

Dossier

El rol de la intel·ligència artificial en la gestió de la informació

Intel·ligència artificial:
una introducció de butxaca





Carlos LOPEZOSA
Facultat d'Informació i Mitjans
Audiovisuals, Universitat de
Barcelona
lopezosa@ub.edu

Jorge FRANGANILLO
Facultat d'Informació i Mitjans
Audiovisuals, Universitat de
Barcelona
franganillo@ub.edu

Article rebut el gener del 2025;
revisat el febrer del 2025

DOI
<https://doi.org/10.60940/itemn78433001>

Intel·ligència artificial: una introducció de butxaca

Resum: La intel·ligència artificial busca crear sistemes capaços de fer tasques els resultats de les quals siguin indistingibles de la intel·ligència humana. Aquesta disciplina ha evolucionat notablement, i avui els models generatius, capaços de crear text, imatges, so i vídeo, han transformat la forma en què accedim i creem contingut. Tot i el seu potencial creatiu, aquestes eines plantegen reptes sobre la veracitat i l'autoria, i mostren limitacions com ara biaixos i errors. En aquest context, els professionals de la informació tenen la responsabilitat social de guiar la ciutadania cap a un ús crític i responsable d'aquesta tecnologia.

Paraules clau: intel·ligència artificial, xarxes neuronals, processament del llenguatge natural, models generatius, formulació d'instruccions.

Inteligencia artificial: una introducción de bolsillo

Resumen: La inteligencia artificial busca crear sistemas capaces de realizar tareas cuyos resultados sean indistinguibles de la inteligencia humana. Esta disciplina ha evolucionado notablemente, y hoy los modelos generativos, capaces de crear texto, imágenes, sonido y vídeo, han transformado la forma en la que accedemos y creamos contenido. A pesar de su potencial creativo, estas herramientas plantean retos sobre la veracidad y la autoría, y muestran limitaciones como sesgos y errores. En este contexto, los profesionales de la información tienen la responsabilidad social de guiar a la ciudadanía hacia un uso crítico y responsable de esta tecnología.

Palabras clave: inteligencia artificial, redes neuronales, procesamiento del lenguaje natural, modelos generativos, formulación de instrucciones.

Artificial intelligence: a pocket-sized introduction

Abstract: Artificial intelligence aims to create systems capable of performing tasks with results indistinguishable from human intelligence. This field has evolved remarkably, and today generative models, which can create text, images, sound, and video, have transformed how we access and create content. Despite their creative potential, these tools raise challenges concerning veracity and authorship, and exhibit limita-

tions such as biases and errors. In this context, information professionals have a social responsibility to guide citizens toward the critical and responsible use of this technology.

Keywords: artificial intelligence, neural network, generative models, prompt engineering, natural language processing.



Sembla que la intel·ligència artificial (IA) ha pasat, en poc temps, de ser una idea de ciència-ficció a convertir-se en un actor quotidià a les nostres vides. Potser per a molts de nosaltres, aquesta irrupció es va fer especialment evident amb el llançament del *ChatGPT* el novembre del 2022.

De sobte, vam tenir a l'abast un sistema capaç de generar textos amb una fluïdesa sorprenent, de respondre preguntes amb agilitat i de mantenir converses que, en alguns casos, resultaven indistingibles de les que podríem tenir amb una altra persona.

Ara, moltes persones ja integren eines basades en IA en la seva rutina diària, ja sigui per redactar missatges, generar idees o, fins i tot, automatitzar certes tasques a la feina. Però, què és exactament aquesta tecnologia que tant està revolucionant el nostre dia a dia?

1. Què és la intel·ligència artificial

Curiosament, la IA no és una novetat tan recent com sembla. De fet, les seves arrels es remunten ara fa uns setanta anys. Una de les primeres conceptualitzacions formals sobre la IA va tenir lloc el 1955, durant un esdeveniment al Dartmouth College nord-americà, on aquesta disciplina es va definir com la ciència i l'enginyeria en la creació de màquines intel·ligents.¹

Aquesta definició, tot i ser força precisa per a la seva època, va evolucionar al mateix ritme que la mateixa tecnologia. Actualment, la podríem definir com la disciplina que busca crear tecnologies capaces de resoldre tota mena de tasques complexes que fins ara requerien intel·

En realitat, la IA té una llarga història i fa temps que està integrada en el teixit de la nostra vida quotidiana, sovint de manera invisible.

ligència humana. Per aconseguir-ho, la IA incorpora, en certs casos, un component d'aleatorietat de forma deliberada, precisament per simular la complexitat i la imprevisibilitat del pensament intel·ligent i creatiu.

De fet, aquesta idea d'un sistema que imiti la intel·ligència humana té un origen anterior a la seva formalització com a disciplina científica. Alan Turing va proposar l'any 1950 el que avui coneixem com el «test de Turing». Aquest procediment consisteix a avaluar, a través d'una conversa, la capacitat d'una màquina per exhibir un comportament intel·ligent equivalent al d'una persona. A través d'un terminal, un jutge humà interroga una persona i una màquina, sense saber quina és quina. Si el jutge no pot distingir, de forma fiable, quines respostes provenen de la màquina i quines de la persona, llavors es pot dir que la màquina ha superat el test i que mostra, per tant, un cert grau d'intel·ligència.²

Per a molts especialistes, aquesta similitud amb el comportament humà és la clau per determinar si un sistema és intel·ligent. Implica que una màquina no només sigui capaç de processar informació, sinó que també pugui comprendre-la, aprendre'n i utilitzar-la per prendre decisions o resoldre problemes de manera autònoma.³

Si prenem aquesta accepció com a punt de partida, trobem una gran quantitat de desenvolupaments, alguns dels quals ja icònics, que,

1. McCarthy, John; Minsky, Marvin L.; Rochester, Nathaniel; Shannon, Claude E. "A proposal for the Dartmouth summer research project on artificial intelligence: August 31, 1955". *AI Magazine*, vol. 27, núm. 4 (2006), p. 12–14. <https://doi.org/10.1609/aimag.v27i4.1904> [Consulta: 18/01/2025].
2. Turing, Alan. "Computing machinery and intelligence". *Mind*, núm. 49 (1950), p. 433–460. <https://doi.org/10.1093/mind/LIX.236.433> [Consulta: 18/01/2025].
3. Boden, Margaret A. *Artificial intelligence: a very short introduction*. Oxford: Oxford University Press, 2018.

tot i semblar llunyans, demostren la capacitat de la IA per abastar contextos diversos. Des del Deep Blue d'IBM, capaç de vèncer el campió mundial d'escacs Garri Kaspàrov l'any 1997, fins al robot aspirador Roomba, un dels primers productes comercials d'èxit per a la llar que va incorporar aquesta tecnologia el 2002, la IA ha demostrat el seu potencial en àmbits ben diversos.

Aquests exemples, però, només són la punta de l'iceberg. En realitat, la IA té una llarga història i fa temps que està integrada en el teixit de la nostra vida quotidiana, sovint de manera invisible. Des que el nostre telèfon mòbil ens guia cap a una destinació, fins quan una xarxa social ens suggereix etiquetar certs contactes en una foto o quan una botiga virtual ens recomana un producte amb un encert gairebé telepàtic, la IA hi és present. Aquestes manifestacions discretes són una prova de la seva magnitud i versatilitat com a disciplina en si mateixa.

2. De l'aprenentatge automàtic als models generatius

La IA és un camp vast, amb diverses branques i aplicacions, entre les quals destaca l'aprenentatge automàtic, una àrea que busca capacitar els sistemes per aprendre de grans quantitats de dades de manera autònoma. En lloc de programar els sistemes amb instruccions explícites, se'ls ensenya a aprendre d'exemples, tal com aprenem els humans en observar patrons al nostre entorn quotidià.

Una mostra d'aquesta capacitat es manifesta en els filtres de correu brossa, que aprenen a distingir entre correus legítims i no desitjats a partir de l'anàlisi de milers de casos. Aquesta funció, aparentment senzilla, demostra la

capacitat de l'aprenentatge automàtic per resoldre tasques complexes de classificació i reconeixement de patrons.

Dins l'aprenentatge automàtic, trobem una tècnica avançada: l'aprenentatge profund. Aquesta metodologia imita la forma en què el cervell humà processa informació i adquireix habilitats. Per aconseguir-ho, i aprendre directament de dades sense processar, utilitza xarxes neuronals, estructures formades per capes interconnectades de neurones artificials. I és sorprenent com aquestes xarxes poden identificar patrons complexos sense que se'ls indiqui específicament què han de buscar.⁴

Per exemple, una xarxa neuronal es pot entrenar a resumir articles de revista. Per fer-ho, se la nodreix amb nombrosos articles i els seus corresponents resums. Sense regles predefinides, el sistema aprèn a identificar elements clau com ara noms, dates rellevants i frases importants. Amb el temps, és capaç de generar resums coherents i precisos de nous articles, i d'extreure i condensar la informació essencial per si mateix.

Aquesta mateixa lògica s'aplica al reconeixement d'imatges. Si tenim una col·lecció de fotos d'animals i volem que un sistema les classifiqui automàticament, l'aprenentatge profund permet que la xarxa neuronal aprengui a reconèixer trets distintius, com les orelles d'un gat o les ratlles d'un tigre. D'aquesta manera, la xarxa pot classificar amb precisió imatges que mai no ha vist d'aquests animals; tot això, sense instruccions específiques sobre què ha de cercar.

Això evidencia l'eficàcia de l'aprenentatge profund per analitzar patrons complexos de dades i aprendre'n de manera independent, un procés que s'assembla a com els humans aprenem a reconèixer formes o predir paraules a partir de referències. Aquest aprenentatge és especialment útil per al processament del

4. Campesato, Oswald. *Artificial intelligence, machine learning and deep learning*. Mercury Learning and Information, 2020.

A diferència d'altres branques de la IA, centrades en l'anàlisi i la classificació de dades, els models generatius busquen crear contingut nou i singular de tota mena: text, imatge, veu, so, música o vídeo.

llenguatge natural, que permet als sistemes reconèixer, interpretar i generar llenguatge humà, amb nombroses aplicacions: des de la traducció automàtica i la síntesi de veu fins a la visió per ordinador i la indexació automàtica d'imatges.

Aquests avenços ens han conduït a la IA generativa, la branca més activa i influent avui dia. A diferència d'altres branques de la IA, centrades en l'anàlisi i la classificació de dades, els models generatius busquen crear contingut nou i singular de tota mena: text, imatge, veu, so, música o vídeo. Per fer-ho, aquests models, primer, aprenen patrons i relacions dins un enorme corpus de dades. A partir d'aquest aprenentatge, la IA és capaç de generar contingut nou que manté l'estil i les característiques observades en les dades originals.

En molts d'aquests models generatius, com ara les xarxes generatives adversàries, hi ha dues *xarxes neuronals* que treballen conjuntament de manera antagònica. Una *xarxa generadora* crea nou contingut a partir de les dades apreses, mentre que una *discriminadora* avalua si el contingut generat s'assembla prou a les dades reals de l'entrenament i intenta distingir entre el contingut real i el generat artificialment. Aquestes xarxes es desafien mútuament en una competició contínua on la generadora intenta crear contingut cada cop més realista, mentre que la discriminadora procura ser cada cop més precisa a l'hora de fer aquesta dis-

tinció. Aquesta interacció constant millora la qualitat del contingut generat.

L'atractiu principal dels models generatius rau en la facilitat d'ús. Per interactuar amb el sistema i produir contingut, no cal dominar cap llenguatge de programació complex. Sovint n'hi ha prou d'utilitzar el llenguatge natural. A través d'instruccions senzilles i naturals —allò que anomenem peticions o *prompts*— el sistema pot entendre i respondre amb una naturalitat comparable a la humana.

2.1. Generació de llenguatge natural

La generació de llenguatge natural és una aplicació destacada de la IA generativa. Aquesta tecnologia es basa en models de llenguatge extensos que, a partir de l'anàlisi estadística de grans corpus de text, assimilen les regles del llenguatge i funcionen com a representacions probabilístiques. Per fer-ho, transformen les paraules en vectors numèrics que capturen el seu significat i la seva funció sintàctica. Aquests vectors, que es projecten en un espai multidimensional, formen mapes semàntics on les paraules amb significats similars es troben properes. A més, aquestes representacions permeten als models transformar les paraules, i canviar-ne trets semàntics, per tal de generar textos originals.

En l'actualitat, hi ha diversos models de llenguatge, com ara GPT (d'OpenAI), PaLM i Gemini (de Google), LLaMA (de Meta) o Claude (d'Anthropic). Tots ells imiten l'escriptura humana de manera tan convincent que resulta difícil distingir un text humà d'un de generat artificialment. De fet, alguns estudis suggereixen que models com el GPT-4 han superat proves similars al test de Turing,⁵ cosa que planteja reflexions importants sobre la naturalesa de la intel·ligència i la creativitat.

5. Jones, Cameron R.; Bergen, Benjamin K. «People cannot distinguish GPT-4 from a human in a Turing test» (2024). Prepublicació. <https://arxiv.org/abs/2405.08007> [Consulta: 18/01/2025].

2.2. Síntesi d'imatges a partir de text

La capacitat de generar imatges a partir d'una descripció textual o, fins i tot, d'una o diverses imatges de referència, és un altre assoliment destacable de la IA generativa. Aquests models aprenen a associar descripcions textuales amb imatges, i construeixen representacions numèriques que codifiquen els trets visuals. De manera similar a com els models de llenguatge aprenen les relacions entre paraules, aquests models aprenen les relacions entre descripcions (textuals o gràfiques) i les seves corresponents representacions visuals. A partir d'aquestes representacions, transformen l'entrada —ja sigui text o una imatge— en un «plànol» que utilitzen per generar matrius de píxels que s'ajustin a la descripció o referència gràfica donada. Gràcies a això, models com ara *DA-VALL·E*, *Midjourney* o *Stable Diffusion* permeten generar imatges realistes o artístiques amb gran facilitat.

2.3. Síntesi sonora i clonació de veu

La IA generativa ha transformat la creació de contingut sonor i avui permet generar tant peces de música i efectes de so com veus sintètiques que fins i tot poden imitar el timbre de veu d'una celebritat.

D'una banda, la síntesi sonora converteix textos en veu parlada, de manera persuasiva. Tot i ser una tecnologia prou coneguda per la seva aplicació en assistents de veu o lectors de pantalla, gràcies a l'aprenentatge profund ara explora noves possibilitats comercials. D'altra banda, la IA generativa també ha revolucionat la creació d'efectes de so i permet generar tota mena de paisatges sonors per a pel·lícules, videojocs o aplicacions de realitat virtual a partir de simples descripcions o instruccions. Finalment, la clonació de veu, una aplicació innovadora de la síntesi de veu, permet replicar veus famoses per fer-les servir en diversos formats

Aquesta tecnologia permet crear vídeos que transformen completament l'aparença i el discurs dels protagonistes, amb un nivell de realisme que fa difícil distingir-los de la realitat.

i entorns, amb aplicacions com *ElevenLabs* o *Speechify*.

Paral·lelament, la IA generativa també està revolucionant la composició de música i permet crear peces originals en diversos estils a partir d'una simple petició escrita. Actualment, hi ha plataformes com *Amper Music*, *Suno* o *Udio* que permeten explorar aquesta tecnologia.

2.4. Hipertrucatge i generació de contingut audiovisual

La manipulació de vídeos, tot i no ser una novetat, ha assolit noves cotes de sofisticació amb l'hipertrucatge (*deepfake*). Aquesta tecnologia permet crear vídeos que transformen completament l'aparença i el discurs dels protagonistes, amb un nivell de realisme que fa difícil distingir-los de la realitat.

Paral·lelament, els models generatius també permeten crear vídeos sintètics a partir de simples indicacions de text. Aquesta tecnologia, basada en la síntesi d'imatges, ha demostrat un gran potencial creatiu i representa un valuós ajut per a artistes audiovisuals. Eines com *Sora* o *Runway* permeten generar seqüències de vídeo ràpidament, sense necessitat de material audiovisual previ.

Totes dues possibilitats estan transformant la producció de contingut audiovisual en àmbits com el cinema, la televisió o la publicitat, i impulsen noves formes de narració i innovació creativa.



3. Formulació d'instruccions

Tal com indicàvem abans, per interactuar amb els sistemes d'IA i obtenir-ne una resposta específica, cal formular peticions (*prompts*), expressades en llenguatge natural. Aquestes peticions acostumen a ser verbals, però la IA multimodal admet altres formats, com ara imatges, dades, codi font o fins i tot la veu.

La qualitat dels resultats depèn de la precisió i l'exhaustivitat de les instruccions. Una petició bàsica retornarà una resposta genèrica, mentre que una petició més detallada retornarà una resposta més precisa, matisada i pertinent. Sobre aquesta premissa, una petició efectiva ha de preveure alguns components essencials:⁶

- Tasca. Assignar una tasca («redacta», «compara», «resumeix», etc.) guia la resposta de manera efectiva.
- Rol. Establir un rol («actua com...») simula situacions específiques i centra els resultats.
- Context. Proporcionar informació rellevant ajuda el sistema a comprendre millor la situació per generar una resposta coherent.
- Propòsit. Indicar l'objectiu de la petició ajuda a obtenir respostes pertinents.
- Exemples. Oferir mostres o comparacions guia el sistema cap a respostes més precises.
- Audiència. Assenyalar a qui va adreçat el resultat, l'ajusta al context de la petició.
- To. Definir l'estil de la resposta (formal, informal, professional, humorístic, etc.), l'adequa al context específic.
- Detall. Com més detallada sigui la petició, més bona qualitat tindrà la resposta.
- Format. Especificar com es vol presentar la resposta (llista, taula, etc.) afavoreix respostes més útils i comprensibles.
- Extensió. Determinar la longitud objectiu permet controlar l'extensió de la resposta per adequar-la a necessitats específiques.

La qualitat dels resultats depèn de la precisió i l'exhaustivitat de les instruccions.

La formulació de peticions és flexible i no calen tots els elements. Per a una resposta eficaç, n'hi ha prou amb els components necessaris en cada cas, dels quals la tasca, el context i el format de sortida desitjat acostumen a ser els prioritaris.

4. Algunes limitacions

Tot i el seu gran potencial, els models generatius presenten limitacions que comprometen la seva fiabilitat. Aquests models, ja siguin generadors de llenguatge, imatges, so o vídeo, no són creadors originals, sinó intermediaris que processen i transformen informació accessible en altres llocs. El procés pot generar errors i inconsistències, que en el cas del text s'anomenen «al·lucinacions», però que en altres formats es poden manifestar com a resultats inexactes, artefactes visuals o construccions sonores incoherents. La causa d'aquests errors rau en les limitacions dels algorismes i de les dades emprades per a l'entrenament. Per això, és fonamental avaluar acuradament les creacions que generen aquests models abans d'utilitzar-les.

A més, la IA generativa també pot mostrar biaixos i reproduir prejudicis i estereotips presents en les fonts d'entrenament, amb independència de si són text o imatges. Tot i que s'intenta mitigar aquest problema ajustant i reentrenant els sistemes, és difícil crear models completament neutrals, ja que els biaixos es poden introduir en diverses fases del desenvolupament i, a més, són problemes socials complexos que no es poden abordar només amb solucions tècniques.⁷ En qualsevol cas, aquests

6. OpenAI. *Prompt engineering*, 2023. <https://platform.openai.com/docs/guides/prompt-engineering> [Consulta: 18/01/2025].

7. Heikkilä, Melissa. «Why it's impossible to build an unbiased AI language model». *MIT technology review*, 8 agost 2023. <https://technologyreview.com/2023/08/08/1077403> [Consulta: 18/01/2025].

biaixos poden provocar resultats inexactes o ofensius o, per contra, una diversitat forçada o una sobreadaptació a les normes socials, és a dir, respostes amb un excés de «correcció política».

Esbrinar com prevenir o corregir les respostes errònies és un repte majúscul. Tot i que es coneix el funcionament bàsic de les xarxes neuronals, el comportament que resulta de la interacció de milions de paràmetres, sumat a un cert component d'atzar deliberat, és complex i impredecible. La presa de decisions per part de la IA, sovint sense una explicació transparent, configura una «caixa negra» que preocupa a la comunitat científica perquè els errors poden tenir un impacte considerable en camps crítics com la salut, el dret o les finances.

5. Consideracions finals

Certament, la IA ha transformat la forma en què accedim a la informació i creem contingut, i planteja reptes i oportunitats per als professionals de la informació. La capacitat de generar textos, imatges, so o vídeo a partir de simples instruccions ha fet que la formulació de peticions adequades esdevingui una competència clau en la gestió d'informació.

Tanmateix, la IA també ha posat de manifest que la informació generada és òrfena, sense una autoritat clara ni una font fidedigna. Davant d'això, els professionals dels arxius, les biblioteques i la documentació tenen la responsabilitat i l'oportunitat d'actuar com a guies experts en la validació i la cura dels continguts artificials, i d'assegurar que la informació rellevant i fidedigna es pugui distingir de la desinformació i dels biaixos.

En definitiva, la IA generativa no ha de ser la fi dels serveis d'informació, sinó una invitació a redefinir i reforçar el seu rol. Per bé que els professionals no poden competir amb la IA en

A més, la IA generativa també pot mostrar biaixos i reproduir prejudicis i estereotips presents en les fonts d'entrenament, amb independència de si són text o imatges.

velocitat de resposta, sí que han de destacar per la capacitat de curar, valorar, verificar i contextualitzar la informació, i han de guiar els usuaris cap a un ús responsable i crític de la IA. Aquest nou context exigeix que els professionals de la informació esdevinguin referents en la formació d'una ciutadania crítica capaç de discernir sobre la qualitat dels continguts i d'utilitzar les eines d'informació de manera ètica i responsable.

Bibliografia

BODEN, Margaret A. *Artificial intelligence: a very short introduction*. Oxford: University Press, 2018.

CAMPESATO, Oswald. *Artificial intelligence, machine learning and deep learning*. Mercury Learning and Information, 2020.

DALE, Robert. «The voice synthesis business: 2022 update». *Natural Language Engineering*, vol. 28, núm. 3 (2022), p. 401–408. <https://doi.org/10.1017/S1351324922000146> [Consulta: 18/01/2025].

HEIKILÄ, Melissa. «Why it's impossible to build an unbiased AI language model». *MIT technology review*, 8 agost 2023. <https://technologyreview.com/2023/08/08/1077403> [Consulta: 18/01/2025].

JONES, Cameron R.; BERGEN, Benjamin K. «People cannot distinguish GPT-4 from a human in a Turing test» (2024). Prepublicació. <https://arxiv.org/abs/2405.08007> [Consulta: 18/01/2025].

MCCARTHY, John; MINSKY, Marvin L.; ROCHESTER, Nathaniel; SHANNON, Claude E. «A proposal for the Dartmouth summer research project on artificial intelligence: August 31, 1955». *AI Magazine*, vol. 27, núm. 4 (2006), p. 12–14. <https://doi.org/10.1609/aimag.v27i4.1904> [Consulta: 18/01/2025].

OPENAI (2023). *Prompt engineering*. <https://platform.openai.com/docs/guides/prompt-engineering> [Consulta: 18/01/2025].

TURING, Alan. «Computing machinery and intelligence». *Mind*, núm. 49 (1950), p. 433–460. <https://doi.org/10.1093/mind/LIX.236.433> [Consulta: 18/01/2025]. ■