

**Algoritmos, sentencias y derecho constitucional comparado:
explorando el uso de jurisprudencia extranjera en Colombia**

Juan Sebastián Anzola Cortés

**Programa de Ciencias Políticas
Facultad de Derecho y Ciencias Políticas
Universidad de La Sabana
Chía, Cundinamarca
Marzo 2021**

**Algoritmos, sentencias y derecho constitucional comparado:
explorando el uso de jurisprudencia extranjera en Colombia**

**Trabajo de grado para optar por el título de politólogo, presentado por:
Juan Sebastián Anzola Cortés**

**Dirigido por:
Ms. Cristian Felipe Rojas Gonzáles
Ing. José Francisco Molano Pulido**

**Programa de Ciencias Políticas
Facultad de Derecho y Ciencias Políticas
Universidad de La Sabana
Chía, Cundinamarca
Marzo 2021**

Agradecimientos

Las deudas que adquiero al entregar este trabajo no son cuantificables, mucho menos podrían ser pagadas en una vida entera. Sin embargo, reconocer a aquellos con quienes las he adquirido es un deber para mí, ya que mis logros también les pertenecen.

Debo resaltar el rol fundamental de mis padres *Adriana y Mauricio*, así como de mis abuelos *Blanca y Pablo* que, con su apoyo incondicional, amor y cariño, han hecho que los caminos que he recorrido estén libres de todo tipo de obstáculos, lo que me ha permitido realizarme tanto personal como profesionalmente. También debo hacer una mención especial a mi tío Mauricio que siempre fue un modelo a seguir por su ética de trabajo y todos los valiosos consejos que me ha brindado.

De igual forma, en la vida universitaria encontré cuatro grandes mentores que han contribuido enormemente para que pueda llegar hasta acá: Diana Carolina Rojas, que desde el primer día ha creído en mí y me ha orientado en los momentos más importantes de mi vida académica y profesional. También, Andrés Agudelo, Cristian Rojas y José Francisco Molano han sido guías, así como voces de aliento y consejo.

En la universidad conocí verdaderos amigos con quienes he compartido momentos increíbles: Caterin, Johana, Sebastián Nieto, Manuel y Daniela...

A todos, gracias infinitas.

Algoritmos, sentencias y derecho constitucional comparado: explorando el uso de jurisprudencia extranjera en Colombia

Juan Sebastián Anzola Cortés

Resumen

Los resultados de esta investigación son el producto de una exploración metodológica con la que se pretende abordar la detección automática de referencias a jurisprudencia extranjera en las sentencias de la Corte Constitucional de Colombia. Para esto se analizaron más de 6000 sentencias de constitucionalidad emitidas entre 1992 y 2020 con un enfoque de minería de datos, aplicando algoritmos de búsqueda de patrones de lenguaje, que fueron reforzados con asistencia humana. Las estadísticas (preliminares) que se presentan a continuación representan los primeros datos consolidados sobre esta práctica en el país y buscan aportar a la generación de conocimiento en el campo del derecho constitucional comparado.

Palabras clave: derecho constitucional comparado, jurisprudencia extranjera, procesamiento de lenguaje natural, minería de texto.

Abstract

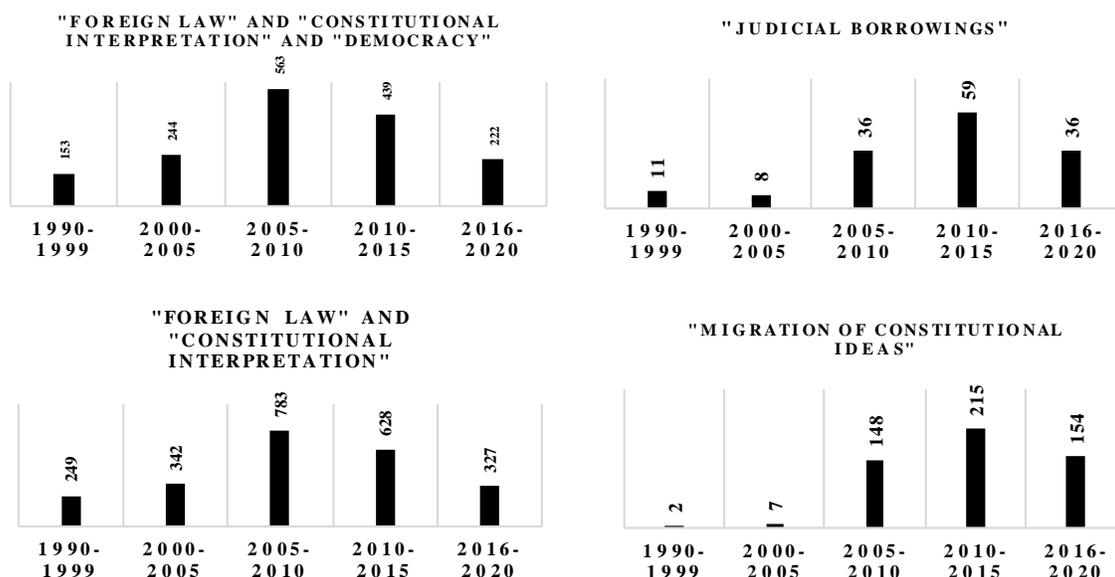
The results of this research are the product of a methodological exploration to address the automatic detection of references to foreign law in the judgments of the Constitutional Court of Colombia. More than 6000 constitutionality decisions issued between 1992 and 2020 were analyzed with a data mining approach, applying algorithms of semantic pattern matching, which were reinforced with human assistance. The data presented in this work are the first consolidated statistics about the use of foreign law in Colombia since the creation of Constitutional Court, over a period of 29 years. This effort aims to contribute to knowledge generation in the field of political science and comparative constitutional law, by allowing researchers to efficiently analyze extensive collections of unstructured data.

Keywords: comparative constitutional law, foreign law, natural language processing, text mining.

Introducción

En el campo del derecho constitucional comparado, el *diálogo entre tribunales constitucionales*¹ es una práctica que ha capturado la atención de académicos a nivel global (Hirschl, 2016). La resolución de conflictos constitucionales, es decir, disputas sobre los acuerdos fundamentales de la sociedad (Rawls, 1979), ha encontrado en la búsqueda y citación de jurisprudencia extranjera una oportunidad para ofrecer a los jueces marcos de referencia que les permitan orientarse en cómo decidir sobre una cuestión particular².

Figura 1. Tendencia de publicaciones académicas relacionadas con los patrones de citación de jurisprudencia extranjera para diferentes términos clave (elaboración propia). Fuente: HeinOnline.



Como fenómeno, los patrones de citación han sido ampliamente estudiados (ver Figura 1) y a pesar de que hay gran variedad de estudios que los abordan, de acuerdo con Ran Hirschl (2016), en la actualidad existe un vacío en cuanto a datos que permitan analizar las tendencias de este *diálogo constitucional*. Uno de los principales retos que se encuentra a la hora de abordar este problema es la naturaleza misma de la tarea, pues el análisis de conjuntos de datos no estructurados de gran volumen, históricamente ha requerido elevadas inversiones de tiempo por parte de los investigadores.

Esta ausencia de datos, así como las barreras existentes de volumen y tiempo tienen implicaciones en los diseños metodológicos y en la teorización sobre la citación de precedentes extranjeros. Aportes como los de Ran Hirschl (2016) desde el punto de vista de los enfoques metodológicos para el estudio profundo de esta práctica, primero, requieren que se conozcan las tendencias y frecuencias en la citación de jurisprudencia extranjera por parte de un tribunal constitucional. La necesidad de resolver

¹ Autores como David Law y Wen-Chen Chang (Law & Chang, 2011) argumentan que la citación de jurisprudencia extranjera no constituye un diálogo entre los tribunales constitucionales, y por lo tanto esta metáfora como referencia a la práctica podría resultar inadecuada.

² Las razones por las que los jueces se embarcan en la práctica de citación de jurisprudencia extranjera son amplias. Acá se presenta una consideración general con el fin de introducir el tema. En la siguiente sección se abordarán con mayor detalle estas cuestiones.

esta cuestión se encuentra justificada por el mismo autor: “Se sabe muy poco sobre el alcance de la práctica en regiones enteras, más específicamente en América Latina, África y grandes partes de Asia...” (Hirschl, 2016, p. 45).

Dicho escenario plantea la necesidad aplicar nuevos métodos de investigación para extraer datos sobre los patrones de citación. En ese sentido, los avances en procesamiento de lenguaje natural – NLP (por sus siglas en inglés) aparecen como un método razonable para resolver este problema³, permitiendo así impulsar la investigación en este campo y otros de las ciencias sociales. Aplicando algoritmos de búsqueda de patrones de lenguaje, este artículo explora cómo con técnicas de NLP se puede detectar automáticamente la presencia de referencias a jurisprudencia extranjera en las sentencias tipo *C*⁴ de la Corte Constitucional de Colombia entre 1992 y 2020.

Para llevar a cabo dicha tarea se definieron dos categorías de reglas o patrones que dan cuenta de una referencia a jurisprudencia extranjera en las sentencias. Empleando el algoritmo GloVe (Pennington, Socher, & Manning, 2014) se modeló un espacio vectorial de 50 dimensiones que dio como resultado una representación de las relaciones semánticas entre el vocabulario usado en las sentencias a partir de calcular sus coocurrencias y estadísticas globales. Este proceso permitió la selección de términos similares tras evaluar los cosenos de similitud entre los vectores de un conjunto de palabras clave y las palabras en el vocabulario. Los términos se clasificaron en términos jurídicos y verbos, y aquellos con mayor similitud fueron el insumo para la búsqueda de patrones.

Así, buscando una respuesta a la pregunta de ¿Cómo analizar efectiva y eficientemente la totalidad de la jurisprudencia emitida por la Corte Constitucional de Colombia para extraer datos sobre los patrones de referencia a jurisprudencia extranjera?, este trabajo presenta los siguientes aportes: (1) el desarrollo de una solución para acceder y detectar de forma sistemática referencias a jurisprudencia extranjera, (2) la presentación de estadísticas preliminares sobre el uso de precedentes extranjeros por parte de la Corte Constitucional entre 1992 y 2020, y finalmente, (3) retos, oportunidades y líneas de investigación futuras.

De una forma más general, también es posible afirmar que este trabajo es una contribución a la exploración y aplicación de técnicas propias de las ciencias de la computación para aportar datos que permitan dar respuestas a problemas de investigación en el campo de las ciencias sociales o el derecho, siendo esta investigación la primera que brinda datos sobre la práctica en Colombia, a partir de un estudio de extensivo de la jurisprudencia de la Corte Constitucional.

El documento se encuentra dividido en 5 secciones que presentan los elementos más frecuentes en el estudio de los patrones de referencia a jurisprudencia extranjera, trabajos similares que han aplicado técnicas de NLP al análisis de textos legales, la propuesta de un solución basada en técnicas de NLP para obtener datos sobre el fenómeno estudiado, la presentación de resultados sobre esta práctica en

³ Diferentes trabajos se han enfocado en las aplicaciones de algoritmos de NLP a diferentes áreas y problemas en el campo legal (Sadeghian, 2018) (Mandal, Ghosh, Ghosh, & Mandal, 2020) (Dyevre, 2020) (Boella, Di Caro, & Leone, 2019) (Nanda, 2019) (Lippi, 2019) (Yamada, Teufel, & Tokunaga, 2019), los más relevantes serán abordados más adelante.

⁴ El enfoque en las sentencias C -de Constitucionalidad- es una delimitación instrumental con el fin de desarrollar una solución sobre un conjunto de datos delimitado. En trabajos posteriores se analizarán las sentencias tipo T y SU.

Colombia, un análisis de los resultados a la luz de los patrones globales de la práctica estudiada y, finalmente, las conclusiones.

El estudio de los patrones de citación de jurisprudencia extranjera

En sistemas políticos con control de constitucionalidad por vía judicial⁵, el control y la resolución de disputas constitucionales generalmente están a cargo de un tribunal supremo. En estos sistemas existe la posibilidad del cambio constitucional por vía interpretativa, la cual tiene implicaciones sociales y políticas profundas. Aquí “(...) interpretes autorizados -los jueces son los intérpretes más importantes- les asignan un significado a las disposiciones constitucionales” (Bernal & Albert, 2016, p. 12)⁶. En este proceso interpretativo, una práctica de los jueces, que se cree cada vez más frecuente, es buscar decisiones judiciales extranjeras que se han enfrentado a problemas jurídicos similares (Perju, 2012).

Estas citaciones, además de ser meros puntos de referencia que permitan orientar al juez sobre un curso de decisión, pueden cumplir otra cantidad variada de roles. Hirschl (2016) plantea que estos van desde ser una autoridad persuasiva hasta propósitos decorativos, pasando por demostrar el éxito de una norma en el extranjero o como elemento de comparación. Sin embargo, que en la interpretación de la constitución los jueces se tomen la libertad de citar jurisprudencia extranjera⁷ abre la puerta a diversas discusiones como:

- La legitimidad democrática de que un juez interprete y, en consecuencia, modifique las normas constitucionales de acuerdo con valores políticos o jurídicos extranjeros (Keamey, 2019) (Halmani, 2012) (Rosenkrantz, 2003).
- Las razones por las que los jueces se embarcan en esta práctica, sean estas jurídicas, sociopolíticas, instrumentales, entre otras (Hirschl, 2016b) (Hirschl, 2016) (Perju, 2012) (Halmani, 2012).
- El alcance de la práctica y las frecuencias con las que las cortes se remiten a jurisprudencia extranjera, así como las áreas del derecho en las que esta práctica es más frecuente (McCrudden, 2000) (Jacobsohn, 2004) (Ferrari, 2019).

⁵ El rol de los tribunales constitucionales en las democracias contemporáneas ha sido objeto de amplios cuestionamientos. El debate gira en torno a las ideas de supremacía legislativa (Waldron, 2018) o la defensa del control de constitucionalidad (Gardbaum, 2016). Posturas como las de Waldron (2018) plantean que el control judicial de constitucionalidad atenta contra la democracia en la medida en que las decisiones tienen un impacto significativo sobre el conjunto general de la sociedad, pero que estas son tomadas por un grupo de personas (los jueces) que no necesariamente las representan y en consecuencia el llamado constituyente primario pierde relevancia en las decisiones que deberían ser colectivas. Por otra parte, la defensa del control de constitucionalidad encuentra su argumento más fuerte en la protección de la sociedad contra riesgos derivados de las dinámicas políticas (Gardbaum, 2016) o la protección de grupos minoritarios contra acciones legislativas que puedan vulnerar sus derechos (Waldron, 2018).

⁶ Carlos Bernal plantea la siguiente distinción: “Las disposiciones constitucionales son los enunciados de una Constitución escrita. Las normas constitucionales son el conjunto de significados que la Constitución escrita expresa o que han sido aceptados por convenciones constitucionales no escritas” (Bernal & Albert, 2016, p. 9-10).

⁷ Este trabajo asume que la citación de un precedente extranjero se realiza de forma voluntaria por parte del juez. Es decir, no existe ninguna disposición que haga obligatorio el uso de estos referentes en la toma de decisiones judiciales. Tampoco se incluye la citación de normas o jurisprudencia relacionada con el ámbito del derecho internacional ni de tribunales supranacionales, por lo tanto, el enfoque está centrado en la referenciación de jurisprudencia entre tribunales en el ámbito nacional.

- La precisión de una variedad de conceptos y metáforas para referirse a la práctica tal como se evidencia en el trabajo de Vlad Perju (2012) al aproximarse a los conceptos de trasplantes, préstamos y migraciones constitucionales o en la crítica de Law & Chang (2011) al concepto de diálogo constitucional global.

Un problema que ha sido constante en las investigaciones frente a los puntos mencionados es que “(...) los supuestos conocimientos universales [sobre esta materia] están basados en un puñado de casos estudiados a menudo, y no siempre son casos o escenarios representativos. Las consideraciones instrumentalistas como la disponibilidad de los datos o la planeación de carrera con frecuencia determinan cuales casos se abordan” (Hirschl, 2016, p. 19). En ese sentido, si bien es natural encontrar limitaciones en estudios que requieren el análisis manual de grandes conjuntos de datos, sobre las conclusiones de los trabajos hasta ahora publicados, Hirschl se cuestiona:

¿Qué tan universales, representativas o generalizables son las lecciones de un cuerpo de conocimiento que casi nunca se refiere a la experiencia constitucional, al derecho ni a las instituciones de más del noventa y cinco por ciento del universo constitucional...? (Hirschl, 2016, p. 254)

Este problema de los casos de estudio es evidente en la concentración de las publicaciones que giran en torno a un grupo pequeño de países. En el libro *Judicial Cosmopolitanism - The Use of Foreign Law in Contemporary Constitutional Systems*, publicado en 2019 y que recopila estudios sobre los patrones de citación alrededor del mundo, se dedican 22 capítulos a Europa, 2 a Asia y 1 a América Latina (Ferrari, 2019)⁸.

Sobre este último grupo de países, en la literatura los casos más prominentes son los de Argentina, Brasil, Chile y México, donde también se observa el problema de las afirmaciones *universalistas*. Este es el caso de la investigadora italiana Francesca Polacchini (2019) que afirma dar cuenta de la frecuencia con la que se hace uso de jurisprudencia extranjera en América Latina, pero en su trabajo cae en una generalización excesiva al establecer tendencias para toda la región con base en el análisis de 4 casos. En contraste, Rosenkrantz (2003) y Zelada (2012) abordan el caso de Argentina y Chile respectivamente, con mayor profundidad y definiendo claramente el alcance de sus estudios.

⁸ De los casos estudiados con mayor frecuencia, resaltan Sudáfrica, Israel y Alemania (Halmani, 2012) (Hirschl, 2016b) (Ferrari, 2019). El interés de la comunidad académica en ellos se da por los contextos políticos de cada uno, siendo Alemania y Sudáfrica casos de transición de modelos autoritarios a democracias constitucionales (Halmani, 2012).

El caso de Israel también llama la atención puesto que no cuenta con una constitución escrita y los jueces en constantemente, más que hacer una defensa de los valores de su sociedad, pueden estar buscando importar valores en función de sus aspiraciones. El profesor Hirschl explica: “En lo que respecta al uso voluntario de fuentes extranjeras en democracias emergentes o en contextos constitucionales disímiles, las decisiones judiciales sobre qué citar y qué no citar no se pueden comprender sin tener en cuenta la perspectiva de las cortes y los jueces constitucionales en cuanto a su propia "identidad" -el esfuerzo por definir quiénes somos "nosotros" como comunidad política y como se expresa de manera pública "la imagen" que queremos proyectar o el "lugar" que ocupamos o deberíamos ocupar en el mundo” (2016b, p. 120).

También es importante mencionar el caso de Estados Unidos, relevante por el debate entre universalistas y originalistas (Halmani, 2012), así como el caso de Canadá que es reconocido como uno de los mayores exportadores de jurisprudencia a nivel global (Hirschl, 2016).

Sobre el caso colombiano, del que se ocupa este trabajo, se encontró una publicación de la investigadora Lina Marcela Escobar (2008) titulada *El Uso del Precedente Extranjero por parte de la Corte Constitucional Colombiana*, que se enmarcó en el estudio global *Cross-Judicial Fertilization*, que en Colombia buscaba estudiar toda la jurisprudencia de la Corte Constitucional entre 1992 y 2008 con el fin de establecer la frecuencia de la práctica en Colombia. Una característica metodológica que debe ser resaltada sobre dicha investigación es su énfasis en la revisión manual de las sentencias con el apoyo de un grupo de investigadores liderados por Escobar.

Esta esperaba entregar sus resultados de investigación a través de diferentes publicaciones de las cuales el único producto de investigación hallado es el ya mencionado. En ese sentido, un aspecto a resaltar sobre dicho proyecto, a pesar de no haber sido concluido -hasta donde se conoce-, es que logró establecer pautas metodológicas para el estudio de la práctica en Colombia. Así, a la hora de abordar el estudio de los patrones de referencia a jurisprudencia extranjera, presentó una serie de indicadores a evaluar. Algunos de estos son:

1. ¿La referencia está en las opiniones dominantes de la sentencia o en un salvamento de voto?
2. ¿Es una mera referencia o una cita textual?
3. ¿En cuales argumentos se usó la cita?

Es, entonces, razonable que una investigación sobre la práctica considere estos indicadores y de igual forma, de acuerdo con la literatura revisada y las discusiones que existen al respecto (McCrudden, 2000) (Jacobsohn, 2004) (Ferrari, 2019) (Hirschl, 2016), tenga en cuenta:

4. ¿Cuáles son las áreas del derecho en las que esta práctica es más frecuente?
5. ¿Cuáles son los jueces que más recurren a jurisprudencia extranjera en sus decisiones?⁹

Sin embargo, en este punto cabe aclarar que, si bien la gama de variables que se pueden estudiar es amplia y el valor que pueden agregar para llevar a cabo estudios cualitativos es alto, es importante tener en cuenta que la mera tarea de detectar referencias o citas de forma automática es en sí misma compleja y en consecuencia el objetivo general de esta investigación es únicamente identificarlas, dejando así para trabajos posteriores la resolución de las cinco variables expuestas anteriormente. Por lo tanto, se busca extraer la frecuencia con la que la Corte se remite a decisiones judiciales de otros países, así como el número serial de las sentencias de la Corte Constitucional colombiana donde este fenómeno se presenta.

Trabajos relacionados

Teniendo en cuenta las limitaciones expuestas en la sección anterior para el estudio de los patrones de citación de jurisprudencia extranjera, a continuación, se presentan investigaciones relevantes directamente relacionadas con los métodos aplicados en este trabajo. Estas investigaciones han

⁹ En el caso colombiano las sentencias de la Corte son desarrolladas tanto en forma y contenido por el magistrado ponente y después se discute la decisión en la sala plena de la Corte con el fin de tomar la decisión final sobre el caso. De igual forma, se da espacio para que los magistrados registren salvamentos de voto en caso de que sus opiniones sean contrarias a las opiniones dominantes en la Corte. En ese sentido, identificar qué magistrados recurren con mayor frecuencia a la citación de jurisprudencia extranjera puede ayudar a descubrir patrones subyacentes, por ejemplo, de uso estratégico de la jurisprudencia extranjera para sustentar una decisión, en contraste con un ejercicio comparativo.

abordado el estudio de textos legales desde una perspectiva computacional con el fin de hacer más eficiente el proceso de extracción de datos relevantes para diferentes tipos de investigaciones en el ámbito jurídico y de las ciencias políticas.

Dentro de las consultas realizadas a diferentes bases de datos académicas no se encontró una aproximación sistemática a la extracción de citas de jurisprudencia extranjera, sin embargo, hay una base amplia de trabajos que han abordado problemas similares en relación con la detección automática de citas o referencias en textos legales (Sadeghian, 2018) (Mandal, Ghosh, Ghosh, & Mandal, 2020) (Dyevre, 2020) (Boella, Di Caro, & Leone, 2019) (Nanda, 2019) (Lippi, 2019) (Yamada, Teufel, & Tokunaga, 2019). De las publicaciones encontradas tres resultan particularmente relevantes:

- Con el fin de lograr la detección sistemática de la referencia a un texto legal, Sannier et al. (2016) crearon un modelo que se basaba en analizar las palabras al costado izquierdo de la referencia. En ese sentido, el modelo predecía la intención de referencia dada una serie de palabras. Para ello, se basaron en el desarrollo de un modelo semántico del texto legal con técnicas de NLP y adicionalmente aplicaron un sistema de clasificación basado en patrones de lenguaje. En su trabajo demuestran que este enfoque es efectivo para la detección de referencias entre textos legales con una precisión del 84.48% en un conjunto de datos con 2585 referencias.
- Posteriormente, los mismos autores analizando la citación de legislación en textos legales abordan el problema a través de un sistema de detección de patrones basados en términos clave. Dichos términos cubren palabras específicas como art, num, clause, títulos de leyes y nombres que deben ser manejados como un todo (Sannier, Adedjouma, Sabetzadeh, & Briand, 2017). La diferencia principal con el enfoque empleado en el trabajo de 2016 es que ese se fundamentaba en las relaciones semánticas, mientras que el de 2017 se basa en encontrar coincidencias con base en expresiones regulares¹⁰ sofisticadas.
- Zhang y El-Gohary (2016) analizaron métodos que permitiesen extraer requerimientos regulatorios de textos legales con el fin de que estos puedan ser verificados de forma automática. Para ello, los autores desarrollaron un método de NLP con un enfoque basado en reglas o patrones de lenguaje manualmente definidos que permitieron extraer información de los textos combinando características semánticas y gramaticales.

Método propuesto

Con el fin de detectar automáticamente la presencia de referencias a jurisprudencia extranjera, se crearon una serie de algoritmos con base en los enfoques propuestos por Sannier et al. (2016), Sannier et al. (2017) y Zhang y El-Gohary (2016). En concreto, se desarrollaron dos algoritmos en el lenguaje de programación Python para ejecutar búsquedas de patrones que toman en cuenta la presencia de términos clave y factores contextuales dados por palabras, símbolos y números. A continuación, se

¹⁰ Las expresiones regulares son una forma de codificar patrones de texto para ejecutar búsquedas de alta velocidad sobre conjuntos de datos no estructurados -por ejemplo, documentos-. Una aplicación común de este tipo de expresiones es encontrar correos electrónicos donde los elementos clave del patrón son: “@” y “.com”. En el caso del estudio en cuestión estos patrones incluían la presencia de palabras como “num” y “art” para representar un patrón de texto que identificara correctamente el numeral de un artículo de una ley.

presenta el enfoque propuesto de la siguiente forma: (1) se abordan los casos de referencia a jurisprudencia extranjera en el año 1992, se expone la estructura de dos patrones o reglas y se presenta (2) el flujo de procesos para la extracción de los datos en las sentencias.

Definición de los patrones

Una aproximación inicial al problema de identificar las referencias a jurisprudencia extranjera fue realizada con base en las sentencias del año 1992. Específicamente se abordó el caso de la sentencia C-599-92¹¹ en la que se encuentran 4 referencias. El objetivo de esta aproximación inicial era identificar las formas en las que el fenómeno se podría presentar en el texto y definir su estructura de forma abstracta a nivel semántico y gramatical con el fin de determinar las tareas a ejecutar por el algoritmo.

Aquí se encontró que existen dos formas en las que una referencia a jurisprudencia extranjera se puede presentar: (a) como citación en el cuerpo del texto y (b) como elemento en la zona de referencias de la sentencia. La Tabla 1 expone los dos casos:

Tabla 1: ejemplos de referencias a jurisprudencia extranjera

Casos A: referencia como citación el cuerpo del texto	Casos B: como elemento en la zona de referencias de la sentencia
<p>A.1. Así mismo, ya más recientemente, <u><i>el Tribunal Constitucional Español ha sostenido</i></u> lo siguiente a propósito de la aplicación del debido proceso a ordenamientos diferentes al penal... (Sentencia C-599-92, 1992)</p> <p>A.2. <u><i>La jurisprudencia norteamericana por su parte afirmó</i></u> sobre el fundamento aportado por la enmienda XIV de la Constitución acerca del debido proceso legal, entendido en sentido procesal... (Sentencia C-599-92, 1992)</p> <p>A.3. <u><i>El Tribunal Constitucional Español ha manifestado</i></u> al respecto... (Sentencia C-599-92, 1992)</p>	<p>B.1. <u><i>5 Tribunal Constitucional Español</i></u> (sic). Sentencias de 30 de enero y 8 de junio de 1981. (Sentencia C-599-92, 1992)</p> <p>B.2. <u><i>8 Tribunal Constitucional Español.</i></u> Sentencia de marzo 8 de 1985. (Sentencia C-599-92, 1992)</p> <p>B.3. <u><i>10 Tribunal Constitucional Español.</i></u> Sentencia de abril 26 de 1990. (Sentencia C-599-92, 1992)</p>

Un análisis de ambos casos deja ver fácilmente que existen elementos tanto comunes, como diferenciales que deben ser tomados en cuenta a la hora de definir la estructura abstracta del patrón de lenguaje. De acuerdo con las referencias expuestas en los casos A y B, se identificaron las siguientes características comunes:

¹¹ Las sentencias del año 1992 fueron analizadas empleando una búsqueda de países y gentilicios. Con los resultados de ese proceso se realizó una búsqueda manual de referencias a jurisprudencia extranjera que permitió identificar los cuatro casos contenidos en la sentencia C-599-92.

- La presencia del país al que se hace referencia, que aparece en forma de gentilicio.
- La presencia de términos del ámbito jurídico como jurisprudencia, tribunal o constitucional.

Los elementos que distinguen entre sí a los casos A y B son:

- En el caso A la presencia de un verbo dado que la expresión está contenida en el desarrollo argumental de la sentencia.
- En el caso B la presencia de un número al inicio de la línea que puede estar solo o contenido en unos paréntesis cuadrados “[]”, dado que la expresión buscada se encuentra en la zona de referencias y aclaraciones a pie de página de la sentencia.

En consecuencia, teniendo en cuenta los elementos distintivos de cada caso se conformaron dos reglas que su vez tienen variaciones, con el fin de lograr una búsqueda amplia que intente capturar la complejidad con la que se pueden expresar las ideas en el idioma español, sin perder los elementos que constituyen los patrones.

Tabla 2: reglas o patrones de lenguaje

Método A	Método B
A1: <Términos jurídicos> + <País o Gentilicio> + <Verbo>	B1: <Símbolo o número> + <Término jurídico> o <País o Gentilicio>
A2: <País o Gentilicio> + <Verbo> o <Términos jurídicos>	
A3: <Verbo> o <Términos jurídicos> + <País o Gentilicio>	

Los métodos A y B, con sus respectivas variaciones cuentan con diferentes variables contenidas en los símbolos de diple “<>”. Estas se pueden clasificar en aquellas que tienen valor semántico y las que no. En el primer grupo se encuentran: <Términos jurídicos>, <País o gentilicio> y <Verbo>; en el segundo grupo se encuentra: <Símbolo o número>. Las variables de carácter semántico se resolvieron¹² de la siguiente forma:

- <País o gentilicio>: se elaboró un listado con la totalidad de países reconocidos en el sistema de Naciones Unidas y se agregaron los gentilicios correspondientes.
- <Términos jurídicos> y <Verbo>: se creó un modelo de lenguaje de las sentencias utilizando el algoritmo de aprendizaje no supervisado GloVe¹³ (Pennington, Socher, & Manning, 2014). Este crea una representación del lenguaje en un espacio vectorial de 50 dimensiones tomando ventaja de las estadísticas globales y conteos de coocurrencia de las palabras usadas por la Corte en sus sentencias. De esta forma a cada palabra única en las sentencias le corresponde un vector y evaluando cosenos de similitud entre estos se pueden obtener los términos que son similares entre sí. Este procedimiento se realizó con el fin de

¹² Resolver la variable implica asignarle un valor específico. Por ejemplo, a medida que el algoritmo se ejecute la variable <País> irá tomando valores como Colombia, colombiano, colombiana, Portugal, portugués o portuguesa.

¹³ Los algoritmos de aprendizaje automático se clasifican en tres tipos de aprendizaje: supervisado, no supervisado y reforzado. Los algoritmos no supervisados, como el empleado en este trabajo, tienen como objetivo descubrir patrones subyacentes en la información que analizan sin ningún tipo de asistencia humana. En este caso, el objetivo del algoritmo es descubrir las relaciones semánticas en el vocabulario de un texto de forma autónoma, bajo el supuesto de que las términos circundantes a una palabra codifican de forma implícita información sobre su significado, por lo que palabras que aparecen en contextos similares tienen un intenciones semánticas similares (Pennington, Socher, & Manning, 2014).

evitar la selección arbitraria de términos de forma manual. Para el proceso de entrenamiento y creación del modelo de lenguaje todos los hiper-parámetros del algoritmo fueron los recomendados por los autores y que vienen configurados por defecto en el repositorio del código¹⁴.

Confirmación de los patrones

Con el modelo de lenguaje se crearon tres listas de palabras que cumplen un rol semántico en los patrones, de acuerdo con las categorías definidas. Si se toman como ejemplo las referencias en la tabla 1, las listas quedarían constituidas de la siguiente forma:

- <País o gentilicio>: como España, español, española, Estados Unidos de América, Estados Unidos, estadounidense, norteamericano, norteamericana.
- <Términos jurídicos>: como tribunal, constitucional, jurisprudencia.
- <Verbo>: como manifestado, afirmó, sostenido.

Estas listas sirven como punto de referencia para filtrar las frases que mencionan a un país y que expresan una referencia a jurisprudencia extranjera. Al momento de buscar un patrón el algoritmo va a ejecutar la siguiente secuencia:

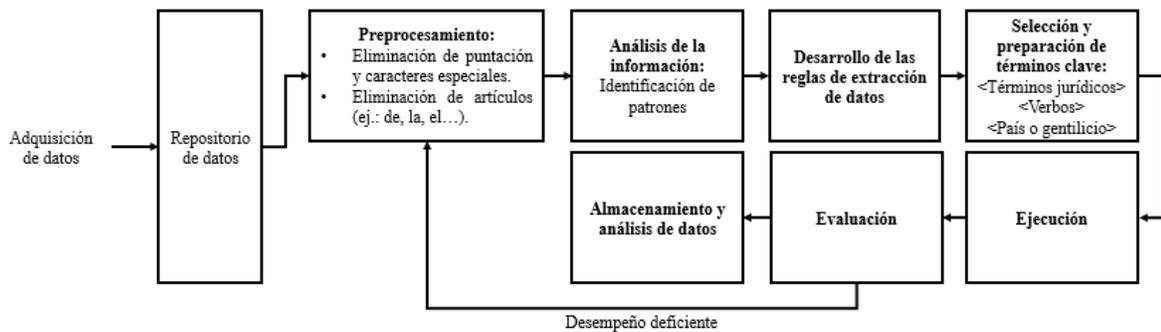
1. Va a buscar cada uno de los países o gentilicios en las sentencias.
2. Si encuentra una coincidencia enseguida va a buscar en las palabras circundantes al país la presencia de <Términos jurídicos> y <Verbos> de acuerdo con el orden establecido en los métodos A1, A2, A3. El algoritmo busca máximo 5 palabras en cada dirección (izquierda y derecha), aunque este parámetro es modificable.
En el caso del método B1 la condición fijada era que en el mismo párrafo estuviesen presentes al inicio <Símbolo o número>, el <País o gentilicio> y el <Término jurídico>, sin tomar en cuenta la distancia entre los términos.
3. Si las palabras circundantes al país coinciden con alguna de las palabras en la lista de <Términos jurídicos> y <Verbos>, la frase se contará como una referencia a jurisprudencia extranjera y se almacenará en una tabla que contiene el serial de la sentencia y el párrafo o la frase donde se confirmó el patrón. En el caso del patrón B1 si todos los elementos estaban presentes de igual forma se almacenaba la información en la tabla correspondiente.

Flujo de procesamiento y análisis

Una vez definidos los dos componentes clave del proceso de búsqueda de patrones, estos pueden ser incorporados en un flujo de procesamiento de datos estándar, que fue adaptado de Zhang y El-Gohary (2016) y siguiendo las pautas establecidas por Rodman (2019) para el preprocesamiento de los datos. Las etapas del flujo de procesamiento abarcan desde la adquisición de los datos, en este caso cada una de las sentencias a analizar, hasta la presentación de resultados, tal como se observa en la figura 2.

¹⁴ Ver: <https://github.com/stanfordnlp/GloVe>

Figura 2: flujo de procesamiento de los datos, adaptado de Zhang y El-Gohary (2016).



A continuación, se especifican los detalles de cada una de las etapas del proceso:

- **Adquisición y repositorio de datos:** la adquisición de las sentencias se realizó mediante una automatización de *web-scraping* utilizando el lenguaje de programación Python y el módulo Selenium (Chapagain, 2019) que entraba a la página web de cada una de las sentencias C emitidas entre 1992 y 2020 y almacenaba los párrafos en archivos Excel que contenían todas las sentencias de un año.
- **Preprocesamiento:** siguiendo las pautas establecidas por Rodman (2019) se aplicaron dos técnicas de limpieza y estandarización de la información que incluyeron la eliminación de la puntuación y de los caracteres especiales. De igual forma se eliminaron todos los artículos de los párrafos (ej.: el, la, de...) debido a que no agregan significado y sí contribuyen en el modelo GloVe a tener una medida de similitud altamente desproporcionada entre las palabras debido a la alta frecuencia con la que los artículos están presentes en el cuerpo de un texto.
- **Análisis de la información y desarrollo de las reglas de extracción de datos:** con base en los análisis de los patrones planteados al inicio de esta sección, en el lenguaje Python se codificaron los métodos definidos (A y B) para analizar cada uno de los párrafos de las sentencias contenidas en el repositorio.
- **Selección y preparación de términos clave:** las palabras de las referencias a jurisprudencia extranjera presentadas en la Tabla 1 se usaron como términos clave para evaluar los cosenos de similitud con todas las palabras contenidas en las sentencias y conformar las listas que se usaron en la ejecución de la búsqueda de patrones. En esta etapa hubo una revisión manual con el fin de evitar que términos no relacionados con la intención semántica de cada lista generaran resultados erróneos al momento de la ejecución de los algoritmos.
- **Ejecución:** la ejecución de los algoritmos se realizó en la plataforma Google Colab Pro que provee un entorno de ejecución para los algoritmos en la nube de Google y que brinda recursos computacionales de alto desempeño como GPUs o TPUs a bajo costo.
- **Evaluación:** en la evaluación de los resultados una limitación de este trabajo es conocer la precisión global de la solución a través de tres métricas convencionales: *accuracy*, *recall* y *F1 score*. De las tres, únicamente se evaluó el *accuracy*, que depende directamente de los resultados que arroja el algoritmo. En el contexto de este trabajo, esta métrica mide la proporción de resultados obtenidos que efectivamente corresponden a

referencias de jurisprudencia extranjera sobre el total de resultados obtenidos, permitiendo obtener un valor de precisión.

Si bien dicha medida es relevante, también es cierto que puede exagerar la noción de éxito o fracaso de la solución. Para superar este problema, las dos métricas complementarias permiten evaluar el desempeño global de los algoritmos en dos pasos. Primero, a través del *recall* se mide la cantidad de resultados positivos sobre el total de resultados esperados. Segundo, con el *F1 score* se obtiene una medida ponderada del *accuracy* sobre el *recall*. Dado que al momento de iniciar la ejecución de la solución no se conocían el total de referencias a jurisprudencia extranjera en las sentencias C entre 1992 y 2020, no fue posible evaluar estas dos métricas pues no se contaba la información necesaria para ello -la cantidad de referencias-.

Por lo tanto, los resultados presentados en este estudio son preliminares en la medida que existe incertidumbre sobre si los datos obtenidos representan o no la totalidad de las referencias a jurisprudencia extranjera en las sentencias C de la Corte¹⁵. Este problema puede ser visto como un dilema en donde: (a) una adquisición inicial de datos manual hubiese excedido cualquier margen de tiempo razonable para concluir la tarea, y (b) una solución inicial basada en métodos automáticos para extraer datos conduciría, inevitablemente, a una situación de incertidumbre puesto que no se podría contar con datos suficientes para evaluar completamente la solución.

En consecuencia, puestos a escoger, la conclusión en este caso es que la eficiencia de la opción *b* con lleva a mejores resultados, debido a que en el marco de una solución basada en ciencia de datos, este trabajo es la primera parte de un círculo virtuoso donde la adquisición inicial de un conjunto de datos producirá un algoritmo con *x* precisión, lo que permite adquirir nuevos datos, mejorar el diseño de la solución y obtener mejores resultados de forma iterativa (Agrawal, Gans, & Goldfarb, 2019).

- **Almacenamiento y análisis de los datos:** tras la ejecución, el algoritmo arroja como resultado un archivo en formato de texto plano con las frases o párrafos que detectó positivamente como referencias a jurisprudencia extranjera y el número de sentencia correspondiente. Estos archivos fueron revisados de forma manual para extraer la métrica de precisión de la solución. Finalmente, se construyeron una serie de reportes en la herramienta Power BI que permite crear visualizaciones interactivas de los datos recogidos y así facilitar su análisis.

Resultados

Las Tabla 3 presenta un resumen de los metadatos que describen el comportamiento de la solución tras la ejecución del flujo completo de procesamiento de las sentencias. Estos resultados, así como las estadísticas sobre los patrones de referencia a jurisprudencia extranjera, fueron extraídos después de un proceso de revisión manual de los resultados. Sobre este aspecto cabe anotar que al momento de codificar las reglas y realizar diferentes experimentos se consideró que una solución óptima debería identificar los casos de referencia a jurisprudencia extranjera en el año 1992, específicamente los de

¹⁵ Es importante resaltar que los resultados de este trabajo constituyen una línea base sobre la cual evaluar actualizaciones o soluciones posteriores, pero este proceso requiere un análisis detallado de las sentencias que se obtengan con la solución actual, con el fin de crear una base de datos precisa con la información suficiente para evaluar tanto el *recall* como el *F1 score*.

la sentencia C-599-92. En ese sentido, el método A1 identificó correctamente los 4 casos en el cuerpo del texto y el método B1 identificó correctamente 2 de 4¹⁶.

Tabla 3. Metadatos de la solución

Número de sentencias analizadas	Número de párrafos procesados	Número de palabras procesadas	
6.786	3.977.198	104.405.026	
Número de sentencias con referencias a jurisprudencia extranjera			Número de referencias a jurisprudencia extranjera
190			473
Precisión de los métodos			
Método	Número de casos identificados	Número de casos confirmados	Precisión (accuracy)
A1	22	22	100%
A2	-	-	-
A3	119	5	4,2%
B1	3.387	446	13,16%
Total	3.528	473	13.40%

Dado que el tribunal estudiado no cuenta con un mecanismo que permita acceder directamente a sus repositorios de información por medio de APIs, ni tampoco es pública una base de datos donde estén contenidos los seriales de cada una de las sentencias emitidas, el proceso de recolección de datos se ejecutó empleando técnicas *web-scraping* que permitieron acceder a las sentencias iterando sobre el número de serie en la estructura de la URL de las sentencias, un total de 5000 veces por año¹⁷. Una evaluación de los resultados de este método reveló que el literal C en la URL puede estar en mayúscula o minúscula (ver nota 14) y que además algunas sentencias solo están disponibles en una de las dos configuraciones. Como resultado, la cantidad de sentencias obtenidas es diferente de la cifra publicada por la Corte Constitucional en su sitio web.

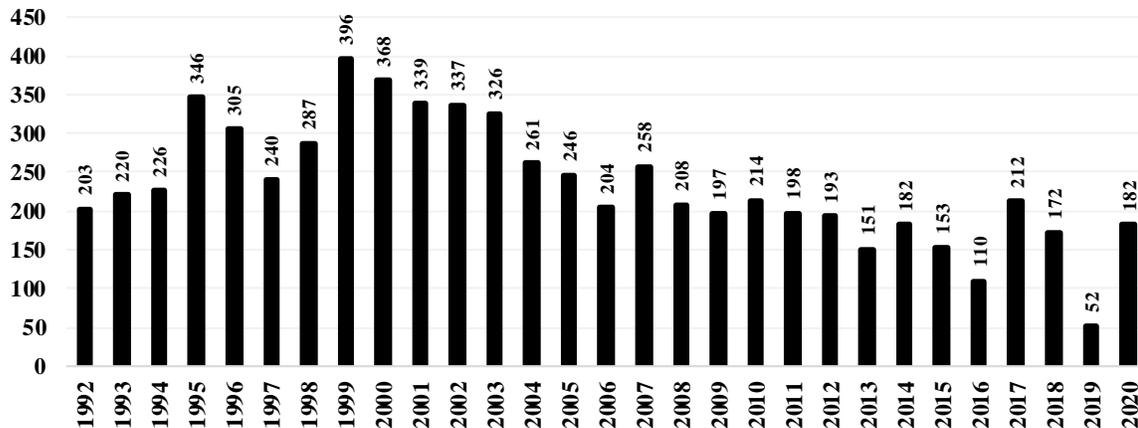
En total se analizaron 6.786 sentencias tipo C de un total de 6.845 sentencias de acuerdo con las estadísticas de la Corte Constitucional, lo que representa una diferencia de 59 sentencias, hecho que es explicado por las particularidades de la URL ya expuestas y lo que parece ser una estructura no

¹⁶ Sobre el método B1 se debe tener en cuenta que la referencia a jurisprudencia norteamericana provenía de una cita de doctrina, lo que de por sí impedía que esa referencia fuera identificada por este método. También, una de las referencias al Tribunal Constitucional Español tenía un error ortográfico, razón por la cual esa referencia tampoco fue identificada. En consecuencia, cuando el algoritmo logró identificar los casos esperados, se procedió a ejecutar la solución completa.

¹⁷ URL ejemplo: <https://www.corteconstitucional.gov.co/relatoria/1992/C-004-92.htm>

estandarizada de archivos y carpetas en los sistemas de información de la Corte. La Figura 3 muestra la cantidad de sentencias procesadas por año¹⁸:

Figura 3: número de sentencias analizadas por año



Con respecto a los casos de referencia a jurisprudencia extranjera identificados por el algoritmo, se debe resaltar que el 94,29% de los resultados confirmados se concentran en el método B1, 4,66% en el método A1 y 1,06% en el método A3. Sobre el desempeño del método B1, que concentra la mayor cantidad de resultados confirmados, una primera hipótesis, de acuerdo con los datos obtenidos, apunta a que la baja precisión de este patrón está relacionada con la ausencia de un parámetro de distancia máxima de los términos circundantes en relación con el país o gentilicio¹⁹, así como que los resultados de este método no correspondían únicamente a la zona de referencia o pies de página de la sentencia, sino también al cuerpo del texto.

Con respecto a los métodos de clase A, se debe tener en cuenta que solo el A1 cumplía con exactitud la estructura de las referencias encontradas en el año 1992 y que los métodos A2 y A3 se plantearon como una alternativa que permitiese capturar más información, aunque no respondían en sí a una estructura o patrón de lenguaje concreto, sino al supuesto de que la estructura de la idea que se expresa con el patrón A1 puede tener diferentes variaciones con respecto a la posición de los verbos y de los términos jurídicos, algo confirmado por el método A3.

Otro aspecto relevante a tener en cuenta sobre los métodos A es que la cantidad de palabras usadas para el entrenamiento del modelo de lenguaje GloVe es baja en contraste con el modelo en inglés desarrollado por los autores de este algoritmo. La cantidad de palabras usadas para crear las estadísticas globales y conteos de coocurrencias en este trabajo está en la magnitud de los cientos de millones, mientras que la cantidad de palabras que los autores emplearon está en la magnitud de los miles de millones (Pennington, Socher, & Manning, 2014).

Este factor tiene un impacto al momento de evaluar los cosenos de similitud de los términos clave con el resto de las palabras en el vocabulario y por lo tanto en la cantidad de palabras en cada lista, lo que a su vez podría afectar la cantidad de veces que el patrón se confirma con los métodos A. En ese

¹⁸ Para una revisión de las estadísticas de la Corte Constitucional ir a: <https://www.corteconstitucional.gov.co/lacorte/estadisticas.php>

¹⁹ La exclusión de un parámetro de distancia máxima se decidió con base en los casos del año 1992 (ver Tabla 1).

sentido para trabajos posteriores es recomendable entrenar el modelo GloVe sobre una base más amplia de textos, por ejemplo, tomando en cuenta también las sentencias tipo T y SU, o incluso ampliando la base de datos a sentencias de la Corte Suprema de Justicia y el Consejo de Estado.

Tabla 4: resultado de la evaluación de cosenos de similitud con base en términos clave²⁰

Verbos		Términos Jurídicos	
Términos clave	Resultado	Términos clave	Resultado
Sostenido	Sostenido	Jurisprudencia	Corporación
Manifestado	Señalado	Tribunal	Jurisprudencia
Afirmado	Explicado	Constitucional	Tribunal
Afirmó	Reiterado	Corte	Corte
	Manifestado	Suprema	Constitucional
	Afirmado	Supremo	Sala
	Dijo		Suprema
	Afirmó		Supremo
	Indicó		
	Sostuvo		
	Explicó		
	Manifestó		
	Advirtió		

Los resultados del modelo lenguaje expuestos en la tabla 4 con respecto a la extracción de términos similares entre sí mostró un desempeño sobresaliente, sin embargo, en esta etapa también hubo un proceso de revisión manual de los términos incluidos por el modelo en cada lista²¹. De igual forma, se determinó de forma empírica que el umbral mínimo para los cosenos de similitud era de 0.75 para verbos y 0.8 para términos jurídicos.

Una conclusión general sobre el desempeño del método propuesto es que, si bien permitió identificar casos de referencia a jurisprudencia extranjera, la cantidad elevada de resultados hace que su función sea la de asistir a un investigador en el proceso de extracción y filtrado de información de las sentencias mediante el uso de patrones de lenguaje. Sobre ese punto se debe tener en cuenta que a pesar de que se requirió un proceso de revisión humana, la cantidad de horas necesarias para filtrar los resultados fue menor a 40, que los resultados analizados de esta forma representan el 0,088% de los párrafos procesados por el algoritmo y que solo el 0.011% de los párrafos daban cuenta de la frecuencia de la práctica. También se debe notar que de 6.786 sentencias en tan solo 190 se identificaron referencias a *jurisprudencia* extranjera, lo que representa el 2,79% de las sentencias.

Finalmente, de acuerdo con los resultados obtenidos, la solución demostró ser útil en aplicaciones diferentes a los de detectar referencias a jurisprudencia extranjera, por ejemplo, identificando casos relacionados con tratados internacionales, casos de citación a tribunales internacionales como la Corte Interamericana de Derechos Humanos, la Corte Internacional de Justicia o el Tribunal Europeo de

²⁰ Los resultados presentados en la tabla fueron los empleados por el algoritmo después de un proceso de revisión manual de los términos.

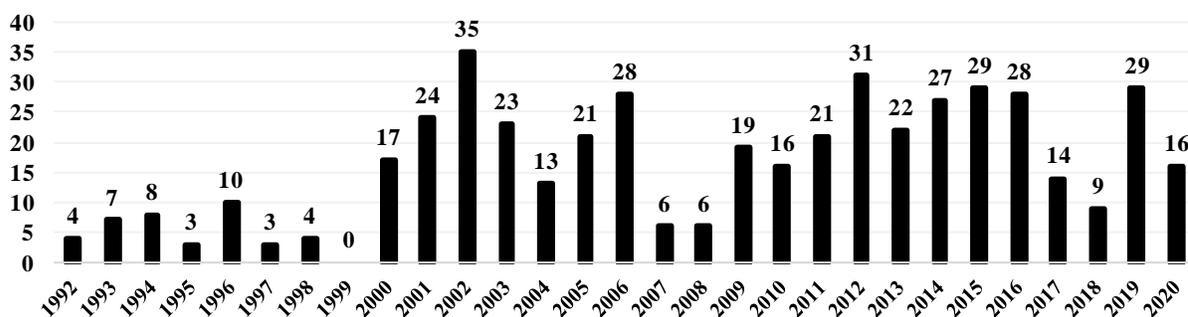
²¹ Los términos que no estaban relacionados con la intención semántica de cada lista son: (a) Verbos: corporación, jurisprudencia, ocasiones; (b) Términos Jurídicos: sostenido.

Derechos Humanos y, eliminando el factor país, podría tener un buen desempeño para obtener autorreferencias de la Corte Constitucional en sus sentencias.

Frecuencia de las referencias a jurisprudencia extranjera en las sentencias tipo C²²

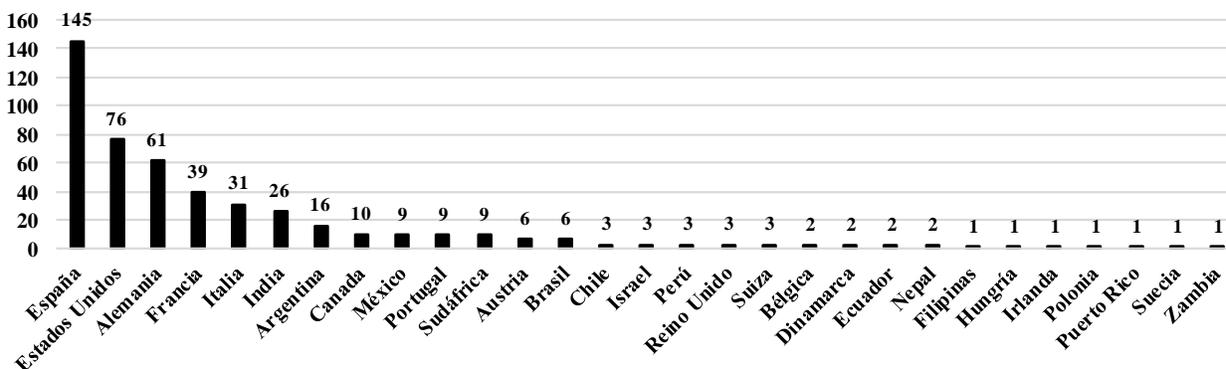
A nivel general los datos arrojados por el algoritmo permitieron identificar que entre el año 1992 y 2020 hubo un total de 473 referencias a jurisprudencia extranjera en decisiones de la Corte Constitucional, distribuidas en 190 sentencias tipo C. Se identificó que el año con la mayor cantidad de referencias fue el 2002 con 35 en total y 1999 fue el único año en el que no se encontró presente esta práctica en ninguna de las 396 sentencias analizadas. Adicionalmente, las cifras muestran que en los primeros años de la Corte Constitucional la práctica no era demasiado frecuente y a partir del 2000 se evidenció un crecimiento considerable en la cantidad de referencias que se ha mantenido relativamente constante a excepción de los años 2007, 2008 y 2018.

Figura 4: Número de referencias a jurisprudencia extranjero por año 1992 - 2020



De igual forma, se encontró que la Corte Constitucional ha recurrido un total de 145 veces a jurisprudencia española, siendo este país el más referenciado, seguido por Estados Unidos, Alemania, Francia e Italia, siendo todos países del hemisferio occidental. Con respecto a países del hemisferio oriental, los más citados por la Corte son India, Filipinas, Nepal y Zambia, siendo India el más citado con un total de 26 citas, estando por encima de países como Canadá. La figura 5 presenta el listado completo de países referenciados en las sentencias tipo C en materia de jurisprudencia.

Figura 5: País de origen de la jurisprudencia extranjera citada

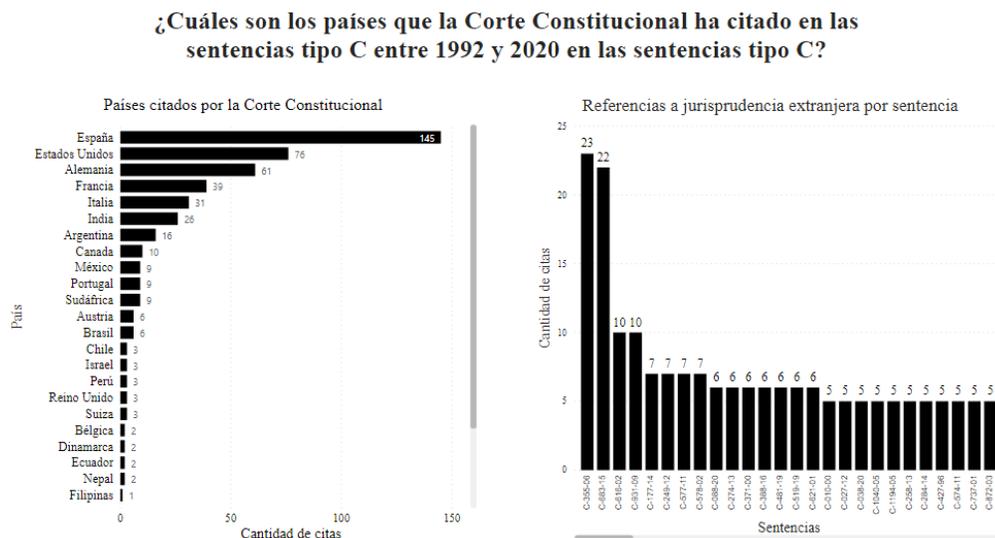


²² Todos los gráficos de este artículo son interactivos. Para acceder a las visualizaciones entre a:

<https://app.powerbi.com/view?r=eyJrIjojZTI0NjU5NDctZmEwZC00YmUxLWJhODMtNDU5NzYzYzAxOTk4IiwidCI6ImFjYUxNjMxLTAwZmUtNDkwZC05MWFilTE2M2VmODcyNjBIZSIsImMiOiR9&pageName=ReportSection>

A nivel de sentencias, se identificó que la que contiene más referencias a jurisprudencia extranjera es la C-355 de 2006 que despenaliza el aborto en Colombia en ciertas circunstancias; los países referenciados en esta sentencia fueron Italia, Portugal, Estados Unidos, Alemania, España, Francia, Brasil, Hungría, Irlanda y Polonia. La siguiente sentencia con mayor cantidad de referencias es la C-683 de 2015, en la cual la Corte aborda el problema de la adopción homoparental, siendo los principales referentes de esta decisión México, Canadá, Francia y Brasil. La figura 6²³ muestra dos gráficos donde se observan algunas de las sentencias con mayor cantidad de referencias y los países de donde la Corte tomó jurisprudencia.

Figura 6: Captura de pantalla de Power BI mostrando en orden descendente las referencias a jurisprudencia extranjera por sentencia



Una variable no contemplada en el alcance de este trabajo, pero que fue tratada dados los resultados obtenidos, es la de los magistrados o jueces que más uso hacen de la jurisprudencia extranjera. Para ello se usó como fuente el dato de magistrado ponente de las sentencias identificadas que fue ingresado de forma manual. Es importante que estos datos no se tomen como una cifra definitiva, pues las referencias identificadas no fueron clasificadas en función de la sección en la que se encuentran en la sentencia. La figura 7 muestra la relación entre magistrados y referencias a jurisprudencia extranjera.

²³ La figura 6 es una captura de pantalla de la herramienta Power BI, accediendo al link en la nota 21 se podrá encontrarla. Una característica relevante es que, dado que los gráficos son interactivos, al seleccionar una sentencia se podrán observar los países citados en esta o, por el contrario, al seleccionar un país se podrá observar las sentencias en las que este aparece.

Figura 5: Relación entre magistrados y citación a jurisprudencia extranjera²⁴

Magistrado Ponente	Número de referencias
MANUEL JOSE CEPEDA ESPINOSA	48
JORGE IVAN PALACIO PALACIO	40
ALEJANDRO MARTINEZ CABALLERO	32
MARIA VICTORIA CALLE CORREA	32
JAIME ARAUJO RENTERIA	29
JUAN CARLOS HENAO PEREZ	29
LUIS ERNESTO VARGAS SILVA	28
ALEJANDRO LINARES CANTILLO	25
CLARA INES VARGAS HERNANDEZ	19
GLORIA STELLA ORTIZ DELGADO	17
GABRIEL EDUARDO MENDOZA MARTELO	14
NILSON PINILLA PINILLA	14
ALBERTO ROJAS RIOS	13
JORGE IGNACIO PRETEL CHALJUB	13
CARLOS GAVIRIA DIAZ	11
MARCO GERARDO MONROY CABRA	11
EDUARDO MONTEALEGRE LYNETT	10
ALVARO TAFUR GALVIS	9
JAIME CORDOBA TRIVIÑO	8
ANTONIO JOSE LIZARAZO OCAMPO	7
MAURICIO GONZALEZ CUERVO	7
RODRIGO ESCOBAR GIL	7
ANTONIO BARRERA CARBONELL	6
CRISTINA PARDO SCHLESINGER	6
FABIO MORON DIAZ	5
HUMBERTO ANTONIO SIERRA PORTO	5
JOSE FERNANDO REYES CUARTAS	5
RODRIGO UPRIMNY YEPES	5
JORGE ARANGO MEJIA	4
DIANA FAJARDO RIVERA	3
LUIS GUILLERMO GUERRERO PEREZ	3
RICHARD S. RAMIREZ GRISALES	3
CATALINA BOTERO MARINO	1
EDUARDO CIFUENTES MUÑOZ	1
JOSE ANTONIO CEPEDA AMARIS	1
MARTHA VICTORIA SACHICA MENDEZ	1
VLADIMIRO NARANJO MESA	1

Colombia y los patrones globales de referencia a jurisprudencia extranjera

Para analizar los resultados sobre las referencias a jurisprudencia extranjera encontrados en las sentencias C de la Corte Constitucional de Colombia en relación con las tendencias globales, es relevante tener en cuenta que la evidencia sobre la práctica a nivel mundial es escasa, que hay una alta variación en la frecuencia y que los factores que explican el por qué de la práctica son variados y difícilmente se puede concluir sobre este aspecto solo analizando las sentencias (Hirschl, 2016). De

²⁴ Los caracteres especiales como tildes fueron eliminados de los nombres como procedimiento de estandarización de los datos. Una visualización interactiva de la tabla relacionado magistrados, países y sentencias se encuentra disponible en la herramienta Power BI.

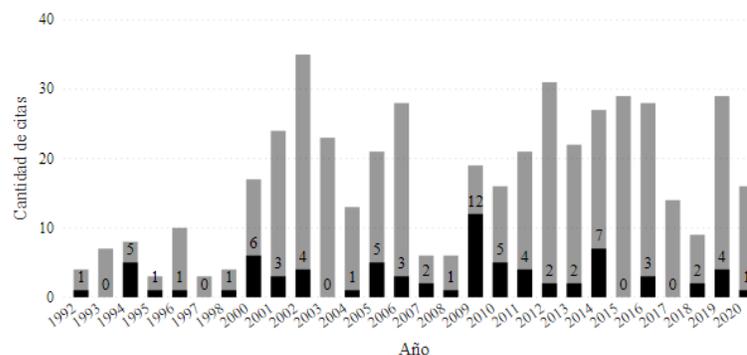
igual forma, se debe tener en cuenta que los datos presentados acá son preliminares y por lo tanto los análisis en esta sección también lo son.

Sobre la frecuencia de la práctica en Colombia se observa que la Corte Constitucional se encuentra en un punto intermedio entre las cortes renuentes a involucrarse en el uso de decisiones extranjeras como Estados Unidos y las que lo hacen más recurrentemente como Canadá y Sudáfrica. Por ejemplo, en un periodo comprendido entre 1995 y 2006 las cortes federales de Estados Unidos citaron un total de 145 decisiones extranjeras (Zaring, 2006), mientras que en el mismo periodo en Colombia hubo 181 referencias en la Corte Constitucional y solo en un tipo de decisión judicial. Por su parte, Calabresi y Zimdahl (2005) reportan que entre 1986 y 2005 la tasa de citación fue del 0,3% en relación con la totalidad de las sentencias emitidas en ese periodo.

A diferencia del caso estadounidense, se sabe que el tribunal constitucional de Sudáfrica acude abiertamente a decisiones judiciales extranjeras y que lo ha hecho en alrededor del 50% de sus decisiones, en contraste con el 2,79% registrado en Colombia; de igual forma, es conocido que en Canadá el tribunal supremo ha recurrido un total de 1500 veces a este tipo de sentencias entre 2000 y 2009, periodo en el cual el tribunal constitucional colombiano hizo uso de esta práctica 196 veces (Hirschl, 2016).

En relación con los tribunales más citados a nivel mundial se encontró que destacan la Corte Suprema de Estados Unidos, Reino Unido y Canadá, así como el Tribunal Constitucional Federal Alemán y el Tribunal Europeo de Derechos Humanos. Sobre los casos de Estados Unidos, investigaciones sobre la relevancia de las decisiones judiciales de este país en el mundo, han revelado que su influencia ha ido decayendo (Hirschl, 2016), aunque en Colombia la Corte Constitucional ha recurrido de manera constante a su jurisprudencia. Sobre este país se encontraron 76 referencias, también que solo en cinco de los años estudiados no se registró ninguna referencia a este país. Adicionalmente, los años en los que más se han hecho referencias a decisiones estadounidenses son 2010 y 2014.

Figura 7: citación de decisiones judiciales estadounidenses entre 1992 y 2020.



Por su parte, los tribunales de Canadá y Alemania, en conjunto, no han tenido una influencia mayor que la de Estados Unidos en las sentencias de la corte colombiana, al menos en cuanto a citas, a pesar de ser dos de los países que más ‘exportan’ jurisprudencia a nivel mundial. En total se registraron un total de 71 referencias a esos países, 61 a Alemania y 10 a Canadá. Sin embargo, un aspecto que coincide con las tendencias globales es que la influencia de estos dos países es creciente, puesto que

el 63% de las referencias se han concentrado entre los años 2010-2020, además son dos de los diez países de los que se analiza mayor jurisprudencia en la Corte Constitucional.

Sobre el caso de España, que merece consideración especial por ser el país al que más se remite la Corte en la búsqueda de decisiones judiciales -con un total de 145 referencias- se puede afirmar que si bien este país no es un referente a nivel mundial en materia de ‘exportación’ de jurisprudencia, en la literatura se encontró una hipótesis que puede explicar este aspecto: los investigadores Gelter y Siems (2013) estudiando patrones de referencia en Europa, encontraron que la similitud de idioma y cultura puede ser un factor que explique por qué un país se remite a otro en búsqueda de soporte para sus decisiones judiciales. Esos dos aspectos son compartidos por Colombia y España dada su historia política, social y económica; sin embargo, esta consideración es especulativa puesto que también sería válido cuestionarse por qué solo España y no otros países de habla hispana. Esas dos interrogantes por el alcance de este trabajo no son abordadas y podrían ser estudiadas por otros investigadores en trabajos futuros.

De forma adicional a la hipótesis de la cultura y el idioma, se encontraron dos hipótesis más que pueden resultar relevantes para entender la práctica en Colombia. Por una parte, se cree que la cantidad de referencias a jurisprudencia extranjera está relacionada con la madurez jurisprudencial de un tribunal, apuntando a que entre más desarrollada esté la jurisprudencia menos probable es que se recurra a otros países en búsqueda de soporte para tomar una decisión; por otra parte, en la literatura se ha encontrado que cuando un tribunal se enfrenta a un caso/tema nuevo o innovador, es más probable que se recurra a otros países en búsqueda de orientación para lograr decidir (Hirschl, 2016).

Con respecto a la madurez, inicialmente los datos parecen controvertir esa hipótesis en la medida en que las referencias a jurisprudencia extranjera son relativamente constantes año a año y con un aumento considerable desde el 2000, un aspecto que se podría abordar con mayor profundidad adicionando a las frecuencias modelado de tópicos/temas de las sentencias en general y de las referencias en el contexto de los argumentos en que son usadas. Finalmente, en relación con los casos nuevos o novedosos, las sentencias C-355-06 y C-683-15 apuntan a confirmar esta presunción, puesto que son las decisiones donde hubo más referencias a jurisprudencia extranjera y abordan dos temas altamente debatidos en la sociedad: el aborto y la adopción homoparental. Estos temas ponen en conflicto o tensión a la cultura dominante en Colombia y la protección de los derechos de las mujeres, de los menores y los no nacidos, así como de minorías históricamente discriminadas como la comunidad LGBT.

Conclusiones

Este trabajo demostró la capacidad que tienen soluciones basadas en el análisis y extracción de datos para contribuir a brindar soluciones a problemas en el área del derecho y de las ciencias políticas, específicamente dando respuesta por primera vez a una cuestión abierta en Colombia desde 2008, año en el que la investigadora Lina Marcela Escobar se propuso responder a la pregunta de ¿Cuántas veces se remite la Corte Constitucional a jurisprudencia extranjera?

Sin embargo, este tipo de soluciones, a pesar de demostrar su utilidad para resolver la tarea propuesta y disminuir la cantidad de tiempo humano requerido para lograrlo, están lejos de ser completamente automáticas, lo que era el objetivo inicial de este trabajo. En ese sentido, la solución propuesta constituye una herramienta de asistencia para el investigador que semi-automatiza las etapas

de análisis de grandes volúmenes de datos no estructurados y reduce así la cantidad de información que debe ser procesada manualmente.

En relación con el punto anterior se debe tomar en cuenta que el tiempo necesario para concluir esta investigación fue de aproximadamente 6 meses en total, desde el momento en que se empezó a escribir el algoritmo para recopilar las sentencias, hasta el momento en que se escribieron estas líneas. Una cantidad de tiempo mucho menor de la que hubiese sido necesaria realizando la tarea de forma manual con uno o más investigadores. Sobre este aspecto se debe tomar en cuenta que finalizada la etapa de desarrollo de la solución, que tomó alrededor de cinco meses, en menos de un día se analizaron más de 6000 sentencias y más de tres millones de párrafos en búsqueda de un grupo de 473 fragmentos que respondían a la pregunta de la frecuencia de la práctica en Colombia, solo tomando como objeto de estudio las sentencias tipo C entre 1992 y 2020.

La solución permitió encontrar, de forma preliminar, que es España la principal fuente de decisiones judiciales a las que se remite la Corte Constitucional de Colombia y que le siguen Estados Unidos, Alemania, Francia e Italia. Así mismo, estos hallazgos permitieron comparar los patrones de citación en las sentencias C con respecto a las tendencias globales, corroborando algunos supuestos como que las cortes de Canadá y Alemania se han convertido en grandes exportadores de jurisprudencia a nivel global. Además, en relación con los datos globales, es posible afirmar que Colombia se encuentra en un punto intermedio entre los países que hacen uso de esta práctica abiertamente y aquellos países renuentes a implementarla.

Finalmente, en relación con las posibles líneas de investigación futuras, se puede decir lo siguiente:

- Los datos constituyen un panorama general de la práctica en Colombia y son un insumo para que tanto politólogos como abogados enfocados en el estudio del derecho constitucional, de la teoría de la democracia, de los Derechos Humanos, entre otros campos, puedan analizar caso a caso la influencia de la jurisprudencia extranjera en las decisiones de la Corte Constitucional y sus efectos políticos y sociales.
- Es necesario reevaluar los resultados presentados ampliando la base de sentencias para entrenar el modelo GloVe, así como crear un conjunto de datos validado y debidamente preparado para evaluar métricas adicionales al accuracy, tal como pueden ser el recall y F1 score.
- Emplear métodos basados en categorizar y segmentar las frases en función de su estructura gramática y semántica – *part of speech tagging* –, con el fin de emplear técnicas más avanzadas de extracción de información basadas en reglas de lenguaje o métodos de aprendizaje profundo, como redes neuronales artificiales recurrentes o transformadores bidireccionales.

Referencias

- Bernal, C., & Albert, R. (2016). *Cambio constitucional informal*. Bogotá: Universidad Externado de Colombia.
- Boella, G., Di Caro, L., & Leone, V. (2019). Semi-automatic knowledge population in a legal document. *Artificial Intelligence and Law*, 227-251.
- Chapagain, A. (2019). *Hands-On Web Scraping with Python: Perform advanced scraping operations using various Python libraries and tools such as Selenium, Regex, and others*. Birmingham: Packt Publishing.
- Corte Constitucional de Colombia. (01 de 01 de 2021). *Corte Constitucional de Colombia*. Obtenido de Sentencias Proferidas por la Corte Constitucional: <https://www.corteconstitucional.gov.co/lacorte/estadisticas.php>
- Dyevre, A. (2020). The promise and pitfall of automated text-scaling techniques for the analysis of jurisprudential change. *Artificial Intelligence and Law*.
- Escobar, L. (2008). El Uso del Precedente Extranjero por parte de la Corte Constitucional Colombiana. *Revista Colombiana de Derecho Internacional*, 391-409.
- Ferrari, G. (2019). *Judicial Cosmopolitanism: the use of foreign law in contemporary constitutional systems*. Leiden: Brill.
- Gardbaum, S. (2016). El principio de separación de poderes y la expansión del control de la constitucionalidad en democracias consolidadas (o ¿Por qué el modelo de supremacía legislativa ha sido prácticamente retirado del mercado?). En C. Bernal, & R. Albert, *Cambio constitucional informal* (págs. 259-320). Bogotá: Universidad Externado de Colombia.
- Halmani, G. (2012). The Use of Foreign Law in Constitutional Interpretation. En M. Rosenfeld, & A. Sajo, *The Oxford Handbook of Comparative Constitutional Law* (págs. 1328-1348). Croydon: Oxford University Press.
- Hirschl, R. (2016). *Asuntos Comparativos: el renacimiento del derecho constitucional comparado*. Bogotá: Universidad Externado de Colombia.
- Hirschl, R. (2016b). En busca de la identidad: uso voluntario de fuentes extranjeras en contextos constitucionales disímiles. En C. Bernal, & R. Albert, *Cambio Constitucional Informal* (págs. 321-420). Bogotá: Universidad Externado de Colombia.
- Jacobsohn, G. (2004). The Permeability of Constitutional Borders. *Texas Law Review*, 1763-1819.
- Kearney, E. (2019). The Utilization of Foreign Law in Domestic Constitutional. *King's Inns Law Review*, 8, 127-168.
- Law, D., & Chang, W. (2011). The Limits of Global Judicial Dialogue. *Washington Law Review Association*, 523-577.

- Lippi, M. P. (2019). CLAUDETTE: an automated detector of potentially unfair clauses in online terms of service. *Artificial Intelligence and Law*, 117-139.
- Mandal, A., Ghosh, K., Ghosh, S., & Mandal, S. (2020). Unsupervised approaches for measuring textual similarity. *Artificial Intelligence and Law*.
- McCrudden, C. (2000). A Common Law of Human Rights? Transnational Judicial Conversations on Constitutional Rights. *Oxford Journal of Legal Studies*, 20(4), 499-532.
- Nanda, R. S. (2019). Unsupervised and supervised text similarity systems for automated identification of national implementing measures of European directives. *Artificial Intelligence and Law*, 199-225.
- Pennington, J., Socher, R., & Mnning, C. (2014). GloVe: Global Vectors for Word Representations. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1532-1543.
- Perju, V. (2012). Constitutional Transplants, Borrowings, and Migrations. En M. Rosenfeld, & A. Sajo, *The Oxford Handbook of Comparative Constitutional Law* (págs. 1304-1328). Croydon: Oxford University Press.
- Polacchini, F. (2019). Foreign Law and Foreign Case Law before Latin American Constitutional Courts and Supreme Courts. En G. Ferrari, *Judicial Cosmopolitanism: The Use of Foreign Law in Contemporary Constitutional Systems* (págs. 827-851). Leiden: Brill.
- Rawls, J. (1979). *Teoría de la Justicia*. México: Fondo de Culura Económica.
- Rodman, E. (2019). A Timely Intervention: Tracking the Changing Meanings of Political Concepts with Word Vectors. *Political Analysis*, 28(1), 87-111.
- Rosenkrantz, C. (2003). Against borrowings and other nonauthoritative uses of foreign law. *International journal of constitutional law*, 269-295.
- Sadeghian, A. S. (2018). Automatic semantic edge labeling over legal citation graphs. *Artificial Intelligence and Law*, 127-144.
- Sannier, N., Adedjouma, M., Sabetzadeh, M., & Briand, L. (2016). Automated Classification of Legal Cross References Based on Semantic Intent. *Requirements Engineering*.
- Sannier, N., Adedjouma, M., Sabetzadeh, M., & Briand, L. (2017). An automated framework for detection and resolution of cross references in legal texts. *Requirements Engineering*, 215-237.
- Sentencia C-599-92, C-599-92 (Corte Constitucional de Colombia 1992).
- Waldron, J. (2018). *Contra el gobierno de los jueces*. Buenos Aires: Grupo Editorial Siglo Veintiuno.
- Yamada, H., Teufel, S., & Tokunaga, T. (2019). Building a corpus of legal argumentation in Japanese judgement documents: towards structure-based summarisation. *Artificial Intelligence and Law*, 141-170.

Zelada, L. (2012). El uso del derecho y la jurisprudencia en los fallos del tribunal constitucional de Chile: 2006-2010. *Revista Chilena de Derecho*, 39(1), 189-224.

Zhang, J., & El-Gohary, N. (2016). Semantic NLP-Based Information Extraction from Construction Regulatory Documents for Automated Compliance Checking. *Journal of Computing in Civil Engineering*, 30(2).