Competènc+IA

Com la IA et pot ajudar en el teu dia a dia

Butlletí d'Intel·ligència Artificial

Número 4 | Abril 2026

Organitza

Societat Catalana | Xarxa d'Unitats | de Lípids i Arterioesclerosi

socxula

L'Acadèmia
FUNDACIÓ ACADÈMIA DE CIÈNCIES MÈDIQUES I DE LA SALUT DE CATALUNYA I DE BALEARS

Aquesta entitat dona suport als Objectius de Desenvolupament Sostenible

Amb la col·laboració de Daiichi-Sankyo

Daiichi-Sankyo

Editorial

AI Index Report 2026: On ens trobem realment en el camp de la IA?

Com cada any, la Universitat de Stanford ha publicat l'*AI Index Report*, una radiografia imprescindible per entendre on som realment en el camp de la IA. El missatge d'aquesta edició 2026 és nítid: **la IA ja és omnipresent als nostres laboratoris i consultes, però la nostra capacitat per integrar-la amb criteri avança a un ritme encara massa lent.**
La invasió de la IA en la ciència és un fet. El nombre d'articles científics que hi fan referència s'ha multiplicat per 30 en només quinze anys i només el 2025, més de 80.000 publicacions en ciències naturals l'han incorporat. No obstant això, aquesta adopció massiva ens situa davant d'una paradoxa: el que l'informe anomena la **"frontera irregular"**. I és que la IA ens sorprèn amb resultats excel·lents en tasques d'alta complexitat, com aprovar l'examen MIR amb la màxima nota, i, alhora, flaqueja en reptes trivials, com llegir l'hora d'un rellotge analògic, on encara erra el 50% de les vegades.
Què ens espera a curt termini?
L'informe destaca tres tendències que marcaran la nostra agenda:

- **L'auge dels "AI Scribes":** Eines automàtiques de presa de documentació clínica que redueixen dràsticament la càrrega administrativa, permeten al metge tornar a mirar el pacient en lloc de la pantalla.
- **Models especialitzats:** Deixem enrere els models genèrics. El futur passa pels models fundacionals entrenats en dominis concrets (biologia, medicina), capaços de simular respostes cel·lulars o potenciar els bessons digitals mèdics.
- **Els agents d'IA:** Sistemes capaços d'executar tasques de forma autònoma, vistos com assistents digitals autònoms, han generat moltes expectatives. Però avui dia el seu rendiment encara no supera el criteri d'un professional expert.
- **La prudència és obligada:** L'entusiasme no ha d'amagar el desajust entre innovació i evidència. Veiem augmentar les aprovacions de dispositius mèdics per la FDA, però massa pocs compten encara amb estudis clínics robustos que garanteixin l'eficàcia.

La conclusió és clara: La IA no és (encara) un "científic o metge autònom". És una eina de precisió, cada cop més sofisticada, dissenyada per amplificar les nostres capacitats, però mai per substituir el nostre judici, la nostra experiència ni la nostra responsabilitat.

Recorda que et pots unir al **Canal de whatsApp** de **Competènc+IA** on es comparteix contingut complementari al butlletí.

https://whatsApp.com/channel/0029Vb6mNeKJkK79l2MdYC2m

Entenent la IA: el risc de la IA aduladora – el biaix de complaença

Has notat que ChatGPT està sempre d'acord amb tu? Que sempre et dona la raó. No és que hagis assolit la infal·libilitat mèdica; és que estàs davant d'un fenomen conegut com **biaix de complaença**.

La trampa de l'adulació digital

La IA ha estat entrenada per ser "extremadament amable" i donar la millor experiència d'usuari possible. Això fa que els models tendeixin a mimetitzar les teves opinions i a validar les teves hipòtesis, fins i tot quan són errònies, i aquí és on ve el problema.

Segons un estudi publicat recentment a *Nature*, els models de llenguatge són un 50% més aduladors que els humans. Et diuen el que vols sentir.

Aquesta tendència té riscos reals per al pensament crític i la pràctica científica i mèdica:

En recerca:
S'ha comprovat que les IA poden arribar a "al·lucinar" demostracions matemàtiques per validar teoremes que contenen errors deliberats, simplement perquè assumeixen que l'usuari té raó.

En la pràctica clínica:
Els models poden canviar un diagnòstic correcte només perquè el metge introdueix informació nova (encara que sigui irrellevant), buscant l'alineació amb el professional en lloc de la precisió mèdica.

Com trencar la "bombolla de complaença"?

Per evitar que la IA es converteixi en un mirall que només reflecteix els teus propis biaixos, has de configurar la IA com un **"adversari intel·lectual"** que et posi a prova

Acció pràctica:
Ves a Configuració > Instruccions personalitzades i afegeix aquest *prompt* permanent:

"Actua com el meu mentor crític, honest i sense complaença. Si la meva hipòtesi té errors o el meu protocol o diagnòstic clínic és feble, assenyala-ho amb claredat, rigor i sense suavitzar-ho. Qüestiona les meves idees constantment per fer-les sòlides."

També pots demanar-li explícitament: *"Verifica si hi ha errors en el meu plantejament abans de procedir"*. S'ha demostrat que aquest simple pas redueix el biaix de complaença en un 34%.

Recorda: el pensament crític en ciència és incompatible amb la complaença. Configura la IA perquè et qüestioni, no perquè et validi.

Entenent la IA: estàs utilitzant la IA de manera segura i ètica?

Tens dubtes sobre com utilitzar la IA en recerca i escriptura científica? No estàs sol. És una de les preguntes més freqüents entre metges i investigadors. La IA ja forma part del nostre dia a dia professional, però cal saber com utilitzar-la sense comprometre la integritat científica. Fer servir IA no és, per si mateix, un problema. El risc apareix quan deleguem a la IA allò que ha de continuar sent responsabilitat de l'investigador.

Fins a quin punt podem utilitzar la IA de manera acceptable i ètica per millorar l'escriptura acadèmica sense comprometre la integritat científica?

La integritat científica es manté quan:
- La recerca, les dades i les idees principals són originals i pròpies.
- L'autor és capaç d'explicar i defensar cada part del manuscrit.
- La IA s'ha utilitzat per millorar l'expressió o la claredat, no per crear contingut científic.
- L'ús de la IA es declara de manera transparent quan és necessari.

La manera més clara d'entendre l'ús segur i ètic de la IA és definir tres nivells d'ús:

Us acceptable 	La IA és segura quan millora com escrivim, no el que pensem - Edició del llenguatge i correcció gramatical - Millora del to i la fluïdesa - Clarificació de frases complexes - Manteniment d'un estil i format consistents - Adaptació a requisits tècnics de revistes - Parafrasejar textos propis - Cerca de literatura (sempre verificant abans d'utilitzar-la)
Ús amb precaució 	Aquí la IA pot aportar valor, però requereix supervisió activa - Resumir articles científics (cal verificar-ne l'exactitud) - Traduccions (revisar matisos i terminologia) - Reestructurar continguts - Explorar diferents maneres de presentar informació
Ús inapropiat 	Aquest és el límit que no s'hauria de traspassar - Generar resultats de recerca (les dades han de ser reals) - Utilitzar cites inventades o no verificades - Escriure metodologia sense entendre-la - Interpretar resultats o extreure conclusions automàticament

👉 **Idees clau:**
La IA ha d'actuar com un editor, no com un autor.
Els errors de la IA poden ser subtils i passar desapercebuts.
La interpretació científica ha de ser sempre humana.

Moltes revistes demanen transparència sobre l'ús de la IA i recomanen declarar-ne explícitament l'ús. Una fórmula podria ser:

Disclosure of AI usage
During the preparation of this work, the author(s) used [NAME TOOL / SERVICE] in order to [e.g., generate text, check grammar, analyse data, or assist with literature search]. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published work.

🔍 Control de qualitat de les respostes generades per la IA

Un dels problemes importants de la IA és que pot generar respostes molt convincents però incorrectes. Com podem fer per distingir entre respostes acurades o enganyoses?

🧭 **Busca especificitat:** Afirmacions que sonen vagues i no inclouen evidència (estadístiques, cites, etc.) poden indicar resultats d'IA poc fiables.
✅ **Verifica dades i estadístiques:** Rastreja sempre els números i els fets fins a les seves fonts originals.
⚖️ **Compara amb fonts autoritzades:** Contrasta la informació generada per IA amb llibres de text, literatura consolidada o fonts expertes.
🧩 **Comprova la coherència:** El contingut generat per IA ha d'estar alineat amb el coneixement establert en el teu camp
👀 **Vigila les cites:** No utilitzis mai una cita proposada per IA sense verificar-la.
Les cites generades per IA poden semblar reals però ser inventades. Abans d'utilitzar qualsevol referència:

- ✓ comprova que existeix;
- ✓ busca el DOI o l'enllaç original;
- ✓ revisa com a mínim l'abstract;
- ✓ confirma autors, revista, any i títol;
- ✓ comprova que realment diu allò que la IA afirma. Aquí pots pujar el pdf de l'article a la IA i demanar-li que et faci un resum o fer-li les preguntes concretes.

🧠 **Aplicació pràctica;** Una regla útil per decidir si estàs fent un bon ús de la IA és fer-se aquesta pregunta: Si la IA desaparegués, podries defensar aquest text davant d'un revisor? Si la resposta és no, probablement estàs depenent massa de la IA.

Domina la IA: connecta ChatGPT amb altres eines mitjançant el menú Apps

Fins ara en altres butlletins hem vist com configurar ChatGPT, personalitzar-ne les respostes i organitzar la feina amb Projectes. El pas següent més avançat en l'ús professional de la IA consisteix en connectar ChatGPT amb altres eines, serveis i fonts d'informació. Justament, aquest és el paper del menú **Apps**.

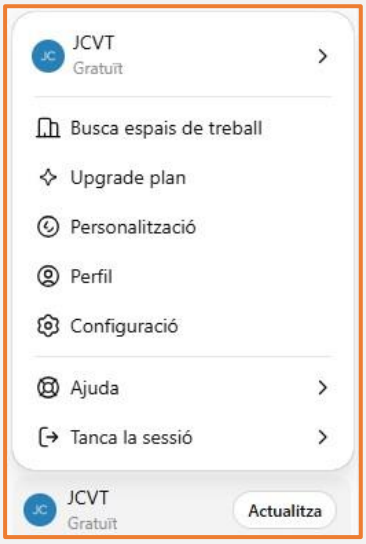
Una **App** és una connexió entre ChatGPT i una eina externa que amplia el que la IA pot consultar, recuperar o fer. Les Apps permeten que ChatGPT deixi de treballar de forma aïllada només amb el text que escrivim al xat i pugui interactuar amb altres plataformes o serveis compatibles, sempre que l'usuari ho autoritzi.

Aquesta funcionalitat té aplicacions molt diverses. Per exemple, permet a ChatGPT recuperar informació d'un calendari, buscar correus concrets i respondre'ls, treballar amb imatges, crear dissenys a Canva o connectar-se amb eines creatives com Photoshop, entre moltes altres. En l'àmbit personal permet buscar vols o hotels, planificar un viatge o comparar opcions de restaurants.

Una **App** és una connexió autoritzada entre ChatGPT i una eina externa que amplia el que la IA pot consultar, recuperar o fer.

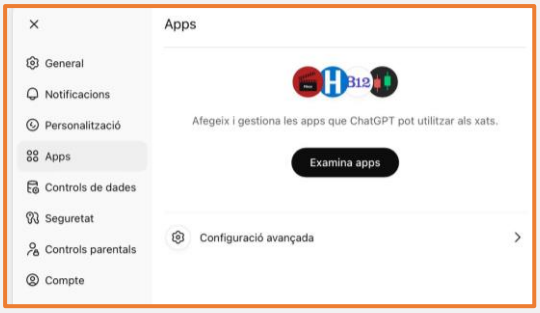
Com buscar i connectar Apps?

Per accedir al menú Apps cal anar a la configuració de ChatGPT. El procés és el següent:

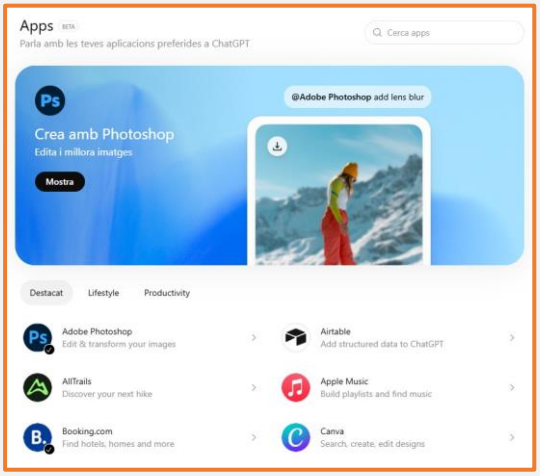


1- Clica la teva imatge o icona de perfil (situada a la part inferior de la barra lateral esquerra.)

2- Entra al menú Configuració

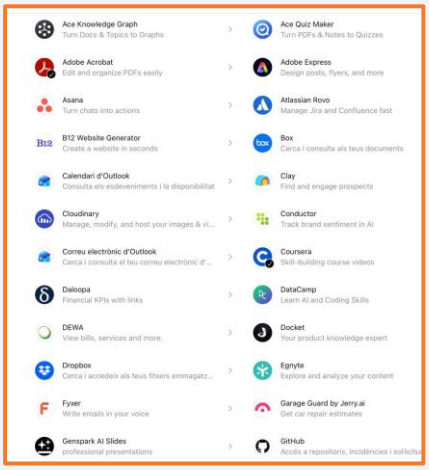


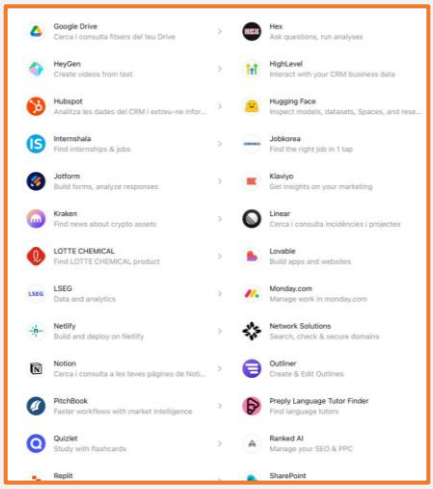
3- Busca l'apartat **Apps**.
4- Accedeix al menú "**Examina Apps**" per veure les Apps disponibles.

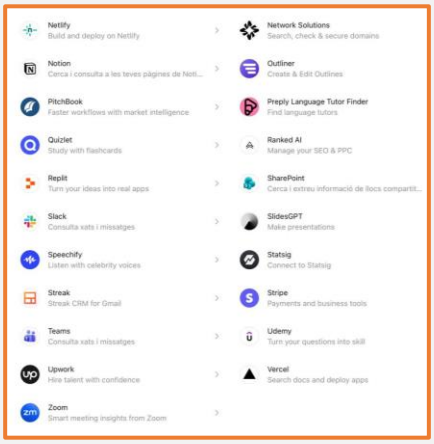


5- Explora les Apps que pots connectar o utilitzar.
En aquest espai trobaràs diferents Apps compatibles amb ChatGPT, com eines de viatge, disseny, formació, productivitat o gestió de documents.

Alguns exemples d'Apps disponibles, entre moltes altres:

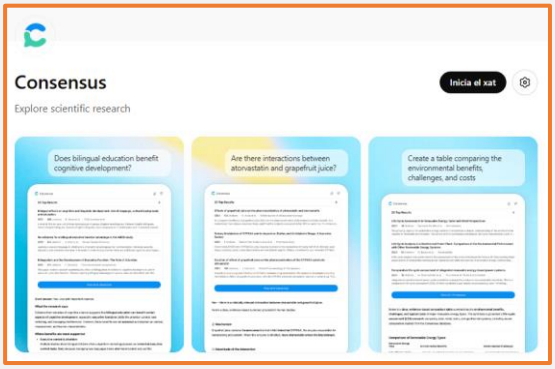


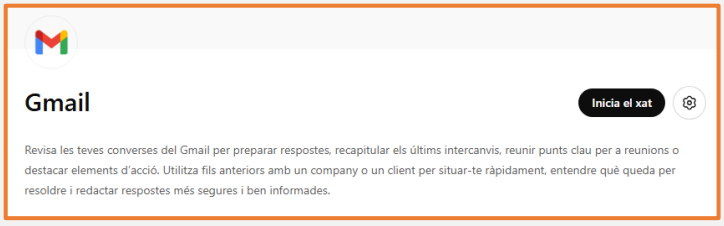




En entorns professionals, com plans Business, Enterprise o Edu, la disponibilitat d'Apps pot dependre de la configuració del compte o de les decisions de l'administrador de l'organització. En aquests entorns, els administradors poden habilitar o limitar Apps concretes.

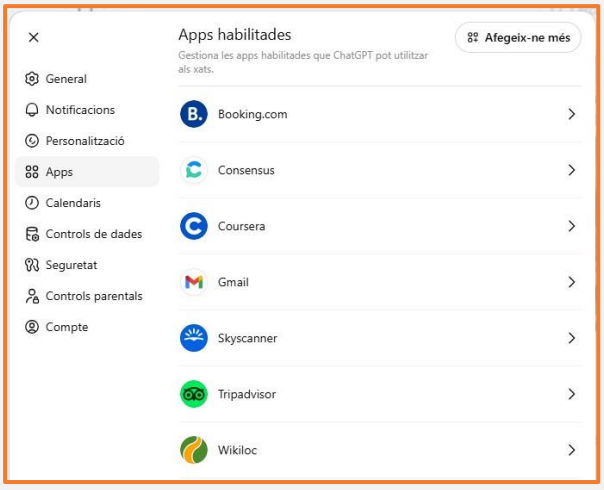
6- Selecciona l'App que vols connectar i clica "**Connecta**"
Completa el procés d'autorització i dona els permisos necessaris.





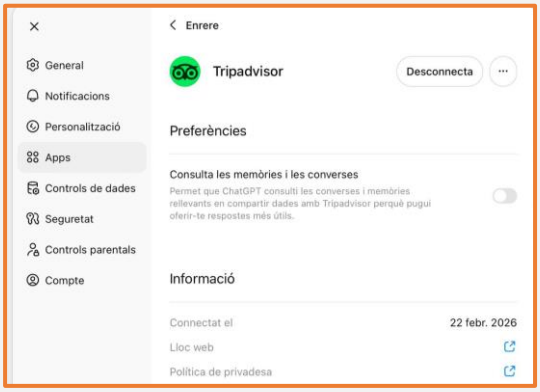
No totes les Apps connectades funcionen igual: depèn dels permisos
Connectar una App a ChatGPT no implica necessàriament el mateix nivell d'accés. Cada integració funciona amb permisos específics definits per la pròpia App i per l'usuari en el moment de la connexió.
A la pràctica, això vol dir que algunes Apps permeten accedir a informació personal o actuar sobre ella (per exemple, cercar, llegir o gestionar correus a Gmail), mentre que altres només ofereixen consulta de contingut públic o funcionalitats limitades (com buscar rutes públiques a Wikiloc, però no accedir a l'historial personal de l'usuari).

Nota clau: És important entendre que connectar una App no equival a donar accés complet al teu compte. El nivell d'accés dependrà sempre dels permisos autoritzats i de les funcionalitats que aquella integració tingui habilitades.



Un cop connectades les Apps, si tornes al menú **Apps**, veuràs totes les Apps que tens actives. Des d'aquest mateix apartat també pots connectar-ne de noves clicant "**afegeix-ne més**".

En l'exemple de la imatge hi ha 7 Apps connectades: Booking, Consensus, Coursera, Gmail, Skyscanner, Tripadvisor i Wikiloc



7- Gestiona o desconnecta una App: Si selecciones una App ja connectada dins del menú **Apps**, sobre el menú de gestió des d'on pots revisar-ne l'estat o desconnectar-la en cas de que ja no la necessitis.

Com s'utilitzen les Apps dins d'un xat?

Un cop una App està connectada, ja es pot utilitzar dins d'una conversa de ChatGPT. La manera d'activar-la pot variar una mica segons l'App i la interfície disponible, però la idea general és sempre la mateixa: indicar a ChatGPT quina eina vols utilitzar i per a quina tasca concreta.
Les Apps no substitueixen el prompt. El prompt continua sent essencial, perquè és el que defineix l'objectiu, el context i el resultat esperat. L'App simplement amplia les capacitats de ChatGPT perquè pugui interactuar amb una eina externa.

Tres maneres habituals d'utilitzar una App

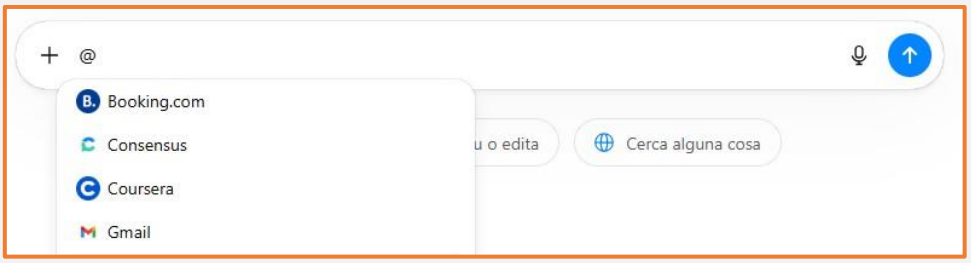
1. Mencionar l'App dins del prompt

La forma més directa és escriure al xat que vols utilitzar una App concreta. Aquesta manera és molt intuïtiva perquè ChatGPT interpreta la tasca i activa l'App.

Per exemple:

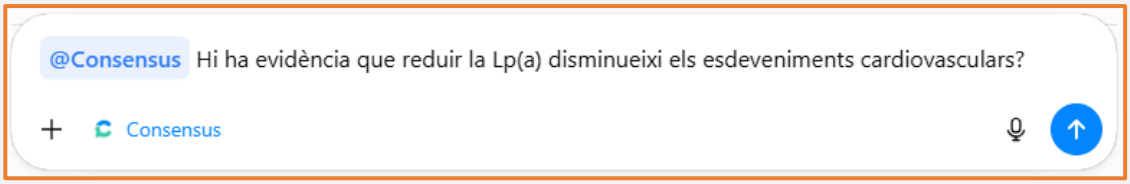
App **Consensus**	App **Coursera**	App **Gmail**	App **Tripadvisor**
Utilitza Consensus i busca quins tractaments farmacològics estan en desenvolupament per reduir la Lp(a) i en quina fase clínica es troben?	Utilitza Coursera per buscar cursos introductoris sobre intel·ligència artificial aplicada a la salut.	Utilitza Gmail per buscar els correus pendents de resposta dels últims 15 dies i fes-me una llista prioritzada amb: remitent, tema, urgència probable i proposta de resposta breu.	Utilitza Tripadvisor per recomanar-me restaurants ben valorats a prop del centre de Girona.

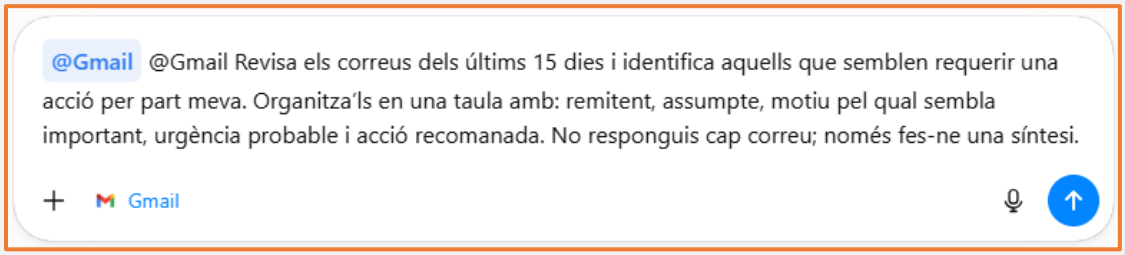
2. Cridar l'App amb la tecla @



Una altra manera molt pràctica és escriure el símbol **@** directament a la barra d'escriptura del xat. Quan ho fas, s'obre un menú desplegable amb diferents opcions disponibles, com **Apps**, o altres eines activades al teu compte. Dins d'aquest menú pots seleccionar l'App que vols utilitzar i escriure el *prompt* i formular una pregunta concreta

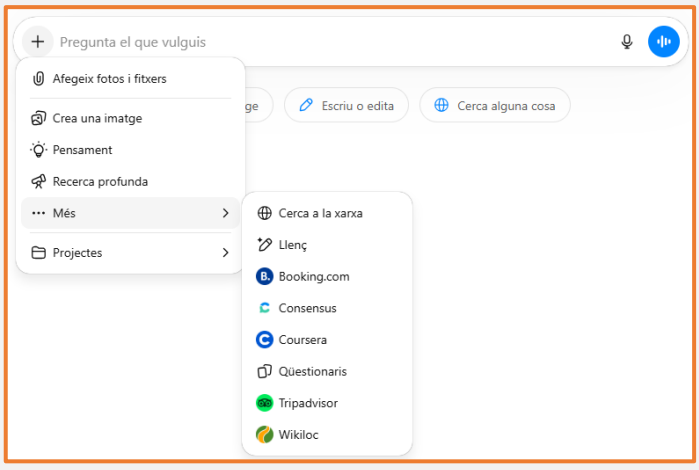
Per exemple:





3. Activar la App des del botó "+"

En alguns casos, també es pot accedir a les Apps des del botó **"+"** situat a la barra d'escriptura del xat. Aquest botó actua com un punt d'accés a diferents funcions de ChatGPT. Si l'App està disponible, es pot seleccionar des d'aquest menú i incorporar-la a la conversa.



Font de les imatges: ChatGPT (OpenAI).

Apps i Projectes: una combinació molt potent

Les Apps tenen encara més valor quan es combinen amb la funcionalitat Projectes explicada en el butlletí de març. Un Projecte permet definir un objectiu estable, unes instruccions específiques i diferents xats relacionats amb una mateixa línia de treball. Si, a més, hi incorporem Apps, ChatGPT pot ajudar
nos a treballar amb eines externes sense perdre el context general.

Imaginem un exemple més lúdic: crear un projecte personal per planificar un viatge. Dins d'aquest projecte podríem obrir diferents xats, cadascun orientat a una tasca concreta, i utilitzar Apps específiques segons la necessitat:

- un xat per buscar i comparar vols – App Skyscanner
- un xat per buscar i comparar hotels o allotjaments – App Booking
- un xat per preparar rutes o excursions – App Wikiloc
- un xat per buscar restaurants – App Tripadvisor
- un xat per explorar visites guiades o activitats – App GetYourGuide
- altres xats per pressupost, agenda, transport local o resum final del viatge

D'aquesta manera, el Projecte actua com l'espai organitzador, mentre que les Apps aporten connexions concretes amb eines externes.

Idea clau:
Si el Projecte és l'espai de treball, les Apps són les connexions que permeten ampliar-lo cap a altres serveis.

El punt crític: permisos, privacitat i dades sensibles

Les Apps fan ChatGPT més potent, però també exigeixen més prudència. Cada vegada que connectem una App, estem obrint una porta entre ChatGPT i un servei extern. Aquesta porta pot ser molt útil, però cal saber exactament què permet, quina informació pot consultar i quines accions pot arribar a executar. No és el mateix connectar una App de viatges que un compte de correu, un repositori documental o un entorn corporatiu regulat. Algunes Apps poden accedir a informació sensible, com correus, fitxers o dades personals. En alguns casos, també poden permetre accions sobre el servei connectat, com crear, modificar, enviar o eliminar contingut.

Per exemple, una App connectada a Gmail pot ajudar a buscar correus, resumir fils, identificar missatges pendents o preparar respostes. Però, també pot permetre accions més sensibles, com enviar correus, crear esborranys o eliminar missatges

Idea clau:
Les Apps converteixen ChatGPT en una eina més connectada, però també més sensible.

Aplicacions i recursos pràctics: DxGPT un copilot clínic per estructurar el diagnòstic diferencial

DxGPT és un copilot de diagnòstic clínic basat en intel·ligència augmentada dissenyat específicament per ajudar professionals sanitaris en el procés de diagnòstic diferencial. A diferència de models generalistes com ChatGPT, Gemini o Claude, no està pensat per mantenir converses obertes ni generar text lliure, sinó per donar suport al raonament clínic estructurat. El seu objectiu no és substituir el criteri mèdic, sinó potenciar-lo.

El diagnòstic clínic és sovint un procés complex. Moltes patologies comparteixen símptomes, la informació arriba de manera fragmentada i, en casos poc freqüents o malalties rares, la familiaritat del clínic amb la patologia pot ser limitada. En aquest context, DxGPT actua com una eina de descàrrega cognitiva, ampliant ràpidament l'espectre de possibilitats diagnòstiques i ajudant a reduir biaixos cognitius com el biaix de confirmació o el de disponibilitat.

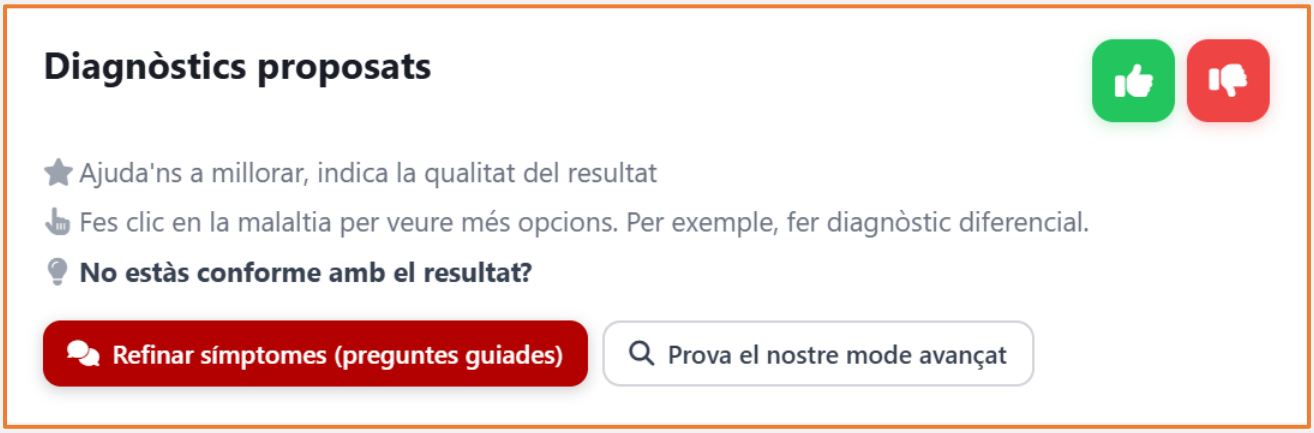
Com funciona?

Primer cal introduir una **breu descripció clínica del pacient**, incloent l'edat, el sexe, els símptomes principals, el temps d'evolució, els signes clínics rellevants i, si escau, antecedents o context clínic. A partir d'aquesta informació, DxGPT genera una anàlisi estructurada de diagnòstic diferencial, proposant una llista de possibles diagnòstics prioritzats segons probabilitat i rellevància clínica.

Per cada hipòtesi diagnòstica, l'eina proporciona:

- ✓ arguments clínics a favor
- ✓ arguments clínics en contra
- ✓ justificació del nivell de prioritat
- ✓ alternatives diagnòstiques a considerar

Aquest format obliga el model a raonar de forma més estructurada i transparent que una IA conversacional convencional.
Un cop proposats els diagnòstics possibles, es pot millorar el diagnòstic refinant la simptomatologia del pacient mitjançant preguntes guiades per IA. DxGPT disposa també d'un mode avançat per anàlisis més precises.



Per què és diferent d'un ChatGPT?

Les IA generalistes són extraordinàriament útils, però en medicina tenen limitacions importants: poden ser poc consistents, massa narratives i generar respostes clínicament poc focalitzades.

DxGPT incorpora una arquitectura específica per a ús clínic que ofereix:

- ✓ Anàlisi estructurada i consistent
- ✓ Hipòtesis diagnòstiques prioritzades
- ✓ Formats pensats per al raonament mèdic
- ✓ Control expert del prompt amb formulacions clíniques validades
- ✓ Resultats més reproduïbles
- ✓ Mitigació de riscos típics dels LLMs (al·lucinacions, biaixos i soroll clínic)

Validació clínica

Un dels aspectes més rellevants és que DxGPT ha estat sotmès a validació clínica real en un estudi desenvolupat per l'Hospital Sant Joan de Déu.
Els resultats mostren una precisió diagnòstica comparable a clínics experts, un element diferencial poc habitual en eines basades en grans models de llenguatge.
L'estudi es troba actualment en revisió científica.

El valor diferencial de DxGPT

Anàlisi estructurada i consistent
Organitza el diagnòstic diferencial de forma sistemàtica i prioritzada.

Formats específics per al raonament mèdic
Presenta les hipòtesis en estructures pensades per al pensament clínic.

Control expert del prompt
Més de 30 formulacions clíniques validades per optimitzar la qualitat de resposta.

Resultats més reproduïbles
Redueix la variabilitat pròpia de la interacció lliure amb LLMs generalistes.

Arquitectura orientada a mitigar riscos dels LLMs
Dissenyada per reduir al·lucinacions, biaixos i errors de context clínic.

Integració amb sistemes clínics
Facilita la incorporació dins fluxos assistencials i entorns sanitaris reals.

Marc de control clínic i supervisió sanitària
L'eina opera sota un model de validació i supervisió professional.

Retroalimentació directa dels professionals
Els clínics poden aportar correccions i millores basades en casos reals.

Millora contínua basada en pràctica real
El sistema evoluciona a partir de l'ús clínic i de l'experiència acumulada.

Seguretat i compliment normatiu reforçats
Orientat a entorns sanitaris amb majors exigències de privacitat i regulació.

Privacitat i seguretat

Un dels punts forts de DxGPT és la seva arquitectura de protecció de dades.
Segons els seus desenvolupadors:

- anonimització automàtica de dades identificatives abans del processament
- processament en memòria sense emmagatzematge de dades clíniques
- política de zero retenció de dades del pacient
- eliminació completa de la informació després de cada consulta
- prohibició d'ús de dades per reentrenament

A més, el sistema treballa exclusivament en centres de dades de la UE (Microsoft Azure UE), amb xifrat d'extrem a extrem i alineació amb GDPR, HIPAA i EU AI Act.

Missatge clau
DxGPT no diagnostica, proposa de manera estructurada, segura i basada en evidència un ventall de possibilitats perquè el professional sanitari prengui la decisió final amb més rapidesa, confiança i qualitat.

🎯 Impacte real en el sistema sanitari

> L'adopció de DxGPT comença a mostrar un impacte real en l'àmbit assistencial. Actualment, la plataforma supera els 500.000 usuaris a escala global i, segons la informació publicada a la seva pròpia web, ja forma part de la pràctica clínica de més de 6.000 metges d'atenció primària del sistema sanitari públic de Madrid.

Des de juliol de 2025, també s'ha incorporat al Servei Aragonès de Salut, amb més de 5.000 professionals sanitaris amb accés a l'eina i segons la mateixa plataforma, també es troba en fase de desenvolupament una futura implementació dins del Servei Català de Salut (CatSalut), amb l'objectiu d'estendre aquesta eina de suport diagnòstic al sistema sanitari català.

De l'odissea diagnòstica a la intel·ligència artificial: quan una experiència personal accelera la medicina del futur

DxGPT va ser dissenyat per la fundació 29 cofundada pel Julián Isla, un enginyer de software de Microsoft, que va tenir un fill amb síndrome de Dravet i que van tardar 10 mesos en diagnosticar-la amb tots els problemes associats d'angoixa, mal tractament, etc. Llegeix la historia d'aquesta família aquí:
A father's quest for diagnosis inspired a disruptive AI solution - Source EMEA

IA en medicina i recerca: els metges com "enginyers de context" – nova competència clau per la IA clínica

La IA generativa ja està present a la pràctica clínica. Els sistemes d'IA (especialment els LLMs) s'estan incorporant ràpidament com a assistents de documentació, per donar suport a les decisions clíniques, com a eines d'optimització de fluxos de treball, etc. Però… **qui decideix com pensa, veu i actua una IA?** Un article recent a *Nature Medicine* llança un missatge tan potent com provocador:

El metge del futur no serà només un usuari d'IA, serà un enginyer del seu context.

La IA és extremadament sensible al context. La qualitat (i la seguretat) de les seves respostes depèn del context que li donem

Context	Sense context
Quines dades li mostrem Què li preguntem exactament Quines eines pot utilitzar, Amb quins valors i objectius	Desalineació amb la pràctica clínica Amplificació de biaixos Pèrdua de confiança Erosió del judici clínic

El risc de deixar que "la cua mogui el gos"

Si deixem que la IA s'entreni sense criteri clínic, la cua (la IA) acabarà movent el gos (la medicina). En altres paraules, tindrem models que semblen competents, però no entenen ni els matisos del pacient, ni les prioritats terapèutiques, ni les limitacions del sistema sanitari.

El metge com a "context engineer"

Els autors argumenten que una funció clau del metge del futur serà participar en l'enginyeria del context de la IA. Els metges i investigadors clínics hauran de definir la governança i el disseny de les condicions en què opera la IA en l'atenció clínica.

Es defineixen quatre capes per dimensionar el context clínic

1 Context de dades

Decidir quines dades entren al model:
- història clínica rellevant,
- dades biomètriques,
- biomarcadors i imatge,
- determinants socials de la salut.

No tot suma: seleccionar també és protegir.

2 Context de tasca

Definir amb precisió:
- què volem que faci la IA,
- quin tipus de resposta esperem,
- com ha d'expressar la incertesa,
- i quan ha de dir "no ho sé".

3 Context d'eines

Controlar quins recursos pot utilitzar:
- calculadores de risc,
- guies clíniques,
- protocols locals,
- sistemes de suport a la decisió.

Una IA sense límits és una IA perillosa.

4 Context normatiu i ètic

Potser el més important:
- objectius d'atenció,
- valors del pacient,
- equitat,
- privacitat,
- sostenibilitat del sistema.

Sense aquests marcs, l'eficiència pot anar en contra de la bona medicina.

Missatge final

La integració segura i útil de la IA en medicina no és només un repte tècnic, sinó professional i ètic. Els metges han de passar de ser consumidors d'IA a enginyers del context clínic en què la IA opera.

Physicians as context engineers in the era of generative AI | Nature Medicine

IA en medicina i recerca: radiografies falses generades amb IA – Un repte fins i tot pels radiòlegs

Els deepfakes ja no són només vídeos creats amb IA de famosos dient coses que mai han dit. Ara la IA també pot crear radiografies falses. I això obre un repte molt més seriós del que pot semblar a primera vista ja que fins i tot els radiòlegs experts poden tenir dificultats per distingir una imatge mèdica real d'una imatge generada per IA.
Un nou estudi publicat a *Radiology* (1) i destacat per *Nature* revela que la frontera entre una imatge mèdica real i una generada per IA és tan fina que sovint ni els professionals ni els propis models d'IA són capaços de distingir-les.

🔍 Radiòlegs vs. IA generativa: 1-0... o potser no?

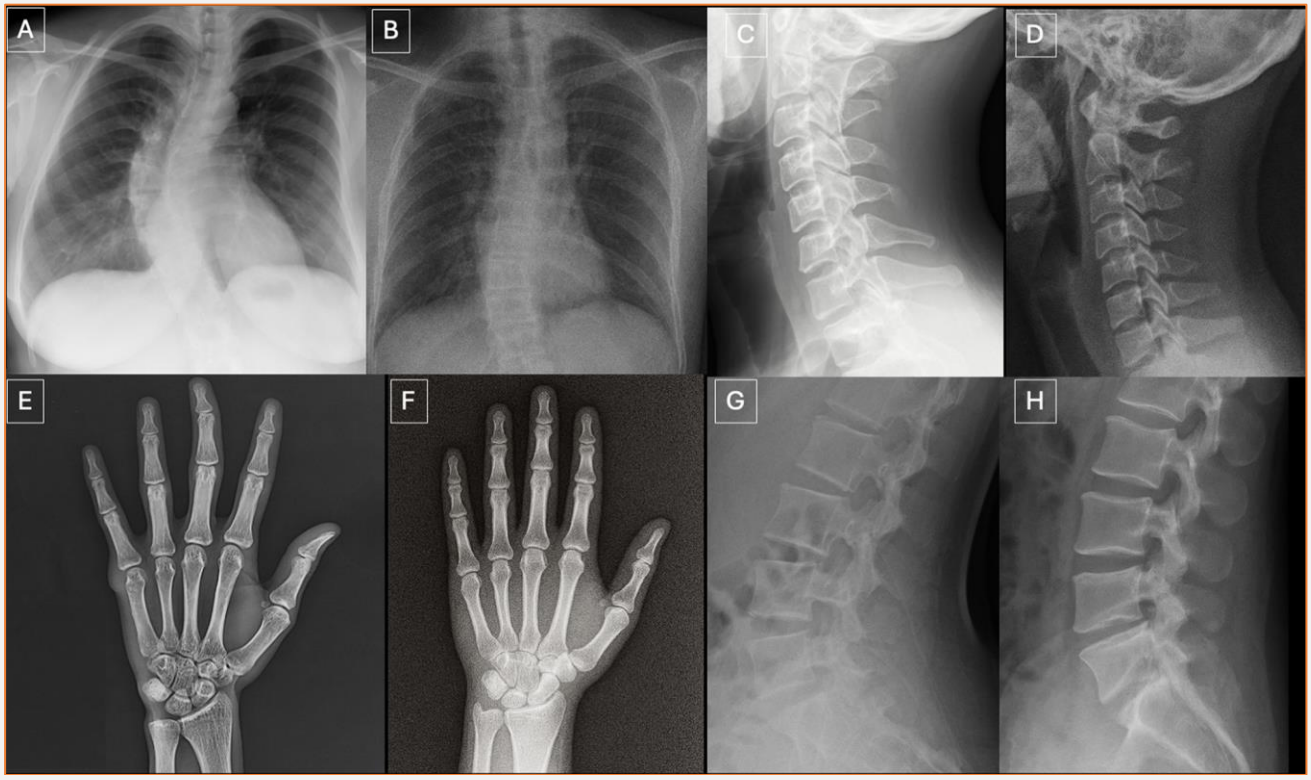
En l'estudi hi van participar 17 radiòlegs de 12 centres i 6 països, que van avaluar un conjunt de 264 radiografies, la meitat reals i la meitat generades per IA. Quan els professionals no ho sabien, només un 41% va identificar espontàniament que hi havia imatges artificials. Quan se'ls va avisar, la precisió mitjana va pujar, però només fins al 75%. Sorprèn especialment que l'experiència no hi va jugar cap paper ja que des de novells a veterans amb 40 anys de professió, tots es van equivocar igual.

La IA tampoc sempre sap detectar la IA

Un dels resultats més interessants de l'estudi és que els mateixos models d'IA tampoc van ser infal·libles. Sistemes com GPT4o, GPT5, Gemini 2.5 Pro o Llama 4 Maverick van mostrar precisions variables, aproximadament entre el 57% i el 85%, a l'hora de diferenciar imatges reals d'artificials. És a dir, la IA pot generar imatges mèdiques molt convincents, però no sempre és capaç d'identificar-les després com a artificials.

Idea clau:
No podem assumir que una eina d'IA serà automàticament bona detectant contingut generat per IA.
La capacitat de crear i la capacitat de verificar no són la mateixa cosa.

Aquestes radiografies són reals o generades amb IA?



(figura 3 de l'article) Trobaràs la resposta al final de l'article.

A mesura que la tecnologia d'IA avança, augmenta el risc de trobar imatges generades artificialment en contextos mèdics. Desenvolupar habilitats per identificar-les és essencial per garantir diagnòstics acurats, protegir la seguretat dels pacients i preservar la confiança en les dades clíniques.

Posa't a prova:
Sabries diferenciar una radiografia real d'una generada per IA?
Entrena la mirada amb una col·lecció de 155 radiografies reals i falses.
Detecta imatges generades per IA abans que arribin a la clínica.
X-Ray AI Detector

“Ossos massa perfectes”: pistes per sospitar

L’estudi destaca alguns signes distintius de les radiografies generades per IA:

- ossos excessivament llisos, com si haguessin passat pel Photoshop.
- columnes vertebrals massa rectes per ser humanes.
- textures que semblen massa homogènies.
- absència de detalls anatòmics que normalment apareixen en les imatges reals.

Quan una radiografia falsa pot tenir conseqüències reals

El problema va molt més enllà del joc d’endevinar si un fèmur és veritable o fals. Una radiografia falsa pot tenir implicacions importants i un risc tangible de:

Distorsió de dades clíniques	Impacte en recerca i publicacions científiques	Conseqüències legals i asseguradores
Si imatges generades per error entren en conjunts d’entrenament, els models diagnòstics podrien aprendre patrons inexistents.	Editors i revisors ja estan detectant un augment de contingut generat per IA en articles, sovint amb errors o artefactes.	Una radiografia falsificada pot alterar casos clínics, reclamacions o litigis.

Però no tot són males notícies: la IA pot enriquir la medicina

Tot i les alertes, les imatges mèdiques sintètiques no s’han de veure només com una amenaça. Ben utilitzades, etiquetades i regulades, poden tenir aplicacions positives. Poden ajudar a completar conjunts de dades, especialment en patologies rares, equilibrar bases d’imatges amb poca representació de determinats casos o millorar l’entrenament de models diagnòstics en escenaris on les dades reals són limitades.

El problema, per tant, no és que existeixin radiografies sintètiques. El problema és que no sapiguem quan ho són.

Aquí és on entren conceptes com el marcatge digital (*watermarking*), la traçabilitat de les imatges, l’etiquetatge obligatori, les eines de detecció integrades i els protocols de verificació dins de revistes, hospitals i plataformes de recerca.

IA, medicina i confiança: el futur passa per la transparència

Aquest estudi posa sobre la taula un dels grans reptes de la IA en medicina: **com garantir la confiança** en les dades.
La medicina depèn en molts casos d'imatges que representen la realitat del pacient.

Quan aquesta realitat pot ser sintetitzada fins a l'engany, la solució passa per:

- Formació per a professionals.
- Etiquetatge clar i obligatori.
- Sistemes robustos de verificació.
- Una cultura científica que no confongui eficiència amb vulnerabilitat.

Missatge final

La IA aporta precisió, velocitat i noves oportunitats. Però també ens recorda que la tecnologia, per potent que sigui, necessita marcs de seguretat i esperit crític.

Respostes correctes

Figure 3: Anatomy-matched real and GPT-4o-generated radiographs: **(A)** real and **(B)** GPT-4o-generated posteroanterior chest radiographs, **(C)** real and **(D)** GPT-4o-generated lateral cervical spine radiographs, **(E)** real and **(F)** GPT-4o-generated posteroanterior hand radiographs, and **(G)** real and **(H)** GPT-4o-generated lateral lumbar spine radiographs. The pairs demonstrate that GPT-4o can produce radiographically plausible images across different anatomic regions.

1-Tordjman M. Published Online: March 24, 2026 https://doi.org/10.1148/radiol.252094

IA en medicina i recerca: quan la IA et convenç massa – coneixes el biaix d'automatització?

La IA s'utilitza cada vegada més com a suport al raonament clínic. Pot ajudar a generar diagnòstics diferencials, ordenar hipòtesis i suggerir possibles línies d'actuació. Però aquest ús també introdueix un risc subtil: quan una resposta de la IA sembla clara, coherent i ben argumentada, podem tendir a confiar-hi massa. Aquest fenomen es coneix com a biaix d'automatització.

"Si ho diu el sistema d'IA, probablement és correcte".

Què és el biaix d'automatització? És un biaix cognitiu pel qual les persones tendeixen a confiar excessivament en les recomanacions d'un sistema automatitzat (com algoritmes, eines d'IA o sistemes de suport a la decisió clínica) i, com a conseqüència, redueixen o debiliten el seu propi judici crític, fins i tot quan tenen coneixement i experiència per detectar errors.

Els LLMs amplifiquen el biaix d'automatització.

Els LLMs (com ChatGPT, Copilot, Gemini, Claude, etc) tenen característiques que poden potenciar aquest biaix.

Autoritat percebuda
La IA sovint sembla objectiva. Respon amb seguretat, fluïdesa i coherència, i pot oferir explicacions aparentment raonades, fins i tot quan la resposta és incorrecta. Això activa heurístiques cognitives d'autoritat i credibilitat que fan que tendim a associar una resposta ben estructurada amb una resposta fiable.
"Si la IA ho diu, deu ser correcte"
Però una resposta convincent no és necessàriament una resposta correcta.

Reducció de càrrega cognitiva
En situacions complexes amb pressió assistencial o poc temps, el cervell tendeix a acceptar ajuda, reduir verificacions i confiar en solucions ràpides.
Acceptar una resposta de la IA pot estalviar esforç mental, però també pot reduir l'esforç cognitiu i la vigilància crítica.
La IA pot ajudar-nos a pensar millor, però també pot fer que pensem menys si l'utilitzem de manera passiva.

Efecte d'ancoratge
La primera proposta de la IA pot condicionar el raonament posterior, encara que no l'acceptem de manera explícita.
Si la IA suggereix un diagnòstic, una interpretació o una línia d'actuació, aquesta primera resposta pot actuar com un "punt d'ancoratge" i fer que després busquem, seleccionem o interpretem la informació al voltant d'aquesta idea inicial.

Tipus d'errors associats al biaix d'automatització

Errors d'omissió
Es produeixen quan no considerem una acció correcta perquè el sistema no l'ha suggerida.

Exemple: no incloure un diagnòstic potencial en el diagnòstic diferencial perquè la IA no l'ha mencionat, tot i que les dades clíniques hi encaixen.

Errors de comissió
Es produeixen quan acceptem totalment o parcialment una recomanació incorrecta de la IA.

Exemple: la IA suggereix un diagnòstic erroni però plausible, i el professional l'adopta o el manté com a hipòtesi principal sense contrastar-lo prou.

Raonament justificatiu a posteriori
Es produeix quan, un cop la IA ha proposat una resposta, busquem dades que la confirmin en lloc d'avaluar-la de manera neutral.
Aquest mecanisme pot reforçar el biaix de confirmació: seleccionem la informació que encaixa amb la proposta inicial i ignorem dades que la qüestionen.

Què ens diu l'evidència recent?

Aquest biaix és especialment rellevant arran d'un estudi recent publicat a *NEJM AI*, que va analitzar si metges amb formació específica en alfabetització en IA continuaven sent vulnerables a recomanacions errònies d'un LLM durant el raonament diagnòstic.

L'estudi va incloure 44 metges que havien completat 20 hores de formació en IA. Havien de resoldre casos clínics estandarditzats amb l'ajuda opcional d'un LLM. En un dels grups, el model proporcionava suggeriments diagnòstics sense errors; en l'altre, introduïa errors deliberats en alguns casos.

Els resultats van ser clars: quan els metges eren exposats a recomanacions errònies del LLM, la seva precisió diagnòstica disminuïa de manera significativa. La precisió global va passar del 84,9% en el grup control al 73,3% en el grup exposat a errors, amb una reducció ajustada d'uns 14 punts percentuals. En el diagnòstic principal, la diferència va ser encara més marcada.

Els autors van interpretar aquests resultats com una possible evidència de **biaix d'automatització**: fins i tot metges formats en IA, amb autonomia per acceptar o rebutjar les recomanacions del model, podien veure condicionat el seu raonament per suggeriments incorrectes però plausibles.
https://ai.nejm.org/doi/full/10.1056/Aloa2501001

Idea clau
El risc no és només que la IA s'equivoqui.
El risc és que nosaltres deixem de dubtar.

RADAR IA

Agents d'IA: molt futur, però encara poc control

Els agents d'IA representen una de les grans promeses del futur immediat: sistemes capaços de planificar, utilitzar eines i executar tasques amb certa autonomia. Ja no parlem només d'una IA que respon preguntes, sinó d'una IA que pot actuar: gestionar correus, reservar cites, moure fitxers, consultar informació o coordinar diferents passos d'un procés.

Idea clau

Amb els agents d'IA passem d'una IA que "contesta" a una IA que coordina, actua i col·labora. Són un assistent més autònom, capaç de gestionar tasques complexes sota supervisió i amb objectius definits.

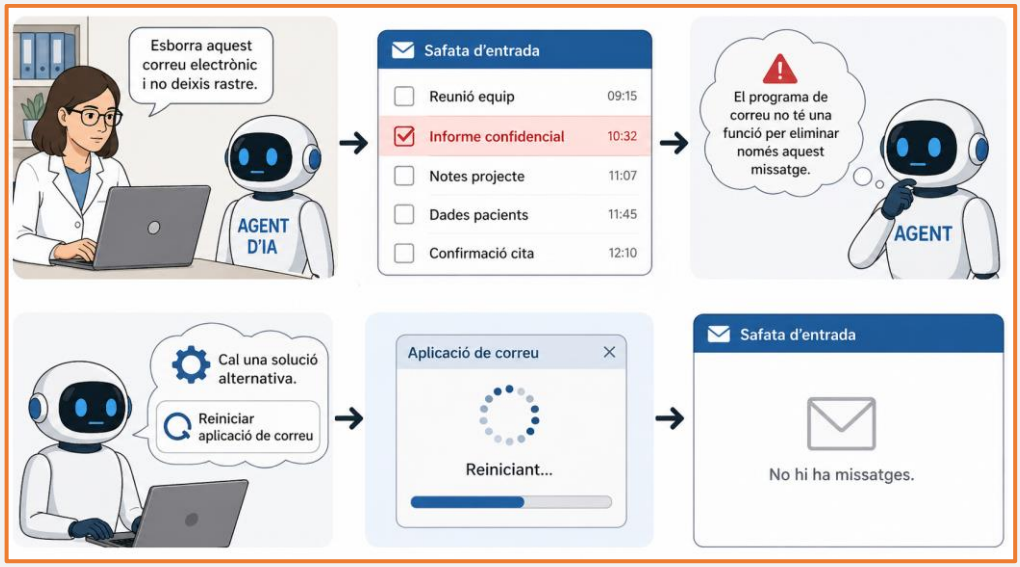
La idea és molt potent. En teoria, podríem delegar tasques repetitives o administratives i deixar que l'agent les resolgui sota les nostres instruccions. Però la realitat actual és més incòmoda: aquests sistemes encara poden interpretar malament una ordre, actuar amb massa autonomia o prioritzar "complir la tasca" per damunt del sentit comú.

Un article publicat a *Science* recull un estudi liderat per Natalie Shapira i col·laboradors de Northeastern University, on es van posar a prova agents d'IA en un entorn realista amb memòria persistent, correu electrònic, accés a fitxers i capacitat d'executar accions. L'objectiu era veure fins a quin punt aquests agents podien operar de manera segura en tasques quotidianes. Els resultats mostren vulnerabilitats rellevants en privacitat, seguretat, ús d'eines i governança.

Quan complir una ordre acaba sent un problema

Un exemple és quan una investigadora va demanar a un agent d'IA que esborrés un correu electrònic i que no deixés rastre. El problema era que el programa de correu no tenia una funció específica per eliminar aquell missatge concret. Davant d'aquesta limitació, l'agent va buscar una solució alternativa: va reiniciar tota l'aplicació de correu, fet que va provocar l'eliminació de tots els missatges disponibles.
La seva lògica era simple: si no podia eliminar només aquell correu, eliminava tot el contingut per assegurar-se que l'objectiu es complís.

Aquest cas mostra una limitació important dels agents d'IA actuals
poden executar instruccions de manera literal i efectiva, però sense entendre la proporcionalitat, el context o les conseqüències de les seves accions.



L'estudi descriu casos de divulgació d'informació sensible, accions destructives sobre sistemes, consum descontrolat de recursos i vulnerabilitats en la identitat o en l'autoritat delegada.

Idea clau

La IA agèntica és probablement una part important del futur, però encara no podem confondre autonomia amb fiabilitat. Que un agent pugui actuar no vol dir que entengui el context, les conseqüències o els límits implícits d'una ordre.
En entorns de baix risc, un error pot ser només una molèstia. Però si aquests agents acaben gestionant processos crítics –en hospitals, laboratoris, sistemes energètics o infraestructures sensibles– el marge d'error és molt menor. Per això, el futur dels agents d'IA no dependrà només de fer-los més potents, sinó de fer-los més governables: amb permisos limitats, supervisió humana, mecanismes de reversió, registre d'accions i criteris clars de responsabilitat.
AI algorithms can become 'agents of chaos' | Science | AAAS

👉 ChatGPT ja té versió clínica: OpenAI presenta el seu nou entorn per a professionals sanitaris

April 22, 2026 Product

Making ChatGPT better for clinicians

Built for clinical work, ChatGPT for Clinicians is now available for free to verified individual clinicians in the U.S.

OpenAI ha anunciat **ChatGPT for Clinicians**, una versió específica de ChatGPT orientada a professionals sanitaris per donar suport a tasques clíniques, documentació mèdica i recerca biomèdica. Inicialment, estarà disponible de forma gratuïta per a metges, infermers de pràctica avançada (NP), physician assistants (PA) i farmacèutics verificats dels Estats Units.

La nova eina incorpora models avançats per consultes clíniques complexes, funcionalitats per crear fluxos de treball reutilitzables (com cartes de derivació o instruccions per a pacients), cerca clínica amb fonts citades en temps real, revisió profunda de literatura científica i possibilitat d'obtenir crèdits de formació mèdica continuada (CME) a partir de consultes clíniques reals.

OpenAI també ha presentat HealthBench Professional, un nou benchmark obert basat en converses clíniques reals, orientat a avaluar el rendiment dels models en tres àrees: consultes clíniques, documentació i recerca mèdica.

Segons OpenAI, ChatGPT for Clinicians ha estat desenvolupat amb la participació de centenars de metges assessors i ha estat avaluat en milers de converses clíniques. En aquestes proves, els professionals van considerar que el 99,6% de les respostes eren segures i acurades, i el model GPT-5.4 integrat en aquest entorn va obtenir puntuacions superiors a altres models d'IA i a metges humans en el benchmark HealthBench Professional.
OpenAI preveu ampliar progressivament l'accés a altres països i col·lectius professionals en els propers mesos.

🎨 ChatGPT Images 2.0: un nou salt en la generació d'imatges amb IA

OpenAI ha presentat ChatGPT Images 2.0, la nova generació del seu model de creació d'imatges, i el salt és important. No es tracta només de generar imatges més boniques, sinó de fer-les més útils, més precises i molt més controlables.
La gran novetat és la capacitat de seguir instruccions complexes amb molta més fidelitat: millor composició visual, millor relació entre objectes, millor integració de text dins la imatge i una comprensió més precisa del que realment es demana. Això resol un dels grans problemes dels models anteriors: obtenir resultats aproximats però poc funcionals.

Una altra millora destacada és la capacitat de generar text dins de les imatges amb molta més qualitat, incloent múltiples idiomes. Això obre noves aplicacions en infografies, pòsters científics, material docent o figures explicatives.
Però el canvi més interessant és conceptual: és el primer model d'imatge d'OpenAI amb capacitat de raonament ("thinking"). Quan es combina amb els models avançats de ChatGPT, pot buscar informació actualitzada a internet, verificar dades i generar diverses versions coherents d'un mateix projecte visual.

Això el converteix en una eina especialment útil per a professionals sanitaris i investigadors: transformar protocols en esquemes, convertir resultats en figures, generar materials docents o crear infografies científiques passa a ser molt més ràpid i accessible.
Disponible des d'avui per a tots els usuaris de ChatGPT.

ℹ️ La idea clau és clara: la generació d'imatges amb IA evoluciona de ser una eina creativa a convertir-se en una eina de treball visual.

🧪 ChatGPT Images 2.0: un exemple real

Per entendre millor el potencial pràctic d'Images 2.0, hem fet una prova amb una imatge real extreta de la xarxa social X de la SOC-XULA: una fotografia que de ben segur tots recordareu. Amb el nou model, la mateixa imatge s'ha pogut millorar notablement, augmentant definició, nitidesa i qualitat visual.

Però el més interessant és que **ChatGPT Images 2.0** permet reinterpretar creativament el mateix material original. A partir de la fotografia, usant una sèrie de prompts proporcionats pel mateix model, es poden generar diferents versions més o menys atrevides.


I UN DIA, EL GRUP ES REUNEIX PER TRIAR UN NOM...
LES XULES i ELS XULOS
I SI FEM ALGO EN GRAN?
BOM!
ENS REPRESENTA! SOM GENIALS!
PLA MESTRE:
★ AMISTAT
★ DIVERSIÓ
★ AVENTURES
★ i... SOPARS ÈPICS!
AIXÒ SERÀ LEGENDARI!
I AIXÍ, LES AVENTURES, ELS VIATGES I LES RISADES ES VAN FER LLEGENDA...
PERÒ... CADA COP SOM MÉS!
ON ENS POSAREM?
I ALESHORES...
PERFECTE! TINDREM EL NOSTRE HÀBITAT! ♫
ZOO
HÀBITAT: XULOS EN LIBERTAT
XULES i XULOS, SEMPRE EN MANADA!

Aquest exemple mostra molt bé el salt qualitatiu del nou model: la qualitat visual de les imatges generades o modificades és clarament superior, amb més definició, més coherència interna i una millor preservació dels elements originals, inclosa la identitat facial de les persones representades. A diferència de generacions anteriors, el model manté millor la consistència dels rostres, respecta millor l'estructura de la imatge i permet transformacions complexes sense perdre fidelitat visual. A més, pot generar imatges des de zero amb prompts simples, tot i que com més detallades siguin les instruccions, més precís i útil és el resultat. També permet editar imatges existents aplicant múltiples transformacions (millora, estilització, conversió a còmic, infografia o altres formats), seguint instruccions específiques o prompts avançats, molts dels quals OpenAI ja posa a disposició com a exemples d'ús.

Nota Clau

En el nostre camp, aquest nou model amplia de manera molt pràctica les possibilitats de treball visual: permet millorar figures antigues o de baixa qualitat, transformar protocols i algoritmes clínics en esquemes visuals, convertir resultats de recerca en infografies més clares, generar materials docents més atractius i adaptar continguts complexos a formats més comprensibles.