Valor probatorio <b>Valor: Alto</b>			
PP 3		De dicha unión procreamos un hijo (s). We had a child under our marriage		

*Ilustración 2. Pantalla de tabla de oposición dialógica y derrotante elaborada por el propio sistema. Sobre ella se realiza el proceso de ponderación probatoria a partir de los pesos asignados a las pruebas y contrapruebas atómicas.*



The screenshot displays the Expert IUS web application interface. At the top, there is a navigation menu with options: TUTORIAL, CATÁLOGOS, PETICIONES, C. ATÍPICOS, and AYUDA. The user is logged in as 'UNAM-III' and is the 'Administrador del conocimiento'. The main content area is titled 'Modificar solución' under the 'Tutorial' section. A message indicates that a solution is being added for a problem: 'I do not know the criteria to assign values'. The 'Datos' section shows the solution name 'Criteria to assign values'. The 'Texto' section contains the following text: 'High Value: Judges use to assign a High value to this type of evidence when the following authenticity conditions are met: They are issued on official paper with the seal and logos either from the Social Security National Institute or from the Social Security for Estate Officials Institute. They are also signed by the official with faculties to do so.' The interface includes a sidebar with expandable sections for 'Pruebas', 'Factores', and 'Tareas'. The bottom of the browser window shows the status bar with 'Listo' and 'Internet'.

***Ilustración 3. Pantalla correspondiente al módulo tutorial que indica las condiciones bajo las cuales se puede atribuir cierto peso epistémico a una prueba atómica.***

http://www.expertius.com.mx - Expert IUS - Microsoft Internet Explorer

UNAM-IIJ Bienvenido(a), Operador judicial Finalizar sesión

Módulo Financiero

Número de expediente:

	Actor		Demandado	
Ingresos	\$ 0	<input type="button" value="Modificar"/>	\$ 11000	<input type="button" value="Modificar"/>
Necesidades	\$ 7762.38	<input type="button" value="Modificar"/>	\$ 5801.47	<input type="button" value="Modificar"/>

Diferencia por cubrir: \$ 7762.38

Asignación al Demandado: \$   Cantidad  Porcentaje

Asignación al Actor: \$

IIJ - UNAM - CONACYT © 2008  
[Ayuda](#) | [Créditos](#) | [Contáctenos](#)

***Ilustración 4. Pantalla correspondiente al módulo financiero desarrollado con base en un estudio empírico para la determinación de la pensión, dependiendo de la información disponible sobre el demandado.***

Entre los inconvenientes que se suelen predicar de este tipo de sistemas expertos es que pueden quedar obsoletos fácilmente cuando ocurren reformas legislativas. Esta crítica es parcialmente correcta en el caso de reformas, si tomamos en cuenta que el conocimiento heurístico compartido respecto de los problemas correspondientes a las normas no derogadas seguirá siendo válido.

Entre sus ventajas se encuentra su transparencia ya que permite conocer la estrategia de decisión del sistema, lo que contrasta con los sistemas basados en redes neuronales, incluyendo *deep* y *machine learning*, en los cuales las operaciones realizadas por el sistema quedan dentro de una caja negra dado que es imposible de conocer y por tanto monitorear el proceso de “razonamiento” de la máquina.

## VI. SISTEMAS DE INTELIGENCIA ARTIFICIAL APLICADA AL DERECHO DESDE EL ENFOQUE BOTTOM UP

### A. *Expertius II*

Este enfoque estuvo basado originalmente en modelos de redes neuronales o conexionistas, que siguen jugando un papel fundamental a través de sistemas tipo *deep* y *machine learning*. Sin embargo, hoy día existen otros marcos teóricos y metodológicos para hacer inteligencia artificial de este tipo, como es el caso de los algoritmos genéticos o los sistemas basados en probabilidad y redes bayesianas.

Siguiendo estas ideas, la inteligencia artificial puede ser dividida en dos grandes grupos, no incompatibles con la clasificación clásica entre simbólica y conexionista y más comprensible que éstas: inteligencia artificial representacional e inteligencia artificial no representacional.

La primera es plenamente compatible con el enfoque simbólico y requiere de la comprensión y modelado de la manera en que los expertos humanos toman sus decisiones. Tal como vimos al exponer *Expertius I* presenta la ventaja de permitir monitorear el “razonamiento artificial”. La segunda, compatible con el modelo conexionista y las otras herramientas de la inteligencia artificial referidas al hablar de inteligencia artificial sin representación simbólica no requiere del modelado del razonamiento de los expertos. Se basa en sofisticados modelos matemáticos capaces de procesar enormes cantidades de información contenidas en bases de datos y recurren

técnicas de minería de datos. Es el caso del sistema *deep blue* que derrotó al campeón de ajedrez Kasparov, que no fue resultado del modelado del razonamiento ajedrecístico de alguna persona que pudiera derrotar al campeón del mundo. Como hemos dicho, el inconveniente de este tipo de sistemas es que los humanos no tenemos acceso a los procesos internos de la computadora y, por tanto, no los podemos comprender y menos monitorear. En el caso particular del derecho esto genera serias dificultades al momento de justificar las razones que están detrás de una decisión jurídica artificial. Uno de los retos actuales consiste en la generación de sistemas capaces de operar con los algoritmos de inteligencia artificial que puedan identificar patrones y sean compatibles con la transparencia de la forma en que la máquina toma las decisiones.

Expertus II también fue desarrollado en el Instituto de Investigaciones Jurídicas de la UNAM, específicamente en su Laboratorio de Constructivismo Jurídico, creado bajo el auspicio del programa Fronteras de la Ciencia del Consejo Nacional de Ciencia y Tecnología. Se ubica dentro de los proyectos orientados a hacer compatibles los enfoques representacional y no representacional (Cáceres y Mansilla, 2019).

El problema de investigación es el siguiente: uno de los problemas centrales de la epistemología jurídica aplicada consiste en el desarrollo de modelos de normatividad epistémica que definan los procesos cognitivos que deben ser seguidos por los operadores jurídicos para fundamentar la determinación de la verdad de proposiciones aseverativas de hechos jurídicamente relevantes. Metafóricamente sería equivalente a la metodología de la investigación científica.

Algunas de las propuestas más conocidas sobre el tema son el uso de probabilidad bayesiana y probabilidad subjetiva (inferencias baconianas). Sin embargo, ninguna de ellas da cuenta de los sofisticados procesos de ponderación que tienen lugar sobre los pesos epistémicos de distintas pruebas atómicas de cada una de las partes,

respecto de cada una de las proposiciones aseverativas de hechos, mismas que, a su vez, deben corresponder a cada uno de los elementos de una estructura normativa.

La probabilidad bayesiana determina de manera lineal la probabilidad de que un evento tenga lugar dada la probabilidad de ocurrencia de otras variables que a su vez tienen cierto valor probabilístico determinado a partir de estadística frecuentista. Este algoritmo no puede resolver los problemas derivados de la oposición o reforzamiento que tiene lugar entre los pesos epistémicos de las distintas pruebas ofrecidas por las partes. Algo semejante ocurre con la probabilidad baconiana, un tipo de probabilidad subjetiva, que es más apta para el modelado del razonamiento de un detective que de un juez.

El objeto de la investigación es simular los procesos cognitivos que realizan los jueces al ponderar los distintos pesos probatorios asignados a las pruebas y contrapruebas ofrecidas por cada una de las partes.

La teoría jurídica computacional incluyó un primer modelo conceptual inspirado en el funcionamiento de las redes neuronales artificiales, con base en el cual se procedió a la realización de un trabajo transdisciplinario entre matemáticas y derecho, cuyo resultado fue considerar a la mente judicial como un espacio vectorial.

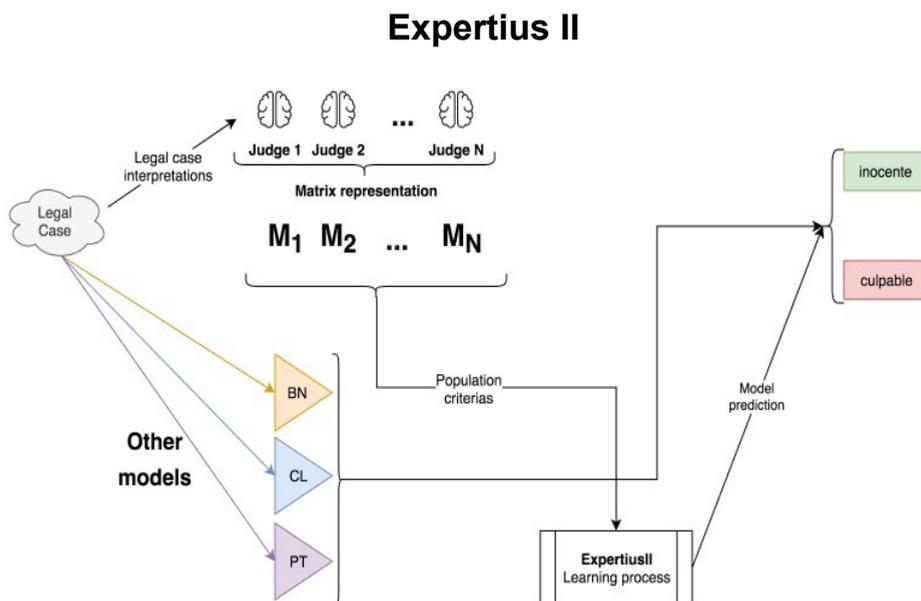
Expertus II puede considerarse un desarrollo más sofisticado que las tablas de oposición dialógica y derrotante elaborado para Expertus I, en el sentido de que el procesamiento de los pesos epistémicos de las pruebas en dichas tablas era poco flexibles y matemáticamente elemental para un sistema que busca modelar los procesos cognitivos de los operadores jurídicos.

Una diferencia fundamental entre Expertus I y Expertus II es que éste ha sido desarrollado partiendo del supuesto de que los procesos cognitivos de los operadores jurídicos son altamente estocásticos y resulta muy difícil, si no imposible, aislar todas las variables que participan en una toma de decisiones judicial. Por esta razón

optamos por la elaboración de un sistema que pudiera identificar los patrones en la toma de decisiones realizadas por una comunidad judicial a la que se le presentan casos de la misma clase, o incluso el mismo caso. En este sentido Expertius II sigue siendo fiel en los supuestos iniciales de Expertius I al considerar la actividad judicial como un proceso de cognición grupal.

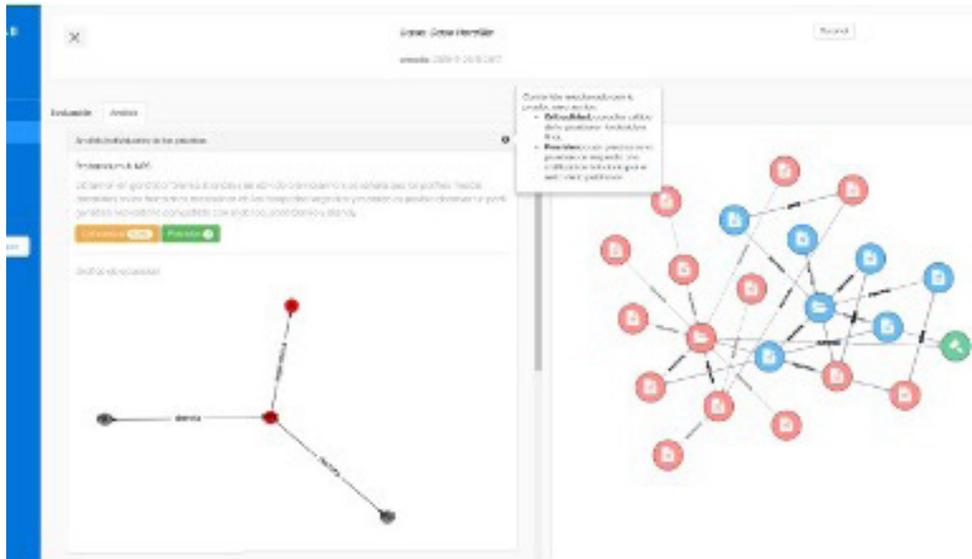
Para la representación gráfica de los patrones de decisión, se optó por la teoría de grafos dirigidos y ponderados, lo que nos permitió representar a las proposiciones aseverativas de hechos como nodos de entrada, cuyo peso epistémico es determinado por el resultado de la ponderación de los pesos epistémicos asignados a cada una de las distintas pruebas atómicas, las cuales, a su vez, son consideradas como nodos de origen en relación con los nodos proposicionales (Ver ilustración 5). De la misma manera, las pruebas atómicas son consideradas nodos destino con respecto a los valores asignados a los parámetros correspondientes a cada uno de los criterios de evaluación especificados para cada tipo de prueba. Dichos parámetros son considerados como nodos de origen con respecto de los nodos correspondientes a cada una de las pruebas atómicas, esto es representado en la ilustración 6.

La gráfica resultante representa el punto de equilibrio resultante de la ponderación vectorial de las pruebas, lo que permite determinar si el sistema satisface o no el estándar probatorio requerido para tomar la decisión final, lo que se muestra en la ilustración 7.



***Ilustración 5. Esquema comparativo de otros modelos de normatividad epistémica y Expertius II.***

## Expertius II: analysis



***Ilustración 6. Grafo emergente del procesamiento de ponderación probatoria. Las flechas corresponden a los arcos que unen a los diferentes nodos, pertenecientes a la oposición dialógica entre las partes, representados con colores diferentes (colores azul y rosa). La decisión final correspondiente al punto de equilibrio vectorial es representada por el mallette coloreado en verde.***

## Expertus II: analysis



***Ilustración 7. Pantalla que muestra la estadística de los patrones de decisión emergentes.***

### ***B. Sistema inteligente para la configuración y visualización del derecho emergente en la era de la globalización***

Planteamiento del problema: Un error epistémico cometido por los filósofos del derecho, que es fuente de inacabables discusiones, consiste en haberse fijado como meta generar teorías que satisfagan las características de las que se ocupan de objetos diferentes al derecho, como es el caso de los objetos matemáticos o geométricos cuyas definiciones son incontrovertibles independientemente de coordenadas espacio-temporales, sin tomar en cuenta que los “objetos” que pretende explicar la teoría del derecho tienen propiedades distintas a las de los matemáticos. Así, mientras la definición de triángulo proporcionada por Euclides permanece inamovible en la filosofía del derecho casi siempre existe una diversidad de teorías en pugna

con el mismo objetivo de dar cuenta de “aquello que hace que el derecho sea lo que es y no otra cosa”.

Una actitud epistémica más sana consiste en asumir que las teorías filosófico-jurídicas constituyen modelos conceptuales que, como cualquier modelo, tratan de sistematizar información importante del dominio que pretenden explicar. Desde esta perspectiva queda claro que la correspondencia entre teoría y dominio, entre *explanas* y *explanandum* debe ser el resultado de un equilibrio reflexivo entre las características de aquello que se pretende modelar y el modelo que lo representa.

Esto es relevante con relación a los modelos kelsenianos relativos a la estructura piramidal de los sistemas normativos, así como a la identificación entre derecho y Estado cuya relación teoría-dominio ha perdido su acoplamiento epistémico debido a que la dinámica de los sistemas normativos ha dejado de corresponder a las propuestas kelsenianas debido al proceso de globalización jurídica. Hoy día ya no es posible identificar con precisión cuáles son el conjunto de normas que forman parte de un sistema dada la diversidad de insumos normativos generados más allá de la producción de los órganos legislativos estatales. Por ejemplo, el *soft law*, las consultas consultivas hechas a organismos internacionales, precedentes internacionales, argumentos tomados por instancias de otros países adaptables al derecho interno, etcétera.

El objetivo del proyecto es modelar y visualizar la conectividad entre resoluciones tomadas por distintos países y la diversidad de insumos jurídicos que sirven para la fundamentación de sus decisiones. Para probar tanto la teoría como la metodología, se tomó como objeto de modelado específico al *ius commune* latinoamericano, es decir, las resoluciones emitidas por los órganos jurisdiccionales de los países pertenecientes al Sistema Interamericano de Derechos Humanos. La finalidad es que el sistema pueda establecer las conexiones entre decisiones y sus fundamentos jurídicos de manera automática, a partir del análisis de los textos digitalizados tanto

de las decisiones como de los insumos normativos empleados para su fundamentación, visualizar la red dinámica emergente de dicha conectividad y hacerla susceptible de medir sus propiedades (Meza et al., 2019).

La teoría jurídica computacional se basó en la teoría hipertextual del derecho y consistió en la elaboración de la categoría discursiva de marcador semántico de conectividad normativa, es decir, aquellas expresiones que cuando aparecen en el texto de una resolución anuncian que a continuación se hará alusión al fundamento correspondiente. Sobre esta base, juntamente con el doctor Iván Vladimir del Instituto de Investigaciones en Matemáticas Aplicadas y Sistemas de la UNAM, aplicando técnicas de *clustering* y teoría de redes se desarrolló un prototipo capaz de generar, representar y medir la dinámica emergente del derecho globalizado, los pasos iniciales para este proceso se hacen en la pantalla mostrada en la ilustración 8. Algunos ejemplos de las métricas de redes reflejadas en la visualización son el cálculo de los *hubs* o los nodos con el mayor grado de conectividad que son los que tienen un mayor grado de influencia sobre la dinámica del sistema, o, por ejemplo, cuáles son las decisiones que tienen menor grado de conectividad y por tanto impactan menos en las decisiones de la comunidad judicial internacional, esto se muestra en la ilustración 9.

La representación de esa conectividad se modifica de manera automática en función de los nuevos casos que se vayan introduciendo a la base, lo que pone de manifiesto que la dinámica del sistema globalizado corresponde a la de un sistema adaptativo complejo, como podemos ver en la ilustración 10.

En términos de inteligencia artificial, el sistema simula los procesos que tendrán lugar en la memoria asociativa de un agente artificial capaz de realizar operaciones de conectividad que ninguna mente humana podría realizar debido a nuestras naturales limitaciones cognitivas.

En términos de teoría general del derecho, la conclusión es que el modelo muestra que hoy día hay un desajuste entre el explanans y el explanandum de la teoría piramidal del derecho y procede ser sustituido por un modelo reticular del derecho el cual, además de dar cuenta de la morfología y dinámica de los sistemas internos, puede insertarlos en el contexto de la globalización del derecho.

sentencias Números **Conectividad** Acerca

### Conectividad Normativa

Visualizador de conectividad de las sentencias de la Corte Interamericana de Derechos Humanos

### Grafos por país

Países  
Argentina

Mínimo de citas  
5

Excluir citas comunes  Expandir artículos  Solo casos de la corte

Dibujar

### Grafos completo

Mínimo de citas  
5

Expandir a artículos  Solo casos de la corte

Grafo sin links comunes

Mínimo de citas  
5

Grafo completo

### Configurar grafo

Expandir a artículos  Solo casos de la corte

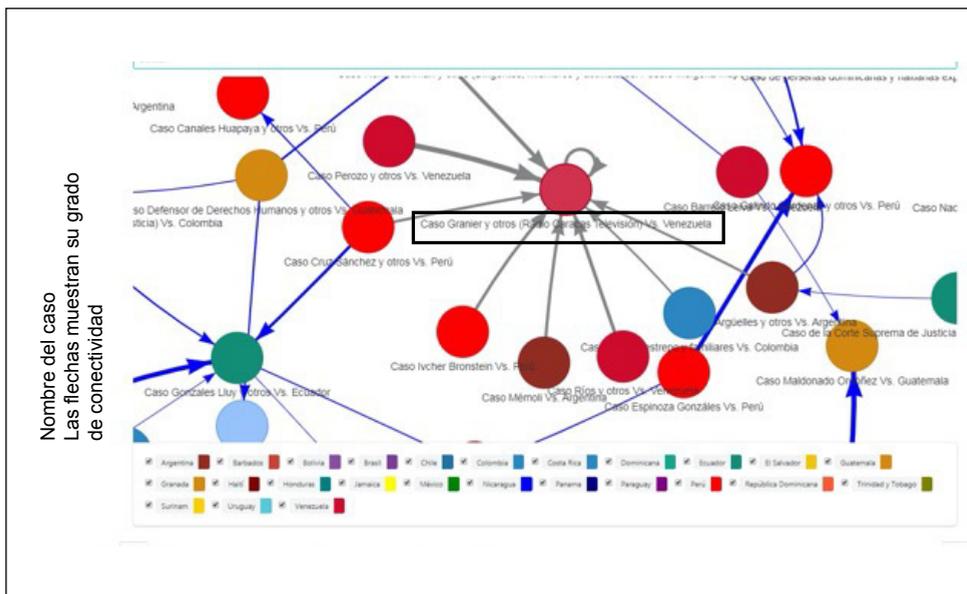
Mínimo de citas  
5

Incluir sentencias  
 Termino a incluir (usar minúsculas)

Excluir sentencias  
 Termino a excluir (usar minúsculas)

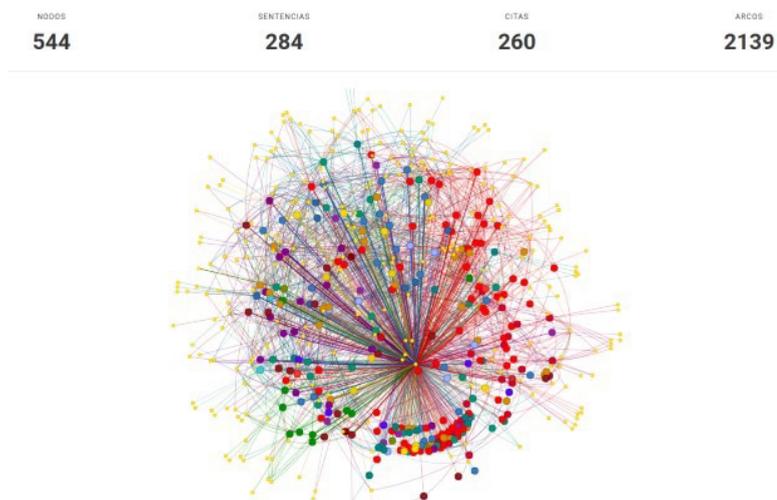
Incluir documentos citados

***Ilustración 8. Pantalla del sistema mostrando la aplicación de la teoría de grafos.***



***Ilustración 9. Representación de una red emergente con los nombres de las disposiciones interconectadas por el sistema.***

# Grafo de conectividad



*Ilustración 10. Representación de una red emergente con los nombres de las disposiciones interconectadas por el sistema.*

## VII. SISTEMAS COMERCIALES DE INTELIGENCIA ARTIFICIAL APLICADA AL DERECHO (ENFOQUE BOTTOM UP)

Además de la investigación académica ligada al desarrollo de teorías jurídicas computacionales, diversos organismos se han abocado al desarrollo de sistemas orientados a auxiliar directamente a los profesionales del derecho. A continuación, se exponen brevemente los más conocidos.

### A. *Ross*

Se trata de un sistema de inteligencia artificial basado en la tecnología de *machine learning* desarrollado por IBM a través de su plataforma Watson, famosa por su capacidad de brindar diagnósticos médicos acertados y haber ganado en el concurso *Jeopardy* en los Estados Unidos. Cuenta con algoritmos para la comprensión de lenguaje natural y se afirma que puede leer 10,000 páginas por segundo y brindar respuesta en tiempo real a la búsqueda de información sobre precedentes, legislación de 9 países, libros, revistas y noticias relevantes para la solución de un caso. También es capaz de predecir y notificar de los riesgos que nuevos precedentes o leyes podrían significar para determinado asunto (*ROSS Intelligence* (“*ROSS*”), 2020).

En 2016 fue contratado por la firma estadounidense *Baker & Hostetler* y desde entonces ha aumentado la demanda de sus servicios por otras firmas de abogados.

### B. *Vincent*

Ha sido desarrollado por el Centro de Innovación de la empresa vLex, a partir de técnicas de *deep learning* y *clustering* (vLex, n.d.).

Cumple algunas funciones semejantes a las de Ross y del Sistema para la Configuración y Visualización del Derecho Emergente en la Era de la Globalización y su objetivo es relacionar los precedentes más relevantes conectados con una sentencia objeto de consulta, representar la relación mediante un grafo dirigido denominado mapa de precedentes y medir el impacto de las citas referenciadas en la argumentación de las decisiones. El tamaño de los nodos que representan a las sentencias es directamente proporcional a su grado de conectividad y por tanto importancia para la decisión consultada.

### C. *Prometea*

Es un sistema tipo *machine learning* supervisado conjuntamente con técnicas de *clustering*, fue desarrollado por el Ministerio Público Fis-

cal de la ciudad de Buenos Aires. Funciona a través de la correlación estadística de palabras clave previamente proporcionadas al sistema (Estevez, Fillotrani y Lejarraga, 2020). Su objetivo es ayudar a los operadores a desahogar el trabajo repetitivo para que puedan dedicar su tiempo a funciones sustanciales. Su operación implica la identificación de patrones de propiedades en una gran cantidad de documentos e identificar su correlación con decisiones específicas.

El sistema ha sido utilizado por la Fiscalía para casos de amparo en materia de derecho a la vivienda, reclamo de taxistas a quienes se les niega la licencia por no satisfacer ciertos requisitos o tener antecedentes penales y localización de fallos archivados de casos similares. También está siendo utilizado por la Corte Constitucional de Colombia para casos que versen sobre salud pública y en la Corte Interamericana de Derechos Humanos en materia de notificaciones relativas a solicitudes de opiniones consultivas (*IA Lab*, 2020).

## VII. EL DERECHO DE LA INTELIGENCIA ARTIFICIAL

Como se dijo en páginas anteriores, la relación entre derecho e inteligencia artificial presenta dos caras: una corresponde a la inteligencia artificial aplicada al derecho, tratada en los apartados anteriores y otra, el derecho de la inteligencia artificial en la que el derecho se encarga de regular las actividades y consecuencias de la inteligencia artificial, con la finalidad de prevenir y contender los posibles efectos sociales negativos que ésta puede llegar a tener.

Con la finalidad de evitar confusiones, es importante distinguir entre la ética de la inteligencia artificial y el derecho de la inteligencia artificial.

Para tales efectos es útil diferenciar entre una concepción intrasistémica y otra extrasistémica del derecho. La primera se refiere a todo aquello que acontece tanto a nivel normativo como en la dinámica social derivadas de la ejecución de las normas jurídicas.

La visión extra sistemática, por otra parte, se refiere a todo aquello que ocurre fuera del marco de intrasistemático, particularmente a los fenómenos sociales que tienen lugar antes de su regulación.

Esta categorización teórica es relevante para comprender, a su vez, dos relaciones entre la ética de la inteligencia artificial aplicada al derecho, una extrasistémica y otra intrasistémica. La primera comprende las discusiones tendientes a prevenir sus efectos perniciosos y la manera de contenderlos, previas a la legislación. Desde el otro punto de vista intersistémico el derecho se convierte en objeto de valoración ética, por tanto, susceptible de ser reformado.

Como también se ha dicho, la inteligencia artificial ha impactado en nuestras vidas haciendo posibles muchísimas cosas que serían impensables sin ella, entre los que se encuentran los desarrollos en inteligencia artificial aplicada al derecho. Sin embargo, también implica riesgos importantes respecto de los cuales el derecho debe estar prevenido para contenderlos con la mayor eficacia posible.

Algunos de los peligros de la inteligencia artificial han llegado a constituir campos comunes. Entre ellos se encuentran: El riesgo de desempleo por el desplazamiento de sistemas inteligentes. En este sentido, en el Foro Económico Mundial de Davos se dijo que en menos de 20 años los 15 países más industrializados perderían 5 millones de trabajos ocupados por robots y en Japón (Geller y Hirschler, 2017), *Nomura Group* estima que para 2035 el 50% de los empleos japoneses serán realizados por robots (Hernández, 2015). Otro gran riesgo es el llamado mercado de datos por virtud del cual nuestra información se convierte en objeto de comercio para las empresas que, a través de sistemas de clasificación, conocen nuestro perfil y hábitos de consumo que explotan a través de técnicas de mercadotecnia individualizada sobre cuya eficacia tenemos poco control con el fin de potenciar a la sociedad de consumo (*La IA Aplicada Al Marketing, Una de Las Innovaciones Del 2021*, 2021). Desde luego el manejo de nuestra información por parte de grupos criminales forma parte de las preocupaciones más propagadas y pre-

ocupantes del uso de estas tecnologías, sobre todo si consideramos que enormes bases de datos hackeadas al Estado se venden en la *dark-web* (Hidalgo Pérez, 2019; DW Español, 2021).

Los sesgos cometidos por los sistemas de clasificación también pueden llevar a actos de discriminación con un sinfín de consecuencias, como pueden ser no conseguir un empleo, o, lo que es peor, ser estereotipados como potencial delincuente por de sistemas predictivos y clasificatorios (Naciones Unidas México, 2022). En un nivel más sofisticado, la inteligencia artificial también implica riesgos para la seguridad nacional. Por ejemplo, se sabe que el gobierno chino ha tenido acceso y cruzado bases de datos de funcionarios públicos norteamericano con el fin de reclutarlos como espías (Véliz, 2022). Paradójicamente todos estos peligros se ven potenciados exponencialmente por el hecho de admitir voluntariamente un sutil espionaje casero gracias a internet de las cosas (Instituto de Ingeniería UNAM, 2018).

No es posible en este trabajo hacer una exposición de los diferentes riesgos sociales que trae aparejada la inteligencia artificial, por lo que he seleccionado algunos por su trascendencia en contra de los derechos fundamentales.

### A. *Armas autónomas*

El término 'armas autónomas' denota a aquellas que son capaces de llevar a cabo una misión sin intervención humana, ni en la toma de decisiones sobre cómo actuar, ni en la ejecución de la tarea militar. En caso de existir con capacidad de acción plena los llamados robots y asesinos constituyen el ejemplo paradigmático de un arma autónoma (Comité Internacional de la Cruz Roja (CICR), 2022).

El término armas semiautónomas denota aquellas en las que la intervención humana está limitada porque parte importante de la tarea corresponde al artefacto. Sería un ejemplo de armas semiautomáticas los drones. Debe destacarse que no todo artefacto que se

comporta de manera autónoma en operaciones militares constituye un arma autónoma. Por ejemplo, los que realizan actividades de vigilancia y reconocimiento.

Existen dos posturas sobre el potencial uso de las armas autónomas:

Argumentos a favor:

- Pueden disminuir las bajas y lesiones de los militares humanos.
- Serían más eficaces en el campo de batalla por su poder de permanencia.
- Tendrían una capacidad de reacción más eficiente que los humanos y no tienen interferencias emocionales en la ejecución de sus tareas.
- No se dejarían llevar por reacciones emocionales que estén más allá del cumplimiento de su misión como puede ser odio racial, venganza, etcétera.
- Son más eficaces que los propios humanos en el cumplimiento de las normas establecidas en el derecho internacional humanitario ya que podrían ser más eficientes en el uso controlado de la fuerza sin usar más allá de la estrictamente indispensable.

En la actualidad existen varios proyectos sobre armas autónomas como parte de la tecnología militar robótica en más de 15 países, entre los que figuran: Estados Unidos, Reino Unido, Israel, Rusia y China.

En particular la fuerza aérea de los Estados Unidos tiene un programa que se prolongará hasta 2049.

En la actualidad operan distintos tipos de robots con distintos grados de autonomía, como son:

- El sistema sensor *Fused Weapon* (SFW) (Kallenborn, 2022).

- El sistema *Phalnox (Mk 15 Phalanx*, 2015), el C-RAM (Arma-  
pedia, 2022) o el *north group grummanX-47 B* de los Estados  
Unidos (*Northrop Grumman X-47, Drone de Ataque y Reconocimien-  
to*, n.d.).
- *El Harpy* de Israel (IAI, 2017).
- El *Taranis* del Reino Unido (Cole, 2020).
- El robot SGR-A1 de Samsung *techwin (Samsung Techwin SGR-A1,  
El Robot Centinela de Corea Del Sur*, 2010).

A pesar de la confianza depositada por algunos en las virtudes potenciales de las armas autónomas existen casos sumamente preocupantes como el ocurrido el 03 de julio de 1988 cuando el buque de guerra USS Vincennes, dotado del sistema AEGIS de defensa antiaérea derribó al vuelo 655 de Irán Air, debido a un error de identificación del radar como un F-14 iraní. La decisión qué se confió el sistema tuvo como resultado la muerte de 290 civiles.

### B. *Riesgo para la democracia*

La empresa *Cambridge Analytics* fue fundada el 2013 por los ultraconservadores Robert Mercer y Steve Bannon, asesor principal de Trump (Kaiser, 2019).

En marzo de 2018 *The New York times*, *The Guardian* y *The Observer* denunciaron que se estaban explotando datos de los ciudadanos indebidamente. Un ciudadano de nombre Christopher Wyli, ex empleado de *Cambridge Analytics* reveló que la empresa había desarrollado una maquinaria para manipular la decisión electoral de los votantes mediante la explotación de unos 50 millones de usuarios de *Facebook*.

En 2018 el canal británico *Channel 4* dio a conocer vídeos obtenidos por un periodista infiltrado que se hizo pasar por político de Sri Lanka, en donde el director ejecutivo de *Cambridge Analytics* dice haber intervenido en la manipulación de procesos electorales de 200 países incluyendo a Nigeria, Kenia, República Checa, India y Argentina.

El desarrollo de *Cambridge Analytics* se atribuye a la derecha o extrema derecha y se ha vinculado como entre otros, al triunfo de Trump, Bolsonaro, del Brexit y de Johnson en Inglaterra.

El propio *Channel 4* también refirió vinculaciones entre *Cambridge Analytics* y el Partido Revolucionario Institucional (PRI) de México.

El éxito de la empresa es el resultado de la combinación de técnicas de inteligencia artificial que explotan la minería de datos y desarrollos en psicología realizados en el terreno de la mercadotecnia conocidos con la expresión micro focalización desarrollados por Michal Kosanski.

Una etapa importante del proceso es el conocido como *Knowledge Discovery in Databases* (KDD) consistente en la recolección de información proveniente de grandes volúmenes de datos, con la finalidad de someterlos a técnicas de inteligencia artificial que, con base en complejos cálculos estadísticos, identifican patrones ocultos en dicha información, que en este caso correspondían a perfiles de los ciudadanos que se buscaron para ser manipulados con mensajes personalizados.

La microfocalización, por su parte, es una es una metodología de la mercadotecnia directa desarrollada en el Centro de Psicometría de la Universidad de Cambridge por parte de Michael Kosinski y David Stillwell.

Para recolectar la información se desarrolló una aplicación en *Facebook* denominada *my personality* con base en las categorías de extroversión apertura mental responsabilidad afabilidad y nerviosismo.

Los datos obtenidos fueron cruzados con otros de la vida cotidiana propios de información personal como la nacionalidad, el sexo, el estado civil, etcétera, que fueron sometidos al tratamiento de minería de datos e inteligencia artificial para obtener patrones con descripciones precisas de las características de los targets.

El siguiente paso consistió en la generación de mensajes personalizados dirigidos a la manipulación de la cognición corporizada y el inconsciente adaptativo de los electores aprovechando la información que se tenía a través de la aplicación.

Un escenario aterrador para la manipulación política implica la posibilidad de agregar a experimentos como el de *Cambridge Analytics* las técnicas de manipulación social conocidas como “Golpe de Estados Suaves” elaboradas para la desestabilización política relacionadas con el *lawfare*, guerra jurídica o guerra judicial que, según Thierry Meyssan, fueron desarrolladas Gene Sharp a solicitud de la CIA con el fin de derrocar gobiernos sin necesidad de intervención militar. La investigación se encuentra documentada en el libro titulado “De la dictadura a la democracia. Un sistema conceptual para la liberación” (Sharp, 2003). De manera sintética las técnicas del golpe de estado suave se distribuyen a lo largo de 5 etapas: la primera consiste en generar y promocionar un clima de malestar social; la segunda implica intensas campañas en defensa de la libertad de prensa y de los derechos humanos acompañadas de acusaciones de totalitarismo contra el gobierno; la tercera etapa implica una lucha activa por reivindicaciones políticas y sociales y la promoción de manifestaciones y protestas violentas; la cuarta etapa corresponde a una guerra psicológica y desestabilización del gobierno creando un clima de ingobernabilidad y la quinta etapa consiste en forzar la renuncia del Presidente mediante revueltas callejeras. Paralelamente se va a preparar el terreno para una posible intervención militar, mientras se desarrolla una guerra civil prolongada y se logra aislamiento internacional del país. Pierre Masen refiere al uso de las técnicas del golpe de estado blando en Irán y en América Latina, particularmente en Honduras Paraguay Brasil y Nicaragua.

### C. *Hipervigilancia estatal: el Big Brother Chino*

Una novela tan famosa como aterradora es 1984 de George Orwell (2020). En ella nos pinta un mundo totalitario en el que todo es con-

trolado por el Big Brother. Desafortunadamente la ciencia ficción se ha vuelto realidad.

Desde el año 2014 el gobierno chino ha implementado un sistema de control social en una treintena de localidades, está basado en inteligencia artificial, el uso de *apps* que centralizan todos los registros digitales de cada ciudadano, el uso de millones de cámaras de reconocimiento facial e incluso artefactos que recogen datos biométricos, como es el caso de los uniformes inteligentes empleados por los niños en las escuelas.

A través de este sistema de recolección masiva, el uso de técnicas de minería de datos y sistemas de clasificación de inteligencia artificial el gobierno chino tiene no solo tiene la capacidad de contar con datos fijos como sino también la posibilidad de hacer inferencias a partir de ellos sobre el comportamiento de sus ciudadanos. En otras palabras, el *Big Brother* chino centraliza la información correspondiente a todas las conductas de sus ciudadanos que considera relevantes, al tiempo que centraliza y administra una gran cantidad de incentivos necesarios para la vida cotidiana y los utiliza para premiar o sancionar a los que clasifica como buenos y malos ciudadanos. Por ejemplo, no hacer fila para subir al autobús, pasar el cruce de la calle con el semáforo para peatones en rojo, gritar en el metro o tirar una colilla en la calle pueden traer aparejadas sanciones como la prohibición de viajar en aviones, viajar en clasificaciones inferiores a las que en condiciones normales se tendría derecho, etcétera. El hecho de comportarse de conformidad con las conductas consideradas buenas por parte del gobierno puede facilitar dar prioridad en citas médicas, mejores créditos para comprar una vivienda, incluso ser mejor calificados en las aplicaciones de citas, comprar ciertos productos, etcétera.

Además de estos incentivos directos, el gobierno chino utiliza mecanismos de presión social basados en el principio de homofilia al dar a conocer públicamente las listas de los ciudadanos buenos y malos. Un ejemplo del proceso de socialización de este sistema es el

que ocurre en medios de transporte público como el tren bala, en el que al inicio de cada recorrido un alta voz emite un mensaje en el que les recuerda a los pasajeros comportarse adecuadamente pero no a ver disminuida su puntuación de su bono social (Garrido-Julve, 2019).

La eficacia de este mecanismo de mentalización social se pone de manifiesto cuando gente entrevistada expresa que el sistema facilita una sociedad mejor, promueve la mutua confianza y una vida social más armónica, todo lo cual compensa su pérdida de la libertad.

Pero el fenómeno presenta otra cara de la moneda si de control político se trata. Así, por ejemplo, el periodista Liu Xiu, sin ninguna notificación, vio restringido el acceso a cosas que antes podía disfrutar al ser acusado de difamación y propagación de rumores por sus reportajes sobre corrupción, sin posibilidad de apelar la decisión. En contraste con esta situación, el gobierno chino premia las manifestaciones explícitas hechos a su favor.

Otro ejemplo de control político del sistema de bonos tiene que ver con la discriminación de la que son objetos los miembros de una pequeña comunidad musulmana en China, que al ser identificados mediante los algoritmos de clasificación del sistema son enviados a centros de reeducación.

Uno de peligros más importantes de este sistema es su replicación por parte de otros regímenes totalitarios como el de Rusia, Arabia Saudita, e incluso para países democráticos que quieran monitorear y controlar mejor a sus ciudadanos. Desafortunadamente, esta propagación se encuentra en fase inicial en Venezuela a través del Carnet de la Patria, aún no obligatorio, pero cada vez más imprescindible pues es necesario para obtener alimentos subsidiados por el estado. En un sistema como este se puede condicionar el acceso a bienes necesarios para la subsistencia como la obtención de gas doméstico, energía eléctrica o agua potable. Las condiciones tecnológicas para que esto sea posible son ya una realidad (Rtve. Play Radio, 2022).

## IX. LA ÉTICA DE LA INTELIGENCIA ARTIFICIAL Y EL DERECHO DE LA INTELIGENCIA ARTIFICIAL

A pesar de los peligros que implica la inteligencia artificial, los esfuerzos para su regulación siguen siendo escasos tanto en el ámbito de la ética como en el jurídico. Entre ellos se encuentran los siguientes (Olvera, 2022):

- The Asilomar AI Principles, *Future of Life Institute* 2017;
- The Montreal Declaration for Responsible AI, *University of Montreal*, 2017 y 2018;
- Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems. *Institute of Electrical and Electronics Engineers (IEEE)* 2017 y 2019.
- The Ethical Principles offered in the Statement on Artificial Intelligence, Robotics and ‘Autonomous’ Systems, *European Commission’s European Group on Ethics in Science and New Technologies*, EGE 2018.
- The ‘five overarching principles for an AI code’, *UK House of Lords Artificial Intelligence Committee’s report, AI in the UK: ready, willing and able?* 2018.
- The Tenets of the Partnership on AI, *a multi-stakeholder organization consisting of academics, researchers, civil society organisations, companies building and utilising AI technology*, 2018).
- Organización para la Cooperación y el Desarrollo Económicos (OECD) (2019).
- Reglamento Europeo (Unión Europea, 2016).
- Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO, 2023).

A pesar de estos indudables avances, la relevancia y riqueza de los riesgos que implica lo que se ha dado en llamar “la cuarta revolución industrial” constituye un tema de investigación en las fronteras de la ciencia sobre los cuales, sin duda alguna, el mundo académico tiene mucho que decir.

## **X. CONSIDERACIONES FINALES**

Hace una generación la posibilidad de que las computadoras pudieran superar el *test* de Turing pertenecía al terreno de la ciencia ficción. Hoy día, no solo pueden responderlo, sino que pueden ofrecer respuestas a problemas que ningún humano podría contestar. Como si se tratara de una novela de ciencia ficción que nos hubiera escandalizado hace algunos años, las computadoras han comenzado a vencernos en los más diversos terrenos: *Deep Blue* al campeón mundial de ajedrez; *Alpha Go* aprendió a generar sus propios heurísticos para vencer a jugadores profesionales de *GO* en un tiempo récord que ningún humano podría haber alcanzado; *Watson* ha derrotado a jugadores humanos en el juego de Jeopardy; el androide Sofía, se ha convertido en una celebridad internacional al responder inteligentemente y en tiempo real a preguntas abiertas y ha obtenido la primera nacionalidad concedida a un no humano; por si fuera poco, en un terreno que apenas hace unos meses habiéramos considerado exclusivo de los seres humanos: el arte, la pintura que ganó un concurso en Colorado fue hecha por un sistema de inteligencia artificial (RT, 2022) y otro muy recientemente ha surgido AudioLM, un sistema desarrollado por *Google* que a partir de fragmentos de unos segundos de una interpretación de voz es capaz de completar una pieza armónica y con texto coherente, irreconocible de una obra humana y que puede hacer lo mismo con fragmentos de una pieza musical de unos cuantos segundos (Entrepreneur, 2022).

Uno de los temas más preocupantes y menos explorados, es la capacidad de la inteligencia artificial para manipular nuestros mecanismos cognitivos más profundos, sin que siquiera lo sospechemos. Ya ha mostrado lo equivocado de nuestras creencias acerca de la democracia como un ejercicio de hombres libres y racionales.

Como es natural en el ámbito de la ciencia y la tecnología, unos avances se irán integrando con otros que en este momento están en la frontera entre presente y futuro, como es el caso de la realidad virtual y el metaverso, capaces de ubicarnos en una o varias “Matrix” alternativas en las que nuestra identidad cognitiva natural corre el riesgo de disolverse para ser sustituida por uno o diferentes avatares.

Hace años nos aterraba pensar en un mundo controlado por un ente omnisciente sobre todos los aspectos de nuestra vida, que ya es una realidad. Hoy día es tecnológicamente posible pensar en la expansión del *Big Brother* chino a nivel planetario, gracias a la interconectividad mundial.

¿De qué manera impactará la computación cuántica en este escenario? ¿qué podemos esperar de sistemas cada vez más autónomos y capaces de aprender unos de otros? ¿qué papel jugará la robótica social en el futuro? ¿qué es lo que sucederá con toda esa fusión? ¿qué lugar tendrá el derecho en la nueva sociedad emergente? ¿su función será sustituida por dioses artificiales que conocen todas nuestras acciones y castigan y premian a voluntad?

Estas son solo algunas preguntas cuyas posibles respuestas aún son insospechadas y los juristas tendremos la gran responsabilidad contribuir a contestar. Pero, para ello, primero necesitamos reinventar nuestra manera de hacer y comprender al derecho, porque no podremos contribuir a resolver problemas de un futuro en ciernes, con supuestos que han sido útiles para una realidad social en vías de extinción.

## XI. FUENTES DE CONSULTA

- Armapedia. (2022). “¿Cómo Funciona un C-RAM? | 4500 balas por Minuto!” <https://www.youtube.com/watch?v=AApkuUx8xGA>
- BELDA, I. (2017). *La inteligencia artificial. De los circuitos al conocimiento*. RBA.
- CÁCERES, Enrique, y MANSILLA, Ricardo. (2019). *Expertius*. <http://expertius.mx/login?next=%2F>
- CÁCERES, Enrique. (2008). EXPERTIUS: A Mexican Judicial Decision-Support System in the Field of Family Law. In E. Francesconi, G. Sartor, y D. Tiscornia (Eds.), *Legal Knowledge and Information Systems* (pp. 78–87). *Frontiers in Artificial Intelligence and Applications*. <https://doi.org/10.3233/978-1-58603-952-3-78>
- Cole, Chris. (2020). “Drone Wars at Ten #3: What’s next? A peek at the future”. *Drone Wars*. <https://dronewars.net/tag/taranis/>
- Comité Internacional de la Cruz Roja (CICR). (2022). *Preguntas y respuestas: Lo que hay que saber sobre las armas autónomas*. <https://www.icrc.org/es/document/preguntas-y-respuestas-sobre-armas-autonomas>
- Dw Español. (2021). “¿Están mis datos a la venta en la darkweb?” <https://www.youtube.com/watch?v=sMG9UDv6eq8>
- Entrepreneur. (2022). “Google desarrolla un sistema de inteligencia artificial capaz de completar audios de voz y música.” <https://www.entrepreneur.com/es/tecnologia/google-desarrolla-un-sistema-de-inteligencia-artificial/436842>
- Estevez, Elsa., Fillotrani, Pablo., Y Linares, Sebastian. (2020). Prometea: Transformando La Administración De Justicia Con Herramientas De Inteligencia Artificial.” *Inter-American Development Bank*. <https://policycommons.net/artifacts/304001/prometea/1220760/>

- Garrido-julve, luis. (2019). “Buen Ciudadano, A La Fuerza: China Acelera Su Plan De Control Social Para 2020.” *El Confidencial*. [https://www.elconfidencial.com/mundo/2019-09-20/buen-ciudadano-a-la-fuerza-china-acelera-su-plan-de-control-social-para-2020\\_2201731/](https://www.elconfidencial.com/mundo/2019-09-20/buen-ciudadano-a-la-fuerza-china-acelera-su-plan-de-control-social-para-2020_2201731/)
- GELLER, Martine., & HIRSCHLER, Ben. (2017). *El impacto de los robots en el empleo, una creciente preocupación en Davos*. REUTERS. <https://www.reuters.com/article/davos-reuni-n-robots-idESKBN15411D>
- HERNÁNDEZ, David. (2015). *Los robots ocuparán la mitad de los empleos en Japón para 2030*. Computerhoy. <https://computerhoy.com/noticias/life/robots-ocuparan-mitad-empleos-japon-2030-37777>
- HIDALGO, Montse. (2019). *Cómo los criminales pueden utilizar la inteligencia artificial*. El País. [https://elpais.com/retina/2019/08/08/talento/1565261484\\_817309.html](https://elpais.com/retina/2019/08/08/talento/1565261484_817309.html)
- IA Lab. (2020). Prometea En El «Mundial» de Inteligencia Artificial. Recuperado el 8 de octubre, 2022, from <https://ialab.com.ar/prometeacumbremundial/>
- IAI. (2017). *HARPY. Autonomous Weapon for All Weather*. Recuperado el 8 de octubre, 2022, from <https://www.iai.co.il/p/harpy>
- Instituto de Ingenieria UNAM. (2018). *El peligro del internet de las cosas*. <http://www.ii.unam.mx/es-mx/AlmacenDigital/Notas/Paginas/internetdelascosas.aspx>
- KAISER, Brittani. (2019). *La dictadura de los datos*. HarperCollins Mexico.
- KALLENBORN, Zachary. (2022). *Applying arms-control frameworks to autonomous weapons*. Brookings. <https://www.brookings.edu/techstream/applying-arms-control-frameworks-to-autonomous-weapons/>
- CASTRO, Angel. (2021) *La IA aplicada al marketing, una de las innovaciones del 2021*. Interactiva. <https://interactivadigital.com/>

- com/opinion-marketing-digital/la-inteligencia-artificial-aplicada-al-marketing-ya-logra-retornos/
- MEZA, Ivan, AGUILAR, Sebastian., GARCÍA-CONSTANTINO, Matias., LÓPEZ, Carmen, Y CÁCERES, Enrique. (2019). Extracción automática de información jurídica de sentencias. In *MADIC. A cinco años de su creación: pasado, presente y futuro* (pp. 41–60). Universidad Autónoma Metropolitana. Unidad Xochimilco. <http://ilitia.cua.uam.mx:8080/jspui/handle/123456789/586>
- Mk 15 Phalanx*. (2015). Retrieved October 8, 2022, from <https://www.deagel.com/Protection Systems/Mk 15 Phalanx/a001641>
- Naciones Unidas México. (2022). *UNESCO pide detener la discriminación a través del mal uso de la Inteligencia Artificial*. UNESCO. <https://mexico.un.org/es/187451-unesco-pide-detener-la-discriminacion-traves-del-mal-uso-de-la-inteligencia-artificial>
- Northrop Grumman X-47, drone de ataque y reconocimiento*. (2008). <https://www.youtube.com/watch?v=tyBEIUzTErw>
- OLVERA, Assuán. (2022). Legislación tecnológica y de la Inteligencia Artificial.
- Organización para la Cooperación y el Desarrollo Económicos (OECD). (2019). *Recommendation of the Council on Artificial Intelligence*. <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>
- ORWELL, George. (2020). *1984*. Penguin Random House.
- ROSS Intelligence (“ROSS”)*. (2020). <https://blog.rossintelligence.com/>
- RT. (2022). *Obra generada por IA gana el primer lugar en un concurso de bellas artes y desata polémica*. <https://actualidad.rt.com/actualidad/440320-obra-generada-ia-gana-concurso-pintura-polemica-eeuu>
- Rtve. Play Radio. (2022). *Puntos por ser buen ciudadano: el crédito social*. Rtve. Play Radio. <https://www.rtve.es/play/audios/>

- la-cuadratura-del-circulo/puntos-buen-ciudadano-credito-social-chino-buenos-malos-ciudadanos-premios-castigos-cuadratura-del-circulo-radio-nacional/6372832/#:~:text=Desde 2020 funciona de manera,e incluso quedar expuestos p  
*Samsung Techwin SGR-A1, el robot centinela de Corea del Sur.* (2010). Retrieved October 8, 2022, from <http://lifeisinfinity.blogspot.com/2010/12/samsung-techwin-sgr-a1-el-robot.html>
- SHARP, Gene. (2003). *De la dictadura a la democracia. Un Sistema Conceptual para la Liberación*. The Albert Einstein Institution. [https://www.academia.edu/33323874/DE\\_LA\\_DICTADURA\\_A\\_LA\\_DEMOCRACIA\\_Un\\_Sistema\\_Conceptual\\_para\\_la\\_Liberación](https://www.academia.edu/33323874/DE_LA_DICTADURA_A_LA_DEMOCRACIA_Un_Sistema_Conceptual_para_la_Liberación)
- UNESCO. (2023). *Ética de la Inteligencia Artificial*. Retrieved October 8, 2022, from <https://unesdoc.unesco.org/ark:/48223/pf0000385082.page=4>
- Unión Europea. (2016). *Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen las normas armonizadas en materia de Inteligencia Artificial (ey de Inteligencia Artificial) y se modifican determinados actos legislativos de la Unión*. <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=celex:52021PC0206>
- VÉLIZ, Carissa. (2022). Privacidad es Poder. Debate.**  
vLex. (n.d.). *Vincent*. <https://vlex.com.mx/p/vincent/>