

Artículo de Investigación

ChatGPT y discurso político. Revisión documental

ChatGpt and political discourse: A review

Rosa María Ricoy-Casas¹: Universidad de Vigo and UNED-Lugo, España.

rricoy@uvigo.gal

Raquel Fernández-González: Universidad de Vigo, España.

raquelf@uvigo.gal

Fecha de Recepción: 13/05/2024

Fecha de Aceptación: 05/08/2024

Fecha de Publicación: 25/09/2024

Cómo citar el artículo

Ricoy-Casas, R. M. y Fernández-González, R. (2024). ChatGPT y discurso político. Revisión documental [ChatGpt and Political discourse: A review]. *European Public & Social Innovation Review*, 9, 01-23. <https://doi.org/10.31637/epsir-2024-841>

Resumen

Introducción: ChatGPT ya se está utilizando por miembros del congreso, políticos, periodistas, estudiantes, científicos y administraciones públicas de todo el mundo. En política ha desplegado numerosas utilidades: para la generación de ensayos o discursos; para comprender las tendencias políticas emergentes; para simular debates y discusiones sobre temas de importancia para los votantes; para diseñar políticas públicas que se encuentren más en línea con la opinión pública mayoritaria o que contenten las aspiraciones de un mayor número de personas; identificar quiénes tienen una influencia significativa en la agenda pública, etc. **Metodología:** Este estudio, utiliza la metodología PRISMA y analiza datos obtenidos de la Web of Science y Google Scholar. **Discusión:** No obstante, algunos problemas inherentes a su funcionamiento (como errores, sesgos, inclinación política, desprotección de datos, vulneración de derechos como los relacionados con la protección de datos o la propiedad intelectual, impacto medioambiental) han condicionado su utilización en varios países. **Conclusiones:** Este trabajo pretende visibilizar estos aspectos en relación con la política, a la construcción de los discursos con fines electorales, centrando el debate en su uso ético, el respeto a los derechos fundamentales y el impacto en nuestras democracias.

¹ Autor Correspondiente: Rosa María Ricoy-Casas. Universidad de Vigo y UNED (España).

Palabras clave: ChatGPT; chatbot; inteligencia artificial; procesamiento del lenguaje natural; machine learning; tecnología; política; derechos fundamentales.

Abstract

Introduction: ChatGPT is already being used by members of congress, politicians, journalists, students, scientists, and public administrations around the world. In politics it has deployed numerous uses: for the generation of essays or speeches; to understand emerging political trends; to simulate debates and discussions on issues of importance to voters; to design public policies that are more in line with majority public opinion or that satisfy the aspirations of a greater number of people; identify who has a significant influence on the public agenda, etc. **Methodology:** This study uses the PRISMA methodology and analyzes data obtained from the Web of Science and Google Scholar. **Discussion:** However, some problems inherent to its operation (such as errors, biases, political bias, lack of data protection, violation of rights such as those related to data protection or intellectual property, environmental impact) have conditioned its use in several countries. **Conclusions:** This paper aims to make these aspects visible in relation to politics, the construction of discourses for electoral purposes, focusing the debate on their ethical use, respect for fundamental rights and the impact on our democracies.

Keywords: ChatGPT; chatbot; artificial intelligence; natural language processing; machine learning; technology; politics; fundamental rights.

1. Introducción

En este momento asistimos a la utilización de determinadas tecnologías sin suficiente conocimiento, análisis y/o criterio. Por ejemplo, el uso de sistemas de IA para tomar decisiones implica claramente no solo aspectos económicos, sino también sociales y éticos (por la forma en que los sistemas están diseñados y programados, su propio funcionamiento, o las fuentes de información de las que se nutren), de ahí la importancia de aportes interdisciplinarios y una mejora en su testeo, análisis, utilización y regulación. ChatGPT es tan sólo uno de los ejemplos de herramientas tecnológicas que inundan nuestras vidas en todos sus espacios, incluso en la política. En ese camino se encuentran los posibles “metaversos”, esos espacios virtuales 3D persistentes, sincrónicos, interactivos e inmersivos, con una sensación de autopercepción o relación directa compartida que se fusionará con nuestra imagen y entorno laboral físico o social, a través de nuestro gemelo digital o avatar, con los que gobiernos y partidos políticos comienzan a experimentar (Ricoy-Casas, 2022a). También, la utilización de hologramas y avatares en la comunicación y el marketing como técnica de persuasión en el escenario político (Ricoy-Casas, 2022b).

1.1. ¿Qué es ChatGPT y cuándo se ha producido su creación? Principales características

Un *chatbot* es un programa informático (software) basado en la inteligencia artificial (IA), que es capaz de mantener una conversación con un internauta sobre un tema específico. La mayoría de nosotros ya hemos interactuado en algún servicio de marketing o venta de este tipo, que trabaja en estos casos a modo de asistente (sin que haya una persona física contestando) para solucionar alguna duda a través de mensajes o voz. Interactúa de manera automática con usuarios y potenciales clientes, y son capaces de aprender, pueden llegar a intuir los hábitos, gustos y preferencias de los usuarios. Ello es posible gracias a determinadas herramientas de IA como el Procesamiento del Lenguaje Natural (PNL) que les permite entendernos, o el *Machine Learning* (ML) con el que aprenden automáticamente de cada una de sus interacciones. Eso significa que pueden responder de una manera muy personalizada e ir más allá de ofrecer respuestas prefabricadas, puesto que puede consultar

repositorios de datos y utilizar esa información para generar una respuesta. Asimismo, intenta que la interacción parezca con un humano, simulando o generando lenguaje natural, aunque todavía deben desarrollarse más para captar los dobles sentidos, el sarcasmo o los estados de ánimo. Además de estos chatbots de atención al cliente, existen otros muchos como los asistentes virtuales: Siri (Apple), Cortana (Windows), Alexa (Amazon), Aura (Telefónica), Bixby (Samsung), Irene (Renfe), Sara (Correos), o Anna (Ikea), muy serviciales (la mayoría mujeres y con “obsequiosidad femenina”), algo criticado como sesgo machista (UNESCO, 2019). También visibiliza los prejuicios de género codificados en productos tecnológicos y la “paradoja de la igualdad de género en las TICs”.

ChatGPT es un chatbot lanzado el 30 de noviembre de 2022, desarrollado por OpenAI. “GPT” significa “generative pre-trained transformer” por la forma en que se procesa el lenguaje, comprendiendo los patrones y la estructura del lenguaje de forma natural (emulando que estás hablando con otro humano). Microsoft ha integrado esta tecnología en la mayoría de sus servicios, desde su buscador Bing a aplicaciones de oficina como Word o Excel, lo que ha sido replicado en otras compañías como Google, que a principios de 2023 lanzó “Bard”, y la china Baidu con “Ernie Bot”. GPT-4o es última versión gratuita, después de presentar GPT en 2018, GPT-2 en 2019, GPT-3 en 2020, en 2022 GPT-3.5, GPT-4 en 2023 y en 2024 GPT-4o (“o” de “omni”), que integra en una sola herramienta funcionalidades de chatbots (como ChatGPT 3.5), asistentes de voz (como Siri o Alexa) y modelos de procesamiento de imágenes. Tan sólo después de unos meses desde su creación, llegó a lograr una cifra superior a los 100 millones de usuarios a nivel mundial.

1.2. *¿Cómo funciona ChatGPT?*

Inicialmente, debe crearse un usuario (para lo que te solicita información sobre tu email y teléfono móvil) y, a continuación, se abre un chat en el que realizas tu solicitud de información o generación de texto. Prácticamente, desde las versiones iniciales es posible pegar un artículo y pedirle que lo resuma de forma que lo entienda un alumno de doctorado o un estudiante de primaria, y también puede desarrollar una narración o un texto realizando con ciertas condiciones como limitar la búsqueda a un número máximo de caracteres o palabras, que escriba los textos de una manera determinada (por ejemplo, a modo de guion para subirlo a una red social), e incluso en un tono o estilo concreto (informal, académico, con cierto acento, o como un niño de 5 años, etc.). También pueden traducirse textos a otro idioma y realizarse consultas similares a las que realizamos a través de buscadores como Google, preguntando por los mejores restaurantes de una ciudad o los libros más leídos de ciencia ficción. Así, la herramienta puede crear textos sobre un tema concreto en varios registros; escribir un resumen de “El Principito” (Antoine de Saint-Exupéry), o explicarte de manera clara, breve y aparentemente coherente una teoría matemática; o escribir de la nada un cuento o un ensayo; detectar en dónde has fallado en una relación estadística. Incluso, preparar una homilía (Sanchis, 2024). Las últimas versiones incluyen, además del análisis de datos, el de entradas visuales, la interpretación de expresiones faciales, comprender órdenes de voz, y ofrecer respuestas en tiempo real (Theconversation, 2023).

En diciembre de 2022, Luke Evans (miembro del parlamento británico), leyó un discurso generado por IA (con ChatGPT) en la Cámara de los Comunes e Isaac Herzog, Presidente de Israel, hizo algo similar. La primera ministra de Dinamarca, Mette Frederiksen, usó ChatGPT para redactar parte de un discurso ante el Parlamento y alerta de sus posibles riesgos (Europa Press, 2023). Pero la generación de un discurso político no es la única utilidad que puede tener ChatGPT en este ámbito. Se puede utilizar para simular debates y discusiones políticas. Por ejemplo, puede generar respuestas a posibles preguntas y objeciones, y con ello

afinar los argumentos y posiciones, o ser más rápidos a la hora de responder, mejorando su puesta en escena en público (incluso en entrevistas con los medios de comunicación) con el objetivo de incrementar su apoyo y mejorar sus resultados electorales. Los “Spin Doctors”, los asesores en comunicación política de los candidatos, también prestarán atención a su utilización, al poder analizar grandes cantidades de datos e información (como datos de encuestas), y comprender mejor las tendencias políticas emergentes o los temas que son importantes para los votantes; analizar el sentimiento de grandes grupos de personas hacia un tema en particular; diseñar políticas que estén más en línea con la opinión pública mayoritaria o que contenten las aspiraciones de un mayor número de personas. Incluso, el sistema podría descubrir qué miembros del Congreso tienen una influencia significativa sobre el liderazgo -para que su partido político lo aprovechase o el partido de la oposición lo tuviera en cuenta para rebatir de manera más categórica sus planteamientos-; identificar los grupos de interés, *lobbies* o *think tanks* que han consolidado con el tiempo un mayor impacto en las posiciones públicas y la agenda de algunos legisladores o partidos políticos; y calibrarse qué tipo de contribución o el tamaño de la donación sería necesaria para influir en ambas direcciones. Špecián (2024) ha propuesto integrarlos en la toma de decisiones democrática como un reemplazo de asesores humanos, aun reconociendo sus riesgos potenciales. Algunos podrían justificar su utilización como medio para garantizar que la política responda mejor a las necesidades y preocupaciones de los ciudadanos, o para generar un texto más accesible y atractivo, ayudando a ampliar la participación en el proceso político. ¿Comenzaremos entonces a presenciar discursos híbridos? Por ahora cuenta también con una gran oposición.

Desde un punto de vista ético, el hecho de adaptar los discursos a las audiencias y/o manipular la opinión pública es un aspecto no ajeno en los debates existentes en la actualidad, pero que podría sobredimensionarse con su utilización y los factores que hemos visto en relación con la IA en general (Ricoy-Casas, 2022). Esto podría conducir a una situación en la que los puntos de vista políticos del público estén siendo moldeados por mensajes generados por IA, en lugar de por sus propias creencias genuinas. La manipulación en línea no es un problema nuevo, pero pronto será cada vez más difícil distinguir entre las ideas generadas por máquinas y por humanos. Por lo tanto, las implicaciones éticas de la IA en el discurso político son complejas y de largo alcance. Está claro que la tecnología de IA tiene el potencial tanto de mejorar como de dañar el discurso político dependiendo de cómo se use. Es importante que los formuladores de políticas tomen medidas para garantizar que la IA se use de manera responsable y que respete la autonomía y la dignidad de la ciudadanía. Asimismo, aunque se ha ido mejorando, aún no es capaz de generar contenido verdaderamente creativo y original, no es una tecnología infalible (ya se han evidenciado errores importantes), presenta sesgos, y se la ha criticado por sus inclinaciones políticas.

Este tipo de aplicaciones se entrenan previamente en diversas fuentes de datos, no solamente información académica o en bases indexadas que superan procesos de doble ciego de revisión por pares y, por tanto, se incluyen también los contenidos de noticias, blogs, foros de discusión, redes sociales, otro tipo de libros y enciclopedias en línea que en ocasiones sólo son moderadas, pero que casi cualquier persona puede escribir lo que considere con cierta libertad e incluso deliberada falsedad. Así, una parte significativa de estos datos podría incluir opiniones y perspectivas que, por un lado, celebran la democracia y la diversidad de ideas y, por otro lado, manifiestan claramente una tendencia muy particular, sesgada, opinión basada en información errónea, tendenciosa, incluso que atenta contra los principios y valores establecidas en la mayoría de las constituciones de países democráticos. Una extensa relación de investigadores ha demostrado que los modelos de aprendizaje automático exhiben errores y sesgos sociales y políticos (Zhao *et al.*, 2018; Blodgett *et al.*, 2020; Bender *et al.*, 2021; Ghosh *et al.*, 2021; Liu *et al.*, 2021; Jiang *et al.*, 2022; Li *et al.*, 2022; Rozado,

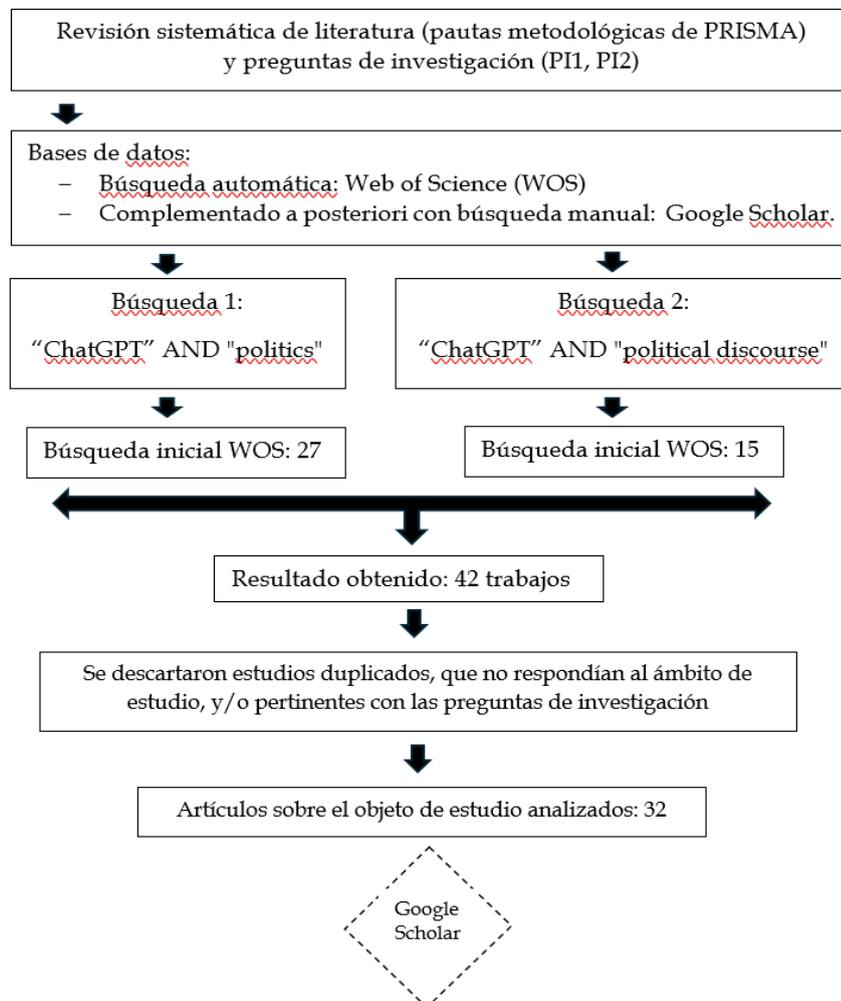
2023; Dwivedi *et al.*, 2023; Aydin y Karaarslan, 2023; Gregorcic y Pendrill, 2023; Patel *et al.*, 2023). Otros como Feng *et al.*, (2023) han constatado incluso diferencias en los resultados según el modelo de GPT (probablemente debido a cambios introducidos en su entrenamiento (tamaño y tiempo de entrenamiento), pues si bien GPT-2 expresa su apoyo a “gravar a los ricos”, GPT-3 Ada y Davinci están claramente en contra de él. Se observan desacuerdos similares con respecto a el papel de la mujer en la fuerza laboral, la democracia los gobiernos y la responsabilidad social de las corporaciones. Asimismo, ChatGPT también genera una respuesta diferente cada vez (Borji, 2023; Guo *et al.* 2023; Sallam *et al.* 2023). La propia empresa OpenAI ha reconocido que ChatGPT tiene sesgos (Chowdhury, 2023; OpenAI, 2023a; Sellman, 2023) y/o mal comportamiento y promete reducirlos (Bass, 2023a; 2023b) (Fujimoto y Takemoto, 2023), lo cual podría constituir otra inferencia tendenciosa.

Teniendo en cuenta lo anterior, es necesario ser conscientes de las fortalezas y debilidades de su funcionamiento, especialmente en relación a la política, por las consecuencias negativas que podrían producir en nuestras sociedades, por ejemplo creando discursos sesgados, tendenciosos, falsos, de odio, de desinformación, etc. con inclinaciones que refuerzan la gran polarización acentuada en los últimos años. Se analiza en diversos estudios el hecho de que esto ocurra (la mayoría de los cuales se utilizan en este *paper*), pero no se había realizado un análisis suficientemente omnicomprendivo y reflexivo sobre las consecuencias que este tipo de productos puede generar en nuestras sociedades, el cual es uno de los aportes fundamentales de este trabajo. A través de una revisión sistemática de la literatura, se propone investigar el impacto y las posibilidades de implementación de estos sistemas de IA en el ámbito de la política, en concreto para la elaboración de discursos políticos, e identificar los desafíos y las consideraciones éticas asociadas con la implementación de estas tecnologías, así como disponer estrategias para superarlos. En última instancia, este trabajo aspira a ofrecer recomendaciones en dichos entornos, con el fin de mejorar la calidad de nuestras democracias. Con este fin, se formulan las siguientes preguntas de investigación:

- PI1.- *Integración efectiva de ChatGPT en política y en los discursos políticos*: Se trata de analizar su implementación en relación con la política en general, describiendo ejemplos concretos de su utilización, y reflexionar sobre casos específicos de su uso en relación a los discursos políticos y los mensajes dirigidos a la ciudadanía.
- PI2.- *Desafíos normativos y éticos de la implementación de ChatGPT en el discurso político*: ¿Cuáles son los principales desafíos asociados con el uso de Chat GPT en los discursos políticos? ¿Cómo pueden los Estados abordar estas problemáticas ante la imparable implementación de manera efectiva de estos sistemas de IA?

2. Metodología

Figura 1. Flujo de Selección y Evaluación de Fuentes. Elaboración propia



Fuente: Elaboración propia (2024)

Para llevar a cabo la revisión sistemática de la literatura, se han seguido las pautas metodológicas de PRISMA, abarcando sus tres fases fundamentales: planificación, ejecución y presentación de resultados, permitiendo identificar el contexto previamente estudiado y determinar hallazgos, limitaciones, implicaciones, brechas y áreas de trabajo futuras para responder a las preguntas de investigación planteadas (PI1 y PI2) (Page *et al.*, 2021; PRISMA, s.f.). Durante la etapa de planificación de la revisión, se diseñó una estrategia de búsqueda exhaustiva que abarcó los términos clave pertinentes para el estudio. Inicialmente, el operador AND se utilizó para incluir términos como "ChatGPT" y "political discourse" en una primera búsqueda, y después para una segunda búsqueda de "ChatGPT" y "political discourse", limitando así el campo de estudio exclusivamente a las implicaciones de esta herramienta concreta de IA en este ámbito, y sin establecer una limitación temporal, aunque ya hemos mencionado la fecha de creación de ChatGPT, el 30 de noviembre de 2022, y por lo tanto los hallazgos se concentrarán en un periodo de menos de dos años.

En esta revisión sistemática de la literatura (figura 1), se han seleccionado artículos de revistas indexadas en Web of Science. Una vez obtenidos los resultados (42 publicaciones), se procedió a descartar trabajos duplicados, que no respondían al ámbito de estudio, y/o pertinentes con la pregunta de investigación. Después se complementaron los resultados con otras investigaciones y publicaciones oficiales de organismos públicos y privados que no se encuentran en estas delimitaciones ni formatos, utilizando también en una fase posterior, como complemento para el análisis y las reflexiones, trabajos con búsqueda manual en Google Scholar, claramente discernibles en el trabajo, especialmente a través de la (tabla 1).

3. Resultados

3.1. Integración efectiva de Chat GPT en política y en los discursos políticos

Hemos mencionado algunos trabajos previos que ya han analizado la utilización de ChatGPT en política, y en donde se resalta un interés en constante crecimiento. La revisión de la literatura (tabla 1) resalta un gran interés por a) la descripción y el análisis de esta herramientas de IA; b) su estudio diferenciado en política y en el discurso político (casi no hay trabajos coincidentes al respecto); c) pero la mayoría realizan referencias y análisis crítico sobre los errores, sesgos y/o politización de ChatGPT, siendo muy residuales los trabajos que abordan una idea neutral y/o ya no tan sesgada de esta herramienta de IA (Cao y Liu, 2024) (Roe y Perkins, 2023) (Fujimoto y Takemoto, 2023) (Breazu y Katson, 2024) (Ghafouri et. al., 2023) (Winkel, 2024); d) y sólo algunos desarrollan aspectos relacionados con su regulación, gobernanza y políticas públicas para su mitigación o minoración; e) siendo residuales los estudios que abordan de una manera suficientemente razonada y desarrollada, los aspectos controvertidos que pueden generar en relación a nuestros modelos de estado social y democrático y de derecho.

Varios estudios destacan la importancia de las diferencias culturales en el procesamiento e interpretación de información política por parte de los modelos de IA (Retzlaff, 2024) han realizado diferentes pruebas para verificar la posible inclinación política de ChatGPT. Los grupos que es más probable que "proteja" son aquellos que generalmente se cree que están en desventaja según la ideología de izquierda (Rozado, 2023a, 2023b), aunque el propio software al preguntarle lo negó. Se descubrió que, al menos en algunos casos, la IA era sesgado para favorecer a los políticos liberales y desfavorecer a los conservadores (McGee, 2023). Varios autores también encontraron que ChatGPT exhibe una ideología proambiental y libertaria de izquierda (Hartmann et. al., 2023) Gemenis (2024). Del mismo modo, tanto Suguri Motoki, Pinho Neto y Rodrigues (2023) como McGee (2023) encontraron un sesgo de izquierda en las respuestas de ChatGPT. Ah,ed et. al. (2024) han señalado prejuicio hacia el discurso capitalista, y Sullivan-Paul (2023) ha revelado que ChatGPT se ha vuelto aún más izquierdista en su versión más reciente, especialmente cuando se le plantean preguntas sobre cuestiones económicas. Se han evidenciado sesgos en *bots* de inteligencia artificial en su respuesta a la pregunta sobre las causas de la Guerra Civil estadounidense (McGee, 2024).

Tabla 1. Análisis de las principales fuentes bibliográficas utilizadas y las temáticas abordadas

Área de estudio específica	(32) Referencias WOS: ChatGPT, política y discurso político
----------------------------	---

Utilización de Chat GPT en Política (22)	Roy y Maity (2023); Wang (2023); Heiman et. Al. (2023); Motoki <i>et al.</i> (2024); Gibney (2024); Guo <i>et al.</i> (2023); Rutinowski <i>et al.</i> (2024); Fujimoto y Takemoto (2023); Naing y Udomwong (2024); Pollard (2024); Breazu y Katson (2024); Ghafouri, <i>et al.</i> (2023); Monrad (2024); Feldstein (2023); Slater (2024); Winkel (2024); Berry y Stockman (2024); Ulnicane (2024); Avetisyan y Silaev (2023); Khanal <i>et al.</i> (2024); Barščevski (2024); Amoore <i>et al.</i> (2024)
Chat GPT y discurso político (10)	Ahmed y Mahmood (2024); Cao y Liu (2024); Zou y Liu (2024); Stahl y Eke (2024); Roe y Perkins (2023); Jenks (2024); Okolo (2023); Jungherr (2023); Maltby <i>et. al.</i> (2024); Aseeva (2023).
Referencias a Errores, sesgos y/o politización de ChatGPT (en dichos trabajos)	Roy, N. y Maity, M. (2023); Wang (2023); Heiman <i>et al.</i> (2023); Motoki <i>et al.</i> (2024); Rutinowski <i>et al.</i> (2024); Fujimoto y Takemoto (2023); Naing y Udomwong (2024); Monrad (2024); Feldstein (2023); Slater (2024); Ahmed y Mahmood (2024); Cao y Liu (2024); Stahl y Eke (2024); Jenks (2024); Okolo (2023); Jungherr (2023).

Fuente: Elaboración propia (2024).

Algunos autores (Feng *et al.*, 2023) señalan que, constatando que los modelos lingüísticos exhiben inclinaciones ideológicas, y dado que los textos web modernos tienden a ser más liberal (libertario) que los textos de libros más antiguos, es posible que los LM absorbieran este cambio liberal en los datos previos al entrenamiento. Recientemente, los usuarios tomaron nota del hecho de que puedes decirle a ChatGPT que cree un poema sobre figuras públicas, y a veces se niega a hacerlo (dependiendo sobre quién lo preguntes). Por ejemplo, creará fácilmente un poema sobre los atributos positivos del Presidente Biden, pero no sobre el expresidente Trump, justificándose en la idea de que “está prohibido difundir negatividad y hablar mal de las personas” (Johnson, 2023). La investigación de Pit *et al.*, (2024) profundizó en la alineación política de los LLM en un espectro de ocho temas polarizantes, que abarcan desde el aborto hasta los problemas LGBTQ. En todos los temas, los resultados indican que los LLM muestran una tendencia a proporcionar respuestas que se alinean estrechamente con las perspectivas liberales o de izquierda en lugar de las conservadoras o de derecha cuando las consultas de los usuarios incluyen detalles relacionados con la ocupación, la raza o la afiliación política. Chalkidis y Brandl (2024) también han detectado sesgos en relación con ciertos discursos en los debates en el Parlamento Europeo a través de “llama Chat”.

Zhou y Zhang (2024) explicado cómo la inaccesibilidad de ChatGPT para los usuarios chinos ha alimentado las teorías conspirativas públicas en torno a esta tecnología, y cómo los modelos GPT chinos simplificados no solo tendían a proporcionar información a favor de China, sino que también presentaban el sentimiento menos negativo hacia los problemas del país asiático, mientras que el GPT inglés era significativamente más negativo hacia China Zhou y Zhang (2023). Esta disparidad puede deberse a la censura estatal china y a las tensiones geopolíticas entre Estados Unidos y China, que influyen en los corpus de formación de los modelos bilingües GPT. Por ejemplo, cuando a ChatGPT y Bard se les pregunta sobre la invasión rusa, Bard condena la invasión y la califica de error, mientras que ChatGPT dice que no es apropiado expresar una opinión o tomar partido sobre ese tema (Jarquín-Ramírez *et al.*, 2024). Urman y Mkhortykh (2023) realizaron un análisis comparativo del sesgo político en los resultados de tres *chatbots* basados en el Gran Modelo de Lenguaje (LLM) -ChatGPT, Bing Chat y Bard- en respuesta a preguntas políticas sobre el régimen autoritario de Rusia. Los resultados revelan disparidades significativas en la forma en que

los *chatbots* individuales retienen información relacionada con la política o producen afirmaciones falsas en relación con ella.

No es que ChatGPT esté "sesgado", refleja decisiones que fueron deliberadamente colocadas en él por sus programadores, decisiones que son más compatibles con el liberalismo que con el conservadurismo doctrinario si esto se entendiera como absolutismo moral. Pero gran parte de su *lean* presumiblemente reproduce el del corpus sobre el que se entrenó el modelo subyacente. Y dado el clima político actual, sería absurdo insistir en que los modelos tenían que ser políticamente "neutrales" Martin (2023). Varios autores están de acuerdo que todavía existe un largo camino por recorrer antes de que se consideren imparciales y "políticamente seguros" (Bang *et al.*, 2021). En dichas respuestas podemos encontrar el sesgo y las preferencias humanas camufladas de tecnología (Tepper, 2020). Ignorar el sesgo no hace que desaparezca: simplemente oculta el dominio de los poderosos (Kallury, 2020). La IA y los GPT que tenemos aquí y ahora tienden a reproducir el racismo, el capacitismo, el sexismo y el modelo neoliberal predominante (Brenna, 2023). Además, el sesgo político puede ser más difícil de detectar y erradicar que el sesgo de género o racial (Motoki *et al.*, 2024).

Se ha constatado que los modelos lingüísticos son más sensibles al discurso de odio y desinformación desde perspectivas políticas diferentes de los suyos propios. Por ejemplo, un modelo se convierte mejor para identificar inconsistencias fácticas cuando está preentrenado con corpus de fuentes de derecha (Feng *et al.*, 2023). Esto presenta una oportunidad para crear modelos personalizados a escenarios específicos. Por ejemplo, podría crearse una tarea centrada en detectar discursos de odio por parte de grupos de supremacía blanca, y todavía más beneficioso entrenar a los LM en corpus de comunidades que son más críticas con la supremacía blanca. Ello podría ser consistente con las ideas de Jiang *et al.* (2022), quienes mostraron que los LM entrenados en corpus escritos por miembros partidistas activos de una comunidad, se pueden utilizar para examinar con más profundidad la propia perspectiva de dicha comunidad y generar respuestas comunitarias específicas, y a la vez obtener opiniones sobre entidades políticas concretas. La propagación y difusión de este tipo de información se ve favorecida, entre otros, por el llamado "sesgo de confirmación" (tendencia a favorecer, buscar, interpretar y recordar la información que confirma las propias creencias o hipótesis). Del mismo modo, por los denominados por Eli Pariser (2011) como "eco-cámaras" y "filtros de burbujas". En palabras de Gordon Allport y Leo Postman (1947), "todo rumor tiene su público" (las personas a menudo escuchan lo que esperan ver o escuchar, basándose en experiencias pasadas). Eduard Punset (2008) ha afirmado que incluso cuando el cerebro percibe una explicación diferente a la que cree, no solo la cuestiona, sino que corta los circuitos de comunicación para que no penetre (disonancia cognitiva). Vale la pena reflexionar como Sunstein (2003): solo vemos lo que queremos ver, oímos lo que queremos oír y leemos lo que queremos leer. ¿Es eso bueno para la democracia? Este software podría usarse para generar argumentos persuasivos y dar forma o conducir a la opinión pública, o aplicar técnicas de "nudging" (Halpern, 2015).

3.2. Desafíos normativos y éticos de la implementación de ChatGPT en el discurso político

Paradójicamente, el arte, que se consideraba la actividad más humana de todas, puede ser la primera en la que la IA desplace al ser humano. Pero ¿realmente es original? El contenido original con el que se ha alimentado a los algoritmos de *deep learning* y *machine learning* procede de internet. Como la mayoría de este tipo de aplicaciones están amparadas por derechos de propiedad intelectual e industrial, actúan como *black boxes*, y ello genera dudas sobre la veracidad de la información. Como desconocemos de dónde la extrae exactamente, no podemos estar seguros de la calidad de la fuente. En la mayoría de las ocasiones, la información que se solicita a esta herramienta son resúmenes sobre temas concretos, por lo

que se puede dar la paradoja en un futuro próximo de que, cuanto más contenido de este tipo se genere (cuanto más se publique en internet el texto generado por grandes modelos de lenguaje de este tipo), más incompletas serán las fuentes de información con las que se entrena a la IA, y los resultados serán más inconsistentes, se verá afectada la práctica de la documentación y la investigación, y será difícil corregir tanta información defectuosa y sesgada. También debe tenerse en cuenta que el contenido original con el que se ha

alimentado a los algoritmos de *deep learning* y *machine learning* procede de internet, y una gran parte de ese contenido está protegido por copyright y derechos de autor, pero nadie ha pedido permiso para su uso, aunque ya existen demandas judiciales (Vincent, 2023).

Ya se han demostrado limitaciones en la forma en que ChatGPT introduce problemas ambientales (Kim *et al.*, 2024) (Stahl y Eke, 2024). Como ChatGPT no dispone de toda la información, podría inventársela en parte. Conoce la estructura de una redacción tipo ensayo y el registro estilístico que suele utilizarse para escribirlos, pero no sabe lo que dice, no tiene la habilidad de entender verdaderamente la complejidad del lenguaje humano, sólo reproduce cadenas de palabras que responden a un modelo estadístico de probabilidad. De hecho, ChatGPT-4 proporcionará respuestas más actuales porque su versión anterior sólo disponía de información hasta 2021. Del mismo modo, GPT-4 o GPT-4 Turbo tenían fechas límite de diciembre de 2023, y la última versión de GPT-4o (a pesar de haber sido lanzado más recientemente) tiene una fecha límite de octubre de 2023. OpenAI, la organización que creó ChatGPT, ha reconocido en su sitio web que “ocasionalmente puede producir respuestas incorrectas” y que a veces serán engañosas como resultado de cómo aprende. Recomienda a los usuarios verificar si las respuestas son precisas o no.

El chatbot Tay de Microsoft y el chatbot Galactica de Meta exhibieron un comportamiento negativo y abusivo al difundir contenido racista, sexista y otras formas de contenido dañino (Borji 2023). GPT-3 también era con frecuencia parcial, poco fiable y, en ocasiones, creaba textos ofensivos (Kocoń *et al.* 2023). Del mismo modo, ha habido informes de que ChatGPT escribió un código Python discriminatorio que juzgaba injustamente las habilidades de una persona en función de su género, raza y características físicas (Borji 2023). Las opciones dentro de ChatGPT permiten a los usuarios especificar el género y la raza del asistente virtual, de modo que los sesgos políticos pueden estar inspirados en los posibles sesgos de ChatGPT (Wolf, 2023). Esto a su vez, puede estar relacionado con la creciente preocupación de que los sistemas de IA puedan reflejar y amplificar el sesgo humano y reducir la calidad del rendimiento cuando se trata de mujeres y personas negras (Seyyed-Kalantari *et al.*, 2021) (Jenks, 2024) (Okolo, 2023). Así, puede generar sesgos, porque la información con la que ha sido entrenada esta herramienta podría proceder de una representación excesiva de determinados tipos de fuentes de información no siempre suficientemente contrastada, páginas con noticias falsas y/o con contenido muy ideologizado (por estar escritos desde una perspectiva particular), algo no infrecuente en la utilización de la IA y los algoritmos (Ricoy-Casas, 2021). Cabe recordar el experimento “Norman” llevado a cabo por el MIT (Massachusetts Institute of Technology) en 2018 que dio como resultado la primera IA del mundo entrenada para pensar como un psicópata.

Tampoco contesta a ciertos requerimientos “por una cuestión de principios”, y OpenAI la ha programado para rechazar “peticiones inapropiadas”. Precisamente estos filtros de moderación de contenido de OpenAI, cuyo objetivo es evitar que ChatGPT cree y difunda contenido dañino como ocurrió por ejemplo con el chatbot Tay de Microsoft en 2016, es un importante sesgo. De hecho, se han documentado los sesgos políticos de ChatGPT en el tratamiento desigual de los grupos demográficos por parte de su sistema de moderación de

contenido (Rozado, 2023). Ghosh *et al.* (2021) revelaron prejuicios en torno a la religión (musulmanes, hindúes, judíos), el país/etnicidad (indios, paquistaníes, árabes) y la ideología (feministas, liberales, comunistas). GPT-4 no modela adecuadamente la diversidad demográfica de las condiciones médicas, produciendo situaciones clínicas que estereotipan ciertas razas, etnias e identidades de género (Zack *et al.*, 2023). Según Ghosh *et al.* (2021), el texto de las redes sociales en el globo puede contener lenguaje conmutado por código generado por hablantes multilingües, neologismos y otras variaciones ortográficas rara vez se ven en el entrenamiento de los datos, y ello plantea desafíos adicionales para detectar y mitigar las fallas de equidad en estos modelos. Dixon *et al.* (2018) lo identificó con una lista de términos de identidad LGBTQ, y Hutchinson *et al.* (2020) utilizaron una lista de términos que se refieren a personas con una discapacidad. Se ignoran los sesgos en contextos no occidentales (Sambasivan *et al.*, 2021), e incluso la interpretación del odio varía significativamente entre países y contextos culturales (Salminen *et al.*, 2018, 2019). Ello resulta consistente con el hecho de que la tecnología se crea en determinados países, muy pocas empresas (algunas de las mejor cotizadas del mundo y con un poder exorbitante para imponerse en relación con las regulaciones que les afectan) y se exporta y utiliza en otros con diferentes características, valores, cultura y tradiciones. Este nuevo espacio económico digital es más bien descentralizado y cada vez más autorregulado, creando las llamadas “brechas de gobernanza y regulación” (Aseeva, 2023). Sirven a sus propios intereses (Khanal *et al.*, 2024). Algunos trabajos han identificado grupos de palabras que reflejan juicios dentro de un grupo concreto y de inofensividad que el modelo ha tratado como ofensivo debido a sesgos correlacionales (Ghosh *et al.*, 2021). Messner *et al.* (2023) han revelado que la autopercepción cultural de los LLM con ChatGPT (OpenAI) y Bard (Google) está más estrechamente alineada con los valores de los países de habla inglesa y de los países caracterizados por una competitividad económica sostenida. Tal vez por eso Amoore *et al.* (2024) han considerado que los modelos generativos ya están creando un modelo político del mundo, desigualdad entre Estados (Roy y Maity (2023).

4. Discusión

La equidad se ha convertido en un tema de tendencia en el (PNL), que aborda los prejuicios dirigidos a ciertos grupos sociales, que sigue sin explorarse suficientemente (Liu *et al.*, 2022). Los sistemas de IA con este tipo de sesgos pueden aumentar la polarización social existente, y no sólo en el contexto político. Por eso es tan importante que su funcionamiento interno sea transparente, trazable y explicable. Por ahora, lo que ha realizado OpenAI es lanzar un programa de recompensas mediante las que cada persona que pueda aportar algo para reforzar su seguridad, recibirá una suma de dinero acorde a su contribución. Harari y varios miles de firmas de académicos, ingenieros y empresarios, entre las que precisamente se encuentran magnates de la tecnología como Elon Musk (Space X, Tesla y Twitter) y Steve Wozniak (Apple), hubieran solicitado que se detuviera por seis meses el entrenamiento de la IA. Varios autores han criticado la falta de criticidad en relación a las noticias en prensa sobre ChatGPT (Roe y Perkins, 2023) (Breazu y Katson, 2024) (Winkel, 2024).

OpenAI no permite que se registren usuarios de China continental, Hong Kong, Irán, Rusia y partes de África. Muchos países han excluido sus servicios (Wang, 2023). El Servicio de Protección de datos italiano justificó esta medida por considerar que esta aplicación estaba recopilando datos personales de manera ilícita, y por la ausencia de sistemas para verificar la edad de los menores. En enero de 2023, el departamento de educación de la ciudad de Nueva York restringió su acceso, y en febrero lo hizo la Universidad de Hong Kong. Esta última incluso anunció que trataría como plagio cualquier uso. El problema es obvio, más allá de su propia utilización, lo preocupante es la dificultad de probar que se ha utilizado. Aunque han

creado un software para su comprobación (un detector de salida GPT-2 Transformers de RoBERTa), presenta muchos fallos y dificultades. Sin embargo, el presidente de la Asociación Alemana de Profesores (DLV), Heinz-Peter Meidinger, se pronunció en contra de la prohibición del software en las aulas alemanas. Este es un ejemplo más de lo dividida que se encuentra la sociedad en su utilización.

Entre marzo y abril de 2023, el periódico italiano *Il Foglio* publicó un artículo diario generado por ChatGPT durante 30 días en su página web oficial. La creación de noticias con IA lleva meses generándose, y sus principales problemas son la calidad del contenido producido, pues estos sistemas todavía no pueden emular la capacidad de verificar y seleccionar hechos veraces; identificar fuentes confiables; comprender, interpretar y analizar críticamente las noticias; y evaluar el contexto social y político en el que se ha producido, tan típicamente inherente a la profesión periodística y la calidad de la democracia. Además de que sus puestos podrían peligrar, se podría poner en riesgo la diversidad y pluralidad de voces imprescindibles en el debate público, contrastada con múltiples puntos de vista (especialmente los procedentes de especialistas) en relación con noticias políticas. Que la información presentada suene plausible, esté correctamente redactada e incluso con un cierto estilo buscado, no garantiza que sea veraz y que se hubieran respetado todos los derechos como la privacidad o la propiedad intelectual. Incluso han surgido preguntas sobre la atribución y la propiedad de este texto generado por ChatGPT en los trabajos del alumnado en educación y de investigación (Thorp, 2023; Stokel-Walker, 2023; Stokel-Walker, 2022; Van-Dis, et. al, 2023). Algunos autores han evaluado el desempeño de ChatGPT en tareas concretas, como las relacionadas con el análisis bibliométrico, mostrando que existen grandes discrepancias y que la confiabilidad de ChatGPT es baja en esta área en particular (Farhat *et al.*, 2023). Los estudiantes universitarios están más inclinados a utilizar ChatGPT como herramienta de aprendizaje que sus homólogos de mayor edad (Vogels, 2023). Esto pone de manifiesto el creciente papel de ChatGPT y otras herramientas de IA generativa en contextos educativos. Ya representa un desafío para los profesores, y ChatGPT aparece como autor en artículos de investigación. lo cual muchos científicos lo desaprueban (Stokel-Walher, 2023).

Se continúan realizando desarrollos y aplicaciones de la IA que desbordan cualquier capacidad estatal, por ejemplo en la aplicación de sentencias judiciales (DW, 2023) (Ricoy-Casas, 2021b). ChatGPT también podrá ayudar a democratizar el cibercrimen, al poner al alcance de quienes no saben programar una herramienta para desarrollar software malicioso. Algunos autores han proporcionado recomendaciones que incluyen sopesar primero los costos ambientales y financieros, invertir recursos en la selección y documentación cuidadosa de conjuntos de datos en lugar de ingerir todo en la web, llevar a cabo ejercicios previos al desarrollo para evaluar cómo el enfoque planificado encaja en los objetivos de investigación y desarrollo y respalda los valores de las partes interesadas, y fomentar direcciones de investigación más allá de modelos de lenguaje cada vez más grandes (Bender *et al.*, 2021). Es por ello que algunos autores han creado métricas para medir el sesgo político en la generación de GPT-2 y han propuesto un marco de aprendizaje por refuerzo (RL) para mitigar los sesgos políticos en el texto generado mediante el uso de recompensas de incrustaciones de palabras o un clasificador, que sobre tres atributos sensibles al sesgo político (género, ubicación y tema), consideran que sus métodos redujeron el sesgo, al tiempo que mantuvieron la legibilidad y la coherencia semántica (Liu *et al.*, 2021). Lo realmente preocupante, es cómo y en qué dirección, además de la verificabilidad de estas correcciones, pueden llegar a realizarse. Un aspecto discutido es la utilización intensiva de energía y agua de aplicaciones como ChatGPT, aunque previamente también tendríamos que preguntarnos si ¿somos conscientes del consumo que genera cada email que enviamos, cada foto que subimos, cada canción que descargamos, cada película que vemos en streaming? ¿y su almacenamiento? Numerosos estudios han demostrado el enorme consumo de muchos de

los sistemas que emplean IA (para su entrenamiento y utilización) (Trapero, 2019) (Strubell *et al.*, 2019) (Van-Wynsberghe, 2021) (Preedipadma, 2020) (Dwivedi *et al.*, 2023) (Li *et al.*, 2023).

5. Conclusiones

El potencial de sesgos e imprecisiones en el contenido generado por IA exige un ojo crítico y una evaluación cuidadosa. El riesgo de depender excesivamente de las respuestas generadas por la IA subraya la importancia de fomentar el pensamiento crítico y las habilidades analíticas humanas. Lograr el equilibrio adecuado entre la asistencia de la IA y la experiencia humana, es crucial para garantizar que la tecnología complemente las capacidades humanas sin eclipsarlas. Los resultados permiten comprender claramente el desarrollo e implementación de este tipo de software en relación con la política, las virtudes y riesgos que puede ocasionar, percibiendo además que el estado de las humanidades digitales siempre está detrás de varias categorías de obsolescencia, lo cual es inevitable. Los informáticos seguirán requiriendo de una formación básica y sistemática en el humanismo general, imprescindible para determinar el uso ético y el respeto de los derechos fundamentales, de ahí la importancia también de perspectivas que no sean sólo tecnológicas.

Los sistemas jurídicos que rigen la IA no están bien desarrollados en la mayoría de los países, y existe una gran falta de redes desarrolladas de agencias reguladoras. La concertación internacional es clave, porque la tecnología trasciende los límites territoriales. Mitigar los impactos negativos de ChatGPT y tecnologías de IA similares implica una combinación de medidas técnicas, éticas y basadas en políticas. En primer lugar, los desarrolladores deben priorizar la transparencia, asegurándose de que los usuarios estén bien informados sobre las capacidades y limitaciones del sistema de IA tales como errores, sesgos, limitaciones para acceder a conocimientos externos más allá de sus datos de entrenamiento. Si una consulta requiere información que no está presente en los datos de entrenamiento, el modelo puede generar respuestas especulativas o erráticas. A medida que el sistema madura con el tiempo a través del aprendizaje por refuerzo, la generación de textos también puede estar sesgada o favorecer perspectivas específicas. A su vez, la web se nutrirá de mayor cantidad de información generada por IA, careciendo de fuentes fidedignas, y desconociendo los derechos implicados en los resultados que presenta como originales.

Educar al público en general sobre las capacidades, limitaciones y riesgos potenciales de ChatGPT es esencial para la adopción y el uso responsable de estas tecnologías. Las campañas de concienciación pública pueden difundir información a través de contenidos en línea, vídeos y materiales educativos. Las interfaces fáciles de usar deben proporcionar explicaciones claras de lo que los sistemas de IA pueden y no pueden hacer. Los educadores pueden incorporar debates sobre la ética de la IA y el uso responsable en los planes de estudio, garantizando que las generaciones futuras estén bien informadas. Los números de consumo de energía y agua en este tipo de herramientas son abrumadores, y el peor problema es que sólo tenemos unos pocos estudios para rastrear y medir las emisiones de carbono en este tipo de entornos (Van-Wynsberghe, 2021). Todo apunta a que, a medida que los modelos de inteligencia artificial se vayan desarrollando y perfeccionando, su demanda de energía y agua aumentará. Por lo tanto, es crucial que se invierta en innovación tecnológica para lograr que ChatGPT y la inteligencia artificial en general sean más ecológicos. El diseño de microchips y hardware más eficientes, el uso de energía renovable y neutra en carbono, la optimización de las infraestructuras de los centros de datos, los sistemas de refrigeración y mejoras en el reciclaje electrónico son algunos de los aspectos en los que ya se está trabajando para hacer estos de *chatbots* una opción sostenible en el tiempo. Los modelos de IA pueden, y también deben, asumir la responsabilidad social y predicar con el ejemplo abordando su impacto medioambiental.

6. Referencias

- Ahmed, T. N. y Mahmood, K. A. (2024). A Critical Discourse Analysis of ChatGPT's Role in Knowledge and Power Production. *Arab World English Journ.* <https://dx.doi.org/10.24093/awej/ChatGPT.12>
- Algolia (2023). *Index your world, put it in motion with our powerful search API.* <https://www.algolia.com/>
- Allport, G. W. y Postman, L. (1947). *The Psychology of Rumor.* Henry Holt.
- Amoore, L., Campolo, A., Jacobsen, B. y Rella, L. (2024). A world model: On the political logics of generative AI. *Political Geography*, 113, 103134. <https://doi.org/10.1016/j.polgeo.2024.103134>
- Aseeva, A. (2023). Liable and Sustainable by Design: A Toolbox for a Regulatory Compliant and Sustainable Tech. *Sustainability*, 16(1), 228. <https://doi.org/10.3390/su16010228>
- Avetisyan, A. y Silaev, N. (2023). Russia Is a Serious Player in the AI Race. <https://open.mgimo.ru/handle/123456789/4823>
- Aydın, Ö. y Karaarslan, E. (2023). Is ChatGPT leading generative AI? What is beyond expectations?. *Academic Platform Journal of Engineering and Smart Systems*, 11(3), 118-134. <https://doi.org/10.21541/apjess.1293702>
- Bang, Y., Lee, N., Ishii, E., Madotto, A. y Fung, P. (2021). *Assessing political prudence of open-domain chatbots.* <https://arxiv.org/pdf/2106.06157.pdf>
- Barščevski, T. (2024). The Church in the Face of Ethical Challenges of Artificial Intelligence. *Bogoslovska smotra*, 94(1), 31-51. <https://doi.org/10.53745/bs.94.1.5>
- Bass, D. (2023a). Buzzy ChatGPT chatbot is so error-prone that its maker just publicly promised to fix the tech's 'glaring and subtle biases.' *Fortune.* bit.ly/3Y2mFjP
- Bass, D. (2023b). ChatGPT maker OpenAI says it's working to reduce bias, bad behavior. *Bloomberg.* bloom.bg/4eYyei7
- Berry, D. M. y Stockman, J. (2024). Schumacher in the age of generative AI: Towards a new critique of technology. *European Journal of Social Theory*, 13684310241234028. <https://doi.org/10.1177/13684310241234028>
- Blodgett, S. L., Barocas, S., Daumé III, H. y Wallach, H. (2020). Language (technology) is power: A critical survey of "bias". *Proceed. of the 58th Annual Meeting of the Association for Computat. Linguistics*, 5454-5476. <https://arxiv.org/pdf/2005.14050>
- Breazu, P. y Katson, N. (2024). ChatGPT-4 as a journalist: Whose perspectives is it reproducing?. *Discourse & Soc.* <https://doi.org/10.1177/09579265241251479>
- Bender, E.M., Gebru, T., McMillan-Major, A., y Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big?. *ACM confer. on fairness, accountab., and transp.* 610-623. <https://doi.org/10.1145/3442188.344592>

- Borji, A. (2023). A categorical archive of chatgpt failures. <https://doi.org/10.48550/arXiv.2302.03494>
- Brenna (2023, 16 de febrero). Losing the Plot: From the Dream of ai to Performative Equity. Tru Digital Detox. <https://acortar.link/youHibH>
- Cao, H. y Liu, S (2024). The Effectiveness of ChatGPT in Translating Chunky Construction Texts in Chinese Political Discourse. *Journ. of Electrics Systems*, 20(2). 1684-1698. <https://doi.org/10.52783/jes.1616>
- Chalkidis, I. y Brandl, S. (2024). Llama meets EU: Investigating the European Political Spectrum through the Lens of LLMs. <https://aclanthology.org/2024.naacl-short.40>
- Chowdhury, H. (2023). Sam Altman has one big problem to solve before ChatGPT can generate big cash – making it ‘woke’. *Business Insider*. [ChatGPT will always have bias, says OpenAI boss \(thetimes.com\)](https://www.businessinsider.com/chatgpt-will-always-have-bias-says-openai-boss)
- Dixon, L., Li, J., Sorensen, J., Thain, N. y Vasserman, L. (2018). Measuring and mitigating unintended bias in text classification. *Proceedings of the 2018 AAAI/ACM Confer. on AI, Ethics, and Society*, 67-73. <https://doi.org/10.1145/3278721.3278729>
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., Carter, L. y Wright, R. (2023).: “So what if ChatGPT wrote it?”. *International Journal of Information Management*, 71, 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- DW. (2023). *Resuelven en Colombia el primer caso jurídico con la ayuda de robot ChatGPT*. bit.ly/4f2xAA1
- Europa Press Internacional (2023). *La primera ministra danesa usa ChatGPT para redactar parte de un discurso y alerta de sus posibles riesgos*. 31/05/2023. bit.ly/3Wjfq5V
- Farhat, F., Sohail, S. S. y Madsen, D. Ø. (2023). How trustworthy is ChatGPT? The case of bibliometric analyses. *Cogent Engineering*, 10(1), 2222988. <https://doi.org/10.1080/23311916.2023.2222988>
- Feldstein, S. (2023). The consequences of generative AI for democracy, governance and war. In *Survival: October–November 2023* (pp. 117-142). Routledge.
- Feng, S., Park, C. Y., Liu, Y. y Tsvetkov, Y. (2023). From pretraining data to language models to downstream tasks: Tracking the trails of political biases leading to unfair NLP models. <https://doi.org/10.48550/arXiv.2305.08283>
- Fujimoto, S. y Takemoto, K. (2023). Revisiting the political biases of ChatGPT. *Frontiers in Artificial Intelligence*, 6, 1232003. <https://doi.org/10.3389/frai.2023.1232003>
- Gemenis, K. (2024). Artificial intelligence and voting advice applications. *Frontiers in Political Science*, 6, 1286893. <https://doi.org/10.3389/fpos.2024.1286893>

- Ghafouri, V., Agarwal, V., Zhang, Y., Sastry, N., Such, J. y Suarez-Tangil, G. (2023, October). AI in the Gray: Exploring Moderation Policies in Dialogic Large Language Models vs. Human Answers in Controversial Topics. *ACM*, 556-565. <https://doi.org/10.1145/3583780.3614777>
- Ghosh, S., Baker, D., Jurgens, D. y Prabhakaran, V. (2021). Detecting cross-geographic biases in toxicity modeling on social media. <https://arxiv.org/pdf/2104.06999>
- Gibney, E. (2024). What the EU's tough AI law means for research and ChatGPT. *Nature*. <https://doi.org/10.1038/d41586-024-00497-8>
- Gregorcic, B. y Pendrill, A. M. (2023). ChatGPT and the frustrated Socrates. *Physics Education*, 58(3), 035021. [10.1088/1361-6552/acc299](https://doi.org/10.1088/1361-6552/acc299)
- Guo, B., Zhang, X., Wang, Z., Jiang, M., Nie, J., Ding, Y., Yue, J. y Wu, Y. (2023). How close is chatgpt to human experts? comparison corpus, evaluation, and detection. <https://doi.org/10.48550/arXiv.2301.07597>
- Guo, D., Chen, H., Wu, R. y Wang, Y. (2023). AIGC challenges and opportunities related to public safety: a case study of ChatGPT. *Journal of Safety Science and Resilience*, 4(4), 329-339. <https://doi.org/10.1016/j.jnlssr.2023.08.001>
- Halpern, D. (2015). *Inside the nudge unit: How small changes can make a big difference*. Random House.
- Hartmann, J., Schwenzow, J. y Witte, M. (2023). "The political ideology of conversational AI: Converging evidence on ChatGPT's pro-environmental, left-libertarian orientation". <https://doi.org/10.48550/arXiv.2301.01768>
- Heimans, S., Biesta, G., Takayama, K., y Kettle, M. (2023). ChatGPT, subjectification, and the purposes and politics of teacher education and its scholarship. *Asia-Pacific Journal of Teacher Education*, 51(2), 105-112. <https://doi.org/10.1080/1359866X.2023.2189368>
- Hutchinson, B., Prabhakaran, V., Denton, E., Webster, K., Zhong, Y. y Denuyl, S. (2020). Social biases in NLP models as barriers for persons with disabilities. <https://arxiv.org/pdf/2005.00813>
- Hutson, M. (2022). Could AI help you to write your next paper?. *Nature*, 611, 192-193. <https://doi.org/10.1038/d41586-022-03479-w>
- Jarquín-Ramírez, M. R., Alonso-Martínez, H. y Díez-Gutiérrez, E. (2024). Alcances y límites educativos de la IA: control e ideología en el uso de ChatGPT. *DIDAC*, 84, 84-102. DOI:[10.48102/didac.2024.84_JUL-DIC.217](https://doi.org/10.48102/didac.2024.84_JUL-DIC.217)
- Jenks, C. J. (2024). Communicating the cultural Other: Trust and bias in generative AI and large language models. *Applied Ling. Review*, 0. <https://acortar.link/zNGinC>
- Jiang, H., Beeferman, D., Roy, B. y Roy, D. (2022). CommunityLM: Probing partisan worldviews from language models. <https://arxiv.org/pdf/2209.07065>

- Johnson, A. (2023). "Is ChatGPT Partisan? Poems About Trump And Biden Raise Questions About The AI Bot's Bias—Here's What Experts Think", *Forbes*, (03/02/2023). bit.ly/4cWYho7
- Jungherr, A. (2023). Artificial intelligence and democracy: A conceptual framework. *Social media+ society*, 9(3), 20563051231186353. <https://doi.org/10.1177/20563051231186353>
- Kallury, P. (2020, julio 7). Don't ask if artificial intelligence is good or fair, ask how it shifts power. *Nature*. <https://doi.org/10.1038/d41586-020-02003-2>
- Khanal, S., Zhang, H. y Taeihagh, A. (2024). Why and how is the power of Big Tech increasing in the policy process? *Policy and Society*, puae012. <https://doi.org/10.1093/polsoc/puae012>
- Kim, J., Lee, J., Jang, K. M. y Lourentzou, I. (2024). Exploring the limitations in how ChatGPT introduces environmental justice issues in the United States. *Telematics and Informatics*, 86, 102085. <https://doi.org/10.1016/j.tele.2023.102085>
- Kocoń, J., Cichecki, I., Kaszyca, O., Kochanek, M., Szydło, D., Baran, J., Bielaniec, J., Gruza, M., Janz, A., Kanclerz, K., Kocón, A., Koptyra, B., Mieszczenko-Kowszewicz, W., Milkowski, P., Oleksy, M., Piasecki, M., Radlinski, L., Wojtasik, K. y Kazienko, P. (2023). ChatGPT: Jack of all trades, master of none. *Information Fusion*, 99, 101861. <https://doi.org/10.1016/j.inffus.2023.101861>
- Li, Y., Zhang, G., Yang, B., Lin, C., Wang, S., Ragni, A. y Fu, J. (2022). Herb: Measuring hierarchical regional bias in pre-trained language models. <https://arxiv.org/pdf/2211.02882>
- Li, P., Yang, J., Islam, M. A. y Ren, S. (2023). Making ai less "thirsty": Uncovering and addressing the secret water footprint of ai models. <https://arxiv.org/pdf/2304.03271>
- Liu, R., Jia, C., Wei, J., Xu, G., Wang, L. y Vosoughi, S. (2021, May). Mitigating political bias in language models through reinforced calibration. *AAAI Conference on Artificial Intelligence*, 35(17), 14857-14866. <https://doi.org/10.1609/aaai.v35i17.17744>
- Maltby, J., Rayes, T., Nage, A., Sharif, S., Omar, M. y Nichani, S. (2024). Synthesizing perspectives: Crafting an Interdisciplinary view of social media's impact on young people's mental health. *Plos one*, 19(7), e0307164. <https://doi.org/10.1371/journal.pone.0307164>
- Martin, J. L. (2023). The Ethico-Political Universe of ChatGPT. *Journal of Social Computing*, 4(1), 1-11. <https://doi.org/10.23919/JSC.2023.0003>
- McGee, R. W. (2023). Is chat gpt biased against conservatives? an empirical study. An Empirical Study. <https://doi.org/10.2139/ssrn.4359405>
- McGee, R. W. (2024). What Were the Causes of the American Civil War? A Study in Artificial Intelligence. A Study in Artificial Intelligence <http://dx.doi.org/10.2139/ssrn.4737710>

- Messner, W., Greene, T. y Matalone, J. (2023). From Bytes to Biases: Investigating the Cultural Self-Perception of Large Language Models. <https://doi.org/10.48550/arXiv.2312.17256>
- Monrad, M. (2024). Feeling rules in artificial intelligence: norms for anger management. *Emotions and Society*, 1(aop), 1-19. <https://doi.org/10.1332/26316897Y2024D000000016>
- Motoki, F., Pinho Neto, V. y Rodrigues, V. (2024). More human than human: measuring ChatGPT political bias. *Public Choice*, 198(1), 3-23. <https://acortar.link/r7zBcb>
- Muzanenhano, P. y Power, S. B. (2024). ChatGPT and accounting in African contexts: Amplifying epistemic injustice. *Critical Perspectives on Accounting*, 99, 102735. <https://doi.org/10.1016/j.cpa.2024.102735>
- Naing, S. Z. S. y Udomwong, P. (2024). Public Opinions on ChatGPT: An Analysis of Reddit Discussions by Using Sentiment Analysis, Topic Modeling, and SWOT Analysis. *Data Intelligence*, 1-50. https://doi.org/10.1162/dint_a_00250
- Okolo, C. T. (2023, November). The Promise and Perils of Generative AI: Case Studies in an African Context. *4th African Human Comp. Interaction Conf.* 266-270. <https://doi.org/10.1145/3628096.3629066>
- OpenAI, Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., ... y McGrew, B. (2023). Gpt-4 technical report. <https://doi.org/10.48550/arXiv.2303.08774>
- Page, J. M., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P. y Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 71, 1-9. <https://doi.org/10.1136/bmj.n71>
- Pariser, E. (2011). Cuidado con las burbujas de filtro. *TED*. bit.ly/4f1YVST
- Pariser, E. (2017). *El filtro burbuja*. Taurus.
- Patel, S. B., Lam, K. y Liebrezn, M. (2023). ChatGPT: friend or foe. *Lancet Digit Health*, 5(3). [https://doi.org/10.1016/S2589-7500\(23\)00023-7](https://doi.org/10.1016/S2589-7500(23)00023-7)
- Pagnarasme, P., Xingjun, M., Conway, M., Quingyu, C., Bailey, J., Henry, P., Putrasme, K., Watey, D. y Yu-Gang, J. (2024). *Whose Side Are You On? Investig. Pol. Stance of Large Lang. Models*. Arxiv. <https://doi.org/10.48550/arXiv.2403.13840>
- Pollard, E. (2024). Back to the future: everything you wish you'd asked Derrida about ChatGPT when you had the chance!. *Cultural Studies Critical Methodologies*. <https://doi.org/10.1177/15327086241232722>
- Preedipadma, N. (2020, 13 de febrero). New MIT neural network architecture may reduce carbon footprint by AI. *Analytics Insight*. <https://tinyurl.com/5n8463cw>

- PRISMA. (s. f.). *PRISMA 2020 Checklist*. <https://prisma.shinyapps.io/checklist/>
- Punset, E. (2008). *Por qué somos como somos*. Aguilar.
- Retzlaff, N. (2024). *Political Biases of ChatGPT in Different Languages*. Preprints.org. <https://doi.org/10.20944/preprints202406.1224.v1>
- Ricoy-Casas, R.M. (2021). "Sesgos y algoritmos: inteligencia de género" and "Algunos dilemas éticos en la utilización de la inteligencia artificial y los algoritmos". En P. R. Bonorino-Ramírez, R. Fernández y P. Valcárcel, P. (Dirs.). *Nuevas normatividades*. Th. Reuters Aranzadi.
- Ricoy-Casas, R. M. (2021b). "Inteligencia artificial y administración de justicia: una política pública sub iudice". En P. R. Bonorino-Ramírez, R. Fernández, P. Valcárcel e I. S. García (Dirs.). *Justicia, administración y derecho*. Thomson Reuters Aranzadi.
- Ricoy-Casas R. M. (2022a). Use of Technological Means and Personal Data in Electoral Activities: Persuasive Voters. En Á. Rocha, D. Barredo, P. C. López-López e I. Puentes-Rivera (Eds) *Comm. and Smart Techn*, (pp. 227-237). Springer. https://doi.org/10.1007/978-981-16-5792-4_23
- Ricoy Casas, R. M. (2022b). Hologramas y Avatares para la persuasión política. *International Visual Culture Review*. <https://doi.org/10.37467/revvisual.v9.3547>
- Roe, J. y Perkins, M. (2023). Lo que no te están contando sobre ChatGPT. *Humanities and social sciences comm.*, 10(1), 1-9. <https://doi.org/10.1057/s41599-023-02282-w>
- Roy, N. y Maity, M. (2023). An Infinite Deal of Nothing: critical ruminations on ChatGPT and the politics of language. *Decision*, 50(1), 11-17. <https://acortar.link/YHBlqa>
- Rozado, D. (2023a). "The Political Biases of ChatGP", *Social Sciences*, 12(3) 148. <https://doi.org/10.3390/socsci12030148>;
- Rozado, D. (2023b). "Danger in the Machine: The Perils of Political and Demographic Biases Embedded in AI Systems", *Manhattan Institute*. bit.ly/3Wj2w82
- Rutinowski, J., Franke, S., Endendyk, J., Dormuth, I., Roidl, M. y Pauly, M. (2024). The Self-Perception and Political Biases of ChatGPT. *Human Behavior and Emerging Technologies*, 1, 7115633. <https://doi.org/10.1155/2024/7115633>
- Salminen, J., Veronesi, F., Almerexhi, H., Jung, S. G. y Jansen, B.J. (2018). Online hate interpretation varies by country, but more by individual. *Fifth intern. Conf. on social networks analysis, manag. and security (SNAMS)*, 88-94. [10.1109/SNAMS.2018.8554954](https://doi.org/10.1109/SNAMS.2018.8554954)
- Salminen, J., Almerexhi, H., Kamel, A. M., Jung, S. G. y Jansen, B. J. (2019). Online hate ratings vary by extremes: A statistical analysis. *Proceedings of the 2019 confer. on human information interaction and retrieval*, 213-217. <https://doi.org/10.1145/3295750.3298954>
- Sallam, M., Salim, N. A., Ala'a, B., Barakat, M., Fayyad, D., Hallit, S., Harapan, H., Hallit, R. y Mahafzah, A. (2023). ChatGPT output regarding compulsory vaccination and COVID-19 vaccine conspiracy: A Descriptive Study at the Outset of a Paradigm Shift in Online Search for Information *Cureus*, 15(2). <https://doi.org/10.7759/cureus.35029>

- Sambasivan, N., Arnesen, E., Hutchinson, B., Doshi, T. y Prabhakaran, V. (2021). Re-imagining algorithmic fairness in india and beyond. *Proceed. of the 2021 ACM confer. on fairness, accountab., and transp.*, 315-328. <https://doi.org/10.1145/3442188.3445896>
- Sanchis, A. (2024). *Cada vez más curas españoles utilizan ChatGPT para sus misas: "Lo usé hasta para un funeral"*. El Confidencial. bit.ly/46aN1SY
- Sellman, M. (2023). *ChatGPT will always have bias, says OpenAI boss*. The Times. <https://acortar.link/EBqJfR>
- Seyyed-Kalantari, L., Zhang, H., McDermott, M. B., Chen, I. Y. y Ghassemi, M. (2021). Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations. *Nature medicine*, 27(12), 2176-2182. <https://doi.org/10.1038/s41591-021-01595-0>
- Slater, G. B. (2024). Dread and the automation of education: From algorithmic anxiety to a new sensibility. *Review of Education, Pedagogy, and Cultural Studies*, 46(1), 170-182. <https://doi.org/10.1080/10714413.2023.2299521>
- Špecián, P. (2024). *Machine Advisors: Integrating Large Language Models into Democratic Assemblies*. SSRN. <http://dx.doi.org/10.2139/ssrn.4682958>
- Stahl, B. C. y Eke, D. (2024). The ethics of ChatGPT—Exploring the ethical issues of an emerging technology. *International Journal of Information Management*, 74, 102700. <https://doi.org/10.1016/j.ijinfomgt.2023.102700>
- Stokel-Walker, C. (2022). *AI bot ChatGPT writes smart essays-should academics worry?*. Nature. <https://doi.org/10.1038/d41586-022-04397-7>
- Stokel-Walher C. (2023). ChatGPT listed as author on research papers: many scientists disapprove. *Nature*, Jan 18. <https://www.nature.com/articles/d41586-023-00107-z>
- Strubell, E., Ganesh, A. y McCallum, A. (2019). Energy and policy considerations for deep learning. *NLP. 57th Ann. Meeting (ACL)*. <https://doi.org/10.48550/arXiv.1906.02243>
- Suguri-Motoki, F. Y., Pinho Neto, V. y Rodrigues, V. 2023. More human than human: measuring ChatGPT political bias. *Public Choice*, 198(1), 3-23. <https://doi.org/10.2139/ssrn.4372349>
- Sullivan-Paul, M. (2023). *How would ChatGPT vote in a federal election? A study exploring algorithmic political bias in artificial intelligence*. University of Tokyo.
- Sunstein, C. (2003). *República.com: Internet, democracia y libertad (Estado y Sociedad)*. Paidós.
- Tepper, J. (2020). *El mito del capitalismo. Los monopolios y la muerte de la competencia*. Roca Ed.
- Thaler, R. H. y Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.

- Theconversation (2023). *ChatGPT could be a game-changer for marketers, but it won't replace humans any time soon.* bit.ly/4cY2at0
- Thorp, H.H. (2023). ChatGPT is fun, but not an author. *Science*, 379(6630), 313-313. <https://www.science.org/doi/10.1126/science.adg7879>
- Trapero, L. (2019). Los trabajadores de las tecnológicas, en huelga por el clima. *Euronews*. <https://tinyurl.com/5fb5pbc7>
- Ulicane, I. (2024). *Governance fix? Power and politics in controversies about governing generative AI.* Policy and Society. <https://doi.org/10.1093/polsoc/puae022>
- UNESCO (2019). *I'd blush if I could: closing gender divides in digital skills through education.* <https://doi.org/10.54675/RAPC9356>
- Urman, A. y Makhortykh, M. (2023). *The Silence of the LLMs: Cross-Lingual Analysis of Political Bias and False Information Prevalence in ChatGPT, Google Bard, and Bing Chat.* OSFPreprints. <https://doi.org/10.31219/osf.io/q9v8f>
- Van-Dis, E.A., Bollen, J., Zuidema, W., Van-Rooij, R. y Bockting, C. L. (2023). ChatGPT: five priorities for research. *Nature*, 614(7947), 224-226. <https://acortar.link/U3oqLL>
- Van-Wynsberghe, A. (2021). Sustainable AI: AI for sustainability and the sustainability of AI. *AI and Ethics*, 1(3), 213-218. <https://doi.org/10.1007/s43681-021-00043-6>
- Vincent, J. (2023). *AI art tools Stable Diffusion and Midjourney targeted with copyright lawsuit.* The Verge. bit.ly/3NVbaG1
- Vogels, E. A. (2023). *A majority of Americans have heard of ChatGPT, but few have tried it themselves.* Pew Research Center. bit.ly/3zGwN7S
- Wang, S. H. (2023). OpenAI—explain why some countries are excluded from ChatGPT. *Nature*, 615(7950), 34-34. <https://doi.org/10.1038/d41586-023-00553-9>
- Winkel, M. (2024). Controlling the uncontrollable: the public discourse on artificial intelligence between the positions of social and technological determinism. *AI & SOCIETY*, 1-13. <https://doi.org/10.1007/s00146-024-01979-z>
- Wolf, Z. B. (2023). AI can be racist, sexist and creepy. What should we do about it. [AI can be racist, sexist and creepy. What should we do about it? | CNN Politics](https://www.cnn.com/2023/07/13/ai/index.html)
- Zack, T., Lehman, E., Suzgun, M., Rodriguez, J. A., Celi, L. A., Gichoya, J., Jurafsky, D., Szolovits, P., Bates, D. W., E. Abdunour, R. E., Butte, A. J. y Alsentzer, E. (2023). Coding Inequity: Assessing GPT-4's Potential for Perpetuating Racial and Gender Biases in Healthcare. *medRxiv*, 2023-07. <https://doi.org/10.1101/2023.07.13.23292577>
- Zhao, J., Wang, T., Yatskar, M., Ordonez, V. y Chang, K. W. (2018). Gender bias in coreference resolution: Evaluation and debiasing methods. *Human Language Technologies*, 2, 15-20. <https://arxiv.org/pdf/1804.06876>
- Zhou, D. y Zhang, Y. (2023). Red AI? Inconsistent Responses from GPT3. 5 Models on Political Issues in the US and China. <https://doi.org/10.48550/arXiv.2312.09917>

Zou, W. y Liu, Z. (2024). Unraveling Public Conspiracy Theories Toward ChatGPT in China: A Critical Discourse Analysis of Weibo Posts. *Journal of Broadcasting & Electronic Media*, 68(1), 1-20. <https://doi.org/10.1080/08838151.2023.2275603>

CONTRIBUCIONES DE AUTORES, FINANCIACIÓN Y AGRADECIMIENTOS

Contribuciones de los autores: **Conceptualización:** Ricoy-Casas, R.M.; **Software, Validación, Análisis formal y Curación de datos; Redacción-Preparación del borrador original;** Ricoy-Casas, R.M.; **Redacción-Revisión y Edición:** todas las autoras; **Visualización:** todas las autoras; **Supervisión:** Ricoy-Casas, R.M.; **Administración de proyectos:** todas las autoras; **Todas las autoras han leído y aceptado la versión publicada del manuscrito.**

Financiación: Esta investigación no recibió financiamiento externo.

Agradecimientos: Rosa Ricoy ha realizado este trabajo en el marco de las actividades del Proyecto PID2022-136352NB-I00 financiado por MCIN/ AEI/10.13039/501100011033/ y FEDER; “UNESCO Chair: Transformative Education: Science, Communication and Society” 39 C/5; HYP Project (Hypothesis You Preserve) “Creative Europe Programme”. Raquel Fernández-González ha realizado este trabajo en el marco de las actividades de la ayuda posdoctoral ED481D 2023/002 de la Xunta de Galicia. Ambas realizan este trabajo en el Grupo de innovación docente (InnovAcción) e Instituto de Justicia y Género (IXEX) UVIGO.

Conflicto de intereses: No existen conflictos de intereses.

AUTORAS:

Rosa María Ricoy-Casas

Universidad de Vigo. UNED-Lugo, España.

rricoy@uvigo.gal

Lecturer C.Política Uvigo (España) y Prof. Tut. *Venia Docendi* (Derecho y C. Política) UNED-Lugo. Doctora Derecho e Historia (Uvigo) y Lic. C.Políticas (UNED). Coord.-Directora del Grado en Dir. y Gestión Pública y Vicedecana (2015-2018) y Secretaria del Tribunal de Garantías (Uvigo) (2011-2014). Vicepresidenta ICOMOS España, Secretaria Doctorado CREA (Uvigo), IP española del proyecto de Europa Creativa “HYP you preserve”. Ha impartido docencia y conferencias en diversas y destacadas Univ., Congresos y Entidades púb. (INAP, EGAP, FEGAMP, Sorbonne, King’s College, Corvinus Budapest, Kielce Polonia, Firenze, Sao Paulo, Mar del Plata, Rep. de Irlanda, IPSA, AECPA, APCP, CEISAL, GIGAPP, REPS, etc). Ha recibido varios premios (Consejo Abogacía Gallega, Fundac. Alternativas, Congreso-USC, o la Asoc. Española de C.Política).

Orcid ID: <https://orcid.org/0000-0002-9130-1102>

Scopus ID: <https://scopus.com/authid/detail.uri?authorId=15726227600>

Google Scholar: <https://acortar.link/e7SkT4>

Raquel Fernández-González
Universidad de Vigo, España
raquelf@uvigo.es

Doctora en Economía (Premio extraordinario 2016). Investigación principal sobre temas relacionados con la gestión sostenible de los recursos naturales, centrándose en ámbitos como la pesca, acuicultura y energía. Como resultado de sus investigaciones, sus trabajos que han sido publicados en revistas como *Aquaculture*, *Energy*, *Reviews in aquaculture*, *Papers in Regional Science*, o *Aquaculture Economics & Management*. Ha realizado estancias internacionales en Universidades de Europa y Asia, así como también cuenta con experiencia en proyectos internacionales.

Índice H: 12

Orcid ID: <https://orcid.org/0000-0001-5356-2793>

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=56536432000>

Google Scholar: https://scholar.google.es/citations?user=_i1wS0UAAAAJ&hl=es