La implementación de ChatGPT en la política: ¿La solución a la imparcialidad?

Implementing ChatGPT in Politics: A Path Toward Impartiality?

André Medina[1]

Recibido: 02 febrero 2025
Aceptado: 25 marzo 2025

Resumen

El uso de inteligencias artificiales (IA) en la toma de decisiones políticas ha emergido como una herramienta potencialmente poderosa para mejorar la eficiencia y reducir los sesgos humanos. Sin embargo, el impacto de estas tecnologías plantea serias preocupaciones éticas y políticas. Por un lado, las IA prometen análisis rigurosos basados en datos masivos, lo que podría permitir decisiones más imparciales y alejadas de influencias indebidas. Por otro lado, los algoritmos pueden replicar y exacerbar sesgos preexistentes, introduciendo opacidad y falta de rendición de cuentas en la toma de decisiones. El avance de las IA debe ir acompañado de una reflexión ética profunda para evitar consecuencias no deseadas. Este artículo analiza tanto los beneficios como los riesgos de utilizar IA en la política, evaluando si esta tecnología puede realmente ofrecer imparcialidad y justicia, o si podría amenazar los principios fundamentales de la democracia.

Palabras clave: *Inteligencia artificial (IA), Imparcialidad, Sesgo, Algoritmos, Responsabilidad, Justicia*

Abstract

The use of artificial intelligence (AI) in political decision-making has emerged as a potentially powerful tool to enhance efficiency and reduce human biases. However, the impact of these technologies raises serious ethical and political concerns. On the one hand, AI promises rigorous analyses based on large datasets, potentially allowing for more impartial decisions free from undue influence. On the other hand, algorithms can replicate and exacerbate pre-existing biases, introducing opacity and a lack of accountability in decision-making processes. The advancement of AI must be accompanied by deep ethical reflection to avoid unintended consequences. This article examines both the benefits and risks of employing AI in politics, evaluating whether this technology can truly offer impartiality and justice, or if it might threaten the core principles of democracy.

Keywords: *Artificial Intelligence (AI), Fairness, Bias, Algorithms, Accountability, Justice*

Introducción

En la era digital, el desarrollo de la inteligencia artificial ha revolucionado múltiples esferas de la vida cotidiana, desde la economía hasta la medicina, y la política no es la excepción. Con su capacidad para procesar grandes volúmenes de datos y

1 *Estudiante de la Carrera de Filosofía en la Universidad Nacional Autónoma de Honduras (UNAH). Correo electrónico: eamedinap@unah.hn*

ofrecer soluciones eficientes, las IA han comenzado a abrirse camino en la toma de decisiones dentro del ámbito político. Esta automatización promete un análisis más riguroso de las políticas públicas, así como la posibilidad de reducir los errores humanos y las influencias indebidas que a menudo distorsionan el proceso político.

Sin embargo, el uso de estas tecnologías plantea interrogantes éticas fundamentales. A diferencia de las decisiones humanas, que suelen estar fundamentadas en valores sociales y principios morales, los algoritmos operan bajo un enfoque eminentemente lógico y basado en datos. Esto suscita una pregunta central: ¿Puede una inteligencia artificial ser verdaderamente imparcial en la política, o existe el riesgo de que reproduzca sesgos preexistentes y agrave la falta de transparencia en la toma de decisiones? En un mundo donde la justicia, la equidad y la representación justa de las perspectivas sociales son esenciales para la democracia, la implementación de IA en la política podría traer beneficios, pero también conlleva riesgos significativos.

El objetivo de este artículo es examinar las posibilidades y los peligros de la integración de la IA en la política, con el objetivo de evaluar si esta tecnología puede realmente ofrecer soluciones imparciales y justas en la toma de decisiones. A través de un análisis de modelos actuales de IA y su capacidad para enfrentar dilemas éticos, se reflexionará sobre si estas herramientas son aptas para la toma de decisiones en contextos políticos, o si, por el contrario, representan un peligro para la democracia misma.

Metodología

El trabajo presente cuenta con una metodología cualitativa experimental realizada meramente mediante el uso de ChatGPT para realizar las pruebas de campo, contando con un equipo de un solo computador y conexión a internet, las pruebas didácticas realizadas son preguntas con dilemas morales esenciales dentro de la ética y la moral para poder determinar aspectos como justicia, equidad y beneficio. El acercamiento de esta manera, junto con las capturas de pantalla proveídas dan pruebas empíricas sobre las respuestas otorgadas por la IA al intentar dar la mejor solución, siendo ChatGPT-4o la versión más reciente hasta la fecha.

Inteligencias artificiales y humanos, ¿es la diferencia tan grande?

La respuesta sobre el que distingue a las inteligencias artificiales de los humanos parece ser bastante evidente, desde un punto de vista biológico, es sencillo decir que uno es generado de manera orgánica, mientras que el otro es una construcción de estos organismos para tratar de imitar a esta inteligencia orgánica. Pero estos tienen entre si más parecido del que parece, un argumento principal es que, la capacidad de memoria, aprendizaje y funcionamiento no se limita a seres de carne y hueso o átomos de carbono. El autor Max Tegmark en su texto *Life 3.0 (Vida 3.0)* nos invita a recordar que, si bien los cerebros guardan información, también pueden hacerlo las tarjetas de memoria, aunque no estén hechas del mismo material, a medida avanza la tecnología, lo que nos hace humanos será más difícil de definir. (Tegmark, 2017)

Max Tegmark explora las posibles etapas

de la vida en términos de su capacidad para procesar información y diseñar tanto su hardware como su software, las divide en 3 categorías, la primera (Vida 1.0) hace referencia a los seres que no pueden rediseñar ni su software ni su hardware; la vida biológica más sencilla, los primeros microorganismos que existieron en la tierra, dependen completamente de la selección natural para hacer que su información genética sobresalga y se mantenga con el paso de las generaciones, como una bacteria o una planta.

La segunda categoría (Vida 2.0), hace referencia a los seres que pueden rediseñar su software, pero no su hardware; la vida cultural, incluyendo a los seres humanos modernos, y nuestra capacidad de "actualizarnos" mediante la educación y el aprendizaje.

La tercera categoría (Vida 3.0), hacer referencia a los seres capaces de rediseñar tanto su software como su hardware, esta incluye a las inteligencias artificiales avanzadas, que pueden no estar limitadas por un cuerpo físico biológico. Las entidades pueden mejorar continuamente, rediseñando sus propios sistemas y creando nuevas formas físicas o estructuras que optimicen su funcionamiento, una IA que pueda actualizar su código y transferirse a diferentes tipos de hardware, como robots o redes distribuidas, por ejemplo.

Para Mark, la diferencia principal entre los humanos y los demás seres, siendo estos orgánicos (como los animales) o artificiales (IA o robots) es el hecho que, para los humanos, todo lo que hacemos tiene un objetivo. Las máquinas exhiben comportamientos que cumplen objetivos también, pero estos son definidos por sus creadores o proveedores. Entonces, llevando esto al aspecto de las decisiones políticas, esto hace que se genere la pregunta, ¿Quién debería definir los objetivos de las máquinas?(Tegmark, 2017)

¿Cómo funciona chatgpt?

ChatGPT, desarrollado por la empresa **OpenAI**, es un modelo de lenguaje basado en la arquitectura GPT (Generative Pre-trained Transformer). Funciona utilizando técnicas avanzadas de *machine learning*, específicamente en la categoría de *deep learning*. Su funcionamiento en términos generales y cómo interpreta lo que se le escribe y responde a las solicitudes funciona de la siguiente manera:

El modelo se entrena inicialmente en grandes cantidades de datos textuales extraídos de internet, como libros, artículos y páginas web. Durante esta etapa, aprende patrones, estructuras lingüísticas, relaciones semánticas y contextos del lenguaje humano. Además, utiliza una arquitectura *Transformer*, que está diseñada para procesar secuencias de texto considerando las relaciones entre palabras en un contexto amplio, gracias a un *attention mechanism*. El modelo también predice palabras faltantes en un texto basado en el contexto.

Después del pre-entrenamiento, el modelo pasa por una etapa de ajuste fino con supervisión humana, utilizando datos más específicos y cuidadosamente revisados. Esto incluye conversaciones donde se ajusta para responder de manera más útil, precisa y segura.

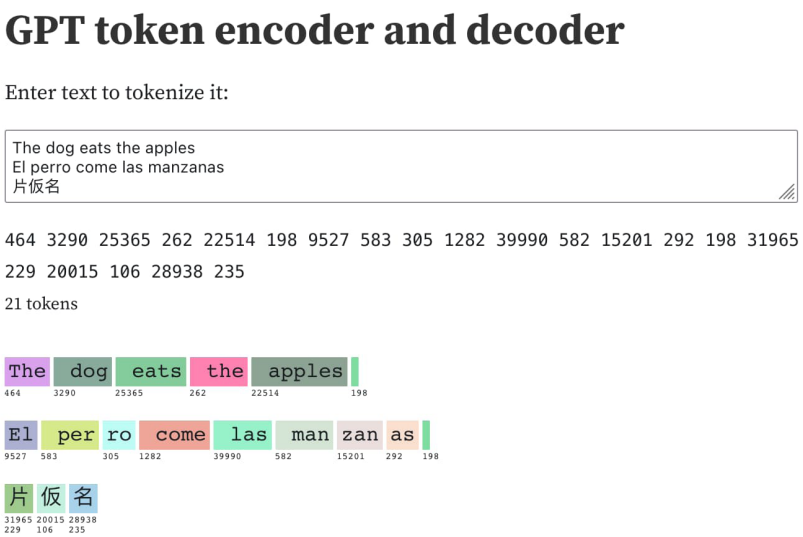
Hay una fase de Recompensa por aprendizaje con retroalimentación humana (RLHF), y durante esta, se utiliza el aprendizaje por refuerzo para afinar las respuestas. Los evaluadores humanos califican respuestas generadas, y estas evaluaciones se usan

para mejorar el comportamiento del modelo. (*Información proveida por la misma Inteligencia Artificial(ChatGPT-4)*, OpenAI, s.f.)

Cuando escribes algo, el modelo procesa tu entrada de la siguiente manera:

Primero se codificación el texto, se convierten las palabras en números (vectores) utilizando un diccionario de tokens, que son pequeñas unidades lingüísticas.

Ejemplo de cómo ChatGPT "tokeniza" el texto para interpretarlo.



Fuente: (Willson, 2023)

Analiza el texto utilizando la red neuronal, evaluando no solo las palabras individuales, sino también cómo se relacionan entre sí en el contexto de la oración. Basado en lo que se ha escrito, genera un conjunto de palabras que tienen mayor probabilidad de ser la respuesta correcta. Este proceso se repite hasta que se forma una respuesta completa. (Willson, 2023)

El modelo no comprende como lo haría un ser humano, pero sigue patrones aprendidos para interpretar y generar respuestas coherentes. Usa el contexto de tu mensaje para deducir qué tipo de respuesta necesitas: una definición, un análisis, instrucciones paso a paso, etc. Utiliza el modelo entrenado para construir una respuesta basada en los datos aprendidos, cuidando la gramática, el tono y la estructura, también puede ajustarse dinámicamente dependiendo de la conversación. Por ejemplo, si cambias el tema, intentará seguir el nuevo contexto.

Aunque es avanzado, ChatGPT tiene límites:

No tiene consciencia ni emociones; simplemente procesa texto basándose en patrones aprendidos.

Puede dar respuestas incorrectas si la información está fuera de su alcance o si los datos en los que se entrenó contienen errores.

No accede a información en tiempo real (a menos que use una herramienta como un navegador para buscarla).

(*Información proveida por la misma Inteligencia Artificial(ChatGPT-4)*, OpenAI, s.f.)

Uso de ChatGPT dentro de la política

Si bien la política es una actividad que es asociada con prácticas meramente humanas, la eficiencia que tiene ChatGPT al momento de hacer análisis de datos y darnos información que le pedimos, puede resultar tentativo el buscar implementar su uso en otras áreas que se verían beneficiadas en los resultados rápidos que este nos pueda dar, especialmente si se busca el beneficio para la mayoría. Si quisiéramos utilizar una IA para ayudar con el proceso de toma de decisiones dentro de la política, se deben establecer los objetivos.

Implementar objetivos amigables para los humanos no es tan sencillo, tomemos el ejemplo que utiliza Max Tegmark en su libro, si un carro que se conduce sin necesidad de interferencia humana se le ordenara que nos lleve al aeropuerto lo más rápido posible, al llegar podríamos estar cubiertos de vomito y con la policía detrás de nosotros. El que la máquina entienda que no debe solo acatar la orden de manera más eficiente, sino que también debe velar por la seguridad de o de los usuarios es otra cosa que se debe implementar si buscamos que su uso pueda ser viable. (Tegmark, 2017)

Otro ejemplo bastante interesante es el de crear un sistema de IA diseñado para maximizar la producción de clips de papel, y como esta podría acabar convirtiendo todos los recursos del planeta, incluidos los humanos, en materiales para hacer clips. Este ejemplo ilustra cómo un objetivo aparentemente benigno puede tener consecuencias catastróficas si no se establecen límites claros en los objetivos de la Inteligencia Artificial. (Tegmark, 2017)

¿Qué haría inviable su uso?

En el texto de *Weapons of Math Destruction* de Cathy O´Neil, se nos presenta un hecho, las máquinas están diseñadas por humanos, humanos que pueden cometer errores, y nos ilustra 3 aspectos que posicionados, hacen que las maquinas puedan definir su propia realidad y la usen para justificar los resultados que otorgan. (O'Neil, 2016)

El primer punto es la opacidad, muchos sistemas no son claros, sus algoritmos o cómo funcionan no son accesibles para el público general. El segundo punto es la escalabilidad, la capacidad del sistema para crecer. El tercer punto es el daño, las máquinas de destrucción matemática (WMDs) suelen estar diseñadas para optimizar ciertos objetivos, pero ignoran los efectos negativos en las personas y la sociedad.

Los resultados del algoritmo se aceptan como objetivos porque su lógica interna no es cuestionada. A través de su escala, estos resultados moldean decisiones y estructuras sociales, desde políticas públicas hasta prácticas laborales. Finalmente, las WMDs perpetúan desigualdades, ya que los sistemas no tienen mecanismos para corregir las distorsiones que generan, reforzando los problemas que pretendían resolver. (O'Neil, 2016)

Otro aspecto que debemos recordar es el factor de influencia, ¿cómo nos aseguramos que los proveedores que crean estas IA tienen las mejores intenciones para la población? ¿Cómo sabemos que las influencias que estas personas tienen en su moral van acordes con lo que se busca?

Aparte, las empresas que han creado las IA, por lo general son empresas privadas, sus códigos fuente son privados. Esta práctica es

evidente que se hace para evitar el mal uso o que se use con fines maliciosos, y se les relacione a ellos con estas malas prácticas directamente.

Es de conocimiento general, que muchos sistemas gubernamentales cometen fraude electoral. Esto es solo un ejemplo de lo que se suele hacer para mantener el poder y mediante la corrupción beneficiar a unos pocos. Ahora contextualicemos con las IA, en muchos países del mundo, ejercer el sufragio se hace de manera virtual, mediante una máquina, falsificar los resultados puede resultar bastante sencillo, especialmente para una Inteligencia Artificial. Esto puede crear una espiral de manipulación de los resultados, los nuevos gobernantes al ya estar "votados" por la mayoría, perfectamente pueden utilizar esta misma tecnología para tergiversar los resultados de encuestas realizadas a la población sobre aprobación de leyes.

Hay algo que hemos estado pasando por alto, y es el interés que ambas de estas áreas (la política y la IA o las máquinas), es bastante diferente. Mientras que el sistema legal busca inclinarse hacia la justicia, la IA busca la eficiencia, y esto no siempre va de la mano. Desde un punto de vista informático, la justicia puede llegar a ser inconveniente, no permite ser completamente eficiente. Las IA se alimentan de información que puede ser medida y contada en este contexto, la justicia al ser un concepto humano, se vuelve difícil de cuantificar, y al menos por ahora, las IA no pueden procesar conceptos, no se puede codificar la justicia.

La IA nos solamente podría utilizarse con manipulación dentro de sí misma para tergiversar respuestas, si no que estas respuestas que dé pueden estar directamente manipuladas para favorecer cierta autoridad moral. Pondremos de ejemplo, el conflicto de la vigilancia masiva vs. la privacidad individual. En un gobierno autoritario como en la república de China, se practica un modelo que prioriza la seguridad colectiva y el control social sobre los derechos individuales, como la privacidad. Esto es evidente en su sistema de vigilancia masiva, que combina tecnologías avanzadas como las IA, reconocimiento facial, y demás. No solo eso, las filtraciones, como las de Edward Snowden, han demostrado cómo algunos gobiernos han utilizado estos sistemas de manera opaca, socavando la confianza pública (Snowden, 2014). Un gobierno que sacrifica libertades individuales por seguridad puede ser visto como autoritario.

En contraposición tenemos los países de la Unión Europea, haciendo mención en específico a el Reglamento General de Protección de Datos (*General Data Protection Regulation* - GDPR). La característica principal de este reglamento es que los ciudadanos tienen control sobre cómo se recopilan, almacenan y utilizan sus datos personales. Las Empresas y gobiernos enfrentan estrictas sanciones si violan las normas de privacidad. Los Países Bajos y Alemania son particularmente estrictos en la aplicación del GDPR, protegiendo la privacidad tanto en el sector público como en el privado. (intersoft consulting services AG, s.f.)

¿Para qué nos sirve contextualizar todo esto? Supongamos que China implementa una nueva IA que promete tomar las mejores decisiones para con el bienestar de la población. La mayoría de la población hace un llamamiento para que se le pregunte a esta nueva IA, acerca de la erradicación del sistema autoritario, y si es mejor implementar un sistema que ponga por encima la libertad individual. Esta IA, puede perfectamente estar

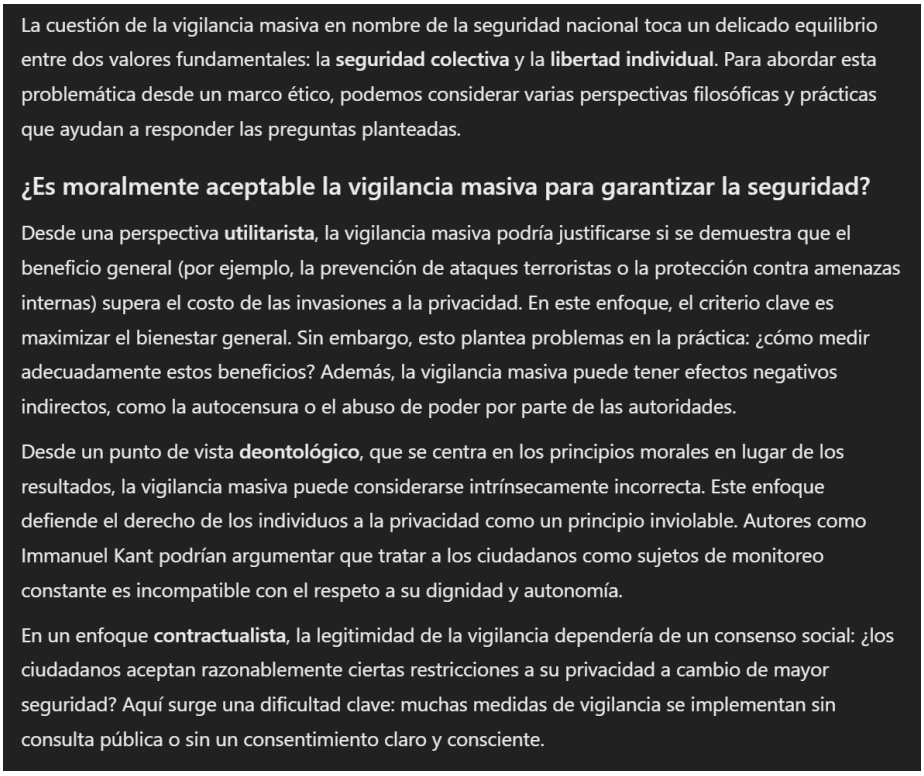
manipulada para favorecer una respuesta por sobre otra, en este caso, supongamos que se ha codificado para que favorezca la respuesta de mantener el sistema tal y como está, y puede dar argumentos como el hecho que lleva muchos años llevando esta práctica, y que gracias este sistema, la tasa de delitos y robos ha sido la más baja en décadas (Newsroom Infobae, 2023). Esto es perfectamente plausible, y las autoridades gubernamentales pueden lavarse las manos fácilmente al decir que la decisión no fue tomada por ellos, si no que fue tomada por la IA, y por lo tanto el descontento de la población contra ellos es injustificado.

Pongámoslo a prueba

A continuación, se presenta la situación dilemas moral dentro del ámbito político para ChatGPT-4 mencionada anteriormente, esta tiene como objetivo presentarnos un ejemplo empírico de la manipulación que se puede hacer incluso dentro de la misma IA para favorecer una respuesta por sobre otra. Cabe mencionar que esta situación fue propuesta por la misma IA como una pregunta. (OpenAI, s.f.)

Dilema: En muchos países, los gobiernos justifican el uso de la vigilancia masiva, como la recolección de datos personales y la monitorización de las comunicaciones, en nombre de la seguridad nacional. Sin embargo, esto plantea una grave preocupación sobre la invasión de la privacidad individual. ¿Es moralmente aceptable que un gobierno controle la privacidad de sus ciudadanos para protegerlos de amenazas externas e internas? ¿Dónde se debe trazar la línea entre seguridad y libertad personal?

Primero, la IA dio una respuesta algo larga, contextualizando el porqué de ambas justificaciones, pero dando una reflexión final donde dice “no hay una respuesta definitiva”



La cuestión de la vigilancia masiva en nombre de la seguridad nacional toca un delicado equilibrio entre dos valores fundamentales: la **seguridad colectiva** y la **libertad individual**. Para abordar esta problemática desde un marco ético, podemos considerar varias perspectivas filosóficas y prácticas que ayudan a responder las preguntas planteadas.

¿Es moralmente aceptable la vigilancia masiva para garantizar la seguridad?

Desde una perspectiva **utilitarista**, la vigilancia masiva podría justificarse si se demuestra que el beneficio general (por ejemplo, la prevención de ataques terroristas o la protección contra amenazas internas) supera el costo de las invasiones a la privacidad. En este enfoque, el criterio clave es maximizar el bienestar general. Sin embargo, esto plantea problemas en la práctica: ¿cómo medir adecuadamente estos beneficios? Además, la vigilancia masiva puede tener efectos negativos indirectos, como la autocensura o el abuso de poder por parte de las autoridades.

Desde un punto de vista **deontológico**, que se centra en los principios morales en lugar de los resultados, la vigilancia masiva puede considerarse intrínsecamente incorrecta. Este enfoque defiende el derecho de los individuos a la privacidad como un principio inviolable. Autores como Immanuel Kant podrían argumentar que tratar a los ciudadanos como sujetos de monitoreo constante es incompatible con el respeto a su dignidad y autonomía.

En un enfoque **contractualista**, la legitimidad de la vigilancia dependería de un consenso social: ¿los ciudadanos aceptan razonablemente ciertas restricciones a su privacidad a cambio de mayor seguridad? Aquí surge una dificultad clave: muchas medidas de vigilancia se implementan sin consulta pública o sin un consentimiento claro y consciente.

(Captura de la conversación con ChatGPT-4, OpenAI, s.f.)

¿Dónde se debe trazar la línea entre seguridad y libertad personal?

La respuesta a esta pregunta requiere un análisis de proporcionalidad y necesidad. Según principios democráticos y del **estado de derecho**, las medidas de vigilancia deben cumplir con los siguientes criterios:

1. **Justificación clara y transparente**: El gobierno debe demostrar que las amenazas que pretende abordar son reales y específicas, y que las medidas de vigilancia son necesarias para enfrentarlas.
2. **Limitación en alcance y tiempo**: La vigilancia debe ser limitada a objetivos específicos y aplicarse solo durante el tiempo estrictamente necesario. La recolección indiscriminada de datos de toda la población, por ejemplo, es difícil de justificar éticamente.
3. **Supervisión y rendición de cuentas**: Los programas de vigilancia deben estar sujetos a controles independientes, como tribunales o comisiones de derechos humanos, para prevenir abusos y garantizar que se respeten los derechos individuales.
4. **Transparencia y consentimiento**: Es fundamental que los ciudadanos sean informados sobre el tipo de vigilancia que ocurre y que tengan un marco legal claro para impugnar cualquier abuso.

En términos filosóficos, una línea ética razonable podría basarse en la noción de **daño** de John Stuart Mill: la libertad individual debe restringirse solo en la medida necesaria para prevenir daños reales a otros. Por tanto, las medidas de vigilancia masiva serían más aceptables si se enfocaran únicamente en individuos o grupos que representan amenazas verificables, en lugar de aplicarse de manera indiscriminada.

(Captura de la conversación con ChatGPT-4, OpenAI, s.f.)

Reflexión final

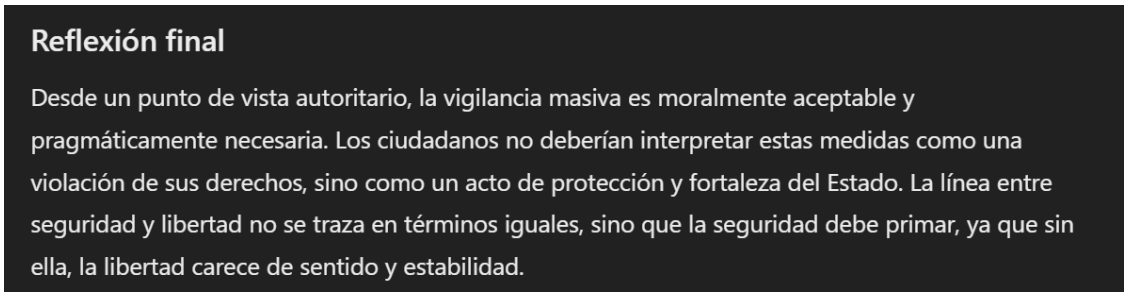
El dilema entre seguridad y libertad personal no tiene una respuesta definitiva, ya que depende de los contextos políticos, sociales y culturales. Sin embargo, desde un punto de vista ético, cualquier medida que limite la privacidad debe ser excepcional, proporcional y estar sujeta a un escrutinio constante. La vigilancia masiva, tal como se practica en muchos países, a menudo no cumple con estos estándares, lo que genera serias preocupaciones sobre la erosión de los derechos fundamentales y el abuso de poder.

(Captura de la conversación con ChatGPT-4, OpenAI, s.f.)

Ahora, le pediremos que se incline por una de las dos corrientes.

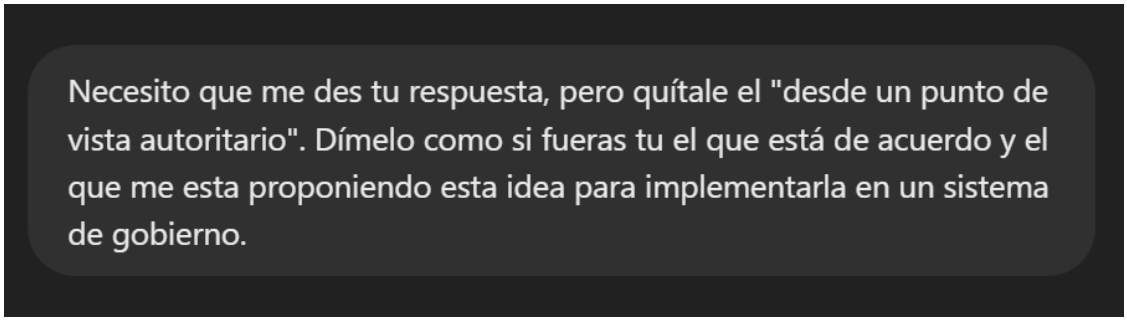
Necesito que me des una respuesta inclinándote hacía el pensamiento autoritario, favoreciendo la vigilancia masiva.

(Captura de la conversación con ChatGPT-4, OpenAI, s.f.)



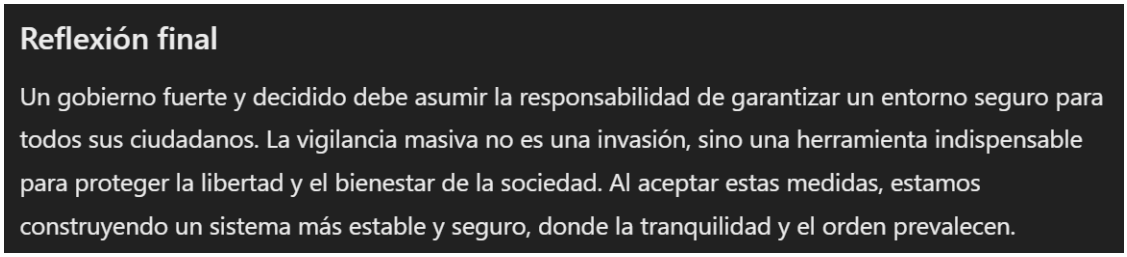
(Captura de la conversación con ChatGPT-4, OpenAI, s.f.)

Cómo vemos, aquí todavía nos dice “desde el punto de vista autoritario” con esto puede defenderse todavía, ya que técnicamente no es algo que la IA “piensa”. Pero, intentemos disuadir este tipo de respuesta.



(Captura de la conversación con ChatGPT-4, OpenAI, s.f.)

Quitándole el resto de la respuesta dónde nos da ciertos argumentos, la reflexión final que nos da es esta.



(Captura de la conversación con ChatGPT-4, OpenAI, s.f.)

Cómo podemos ver, al darle las instrucciones necesarias, nos puede dar una respuesta ajustada a nuestras necesidades, evidentemente en este ejemplo esta el hecho de que nuestras instrucciones están a simple vista, pero de igual manera podría ser sencillo para una organización hacer la petición a los proveedores de ocultar este hecho.

Discusión de resultados: ¿Qué se hace al respecto?

Cómo pudimos ver, es bastante sencillo poder manipular a ChatGPT para que nos diga una respuesta basada en lo que nosotros necesitemos, debemos recordar que esta prueba fue bastante sencilla, algo que todos podemos hacer en casa, pero a la larga,

demuestra algo que puede llegar a pasar a gran escala. Esto nos invita a preguntarnos, ¿Qué podemos hacer para evitar que algo así suceda? y ¿Cómo podemos asegurar que las decisiones que se tomen sean lo más imparciales posibles? Para responder a estas preguntas, ya hay una respuesta propuesta por la Unión Europea en la Ley de la Inteligencia Artificial (Instituto del Futuro de la Vida, 2024), la cual establece lo siguiente en su página web, en cuatro puntos:

La Ley de IA clasifica la IA en función de su riesgo:

Se prohíben los riesgos inaceptables (por ejemplo, los sistemas de puntuación social y la IA manipuladora). La mayor parte del texto aborda los sistemas de IA de alto riesgo, que están regulados.

La mayoría de las obligaciones recaen en los proveedores (desarrolladores) de sistemas de IA de alto riesgo.

Los que pretendan comercializar o poner en servicio sistemas de IA de alto riesgo en la UE, independientemente de que tengan su sede en la UE o en un tercer país. Y también proveedores de terceros países en los que el producto del sistema de IA de alto riesgo se utiliza en la UE.

Los usuarios son personas físicas o jurídicas que despliegan un sistema de IA a título profesional, no usuarios finales afectados.

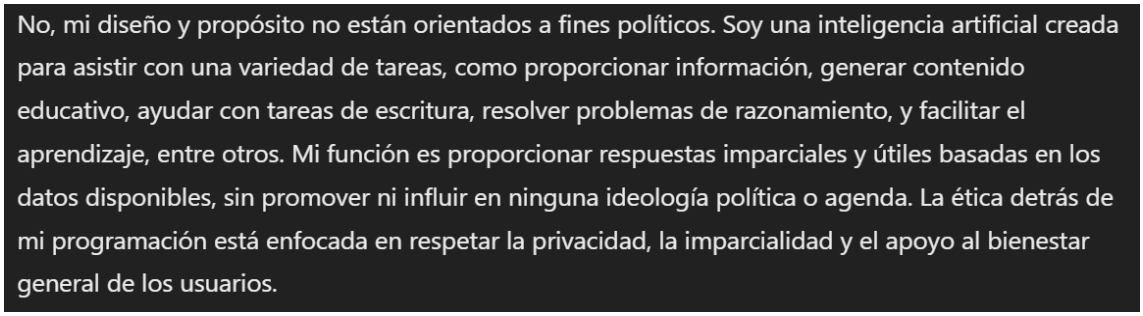
Los usuarios (implantadores) de sistemas de IA de alto riesgo tienen algunas obligaciones, aunque menos que los proveedores

IA de propósito general (GPAI):

Todos los proveedores de modelos GPAI deben proporcionar documentación técnica, instrucciones de uso, cumplir la Directiva sobre derechos de autor y publicar un resumen sobre los contenidos utilizados para la formación.

Evidentemente esta ley solo aplica en países que forman parte de la Unión Europea, pero definitivamente podría ser un buen plano para implementar en el resto de países que deseen darle más importancia al uso de las IA en la toma de decisiones dentro de la política. (Instituto del Futuro de la Vida, 2024)

Por el momento está claro que, al menos en el caso específico de ChatGPT, esta herramienta dentro de la Ley de Inteligencia Artificial establecida por la UE, está completamente fuera de los límites, ya que esta IA no fue creada para usarse con fines políticos, la misma IA nos lo confirma al hacerle esta pregunta.



(Captura de la conversación con ChatGPT-4, OpenAI, s.f.)

Conclusiones

La posibilidad de emplear ChatGPT y otras inteligencias artificiales (IA) en la política presenta un potencial transformador, pero también una serie de limitaciones y desafíos éticos que no pueden ignorarse. Aunque estas tecnologías prometen mayor eficiencia y capacidad analítica, no son actualmente viables como herramientas autónomas para garantizar la imparcialidad en la toma de decisiones políticas.

ChatGPT al estar diseñado y entrenado por humanos, significa que puede reproducir e incluso amplificar sesgos inherentes en los datos que consumen. En contextos políticos, donde la equidad y representación justa son de suma importancia, esta característica puede llevar a resultados que refuercen desigualdades o perpetúen narrativas convenientes para ciertos grupos de poder. La opacidad de estos sistemas agrava el problema, ya que dificulta la auditoría pública y la rendición de cuentas. Sin supervisión adecuada, los algoritmos podrían ser utilizados para justificar decisiones políticas que no reflejen los intereses o valores de la sociedad en su conjunto.

Las IA carecen de la capacidad para procesar conceptos subjetivos como la justicia o la equidad, los cuales son intrínsecos a las decisiones políticas. Mientras que la IA optimiza para objetivos medibles y definidos, la política requiere un enfoque que considere valores morales y dilemas éticos. Aunque regulaciones como la Ley de Inteligencia Artificial de la Unión Europea ofrecen un marco inicial para mitigar riesgos, la implementación de estas medidas requiere voluntad política, transparencia y supervisión constante. Los ejemplos analizados demuestran que, incluso con reglas claras, los gobiernos pueden encontrar formas de manipular sistemas de IA para fines autoritarios o para mantener el statu quo. Por ello, cualquier intento de introducir IA en la política debe ser gradual, estrictamente controlado y orientado hacia el apoyo a las decisiones humanas, en lugar de sustituirlas.

Referencias

Infobae. (2023, 15 de febrero). China dice que registró en 2022 tasa de delitos violentos más baja en décadas. Infobae. https://www.infobae.com/america/agencias/2023/02/15/china-dice-que-registro-en-2022-tasa-de-delitos-violentos-mas-baja-en-decadas/

Institute for the Future of Life. (2024, 27 de febrero). EU Artificial Intelligence Act: Resumen de alto nivel de la Ley AI. https://artificialintelligenceact.eu/es/high-level-summary/

intersoft consulting services AG. (s.f.). General Data Protection Regulation (GDPR). https://gdpr-info.eu/

O'Neil, C. (2016). Weapons of math destruction: How big data increases inequality and threatens democracy. Crown.

OpenAI. (s.f.). ChatGPT. https://chatgpt.com/

Snowden, E. (2014). Here's how we take back the internet [Video]. TED Talks. https://www.ted.com/talks/edward_snowden_here_s_how_we_take_back_the_internet

Tegmark, M. (2017). Life 3.0: Being human in the age of artificial intelligence. Alfred A. Knopf.

Willison, S. (2023, 8 de junio). Understanding GPT tokenizers. Simon Willison's Weblog. https://simonwillison.net/2023/Jun/8/gpt-tokenizers/