



De la Intel·ligència Artificial a la immortalitat de l'ànima

N. 269. Setembre del 2023. Suplement del Quadern CJ n. 234
Cristianisme i Justícia. Roger de Llúria, 13, 08010 Barcelona
93 317 23 38 • info@fespinal.com • www.cristianismeijusticia.net

A finals del 2022 es va donar a conèixer al gran públic l'anomenat ChatGPT, un *software* programat per reproduir el llenguatge humà amb la capacitat de donar resposta a qualsevol pregunta que se li plantegi. El fet que aquest *software* hagi après a donar respostes no prefixades —gràcies a un entrenament algorítmic mitjançant xarxes neuronals i a la quantitat d'informació que ha arribat a digerir— ha posat l'expressió intel·ligència artificial (IA) en boca de tothom. Tenim una eina capaç d'escriure des d'un poema romàntic inèdit a l'estil de Goethe fins a un treball acadèmic comparant dos autors que fins ara no s'havien estudiat mai junts, i fer-ho millor que el 99% dels mortals. Les preguntes que això ens suscita són moltes: és realment intel·ligent, la IA? On radica la novetat? Quines con-

seqüències pot tenir la seva popularització? Ajudarà a millorar el nostre món? Què en podem esperar? A continuació avaluem algunes respostes diferents de les que ens donaria ChatGPT.

El problema no resideix en la veritat sinó en la versemblança

Un dels debats més importants al voltant de ChatGPT —i de tots els models de llenguatge extens (LLM, per les seves sigles en anglès: *Large Language Models*)— no és si el que diu és veritat o mentida, si s'equivoca molt o poc, sinó que ha assolit una versemblança total, de manera que la conversa amb el xat no es pot diferenciar d'una conversa real. El teòric Lopez de Mántaras ho escrivia fa

poc: el problema amb ChatGPT és el seu antropomorfisme, que ens fa caure en la falsa sensació de realisme. Ens el creiem, no perquè digui la veritat o s'hi acosti, sinó perquè imita a la perfecció la conversa humana, les seves respostes estan molt ben escrites i transmet credibilitat. Però tinguem sempre present que la IA ni sap per què sap el que sap, ni entén el que diu o respon. La seva resposta busca simular el llenguatge humà, i per tant després d'una frase com: «El millor jugador de futbol de la història és...» ens respondrà el que hagi après de frases similars escrites a internet, afegint a la frase una paraula darrere l'altra seguint criteris probabilístics i garantint que el que diu tingui sentit. Fins i tot semblarà que s'ha forjat una opinió pròpia. Però el que hi ha al darrere és, com ha descrit la professora de lingüística computacional Emily Bender, un lloro estocàstic.

Certament, davant d'una eina com aquesta la confusió està garantida, i la possibilitat d'utilitzar-la traient profit de les nostres vulnerabilitats encara ho està més. La IA augmenta de manera notable el que el món digital i les xarxes socials ja van iniciar: la postveritat. Les possibilitats de manipulació, d'utilització no contrastada d'informació, de xantatge emocional, ens obliguen a una formació ciutadana que, d'una banda, tingui criteris de discerniment per poder posar sota sospita el que llegeix, i de l'altra que hagi interioritzat prou el sentit comú per distingir si s'està sent o no manipulat en la informació que es rep. Ens caldrà, en definitiva, ser més escèptics i menys confiats.

La irrupció de la intel·ligència artificial pot donar lloc a un gir epistèmic definitiu en la nostra escala de valors en la qual la veritat deixa de tenir importància

i el que compta és la versemblança, és a dir, l'aparença de veritat.

La *insostenibilitat* material

Gran part del debat generat al voltant de la IA té a veure amb el perill dels seus biaixos. Ens preocupa que, essent la IA una caixa negra (no sabem per què diu el que diu o fa el que fa) acabi reproduint les discriminacions que es donen *de facto* en la societat. Però és evident que si la IA és racista o masclista ho serà perquè la nostra societat ho és. Ni la IA és autònoma ni podrà substituir el nostre judici moral. Encara menys millorar-lo. La preocupació pel biaix dels algorismes —però sobretot de les dades de què es nodreix— participa d'una gran ingenuïtat tecnooptimista: la creença que la tecnologia ha vingut per resoldre els nostres dilemes morals, per fer-nos millors persones. No podem esperar més de la IA del que esperem de nosaltres mateixos.

La segona preocupació més estesa té a veure amb el fet que la IA pugui substituir definitivament el treball humà. Un article recent publicat per *Fortune* exposava el que es coneix com *la paradoxa de la productivitat*: el canvi digital introduït des de mitjans dels anys noranta, contràriament al que s'esperava, no ha provocat grans augments de productivitat. És a dir, disposar de tecnologia digital no ens fa més eficients ni elimina realment llocs de treball, només els transforma. La pregunta que ens hem de fer més aviat és qui es beneficia d'aquestes tecnologies i de quina manera permeten acumular cada cop més poder en menys mans.

El que realment hauria de ser motiu de debat, i és un dels temes de què menys s'ha parlat, és la insostenibilitat de la IA.

Quan comentes la qüestió amb informàtics, veus que el que més els preocupa és la capacitat computacional necessària per entrenar i mantenir en marxa la IA: els càlculs utilitzats per entrenar xarxes neuronals s'han multiplicat per cent milions els últims deu anys. A nosaltres ens sembla innocu, gairebé màgic, però demanar a ChatGPT que t'expliqui l'acudit més divertit que sap implica una utilització de servidors, un consum d'energia i un nombre de càlculs que en fan inviable l'ús extens i universal. Un divertiment com aquest es converteix en un cost computacional impossible d'assumir a gran escala.

Una lectura antropològica i teològica

La lectura més interessant del fenomen sempre cau del costat humà. I nosaltres, com ens relacionarem amb la IA? Fa poc vam veure amb alumnes d'un seminari sobre pensament científic un capítol de la sèrie *Black Mirror* titulat «Be right back». Té exactament deu anys. Una dona jove perd la seva parella. Una empresa ha desenvolupat una tecnologia que permet, recuperant la informació de la persona en vídeos, fotos i xarxes socials, simular una conversa per xat amb el mort, proposta que més endavant es converteix en una conversa telefònica i, finalment, en un robot idèntic a la persona morta. Com podeu imaginar, el suposat retorn del mort no arriba a satisfer en cap moment la persona viva, més aviat al contrari, l'aïlla, la desestabilitza i la incapacita per reprendre la seva vida. L'enfonsa encara més en un pou sense llum.

Models LLM com els del ChatGPT permeten, si se'ls dona la informació cor-

responent, simular una conversa per xat amb qualsevol persona morta, per exemple Kennedy o Freddie Mercury. Altres eines ja desenvolupades poden agafar les seves veus en talls de tres segons i simular a partir d'aquí una conversa sonora real. I ara què? Si això ja és possible, estem obligats a desenvolupar-ho? Imagineu la quantitat de diners que podria aconseguir un reclam com aquest: torna a parlar amb els teus morts per telèfon! Quanta gent pagaria per aquest servei? Estem a pocs mesos de saber-ho per macabra que ens sembli la idea.

En aquest sentit, la millor reflexió me la va proporcionar un alumne: «A la noia no se li ha donat ni l'oportunitat de superar el dol». Aquí rau la qüestió més rellevant. La tecnologia ens permet fugir i no haver d'enfrontar-nos al que ens fa por, ens amenaça o significa un repte vital. Però això parteix d'una mala comprensió de l'antropologia humana. Recordant Hölderlin, «on hi ha adversitats neix el que ens salva». Volem estalviar a l'humà haver de ser humà. No fem cap favor a ningú negant-li la possibilitat del dol en allò que forma part central d'una vida. La mort d'un ésser estimat pot ser la pitjor de les desgràcies, però la vida va exactament d'això. Com afirma Josep Maria Esquirol, no es tracta de tancar la ferida infinita que ens constitueix, sinó d'aprendre a «acompanyar i respondre al seu excés». Amb cada crossa tecnològica que afegim al nostre dia a dia ens empetim i ens fem més febles, més incapaços; en definitiva, menys humans.

No hi ha cap dubte que la IA participa d'un cert gnosticisme. Es tracta de la vella heretgia cristiana segons la qual la salvació es pot assolir a través del coneixement i la il·luminació. Alhora, es rebutja el món material i fins i tot el cos,

per imperfecte. El culte a les dades, la possibilitat de transcendir la nostra vida mortal buidant-nos «al núvol» és, de fet, la versió contemporània de l'heretgia gnòstica. I és sobretot la il·lusió, encara que sigui a base de *bits*, segons la qual un cop el cos fineixi la nostra ànima es perpetuï. Serà la IA la porta d'entrada a la immortalitat de l'ànima? La *tecnolatria* com a religió és el gnosticisme del segle XXI. Sigui com sigui, no és el mateix la resurrecció de la carn que la immortalitat de l'ànima. La primera no és possible ni ho serà mai, la segona ja la tenim a l'abast gràcies a un xat que pot fer creïble la conversa amb persones mortes. Valdria més que ens en protegíssim. En temps de secularització estesa, on res està proveït del sagrat, la possibilitat de transcendència s'ha traslladat al camp digital. Busquem la salvació a través de la tecnologia. No l'obtindrem. El que tindrem, més aviat, serà un malson de confusió, la desconexió del nostre ésser natural i la incapacitat d'entendre la vida com a éssers finits que som.

Regular la IA i dotar-nos d'eines per al discerniment

Ningú nega que la IA pot ser molt útil en aplicacions de diagnòstic mèdic, de predicció climàtica o de prevenció d'accidents en la conducció. Ara bé, no a qualsevol preu. Expressions com «la IA ha vingut per quedar-se» o «la IA és neutra, tot depèn de l'ús que en facis» formen part d'una mirada miop sobre el que

és la tecnologia per a les nostres vides i el que implica. La fascinació que genera la tecnologia sovint és acompanyada d'un discurs ingenu alimentat per l'interès de tants diners invertits que busquen rendibilitat. La bona notícia: encara som a temps de limitar-ne les pitjors conseqüències evitant tot negoci amb les vulnerabilitats de l'ànima humana. La mala notícia: un cop la tecnologia està desenvolupada transforma el nostre medi i la seva simple existència ja ha canviat la manera que tenim de relacionar-nos amb el món.

L'antídoto personal serà cultivar la capacitat de discerniment. Saber el que fem i per què ho fem: «On vaig i a fer què», que es demanava sant Ignasi. Només la nostra intencionalitat, el fet que les nostres accions i paraules estiguin dotades d'una orientació, ens pot fer avui diferents a les màquines. I en l'àmbit social val més que comencem a regular amb legislacions i protocols ètics cada nou desenvolupament digital, perquè hi ha molts milions invertits que esperen el seu retorn econòmic a punt d'aprofitar-se de totes les nostres febleses. Que vols tornar a parlar per WhatsApp amb la teva germana que va morir l'any passat? Ho tenim a un clic. Deixarem que això es desenvolupi com a negoci? Espero que siguem prou intel·ligents, perdó, prou humans, per no permetre que passi.

Xavier Casanovas Combalia
Professor de la Càtedra d'Ètica d'IQS
(URL)