

Enseñar inteligencia artificial y ciencia de datos en la carrera de abogacía

David Mielnik⁽¹⁾

✉ david.m@utdt.edu

Lo que no puedo crear, no puedo entender.

Richard P. Feynman

1. Introducción

La progresiva digitalización y publicación de leyes, sentencias y decisiones administrativas en portales oficiales gratuitos nos ha provisto de una inusitada riqueza de información jurídica, a la que anteriormente solo podían acceder agencias estatales, grandes estudios jurídicos o centros de investigación capaces de solventar una suscripción a bases de datos especializadas. Los méritos de esta decisión para la transparencia institucional son muchos y han sido ampliamente reconocidos (Elena y Pichón Rivière, 2014). Sin embargo, con ellas, el derecho ha empezado a presentar los desafíos propios de la era de la información: hoy en día los tribunales publican literalmente decenas de miles de decisiones, que se suman a las también cuantiosas leyes, reglamentos y resoluciones administrativas, para dar como resultado un volumen de información que ningún intelecto humano está genuinamente en condiciones de abarcar. Eso puede ser especialmente abrumador para las y los profesionales más jóvenes, quienes, recién graduados, están todavía desarrollando las herramientas analíticas que permiten a sus colegas de mayor experiencia identificar eficientemente las doctrinas y argumentos más relevantes.

(1) Abogado (Universidad Torcuato Di Tella –UTDT–). Máster en Derecho Penal (UTDT). Profesor de Análisis Computacional del Derecho y Tecnología e Innovación Legal (UTDT).

En trabajos anteriores he argumentado que los avances en inteligencia artificial (IA) y ciencia de datos —la disciplina dedicada a desarrollar métodos para adquirir, procesar y explotar grandes volúmenes de información con el objetivo de descubrir nuevo conocimiento (Iguar y Seguí, 2017)— nos permiten dar cuenta del laberíntico paisaje en el que se está convirtiendo el derecho moderno, desde una perspectiva novedosa, nunca antes alcanzada. Entre otras cosas, mostré que podemos utilizar inteligencia artificial (IA) para realizar diversas tareas jurídicamente valiosas que antes resultaban imposibles como identificar las propiedades y antecedentes de un caso que más pueden inclinar la decisión en uno u otro sentido (Mielnik, 2022a); descubrir los principales tópicos que llegan a conocimiento de los tribunales (Mielnik y Altszyler, 2023); o cuantificar el grado en el que los casos se deciden con perspectiva de género o se expresan en lenguaje claro (Mielnik, 2022b).

En esta oportunidad quisiera detenerme en un desafío diferente: el de cómo transmitir estos conocimientos a estudiantes de abogacía para que ellas y ellos puedan utilizar por sí mismos herramientas como estas para amplificar sus capacidades de análisis legal. En efecto, la carrera de Derecho suele ser vista como “muy tradicional” y no nos resulta tan natural, quizás, pensar que una/o futura/o abogada/o puede graduarse dominando tanto los códigos de leyes como los informáticos. Al mismo tiempo, la introducción de tecnología en la carrera puede verse con cierta desconfianza en la creencia de que requiere recursos computacionales a los que no todos pueden acceder. Ambos preconceptos son afortunadamente erróneos, como intentaré mostrar en las líneas que siguen. Antes de eso, déjenme dedicar unas líneas a explicar brevemente algunas de las posibilidades que la inteligencia artificial —y la tecnología computacional, en general— puede abrir para las/os abogadas/os presentes y futuras/os.

2. Estudios legales computacionales

En los últimos cinco o diez años, la disponibilidad más amplia de recursos informáticos cada vez más poderosos ha llevado a investigadoras e investigadores de derecho a adoptar nuevas tecnologías para enfrentar los desafíos que se les presentan. El uso de estas tecnologías ha permitido el descubrimiento de conocimiento jurídico a partir de la explotación de grandes cantidades de información —algo que hasta hace poco parecía imposible—. Este abordaje interdisciplinario suele ser denominado análisis computacional del derecho (ACD) y combina la tecnología de la información con saberes técnico-jurídicos para identificar patrones y regularidades que están diseminados en

textos legales (leyes, sentencias, doctrina, etc.) y que son difíciles de percibir a simple vista.

¿En qué consiste exactamente el ACD? Si dejamos de lado los algoritmos y los bits por un momento, el análisis computacional del derecho es básicamente una forma de lectura: igual que usamos nuestros sentidos y capacidad de razonamiento para analizar textos en busca de conocimiento, el ACD también es un método para extraer información a partir del análisis del discurso legal. La diferencia es que parte de este proceso se realiza fuera de nuestro cerebro, en una computadora, a la que podemos delegar ciertas tareas específicas que pueden abarcar volúmenes de información que están, por su magnitud, por encima de las que es razonable estudiar con técnicas analíticas tradicionales.

Para utilizar un ejemplo con el que he ilustrado esta idea en el pasado (Mielnik, 2022, pp. 4-5), supongamos que queremos abordar la pregunta por el grado con el que los tribunales penales han adoptado una mirada con perspectiva de género en el juzgamiento de los casos que llegan a sus estrados. El modo más habitual de acercarse a esa interrogante consiste —al menos en parte— en evaluar cualitativamente algunas decisiones recientes, seleccionadas según algún criterio de relevancia razonable, intentar dilucidar la tendencia general y formular consideraciones críticas.

Por supuesto, no hay nada incorrecto en ese enfoque. No obstante, ofrecer un relevamiento genuinamente exhaustivo del estado de la cuestión —lo que demandaría el examen de decenas de miles de decisiones— probablemente nos resultaría prohibitivamente costoso en términos de tiempo y energía. Así, una manera complementaria de abordar el tópico mediante técnicas computacionales podría consistir en el entrenamiento de un modelo de inteligencia artificial capaz de identificar por nosotros **todos** los fallos que de alguna manera han mirado el derecho en clave no discriminatoria, reservándonos la responsabilidad de interpretar los resultados del procesamiento, ponerlos en el contexto de nuestros saberes jurídicos y presentarlos mediante gráficos que permitan resumir y comunicar eficientemente nuestras conclusiones.

Por cierto, debo enfatizar que este tipo de análisis no tiene vocación de reemplazar las metodologías tradicionales de análisis legal, sino de complementarlas. En efecto, el procesamiento de grandes volúmenes de información jurídica permite observar y dar cuenta del derecho desde una escala nunca antes alcanzada, pero no provee una comprensión detallada de cada texto. La lectura tradicional, por otra parte, aporta ese entendimiento y

detalle, pero consume tiempo y recursos mentales finitos, y está, por lo tanto, limitada a trabajar con un número relativamente pequeño de documentos seleccionados. Nada impide, entonces, que combinemos cada abordaje con el fin de potenciar sus respectivas fortalezas y minimizar sus debilidades.

Desde el punto de vista metodológico, la propuesta central de los estudios legales computacionales, como el ACD, radica en la adopción de tecnologías emergentes, como el procesamiento del lenguaje natural (PLN) y el aprendizaje automático (ML, por sus siglas del inglés: *machine learning*), que son algunas de las caras más famosas de lo que hoy conocemos como inteligencia artificial. Esas tecnologías —que, a su vez, forman parte de una caja de herramientas más amplia que llamamos **ciencia de datos**— pueden ser utilizadas para construir conocimiento mediante la explotación de grandes volúmenes de información, de formas, algo que hasta hace no mucho parecía imposible. En este sentido, el ACD es una forma de ciencia de datos **aplicada**, que combina tecnología con saberes jurídicos para identificar patrones y regularidades que, si bien de alguna manera **ya están presentes** en los textos legales, se encuentran diseminados a lo largo de miles y miles de textos y nos resultan, por ese motivo, inaccesibles a simple vista.

Déjeme ofrecer dos palabras sobre estas tecnologías, que he discutido más en detalle en algunos trabajos previos (Mielnik, 2022a, pp. 6-12; 2022b, pp. 179-181; 2023), a cuya lectura invito para complementar esta breve nota. Por un lado, el procesamiento del lenguaje natural (PLN) engloba diversas técnicas informáticas diseñadas para lidiar con las complejidades y matices del lenguaje humano, escrito y oral. En el marco del análisis computacional del derecho, es el área de la ciencia de datos que posibilita que las computadoras dejen de ver a los archivos que contienen el texto de sentencias o leyes como meras cadenas de caracteres y comiencen a identificar las estructuras —palabras, oraciones, párrafos— que hacen de esos documentos un discurso jurídico.

El aprendizaje automático (ML), a su turno, es la tecnología detrás del mayor proceso de maduración que ha atravesado la inteligencia artificial en los últimos años. Muy sintéticamente, la característica de los programas creados bajo este paradigma es que son capaces de “aprender” de los datos a los que son expuestos y regular sus propios parámetros de funcionamiento interno para así modelar los patrones subyacentes a la información, con el objeto, entre otros muchos, de clasificarla en variables discretas o continuas (tareas de “aprendizaje supervisado”) o encontrar agrupamientos relevantes (típicamente, tareas de “aprendizaje no supervisado”).

Juntos, el procesamiento del lenguaje natural y el aprendizaje automático nos permiten construir funciones y algoritmos programáticos capaces de transformar información **desestructurada** —como la que está latente y diseminada a lo largo del texto de las sentencias o leyes— en datos **estructurados**, que podemos cuantificar y reutilizar para, por ejemplo, aproximar estadísticas sobre el funcionamiento judicial, inferir la distribución temática de las sentencias o identificar las propiedades de un caso que mejor pueden explicar la decisión del tribunal que lo resuelve.

3. Por qué enseñar inteligencia artificial y ciencia de datos en abogacía

Debo reconocer que, puesto frente a esta pregunta, mi primera reacción es responder “¿por qué no?”. Después de todo, tanto la IA como la explotación de datos son áreas de la investigación en expansión que desde hace tiempo atraviesan casi todas las disciplinas humanas, incluyendo las ciencias sociales (Grimmer y Stewart, 2013; Monroe *et al.*, 2008). Así, no hay dudas de que esta clase de técnicas y metodologías pueden contribuir a la formación en derecho. Sin embargo, la abogacía es una carrera de por sí extensa, con programas muchas veces ya saturados. Voy a ofrecer, entonces, a modo de ejemplo, algunos beneficios directos que advierto en la enseñanza de la IA y la ciencia de datos específicamente para futuras/os abogadas/os.

En primer lugar, una estudiante avanzada de derecho podría, por ejemplo, utilizar un algoritmo de aprendizaje automático **no supervisado** para inferir los tópicos que mayor presencia tienen en una colección de textos —v. *gr.*, fallos, leyes o piezas de evidencia documental—. La utilidad que esta clase de sistemas de inteligencia artificial puede tener para juristas fue discutida, entre otros, en el artículo “Inteligencia artificial aplicada al estudio del derecho” (Mielnik y Altszyler, 2023), donde se realizó un análisis computacional retrospectivo de la jurisprudencia de la Cámara Federal de Casación Penal, desde el comienzo de su actividad jurisdiccional, en 1992, hasta finales de 2022. Entre muchas otras cosas que algoritmos como el allí descrito pueden realizar —tales como indagar en el impacto de la perspectiva de género en las resoluciones, evidenciar el surgimiento y extinción de doctrinas judiciales o explorar el impacto que la pandemia de COVID-19 tuvo en la práctica tribunalicia—, obtener un modelo de los tópicos presentes en una colección reciente de fallos de un tribunal cualquiera puede permitir a la estudiante del ejemplo conocer, de manera objetiva, cuáles son los principales temas que se están litigando en un momento determinado.

Así, ella podría desarrollar una “taxonomía de la práctica” (Carter *et al.*, 2020): un registro empírico de los temas, doctrinas y discusiones que más presencia tienen en la realidad del trabajo cotidiano del tribunal, más allá de la competencia material que surge de las leyes, o de lo que puedan sugerir, en abstracto, los libros de texto. Esto le permitiría entender mejor cómo se están aplicando las leyes en la práctica y le daría una visión más completa de los problemas legales en juego. Contaría, entonces, con información de altísimo valor práctico para orientar sus estudios o prepararse para exámenes de ingreso a alguna de las instituciones que componen el servicio de administración de justicia y, además, habría adquirido habilidades que pueden serle especialmente útiles también en otras áreas de la práctica legal, como el análisis de testimonios o piezas documentales.

Más aún, armada con conocimientos sobre análisis computacional del derecho, la estudiante podría también utilizar algoritmos de aprendizaje **supervisado** —habitualmente conocidos como algoritmos predictivos— no solo para intentar generar un modelo que le permita pronosticar cómo podría pronunciarse un tribunal, en base a sus precedentes, sino, más en particular, para inferir los argumentos y propiedades de un caso más fuertemente correlacionado con un cierto tipo de decisión. La efectividad de este tipo de algoritmos para analizar la jurisprudencia nacional fue estudiada en el artículo “Análisis computacional del derecho argentino” (Mielnik, 2022a), y ellos también formaron parte de la metodología empleada en otros trabajos computacionales que se enfocaron en los precedentes de la Corte Suprema de Estados Unidos (Katz *et al.*, 2017) y del Tribunal Europeo de Derechos Humanos (Medvedeva *et al.*, 2022).

Quiero enfatizar el poder nivelador que puede tener la adopción de esta clase de métodos por parte de las y los profesionales del derecho más jóvenes. En efecto, con el tiempo, cualquier profesional del derecho se va formando una idea, una suerte de “representación mental” de los temas, doctrinas o argumentos que más chances de éxito tienen para respaldar sus pretensiones ante los distintos tribunales en los que le toca litigar. Este efecto ha sido incluso demostrado empíricamente (González Bertomeu, 2020). Ese conocimiento es extraordinariamente valioso pero, a su vez, muy costoso de conseguir, y en el fondo solo está basado en la experiencia subjetiva y necesariamente limitada de cada litigante.

En cambio, el uso de un algoritmo de inteligencia artificial podría permitirnos, en cierta medida, inferir el mismo tipo de consideraciones, pero basadas en razones objetivas, derivadas del

procesamiento de todos los precedentes relevantes, y de esa forma “nivelar” el campo de juego para los recién graduados. Por supuesto, este tipo de algoritmos solo es capaz de establecer correlaciones estadísticas entre, pongamos, argumentos y decisiones (entre otras cosas). Así, si bien podemos ahorrar tiempo y esfuerzo, hay que destacar que el procesamiento informático, en sí mismo, todavía no nos dice nada sobre la existencia de una relación de causalidad entre las propiedades de un caso y su resolución (si es que ella existe). Ese tipo de explicaciones deberán necesariamente ser completadas con saberes jurídicos, que el profesional o la profesional debe dominar para que este tipo de metodologías tenga algún sentido genuino.

Ahora bien, más allá de las tareas específicas que una o un estudiante de derecho podría aprender a delegar en un sistema inteligente (de los que referí apenas un par, solo a título ejemplificativo), hay diversas razones pedagógicas que, en mi opinión, hacen del estudio del análisis computacional del derecho una buena idea. La más importante, creo, es que a lo largo de sus carreras profesionales, las y los estudiantes actuales inevitablemente deberán enfrentarse con tecnologías como la inteligencia artificial, que hoy mismo está protagonizando lo que algunos ya llaman la Cuarta Revolución Industrial (Velarde, 2020), y que está llamada a proponer enormes desafíos para nuestra sociedad. ¿Cómo evitar que una IA nos discrimine por nuestra raza o género, como lo ha hecho en el pasado, cuando buscamos empleo o solicitamos acceder a programas de libertad condicional? ¿Cuál es la responsabilidad de un/a médico/a que emitió un diagnóstico equivocado, basado en las predicciones de un algoritmo? ¿Cuál es la responsabilidad de un/a abogado/a por utilizar IA? ¿Y de **no** utilizarla, hacerlo habría mejorado las chances de éxito de su cliente? Estas son solo algunas de las clases de preguntas que ya, hoy mismo, nos estamos haciendo (o deberíamos). Y son solo el principio.

En un escenario como este, tener conocimiento de primera mano acerca de cómo funciona una tecnología disruptiva potencialmente tan útil, pero también peligrosa, ofrecerá ventajas obvias. Especialmente, porque de otro modo es difícil encontrar analogías que nos permitan entender y explicar su funcionamiento si nunca vimos cómo funciona por dentro un algoritmo predictivo o carecemos al menos de intuiciones acerca de cómo es el proceso de aprendizaje de una máquina. Como en tantas otras cosas, no es lo mismo leerlo en un libro o que nos lo cuenten, que hacerlo nosotros mismos, con nuestras manos. Como solía recordar el premio nobel Richard Feynman a sus estudiantes: “Lo que no puedo crear, no puedo entender”. Las y los estudiantes –nativos digitales– lo saben perfectamente.

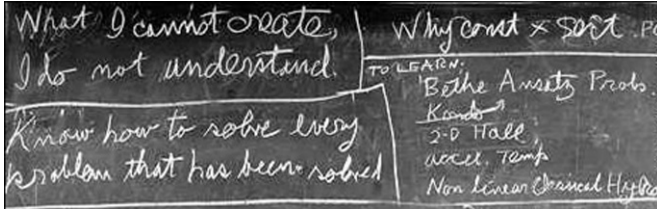


Figura 1. El pizarrón de Richard Feynman, con la expresión en inglés “What I cannot create, I do not understand” (“Lo que no puedo crear, no puedo entender”).

Fuente: California Institute of Technology.

Por lo demás, educar a generaciones de estudiantes de abogacía en la importancia de los datos judiciales de calidad, y en su uso responsable, seguramente contribuirá a preparar profesionales que demanden, cada vez con mayor intensidad, que el Estado cumpla con su obligación de hacer pública y accesible esa clase de información —un compromiso hoy solo cumplido a medias, indudablemente, al menos en parte, porque son pocas/os los profesionales capacitados para aprovecharla—.

4. Las (creencias en torno a las) barreras de acceso al análisis computacional del derecho

Quisiera detenerme ahora en algunas creencias en torno a la posibilidad de enseñar tecnología en la carrera de derecho. Hace algunos años, quizás las más importantes objeciones en contra de incluir en la carrera de derecho cursos vinculados con programación, datos o incluso inteligencia artificial tenían que ver con la falta de conexión aparente con esas disciplinas, lo que rápidamente puede llevar a creer que se trata de habilidades, como mínimo, innecesarias, que solo sobrecargarían una carrera ya de por sí extensa y exigente.

Mi impresión es que este tipo de objeciones no pueden ser mantenidas en la actualidad. La velocidad con la que se generan y distribuyen datos —que da como resultado el fenómeno que habitualmente llamamos *big data*—,⁽²⁾ así como la adopción cada vez más masiva y acelerada de tecnologías con poder disruptivo —con la inteligencia artificial a la cabeza, pero también incluyendo la computación en la nube, la *blockchain*, etc.— han impactado en casi todos los aspectos de la vida en sociedad y de esa manera, naturalmente, en el derecho. Hoy en día, en efecto, muchas de las preocupaciones centrales para cada rama del derecho están atravesadas por las nuevas tecnologías: desde la pregunta por la propiedad intelectual de

(2) Es el término con el que habitualmente nos referimos a una cantidad de información tan vasta o compleja que resulta difícil o imposible de analizar utilizando herramientas y técnicas tradicionales. Así lo define también el Observatorio Nacional de Big Data, dependiente de la Subsecretaría de Tecnologías de la Información y las Comunicaciones, en su sitio web oficial: argentina.gob.ar/grupo-de-trabajo/observatorio-nacionalbig-data/observatorio-big-data/que-es-big-data

obras generadas mediante IA o el estatus legal de los contratos autoejecutables (“*smart contracts*”) en el derecho civil, hasta la persecución de ciberdelitos en el derecho penal, pasando por los desafíos que plantea la evidencia digital para el derecho probatorio en general, la tecnología *blockchain* para el derecho notarial y el almacenamiento descentralizado de datos para el derecho internacional.

Es cierto que muchas de esas interrogantes ya se incluyen en los programas de estudio de diversas facultades de derecho argentinas pero, muchas veces, las discusiones dejan rápidamente en evidencia un conocimiento apenas superficial de las tecnologías subyacentes. Eso empobrece el debate o, directamente, nos hace incurrir en errores conceptuales. El ejemplo más ubicuo aparece en la discusión jurídica sobre los sesgos de la IA –un problema urgente y real–, una discusión que muchas veces ignora por completo el proceso de aprendizaje automático de los algoritmos que están detrás de esa tecnología, y nos puede llevar a pensar que se trata de artefactos que surgen de que el o la programadora del sistema haya trasladado sus propios prejuicios al sistema, en lugar de ubicar el riesgo en los conjuntos de datos con los que el modelo es entrenado.

También se escucha habitualmente que “los algoritmos de IA funcionan como cajas negras, inescrutables”, una afirmación que tal vez sea válida para cierto tipo de arquitecturas, pero que ignora la existencia de una enorme cantidad de sistemas transparentes, trazables y en gran medida explicables. Incluso redes neuronales con ciertas estructuras especialmente complejas –habitualmente citadas como paradigma de algoritmos indecifrables– cuentan con mecanismos de interpretación cada vez más aceptables (Angelov *et al.*, 2021).

Otra objeción a la inclusión del análisis computacional del derecho dentro de la currícula de una carrera de abogacía gira en torno a lo que se percibe como una insuperable barrera técnica de entrada: acceder a disciplinas como la ciencia de datos o la IA requiere que las y los estudiantes (y sus docentes!) aprendan a programar, y hacerlo –se sostiene– exige internalizar un conocimiento arcano, reservado solo para los iniciados de carreras como la ingeniería o la informática. Eso podía ser verdad hace 30 o 40 años, cuando la programación se realizaba casi exclusivamente mediante lenguajes de “bajo nivel” –como el ensamblador o el lenguaje C original–, los cuales, por su cercanía al código de máquina, exigían, entre otras cosas, un conocimiento profundo de la arquitectura y el funcionamiento de una CPU, la gestión de la memoria RAM y la manera de mover cada *byte* de una ubicación del *hardware* a otra mediante instrucciones francamente incomprensibles.

En contraste, la programación moderna –al menos, en lo que puede ser de interés para una o un abogada/o, presente o futuro– emplea lenguajes de “alto nivel” como Python, Julia o R, que sintácticamente guardan muchas similitudes con el lenguaje natural (típicamente el inglés), y realizan por nosotros la mayor parte de las operaciones subyacentes, lo que nos permite concentrarnos en lo importante: la lógica de nuestro programa y el procesamiento de la información que nos interesa.

A su vez, la comunidad de desarrolladores/as ha producido y mantiene miles de “bibliotecas” o “módulos” de funciones que nos eximen de tener que “reinventar la rueda” en nuestros propios códigos, y de esa manera, nos permiten pensar en la programación más bien como un juego de Lego, en el que nuestra tarea, como analistas computacionales del derecho, es combinar las piezas con creatividad, inteligencia y un propósito jurídico claro.

Por otra parte, si bien debería estirar mucho los conceptos para sostener que escribir un programa informático se parece a redactar un contrato o una demanda, sí estoy preparado para afirmar que hay habilidades que se enseñan en la carrera de derecho que posicionan a sus estudiantes en un lugar privilegiado para aprender programación. Tengo en cuenta, por un lado, que la programación es en el fondo una forma de lógica formal aplicada –algo que también se enseña en cualquier curso de introducción al pensamiento científico–. Si uno le pregunta a un/a especialista, seguramente nos dirá que la programación ocurre un 80% en nuestra mente –en el diseño del programa y en la comprensión del flujo de la información y sus transformaciones– mientras que el código representa tan solo un 20% del proceso –la traducción de esas ideas a un lenguaje comprensible por la computadora–.

En la carrera de derecho, por su parte, estamos permanentemente expuestos (y finalmente acostumbrados) a enfrentarnos a extensos planteos lógicos, que se van desarrollando desde ideas-principio más bien básicas y de sentido común, hasta plasmarse en la solución de un caso, propuesta y defendida por una parte, o adoptada por un tribunal. Los escritos de los litigantes o las sentencias judiciales, por caso, contienen árboles de decisión muy complejos, a veces plagados de bifurcaciones anidadas que podrían formalizarse como cláusulas condicionales. Así, una habilidad básica que desarrolla cualquier jurista es, precisamente, la de poder seguir la lógica de enunciados muy sofisticados, filtrar mentalmente los artificios retóricos y comprender el flujo de la información a lo largo de desarrollos argumentales que suelen extenderse por varias páginas de texto.

Esa habilidad es absolutamente central para la programación (al menos para el 80% de lo que implica desarrollar una función informática), por lo que el esfuerzo pedagógico rápidamente puede enfocarse en aprender los fundamentos del código, que son

realmente muy simples y accesibles. En ese sentido, hay que advertir también que las instrucciones y estructuras que una o un analista computacional del derecho debe poder conocer para escribir sus propias funciones y algoritmos es solo un subconjunto, y por cierto bastante reducido, de aquellas que necesita aprender una programadora profesional para tener éxito en una carrera especializada. Este es otro punto a resaltar: el objetivo de la formación en ciencia de datos e IA para futuras abogadas y abogados es que puedan aprovechar esas herramientas, en la medida en que les sean útiles para extraer y trabajar con información jurídica a la que, de otra manera, no tendrían acceso. Para eso deben conocer suficientes fundamentos de programación y desarrollar ciertas intuiciones básicas sobre esas técnicas y métodos; pero en ningún caso se debería esperar, ni es necesario, que las dominen en el sentido en que lo haría una graduada de ingeniería en informática o similar.

Consideremos el siguiente código de Python (un lenguaje de muy alto nivel, usado frecuentemente en ciencia de datos). Creo que cualquiera podría interpretar la función "juicio_abreviado", de la figura 2, que hace exactamente lo que parece: verifica si (*if*) el monto pretendido en una imputación es igual o menor a 6 años, y, en ese caso, advierte que el juicio abreviado (art. 431 bis CPPN) es una opción procesal. Este código realmente funciona si uno lo ejecuta en una computadora, y tranquilamente podría formar parte de un programa que nos ayude a clasificar casos, de acuerdo a si pueden ser adjudicados mediante el procedimiento del juicio abreviado previsto, al menos, en el código de procedimientos del ordenamiento penal federal.

Figura 2. Código Python (alto nivel).

Fuente: elaboración propia.

```
def juicio_abreviado(imputacion):
    if imputacion.monto_pena_pretendido <= 6:
        return "Juicio abreviado disponible"
```

En comparación, la misma función en el lenguaje de bajo nivel C (uno de los más antiguos) se escribiría como muestra la figura 3: una manera, creo, mucho menos intuitiva pero, afortunadamente, innecesaria actualmente. En efecto, en 2023 nadie hace en la práctica ciencia de datos en lenguaje C.

Figura 3. Código equivalente al de la figura 2, pero escrito en el lenguaje C (bajo nivel).

Fuente: elaboración propia.

```
#include <stdio.h>
#include <stdlib.h>

typedef struct {
    float monto_pena_pretendido;
} Imputacion;

char* juicio_abreviado(Imputacion imputacion) {
    if (imputacion.monto_pena_pretendido <= 6) {
        return "Juicio abreviado disponible";
    }
}
```

Una tercera objeción a la enseñanza de IA y ciencia de datos en las aulas de derecho, finalmente, podría señalar que este tipo de tecnologías requiere que las y los estudiantes posean computadoras poderosas, a las que no todos están en condiciones de acceder.

En la actualidad, sin embargo, existen múltiples plataformas en línea que permiten el uso colaborativo y/o educativo gratuito de herramientas para análisis de datos e inteligencia artificial, como Google Colaboratory o Kaggle, entre otras. Estas plataformas proporcionan acceso a *hardware* en la nube, con capacidad suficiente para ejecutar cualquier algoritmo o función que un analista computacional del derecho pueda necesitar –en efecto, sus prestaciones superan ampliamente a las que se pueden encontrar en la mayoría de las computadoras personales comerciales–. Dado, entonces, que los programas que escribimos en esas plataformas se ejecutan en el *hardware* virtual que ellas nos facilitan, utilizar estas herramientas solo requiere de una conexión a internet y una cuenta gratuita. Así, por caso, las *notebooks* que se distribuyen gratuitamente a través de programas como Conectar Igualdad son más que suficientes para aprovechar todas las herramientas necesarias para trabajar con IA y ciencia de datos en las aulas de derecho.

Con estas plataformas educativas en línea, las y los estudiantes de abogacía tienen la oportunidad de trabajar con las mismas herramientas que utilizan los profesionales del área. Como expliqué más arriba, entonces, lejos de profundizar brechas digitales, la enseñanza de la ciencia de datos puede contribuir a **igualar** las oportunidades de estudiantes y graduados/as con acceso a más y menos recursos materiales.

5. Palabras finales: mi experiencia enseñando ciencia de datos e IA a estudiantes de derecho

En función de las múltiples ventajas que esta clase de metodologías, técnicas y saberes pueden tener para abogadas y abogados prospectivos –y consciente también de los desafíos que puede entrañar su enseñanza– durante 2022 diseñé el programa del curso “Análisis computacional del derecho” (ACD), que desde ese mismo año dicto tanto en la carrera de abogacía (grado), como en la maestría en derecho penal (posgrado). Quisiera concluir esta presentación reseñando brevemente su estructura, funcionamiento y resultados, en el entendimiento de que ellos pueden resultar útiles para llevar adelante iniciativas similares –y seguramente superadoras–.

“Análisis computacional del derecho” es una materia pensada para estudiantes de 3^{ro} a 5^{to} año de la carrera, o estudiantes

de posgrado, única, al menos de momento, en Argentina y Latinoamérica. Surge de una observación doble: por un lado, hay numerosas tecnologías disruptivas que están cambiando prácticamente todos los paradigmas de la vida en sociedad –la inteligencia artificial, la automatización, la *blockchain*, etc.– y que esos cambios suponen grandes desafíos para los y las abogadas, presentes y futuros, que requieren que también cambiemos la manera en la que concebimos una carrera de derecho y las herramientas con las que un abogado o una abogada tienen que salir al mercado laboral.

Por otro lado, la materia parte también de la observación –que he sugerido más arriba– de que el derecho mismo se está convirtiendo en un fenómeno de *big data*, donde la cantidad de información jurídica en bruto, esto es, sentencias, leyes, normativa, tratados, así como las tareas que exige la profesión –la diligencia debida, los procedimientos de *compliance*, etc.–, se van haciendo tan grandes y complejas que demandan cada vez más conocimientos técnicos especializados que permitan, precisamente, lidiar con esos grandes volúmenes de información.

Para lidiar con esos desafíos, las clases de ACD tienen dos componentes principales: de un lado, la discusión de algunos de los artículos recientes que más impacto han tenido en el área interdisciplinaria de la ciencia de datos, la inteligencia artificial y el derecho. Lo más innovador, sin embargo, y lo que probablemente sea lo más atractivo del curso, es que invita a los y las estudiantes a animarse ellos mismos a desarrollar habilidades de programación y fundamentos de ciencia de datos para ponerlos a disposición de la búsqueda de conocimiento nuevo y a la posibilidad de responder preguntas jurídicas desde una perspectiva nunca antes posible. Así, el curso adopta como filosofía el “aprender haciendo” y brinda las herramientas para que las y los estudiantes puedan estudiar no solo una, cinco, diez, veinte o cincuenta sentencias, sino miles, decenas y centenas de miles de sentencias, decisiones, leyes, tratados, etc. Eso da una perspectiva única que permite responder preguntas que antes quizás ni nos animábamos a formular dado lo improbable de obtener una respuesta.

El curso de ACD comienza con ejercitación básica en uno de los lenguajes de programación más accesibles (Python) y, en parte por eso, uno de los más utilizados para hacer ciencia de datos hoy en día. Una vez que esos fundamentos están explorados, son aplicados a la búsqueda de respuestas jurídicas. En efecto, el objetivo del curso no es formar programadores y programadoras expertas, sino utilizar esta clase de herramientas para estudiar derecho, que es lo que estudiantes y profesionales del derecho hacemos. Así, la parte más importante del curso

está dedicada a realizar, clase a clase, diferentes proyectos de investigación que permitan poner en práctica estas técnicas y metodologías.

Entre otras cosas, indagamos en los temas centrales, por ejemplo, a los que dedicó su atención la Procuración General de la Nación antes, durante y después de la pandemia, mediante el procesamiento de los dictámenes que publica el organismo. También hicimos estudios sobre las maneras en las que podemos extraer información legal de sentencias, cómo presentarla a través de gráficos y figuras y cómo interpretar estos últimos.

Replicamos, asimismo, estudios del impacto de la normativa sobre perspectiva de género y protección y promoción de los derechos de las mujeres en la jurisprudencia. Se trata, en efecto, de un tema central del que hoy en día hablamos permanentemente, pero respecto del cual tenemos pocas herramientas claras para pensar, empíricamente, cuánto ha contribuido genuinamente en la práctica judicial cotidiana. En ese estudio analizamos más de 19.000 sentencias de los tribunales orales en lo criminal de la ciudad de Buenos Aires y generamos, junto con los estudiantes, un algoritmo que permite distinguir, bajo cierto criterio que consideramos apropiado, una sentencia con perspectiva de género de una que no lo tiene. Finalmente, el curso culminó con el entrenamiento por parte de las y los estudiantes de un modelo de inteligencia artificial, un algoritmo de *machine learning*, que aplicamos a sentencias de la Cámara Federal de Casación Penal para hacer inferencias sobre su doctrina reciente.

La evaluación final de la materia, que está orientada a la investigación legal, consistió en la elaboración de un proyecto de investigación para el que cada estudiante (o equipo) seleccionó libremente alguna pregunta de relevancia jurídica, que buscó responder a partir de las metodologías y técnicas abordadas. Entre muchos otros, uno de los trabajos, efectuado sobre 300 fallos penales, indagó sobre los objetos más frecuentemente admitidos por los tribunales como "armas impropias", proponiendo a partir de ese estudio empírico una forma de precisar los tipos penales en los que ese elemento es relevante y siempre blanco de objeciones constitucionales por vaguedad. Otro trabajo estudió el modo en el que la totalidad de las constituciones de Latinoamérica, actuales e históricas, se refieren a los derechos de los pueblos originarios, encontrando algunos resultados sumamente interesantes e inesperados, dada la hipótesis de investigación. En otro, las y los estudiantes analizaron computacionalmente 80 resoluciones de la IGJ y sistematizaron, con un alcance sin precedentes en la literatura, las razones más habituales por las que el registro público rechaza la inscripción de una sociedad comercial.

Los menciono porque, más allá de la contribución específica que cada uno pueda hacer a la ciencia del derecho, a mi modo de ver ellos ponen de resalto el aspecto más interesante de ofrecer un curso como "Análisis computacional del derecho". Lo que tienen en común trabajos como los reseñados, en efecto, es que en ellos los estudiantes se pudieron hacer preguntas que, más allá de las respuestas que encontraron por sí mismos, antes quizás ni siquiera se habrían atrevido a imaginar por la sencilla razón de que carecían de métodos o técnicas para alguna vez encontrar una respuesta.

Esa es la verdadera riqueza de un curso como este: empoderar a los estudiantes y las estudiantes con herramientas para responder preguntas jurídicas desde una perspectiva nunca antes accesible y, sobre todo, para poder hacerse nuevas preguntas, que los docentes probablemente ni siquiera imaginamos □

Referencias

Angelov, P. P.; Soares, E. A.; Jiang, R.; Arnold, N. I. & Atkinson, P. M. (2021). Explainable artificial intelligence: an analytical review. *WIREs Data Mining and Knowledge Discovery*, 11(5). DOI:10.1002/widm.1424

Carter, D.; Brown, J. & Rahmani, A. (2016). Reading the High Court at a distance: topic modelling the legal subject matter and judicial activity the High Court of Australia, 1903–2015. *The University of New South Wales Law Journal*, 39(4), pp. 1300-1354. ssrn.com/abstract=3564387

Elena, S. y Pichón Rivière, A. (2014). Recomendaciones para implementar una política de datos abiertos en el Poder Judicial. *CIPPEC*. cippec.org/wp-content/uploads/2017/03/1344.pdf

González Bertomeu, J. F. (2020). Different ways of losing: public defenders (and private counsel) at the Supreme Court of Argentina. *Law & Society Review*, 54(2), pp. 354-390. DOI: 10.1111/lasr.12473

Grimmer, J. y Stewart, B. M. (2013). Text as data: the promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, 21(3). web.stanford.edu/~jgrimmer/tad2.pdf

Igual, L. y Seguí, S. (2017). *Introduction to data science*. Springer.

Katz, D. M.; Bommarito, M. J. & Blackman, J. (2017). A general approach for predicting the behavior of the Supreme Court of the United States. *Plos One*, 12(4). journals.plos.org/plosone/article?id=10.1371/journal.pone.0174698

Medvedeva, M.; Wieling, M. & Vols, M. (2022). Rethinking the field of automatic prediction of Court decisions. *Artificial Intelligence and Law*, 31, pp. 195-212. doi.org/10.1007/s10506-021-09306-3

Mielnik, D. (2022). Análisis computacional del derecho argentino. *Revista Argentina de Teoría Jurídica*, 23(1), pp. 1-35. revista juridica.utdt.edu/ojs/index.php/ratj/article/view/442/352

----- (2022). Ciencia de datos y análisis del derecho penal. *Revista de Derecho Penal y Criminología*, 12(11).

Mielnik, D. y Altszyler, E. (2023). Inteligencia Artificial aplicada al estudio del derecho: análisis computacional de la jurisprudencia de casación penal. En J. G. Corvalán (Coord.), *Tratado de Inteligencia Artificial y Derecho*. La Ley. En prensa.

Monroe, B. L.; Colaresi, M. P. & Quinn, K. M. (2008). "Fightin" Words: Lexical Feature Selection and Evaluation for Identifying the Content of Political Conflict. *Political Analysis*, 16(4). languagelog.lidc.upenn.edu/myl/Monroe.pdf

Velarde, G. (2020). Artificial Intelligence and its impact on the Fourth Industrial Revolution: a review. *ArXiv*. doi.org/10.48550/arXiv.2011.03044