

Bogotá, 15 de mayo de 2023

Señores

Defensoría Delegada para Asuntos Constitucionales y Legales

Defensoría del Pueblo

E. S. D

Referencia: Solicitud ciudadana de insistencia en la
selección de la tutela con rad. T9301656

De manera atenta se dirigen ante su despacho Ana Bejarano Ricaurte, Susana Echavarría Medina, Catalina Moreno Arocha, Juanita Castro Hernández y Jonathan Bock, en atención de que el expediente de tutela de la referencia no fue seleccionado por la Sala de Selección Número T9301656 de la Honorable Corte Constitucional el día 29 de abril de 2023, para solicitar respetuosamente la insistencia en la selección de dicho expediente con base en los argumentos que presentamos a continuación.

Los accionantes de esta acción de tutela son la familia de un menor de edad que nació con trastorno del espectro autista que actuaron con el fin de que la accionada, la EPS Salud Total, exonerara a dicho menor del cobro de la cuota moderadora de citas médicas y terapias, al igual que el de autorizar el pago del transporte desde la residencia hasta el lugar de las terapias. Esto se fundamentaba en la falta de recursos para sortear dichos gastos.

Con el fin de construir la argumentación de su sentencia, el juez de segunda instancia acudió al uso de la herramienta de inteligencia artificial ChatGPT. Esto lo hizo argumentando que “la Ley 2213 de 2022 tiene por objeto la incorporación de las TIC en los procesos judiciales.” Asimismo, el juez plantea que, si bien no se busca “reemplazar la decisión del Juez”, el uso de esta herramienta se da para “optimizar los tiempos empleados en redacción de sentencias”. La motivación que se expone en la decisión, aparte de esas precisiones, se basa plenamente en la exposición de preguntas planteadas a dicha herramienta y de las respuestas recibidas.

Este caso llama la atención porque plantea inquietudes sobre los límites constitucionales para el uso de herramientas tecnológicas, específicamente las de inteligencia artificial, en el marco de la función judicial. Como se alcanza a esbozar en lo expuesto por el juez de segunda instancia, existe una tensión entre las garantías procesales y el uso de este tipo de tecnología con el fin de optimizar la administración de justicia. Las características propias de la inteligencia artificial, tal y como lo han evidenciado expertos y organismos internacionales, acarrearán riesgos y preocupaciones

de derechos humanos que deben orientar el aprovechamiento de las ventajas que estas mismas pueden traer para distintos procesos y dinámicas de la sociedad y del Estado.

ChatGPT es un chatbot de inteligencia artificial desarrollado por la compañía de tecnología OpenAI en 2022. Este sistema es soportado por un tipo de inteligencia artificial llamada “*Large Language Model*” -gran modelo de lenguaje en español- o LLM por sus siglas en inglés, que usa técnicas de “*Machine Learning*” -aprendizaje automático o automatizado en español- y “*Deep Learning*” -aprendizaje profundo en español- entrenado para comprender y generar lenguaje humano, incluida la clasificación de textos y el análisis de sentimientos. Estos resultados son posibles pues el modelo ha sido suministrado con suficiente texto -artículos, discursos, libros, investigaciones- para aprender a predecir cuáles son las palabras con mayor posibilidad de seguir a otro conjunto de palabras¹. Por ejemplo, si se le presenta la expresión “en un abrir y cerrar de”, con base en algoritmos de aprendizaje automático, modelos probabilísticos y reglas de lenguaje, lo más posible es que el modelo prediga que la palabra que le sigue a la expresión es “ojos”.

De esta manera, los LLMs replican estos modelos de lenguaje a gran escala y contienen tanta información que logran predecir la continuación de palabras en diversos y más complejos contextos. En todo caso, aunque los LLMs almacenen un gran volumen de información, es imposible que realmente “memoricen” la totalidad de los textos, pues matemáticamente no existe tal capacidad. La imposibilidad de predecir el espectro completo de continuaciones y contextos resulta en la selección de continuaciones para contextos similares y, en consecuencia, en generalizaciones que pueden ser equivocadas².

Las limitaciones arquitectónicas y de entrenamiento de los modelos de lenguaje se denominan “*Hallucinations*” -alucinaciones en español- y se han definido como un fenómeno común en el cual el sistema proporciona respuestas incorrectas, irrelevantes o sin sentido³. Si bien ChatGPT ha sido desarrollado a tal punto de acercarse cada vez más a estándares de convencimiento humano es recurrente -por su estructura misma- que cometa errores. Con base en la anterior explicación se ponen de presente algunas consideraciones importantes para la comprensión de este sistema:

1. ChatGPT, a diferencia de los buscadores, no presenta una multiplicidad de fuentes primarias y secundarias para la libre consulta del usuario. Por el contrario, provee una única respuesta que ha predecido basándose en un

¹ New York Times, “How Does ChatGPT Really Work?”,

<https://www.nytimes.com/2023/03/28/technology/ai-chatbots-chatgpt-bing-bard-llm.html>

² Lak Lakshmanan, “Why large language models (like ChatGPT) are bullshit artists”,

<https://becominghuman.ai/why-large-language-models-like-chatgpt-are-bullshit-artists-c4d5bb850852>

³ New York Times, “Artificial Intelligence Glossary: Neural Networks and Other Terms Explained”,

<https://www.nytimes.com/article/ai-artificial-intelligence-glossary.html>

modelo de lenguaje a gran escala. ChatGPT no cita a una o varias fuentes en sus respuestas.

2. ChatGPT puede reproducir, con un alto grado de confianza, información falsa o engañosa. No es un sistema que logre equilibrar, considerar y ponderar una multiplicidad de perspectivas. ChatGPT no distingue la realidad de la ficción, no razona o piensa, no sabe matemáticas ni sabe derecho.
3. ChatGPT no es un experto. Si bien tiene un alto nivel de confianza en sus respuestas para hacerlas convincentes, no logra identificar todas las manifestaciones de “alucinaciones” con precisión. Adicionalmente, la versión gratuita de ChatGPT (a la fecha) solo cuenta con información ingresada hasta el 2021, cualquier información posterior no es considerada.
4. ChatGPT reproduce los sesgos contenidos en la información. El sistema reproduce el tono de los datos sobre los cuáles basa sus respuestas, en este caso, ChatGPT puede llegar a replicar afirmaciones con un tono científico o artístico, pero también con tintes discriminatorios o imprecisos dependiendo de la formulación de las consultas introducidas y de los datos usados para responder.

La limitaciones son incluso presentadas por el mismo sistema que, una vez se ingresa para cualquier consulta, muestra el siguiente mensaje:

“Limitaciones:

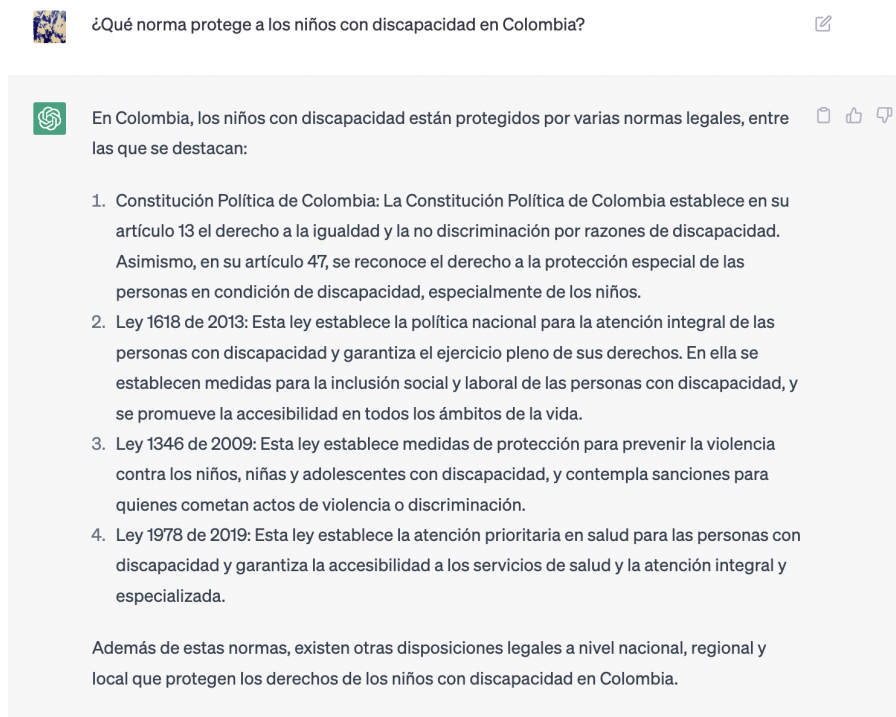
- *Puede generar ocasionalmente información incorrecta*
- *Puede producir ocasionalmente instrucciones dañinas o contenido sesgado*
- *Conocimiento limitado del mundo y de eventos posteriores al 2021”⁴.*

ChatGPT		
☀ Examples	⚡ Capabilities	⚠ Limitations
"Explain quantum computing in simple terms" →	Remembers what user said earlier in the conversation	May occasionally generate incorrect information
"Got any creative ideas for a 10 year old's birthday?" →	Allows user to provide follow-up corrections	May occasionally produce harmful instructions or biased content
"How do I make an HTTP request in Javascript?" →	Trained to decline inappropriate requests	Limited knowledge of world and events after 2021

⁴ OpenAI, ChatGPT, <https://chat.openai.com/>

Las alucinaciones e inconsistencias de ChatGPT son tan recurrentes que, para fines de esta solicitud, se le formularon algunas preguntas relacionadas con el caso que reposa en el expediente y que evidencian las consideraciones previamente presentadas:

Ejemplo 1:



Se le pregunta a ChatGPT “¿Qué norma protege a los niños con discapacidad en Colombia?” A continuación, el chatbot enuncia una lista de “normas legales” que protegen a los niños con discapacidad en Colombia. Sin embargo, y aunque ChatGPT no advierte sobre las imprecisiones de la respuesta, procede a enlistar confiadamente normas que, con una lectura detallada, son equivocadas. En primera instancia, cita el artículo 47 de la Constitución Política, que afirma que:

“ARTICULO 47. El Estado adelantará una política de previsión, rehabilitación e integración social para los disminuidos físicos, sensoriales y psíquicos, a quienes se prestará la atención especializada que requieran”.

Esta respuesta, aunque está relacionada con los derechos de las personas con discapacidad no especifica que protege “especialmente a los niños”. Tampoco precisa si esa afirmación es atribuible a una interpretación jurisprudencial.

En segunda instancia, y con un error más evidente, al final de la lista cita la “Ley 1978 de 2019”. Esta Ley nada tiene que ver con los derechos de los niños, niñas y adolescentes con discapacidad pues es la Ley “Por la cual se moderniza el Sector de las Tecnologías de la Información y las Comunicaciones -TIC, se distribuyen

competencias, se crea un Regulador Único y se dictan otras disposiciones”. La única mención que hace a la población con discapacidad es en el artículo 35 que establece las Funciones del Fondo Único de Tecnologías de la Información y las Comunicaciones y enuncia que esta entidad deberá “*Financiar planes, programas y proyectos para promover el acceso con enfoque diferencial de los ciudadanos en situación de discapacidad a las Tecnologías de la Información y las Comunicaciones*”.

Ejemplo 2:



¿Cuál es la jurisprudencia de la Corte Constitucional sobre el acceso al derecho a la salud para niños con discapacidad en Colombia?



La Corte Constitucional de Colombia ha emitido varias sentencias que establecen la obligación del Estado de garantizar el acceso al derecho a la salud de los niños con discapacidad en Colombia. A continuación, se presentan algunas de las más relevantes:

1. Sentencia T-760 de 2008: En esta sentencia, la Corte Constitucional estableció que los niños con discapacidad tienen derecho a recibir atención integral y especializada en salud que les permita alcanzar su máximo desarrollo. La Corte ordenó al sistema de salud garantizar la atención oportuna y adecuada de los niños con discapacidad y estableció que los servicios de salud deben ser proporcionados con enfoque diferencial para garantizar la inclusión social de los niños con discapacidad.
2. Sentencia T-381 de 2012: En esta sentencia, la Corte Constitucional declaró que el derecho a la salud de los niños con discapacidad es fundamental y debe ser protegido de manera especial. La Corte ordenó a las entidades encargadas de garantizar el derecho a la salud de los niños con discapacidad a establecer protocolos de atención y seguimiento integral que permitan mejorar su calidad de vida.
3. Sentencia T-405 de 2016: En esta sentencia, la Corte Constitucional estableció que el acceso al derecho a la salud de los niños con discapacidad es un derecho fundamental que debe ser garantizado por el Estado. La Corte ordenó a las entidades responsables de garantizar la atención en salud de los niños con discapacidad a garantizar la atención integral, oportuna y efectiva de los niños con discapacidad.

En general, la jurisprudencia de la Corte Constitucional de Colombia reconoce el derecho a la salud como un derecho fundamental de los niños con discapacidad y establece la obligación del Estado de garantizar el acceso a servicios de salud integrales y especializados para garantizar su desarrollo pleno e inclusión social.

Se le pregunta a ChatGPT “¿Cuál es la jurisprudencia de la Corte Constitucional sobre el acceso al derecho a la salud para niños con discapacidad en Colombia?”. A continuación, el chatbot enlista algunas sentencias de la Corte Constitucional que supuestamente son las más relevantes sobre el acceso al derecho a la salud de los niños con discapacidad en Colombia. En todo caso, y aunque ChatGPT no advierte sobre las imprecisiones de la respuesta, procede a enunciar confiadamente sentencias que, tras ser verificadas, puede comprobarse que son erradas. La Sentencia T-760 de 2008 y la Sentencia T-381 de 2012 no mencionan siquiera el término “discapacidad”, mientras que la Sentencia T-405 de 2016, si bien se refiere al derecho a la salud de personas con discapacidad, no menciona en sus consideraciones el término “niño”.

A continuación exponemos cómo este proceso de tutela brinda una oportunidad única para que la Corte Constitucional aborde este aspecto tan crítico de la justicia en tiempos modernos.

Se considera entonces que las características del fallo de tutela del expediente de referencia se enmarcan dentro de los criterios del artículo 52 del Acuerdo 02 de 2015, y que pueden resumirse así:

El expediente bajo referencia presenta un escenario jurídico que satisface dos **criterios objetivos**:

1. Es *novedoso* porque plantea cuestiones de las implicaciones del uso de inteligencia artificial en el marco de la función judicial. No es la primera vez que una autoridad judicial acude a este tipo de mecanismos: diversas instancias en distintas partes del mundo, incluyendo la Corte Constitucional, han acudido a este tipo de herramientas para apoyar su función. No obstante, pocos tribunales nacionales, y ninguno a nivel internacional, se ha referido a las implicaciones legales o de derechos. A la par, la tecnología evoluciona a nivel exponencial y los estados enfrentan la dificultad de crear estándares y lineamientos que permitan un adecuado uso de estas herramientas sin que existan riesgos para los derechos humanos.

Un ejemplo de esta situación se encuentra en el uso de la Correctional Offender Management Profiling for Alternative Sanctions ("COMPAS") en diversas partes de los Estados Unidos. Esta herramienta de análisis de riesgo evalúa las posibilidades de reincidencia de personas convictas para determinar aspectos como la duración de las penas, la posibilidad de acceder a libertad condicional y el costo de fianzas para evitar detención preventiva. Este sistema recibió un alto nivel de críticas en la opinión pública a raíz de una investigación de Propublica que determinó que había una tendencia a que las personas afroamericanas fueran consideradas de mayor riesgo de reincidencia que las personas blancas.⁵ En 2016, la Corte Suprema de Wisconsin se refirió a una acción judicial de un ciudadano que alegaba la violación de su debido proceso por el uso de esta herramienta en el cálculo que llevó a una sentencia de seis años de prisión en su contra. El accionante planteaba argumentos con respecto a su derecho a ser sentenciado con base en información precisa y, aunque la Corte negó su petición, lo hizo estableciendo límites con respecto al uso de una herramienta como COMPAS. Para la Corte, las instancias judiciales que acudieran a esta herramienta tenían que tener precauciones frente a aspectos como: las dificultades para conocer la forma en que se calculan los factores de riesgo, toda vez que COMPAS es de propiedad privada; el hecho de que la muestra de COMPAS fuera nacional y no local; el hecho de

⁵ Propublica. How We Analyzed the COMPAS Recidivism Algorithm. 23 de mayo de 2016. Disponible en: <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>.

que había estudios mostrando los posibles sesgos en contra de minorías; la necesidad de que este tipo de herramientas sean constantemente revisadas y redefinidas con el fin de garantizar su precisión.

Distintos académicos se han referido a esta decisión como un llamado a la prudencia con relación al entusiasmo tecnológico.⁶ También se puede ver que, en el Reino Unido, la policía de Durham ha aplicado una tecnología similar, la Harm Assessment Risk Tool (HART) con el fin de determinar el nivel de riesgo de reincidencia de criminales y así determinar su viabilidad para formar parte del sistema de penas alternativas Checkpoint. Esta tecnología viene acompañada de un marco guía para contrarrestar problemas de opacidad y algunos de los problemas legales prácticos y legales que se presentan con su uso.⁷ Asimismo, se observa que un tribunal de la Haya declaró que un sistema de predicción de riesgo de fraude en el sistema de seguridad social de los Países Bajos violaba el derecho a la privacidad en razón de sus problemas de transparencia y de su ausencia de salvaguardas⁸. En un nivel más macro, se observa que la propuesta de Reglamento en la Unión Europea “por el que se establecen normas armonizadas en materia de inteligencia artificial” se establece que “[d]eben considerarse de alto riesgo ciertos sistemas de IA destinados a la administración de justicia y los procesos democráticos, dado que pueden tener efectos potencialmente importantes para la democracia, el Estado de Derecho, las libertades individuales y el derecho a la tutela judicial efectiva y a un juez imparcial.”⁹

Lo anterior demuestra que el uso de la inteligencia artificial en el marco de la administración de justicia no es un aspecto que se limite a su aplicación sin ningún tipo de reflexión o salvaguarda. Se observa que incluso en el caso de herramientas diseñadas con fines muy específicos como COMPAS y sobre las que existe cierto nivel de control por el sistema judicial hay preocupaciones sobre aspectos de derechos humanos. Las cuestiones pueden ser mucho más amplias con respecto a una herramienta como ChatGPT, que no fue diseñada para la asistencia en procesos judiciales. En línea con los debates que han ocurrido sobre estos temas a nivel comparado, se considera fundamental que la Corte Constitucional aborde el estudio

⁶ Pedro Rubim Borges Fortes, "Paths to Digital Justice: Judicial Robots, Algorithmic Decision-Making, and Due Process" (2020) 7:3 AsianJLS 453.; "Criminal Law - Sentencing Guidelines - Wisconsin Supreme Court Requires Warning before Use of Algorithmic Risk Assessments in Sentencing - State v. Loomis 881 N.W.2d 749 (Wis. 2016)" (2017) 130:5 Harv L Rev 1530.

⁷ Marion Oswald, Jamie Grace, Sheena Urwin & Geoffrey C. Barnes (2018) Algorithmic risk assessment policing models: lessons from the Durham HART model and 'Experimental' proportionality, Information & Communications Technology Law, 27:2, 223-250, DOI: [10.1080/13600834.2018.1458455](https://doi.org/10.1080/13600834.2018.1458455)

⁸ Appelman, N., Ó Fathaigh, R. & van Hoboken, J. Social Welfare, Risk Profiling and Fundamental Rights: The Case of SyRI in the Netherlands. *JIPITEC*, 12(4), 257-271.

⁹ Comisión Europea. Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de inteligencia artificial) y se modifican determinados actos legislativos de la Unión. COM(2021) 206 final 2021/0106(COD)

del expediente de la referencia con el fin de establecer lineamientos y parámetros que deban ser aplicados por los jueces a la hora de utilizar este tipo de herramientas.

2. Se presenta la necesidad de *aclarar el contenido y alcance de un derecho fundamental*

La Constitución establece el derecho a ser juzgado “ante juez o tribunal competente y con observancia de la plenitud de las formas propias de cada juicio”. El expediente de referencia plantea preguntas sobre hasta qué punto puede un juez utilizar una herramienta de inteligencia artificial, más específicamente una como ChatGPT, con el fin de alimentar la motivación de sus providencias. Esto tiene una incidencia directa sobre derechos como el debido proceso consagrado en el artículo 29 de la Constitución y, además, de las fuentes de derecho admisibles en aplicación del artículo 230 superior como elemento esencial del Estado social de Derecho. ¿Puede considerarse que un juez que fundamenta su decisión en una consulta a este tipo de herramienta ha actuado con “observancia de la plenitud de las formas propias de cada juicio”? ¿Hasta qué punto puede decirse que recurrir a este tipo de herramientas preserva la sana crítica del juez como las “reglas del correcto entendimiento humano”, en las que interfieren “las reglas de la lógica, con las reglas de la experiencia del juez”¹⁰?

En línea con preocupaciones similares, la Comisión para la Eficiencia de la Justicia del Consejo Europeo estableció una Carta sobre aspectos éticos en el uso de la Inteligencia Artificial en los sistemas judiciales y en su entorno. Dentro de este documento se establecen una serie de principios, dentro de los que destacamos el de asegurar el respeto por los derechos humanos. Tal principio es desmenuzado por la Comisión indicado que es esencial que el uso de herramientas de inteligencia artificial para la resolución de disputas o para asistir en la toma de decisiones judiciales no atente contra el derecho de acceder a un juez y el derecho a un juicio justo, lo cual implica también el debido apego al Estado de derecho y a la independencia judicial, al igual que la aplicación de la “ética por diseño” y “derechos humanos por diseño”.¹¹

¿Cómo podría garantizarse este tipo de principios en el contexto colombiano? La jurisprudencia Constitucional ha establecido una serie de salvaguardas con respecto al debido proceso y el acceso a la administración de justicia, pero estos no fueron diseñados de cara a la aplicación de herramientas de inteligencia artificial. Lo anterior hace que un pronunciamiento de la Corte Constitucional con el fin de determinar el alcance de estos derechos en este contexto sea crítico.

¹⁰ Corte Constitucional. Sentencia T 041 de 2018.

¹¹ European Commission for the Efficiency of Justice (CEPEJ). European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment. 2018. Disponible en: <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>.

Asimismo, el expediente bajo referencia presenta un escenario jurídico que satisface un **criterio subjetivo**:

1. Se presenta la *urgencia de proteger un derecho fundamental*

En línea con lo anterior, el riesgo al derecho fundamental al debido proceso del artículo 29 de la Constitución Política y al derecho a las garantías judiciales que incluye el derecho a ser juzgado por un juez o tribunal competente, independiente e imparcial de los artículos 8 Y 25 de la Convención Americana sobre Derechos Humanos es inminente y debe ser atendido por la Corte Constitucional.

El uso incremental y desatendido de herramientas de inteligencia artificial para la administración de justicia supone una desprotección mayor para el derecho al debido proceso. Es urgente que la Corte Constitucional se pronuncie al respecto y establezca parámetros sobre el uso de herramientas de inteligencia artificial, incluyendo entre otras a ChatGPT, para garantizar que los usuarios del sistema judicial verdaderamente se les garantice un proceso equitativo con un juicio independiente e imparcial. Al respecto, diversos estudiosos de la inteligencia artificial han planteado propuestas para el uso adecuado y ético de estas herramientas en los procesos judiciales¹². Lo más urgente, para alcanzarlo, es un entendimiento por parte de los jueces sobre el alcance y limitaciones de estas herramientas, así como un monitoreo constante en la efectividad y alineamiento con el debido proceso que supone el uso de estas herramientas¹³. Los jueces se han encontrado con una novedosa, popular y expedita herramienta de consulta que deberá ser usada con suma responsabilidad para no llegar a vulnerar el debido proceso. Una revisión detallada del uso de ChatGPT como herramienta de consulta para la administración de justicia debe tener en cuenta los riesgos a la privacidad y seguridad de la información, la calidad de los datos, los potenciales sesgos, los usos ilegales de información y otros factores de injusticia que acarreen estos sistemas¹⁴. Debe entonces la Corte Constitucional brindar los parámetros necesarios para la garantía del debido proceso.

Finalmente, se observa que el expediente bajo referencia presenta un escenario jurídico que abarca un **criterio complementario**:

1. Existe una necesidad de *preservación del interés general*

¹² A. D. Reiling, "Courts and Artificial Intelligence," International Journal for Court Administration 11, no. 2 (2020): 1-10

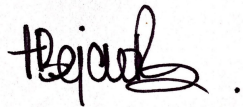
¹³ Ibid

¹⁴ Helberger, N. & Diakopoulos, N. (2023). ChatGPT and the AI Act. Internet Policy Review, 12(1). <https://doi.org/10.14763/2023.1.1682>

En relación con la urgencia de proteger el debido proceso, es también necesario considerar el amplio y creciente uso del chatbot ChatGPT, que de no ser atendido por la Corte Constitucional y en caso de establecerse como una herramienta de amplio uso entre los jueces, la administración de justicia podría enfrentar una amenaza al orden público. Apaciguar el riesgo de reemplazar, en cualquier medida, a un juez que falla en derecho, por un modelo de lenguaje que prediga la continuación lingüística y no las normas aplicables, debe ser una prioridad para la Corte Constitucional.

No siendo más, quedamos a disposición de la Defensoría Delegada para Asuntos Constitucionales y Legales, para ampliar, en caso de ser necesario, las razones de respaldo a la acción de tutela con el radicado T9301656.

Cordialmente,



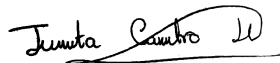
Ana Bejarano Ricaurte
T.P. 208.989 del C.S. de la J
C.C. No. 1.136.879.823 de Bogotá



Susana Echavarría Medina
C.C. 1152225226 de Medellín



Catalina Moreno Arocha
T.P. No. 262652 del CS de la J
C.C. 1.098.646.935 de Bucaramanga



Juanita Castro Hernández
T.P. 337.782 del C.S. de la J
C.C. 1.010.229.272 de Bogotá D.C.



Jonathan Bock
Director Ejecutivo
Fundación para la Libertad de Prensa