

A critical reflection on the treatment of AI system's 'agency' in the (Spanish) media

Ariel Guersenzvaig*, Ramon Sangüesa

Published: 30 November 2022

Abstract

For some time now the reception and communication of the achievements of artificial intelligence systems in the media seem to revolve around a number of similar tropes and rhetorical structures. The attribution of 'agency' and intention to the artificial systems created, the granting of superhuman powers to these systems, and the supposedly logical consequence of the elimination of domains in human practice are some of the common points in these reports on the media. Recent cases related to DALL-E and GPT-3 are analyzed and commented on in this paper. Several problems related to hyperbolic and misplaced reporting are discussed.

Keywords:

AI; science journalism; philosophy of AI; creativity; ML

1 Introducción

Con cada vez mayor frecuencia leemos en los medios generalistas acerca de las increíbles y cuasi mágicas capacidades de la inteligencia artificial (IA). Gracias a ella, los robots serán más eficientes e inteligentes que nosotros y, a la larga, terminarán por reemplazarnos, leemos una y otra vez. Las fabulosas capacidades de estos sistemas irían desde la detección de todo tipo de enfermedades al cálculo del riesgo de reincidencia de un recluso que solicita la libertad condicional. Esto último es precisamente para lo que se usa el software *RisCanvi*, en la Justicia catalana y cuyos resultados se integran en los informes que los jueces consideran para decidir sobre esta cuestión [22]. Sistemas de este tipo, que podemos llamar *clasificatorios*, ya se utilizan para analizar, por ejemplo, una foto de un lunar en la piel para etiquetarlo como potencialmente “benigno” o “maligno” [1]. A pesar de que sus resultados en ocasiones puedan ser prometedores, estas tecnologías también son objeto de serias críticas por parte de académicos e investigadores ya que producen resultados sesgados y discriminatorios, son opacos en su funcionamiento, dificultan la atribución de responsabilidad cuando las cosas salen mal y ejecutan

tareas que quizás no se deberían delegarse a una máquina [5]. A veces, estos sistemas ni siquiera funcionan bien desde el punto de vista técnico y sus resultados son incluso peores que los que se podrían obtener mediante el azar. La propia Comisión Europea recoge estas y otras preocupaciones en el primer marco normativo común, la “*AI Act*”, que actualmente se encuentra en sus estadios finales de aprobación por parte del Parlamento Europeo. Si bien la norma reconoce el potencial rol beneficioso que la IA puede tener, también considera de *alto riesgo* a las tecnologías de IA empleadas en, entre otros ámbitos, educación o administración de justicia [24].

Aquí, sin embargo, queramos concentrarnos en las reacciones que han suscitado recientemente otro tipo de sistemas y en el tratamiento que se les da habitualmente por parte de los medios de comunicación generalista. Nos referimos a los sistemas de inteligencia artificial con los que se *generan* cosas nuevas, a los que llamaremos sistemas *generativos*. Mientras que un sistema clasificatorio se usa para trabajar con lo existente, los sistemas generativos se usan para “crear” algo que no existía hasta que es *artificialmente* generado por un sistema computacional.

Tomando como punto de partida estos sistemas generativos, realizaremos una reflexión epistemológica, basada en argumentos teóricos y argumentación analítica, acerca de cómo divulgar los avances que se dan en este campo científico-técnico de modo que esta divulgación contribuya al conocimiento y no a aumentar expectativas infundadas y a la propia *agenda* de las empresas promotoras de estos sistemas.

2 GPT-3 y DALL-E en los medios

Dos conocidos ejemplos de sistemas generativos son *GPT-3* [17] y *DALL-E* [18], ambos desarrollados por la empresa OpenAI. Estos han sido comentados desde su aparición en 2020 en incontables telediaros y programas de radio de todo el mundo, en los principales periódicos nacionales y en infinidad de otros medios.

El software *GPT-3* sirve para generar textos y responder a preguntas complejas (“¿Cuál es la razón de la guerra de Ucrania?”) o a órdenes concretas (“Escribe la historia del imperio persa con perspectiva de género”). Con el software *DALL-E*, en cambio, se pueden generar imágenes. OpenAI, la compañía detrás de ambos sistemas, ofrece en diversos *preprints* [19] ejemplos de algunos resultados obtenidos con *DALL-E* a partir de peticiones tan variadas como solicitar el dibujo de “una nutria marina al estilo de ‘La joven con un pendiente de perla’ de Johannes Vermeer” o una “foto de un astronauta a caballo”. Las imágenes resultantes son ciertamente notables y chocantes. Al menos ante el ojo inexperto, pasan por un dibujo realizado por un ser humano. *GPT-3*, por otro lado, ya se ha usado para escribir artículos y columnas de opinión [23]. (No nos

Guersenzvaig, Ariel
Elisava Facultad de diseño e ingeniería, UVic-UCC
Barcelona, España
Correspondence: aguersenzvaig@elisava.net

Sangüesa, Ramon
Universitat Politècnica de Catalunya
Barcelona, España
Email: ramon.sanguesa.i@upc.edu

interesa abordar ahora la cuestión empírica de si los resultados obtenidos tienen la misma excelencia técnica de una ilustradora o un guionista profesional. Para nuestra discusión asumimos que los dibujos son tan perfectos como dicen los medios.)

No es raro encontrar diarios de reputación mundial que al informar sobre estos sistemas incluyen estrambóticos titulares que van desde “IAs que se enamoran” [16], “que son creativas, escriben y pintan” [13], que “dominan” el lenguaje [11] o que directamente “lo pueden todo” [10]. Dadas las limitaciones de espacio, esperamos que el lector se conforme con estos pocos ejemplos.

2.1 ¿Pinta la IA mejor que un ser humano?

Llegamos a la primera cuestión que nos interesa abordar: ¿“Pinta” la inteligencia artificial por sí sola? tal como se sugiere en los medios. ¿GPT-3 realmente “escribe” o “entiende” en el sentido que “escribir” y “entender” tienen cuando lo hacen los humanos?

Una primera cuestión obvia es que estos sistemas ni se hacen a sí mismos ni se activan por sí solos sino que requieren de equipos muy numerosos de personas que los diseñen, desarrollen, entrenen, mantengan y pongan en marcha. De la misma manera que para que *Deep Blue* ganara a Garry Kaspárov hizo falta que sus creadores lo entrenaran con gigantescos *data sets* conformados por infinidad de partidas de ajedrez jugadas por humanos y reajustaran sus estrategias, reprogramándolas de forma reiterada, también sucede que para que *DALL-E* ilustre al estilo Kandinsky hacen falta no sólo Kandinsky y sus pinturas sino también las de sus seguidores e imitadores. Resulta ineludible contar con productos del ingenio humano como base para estos *nuevos* productos y dedicar horas y horas de trabajo muy especializado para estabilizar el funcionamiento de este software. El personal de ingeniería que desarrolla *GPT-3* prepara y monitoriza a este software con cientos de miles de textos para que cuando alguien lo ponga en marcha genere como *output* una serie de palabras que nos parezca, a nosotros, una oración coherente. Dentro del sistema, no obstante, esta oración no será estrictamente una oración sino simplemente un vector probabilístico [12]. ¿Qué palabra (o sílaba incluso) debe ser la siguiente? ¿Qué combinación de palabras es la más probable en función de lo que se le ha solicitado?

Atribuyéndoles capacidad de entendimiento, algunos insisten en que estos sistemas han pasado pruebas de comprensión lectora (algo que según informes podría haber ocurrido [25]). Pero es preciso hacer una salvedad: estas pruebas están diseñadas para el fin específico de medir lo que un *ser humano* es capaz de entender, no para medir lo que entiende una máquina. Aplicar estas pruebas a un sistema computacional es un grave error metodológico. No es posible afirmar en base a la superación de estas pruebas de comprensión lectora que un sistema algorítmico tenga la capacidad de entender lo que procesa, ni en relación a captar el significado de las palabras ni a la comprensión global del texto. ¿Recuerdan los videos conspiracionales en los que se exprimía zumo de naranja en un test rápido de antígenos y este terminaba dando un resultado positivo? Para los *conspiranoicos* que distribuían estos videos por la red esto demostraba que los tests eran un engaño o directamente que el COVID no existía. Pero los tests rápidos de antígenos están diseñados para comprobar fluidos con un pH como la saliva, no para zumos cítricos, que tienen un pH ácido fuera de ese rango. El resultado en este caso no es ni siquiera un falso positivo, ya que la prueba con zumo queda invalidada metodológicamente. Utilizar con un sistema de IA una prueba de comprensión lectora pensada para humanos adolece de un problema metodológico parecido.

Un sistema de IA *no entiende*, en un sentido humano, nada de lo que escribe. Recientemente algunas investigadoras se han referido a estos sistemas como *loros estocásticos* [4]. En este

sentido, GPT-3 entiende tanto de lo que escribe como un loro que canta el tango “El día que me quieras”.

Hay otra cuestión importante que suele dejarse de lado en favor de la fascinación. El problema aquí es lo que algunos filósofos [9] denominan la *falacia mereológica*: confundir una parte por el todo. El arte, la ilustración, la redacción de entradas de enciclopedias, la literatura, la composición de canciones, o el periodismo son *prácticas* humanas y, como tales, son actividades humanas cooperativas cuyos resultados no son equiparables a la totalidad de la práctica [8]. Una práctica es mucho más que sus resultados. La literatura, como práctica, no es equiparable a una novela, ni siquiera es la suma de todas las novelas. Es mucho más que eso. Es una tradición, o varias, un espacio de diálogo y disenso acerca de las excelencias de la práctica, acerca de sus medios y sus propósitos. El ajedrez es mucho más que un jaque mate. Cuando Deep Blue derrotó a Kaspárov, el ajedrez no se acabó. ¿Por qué habría de acabarse con DALL-E el arte de la ilustración?

2.2 Atribución indebida de agencia

Que la mayoría de las personas se maravillan e impresionan con la fluida comunicación que podemos tener con un sistema artificial es algo que sabemos desde las pruebas con el *chatbot* Eliza, hace ya más de 50 años. Parece, sin embargo, que no hemos aprendido la lección. Joseph Weizenbaum, su creador, quedó perplejo ante la reacción de los usuarios de ese rudimentario sistema de diálogo (que, no es detalle menor, se presentaba como psicoterapeuta). De hecho, se puede reaccionando ante esa crédula fascinación y, casi, sumisión, Weizenbaum se cuestionó cómo era posible que sus usuarios humanos le concedieran semejante capacidad de comprensión y escucha [26]. Es más, los usuarios se sentían entendidos y acompañados por lo que era un sistema que prácticamente jugaba con las mismas palabras de las preguntas y declaraciones que los “pacientes” escribían a través del terminal, algo enormemente más sencillo que lo que ahora hacen *GPT-3* o *DALL-E*. Pasan los años y volvemos otra vez a caer en una deriva que parece muy humana, la de adscribir un enorme grado de agencia, autonomía e inteligencia a nuestras propias creaciones.

La noción de *agencia*, entendida en un sentido fuerte, indica la autonomía de un agente para definir, perseguir y conseguir sus propios objetivos [21]. Seguramente alguna vez hemos oído a algún conductor (o a nosotros mismos) exclamar “¿por qué ese BMW se cruza así a 160 km/h?”. Esta atribución de agencia —como si el coche por sí solo *decidiera* cruzarse— no es realmente problemática mientras no perdamos de vista que esta atribución de agencia es solo un recurso discursivo para referirnos a lo que los humanos hacemos con las cosas. Es decir delegar la ejecución de tareas: dejar que ellas *hagan cosas por nosotros*. En este sentido la agencia de la tecnología se trata de una agencia *débil*. Sobra decir que, a su vez, en la interacción con la tecnología los humanos cambiamos de manera profunda en nuestras maneras de ser, percibir, actuar y estar en el mundo. El conocido Principio de Churchill recoge esta dinámica: “damos forma a nuestros edificios, luego ellos nos dan forma a nosotros”.

A diferencia de lo que sucedía en tiempos de Weizenbaum, hoy en día tanto los creadores de estos espectaculares sistemas como los propios departamentos de comunicación de las empresas que los desarrollan parecen desvivirse para conseguir que demos por supuesta una agencia fuerte de las máquinas y que concedamos que tienen ya algo cercano a la omnipotencia.

Esta perspectiva llega al público general y contribuye a asentar un relato de progreso imparables, inexorable y unidireccional, no sólo ya hacia máquinas con el mismo nivel de agencia que los humanos sino directamente sobrehumanas y no sólo en tareas muy

concretas sino de manera general. En cambio, este es un debate que en la propia disciplina de la inteligencia artificial lleva generando periódicamente desde su inicio no pocas discusiones y enfrentamientos. Y periódicamente la inteligencia artificial ha caído en “inviernos” paralizantes, víctima del propio cultivo de elevadísimas expectativas. De repente, se acaba el dinero para la IA porque esta no cumple con lo prometido [20].

Un mínimo conocimiento de la historia de la disciplina y de los valores del entorno empresarial que la ha aprovechado, nos debería llevar a adoptar una cierta distancia de cautela en nuestros juicios, o, como mínimo, a no entusiasmarlos tan fácilmente. Sin embargo, abundan las actitudes acríticas y crédulas. Estas son especialmente perniciosas cuando provienen de medios de comunicación serios, que justamente deberían estar para hacer pensar a la ciudadanía más allá de sus primeras impresiones intuitivas. Nos resulta inadmisible que desde la sección de tecnología de medios especializados o de diarios de mucho prestigio no se mantengan unos mínimos estándares en cuanto al vocabulario empleado que atribuye a estos sistemas una agencia y autonomía que no tienen (e.g., “la IA dibuja y es creativa”). Claro, el papel del titular llamativo desde siempre y el *clickbait* necesario para sobrevivir ahora no son desdeñables para entender estas prácticas comunicativas tan empobrecedoras del debate público.

2.3 El problema de las fuentes

Tampoco se alcanzan unos mínimos estándares de calidad cuando estos mismos medios no consideran seriamente qué tipo de fuentes pueden contar como evidencias para que hablar de “avances históricos” esté justificado. Lo que se ha escrito en las últimas semanas sobre *DALL-E* (o sobre el modelo *540B* desarrollado por Google, similar en operativa a *GPT-3*), está basado en *papers* internos de las propias empresas. O, como mínimo, las referencias que son aducidas en la mayoría de las comunicaciones en ámbitos generalistas son así. Se trata de informes internos que no han pasado por los canales habituales de la comunicación científica ni han sido sometidos al escrutinio de la revisión por pares de las revistas especializadas, ni han aportado información verificable acerca de la manera en que estos sistemas han sido entrenados, ni en base a qué datos. A veces incluso, estas empresas evitan las críticas negando el acceso a prestigiosos académicos [15], una práctica reñida con las buenas prácticas científicas.

Los analistas y medios que anuncian el fin del arte o del periodismo renuncian a requerir unas mínimas evidencias objetivas y sucumben a un papanatismo que no sólo no se cuestiona nada sino que además copia, pega y reproduce con domesticada ingenuidad los materiales distribuidos por los departamentos de RRPP y comunicación de las empresas que están detrás de estos sistemas, convirtiéndose en un coro que amplifica un mensaje de origen empresarial que responde a una clara estrategia de marketing.

Dicho de otra manera, una empresa nos cuenta que ha llegado a Marte pero no nos dice cómo y como única evidencia nos muestra la foto de una roca guardada en una caja fuerte, que no se puede abrir porque, claro, es secreto industrial. Un coro de niños canta alabanzas desde sus páginas y pantallas: “Histórico, ¡Google ha llegado a Marte! Nada volverá a ser igual.”

La acusación de escepticismo (tal como le sucedió a uno de nosotros en conversaciones en Twitter) es la estrategia más al alcance de la mano de los crédulos cuando se les cuestiona este proceder. Sin embargo, tal como recomienda la anteriormente citada Bender [3], debemos rechazar que se nos etiquete así y, con ello, se nos saque de la discusión. Aceptar que se nos relegue al rincón de los escépticos es aceptar que quienes enmarquen y participen en el debate sean únicamente los crédulos. En una

cuestión científica o técnica, uno es escéptico cuando cuestiona una afirmación al encontrarse ante evidencias que, por una razón u otra, no acepta como válidas o suficientes. En el caso de *DALL-E* simplemente, como científicos, no aceptamos que un *paper* corporativo constituya, por sí solo, una evidencia que justifique afirmaciones periodísticas tan ditirámicas. Si la industria del tabaco y del petróleo nos han enseñado algo es que los “estudios” corporativos internos no son evidencias científicas hasta que transitan por los canales científicos apropiados. Claro está que esto tampoco es una garantía completa, pero eso es otra cuestión. Lo central aquí es que la acusación de escepticismo no es más que un sofisma cuando proviene de quien no blande artículos revisados por pares sino documentación corporativa sin verificación externa. ¿Acaso nos llamamos escépticos por rechazar el nulo valor evidencial de los positivos obtenidos mediante zumo de naranja?

Son aquellos que dicen que han llegado a Marte sobre quienes recae el peso de la prueba. Son ellos quienes deben aportar evidencias legítimas que sustenten sus afirmaciones y ofrezcan suficientes datos para que la comunidad técnica y científica pueda corroborarlas y para que los medios periodísticos, eventualmente, puedan divulgarlas una vez corroboradas.

Si la cosa quedara aquí, se trataría de una cuestión de estándares profesionales del periodismo y por tanto sería algo fuera de nuestra competencia. Pero la cosa no queda aquí, y por eso nos importa el asunto. Cada uno de estos artículos contribuye a normalizar y consolidar una narrativa que de manera sistemática exagera las capacidades reales de estas tecnologías y les atribuye un rol en la sociedad que no tienen ni quizás deban tener (este es precisamente un debate importante, necesario y urgente). Este *hype*, esta hipereexpectación acerca de las posibilidades de la IA, contribuye a nutrir y asentar imaginarios tecnológicos que inequívocamente benefician a las empresas desarrolladoras y comercializadoras de estas tecnologías. Por eso resulta tan pernicioso cuando los medios reproducen la idea de que un sistema de software que implementa técnicas de IA actúa autónomamente y por tanto entiende conceptos, escribe como un guionista, pinta como un artista, es creativa, empática y hasta tiene un grado embrionario de consciencia.

En la intersección de las ciencias de la computación, las ciencias cognitivas, la filosofía y las humanidades sigue sin haber consenso acerca de hasta qué punto las afirmaciones acerca de la inteligencia o la superinteligencia (entendida como la situación en la que un sistema no solo aprende sino que aprende cómo aprender [6]) de las máquinas están justificadas [14]. Incluso aceptando la posibilidad teórica de la superinteligencia es poco controversial afirmar que a la técnica le falta bastante para llegar a ese momento [6]. Es más, hasta los significados de las nociones de “inteligencia” y “consciencia” son vigorosamente debatidos. Sin embargo, cada vez que los medios mencionan a las “inteligencias artificiales” (sic) que “superan” a los humanos se pierde de vista que, con el estado del arte actual, la atribución de agencia no debe ser más que un recurso discursivo para hablar de aquello que nosotros hemos delegado en ellas. Lo nocivo es que se repite un argumento de persuasión y seducción que será utilizado en beneficio propio por parte de los vendedores de estas tecnologías. De este halo persuasivo se benefician especialmente las empresas desarrolladoras de sistemas clasificatorios, cada vez más ubicuos.

3 Conclusión

El mecanismo del beneficio retórico es sencillo de imaginar: si un sistema de IA es capaz de leer y comprender como cualquiera de nosotros, de redactar una entrada de la Wikipedia, o incluso de pintar como Vermeer, de jugar al ajedrez como Kaspárov o de

escribir y lograr emocionar como un guionista de Hollywood, ¿por qué no habría de ser también capaz de asignar recursos sanitarios, despedir y contratar trabajadores, decidir quién recibe un préstamo, quién puede pasar una frontera, quién copia en un examen o qué personas y barrios deben ser vigilados para combatir la criminalidad? Estos no son meros escenarios futuristas, todo esto ya está ocurriendo a pesar de la fragilidad técnica de estos sistemas y los altísimos riesgos y daños que conlleva su uso, especialmente en situaciones de injusticia estructural que afectan a colectivos vulnerables y personas tradicionalmente discriminados [5]. Pero está ocurriendo no por la decisión autónoma de estos sistemas que usan técnicas de inteligencia artificial, sino porque alguien en un puesto jerárquico de una empresa o institución pública ha decidido integrarlos en estos procesos, a menudo reemplazando el juicio humano, que si bien es claramente imperfecto, es más transparente que el de estos sistemas y permite un mejor escrutinio externo de la toma de decisiones y sus justificaciones.

Nada de lo que “hace” un sistema que implementa técnicas de IA es verdaderamente análogo a lo que “hace” un ser humano. El resultado obtenido puede resultar idéntico (un dibujo o un texto, en los casos que hemos analizado). Sin embargo, a no ser que estemos dispuestos a adoptar una visión reduccionista que afirme que los seres humanos no somos más que sistemas estadísticos de toma de decisiones, el proceso creativo por el cual un dibujante llega a un dibujo o una escritora a un texto es radicalmente distinto del empleado por las máquinas para llegar a estos resultados [2].

Para acabar, nada impide, en principio, que *GPT-3* y *DALL-E* generen resultados de altísima calidad técnica. De ser así, nos encontraremos ante nuevas y maravillosas muestras del ingenio humano. A menos que estos sistemas resulten en nuevas formas de perpetuar prejuicios discriminatorios y exacerbar injusticias estructurales existentes, debemos celebrarlos y felicitar con admiración a montones de personas que con su creatividad y trabajo los han hecho posibles. No obstante, habrá que seguir teniendo presente que la acción humana es una condición necesaria para la existencia y funcionamiento de estos sistemas y de muchos otros como, por ejemplo, el empleado por el pianista Marco Mezquida, quien utilizó un sistema que reaccionaba a sus improvisaciones con nuevas improvisaciones durante el arranque del Sonar AI+MUSIC Festival del 2021 en Barcelona [7]. Hablar de “IAs que pintan, escriben o hacen música mejor que los humanos”, no solo es absurdo y profundamente ingenuo, sino que también es una falta de reconocimiento hacia todos esos humanos que diseñan, desarrollan, entrenan, ponen en marcha y mantienen estos sistemas.

4 Referencias

- [1] Adamson, A. S. and Smith, A. Machine Learning and Health Care Disparities in Dermatology. *JAMA Dermatology*, 154, 11 (2018), 1247-1248.
- [2] Amabile, T. M. Componential Theory of Creativity. In *Encyclopedia of Management Theory* ed. Sage, LA, 2013.
- [3] Bender, E. M. Tweet 2:18 AM · Apr 18. 2022. Disponible en: <https://bit.ly/3I9F0n7>
- [4] Bender, E. M., Gebru, T., McMillan-Major, A. and Shmitchell, S. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In *FAccT '21* <https://doi.org/10.1145/3442188.3445922>
- [5] Benjamin, R. *Race After Technology: Abolitionist Tools for the New Jim Code*, ed. Polity, Cambridge, 2019.
- [6] Bryson, J. *The Past Decade and Future of AI's Impact on Society*. <https://bit.ly/3GrVTZc>
- [7] Festival, S. Marco Mezquida 'Piano + AI'. 2021. Disponible en: <https://www.youtube.com/watch?v=dHh4vP5T6VM>.
- [8] Guersenzvaig, A. *An inquiry into the practice of Designed*. Mimesis International, Forthcoming.
- [9] Harre, R. O. M. Behind the Mereological Fallacy. *Philosophy*, 87, 341 (2012), 329-352.
- [10] Heller, P. *Ein algorithmischer Alleskönner*. *Frankfurter Allgemeine*, 2020. Disponible en: <https://www.faz.net/>
- [11] Johnson, S. A.I. Is Mastering Language. Should We Trust What It Says? *New York Times Magazine*, 2022.
- [12] Jurafsky, D. and Martin, J. H. *Speech and Language Processing*, 2nd ed. Pearson., Upper Saddle River, 2009
- [13] Llaneras, K. ¿Crees que la inteligencia artificial no es creativa? Mira cómo escribe y cómo pinta. *El País*, 2022. Disponible en: <https://bit.ly/3GIwvRM>
- [14] Marcus, G. *Rebooting AI: Building Artificial Intelligence we can Trust*, ed. Pantheon Books, New York, 2019.
- [15] Marcus, G. and David, E. Experiments testing GPT-3's ability at commonsense reasoning: results. 2022. Disponible en: <https://bit.ly/3G0xFUg>
- [16] Metz, C. When A.I. Falls in Love. <https://nyti.ms/3Q4aUTZ>
- [17] OpenAI.API. 2022. Disponible en: <https://openai.com/api/>
- [18] OpenAI. DALL-E 2. 2022. Disp. <https://openai.com/dall-e-2/>
- [19] Ramesh, A., Dhariwal, P., Nichol, A., Chu, C. and Chen, M. *Hierarchical Text-Conditional Image Generation with CLIP Latents*. City, 2022. Disponible en: <https://bit.ly/3CcvzQ3>
- [20] Russell, S. J. and Norvig, P. *Artificial Intelligence: A Modern Approach*, 3rd ed. Prentice Hall, Upper Saddle River, 2010.
- [21] Sangüesa, R. and Guersenzvaig, A. AI as a Design Material: Dealing with New Agencies. *Temes de Disseny*, 2019, 6-25.
- [22] Soler, C. *RisCanvi. La predicció de la reincidència i la conducta violenta a les presons*. Direcció General de Serveis Penitenciaris, Departament de Justícia, Generalitat de Catalunya, Barcelona, 2018.
- [23] The Guardian. A robot wrote this entire article. Are you scared yet, human? 2020. <https://bit.ly/3G4TYZ5>
- [24] UE. What is the EU AI Act? 2022 <https://bit.ly/3jDUoOA>
- [25] VV.AA. Google Research *PaLM: Scaling Language Modeling with Pathways*. City, 2022. <https://bit.ly/3jCyWcl>
- [26] Weizenbaum, J. *Computer power and human reason: from judgment to calculation*, ed. W.H. Freeman, SF, 1976.



© 2022 by the authors. This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.