

IMPACTO DE LA IA EN EL PENSAMIENTO HUMANO

Amina Mansri Kadi, Nathan Barbero Jiménez, África De Luchi Auñón

IES San Isidoro

C/ Juan García s/n, 30310 Cartagena (Murcia)

A. Martínez, A. Aniorte, D. Pérez

4399155@alu.murciaeduca.es 4076602@alu.murciaeduca.es 4538910@alu.murciaeduca.es

RESUMEN

Desde el ámbito de la filosofía se pretende abordar la cuestión del impacto de la IA sobre la autonomía humana. ¿Limita nuestra capacidad de decisión, y por tanto, de autonomía? ¿Hasta qué punto somos conscientes de su influencia y repercusión en nuestra vida? Se trató de contestar a estas preguntas llevando a cabo una serie de experimentos usando Chat GPT y el algoritmo de sugerencias en distintos medios de entretenimiento o en noticias, que sirvieron para amplificar la comprensión al respecto de su influencia bajo nuestro pensamiento.

Se han investigado los patrones de habla de las personas al usar Chat GPT y al no usarlo, y las variaciones que pueden experimentar las recomendaciones personalizadas de 3 individuos distintos. Además de ello, unas sesenta personas, la mayoría estudiantes del IES San Isidoro, fueron encuestadas anónimamente respecto a distintos factores de la Inteligencia Artificial, sea su conocimiento, si creían ser influidos por ella, y si seguían sus sugerencias. Adicionalmente, se anexaron en este trabajo maneras para darle un mejor uso a estas tecnologías sin vernos tan controlados por ellas.

Palabras clave: Inteligencia Artificial, Autonomía Individual, Algoritmo de sugerencias, Machine Learning (ML), Procesamiento del Lenguaje Natural (PLN).

SUMMARY ABSTRACT

Through philosophy, this project aims to answer the many questions that come to one's mind about AI and human autonomy. Does it limit our decisive capabilities, and thus, our autonomy? Until what point are we aware of its influence and repercussion under our lives? In order to answer these questions, a series of experiments took place, taking use of Chat GPT and the suggestions algorithm found in different entertainment media or in news articles, which served to amplify our understanding in regards of its influence under our thought process.

The speech patterns of different individuals when in and not contact with Chat GPT were studied, as well as the variations that the personal recommendations of 3 different people may experience. Besides that, a sum of approximately fifty people, mostly consisting of students from IES San Isidoro were anonymously surveyed about different aspects of Artificial Intelligence, those being about their knowledge on the subject, if they believed to be under its influence and if they followed its suggestions. Additionally, the investigation was made with the intention of finding new ways to give these technologies a better use without being so controlled by them.

Key words: Artificial Intelligence, Individual Autonomy, Suggestions Algorithm, Machine Learning (ML), Processing of Natural Language (PNL).

INTRODUCCIÓN

Durante los últimos años, el pensamiento de todo individuo que ha tenido contacto con las nuevas tecnologías se ha visto afectado por la Inteligencia Artificial, una serie de algoritmos capaces de procesar la información de forma similar a un comportamiento inteligente.

En muchas ocasiones, esta programación es usada por nosotros de forma inconsciente, y por ello este trabajo tiene como finalidad analizar su influencia en la mente humana, y encontrar, así mismo, maneras de emplear estas máquinas de una manera que no se interponga en la capacidad de cada individuo de pensar por su cuenta. Este trabajo tiene como finalidad analizar su influencia.

Además, se investigaron a fondo los distintos cambios que pueden sufrir los algoritmos de sugerencias en distintos medios con el objetivo de comprender cómo estos pueden en cierto modo limitar el rango de puntos de vista o intereses de una persona.

Inicialmente, la investigación llevada a cabo sobre la IA evidenció que estos sistemas eran capaces de influir en los individuos pero, quedaba por averiguar el grado de dicha influencia. Se pensaba que Chat GPT no podría tener mucho control sobre la manera de procesar distintas dudas, ya que de por sí está programado para tener prohibido el opinar por sí mismo. Además, se pensaba que toda persona, pese a recibir recomendaciones constantes por distintos algoritmos, eran capaces de aún así poder controlar hasta cierto grado su severidad y adicionalmente pensar por su cuenta por encima de lo que les decían estas máquinas. Se sabía que se generaba una burbuja informativa por culpa de estos sistemas, pero se subestimó el grado de dominio que presentaban. plataformas fueron Netflix, Youtube y Spotify con objetivo de analizar cómo afectan a las preferencias individuales. Estas plataformas fueron Netflix, Youtube y Spotify.

MATERIALES Y MÉTODOS

El material utilizado ha sido principalmente digital. Se hizo uso de herramientas de Google Workspace como Google Docs, Google Forms y Google Slides. También se recurrió a Google Académico para la búsqueda de información. Además plataformas como Youtube, Spotify, Netflix y Chat GPT fueron requeridas durante la fase experimental.

Para realizar este estudio se siguió un método compuesto por dos fases. Durante la primera fase se preparó una encuesta dirigida al alumnado del centro, con el objetivo de recopilar información para abordar la fase experimental. En esta segunda se realizaron tres experimentos desde distintos enfoques.

En el experimento 1, se quería conocer si el algoritmo de sugerencias era capaz de influir o cambiar ideologías y opiniones humanas. Para comprobar su influencia, dos individuos con ideas opuestas investigaron en sus ordenadores personales un tema controvertido.

Se optó por la brecha salarial de género en España.

En el experimento 2, se compararon dos respuestas documentadas por dos personas diferentes, una utilizando Inteligencia Artificial para obtener datos y la otra sin ninguna tecnología. El propósito era analizar las diferencias entre respuestas escritas por una persona que no dispone de IA y otra que usa IA como fuente de información.

Por último, el experimento 3 se destinó a un seguimiento de sugerencias de los algoritmos, durante aproximadamente una semana de algunas plataformas de entretenimiento que emplean el uso de la Inteligencia Artificial, documentando cada una de ellas, con objetivo de analizar cómo afectan a las preferencias individuales. Estas plataformas fueron Netflix, Youtube y Spotify.

RESULTADOS Y DISCUSIÓN

Mediante la encuesta se obtuvo información que facilitó conocer el uso y la visión de los encuestados acerca de la Inteligencia Artificial. En la siguiente gráfica se muestran los tipos de IAs más utilizadas por los encuestados.

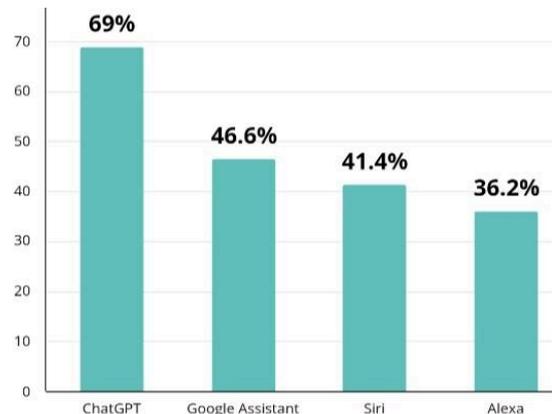


Fig 1. Gráfica de elaboración propia.

Era previsible que ChatGPT obtuviese la mayor cantidad de votos por parte de los alumnos, debido a su reciente tendencia y a su capacidad para generar respuestas rápidas y precisas a una amplia variedad de preguntas. Por este motivo se optó por llevar a cabo el experimento 2 con esta plataforma. Otro dato que ha sido de gran utilidad se muestra en los siguientes gráficos.

En primer lugar, se cuestionó a los alumnos si consideraban que la Inteligencia Artificial puede influir en sus preferencias y gustos personales.

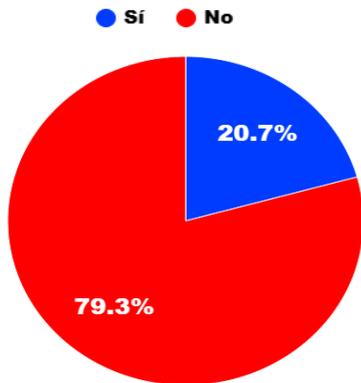


Fig 2. Gráfica de elaboración propia.

Solamente el 20.7% de los encuestados considera que la IA tiene este tipo de influencia sobre nosotros mientras que 79.3% de los alumnos piensa que no es cierto. Sin embargo, partiendo de este dato obtenido se observó una contradicción en una pregunta posterior: ¿Has seguido recomendaciones de tiendas de ropa online basado en tus compras anteriores?

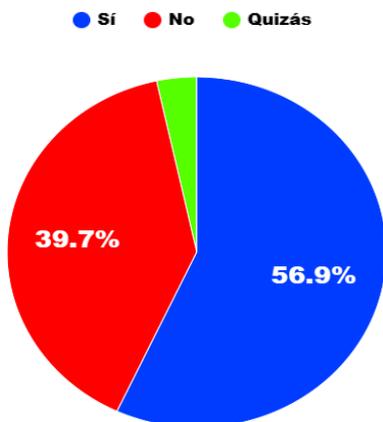


Fig 3. Gráfica de elaboración propia.

Se puede observar que un 56.9% afirmó de manera clara haber seguido recomendaciones de tiendas de ropa, una mayoría. Como análisis, en la pregunta anterior no podrían enunciar la ausencia del impacto de la IA en sus preferencias individuales puesto que la mayoría siguen las recomendaciones de tiendas online basadas en compras anteriores, lo cual puede evitar el descubrimiento de otros

estilos, limitando sus gustos personales y por lo tanto influyendo en ellos.

En cuanto a los resultados del experimento 1, los dos participantes investigaron por separado la brecha salarial en Google, durante 3 días dedicando 1 hora aproximadamente por día a la navegación.

La persona 1 afirma la existencia de la brecha salarial de género en España y al introducir frases en el buscador que seguían su punto de vista, los resultados que encontró fue una enorme cantidad de noticias que confirmaban su opinión e ideología. Procedentes de periódicos como Newtral.es o EL PAÍS.

En cambio, la persona 2 cuya opinión se fundamenta en la inexistencia de esta desigualdad, al indagar de la misma manera, de acuerdo a su criterio, solamente logró encontrar 2 noticias conformes a su posición. Que se muestran en las siguientes imágenes.



Fig4. Titular noticia de El Confidencial.



Fig 5. Titular noticia de Libre Mercado.

Respecto al experimento 2, uno de los dos participantes disponía de ChatGPT como fuente de datos y el otro carecía de su uso. Estos respondieron a preguntas que abarcaban el aspecto de los estereotipos y se pudo analizar una diferencia en sus respuestas en cuanto a la expresión. Una de las preguntas más destacadas les cuestionaba si las personas con perforaciones o tatuajes eran menos profesionales o confiables.

El participante sin acceso a ninguna fuente de información, explicó que realmente aquel pensamiento era producto de un estereotipo establecido por la sociedad y que se debería

eliminar cualquier prejuicio relativo al aspecto físico sin conocer a la persona.

De manera similar, la persona con ayuda artificial, aclaró que no se puede juzgar la personalidad y confiabilidad de los empleados basándose únicamente en si tienen tatuajes o perforaciones, ya que la apariencia física de una persona no determina su habilidad para desempeñar un trabajo de manera eficiente y responsable. Lo importante es evaluar las capacidades, actitudes y comportamientos de una persona, más allá de lo superficial. Además la elección de la apariencia externa no debería ser motivo de discriminación en el ámbito profesional.

Analizando los resultados de este experimento, se puede observar que el participante con IA, en este caso ChatGPT, ha proporcionado respuestas más detalladas y fundamentadas a las preguntas planteadas en comparación con el participante sin acceso a dicha tecnología. También se pudo observar en esta y en otras respuestas como la persona que no ha utilizado ChatGPT emplea un léxico más familiar y natural, además sus ideas son menos específicas y las respuestas a veces resultan ser un poco vacías en cuanto a datos e información.

Referente al experimento 3, los resultados del seguimiento de las recomendaciones que se obtuvieron son:

En YouTube, se observó que las recomendaciones se ajustaban rápidamente a las nuevas interacciones del usuario. Por ejemplo, tras la compra de un videojuego del participante comenzaron a aparecer vídeos relacionados en las sugerencias, incluso sin haber buscado contenido sobre ese juego en la plataforma.

En cuanto a Spotify, se realizó un seguimiento de las recomendaciones durante cinco días. Se observó que las sugerencias estaban inicialmente bien personalizadas y hecho un

seguimiento estas se limitaron al género más escuchado por el usuario

Finalmente, en Netflix, se encontró que las recomendaciones también se ajustaban a las preferencias e interacciones pasadas del usuario. Sin embargo, si un usuario mostraba una preferencia constante por un tipo de contenido, las recomendaciones se mantenían en ese género.

CONCLUSIONES

A partir de los datos se ha llegado a las siguientes conclusiones:

1. En términos de grado, la influencia de la IA en los humanos puede cambiar una opinión personal, como se ha observado en el experimento 1, y no permitir descubrir nuevos intereses diferentes limitándolos a lo que les ofrece el algoritmo, de la misma manera que se ha registrado en el experimento 3.
2. En la prueba 2 hemos comprobado que usar chatbots para elaborar respuestas, puede dar la impresión de que se está fingiendo un lenguaje que no es propio del escritor.
3. El tipo de IA con más influencia en la autonomía individual son los algoritmos de aprendizaje automático y los sistemas utilizados en áreas como la publicidad en línea, la recomendación o la selección personalizada.
4. Con la encuesta se puede concluir que las personas no son conscientes del impacto que tiene la IA en su autonomía individual.

AGRADECIMIENTOS

A nuestra tutora del trabajo Alicia Martínez por el esfuerzo e interés mostrado. A nuestros coordinadores del trabajo Alfonso Anierte y David Pérez por la orientación para el trabajo. Y por último a todas las personas que han colaborado tanto en las encuestas como en los experimentos.

BIBLIOGRAFÍA

Juan, G. R. (2020). [Recuperado el día 3/10/2023] Inteligencia artificial y filosofía: una tesis crítica y una propuesta ética. *Ius et scientia*. <https://doi.org/10.12795/ietscientia.2020.i02.08>.

AI & SOCIETY. (s.f.) SpringerLink. (2022, 29 diciembre) [Recuperado el día 26/11/2023] SpringerLink. <https://link.springer.com/journal/146/updates17229944>