

INNOVACIÓ O RETROCÉS?

UNA MIRADA ALS LLM DES DE LA FILOSOFIA DE LA TECNOLOGIA

Núria Vallès Peris

UAB – Universitat Autònoma de Barcelona

Entendre com funciona la intel·ligència artificial generativa ens dona l'oportunitat de reflexionar i discutir sobre com aquesta tecnologia pot afectar la nostra comunicació.

Aquesta conversa és important per decidir junts com volem que es desenvolupin o es limitin les potents tecnologies d'IA.

A

B

La pregunta inicial d'aquest capítol és irònica, ja que l'oposició entre «innovació» i «retrocés» és un fals dilema: d'una banda, presenta només dues opcions com a úniques, sense considerar altres possibles respostes; i, d'altra, presenta aquestes dues opcions com a mútuament excloents, sense tenir en compte que «innovació» i «retrocés» també podrien anar juntes.



No podem aturar la innovació.
Ara bé, tampoc hem d'assumir que tot és inevitable.

HEM DE PODER DISCUTIR SOBRE QUINES SÓN LES TECNOLOGIES DESITJABLES PER AL BÉ COMÚ,

AQUELLES QUE MILLOR RESPONEN A LES NECESSITATS DE LES PERSONES I EL MÓN EN QUÈ VIVIM.

Aquesta discussió no s'hauria de mantenir desde posicions binàries; la valoració de les tecnologies depèn del criteri, entorn i punt de vista dels actors involucrats (responsables polítics, desenvolupadors, usuaris, etc.).



PER TANT, QUALSEVOL VALORACIÓ HAURIA DE SER SEMPRE DISCUTIBLE.

El *hype* de la IA

P.2

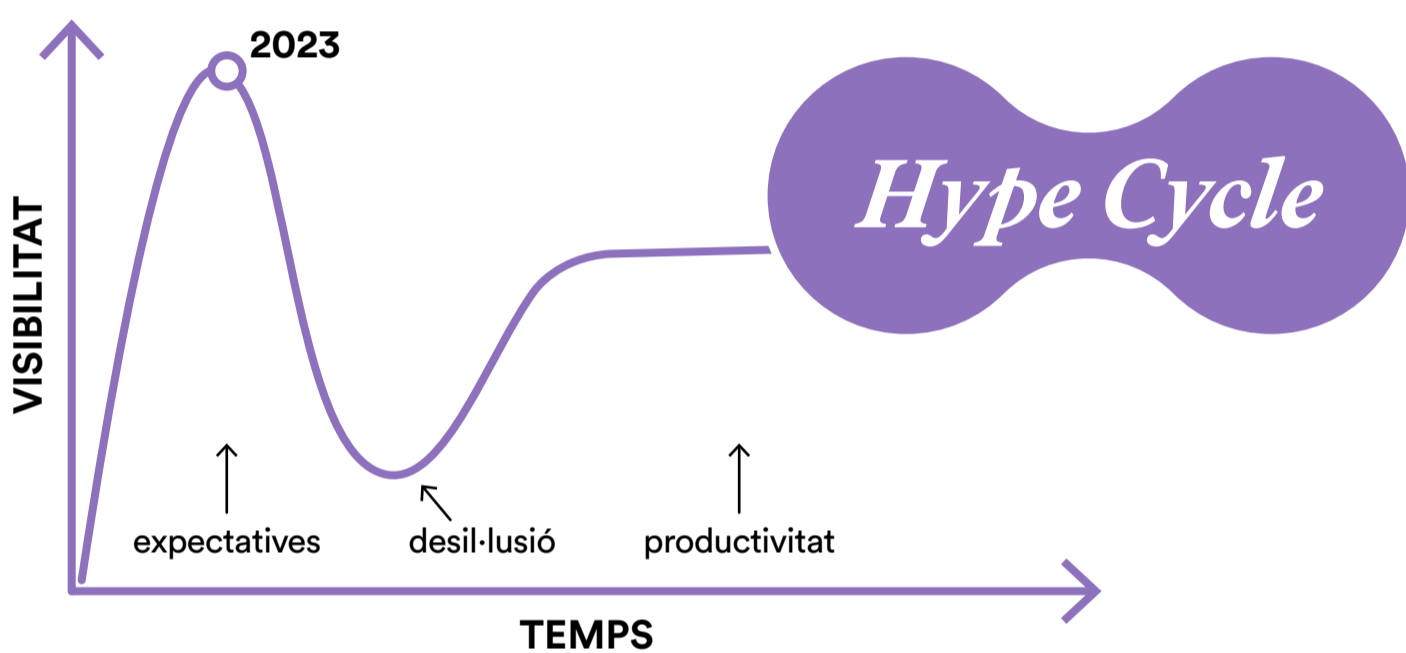


Amb el ràpid desenvolupament de la IA generativa, com ChatGPT, i la seva progressiva aplicació en múltiples àmbits de la nostra vida, han proliferat també els debats al seu voltant.

Es parla d'una revolució històrica que transformarà la salut, l'educació, el treball, l'art...

Alhora, també s'apunten una sèrie de riscos molt greus, com la desaparició de la creativitat i d'altres qualitats humanes o la substitució dels humans per màquines.

Ara bé, tot i l'interès que generen aquests tipus de debats, la seva principal característica és la tendència a l'**exageració** i l'**especulació**, el que es coneix en anglès com a *AI hype*.



Tecnoutòpiques o tecnofòbiques?

UNA DE LES CONSEQÜÈNCIES DE L'*AI HYPE* ÉS QUE LES POSICIONS ES POLARITZEN ENTRE TECNOUTÒPIQUES I TECNOFÒBIQUES: **INNOVACIÓ O RETROCÉS?**



Hem d'abraçar acríticament el desenvolupament de qualsevol tecnologia en nom de la innovació i el progrés, fent ulls clucs a les seves conseqüències?

O, al contrari, l'única solució és aturar màquines, frenar el desenvolupament tecnològic i tornar a una societat preindustrial?

Ni una cosa ni l'altra. Però no és senzill reflexionar i discutir de manera raonada en l'escenari de constant exageració i manca de transparència de les grans empreses tecnològiques.



En aquest impàs, la **filosofia de la tecnologia** és una mena de **caixa d'eines plena de conceptes** que ens poden ser d'utilitat per pensar críticament, més enllà de la publicitat i la fantasia, sobre ChatGPT o altres tecnologies.

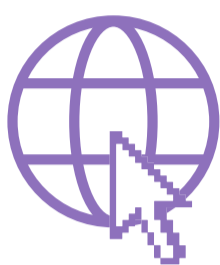
La tecnologia resol tots els nostres problemes?

P.3

Tot i els beneficis potencials de la IA generativa, cal reconèixer que el seu desenvolupament va acompanyat de la narrativa (discutible) que la tecnologia pot resoldre tots els problemes.



D'AIXÒ SE'N DIU SOLUCIONISME TECNOLÒGIC



El solucionisme tecnològic, també conegut com a *imperatiu tecnològic*, considera que certs problemes complexos (com el canvi climàtic, les epidèmies o l'educació) tenen una solució senzilla a partir de l'aplicació de tecnologies i algorismes.

I AIXÒ OBRE EL DEBAT SOBRE LA RESPONSABILITAT

Les discussions sobre l'adopció de tecnologies sovint tendeixen, sobretot en el cas de la IA, a centrar-se en enlluernadores (i hipotètiques) aplicacions futures.

Com a conseqüència, la tecnologia pot arribar a percebre's com una força imparable a la qual no queda cap altra opció que adaptar-se.

Sense una anàlisi del desenvolupament de les aplicacions actuals, aquesta percepció no té en compte les necessitats de la població ni els riscos que aquestes tecnologies puguin suposar.



2

Aquesta dinàmica especulativa responsabilitza exclusivament la tecnologia d'una sèrie de fenòmens socials, quan, en realitat, aquests fenòmens són fruit d'una sèrie de decisions i dinàmiques que es produeixen en diversos àmbits.

La responsabilitat de la tecnologia

Per exemple, quan els informatius ens diuen que una empresa ha acomiadat una quantitat important de la plantilla per culpa de la utilització de la IA generativa:



Es responsabilitza la tecnologia del fet que molta gent es quedi sense feina, com si la tecnologia sola hagués decidit el destí d'un munt de persones.

Però...

Hi ha altres actors, a part de la IA generativa, que puguin haver incidit en aquesta decisió?

Era possible una altra alternativa?



És evident que la responsabilitat del que ha succeït no és exclusiva de la IA. Aquest tipus de qüestions ens porten a fer-nos altres preguntes i reflexions.

Són *neutrals*, les tecnologies?

P.4



Des de la visió tradicional de la ciència i la tecnologia, se suposa que els descobriments científics condueixen a innovacions tecnològiques i generen progrés social.

Des d'aquesta visió, es creu que la ciència és independent i es pot distingir de qüestions polítiques, interessos econòmics o altres àmbits de l'activitat humana.

AQUESTA ÉS LA VISIÓ DE LA FILOSOFIA RACIONALISTA O NEOPOSITIVISTA

A finals dels anys 60 i principis dels 70 del segle xx, la imatge tradicional de l'autonomia i la neutralitat del coneixement científic va entrar en un període de profunda crisi.



Ni bones, ni dolentes

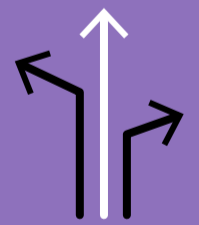
En aquesta crisi es van posar de manifest algunes de les conseqüències més negatives del desenvolupament científic i tecnològic, i els riscos que suposaven per a la salut humana.



VAN APARÈIXER:

- el moviment ecologista,
- el moviment antinuclear,
- el moviment feminista, i
- en general, diversos grups i associacions que qüestionaven explícitament els estrets vincles entre la investigació científica, l'estament militar i la indústria.

Paral·lelament, des de la filosofia, la història de la ciència i la sociologia del coneixement es van començar a desenvolupar aproximacions alternatives a la concepció tradicional de la ciència i la tecnologia.



EN L'ÀMBIT DEL DESENVOLUPAMENT TECNOLÒGIC, AQUESTES PERSPECTIVES ALTERNATIVES CONSIDEREN QUE LES TECNOLOGIES NO SÓN BONES NI DOLENTES, PERÒ... (I AQUEST PERÒ ÉS MOLT IMPORTANT) TAMPOC SÓN NEUTRALS.

Què vol dir que les tecnologies *no són neutrals*?



Vol dir que tenen efectes i porten inscrits una sèrie de valors.

Com identificar aquests valors en la IA generativa?

P.5

Qüestionar la neutralitat de la ciència i la tecnologia difumina la frontera entre ciència i política, natura i societat, natural i artificial...

La IA generativa no es pot comprendre de manera aïllada ja que està estretament relacionada amb l'entorn social, polític, econòmic i cultural.

Així, qualsevol dispositiu tecnològic és més que un simple artefacte: és un entramat sociotècnic; o, com Bruno Latour l'anomena, un «actor-xarxa».

Quan es dissenya una eina tecnològica, aquesta també incorpora una sèrie de valors, fruit del seu context.



Aquests valors «incrustats» configuren el guió que ens indica com comportar-nos com usaris.

Un exemple:

Quan anem a una màquina de cafès i surt un got de plàstic, el guió de comportament del got ens diu:



**utilitza'm
i llença'm**

Aquest missatge implica una sèrie de valors relacionats amb la producció de plàstics i el desenvolupament industrial.

Ara bé, si substituïm aquest got per una tassa de ceràmica que hem portat de casa, el guió de la tassa ens dirà:



**utilitza'm,
neteja'm i
reutilitza'm**

En aquest altre missatge s'integren una sèrie de valors relacionats amb la reducció de residus.

Analitzar el guió de comportament i els valors inscrits en ChatGPT no és tan simple com en el cas del got i la tassa de cafè. Ara bé, les preguntes que ens fem per encetar el debat són similars a les preguntes de l'exemple proposat.



1. Com s'ha fet ChatGPT?

Sabem quins són els recursos naturals necessaris per produir-lo? El seu desenvolupament i ús tenen algun impacte ambiental? D'on i com s'han obtingut les dades necessàries per entrenar-lo? Quines persones han participat en el seu disseny? I en el seu entrenament? Els diversos grups de persones relacionades amb la producció de ChatGPT, tenen les mateixes condicions de vida i treball?

2. Quin tipus d'accions prescriu ChatGPT?

Per al seu ús en l'àmbit educatiu, per exemple, quines accions relacionades amb l'estudi i l'aprenentatge ens anima a emprendre? Aquestes accions tenen algun tipus d'efecte sobre altres col·lectius? La prescripció d'unes accions implica que altres accions es descarten: quines accions es deixen de fer a l'aula quan s'utilitza ChatGPT? Són accions que volem mantenir en les nostres vides o les podem descartar? L'anàlisi seria la mateixa si examinéssim altres aplicacions d'IA generativa? Aquesta reflexió seria la mateixa en tots els àmbits, per exemple, en salut per a la realització de diagnòstics mèdics?

REFERÈNCIES PER SABER-NE MÉS

- a. Aibar, E. (2023). *El culto a la innovación. Estragos de una visión sesgada de la tecnología*. Barcelona: Ned Ediciones.
- b. Crawford, K. (2023). *Atlas de IA*. Barcelona: Ned Ediciones.
- c. Domènech, M. i Tirado, F. J. (1998). *Sociología simétrica. Ensayos sobre ciencia, tecnología y sociedad*. Barcelona: Gedisa.
- d. Latour, B. (2023). *Habitar la Terra*. Barcelona: Arcàdia.