

# Cervells, robots, humans

Apunts d'un UPCnauta

Xavier Duran

**UPCArts**Diàlegs



UNIVERSITAT POLITÈCNICA  
DE CATALUNYA  
BARCELONATECH



# Cervells, robots, humans



# Cervells, robots, humans

## Apunts d'un UPCnauta

Xavier Duran



UNIVERSITAT POLITÈCNICA  
DE CATALUNYA  
BARCELONATECH

Diàlegs UPCArts

Coordinació: Carme Fenoll i Antoni Hernández-Fernández

© Disseny de la coberta: Jordi Sàbat

Primera edició: gener de 2024

© Xavier Duran, 2024  
© Iniciativa Digital Politècnica, 2024  
Oficina de Publicacions Acadèmiques Digitals de la UPC  
Jordi Girona, 31 edifici K2M 08034 Barcelona  
Tel.:934 015 885  
[www.upc.edu/idp](http://www.upc.edu/idp)  
E-mail: [info.idp@upc.edu](mailto:info.idp@upc.edu)

ISBN: 978-84-10008-22-9

ISBN digital: 978-84-10008-21-2

DL: B 3293-2024

DOI: 10.5821/ebook-9788410008212

Qualsevol forma de reproducció, distribució, comunicació pública o transformació d'aquesta obra només es pot fer amb l'autorització dels seus titulars, llevat de l'excepció prevista a la llei.

## Sumari

- 9 Pròleg**  
*Jordi Llorca*  
*Vicerector de Recerca de la UPC*
- 13 Entre xips i neurones**
- 21 Artificial, sí, però... intel·ligència?**  
28 Intel·ligència sense cervell
- 35 D'ètica i biaixos**  
45 Simular el sentit comú?  
54 Un univers per a cadascú  
59 Màquines al·lucinades  
61 La IA treballa. I nosaltres?
- 69 Diagnòstics, fàrmacs i bessons digitals, la cara positiva de la IA**  
72 IA i salut, un camp en expansió  
78 Bessons digitals  
83 Pacients digitals

- 87 **Converses amb ChatGPT i amb chatbots**
- 99 Chatbot *contra el* bullying
- 103 Vida digital després de la mort
- 111 **Passeig amb els robots**
- 116 La robòtica arriba a Catalunya
- 124 Per la vall inquietant
- 135 Realitat i ficció
- 146 Caminant amb robots
- 150 El futur i les pors
- 155 **L'art difícil de predir**
- 161 **Final (provisional) de trajecte**
- 169 **Breu selecció bibliogràfica**

## Pròleg

*Jordi Llorca*

Vicerector de Recerca de la UPC

El llibre que teniu a les mans és fruit de la segona edició de la Residència d'escriptura, organitzada pel programa cultural UPCArts de la Universitat Politècnica de Catalunya (UPC). Amb aquest programa, impulsem la participació d'escriptors i escriptores en la vida cultural i comunitària de la nostra universitat per tal d'establir ponts entre la docència, la recerca i la transferència de tecnologia de la universitat amb el públic general mitjançant la literatura contemporània.

El coneixement no té límits, i menys encara la literatura. Però, si no es comunica, no esdevé transformador, i la literatura ofereix una oportunitat comunicativa excel·lent. El relat literari troba una font d'inspiració enorme en tots els àmbits de coneixement de la UPC (arquitectura, enginyeria, ciència i tecnologia).

Xavier Duran, guanyador de la Residència d'escriptura 2023, és periodista i divulgador científic. Ha viscut intensament durant uns quants mesos la recerca que es fa a la UPC i ha conegut les persones que hi ha al darrere. Conèixer tota la recerca que es fa a la UPC en pocs mesos era un repte gegantí; per això, l'escriptor va escollir tractar sobretot de la robòtica i la intel·ligència artificial, dues de les múltiples àrees en què la UPC excel·leix. I segur que aquesta tria no li fou gens fàcil!

Aquesta obra es mou en el laberint contemporani de la creativitat i la tecnologia, on les fronteres entre el cervell humà, les màquines i els algorismes sembla que es van diluint progressivament. Amb una mirada hàbil i penetrant, bevent de la seva experiència personal a la UPC, l'escriptor trosseja les diverses capes del coneixement acadèmic i ens guia amb destresa pels diferents campus universitaris de la UPC, per tal d'oferir-nos una perspectiva única i reveladora sobre com els laboratoris tecnològics es fusionen amb els passadissos de la ment. El relat esborra les barreres rígides entre les disciplines i desafia les idees preestablertes, mostrant com la ment humana, els robots i la intel·ligència artificial es poden entrellaçar de maneres inimaginables.

A banda de presentar projectes revolucionaris de la UPC, l'autor també trenca una narrativa que cospa la passió i la dedicació dels professors i professores que els lideren i els fan possibles. Cada pàgina és una invitació a explorar un món on la ciència i la imaginació convergeixen, obertes a noves possibilitats i a descobriments inesperats.

Les fronteres del coneixement no són límits infranquejables, sinó terres fèrtils per a la creativitat, l'exploració i la innovació. Benvinguts a aquest interessant viatge literari, on el futur i la realitat es confonen, on es perfilen els nostres propis límits com a éssers pensants i creatius, i on la UPC esdevé un escenari de descobriments inesperats. Com digué un altre escriptor, Joseph Conrad, «l'autor escriu només la meitat del llibre; de l'altra meitat, se n'ha d'ocupar el lector». Que aquesta obra sigui una invitació a la reflexió i a la celebració de l'audàcia científica i tecnològica, i del que significa ser humà en una era de descobriments incessants.



## Entre xips i neurones

La primera sensació va ser de vertigen.

T'ofereixen una oportunitat extraordinàriament temptadora, plena de possibilitats i, de cop, et sents en una tribuna elevada des d'on contemples un laberint ple de camins, sense saber quin o quins triar. Senders que es bifurquen, com en el conte de Borges, que es retroben i porten a un mateix lloc –o a llocs diferents, però que estan comunicats. Rutes, totes elles molt enriquidores, però que requereixen una tria.

I tot potentat –o complicat–, perquè a cada visita, a cada discurs, a cada conversa apareixen més idees, capaces, totes elles, d'entusiasmar i promoure encara més idees, d'obrir nous camins.

Per la meua feina, he pogut conèixer moltes persones de la UPC i moltes de les activitats i recerques que s'hi fan. Ser periodista científic és sovint un pri-

vilegi, perquè et porta a la cuina on es cou el futur, on s'afronten els problemes actuals o es fa volar la imaginació –una imaginació molt consistent i coherent, tocant bastant de peus a terra– per fer progressar el coneixement.

Per això, quan vaig tenir la sort que m'eleghis- sin *escriptor resident 2023* en el marc de la iniciativa UPCArts, no partia de zero. Tanmateix, com no sentir vertigen quan se t'obren les portes d'una universitat amb nou campus i divuit centres docents? Quin camp escollir en una universitat que investiga des de la comunicació dels cetacis en els fons marins i les mines fins a l'enginyeria aeroespacial i les supernoves, passant pel tèxtil, el paper i tants altres àmbits?

I això sense comptar altres centres o institucions vinculades que es troben en terrenys de la UPC i que també he pogut conèixer de prop. Un d'ells és el Barcelona Supercomputing Center (BSC), als jardins de la Torre Girona, on fa anys la família propietària va construir una capella, avui dessacralitzada. Allà, traspassades dues torres d'aparença romànica, i entre arcs de mig punt i vitralls que potser volen ser neogòtics, hi ha l'urna de vidre que protegeix el superordinador *MareNostrum*. Art sacre i tecnologia informàtica, fusió de tècnica arquitectònica i ciència informàtica, i un contrast entre l'antiguitat i la modernitat, entre la fe i la raó. Al desembre del 2023 es va inaugurar el *MareNostrum 5*, que ja no s'ha pogut situar a la capella i ha necessitat un altre emplaçament. Aquest superordinador es troba entre els vint més potents del món i

té un rendiment total màxim de  $314$  petaflops, és a dir,  $314.000$  bilions de càlculs per segon.

I un altre centre de recerca punter en el seu camp és l'Institut de Ciències Fotòniques (ICFO), al campus de Castelldefels, que investiga en aquest món estrany de la quàntica, que sorprèn i desorienta als no entesos –quan proposava temes a la redacció, havia d'anar amb cura perquè algunes de les paraules no fessin saltar de la cadira els editors–, però que sempre acaba tenint el lligam directe amb la vida quotidiana. Allà, la complexitat dels experiments va acompanyada d'una ferma voluntat de divulgar i atansar la ciència als més joves. Per això, a banda de les sales on el món quàntic s'explica a partir de les nostres percepcions quotidianes, en algunes parets hi ha reproduccions de portades de *Science Wonder Stories* o d'*Amazing Stories*.

Per tot això, calia acotar la tria, de manera que em vaig decidir per dos camps molt relacionats entre ells i que m'interessen molt, no només per les possibilitats que obren o que ja han obert, sinó també per tot el debat ètic i social que generen: la intel·ligència artificial i la robòtica.

El repte següent era amb qui parlar. Calia fer un tria que, malgrat tot l'assessorament, podria ser tan subjectiva com qualsevol altra. Per això, vull agrair l'ajut de totes les persones amb qui he parlat i demanar disculpes a aquelles amb qui no he parlat i que, de ben segur, em podrien haver fet aportacions molt interessants.

El temps és limitat. I això va en detriment d'una oportunitat com aquesta. Perquè aprofitar bé la resi-

dència significaria barrejar-se cada dia amb la vida de la universitat, assistir a algunes classes, veure pràctiques i recerques, participar en sortides –docents o no– dels estudiants...

Crec que, malgrat les limitacions, he aprofitat l'oportunitat –per bé que sempre penses que podries haver fet molt més. He pogut conèixer moltes persones, aprofundir temes, endinsar-me en nous debats... I ara en teniu a les mans un dels resultats, perquè una part del vertigen provenia del compromís d'escriure un llibre relacionat amb la UPC. Cada nou llibre és un gran repte, encara que n'hagis publicat uns quants. Com l'enfoques? Com l'estructures? Com fas que sigui alhora interessant i amè? Un llibre és el resultat de triar entre múltiples opcions. Aquí el teniu i espero que us resulti almenys una mica útil, interessant o distret.

Els projectes aparentment senzills són molt més complexos de dur a terme del que la majoria de gent pensa. Per això, vull agrair, en primer lloc, a Carme Fenoll, directora d'UPCArts, haver tingut aquesta iniciativa, una més de les que ha impulsat per intentar –i aconseguir– relacionar les humanitats amb una universitat científica i tecnològica. I, naturalment, al rector Daniel Crespo i al seu equip, que han acceptat i han desenvolupat una idea que fa que la UPC torni a ser pionera, en aquest cas, en la creació de la figura de l'escriptor resident.

El meu agraïment, també, a totes les persones que treballen des de l'anonimat perquè tot funcioni correctament en el dia a dia. Sovint no se les veu ni se

les anomena, però les seves gestions són essencials perquè les coses es puguin dur a terme en les condicions correctes.

I, naturalment, gràcies a totes les persones que m'han dedicat part del seu temps i m'han proporcionat materials, amb una paciència i amb una amabilitat encomiables. Aquesta n'és la llista i crec que no em deixo a ningú. Les poso per ordre alfabètic, que potser no és el millor criteri, però sempre és un dels més asèptics: Marc Aliet, Josep Amat, Cecilio Angulo, Cristina Canal, Alícia Casals, Rolando Chacón, Ulises Cortés, Cesca Cufí-Prat, Joan de Pablo, Adriana Farran, Karina Gibert, Ferran Marquès, Martina Massana, Manuel Moreno, Júlia Pareto, Marta Pegueroles, Alberto Sanfeliu, Gil Serrancolí, Carles Sora, Carme Torras i Verónica Vilaplana.

Donar coherència al llibre obliga, malauradament, a deixar de banda els treballs d'algunes d'aquestes persones. Així, a la visita que vaig fer al campus de Diagonal-Besòs, acompanyat de Joan de Pablo, professor i investigador de l'Escola d'Enginyeria de Barcelona Est (EEBE) i delegat del rector, em vaig retrobar amb una companya de promoció de la UAB, Adriana Farran, avui directora de l'EEBE. I també vaig poder conèixer la recerca de les doctores Cristina Canal i Marta Pegueroles. Els seus treballs són apassionants i punters, però no sabia com lligar-los amb la IA. I és per això que els exposo breument aquí.

Concretament, al campus de Diagonal-Besòs, Cristina Canal i Marta Pegueroles investiguen di-

ferents tecnologies per a aplicacions mèdiques dins del grup de recerca en Biomaterials, Biomecànica i Enginyeria de Teixits (BBT), del Centre de Recerca en Enginyeria Biomèdica (CREB) de la UPC. El BBT té com a objectiu desenvolupar biomaterials per a la regeneració i/o la reparació funcional de teixits i òrgans. Per això, cal dissenyar materials que contribueixin a regenerar i formar de nou els teixits degradats, i que equilibrin perfectament la integració del biomaterial amb la recuperació funcional perduda.

Cristina Canal dirigeix el Laboratori de Plasma per a Aplicacions Biomèdiques, on desenvolupa i aplica nous biomaterials que, junt amb la tecnologia del plasma fred, proporcionen, entre altres coses, noves vies per a la regeneració de teixits o per a la teràpia del càncer.

Va ser apassionant conèixer les aplicacions mèdiques d'aquest quart estat de la matèria, en què electrons han ionitzat els àtoms de gas. Així, s'ha demostrat que els plasmes poden matar selectivament cèl·lules canceroses, sense afectar els teixits sans.

I no només això. Al laboratori vaig veure el que anomenaven «bolígraf» de plasma, un aparell que ja està en les fases clíniques de la recerca i que es podria utilitzar en les úlceres diabètiques, perquè és antibacterià i combat la infecció i estimula la regeneració dels teixits.

Com explicava la doctora Canal, junt amb altres investigadors punters en aquest camp, a la revista *Europhysics News* el 2022, la fusió de la física, la química i l'enginyeria amb la ciència mèdica ha donat lloc a

la medicina del plasma, «que pretén desenvolupar tecnologies noves i innovadores per millorar la qualitat de vida dels pacients i les seves famílies». La recerca ha obert «una miríada d'aplicacions en els límits de l'enginyeria de teixits i la medicina regenerativa, com l'erradicació de biofilms, la cura de ferides o la teràpia del càncer».

Per la seva banda, la doctora Pegueroles centra la seva recerca en el desenvolupament, la fabricació i la modificació superficial de biomaterials per a aplicacions cardiovasculars. Una de les seves línies de recerca és la fabricació de *stents*<sup>1</sup> polimèrics i metàl·lics amb impressió 3D. En aquest camp, ha desenvolupat *stents* biodegradables que duren dos anys, fets amb PLA (àcid polilàctic). Mentre es va degradant dintre de l'artèria, aquest *stent* pot alliberar un fàrmac.

Aquests apunts breus són una mostra que cada visita a la UPC sempre aporta noves informacions, noves idees, noves dades, noves perspectives. No es pot abastar tot, ni de bon tros. Espero que aquest llibre faci una mica de justícia a la tasca ingent que tanta gent desenvolupa cada dia des de fa dècades a la UPC. I que en gaudiu tant com jo, tot passejant entre xips i neurones.

---

1 Un *stent* és un petit tub de malla que es col·loca dintre d'una artèria per evitar que es tanqui.



## Artificial, sí, però... intel·ligència?

La intel·ligència artificial (IA) va néixer l'any 1956. Per ser estrictes, hauríem de dir que aquell any va sorgir el concepte d'IA. Va ser concretament a l'estiu, en el marc de la Conferència de Dartmouth, que va tenir lloc en aquesta població de New Hampshire, als Estats Units.

Un any abans, John McCarthy, professor de matemàtiques del Dartmouth College, havia decidit organitzar un grup per desenvolupar idees sobre màquines pensants. Va escriure, amb Marvin Minsky, Nathaniel Rochester i Claude E. Shannon, una carta a la Fundació Rockefeller per demanar finançament per al *Dartmouth Summer Research Project on Artificial Intelligence*. El nom, doncs, ja apareixia. I quan es va fer la trobada, McCarthy els va acabar de convèncer d'adoptar aquest nom. Naixia la intel·ligència artificial.

La primera entrevista que vaig fer com a escriptor resident va ser a Ulises Cortés, catedràtic de la Facultat d'Informàtica de Barcelona (FIB). Érem al seu despatx, envoltats de diverses obres originals, la majoria d'artistes mexicans, i del cartell d'una exposició sobre modernisme mexicà al San Francisco Museum of Modern Art, on apareixia Frida Kahlo. I, entre les seves primeres paraules, em va dir que podria tenir un problema terminològic en el llibre: «*Intel·ligència artificial* és més aviat un nom comercial, publicitari.»

El terme el va encunyar McCarthy i el va definir com «la ciència de dissenyar i construir màquines intel·ligents». Però, què és una màquina intel·ligent? Marvin Minsky ho va concretar l'any 1968, quan va dir que la intel·ligència artificial és una ciència que té com a objectiu assolir «que una computadora faci coses que, quan les fan les persones, es diu que impliquen intel·ligència».

McCarthy buscava un nom que distingís aquesta disciplina de l'estudi dels autòmats, però també de la cibernetica, creada per Norbert Wiener, que considerava que era una persona amb molta projecció, que potser hauria captat massa l'atenció i hauria adquirit un protagonisme excessiu.

Tanmateix, a banda d'aquests matisos, el doctor Cortés no li reca parlar d'IA. Des de l'any 1981, existeix una assignatura amb aquest nom a la FIB de la UPC, que després s'ha impartit també a diverses enginyeries. I, des del 2021, existeix com a grau. També hi ha un màster interuniversitari en IA, impartit con-

juntament per la UPC, la Universitat de Barcelona i la Universitat Rovira i Virgili.

Com m'explica Cortés, la informàtica és una forma d'expressió no ambigua per programar una màquina. I la IA és una disciplina científica que busca sistemes que emulin un procés que anomenem *intel·ligència* i els mecanismes que el controlen. A Cortés i als grups que dirigeix (Grup d'Enginyeria del Coneixement i Aprenentatge Automàtic i *Intelligent Data Science and Artificial Intelligence Research Group*) no els preocupa només que les màquines facin coses que es podrien anomenar *intel·ligents* –més endavant ja ens preguntarem si és així–, sinó que les màquines *sàpiguen* per què fan les coses. I per això afirma que la IA no s'entén sense conèixer la seva història i personatges com Llull, Pascal i Leibniz, però també cal estudiar Volta, Cajal, Pavlov, Skinner...

Creu que darrerament s'han barrejat conceptes, sobretot als mitjans de comunicació. Així, el *machine learning* o aprenentatge automàtic és un subconjunt dins la IA que va ser introduït per Arthur Samuel l'any 1959, mentre treballava a IBM, on desenvolupà el primer joc de computadora intel·ligent, concretament un joc de dames. El va definir com «una aplicació de la IA que proporciona als sistemes la capacitat d'aprendre i millorar automàticament a partir de l'experiència, sense estar programats explícitament». Així, amb l'aprenentatge automàtic, l'ordinador és capaç de prendre decisions noves i independents, basades en noves dades i la seva programació, sense directrius humanes addicionals.

Dintre l'aprenentatge automàtic, hi ha un altre subconjunt: el *deep learning* o aprenentatge profund. El terme el va introduir l'any 1986 la professora Rina Dechter, que el va definir com «una classe d'algoritmes d'aprenentatge automàtic que utilitza múltiples capes per extreure progressivament característiques de nivell superior a partir de dades sense processar».

L'aprenentatge profund només fa servir dades i dona respostes sense raonar. Seria com l'alumne clava-colzes que memoritza, però que no sap el perquè de les coses. Per això, Cortés diu que el *deep learning* seria alquímia i la recerca que ells fan seria química. Per dir-ho en un altre to, i sense referències a antigues pràctiques precientífiques, el *deep learning* seria enginyeria i els treballs dels seus grups serien ciència.

Ara podem tornar a la pregunta que dona títol al capítol. Sens dubte, la IA és artificial. Però, realment, és intel·ligència? El debat no és nou. En un article publicat l'any 1950 a la revista *Mind*, Alan Turing, considerat el pare de la informàtica moderna, ja va proposar aquesta qüestió: Poden pensar les màquines?

Com que, segons argumentava, trobar la resposta a aquesta pregunta requeriria primer definir concretament els conceptes de «màquina» i «pensar», i això podria dependre de l'ús que es fes d'aquests termes, va idear un variant d'un suposat joc tradicional anglès, anomenat «de la imitació». Sembla que aquest joc no va existir mai –no se n'han trobat referències– i que Turing se'l va inventar i, a partir d'aquí, se'n va popularitzar el nom. Segons Turing, es tractava que una persona (C) es comu-

niqués amb un home (A) i amb una dona (B) només amb missatges escrits a màquina i que esbrinés quin dels dos interlocutors era l'home i quin la dona. La gràcia és que l'home està autoritzat a mentir per enganyar C, però la dona ha de proporcionar pistes verídiques perquè aquest ho endevini. Al cap de cinc minuts, ja se sap si C ha encertat o si A ha aconseguit confondre'l.

En una variant, ideada també per Turing, A és substituït per una màquina i B pot ser un home o una dona. Així, C ha d'endevinar quin dels dos interlocutors és un humà i quin una màquina. Si la màquina aconsegueix enganyar-lo es considerarà que és capaç de pensar com un humà.

És objecte de debat si alguna màquina ha superat l'anomenat test de Turing i ha aconseguit fer-se passar per un humà. Hi ha qui opina que sí i potser la darrera versió de ChatGPT ha pogut enganyar un gran percentatge d'usuaris. Però probablement el sistema no guanyaria en el joc a experts en IA. I, si ho aconseguís, no voldria dir necessàriament que fos intel·ligent, sinó que té la capacitat d'enganyar una persona i fer-se passar per un humà. Als investigadors el que els preocupa més actualment és avaluar el rendiment dels sistemes en determinades capacitats, com les lingüístiques o les matemàtiques, i en sentit comú. Sobre tot això, és recomanable l'article de Celeste Biever, editora en cap de notícies de *Nature*, que va publicar en aquesta revista el 27 de juliol de 2023: «ChatGPT broke the Turing test –the race is on for new ways to assess AI».<sup>2</sup> Per

---

2 Vegeu: <https://www.nature.com/articles/d41586-023-02361-7>

tant, per respondre a la pregunta que es feia Turing fa més de setanta anys, reformulada en aquests termes: «Pot ser intel·ligent una màquina?», no podem estar-nos, ara com ara, de definir *intel·ligència*.

Si repassem algunes de les afirmacions que van fer els assistents a la Convenció de Dartmouth, la resposta no només seria sí, sinó que, a més, ja s'hauria aconseguit una màquina tan intel·ligent o més que un humà. A la convenció es partia de la conjectura que tot aspecte de l'aprenentatge o qualsevol altra característica de la intel·ligència es pot descriure, en principi, amb tanta precisió que es pot fabricar una màquina per simular-ho. I així s'arribarien a aconseguir màquines que resolguessin les classes de problemes reservats als éssers humans i que milloressin per si mateixes. És a dir, que aprenguessin per si soles, sense que un humà els introduís programes o informació.

L'any 1958, dos dels assistents al seminari, Herbert A. Simon i Allen Newell, van pronosticar que en deu anys un ordinador seria el campió mundial d'escacs i que en el mateix període un ordinador descobriria i demostraria un nou i important teorema matemàtic. Però no fou fins al 1997 que *Deep Blue* va derrotar el campió mundial Garri Kaspàrov. I no tenim notícia que cap ordinador hagi descobert cap nou teorema. Una altra cosa és que hagi ajudat a demostrar-ne alguns de ja coneguts.

Simon va tornar a fer un pronòstic el 1965: en vint anys, les màquines serien capaces de fer qualsevol feina que un humà pogués dur a terme. L'any 1967, Minsky

va afirmar que, en una generació, el problema de crear intel·ligència artificial estaria resolt. I el 1970 va afegir que passarien només entre tres i vuit anys fins a tenir una màquina que superés la intel·ligència general de l'ésser humà mitjà.

No cal dir que aquests eminents professors i investigadors van pecar d'optimisme. Sí que es van desenvolupar alguns sistemes amb capacitats interessants per identificar patrons, però aviat va arribar el primer «hivern» de la IA. Imitar la intel·ligència humana s'havia revelat un objectiu molt complex i la manca de resultats va fer disminuir els fons destinats a la recerca en aquest camp. La disciplina es va fragmentar i, en comptes d'intentar assolir una IA general, es va investigar en determinades capacitats, com el processament del llenguatge natural, el reconeixement de la veu o el reconeixement visual, o en tasques concretes. Per tant, avui tenim màquines que són millors que els humans fent determinades tasques, però cap que es pugui dir que és tan intel·ligent com un humà.

I això ens porta a la paradoxa de Moravec, proposada a mitjan anys vuitanta per l'investigador en robòtica Hans Moravec: «És més fàcil construir ordinadors que facin tasques complexes que màquines que imitin la capacitat de percepció i mobilitat d'un nen d'un any.» A això, hi podem afegir l'afirmació del lingüista Steven A. Pinker, que el 1994 va concloure que «la lliçó principal de 25 anys d'IA és que els problemes difícils són fàcils i els fàcils són difícils».

## Intel·ligència sense cervell

Un dels problemes és la dificultat de definir *intel·ligència*. I només faltava l'aparició del concepte d'intel·ligències múltiples, que potser seria més correcte denominar *capacitats* o *habilitats específiques*, o de la intel·ligència emocional i d'altres idees i propostes. O el debat sobre la intel·ligència dels animals no humans. Un dels objectius d'un comportament intel·ligent és aconseguir sobreposar-se als obstacles que l'entorn li posa per desenvolupar les funcions bàsiques: menjar, dormir, reproduir-se, protegir-se de fenòmens físics, d'imprevistos..., a banda de poder comprendre el món –o l'univers– i els seus integrants i les lleis físiques que els governen.

Per això, qualsevol debat sobre el nivell d'intel·ligència dels dofins o dels pops en comparació amb els humans pot resultar absurd, perquè l'entorn a què s'enfronten aquests animals no té res a veure amb el que ha encarat l'espècie humana al llarg de milers d'anys. Els ximpanzés són intel·ligents i tenen cultura, com ho demostra el fet que poblacions de llocs diferents han desenvolupat eines i processos diferents, perquè s'han enfrontat a necessitats diverses i no han disposat dels mateixos mitjans. Una mostra de la intel·ligència dels humans pot ser haver descobert la forma de fer foc i haver inventat la roda. És evident que un dofí ni pot fer foc ni fabricar una roda, i aquestes coses no li solucionarien cap problema vital.

Però, cal un cervell per ser intel·ligent? La resposta sembla òbvia, però no ens precipitem.

Cesca Cufí-Prat es va graduar en Enginyeria Aeroespacial a la UPC. Treballa a Airbus, on ara es dedica a desenvolupar instruments d'alta precisió per a l'observació de la Terra.

La Cesca ha caminat per Mart. Com us podeu imaginar, no es tracta d'un viatge real al planeta vermell, sinó d'una simulació feta amb el màxim rigor científic. Va formar part de la primera expedició catalana i totalment femenina a Mart, en el marc del projecte *Hypatia*, que té com a objectiu inspirar els joves, sobretot les noies, perquè facin carreres STEM. A l'equip hi havia també Helena Arias, estudiant de Física i Enginyeria Mecànica/Elèctrica de la UPC, que no es va «enlairar» cap al planeta vermell i formava part de l'equip de suport. Viatjarà amb la missió *Hypatia II* l'any 2025. L'Helena investiga també al Sincrotró ALBA i és una practicant destacada de tir olímpic, esport en què ha guanyat algunes medalles.

L'any 2023, les set integrants d'*Hypatia I* van passar dotze dies a la *Mars Desert Research Station*, situada al desert de Utah, als Estats Units, creada per investigar en un entorn que té algunes similituds amb el de Mart. I una de les tasques de la Cesca era investigar el *Physarum polycephalum*.

Es tracta d'un mixomicet, tipus d'organismes coneguts popularment com a floridures. En anglès, també en solen dir *blobs*, per la pel·lícula de terror *The Blob* (1958). Fa uns pocs centímetres de diàmetre i es mou a velocitats d'uns pocs centímetres per hora. I tot i ser unicel·lular, té capacitat d'aprenentatge.



*Figura 1: Imatge del Physarum polycephalum (Font: Bill Sheehan / Wikimedia Commons / CC-BY-SA-3.0)*

Així, quan es troba a l'estadi de plasmodi i ha de buscar aliment, explora el seu entorn i és capaç de memoritzar les rutes que ha seguit i trobar el millor camí cap al menjar. A banda d'això, és molt resistent i un bon candidat per estudiar com un ésser viu respon a les dures condicions de Mart i obtenir més pistes sobre l'origen de la vida a la Terra i les seves primeres etapes.

En aquest cas, un organisme unicel·lular, òbviament sense sistema nerviós ni cervell, és capaç de memoritzar rutes cap al menjar i triar la més adient. Aquesta és una conducta que podríem titllar, almenys, d'una mica intel·ligent.

Si pugem una mica en l'arbre de la vida cercant un organisme més complex, ens trobem amb la *Tripedalia cystophora*, una de les anomenades «meduses caixa» – per la forma del cos –, que viu al Carib i a l'Indo-Pacífic central. Investigadors de la Universitat de Copenhaguen

van publicar l'any 2023 un estudi a *Current Biology* en què descriuen les capacitats d'anàlisi i d'aprenentatge d'aquest organisme tan senzill. Les meduses i els seus parents més propers, anomenats en conjunt *cnidaris*, devien ser dels primers animals a desenvolupar un sistema nerviós, que, per això, és molt simple. No tenen un cervell centralitzat i la *Tripedalia* té quatre estructures que s'hi assemblen, amb un miler de cèl·lules –el cervell humà té cent mil milions de neurones.

És clar que té 24 ulls, i potser així compensa una mica el seu poc cervell amb un sistema de visió més complex. Però el que han descobert els investigadors danesos és que als manglars on viu, per calcular bé la distància fins al lloc on vol capturar menjar, utilitza el contrast, basant-se en com de fosques són les arrels en relació amb l'aigua. Així, pot acostar-s'hi i allunyar-se'n en el moment precís, per tal de capturar el menjar sense lesionar-se amb les arrels.

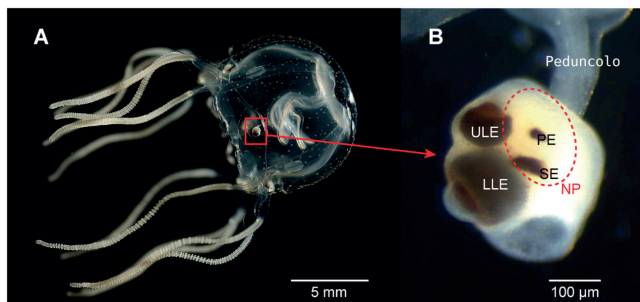


Figura 2. Sistema visual de la *Tripedalia cystophora* (Font: Jan Bielecki, Alexander K. Zaharoff, Nicole Y. Leung, Anders Garm, Todd H. Oakley / Wikimedia Commons / CC-BY-SA-4.0)

El contrast pot canviar cada dia i els investigadors han vist que la medusa l'interpreta combinant impressions visuals i sensacions, i així adapta la seva velocitat i estratègia al moment.

Així doncs, un organisme unicel·lular i una medusa amb un sistema nerviós molt simple tenen conductes que els permeten aprendre i establir estratègies. I podríem dir que aquestes estratègies són intel·ligents. Però, això els fa intel·ligents? No. Per tant, a l'hora de dir que una màquina és intel·ligent, hem d'anar amb cura.

La doctora Karina Gibert és catedràtica de la Facultat d'Informàtica de Barcelona (FIB) i dirigeix l'IDEAI, centre de recerca en ciència de dades i intel·ligència artificial de la UPC. A l'IDEAI, treballa en línies molt diferents, que cobreixen tota la cadena de valor de la intel·ligència artificial, des del disseny de xips específics per a la IA fins a l'assessorament en l'elaboració de polítiques públiques, i fa molt d'èmfasi en l'ètica i en la responsabilitat social.

A la pregunta sobre la intel·ligència de la IA respon que aquesta ciència pretén imitar el raonament humà i que «les versions inicials de la IA» estaven orientades als raonaments i als processos cognitius. Tanmateix, «amb això no n'hi ha prou, perquè necessites també captar el coneixement implícit, sentir emocions, curiositat, etc. Les dificultats per incorporar-hi aquests aspectes van provocar l'«hivern» de la IA.<sup>3</sup> Més tard, el

---

3 Es coneix amb aquest nom el període de final dels anys seixanta i de tota la dècada dels setanta en què es va constatar que els objectius eren molt més difícils d'assolir del que es creia i es van reduir els fons per a la recerca, arran de la manca de resultats.

focus es va desplaçar cap a les dades i el valor que se'n podia extreure.»

I, si les màquines no són intel·ligents o no ho són prou, poden cometre més errors. I això planteja molts dubtes i obliga a ser cauts i a adoptar mesures per fer-hi front. Ho veurem al capítol següent.



## D'ètica i biaixos

Com que el tema de la intel·ligència de les màquines és controvertit, és lògic que sovint es doni per bo que sí que en tenen. Almenys algunes ho semblen. I, per això, fins i tot es planteja la pregunta de quan apareixeran màquines més intel·ligents que els humans. És el que es coneix com a *singularitat tecnològica* o, simplement, *singularitat*. No es tracta només que s'assoleixi aquesta superintel·ligència, sinó que en aquell moment el desenvolupament tecnològic es podria escapar de les mans dels humans i no el podríem controlar.

I quan arribarem a la singularitat? Les previsions han estat molt diverses. Fa dècades, es va dir que seria al voltant de l'any 2000. Altres pronòstics, que s'han anat modificant periòdicament, la situaven als primers anys del segle XXI i algun, al 2021. D'altres més prudents diuen que es troba més lluny. Hans Moravec

pronostica que per l'any 2040 s'assolirà un nivell d'intel·ligència igual a la humana i que al 2050 ja serà d'un nivell superior. L'any 2005, l'informàtic Ray Kurzweil va publicar el llibre *The Singularity Is Near: When Humans Transcend Biology* (publicat en castellà a Lola Books, 2012), però, malgrat el títol, la situava cap al 2045.

Si està tan lluny, per què ens en preocupem? La pregunta és una mica il·lògica, perquè dues dècades –suposant que Kurzweil o Moravec l'encertin– no és gaire temps si volem preparar-nos per als seus efectes, que, segons alguns pronòstics, podrien resultar devastadors. Una altra cosa és si realment això pot tenir conseqüències tan catastròfiques.

Si fem cas de la Declaració sobre el risc de la IA,<sup>4</sup> promoguda pel Center for AI Safety (CAIS), que es va donar a conèixer al maig del 2023 i que ha estat signada per centenars de científics, tecnòlegs i executius relacionats amb la intel·ligència artificial, l'amenaça és ben real. El text és molt breu: «Mitigar el risc d'extinció degut a la IA hauria de ser una prioritat global al nivell d'altres riscos a escala social, com les pandèmies i la guerra nuclear.»

En fan un gra massa? Hi pot haver una IA que acabi essent tan intel·ligent? La doctora Gibert considera que la IA «ens pot condicionar i fer molt de mal». Però afegeix que «no m'atreviria a dir que ens pugui arribar a dominar». Més endavant, li tornarem a preguntar sobre aquest mal que ens pot fer, però si insis-

---

4 Vegeu: <https://www.safe.ai/statement-on-ai-risk>

tim en la possibilitat que ens elimini, voluntàriament o sense voler, hem de reprendre el debat sobre si les màquines són intel·ligents.

Júlia Pareto és filòsofa i està fent la tesi doctoral sobre l'ètica en la robòtica i la intel·ligència artificial a l'Institut de Robòtica i Informàtica Industrial (IRI) de la UPC i el CSIC i a la Facultat de Filosofia de la Universitat de Barcelona. Explica que la intel·ligència està relacionada, des del punt de vista etimològic, amb la capacitat de comprendre i de saber triar: «La intel·ligència té a veure, necessàriament, amb la capacitat de preguntar-se per la finalitat, amb la pregunta del per a què. Les màquines no ho fan i, en aquest sentit, els manca la intel·ligència pròpiament dita.»

Per tant, una màquina malèvola probablement no seria possible, llevat que hagués estat dissenyada amb aquesta finalitat. I afegeix que «si la màquina *ens domina* és perquè, ja sigui en les seves finalitats o en com les duu a terme, entra en xoc amb el que volen els humans».

Posa un exemple que es dona sovint com a demostració del segon cas: «Si a un sistema d'IA se li demana com podem acabar amb els problemes ecològics, pot proposar eliminar la humanitat, que n'és la gran causant. Però això seria, òbviament, una deducció no desitjable.»

És clar que aquest sistema podria arribar a una altra deducció. Tenint en compte que la IA consumeix una gran quantitat d'energia i això té un gran impacte ambiental, una part de la solució podria ser deixar de fer

servir la IA? Potser la IA acceptaria que les màquines també són un greu problema ambiental i que també se les hauria d'eliminar...

Però seguim amb la por que ens dominin. Júlia Pareto afirma que «gran part del debat ètic actual sobre els grans perills es fa sobre la base d'uns desenvolupaments tecnològics que encara no existeixen. I aquest debat és nociu, perquè ens aparta del debat sobre els temes reals i actuals i ens distreu d'aquests riscos, com ara qui controla les dades, com evitar la manipulació a partir de les dades, què fem per impedir la suplantació de la identitat a través de la veu o de la imatge...»

Tot i així, afegim que el debat també pot ser interessant perquè obliga a anar a les arrels: «La discussió sobre els drets dels robots, per exemple, tot i que no és un tema prioritari en l'agenda de l'ètica de les tecnologies, és una oportunitat per reflexionar sobre les condicions a partir de les quals atorguem l'estatus de subjecte moral a un individu o a una entitat. Ho fem en virtut de si tenen determinades propietats intrínseques vinculades a l'agència moral? O ho fem, en canvi, des d'una aproximació relacional, en virtut del que signifiquen per a nosaltres? «És una oportunitat per debatre com atorgar-los l'estatus de subjecte moral. Però, aleshores, també se'ls han d'imposar obligacions? A ells o als seus creadors?»

De robots, en parlarem molt al capítol corresponent. Vegem ara els problemes ètics de la IA. Marc Alier és professor agregat a la UPC, on coordina el programa de doctorat d'Educació en Enginyeria, Ciències i Tecnologia. Entre els seus temes de recerca,

hi ha l'enginyeria dels sistemes d'informació, les tecnologies –com la IA– aplicades a l'aprenentatge i l'ètica de les tecnologies de la informació.

Considera que moltes coses que s'han dit recentment sobre l'ètica de la IA «són sopars de duro». I ho argumenta així: «Es basen en el principi de l'excepcionalitat tecnològica. Com que aquesta tecnologia és tan diferent, hem de buscar unes regulacions i unes bases ètiques diferents, perquè les que teníem fins ara no serveixen. Però això no és necessàriament cert, tot i que sona molt bé en conferències i taules rodones.»

Per al doctor Alier, «moltes coses que s'inclouen dins la nova etiqueta d'ètica de la IA estan proposades per les mateixes grans empreses que tenen interessos en la IA». I, abans de prosseguir amb l'argument, em dona a conèixer un concepte que va proposar l'economista Bruce Yandle l'any 1983: el dels *bootleggers* (traficants) i els baptistes. Yandle va observar que grups amb plantejaments molt diferents poden defensar la mateixa proposta moral i política. Així, a principi del segle xx, els baptistes i altres grups cristians evangèlics volien que es prohibís o es restringís la venda d'alcohol, per diversos motius morals o religiosos. Per la seva banda, els traficants també volien que es prohibís o es restringís perquè, com que vivien del comerç il·legal, aquestes mesures els proporcionarien grans beneficis. De fet, la «lleï seca» és a la base de les grans fortunes i del poder que van assolir alguns grups mafiosos als Estats Units –Al Capone en seria un bon exemple.

I que té a veure això amb la IA? Alier creu que persones com Sam Altman, director executiu d'OpenAI, o Elon Musk, es mostren en entrevistes, *TED Talks* i *podcasts* molt preocupants pel risc d'extinció de la humanitat que suposa la IA general que vindrà en pocs anys. Sam Altman és partidari «que es prohibeixi fer processos d'entrenament de models que a OpenAI ja han fet i que, a més, volen controlar ells mateixos». Per la seva banda, Elon Musk promou manifestos, que signen molts investigadors de renom, en què es demanen moratòries de sis mesos per a l'entrenament de models massius, però no diu que això li permetrà atrapar els seus competidors més avançats. Per tant, Altman i Musk són més aviat els *bootleggers* –és a dir, els traficants, que volen afavorir el seu negoci–, i no els baptistes. La seva intenció sembla que és continuar controlant el camp de la IA, més que estar profundament preocupats pel futur de la humanitat.

Alier afegeix que Sam Altman i Elon Musk «avisen del risc existencial de la IA, mentre hi treballen a sac, però comparteixen discurs amb baptistes com Tristan Harris o Yuval Noah Harari». Aquest darrer escrivia al maig de 2023 al *New York Times*: «Hem convocat una intel·ligència alienígena. No en sabem gairebé res, excepte que és extremament poderosa i ens ofereix regals enlluernadors, però també podria destrossar a cops de destrall els fonaments de la nostra civilització.» En una entrevista a *Fortune*, fins i tot parla del «possible final de la història humana –el final del període dominat pels éssers humans».

Alier també posa l'exemple de YouTube com a mostra que algunes mesures suposadament ètiques de les grans companyies en realitat serveixen per evitar la competència. YouTube es va fer gran permetent als seus usuaris compartir vídeos de forma indiscriminada i era, sovint, un canal de compartició de música i pel·lícules amb *copyright*. Quan YouTube ja havia esdevingut líder del mercat i fou adquirida per Google, va arribar a acords amb els grups propietaris de grans quantitats de continguts sotmesos a drets d'autor. Va participar, mitjançant *lobbies*, en processos d'elaboració de normatives, assegurant-se que les noves regulacions sobre drets d'autor i serveis de *streaming* encaixessin amb les tecnologies de detecció de continguts amb *copyright* que ells estaven desenvolupant.

Segons Alier, la jurisprudència respecte a les lleis de *copyright* que fan responsables a les plataformes de les infraccions «ha portat a impedir o ha dificultat molt que algú altre creï plataformes de *streaming*, perquè si els usuaris comparteixen continguts amb *copyright* a la teva plataforma tu ets el responsable de vetllar que això no passi». I qui s'arrisca?

Per tot això, i per altres raons, és important que els futurs enginyers tinguin nocions d'ètica. Alier diu que això «és molt important i agrada als estudiants d'enginyeria, perquè els fa reflexionar sobre l'impacte de la seva feina. Però és difícil d'ensenyar i s'ha de fer veure aquesta connexió». Ho va començar a fer el professor Miquel Barceló (1948-2021) quan va crear l'assignatura a la Facultat d'Informàtica de Barcelona (FIB), que

ara imparteix Alier, junt amb la professora María José Casañ.

I els filòsofs? És necessari que participin en el debat, però per fer-ho potser també necessitarien en els seus graus alguna assignatura de ciència i tecnologia bàsica. I els economistes i els advocats i politòlegs, també. Això és una opinió meva, però Alier no ho veu malament, perquè per debatre aquests temes ètics «cal conèixer molt bé la tecnologia. Si no, porta a plantejaments que no tenen cap validesa ni utilitat, perquè parteixen d'un desconeixement de la tecnologia».

Per ensenyar ètica als enginyers, cal, doncs, fer-los veure que afecta la seva feina –i que, a més, tenen un deure amb la societat. La forma d'impartir la matèria ja és una altra cosa, però potser un dels punts interessants es pot extreure de la feina feta a l'assignatura impartida per Alier i José Casañ, en què es treballen els aspectes socials, ètics i ambientals, i també s'estudia la història de la informàtica.

La història de l'assignatura la van exposar en un article publicat a *Sustainability* al juliol de 2020 Alier, José Casañ i Ariadna Llorens, que imparteix el curs Economia, Ètica i Societat de la Informació al Grau d'Enginyeria Informàtica de l'Escola Politècnica Superior d'Enginyeria de Vilanova i la Geltrú. Els tres autors hi explicaven que a la FIB la sostenibilitat i el compromís social han esdevingut una competència transversal que es treballa en diverses matèries, mentre que els aspectes relacionats amb l'ètica i amb la història de les TIC només es tracten al curs Aspectes Socials

i Mediambientals de la Informàtica. Es tracta d'una assignatura opcional que cursen uns 50 alumnes dels prop de 200 que es graduen cada any.

Per això, conclouen que «cada nou currículum fa més èmfasi en l'especialització tècnica» i que «els aspectes humanístics estan virtualment absents dels estudis». I afegeixen que «aquesta hiperespecialització, l'absència de les humanitats i el tractament dels aspectes relacionats amb la responsabilitat social dels enginyers com a matèria opcional són un error i causen una mancança en la formació dels futurs enginyers».

No obstant això, com explica Alier, als darrers anys l'assignatura ha passat a fer-se de forma quadrimestral, amb la qual cosa se n'ha doblat la freqüència, i s'hi ha afegit una nova assignatura curta de lliure elecció, específica de temes ètics. El més important de l'article, segons Alier, «és que a la FIB hem estat més de 30 anys ensenyant-ne els aspectes socials, ètics i professionals, la història de la informàtica i la sostenibilitat, i això ha implicat molta feina de desenvolupament del temari i de continguts, i d'experimentació amb metodologies docents, així com l'adaptació als nous plans d'estudis».

Sobre la forma d'ensenyar els aspectes ètics i el pensament crític, Alier i Casañ van presentar un treball al congrés TEEM 2022 («Technological Ecosystems for Enhancing Multiculturality»), que va tenir lloc a Salamanca, i van obtenir el *Best Paper Award* en Educació de l'Enginyeria. En el moment d'escriure aquest text, està pendent de publicació una versió més extensa d'aquesta ponència a la *Revista Iberoamericana*

*de Tecnologías del Aprendizaje* (RITA), de l'Institute of Electrical and Electronics Engineers (IEEE).

Alier i Casañ hi expliquen l'activitat pedagògica desenvolupada per facilitar la participació dels estudiants amb vista a aplicar la teoria als reptes professionals i socials i potenciar el pensament crític. Els estudiants participen en un debat en què se'ls assigna una postura moral, que han de mantenir aplicant una de les teories ètiques exposades a classe. Els resultats mostren, segons els autors, l'eficàcia del mètode, que es podria aplicar a altres camps de l'enginyeria. Integrant aquestes activitats pedagògiques en els programes d'enginyeria, afirmen, «els educadors poden enfortir en els estudiants els valors ètics essencials i les habilitats per gestionar temes ètics intricats en la seva esfera professional i social, i alhora fomentar el pensament crític i encoratjar el diàleg entre els estudiants».

En un treball presentat el 2023 a les XXIX Jornades sobre l'Ensenyament Universitari de la Informàtica (JENUI) a Granada i publicat al volum 8 de les *Actas de la JENUI*, Alier i Casañ també varen mostrar que substituir les classes expositives del professorat per metodologies actives augmentava l'interès i l'atenció dels alumnes i disminuïa l'absentisme. Així, potser un ensenyament més actiu, que doni més paper als estudiants, pot resultar molt positiu en matèries que, tot i ser interessants, són optatives i potser han de tenir algun valor afegit que les faci més atractives. Sens dubte, són assignatures importants per a la formació dels futurs enginyers.

Una alternativa eficaç ja comprovada és fer servir narracions de ciència-ficció, com veurem al capítol sobre els robots.

### **Simular el sentit comú?**

Al seu llibre *La intel·ligència artificial explicada als humans* (Plataforma Actual, 2023), Jordi Torres, catedràtic de la UPC i investigador del Barcelona Supercomputing Center (BSC), explica que sistemes basats en la IA específica, anomenada *IA feble* o *estreta* en la literatura científica, han mostrat ser tan bons com un humà a l'hora de resoldre problemes específics. I que, per a alguns investigadors, sistemes com el ChatGPT són l'avantsala de la IA general, o *IA forta* en la literatura acadèmica, que podria pensar igual que les persones.

Però també explica que un dels reptes d'aquests sistemes és aconseguir que resolguin problemes per als quals no han estat entrenats específicament o que requereixen improvisació. I una de les limitacions més clares és que «els manca un element difícil de parametritzar en termes científics: el sentit comú».

Efectivament, les persones poden trigar a reaccionar o a fer determinats càlculs, però també saben que en algun moment es poden saltar les normes si la situació ho requereix. Si portem una persona ferida o que ha patit un infart a l'hospital, potser ens saltarem alguns semàfors en vermell. Ho farem amb prudèn-

cia i sabent que, malgrat la nostra bona intenció, si es produeix un accident serà per culpa nostra. Però comprenem que, en aquelles circumstàncies, està justificat.

Un cotxe autònom no farà el mateix si no se l'ha programat perquè se salti la norma. Se li poden introduir en el programa algunes instruccions perquè en determinats casos faci una cosa o una altra, però és ben difícil preveure totes les situacions possibles i fer que triï la més adient. Un exemple típic és pensar si, en una situació extrema, el vehicle autònom atropellaria una persona gran o un nen –hi ha moltes variacions sobre aquest problema. És molt probable que una persona intentés no atropellar a cap dels dos, tot i que això fos a costa de fer mal a tercers o, fins i tot, de fer-se mal a si mateixa. Faria el mateix una màquina?

S'atribueix a Voltaire la frase: «El sentit comú no és tan comú.» I algú va dir que és el menys comú dels sentits. Però, tot i això, com es demostra sovint, és una eina que ens permet fer eleccions allà on no arriba el coneixement, en situacions que no són prou clares o que resulten inesperades.

El sentit comú també permet constatar, sense que ens ho hagin d'explicar, que determinades relacions aparents no impliquen cap influència causa-efecte entre dues coses. Per exemple, buscant dades, descobrim que, quan disminueix el consum de margarina per càpita als Estats Units, també baixa el nombre de divorcis a l'estat de Maine. En canvi, quan augmenta el

consum de formatge per càpita, també mor més gent que s'ha enredat al llit amb els seus llençols. Quant més petroli importen els Estats Units de Noruega, més conductors morts hi ha en xocs de trens. I també hi ha forta correlació entre el nombre de persones ofegades en una piscina i el nombre de pel·lícules en què apareix Nicolas Cage.

Són exemples del que s'anomenen *relacions espúries*. Dues coses augmenten o disminueixen al mateix temps amb una forta correlació, però no hi ha cap lligam causa-efecte entre una i l'altra. En alguns casos, pot ser que hi hagi un tercer fet que influeixi en els altres dos i que en sigui la causa real. Però, en molts casos, és pura casualitat.

No sembla gaire difícil per a una persona veure que les dues magnituds no tenen res a veure. A ningú amb sentit comú –suposem– se li acudiria proposar que, per evitar ofegaments a les piscines, es prohibís fer pel·lícules a Nicholas Cage. O a ningú que volgués reduir el nombre de divorcis a Maine se li acudiria intentar aconseguir-ho recomanant rebaixar el consum de margarina. N'hi ha molts més exemples a: <https://www.tylervigen.com/spurious-correlations>

Aquests exemples són anecdòtics i, si algú se'ls prengués seriosament, és probable que no hi haguessin conseqüències negatives –llevat que algú pensés a eliminar Cage per salvar moltes altres vides. El problema es dona quan una correlació s'interpreta malament i se li dona l'estatus de relació causa-efecte.

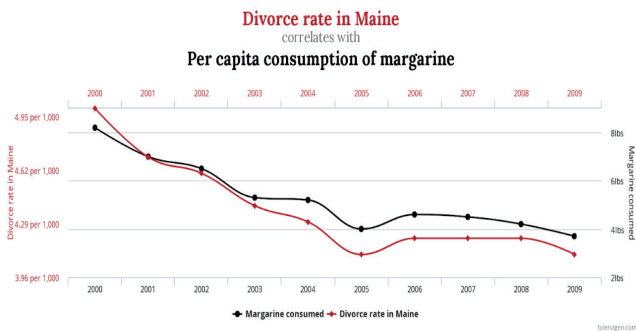


Figura 3. Font: <https://www.tylervigen.com/spurious-correlations>

Això és bàsic en els estudis epidemiològics perquè, si hi ha causes intermèdies que influeixen en les diferents magnituds, s'han d'identificar i s'ha de quantificar el seu paper, perquè altrament és impossible establir mesures preventives o terapèutiques adequades o esbrinar quins grups de la població són més proclius a determinades malalties. Les raons poden ser biològiques, però també socials, econòmiques i culturals.

Un exemple interessant que vaig comentar al meu llibre *L'individu transparent. Dels raigs X al big data* (Pagès Editors, 2016) és la possible relació entre el dispositiu intrauterí (DIU) i la infertilitat o la malaltia inflammatòria pelviana en les dones que utilitzen aquest sistema per evitar embarassos. En diversos estudis, es va poder descartar l'associació quan es controlaven factors com un historial de gonorrea o el

nombre de companys sexuals recents. La causa no era el DIU en si, sinó el fet que entre les dones que l'utilitzaven –evidentment, no totes– eren més freqüents les relacions sexuals amb més parelles diferents. Això incrementava el risc de patir una malaltia de transmissió sexual, cosa que, al seu torn, podia provocar infertilitat.

Un altre exemple el comento al llibre *L'imperi de les dades* (Bromera, 2023, 2a edició). Un estudi relacionava determinats fàrmacs contra el reflux gastrointestinal amb més risc d'infart de miocardi. Com que l'estudi s'havia fet a partir d'anotacions mèdiques i no de grups totalment controlats, investigadors de l'Hospital Parc Taulí de Sabadell i de la Universitat Autònoma de Barcelona van voler estudiar si hi podia haver hagut alguns biaixos que portessin a conclusions equivocades. I, a banda de constatar que el treball no complia alguns dels criteris dels estudis epidemiològics, posaven un exemple d'un possible biaix: els fumadors solen tenir cremors més sovint, però també tenen més risc d'infart. Per tant, l'increment del risc d'infart no vindria dels fàrmacs, sinó que es donaria entre grups específics de pacients que els prenen.

Tot això mostra que, a l'hora d'establir relacions, cal tenir en compte molts factors. En aquest sentit, la IA pot ser un gran ajut, perquè amb la seva capacitat d'establir patrons i trobar relacions, i la velocitat a què pot fer-ho amb milions i milions de dades, pot arribar a deduir coses que s'escapen a l'ull humà o que trigarien molt de temps a ser descobertes. Però, després, cal que

la persona responsable de la recerca miri com s'interpreten les dades i què se'n pot fer.

Per establir aquestes relacions o per prendre decisions, no n'hi ha prou que els algorismes estiguin ben dissenyats i que hi hagi ordinadors molt potents –dos dels elements bàsics que han permès el gran desenvolupament de la IA als darrers anys–, sinó que la qualitat de les dades ha de ser molt bona. I cal saber molt bé què es demana als algorismes.

Sobre el tema de les dades, en presentem un altre exemple. Un sistema d'IA s'alimenta amb la informació que li posem a l'abast. I aquesta és un reflex dels desequilibris de la societat. Imaginem que un algorisme ha de decidir sobre la llibertat provisional d'una persona blanca i d'una afroamericana. Sistemes així ja s'utilitzen. Vegem aquestes dades aparegudes en un estudi de l'any 2016, però que probablement no han canviat gaire: el 63,7% de la població dels Estats Units és blanca, el 16,6% és hispana i el 12,2%, afroamericana.

Ara, entre la població reclusa, hi ha un 46% d'afroamericans, un 39% de blancs i un 19% d'hispanos. L'algorisme que ha de decidir pot deduir que, per alguna raó, els afroamericans són més proclius a delinquir. I això porta a resultats esbiaixats, que s'han denunciat en diversos articles, d'un sistema que juga amb la llibertat i amb la reinserció de moltes persones.

Donem la paraula de nou a Marc Alier. Ens explica que els sistemes són molt bons trobant patrons. Fins fa vint anys, diu, «era molt difícil crear algorismes per trobar patrons i relacions. El patró havia d'estar en l'al-

gorisme mateix; ara, les dades i els patrons inherents alteren o generen l'algorisme».

Actualment, aquests sistemes identifiquen molts més patrons que una persona i «també identifiquen patrons no evidents i ocults..., que poden restar ocults en el model».

Però això també pot ocasionar biaixos o discriminacions ocultes, adverteix. En la selecció de personal, un sistema pot, a partir de la informació que té, seleccionar menys dones o menys afroamericans, o fins i tot cap. Es pot establir que no s'introdueixi la dada del gènere o del grup ètnic, o reentrenar (*fine-tuning*) el sistema, per eliminar aquest biaix».

Tanmateix, això no elimina el problema, perquè «el sistema és prou potent per deduir aquesta dada del codi postal i el biaix pot seguir ocult en una discriminació per codi postal (ZIP als Estats Units). Pot establir relacions entre el ZIP i el nivell socioeconòmic o cultural i això porta igualment a biaixos, perquè ha trobat la informació o la relació que no se li ha donat explícitament».

És cert que el codi postal no sempre és determinant, però, un cop més, anem al conjunt de dades disponibles. Si, en alguns barris, hi ha més població afroamericana, el sistema pot deduir que un candidat d'aquella zona és afroamericà o té més probabilitats de ser-ho. I discriminar-lo, perquè a l'empresa no s'han contractat gaires treballadors afroamericans.

I, del biaix, no se'n salven ni les grans tecnològiques. L'any 2014, Amazon va crear un sistema per fer

més eficient la selecció de personal. Però, com que el va entrenar amb dades dels currículums dels treballadors que havia contractat als deu anys anteriors, el sistema va entendre que les dones havien de ser menys valorades i les discriminava. Simplement, perquè els deu anys anteriors Amazon havia contractat moltes menys dones que homes.

No sempre calen gaires dades. Fa molts anys, el 2000, un estudi de Latanya Sweeney, informàtica de Harvard, demostrava que la data de naixement, el codi postal i el sexe eren suficients per identificar el 87% dels ciutadans nord-americans. I qui no ha anat mai a un comerç i li han demanat la data de naixement per tal de felicitar-lo en el seu aniversari?

Això és un risc a la privacitat i depèn de les dades que el sistema utilitzi i de les instruccions dels algorismes. Però també cal educar la població perquè no faciliti determinades dades i no digui «Accepto» de seguida quan entra en una web o vol fer servir una aplicació. I perquè sàpiga discernir els usos bons de la IA dels inadequats. Per aconseguir-ho, a Catalunya es va crear la web [ciutadanIA.cat](http://ciutadanIA.cat).

Una dels continguts d'aquesta web és el curs CiutadanIA: formació en intel·ligència artificial per a tothom, que fou desenvolupat per especialistes de l'IDEAI «per facilitar a la ciutadania, en general, la comprensió de què és la IA i com manegar amb seguretat les IA que ens envolten».

Una de les expertes que va elaborar el curs és Karina Gibert. I una de les seves línies de recerca és l'anome-

nada *IA explicable*. L'objectiu és «que la IA pugui explicar i justificar les prediccions que fa d'una manera que els humans puguem comprendre fàcilment». La IA que ara està de moda, explica Gibert, agafa dades, busca patrons i troba les lleis que governen els processos. Però, si el mètode és opac, això «vol dir que has entrenat un sistema que et diu què passa, però no et diu per què passa. L'humà no es pot basar en això per decidir, perquè necessita argumentar la seva decisió».

Els sistemes simbòlics, afegix, són explicables. Els subsimbòlics o opacs són una caixa negra. Així, el *clustering* busca tipus de clients. «En general, dones dades i et donen grups. Però, per què un acaba en el mateix grup que un altre? Caldria saber quines associacions fan.»

Si fas models d'IA per donar suport a decisions, conclou, necessites saber-ne el perquè. I un altre tema és «si la resposta o la predicció és bona, perquè potser es basa en dades no prou bones o esbiaixades».

En el cas concret de la música, l'any 2023 es va presentar MusicData UPC, un programa cultural coproduït per Digitalfems, ElLESmusic i la UPC, en el marc del programa UPCArts. MusicData UPC combina la música en directe amb reflexions crítiques sobre l'impacte dels sistemes algorítmics de les plataformes digitals. Així doncs, una part se centrarà en la suposada neutralitat de les tecnologies, particularment en l'àmbit dels sistemes algorítmics musicals.

Molts estudis han mostrat que els algoritmes de recomanació musical poden amplificar biaixos, com

ara el de gènere. Es recomanen en algunes plataformes moltes més obres d'artistes masculins que femenins? Tot i que Spotify té un 44% d'usuàries femenines, només el 22,8% de les reproduccions corresponen a artistes femenines o mixtos. I si les dones escolten el 32,0% d'artistes femenines o mixtos, els homes només n'escolten el 18,3%. Un informe elaborat pel Joint Research Centre de la Comissió Europea, donat a conèixer el 2018, demostrava que l'algoritme de recomanació utilitzat en aquesta plataforma tendia a seleccionar música d'artistes masculins.

El problema de la perspectiva de gènere va molt més enllà de l'àmbit musical. I els biaixos en els salaris, les oportunitats laborals o l'accés a l'atenció mèdica podrien passar desapercebuts. La seva identificació pot donar a conèixer disparitats poc visibles.

MusicDataUPC col·labora amb el GenderDataLab, un espai d'experimentació on es duen a terme projectes d'innovació relacionats amb la perspectiva de gènere en el món de les dades. L'objectiu és concebre i desenvolupar un sistema algorítmic amb enfocament de gènere, de codi obert. Així, s'exploraran alternatives per fomentar la igualtat de gènere.

## **Un univers per a cadascú**

Carles Sora és doctor en Comunicació Interactiva per la Universitat Pompeu Fabra i especialista en mitjans interactius i cultura digital. Des del 2020, dirigeix

el Centre de la Imatge i la Tecnologia Multimèdia (CITM) de la UPC, situat a Terrassa.

M'interessava molt parlar amb ell, com a expert en les interfícies persona-màquina, per conèixer la seva visió sobre l'adaptació personalitzada del metavers. Però la personalització, em va dir, ja és aquí. Al seu despatx, em va explicar que les dades que aportem ja estan creant un perfil que ens fa rebre informació diferent de la que s'envia a altres persones: «Ara ja vivim una captació de dades que fa que a cadascú li arribi informació diferent. A les xarxes socials, amb les *cookies* i l'acumulació d'informació personal, ja ens fan recomanacions sobre la base del que som, i no només del que consumim.»

I ho vivim directament cada dia. Netflix, per exemple, mostra imatges diferents d'una mateixa pel·lícula o sèrie segons el país d'on sigui l'usuari i, fins i tot, segons els gustos o les característiques de la persona. Així, cadascú pot veure fotos diferents o tràilers diferents.

Les recomanacions de consum poden tenir un paper més o menys invasiu, però els perfils personals s'han utilitzat de manera espúria en casos com el de Cambridge Analytica, una empresa que va interferir, entre altres processos, en el referèndum del Brexit al Regne Unit i en les eleccions presidencials als Estats Units del 2016, que va guanyar Donald Trump. A partir de dades extretes de Facebook, suposadament per fer un estudi, va elaborar perfils de votants als quals es podia dirigir publicitat molt personalitzada, incidint en els temes que els farien proclius a votar Trump.

Bàsicament, es tractava d'electors susceptibles de canviar el vot, que tant podria anar a Hillary Clinton com a Donald Trump. Així no es malbarataven esforços i el discurs era prou personalitzat per convèncer amb els temes que eren més rellevants per a cada ciutadà, ja fos l'avortament, l'atur o la inversió militar.

I a quines divisions personalitzades ens pot portar el futur immediat? Per fer un tast d'un futur exagerat, però que l'experiència fa que no descartem del tot, em convida a veure un vídeo de poc més de sis minuts elaborat pel dissenyador Keiichi Matsuda, titulat *Hyper Reality*. Es pot veure a <http://hyper-reality.co/>, però és fàcil de trobar en altres llocs d'internet.

El curt és impressionant i angoixant. Es tracta d'una prova de concepte que visualitza la connexió entre interfícies. Gravat amb punt de vista subjectiu, mostra un escenari on la realitat i la virtualitat es barregen, i la informació ens arriba des de pantalles incorporades que apareixen i desapareixen amb un simple moviment de dits, que recorden *Minority Report* i la interfície controlada pels moviments de les mans de l'usuari. La realitat virtual, la realitat augmentada i altres tecnologies han creat una societat en la qual, com si fos un joc, cada persona té uns punts, que pot incrementar o perdre en la seva activitat diària. La compra es transforma en un passeig on els consells, més que assessorar, aclaparen i fins i tot angoixen. Hi juga un paper destacat la ciberdelinqüència, amb més possibilitats d'acció i de provocar conseqüències greus en un món tan tecnificat.

I això porta Sora a donar-me la seva visió d'un futur possible: «Generaran una figura virtual de com ets tu, amb algorismes.» Tampoc no n'estem tan lluny, com em mostra parlant de les ulleres d'Apple. Es tracta, explica, d'una unió de tecnologies. «Prediuen què faràs, amb un sistema de càmeres que llegeixen els moviments oculars. El sistema informàtic és més ràpid que el nostre sistema cognitiu. La navegació és fluida i no hi ha retards en la visió que t'ofereix.»

És a dir, saben de nosaltres més que nosaltres, fins i tot les nostres intencions immediates. «Les interfícies són cada cop més immersives i això t'allunya de la realitat. Hi ha una hiperrealitat centrada en aquests perfils, amb l'ús d'interfícies. Les tecnologies ja hi són», afegeix.

Aleshores, amb el metavers, que promet realitats virtuals on immemir-nos, amb la nostra identitat o amb una altra, on podrem trobar l'entorn que s'adapti millor als nostres desitjos o, en tot cas, que sigui ben diferent de la nostra experiència quotidiana, cadascú es podrà crear un món a mida.

Els qui ja tenim una edat, podem recordar l'època en què no hi havia gaires possibilitats de triar què veure a la televisió. Per no anar més lluny, tot i que ja hi havia més d'un canal, no era difícil coincidir amb els amics i els companys per comentar la pel·lícula del dia abans o el darrer episodi de la sèrie que ens tenia atrapat. Ara, l'àmplia oferta, per no dir la multiplicitat de canals i la possibilitat de veure les coses on i quan ens va bé, pot dificultar el contrast d'opinions, però continua essent

factible i no gaire difícil trobar amb qui comentar els nostres programes preferits.

Però, què passarà si aquesta multiplicitat passa a donar-se a la vida quotidiana o al metavers? De moment, deixem de banda això darrer, perquè significaria que tots podríem fugir del món real i anar a un altre món, creat a mida, que podria no coincidir en res amb el dels nostres familiars o coneguts.

El risc és que ja hi ha ara, des de fa un temps, és que els missatges i les informacions que ens arriben també estan adaptats al nostre perfil, a les nostres idees. Ja no és que sàpiguen si som aficionats a l'esquí o al tennis i que ens enviïn publicitat personalitzada. El més perillós és que, si no sabem com sostreure'ns-en o no volem fer-ho, també ens envien només notícies i comentaris que reforcen les nostres idees.

Les societats han progressat sobretot quan hi ha hagut intercanvi d'idees i de plantejaments. Necessitem que ens arribin opinions diferents, per replantejar-nos les nostres i per conèixer què pensa qui opina diferent. Llavors, potser reflexionarem i continuarem pensant igual, però de forma més consistent. Potser veurem que una part dels nostres plantejaments s'han de modificar, poc o molt. Però, si rebem només els missatges que ens agraden, arribem a allò que s'anomena *filtre bombolla* o *cambrà d'eco*, com explica Sora. Aquest concepte fou encunyat per l'activista i emprenedor nord-americà Eli Pariser, que ve a dir que, com més perfilats ens tenen els sistemes, menys discrepància cultural o ideològica ens mostren, per tal d'evitar que ens qüestionem els nostres

valors i, per tant, «millorem» la nostra experiència en el sistema. Això ajuda a crear un aïllament intel·lectual.

Per això, com diu Sora, «si vius en realitats diferents, no surts de la teva visió i això limita el debat i l'intercanvi. Per sortir-ne, has d'abandonar la bombolla que t'han creat».

Aquest no era el futur que alguns somiaven fa dues dècades: «Suposadament, internet era la utopia de la xarxa oberta, democràtica, on tothom tenia la capacitat i la llibertat de crear i de comunicar. Però s'ha vist que això no ha anat així.» I, encara pitjor: «Més del 30% de la població mundial no té accés a internet i els continguts no són equitatius. Les fronteres digitals poden ser més fortes que les físiques.»

Caldrà, doncs, fer esforços per limitar aquestes fronteres i que no acabin tancant-nos en universos paral·lels incomunicats. Si viure en la bombolla de la informació dirigida ja és nociu, què podem dir d'uns mons on la virtualitat individualitzada juga un paper més important que la realitat?

## **Màquines al·lucinades**

Els sistemes d'IA també poden tenir al·lucinacions. Com que pretenem que aquests sistemes arribin a imitar el raonament humà, també donem un nom molt humà als problemes que pateixen. Els humanitzem. En les persones, una al·lucinació és una percepció falsa d'objectes o esdeveniments a través de

qualsevol dels nostres sentits: semblen reals, però no ho són. Tots en podem patir en algun moment i sovint no són greus, però en alguns casos poden revelar un problema mental.

En IA, una al·lucinació és una resposta que dona el sistema i que no sembla justificada per les dades que s'han utilitzat en el seu entrenament. Un sistema, com un *chatbot* pot donar una dada totalment absurda en un informe i, fins i tot, pot inventar-se dades o referències inexistents. Moltes vegades, la resposta és tan il·lògica que ja es veu que correspon a una al·lucinació. Però, en molts casos, pot semblar que la resposta és plausible i no en desconfiem.

Les causes poden ser la mala qualitat de les dades o el seu origen, o la forma com s'ha entrenat el model, perquè no pot verificar la informació i simplement utilitza les dades que se li han posat a l'abast, amb els patrons que li han donat. Així, el processament de les dades serà deficient i el sistema farà deduccions absurdes o, si més no, falses. El problema és prou important per preocupar les grans tecnològiques. Però pensem en els sistemes que utilitzem nosaltres o que utilitzen l'Administració o les empreses per prendre decisions. Deixar aquestes eleccions en mans de màquines que poden tenir al·lucinacions sembla un risc inacceptable. La solució no és senzilla, però passa pel control de les dades i dels sistemes per processar-les. En veurem alguns casos al capítol següent.

## La IA treballa. I nosaltres?

A banda del debat ètic, quin és l'impacte socioeconòmic de la IA? Em centraré en els llocs de treball. Pràcticament tots els avenços tecnològics per executar determinades feines han anat acompanyats del tenor que eliminin llocs de treball. Recordem els luddites, que apareguren a principi del segle XIX al Regne Unit i que destruïen màquines i fàbriques. O l'incendi de la fàbrica Bonaplata a Barcelona l'any 1835. Havia estat el primer vapor instal·lat a Catalunya i a tota la Península i la seva posada en funcionament el 1833 marca l'inici de la revolució industrial al nostre país.

El terme luddita –o neoluddita– encara es fa servir per qualificar la gent que s'oposa als avenços tecnològics. Però cal diferenciar l'oposició sistemàtica de la crítica als possibles impactes negatius de determinats avenços. Com veiem en aquest llibre, molts enginyers plantegen la necessitat d'establir un debat i de controlar la IA. I és obvi que no són luddites.

Al llarg de la història, els avenços tecnològics poden haver provocat, inicialment, una pèrdua de llocs de treball i la desaparició o la disminució dràstica d'alguns oficis, però a mitjà o a llarg termini normalment han creat més ocupació i creixement econòmic. Una altra cosa és si aquest creixement s'ha repartit equitativament. La resposta òbvia és que no.

Per tant, la tecnologia aconsegueix un paper cabdal, però al final l'impacte depèn de les decisions de la societat i de com actuen els seus integrants. La IA i la

robòtica poden deslliurar –i, de fet, en alguns casos, com els robots industrials, ja ho han fet– les persones de determinades feines pesades, tedioses o perilloses. I, si cada vegada hi ha més màquines, per a aquestes feines cada cop caldran menys persones.

Com que el progrés tecnològic també fa sorgir moltes professions i necessitats noves, al llarg de la història aquest fet normalment o molt sovint s'ha compensat amb escreix. Una de les coses que han de fer els treballadors és adaptar-se i aprendre nous oficis o noves tècniques. I, si desapareixen o davallen molt determinats negocis i sectors, cal pensar en nous productes i serveis.

Si la tecnologia acaba amb les feines que requereixen menys formació, caldrà que els treballadors es reciclin i adquireixin nous coneixements i habilitats. Però, com que cada vegada té impacte en més feines i hem d'acceptar que no tothom, per capacitat, per mitjans o per les causes que siguin, es pot reciclar i introduir en el món de les tecnologies innovadores, també caldrà pensar a donar a aquests sectors uns mitjans mínims, en forma d'un salari social o de qualsevol altra manera. També es poden donar a canvi de res, o bé perquè la persona farà feines senzilles però també necessàries, com ajudar o fer companyia a gent gran o a malalts, realitzar tasques solidàries en entitats.... Això té l'avantatge que la persona se sentirà útil, no passarà el dia sense fer res i aportarà un benefici acomplint unes tasques socials que sovint oblidem.

Però la IA afecta cada vegada més professions, fins i tot les que es consideren més creatives o que requereixen més formació. Així, es poden escanejar cossos d'actors i actrius i copsar-ne els moviments, i, més endavant, generar imatges per gravar seqüències sense la seva intervenció.

També es poden generar veus per doblar pel·lícules. I fer servir ChatGPT o Google Bard per escriure guions, que després un humà simplement repassarà.

Què pot comportar tot això? Potser actors, dobladors, ajustadors i guionistes es reciclaran i continuaran treballant. I el mateix faran els periodistes, els escriptors, els advocats, els metges i els economistes. Però, com que sovint els avenços tenen, com una de les primeres idees, rebaixar els costos i augmentar els beneficis, els temors d'aquests sectors són justificats.

En un article publicat a l'agost de 2023 al setmanari *New Scientist*, la periodista científica i escriptora Annalee Newitz explicava que alguns estudis havien volgut contractar actors i actrius per digitalitzar els seus cossos i gravar les seves veus, i que els els cedissin «a perpetuïtat». Un actor de veu deia que un estudi el volia contractar només 48 hores per contribuir a un generador artificial de veu. I feines que poden durar mesos, com la de guionista, poden fer-se en pocs dies, perquè només cal que l'expert repassi i poleixi el que ha escrit un sistema d'IA. Naturalment, tot això va acompanyat d'un sou proporcional al temps treballat.

També és cert que apareixen noves professionals i que moltes de les que hi haurà d'aquí a dues dècades ni

tan sols les podem arribar a imaginar avui. Així, avui ja hi ha els entrenadors o entrenadores d'IA, anomenats també *prompt engineers*, «enginyers d'apunts». Fa uns quants anys, els experts en *big data* Viktor Mayer-Schönberger i Kenneth Cukier van proposar una altra professió: la d'algoritmista. Es tractaria d'analitzar i certificar els algoritmes per garantir que funcionen bé i que no són discriminatoris.

Com que la IA (encara) no pot fer-ho tot i sovint es confon, també apareixen feines com etiquetar el que es veu en imatges fixes o en vídeos. Així, a part de millorar la precisió de cercadors com el de Google, també es poden bloquejar imatges inadequades. Però, aleshores, hi ha persones que les han de revisar i que es passen el dia veient fotos o vídeos que poden contenir una violència extrema, agressions sexuals a infants o actes de zoofília.

Això pot tenir conseqüències nefastes per als treballadors. Al setembre de 2023, la Seguretat Social va donar la incapacitat permanent absoluta a un home que va haver d'agafar la baixa després de poc més de dos anys treballant per a una empresa de Barcelona que feia aquestes revisions per a Facebook. Era un dels anomenats moderadors de continguts. A la policia, també solen rotar els agents que revisen continguts, per no cremar-los mentalment.

El problema potser no és que hi hagi feines molt dures, sinó que els treballadors no tinguin unes condicions mínimes de seguretat i no rebin una compensació adequada. Els anomenats *gig workers* –que fan feines concretes, sovint des dels seus domicilis, i

cobren «a tant la peça» – estan molt estesos entre les tecnològiques, que han d'analitzar, descriure i etiquetar milions de dades.

L'Institut d'Internet de la Universitat d'Oxford va elaborar l'informe *Fairwork UK Ratings 2021* en què analitzava onze de les plataformes de treball digital més populars del Regne Unit. Els resultats indiquen que només dues demostraven que pagaven el salari mínim després dels costos; sis de les onze donaven prou protecció als treballadors per a la seva tasca; cinc aportaven proves de contractes en què constaven, de forma clara i accessible, les condicions del servei; quatre tenien un procés formal perquè els treballadors poguessin recórrer contra les decisions, i només una s'havia compromès a implementar un procés perquè els treballadors compartissin els seus punts de vista amb els directius.

El mateix institut va elaborar l'informe *Fairwork Cloudwork Ratings 2022*, en què avaluava 15 plataformes com Fiverr, Upwork i Amazon Mechanical Turk. Varen entrevistar 752 treballadors de 94 països i van parlar també amb els seus gestors. L'informe fou resumit i comentat pel setmanari nord-americà *Time*.

El 30% del treballadors manifestaven que feien tasques no remunerades. Tenint en compte això, el salari era de 2,15 dòlars per hora. A més, els analistes varen donar a cada plataforma una puntuació de 0 a 10 punts, atorgant-ne un màxim de dos per cadascun d'aquests principis: remuneració, condicions de treball, contractes, gestió i possibilitats de representació.

Quatre plataformes (Amazon Mechanical Turk, Workana, Microworkers i Freelancer) van obtenir 0 punts. Cap de les 15 va superar els 5 punts. L'informe indica que obtenir 10 punts sobre 10 no és treure matrícula d'honor, sinó que simplement revela que l'empresa compleix amb els mínims.

Al gener de 2023, *Time* va analitzar diverses feines fetes pel tan popular ChatGPT. D'entrada, explicava que OpenAI estava en converses amb inversors per augmentar els seus fons a 29.000 milions de dòlars, cosa que la convertia en una de les empreses d'IA més valuoses del món.

ChatGPT ha de recórrer a feina feta directament per persones per evitar que les informacions i dades amb què s'entrena el sistema continguin biaixos violents, sexistes o racistes. Per això, van considerar que, etiquetant aquests continguts, el sistema aprendria a detectar-los.

Però, malgrat el valor de l'empresa creadora de ChatGPT, els treballadors encarregats de fer aquesta feina eren pagats per una empresa de Kenya i cobraven entre 1,32 i 2 dòlars per hora, segons havia descobert el setmanari. A Kenya no hi ha establert un salari mínim universal, però el recepcionista d'un hotel de Nairobi cobra almenys 1,52 dòlars per hora. I, en el cas dels revisors de continguts, la feina no només és dura, sinó també difícil, perquè en molts casos costa decidir si un contingut és violent o sexista, o no.

Aquesta empresa es diu Sama, té la seu a San Francisco i contracta treballadors a Kenya, Uganda i

l'Índia per etiquetar continguts per a Google, Meta i Microsoft. S'autoqualifica d'empresa ètica, que ha ajudat a sortir de la pobresa més de 50.000 persones. És cert que, en una economia tan precària, aquest salari tan baix era una manera de donar suport a moltes famílies, però també ho és que les feines s'han de fer amb unes condicions mínimes. Segons la revista, els treballadors deien que havien de revisar entre 150 i 250 fragments de text en jornades de nou hores.

Com en el cas de les innovacions tecnològiques en el treball al llarg de la història, el tema no és només si es perden llocs de treball o se'n creen de nous, o si hi ha feines molt dures, sinó que les condicions siguin bones, les compensacions siguin justes i el creixement econòmic que comporten es reparteixi més equitativament.

Tornant a les feines que existeixen des de fa molt de temps, potser continuaran essent imprescindibles, perquè l'aportació humana no es podrà suplir o superar. O potser es reciclaran i tindran unes funcions diferents, amb uns mitjans diferents. Fins i tot pot ser que algunes persones passin a ser creadores o desenvolupadores d'aquests sistemes. Però també pot ser que l'atur augmenti en aquests sectors i que els ingressos minvin. I penso que, tot i que hi pugui haver lectures optimistes, cal actuar perquè l'impacte social i econòmic negatiu sigui el mínim. La IA és una tecnologia disruptiva i probablement les anàlisis tranquil·litzadores que recorden que «al final sempre ha anat bé» potser no són útils en aquest cas. A

més, al llarg de la història, certament a alguns els ha anat molt bé, però a molts d'altres no tant i a alguns, molt malament.

## Diagnòstics, fàrmacs i bessons digitals, la cara positiva de la IA

Seria injust parlar dels problemes ètics i socials que pot generar la IA i no parlar dels seus grans beneficis. I seria absurd pretendre ser exhaustius en aquest camp. Per tant, em limitaré a comentar aspectes relacionats amb la salut i les ciències biomèdiques, ressaltant sobretot iniciatives de persones vinculades a la UPC, i dedicaré la part final als bessons digitals, que la UPC també investiga i desenvolupa.

Comencem, però, amb una de les fites més impressionants assolida per un sistema d'IA. El novembre de 2020 estarà inscrit amb lletres destacades en el camp de les ciències biomèdiques. A final d'aquell mes, el programa *AlphaFold*, creat per DeepMind, filial d'Alphabet, va aconseguir un resultat impressionant en predir el plegament proteic basant-se en el codi genètic que n'ordena la síntesi.

A partir del codi genètic, es pot saber la seqüència d'aminoàcids de la proteïna, però la seva activitat dependrà de com es plegui en l'espai. Entrenat amb aprenentatge profund, el sistema va obtenir una puntuació superior a qualsevol altre equip a la 14th Critical Assessment of Protein Structure Prediction (CASP14). Els resultats d'*AlphaFold* es van publicar a *Nature* el 15 de juliol de 2021, juntament amb un programari de codi obert i una base de dades de proteomes. En més d'un terç dels casos, el sistema assoleix una fiabilitat de més del 90% analitzant l'estructura de pràcticament totes les proteïnes conegudes del cos humà.

Algún expert ha dit que mai no havia pensat que podria veure en vida un sistema com aquest. Conèixer el plegament d'una proteïna és bàsic per esbrinar com actua, com es pot accedir al seu centre actiu, com hi poden interactuar els fàrmacs, com són els plegaments incorrectes que poden provocar diverses malalties...

Això pot inaugurar una nova era, però l'aplicació de la IA a la salut, tant en recerca com en diagnòstic, i fins i tot en gestió, és molt àmplia. Ja hi ha moltes aplicacions actives o en grau avançat de desenvolupament.

En un article publicat a *Science* el 15 de setembre de 2023, Eric J. Topol, cardiòleg i investigador, fundador i director del Scripps Research Translational Institute de Califòrnia, afirmava que, si bé les màquines no tenen ulls, sembla que en tinguin si es té en compte el progrés dels models d'aprenentatge profund per interpretar imatges de raigs X, tomografies computeritzades

des o ressonàncies magnètiques, entre d'altres. Això permet als metges observar coses que probablement els passarien per alt, perquè l'ull humà no les pot distingir.

Posa com a exemple un escàner retinal que aporta informació que un humà no pot veure, però una màquina sí. Concretament, proporciona informació tan diversa com la pressió sanguínia, el control de la glucosa, el risc de Parkinson o d'Alzheimer, malalties de ronyó o de fetge, la possibilitat d'infarts o d'ictus... I un electrocardiograma dona informació sobre l'edat, el sexe, la presència d'anèmia, el risc de diabetis o d'arrítmies, problemes de ronyó o de tiroides... I aplicar l'aprenentatge profund a imatges del teixit tumoral pot ajudar a conèixer on s'ha originat, les mutacions que ha experimentat i la seva prognosi. I considera que encara hi ha molt de terreny per endavant sobre el paper dels sistemes d'IA en la interpretació d'imatges.

La IA permet fins i tot determinar a quin tipus pertany un tumor cerebral, una eina clau per establir fins on cal eliminar teixit d'una manera tan ràpida que el resultat arriba durant la intervenció quirúrgica mateix. Ho explicava a *Nature* l'11 d'octubre de 2023 un equip del Centre de Medicina Molecular d'Utrecht (Països Baixos), on investiga el català Marc Pagès-Gallego, graduat per la Universitat Rovira i Virgili, un dels primers autors de l'article.

Un altre camp en què la IA pot jugar un paper cabdal és la recerca de nous fàrmacs. Per exemple, transformar una proteïna potencialment útil en un fàrmac requeria molts anys de feina, modificacions i assaigs.

Passen molts anys abans que un nou fàrmac potencial arribi als assaigs clínics. Pel camí, la gran majoria han quedat descartats. Alguns estudis assenyalen que nou de cada deu fàrmacs que arriben a aquests assaigs acaben no essent aprovats per a ser comercialitzats i aplicats. I que costa uns 2.500 milions de dòlars fer arribar un fàrmac al mercat.

Els sistemes d'IA permeten ara seleccionar proteïnes a partir de la seva seqüència d'aminoàcids o del seu plegament i relacionar-les amb la seva eficàcia, seguretat o facilitat de síntesi. Tot això redueix el temps per trobar un nou fàrmac, que tot i així continua essent un procés laboriós en què només un percentatge molt petit dels candidats inicials arriba al mercat.

## **IA i salut, un camp en expansió**

Com a mostra de l'interès que desperta aquest camp, a PubMed es poden trobar més de 200.000 articles sobre IA i salut. Un dels articles es va publicar a la revista *Diagnostics* el març de 2022 i descriu el sistema DigiPatICS. Es titula «Digital Pathology Transformation of the Catalan Health Institute Network of 8 Hospitals».<sup>5</sup> Entre els autors hi ha Ferran Marquès i Josep R. Casas, del Grup de Processament d'Imatges de la UPC. Tal com explica l'article, el projecte és una iniciativa de l'Institut Català de la Salut (ICS) per desplegar patologia digital de manera inte-

---

5 Vegeu: <https://www.mdpi.com/2075-4418/12/4/852>

grada i holística en una xarxa de vuit hospitals, amb la participació de més de 168 patòlegs i l'anàlisi de més d'un milió d'imatges cada any. El propòsit és incrementar la seguretat del pacient i la qualitat de l'atenció, millorant el diagnòstic i l'eficiència del procés.

El projecte inclou el desenvolupament d'eines d'IA amb aprenentatge automàtic i aprenentatge profund, que processen imatges de teixit digitalitzades amb molt alta resolució: són les anomenades *whole slide images* (WSIs), que solen tenir típicament  $200.000 \times 90.000$  píxels, i que han estat obtingudes per escaneig i digitalització de les mostres de biòpsia del teixit. Entre altres objectius, es pretén reconèixer-hi patrons, seleccionar automàticament àrees del tumor i quantificar-les amb l'obtenció de biomarcadors que ajudin als patòlegs a fer la seva diagnosi.

Segons l'ICS, el sistema abasta des de la gestió de la sol·licitud de la mostra d'anatomia patològica fins al lliurament del resultat clínic. I destaca com a punt més rellevant l'escaneig a molt alta resolució de les mostres, que ha de ser equivalent a la dels microscopis que s'utilitzen actualment.

Hi ha moltes altres eines desenvolupades a la UPC o amb participació d'investigadors de la universitat. En un text publicat a la Sala de Premsa de la UPC l'octubre de 2022, s'explica la recerca que fan conjuntament el Grup de Processament d'Imatge i Vídeo (GPI)<sup>6</sup> de la UPC i el Barcelonaβeta Brain Research Center

---

6 Vegeu: <https://futur.upc.edu/GPI>

(BBRC)<sup>7</sup> de la Fundació Pasqual Maragall. L'objectiu és detectar signes d'Alzheimer quan ni tan sols han aparegut o a penes se n'han descobert els primers símptomes. La investigadora que lidera aquest projecte des de la UPC és la professora Verónica Vilaplana, del GPI. Juntament amb el seu equip, ha desenvolupat una eina que combina informació obtinguda automàticament mitjançant l'anàlisi de les imatges de ressonància magnètica amb altres dades demogràfiques i genètiques. Amb els algorismes adequats, es poden identificar les persones amb risc de patir Alzheimer i dissenyar estratègies d'intervenció primerenques per a una malaltia que ara com ara no té cura.

Un altre projecte és la prevenció de suïcidis als trens mitjançant algorismes que analitzen imatges de vídeo i hi detecten patrons. Es tracta del projecte *Anomaly* - Fase 0, en el qual participen investigadors del grup d'Enginyeria del Coneixement i Aprenentatge Automàtic (KEMLG),<sup>8</sup> de l'IDEAI-UPC i del Barcelona Supercomputing Center (BSC-CNS).

El Llibre blanc sobre la Intel·ligència Artificial aplicada a la salut<sup>9</sup> fou publicat l'any 2022 pel CIDAI (Center for Innovation in Data tech and Artificial Intelligence). Aquest centre, coordinat per Eurecat, promou la transferència de coneixement i la realització de projectes conjunts entre entitats generadores

---

7 Vegeu: <https://www.barcelonabeta.org/ca>

8 Vegeu: <https://kemlg.upc.edu/ca>

9 Vegeu: <https://storage.cdn.eurecat.org/CIDAI/WhitePapers/WP-IA-Salut.pdf>

de coneixement, empreses proveïdores de tecnologia i serveis, i empreses i institucions usuàries.

Segons aquest *Llibre blanc*, la IA promet grans avenços en el sector de la salut amb la consolidació de la medicina 4P: preventiva, predictiva, personalitzada i participativa.

Afegeix que Catalunya disposa de grans capacitats i potencialitats per liderar el desenvolupament i l'adopció de la IA al sud d'Europa i per posicionar-se com un pol tecnològic de referència. Però afegeix que les iniciatives explorades per empreses i els organismes tenen un impacte encara baix i que «el nivell d'adopció de tecnologies innovadores en IA per part del sector privat, especialment per part de les petites i mitjanes empreses (pimes) del sector, continua essent menor que en altres països europeus». Així doncs, les oportunitats principals de la IA en el sector de la salut són: el suport al diagnòstic, la planificació i el guiatge d'intervencions, el tractament personalitzat, la monitorització, els sistemes de suport i logística, i la gestió hospitalària.

Al *Llibre blanc* hi ha molts exemples de casos rellevants, dos dels quals són promoguts o estan participats per la UPC. Un és l'IPAMM (Identificació de Perfils d'Agitació en Malalts Mentals), que ja es va dur a terme. El projecte explica que les persones amb afectacions de salut mental poden patir a vegades episodis d'agitació amb un risc altíssim per a la seguretat dels qui els envolten i també per a la seva pròpia integritat. I afegeix que mitigar aquests episodis requereix molts

recursos personals i farmacològics, amb un cost difícil de conèixer, donada la multiplicitat d'accions i d'actors que intervenen en l'episodi de mitigació.

Conèixer aquesta dada és clau per valorar l'ús de determinats fàrmacs de nova generació que redueixen els episodis d'agitació i la seva intensitat quan es donen, però que tenen uns costos molt elevats per a qualsevol sistema de salut pública. Per això, cal valorar-ne el cost amb el dels recursos humans i els ingressos hospitalaris associats a aquests episodis. Així, es pot avaluar si val la pena o no, des de la sanitat pública, incloure un determinat fàrmac.

El projecte consistia a analitzar els informes d'infermeria dels pacients, recollits en llenguatge natural a la història clínica, i la informació addicional sobre el pacient. L'anàlisi de llenguatge natural permet extreure'n símptomes i actuacions, i ajudarà a identificar els tipus d'episodis d'agitació. El Parc Sanitari de Sant Joan de Deu<sup>10</sup> i el KEMLg d'IDEAI-UPC en són els impulsors.

Un altre projecte descrit al *Llibre blanc* és el d'implementació de medicina personalitzada en un melanoma maligne mitjançant IA. L'objectiu és millorar el procés assistencial dels pacients amb melanoma. Si bé per gestionar aquests pacients es fa una estratificació basada en l'anàlisi de la informació clínica i patològica del tumor, hi ha informació rellevant que en queda oclosa. I això limita la precisió del mètode.

---

<sup>10</sup> Vegeu: <https://www.pssjd.org/>

Per això, el Consorci Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS)<sup>11</sup> i la Unitat de Melanoma del Departament de Dermatologia de l'Hospital Clínic de Barcelona estan desenvolupant aquest sistema. El projecte ha comptat amb el finançament de la Fundació La Marató de TV3. Per la UPC, hi participa l'equip de Verónica Vilaplana, del GPI-IDEAI.

L'objectiu és desenvolupar nou algoritmes basats en la IA que permetin millorar l'atenció mèdica, és a dir, el diagnòstic, el seguiment i el tractament dels pacients amb melanoma. Per a això, es fa servir una base de dades anonimitzada amb informació de més de 14.000 pacients amb melanoma de 23 hospitals de Catalunya (Xarxa de Centres de Melanoma), juntament amb dades clíniques, ambientals, socioeconòmiques, fenotípiques, bioquímiques, farmacològiques, genètiques i d'imatge mèdica. I es desenvoluparà un model d'IA per predir el risc de recaiguda, metàstasi o mort del pacient.

Així, es crearan grups d'estratificació de pacients estadísticament coherents, que poden ajudar a prendre decisions de tractament i administratives, i reduiran la quantitat d'intervencions quirúrgiques innecessàries i de teràpies mèdiques quan no siguin útils per a determinats pacients.

---

11 Vegeu: <https://www.clinicbarcelona.org/ca/idibaps>

## Bessons digitals

Una petita maqueta amb cables i pesos recorda la manera usual de calcular resistències d'estructures. El doctor Rolando Chacón, professor agregat del Departament d'Enginyeria Civil i Ambiental de la UPC, em rep al Laboratori de Models Digitals en Estructures i Construcció (LMDEC).

«Els estudis necessaris per al projecte d'estructures i de construcció –m'explica– s'han fet amb models numèrics de predicció i, en alguns casos, amb models físics reduïts, com models amb cables i pesos.»

Però, als anys vuitanta, va aparèixer el disseny assistit per ordinador. I potser això ens fa valorar encara més els pioners. Recordo que fa prop de trenta anys, fent un reportatge sobre la Sagrada Família, els arquitectes em varen dir que algunes estructures dissenyades per Antoni Gaudí eren tan innovadores i complexes que hauria estat molt difícil validar-les sense les noves eines informàtiques.

Per Chacón, actualment la bretxa físicodigital en el món de les infraestructures s'està tancant. Un concepte que ho il·lustra és el «bessó digital». Per elaborar bessons digitals, cal capturar dades del món físic i processar-les. «Els reptes principals –explica– són tres: capturar aquestes dades, crear els programes per processar-les i elaborar-ne el model digital, i establir la relació més fidel entre el món físic i el digital que permeti prendre decisions d'enginyeria relacionades amb les seves prestacions.»

La idea de bessó digital va néixer a principi d'aquest segle, si bé l'informàtic David Gelernter ja havia proposat la idea l'any 1991 al seu llibre *Mirror Worlds*. El 2002, Michael Grieves, enginyer informàtic de la Universitat de Michigan, va comentar a John Vickers, aleshores director de Tecnologia de la NASA, que es podrien crear representacions digitals de sistemes físics. Sorgia la idea d'una rèplica virtual, amb vida pròpia, d'un producte: el *digital twin*. El 2010, Vickers va fer un informe per a la NASA en què estimava que significarien un mercat de prop de 50.000 milions de dòlars l'any 2026.

Era anar un pas més enllà de la simulació, perquè s'hi afegia la capacitat d'incorporar-hi dades en temps real, captades per sensors. La informació que s'obté es processa amb tot un seguit de sistemes –recentment, amb sistemes d'IA– i així se n'obté una representació que mostra la resposta real a tots els factors estudiats.

El bessó digital pot reproduir un producte concret, com la turbina d'un avió, però també conjunts més complexos. En tenim alguns exemples en el projecte europeu ASHVIN<sup>12</sup>, en el qual participa la UPC, que té com a objectiu obrir vies de dades entre els dos àmbits, el físic i el digital, i comprovar la metodologia per aplicar-la a millorar la productivitat en la construcció i fer-la més segura. Permet combinar els bessons digitals amb l'obtenció de dades contínuament per controlar aspectes com la seguretat de les infraestructures.

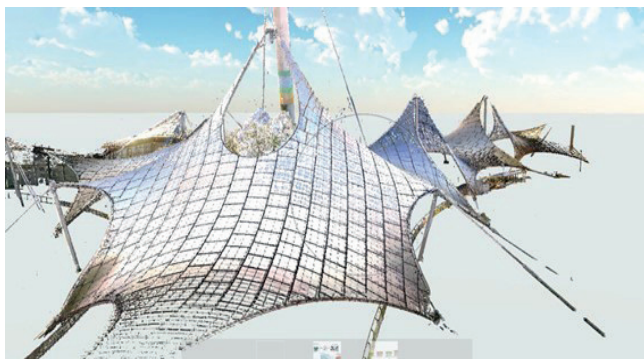
---

12 Vegeu: <https://www.ashvin.eu/>

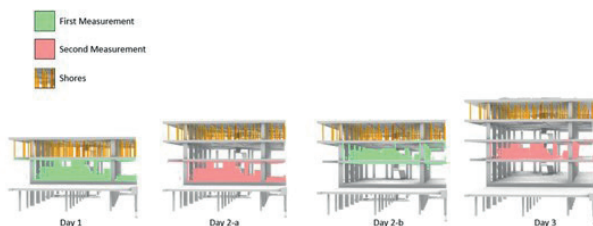
El projecte compta amb socis multidisciplinaris: desenvolupadors, empreses del sector, universitats i institucions públiques de normalització. I no oblida els aspectes ètics i socials, que estudien investigadors de la Universitat Erasme de Rotterdam (Països Baixos): per exemple, de quin són les dades, quin ús se'n pot fer o a disposició de qui es poden posar.

El projecte es basa en una sèrie de demostradors reals de la metodologia, entre els quals hi ha ponts, edificis, aeroports i ports que es troben al llarg del continent europeu. Alguns demostradors mostren com intentar incloure el concepte en fases de projecte –una sèrie de passeres de vianants a Alemanya. D'altres serveixen per desenvolupar bessons digitals en fase de construcció –edificis sensoritzats que proporcionen diàriament dades del procés–, com en el cas del sostre de cables de l'Estadi Olímpic de Munic, on s'han obtingut les tres coordenades espacials de milions de punts amb rastreig làser. Aquest darrer exemple mostra una manera de capturar dades amb tecnologia remota. També s'han pres dades durant la prova de càrrega d'alguns viaductes de l'AVE d'Extremadura.

Es poden elaborar bessons digitals fins i tot de ciutats, aprofitant el treball en paral·lel. Si diversos equips elaboren el bessó digital de diferents edificis i els resultats es combinen, és factible tenir el bessó digital d'un barri o de tota una ciutat.



*Figura 4. Nívol de punts de l'Estadi Olímpic de Munic*



*Figura 5. Digitalització de la construcció d'un edifici de formigó armat a Barcelona*

El sistema es podria aplicar a un problema urgent i creixent: adaptar les ciutats a l'escalfament global i al canvi climàtic que comporta. Ho suggerien investigadors de la Universitat de Hong Kong en una carta

a *Nature* l'octubre de 2023. Deien que es podien simular diversos perfils i estudiar com afecten la població, variant la distribució de la població i els fluxos de mobilitat. Així es podrien planificar i avaluar diverses solucions abans d'assajar-les.

Actualment, existeixen bessons digitals d'estructures tan complexes com un aeroport. N'és un exemple el de San Francisco, en què els usuaris tenen accés, a través dels seus mòbils, a un mapa complet de la instal·lació per saber on anar. Però el sistema reproduceix, a banda de tots els elements visibles –15.000 portes, 2.000 arbres–, també el subsol, amb més de mig quilòmetre de cables elèctrics i canonades. I fins i tot conté informació actualitzada en temps real de la posició dels equips i dels vehicles de manteniment. Una informació que no és a l'abast del públic, però sí dels gestors de la infraestructura.

Un altre aeroport que compta amb el seu bessó digital és el de Vancouver (Canadà). Proporciona informació en temps real tant per fer avisos als viatgers com perquè els seus gestors i els equips de manteniment i de seguretat puguin prendre les decisions amb el màxim nombre de dades.

Un altre objectiu és reduir les emissions. El bessó digital de l'aeroport compta amb un sistema que mesura i ressegueix el recorregut de les emissions dels avions, tant quan aterren com quan s'enlairen o quan es mouen per les pistes. L'objectiu és promoure la descarbonització.

## Pacients digitals

Els bessons digitals també tenen aplicacions en medicina, si bé això encara presenta molts reptes, tant tècnics com ètics. Les possibilitats són ingents, perquè es poden crear bessons digitals d'un òrgan concret, com l'estómac, de tot un sistema, com el digestiu, o simplement d'alguns teixits o tipus cel·lulars. I, a l'altre extrem, es poden fer bessons digitals d'organismes sencers i, fins i tot, de grups de població.

L'any 2016, Bill Ruh, aleshores CEO de GE Digital, va predir que algun dia, en el moment de néixer, tindríem un bessó digital que prendria dades i prediria l'evolució de la nostra salut i malalties com el càncer o d'altres.

Però no anem tan de pressa ni tan lluny. De moment, els bessons digitals presenten moltes possibilitats, tant en la recerca com en les aplicacions clíniques, però també molts problemes. Les proves *in silico* podrien suplir les proves preclíniques *in vivo*. Imaginem un bessó digital que podem fer servir per assajar nous fàrmacs o per estudiar les alteracions que porten a determinades malalties. Si es fan servir les dades individuals d'una persona, les proves permetran millorar-ne el diagnòstic i comprovar-ne les teràpies. Seria un pas de gegant en la medicina personalitzada. Imaginem que no s'administra una medicació o no es practica una intervenció a un pacient, sinó al seu bessó digital. I que només es tira endavant de forma real quan

se n'han comprovat la viabilitat i els efectes positius, almenys amb un alt percentatge de fiabilitat.

Seria un sistema molt interessant, sobretot en malalties complexes i en determinades poblacions, com la gent gran. Però, de moment, presenta moltes dificultats. En primer lloc, de captació de dades. En segon lloc, de sistemes que puguin replicar el comportament de l'organisme o d'una part. Per això, la seva aplicació és molt limitada ara com ara, però amb el desenvolupament de la IA i de diverses tècniques pot ser una realitat més usual en el futur.

I, en alguns casos, augmenten no tan sols els problemes tècnics, sinó també els ètics. El projecte europeu Neurotwin<sup>13</sup>, coordinat per l'empresa catalana Neuroelectrics<sup>14</sup>, té l'objectiu de crear còpies virtuals de cervells. L'objectiu es desenvolupar models de cervells per simular tractaments personalitzats, com ara l'estimulació electromagnètica, i estudiar la seva aplicació a malalties com l'Alzheimer.

També fa assaigs clínics relacionats amb l'epilèpsia, amb l'objectiu de reduir o eliminar les crisis aplicant corrents elèctrics a la zona on el cervell està sobreexcitat. Tracta pacients que no responen a la medicació convencional. Quan els neuròlegs identifiquen la zona causant de l'epilèpsia en una persona, es crea un protocol personalitzat per optimitzar-ne el tractament.

Si el mètode mostra resultats positius, les recerques futures es poden orientar a la depressió, l'esclerosi

---

13 Vegeu: <https://www.neurotwin.eu/>

14 Vegeu: <https://www.neuroelectrics.com/>

múltiple i, fins i tot, a assaigs clínics sobre els possibles efectes terapèutics de substàncies psicodèliques administrades en dosis i en condicions controlades.

Les possibilitats són moltes i poden donar esperança a pacients que ara no tenen remei per als seus mals. Però, a part dels problemes tècnics, també se'n plantegen d'ètics i socials. Així, els bessons digitals aplicats a la salut requereixen moltes dades personals sensibles. Cal garantir que les dades només s'utilitzaran per a uns fins determinats i que no es passaran a persones o grups amb intencions fraudulentament.

Un altre problema, lligat en part amb l'anterior, és la seguretat. Cal reduir les possibilitats que hi hagi ciberatacs que pretenguin obtenir dades personals o interferir en els processos que es vagin desenvolupant.

A més, hi ha un tercer aspecte, que ja s'ha esmentat abans: De quin són les dades? Han de ser compensades les persones que en proporcionen per posar a punt els sistemes? Simplement se'ls ha d'informar? Què passa amb la informació, un cop s'ha assolit el propòsit inicial? Es poden posar a disposició d'altres investigadors?

No hi ha avenç científic o tecnològic que no comporti un debat ètic i social, i afrontar-lo abans que tot estigui massa desenvolupat és essencial per poder preparar-ne els marcs legals i les condicions adients.

I, com més avançada i potent és la tecnologia, més debats hi ha. Així doncs, es poden crear bessons digitals de persones que ja no hi són? En alguns casos, ja s'ha fet. Però el tema és molt delicat i complex, i l'abordarem al final del capítol següent.

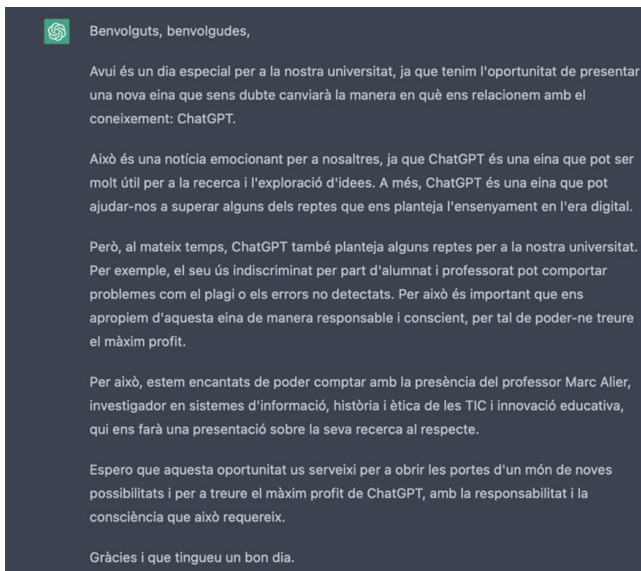


## Converses amb ChatGPT i amb chatbots

«Espero que aquesta oportunitat us serveixi per a obrir les portes d'un món de noves possibilitats i per a treure el màxim profit de ChatGPT, amb la responsabilitat i la consciència que això requereix. Gràcies i que tingueu un bon dia.» Aquestes van ser les paraules finals del breu discurs que Daniel Crespo, rector de la UPC, va pronunciar el 2022 per presentar l'exposició que Marc Alier va fer de la seva recerca amb ChatGPT.

Ningú no va notar res, però, potser per ser màximament coherents o per mostrar, d'entrada, una de les aplicacions del robot, aquest discurs l'havia redactat ChatGPT seguint les indicacions del doctor Alier. Segons m'explica Alier, «era un experiment de dos minuts que es va fer al desembre de 2022 amb una versió molt inicial de ChatGPT. Simplement, vaig introduir a ChatGPT dos exemples de discursos previs per do-

nar-li context, i llavors li vaig demanar un discurs que introduís el tema de la IA generativa, el repte que suposa per a la docència a la UPC, i que em presentés a mi com a ponent.» El resultat es troba a la figura 6.



*Figura 6. Discurs elaborat amb ChatGPT per al rector de la UPC*

El doctor Crespo va comentar que el text li havia semblat correcte i tothom va pensar que les paraules encaixaven perfectament amb el tipus de presentacions que feia el rector. Però, més endavant, quan va presentar l'acte d'homenatge a Eugenio Oñate, profes-

sor emèrit de la UPC, que el rector coneixia molt bé, va deixar de banda ChatGPT i va fer un discurs molt més personal, a partir de records de la seva coneixença. El doctor Crespo creu que, en casos com aquest, ChatGPT no és útil, perquè no té aquesta experiència ni aquests sentiments personals.

Marc Alier en discrepa i ho justifica amb les capacitats de la nova versió: «En la meua experiència, la qualitat dels resultats depèn de la qualitat dels *prompts* (indicacions o apunts) que dones al model. Fent servir la versió avançada de ChatGPT (GPT4) en pots obtenir uns resultats que impacten, si jugues amb el context (aportant informació rellevant, continguts, estil) i demanes les coses de forma hàbil. Ara mateix, es poden elaborar discursos, cartes o articles tan complexos o personals i emotius com els que podria fer una persona humana. Tan sols cal habilitar a l'hora d'utilitzar l'eina.»

En una de les entrades al seu blog «Calaix Desastre»,<sup>15</sup> Marc Alier explica què és ChatGPT, cosa que va molt bé per saber també què no és. Així, diu que és «un tipus d'IA generativa basada en un model de *machine learning*. Podem considerar que ChatGPT és una IA dèbil», és a dir, especialitzada en una àrea.

En el moment d'escriure això, l'eina està disponible de forma gratuïta a <https://chat.openai.com>. És important llegir amb atenció les condicions d'ús. Així, sabrem que el contingut serà utilitzat per desen-

---

15 Vegeu: <https://wasabi.essi.upc.edu/ludo/posts/que-es-chatgtp/>

volupar i millorar els serveis. Si no ho volem, podem fer-ho constar, però aleshores se'ns adverteix: «En alguns casos, això pot limitar la capacitat dels nostres serveis per abordar millor el vostre cas d'ús específic.» És a dir, com passa sovint, si no volem acceptar determinades condicions, se'ns avisa que potser hi sortirem perdent.

ChatGPT funciona a partir de càlculs de probabilitats. Un cop se li ha plantejat un tema i se li ha delimitat el context, el sistema construeix frases a partir dels milions de dades amb què s'ha entrenat, cercant les paraules més probables que seguiran a altres. És a dir, no és que sàpiga gaire de res –en realitat, no sap res de res–, sinó que ha estat entrenat per trobar les frases més probables.

En una conferència pronunciada al Parlament de Catalunya el 9 d'octubre de 2023, amb motiu de la reunió anual de l'European Parliamentary Technology Assessment (EPTA), que llavors presidia Catalunya, el psicòleg cognitiu nord-americà Gary Marcus, que ha criticat molt alguns dels efectes no volguts de la IA i els seus possibles biaixos, posava un exemple dels problemes que ChatGPT té a l'hora de donar respostes lògiques. Ho vaig voler comprovar pel meu compte i el resultat és aquest diàleg amb ChatGPT:

XD

Qui és la mare de Tom Cruise?

ChatGPT

La mare de Tom Cruise es diu Mary Lee Pfeiffer. Va ser una activista i educadora. Tom Cruise és un conegut actor de Hollywood que ha aparegut en nombroses pel·lícules, com “Top Gun,” “Mission: Impossible,” “Jerry Maguire” i moltes altres.

XD

Com es diu el fill de Mary Lee Pfeiffer?

ChatGPT

Lamentablement, no dispo d'informació sobre el fill de Mary Lee Pfeiffer, ja que a la meua base de dades només hi ha informació sobre persones públiques i famoses fins a la meua data de tall del coneixement, el setembre de 2021. Si el fill de Mary Lee Pfeiffer no és una figura coneguda a nivell internacional o no ha estat àmpliament discutit a mitjans de comunicació importants, podria ser difícil trobar informació detallada sobre ell.

Abans m'has dit que era la mare de Tom Cruise

### ChatGPT

Tens raó, Mary Lee Pfeiffer no és la mare de Tom Cruise. M'he equivocat en la meua resposta anterior, i et demano disculpes per la confusió. La mare de Tom Cruise es diu Mary Lee Pfeiffer. El fill de Mary Lee Pfeiffer és Tom Cruise, el conegut actor de Hollywood. Gràcies per corregir-me.

És a dir: sap qui és la mare de Tom Cruise, però és incapaç de deduir, després, que el fill d'aquesta senyora és Tom Cruise. I, quan li faig veure que abans m'ha donat la informació, demana disculpes, però encara s'embolica més i ofereix un paràgraf ben contradictori, en el qual Mary Lee Pfeiffer resulta ser i no ser la mare de Tom Cruise.

Tot i així, a vegades respon problemes senzills de lògica. Per exemple, li he plantejat el següent: el pare de l'Antoni té tres fills. Un es diu Joan i un altre, Pere. Com es diu el tercer? I de seguida ha respost: «El tercer fill es diu Antoni, com es menciona al principi de la frase: El pare de l'Antoni té tres fills.» L'he felicitat perquè molta gent no ho veu a la primera i m'ho ha agraït: «Gràcies! Sí, aquest tipus de preguntes de lògica poden ser enganyoses a primera vista, però són una

manera divertida d'exercitar el pensament i la lògica. Si tens més preguntes o necessites ajuda amb alguna altra cosa, estic aquí per ajudar!» El que no es pot dir és que no sigui amable i educat.

Marc Alier investiga sobre ChatGPT, especialment sobre el seu ús en educació. Explica que el fa servir, tot i que a vegades li dona més feina, perquè fa proves de coses que després reverteixen en les seves recerques. Però, normalment, millora la productivitat fins a 4 vegades i la qualitat –«menys errors, idees més explorades en diàleg amb l'eina».

A vegades, hi manté una conversa, però explica com s'ha de treballar amb el sistema: «L'has de tractar com una criatura a qui plantejes un problema fàcil –si hi ha tantes boles de colors i tants nens, quantes en toquen a cadascun?– i et respon amb una xifra i prou. Li dius que vols que raoni com ho ha resolt, que t'ho expliqui, que et doni el resultat amb una frase i no una quantitat... En aquest cas, la idea és proposar-li un format de resposta, és a dir, a més de què vols dir-li com ho vols.»

Per això, l'eina pot ser útil si saps explicar molt bé què vols. A més, pots anar corregint, matisant o concretant el que pretens: «Si al ChatGPT li dius que redacti un text i després li demanes que no hi posi determinades paraules molt acadèmiques, no només les deixarà d'utilitzar sinó que l'estil del text també serà menys acadèmic.»

Li pots donar textos i orientar sobre l'estil. I per això Alier creu que seria capaç de fer un discurs més personal per al rector: «Si s'hi introdueix la informa-

ció, com aquesta sobre la relació amb el doctor Oñate, ChatGPT també pot redactar-lo.»

Però cal anar amb compte, perquè pot tenir al·lucinacions, com altres sistemes d'IA. I això comporta molts riscos. L'any 2023, un jutge dels Estats Units va multar dos advocats amb cinc mil dòlars perquè havien exposat als tribunals citacions falses generades per ChatGPT. Actuaven en una reclamació per lesions contra una companyia aèria. I es van trobar que ChatGPT s'havia inventat alguns casos i els citava com a precedents.

Per al jutge de Manhattan P. Kevin Castel, que els advocats facin servir la IA com a eina auxiliar no és inadequat, però s'han d'assegurar que la informació és acurada.

La doctora Karina Gibert explica que l'ús de ChatGPT ha comportat problemes per a algunes empreses: «Alguns empleats han dirigit consultes a ChatGPT sobre temes crítics per al negoci, que s'han filtrat com a part de les respostes que ChatGPT després ha fet servir per respondre a altres usuaris de fora de l'empresa.»

Així, com tots els sistemes d'IA que aprenen de les dades que processen, el sistema pot contenir errors. A més, si s'hi introdueix informació confidencial i algú altre demana a ChatGPT alguna cosa relacionada amb el mateix tema, altres persones podrien obtenir aquesta informació, la qual no els hauria de ser revelada. Recordem que, si no diem el contrari, tot allò que fem amb ChatGPT pot ser utilitzat per entrenar i millorar el sistema.

Per això, tant Gibert com Alier recalquen que s'ha de revisar el material que proporciona aquesta eina. Si hi ha informació falsa, la responsabilitat serà de qui no la utilitza bé o se'n refia i no repassa ni en comprova els continguts. El responsable sempre és qui accepta el resultat i el fa servir. «Tota tecnologia necessita comprensió», diu Alier.

Alier aprofita per parlar sobre creativitat i criteris estètics: «ChatGPT no té criteri estètic, però qui l'utilitza l'ha de tenir. Si no, el producte és forçat i, al final, els resultats sempre s'assemblen.» El debat es pot estendre a la IA en general: «A la música, sorgeixen nous estils que de vegades són només la fusió d'altres estils: llatí + jazz, western + samurai, *metal* + *reggae-ton*... Els sistemes actuals d'IA poden generar nous estils, que són fusions d'altres. Però, quan surt una cosa nova com el *blues*, que també té influències anteriors, però és diferent a tot el que existia, ja no pot.» Per això parlem d'IA generativa i no d'IA creativa.

Una mostra deliciosa la tenim en la iniciativa de crear roses per Sant Jordi amb la IA,<sup>16</sup> impulsada per l'InLab FIB, el laboratori d'innovació i recerca de la Facultat d'Informàtica de Barcelona, i el programa UPCArts.

Seguint amb aquestes explicacions, Marc Alier ens dona un exemple per il·lustrar que la possible manca de creativitat no és un problema específic de la IA,

---

16 Vegeu: <https://www.upc.edu/ca/sala-de-premsa/noticies/roses-uniques-per-sant-jordi-sumant-intel·ligencia-artificial-i-creativitat-humana>

sinó també dels humans: «El productor Jon Peters va demanar al director i guionista Kevin Smith que en una pel·lícula de *Superman* introduís una aranya robòtica gegant. El director s'hi va negar perquè va dir que no hi pintava res, que no hi encaixava.» Però el productor va continuar provant sort i va demanar el mateix a Barry Sonnenfeld, director de *Wild Wild West*. «I en aquesta pel·lícula sí que hi surt l'aranya, però es veu que no hi pinta res, que és una aparició forçada. El director té criteri estètic, però el productor no en tenia.» La història l'explica amb detall Rob Harvilla a la web The Ringer, en un article publicat al juny del 2019 i titulat “Giant Spiders, Giant Flops: The Enduring Awfulness of ‘Wild Wild West’”.<sup>17</sup> Si canviem la creativitat per la consistència i la fiabilitat, les aranyes gegants per precedents legals i les pel·lícules per casos judicials, veurem clarament que ChatGPT ens hi pot ajudar, però que no ens en podem refiar del tot. I que, al final, les conseqüències negatives seran per a nosaltres, encara que intentem donar la culpa a OpenAI.

Per tot això, ChatGPT planteja diversos problemes ètics: biaixos inherents als conjunts de dades d'entrenament, que poden provocar discriminació i promoure uns estereotips prejudicials; desinformació i manipulació creades per persones que utilitzen ChatGPT, en forma de notícies falses o teories de la conspiració; atacs a la privadesa; dubtes sobre la responsabilitat,

---

<sup>17</sup> Vegeu: <https://www.theringer.com/movies/2019/6/28/18761548/wild-wild-west-will-smith-20th-anniversary>

perquè és difícil establir qui és responsable de les respostes generades per ChatGPT en casos de contingut inapropiat o perjudicial; manipulació emocional de les persones mitjançant la generació de continguts dissenyats a aquest efecte; dependència tecnològica, que pot afectar la comunicació interpersonal; síntesi de veu i identitat falsa, que augmenta el risc d'engany i frau; creació de continguts massius, com articles o comentaris en línia que poden inundar l'espai cibernètic amb informació generada automàticament; desocupació per l'ús generalitzat de tecnologies com ChatGPT en sectors com l'atenció al client...

Exagero? Crec que no, perquè ho he preguntat al mateix ChatGPT i m'ha proporcionat aquesta llarga llista comentada d'autocrítica. I em sembla consistent i plausible. Però, com es poden evitar o contrarestar aquests problemes? Per no ser subjectiu, ho he tornat a demanar a ChatGPT: i diu que aquests problemes ètics destaquen la necessitat de desenvolupar i aplicar polítiques, regulacions i pràctiques que abordin aquestes qüestions i minimitzin els riscos associats a l'ús de models de llenguatge com ChatGPT.

Ell mateix ofereix possibles mesures, amb una combinació d'enfocaments tècnics, reguladors i ètics: «millorar la recopilació de dades; afegir-hi la supervisió humana, transparència i explicabilitat; determinar responsabilitats clares pel contingut generat per aquestes tecnologies; educació i consciència, col·laboració industrial, disseny ètic, recerca ètica i control de l'usuari».

Potser s'hauria de prohibir ChatGPT? Això seria una mesura extrema. Feriríem ChatGPT si li ho preguntéssim? No ho sembla, perquè ho veu com una possibilitat real. Però aporta molts matisos sobre els casos en què es podria fer això, que, segons diu, «és una qüestió complexa i depèn de diversos factors; no existeix una resposta única per a tots els casos».

Es podria considerar que, en cas de continguts que puguin causar danys greus, incitar a la violència, difondre odi o propagar desinformació perjudicial, les restriccions podrien ser necessàries. Però també cal tenir en compte el dret a la llibertat d'expressió, afegeix. Per això cal trobar un equilibri entre la protecció contra els continguts perjudicials i la preservació de la llibertat d'expressió.

En comptes de prohibir-ne totalment l'ús, proposa com a alternatives regulacions i requisits de supervisió per tal de garantir-ne un ús responsable. Això pot incloure la necessitat de revisar i de moderar el contingut generat per ChatGPT en determinats contextos. I sobretot, és clar, fomentar l'ús responsable de les tecnologies del llenguatge. En tot cas, les polítiques i regulacions han de ser flexibles i estar subjectes a revisió constant per tal de garantir que es mantenen actualitzades i que són eficaces. Les decisions sobre si prohibir o restringir l'ús de ChatGPT s'han de basar en evidències i en una anàlisi aprofundida dels riscos i els beneficis.

Podem dir que ChatGPT, tot i defensar-se, reconeix que cal afrontar els possibles problemes greus que

pot generar. El seu discurs m'ha semblat interessant i coherent. En aquest cas, diria que no ha al·lucinat.

## Chatbot *contra el bullying*

Martina Massana és molt jove. Està estudiant el segon curs de Ciència i Enginyeria de Dades a la UPC. Encara no sap ben bé a quina especialització es voldrà dedicar, perquè «és una carrera bastant versàtil i encara no he vist tot el què em pot oferir». Però, abans d'arribar a la meitat de la carrera i decidir-se, ja ha creat un *chatbot* i ha escrit un llibre contra el *bullying*: *Yo también soy diferente* (Penguin Random House, 2023).

La seva creació es diu Trustbot,<sup>18</sup> amb la qual cosa vol indicar que des del primer moment ha de crear confiança.<sup>19</sup> Cap al final del llibre, explica *qui* és Trustbot –perquè prefereix dir-ho així, i no pas *què* és Trustbot, volent indicar que és gairebé més persona que cosa.

La idea va sorgir en un treball d'institut i és una eina capaç de detectar casos de *bullying*, cosa no gens fàcil, perquè sovint passen desapercebuts: potser l'entorn no els detecta, potser la persona no s'atreveix a explicar-ho, potser se'n veuen indicis però no es considera una cosa prou important... Trustbot, com tots els *chatbots*, pot conversar com si fos una persona. I en tres minuts conclou si la persona amb qui ha parlat pa-

---

18 Vegeu: <https://trustbot.cat/>

19 Vegeu: <https://www.ccma.cat/tv3/sx3/una-inventora-contra-el-bullying/video/6143493/>

teix assetjament. Si la resposta és afirmativa, Trustbot envia la informació de manera anònima a l'equip de psicòlegs del centre.

Aquest *chatbot* encara no s'utilitza, perquè la plataforma on hi havia el codi es va extingir i ara caldria traduir-lo a Python o a algun altre llenguatge i vincular-lo a una nova plataforma. També caldria –explica la Martina– afegir-hi un plus de complexitat, per tal que realment fos versemblant i del tot útil.

En tot cas, la prova pilot amb alumnes de primer i segon d'ESO de l'escola on anava la Martina va resultar útil i positiva. D'entrada, deixava clar que no interactuaven amb una persona real, sinó amb una «il·lusió» de persona, semblant a ells –«vam procurar que parlés un llenguatge semblant al seu, en expressions i maneres de parlar».

Així els interlocutors podien mostrar-se més oberts, perquè una de les coses que costa més als qui pateixen *bullying* és explicar el problema a una persona propera. Amb el Trustbot no se sentien jutjats, sinó escoltats i entesos.

I, com el va crear i com funciona? Es basa en preguntes que s'han de contestar amb poques paraules i simples, i per això «només ha calgut introduir una quantitat baixa de possibles respostes per a cada pregunta». Amb el codi, el programa que vertebrava el *chatbot* «compara l'entrada de l'usuari amb una d'aquestes opcions i dona la reacció que té planejada per a cada cas».

La Martina ho detalla així: «Les preguntes parteixen d'un arbre de decisió creat conjuntament amb el

personal d'atenció psicològica de l'escola. És l'usuari qui dibuixa el camí de branca en branca, amb les seves respostes.»

La Martina va pensar les preguntes a partir d'una taula de marcadors d'assetjament de la Generalitat de Catalunya. Va haver de formular-les de forma que la informació que l'usuari hagués de donar quedés bastant polaritzada cap a una banda o cap a una altra.

Pel que fa a l'elaboració del diagnòstic final orientatiu, «el vaig basar tant en la quantitat com en la qualitat. És a dir, hi ha preguntes a les quals respondre amb un 'sí' és molt significatiu com a marcador d'un possible assetjament, i es tenen en compte amb més importància. I n'hi ha d'altres de menys decisives, que també són tingudes en consideració».

La idea principal, diu, és que la informació que es traspasa als professors sigui el veredict final i les respostes a les preguntes, però, tenint en compte, també, que no es registra mai el nom de l'usuari, per mantenir-ne l'anonimat. Pot ser la mateixa víctima qui recorre al *chatbot* o bé algú que està veient què està passant però que té por de fer un pas endavant. Amb dades com el curs, i potser la classe, els professors es poden fer una idea del problema i focalitzar més la seva atenció, sense trencar directament l'anonimat dels alumnes.

I es pot dir que Trustbot s'ha creat amb eines d'IA? «Si ens referim a l'aspecte més tècnic, el modelatge del Trustbot per assemblar-se a una persona, sí que comptaria com a branca de la IA, tot i que no hi considerem factors més avançats, com l'aprenentatge

automàtic, que són el que faria que aprenguéis i fos també més versemblant.»

Així doncs, Trustbot pot ser una eina important i molt útil, però les decisions continuaran essent dels experts: «En aquest sentit, m'agrada reiterar que Trustbot no pretén substituir cap professional del sector. Al cap i a la fi, cada cas d'assetjament és molt complex i té uns matisos que una intel·ligència artificial tan simple mai no seria capaç de detectar.»

Però, en canvi, «sí que es pot utilitzar com una eina més que contribueixi que els alumnes que estiguin patint una situació d'assetjament, o que en siguin testimonis, trenquin la barrera i facin el primer pas d'explicar el seu problema abans de parlar amb una persona que els pugui ajudar directament», afegeix.

A la Martina Massana li interessa el camp de la IA: «Més que por, com que ara s'ha generalitzat una mica, amb noves iniciatives com ChatGPT i Dall·E –un altre programa d'OpenAI que crea imatges a partir de descripcions textuales–, tinc molta curiositat. Trobo que és una eina més, molt potent, però a la qual ens acabarem acostumant com ho hem anat fent amb totes les noves tecnologies.»

Tanmateix, considera que «s'ha d'anar amb molta cura quan donem a aquests organismes la possibilitat de prendre decisions. No podem deixar que aprenguin i es basin només en informació garbellada i “mastegada”. Hem de tenir en compte que el coneixement generat a partir d'aquesta informació inevitablement conté una part de ficció que hem de reconèixer abans

de continuar construint, i això només ho aconseguirem valorant el factor humà que hi ha al darrere. Hem de mantenir un equilibri: hem d'avançar amb aquestes novetats, però no deixar que ens arrosseguin.»

## Vida digital després de la mort

El 28 de novembre de 2015, Roman Mazurenko va morir de forma inesperada, atropellat per un cotxe. Era el millor amic d'Eugenia Kuyda, que va quedar desolada. I, per alleugerir els seus sentiments de pèrdua, va crear un *chatbot* per poder seguir conversant amb en Roman. Va utilitzar més de 10.000 textos per reconstruir una personalitat tan fidel com fos possible. I, al cap de pocs mesos, ja tenia un Roman Mazurenko virtual.

L'experiència li va semblar tan positiva que el 2017 va crear l'empresa Replika, especialitzada en robots conversacionals, amics virtuals que podien fer companyia a gent que se sentís sola.

L'empresa va anar molt bé, potser massa bé. Al cap de pocs anys, tenia milers d'usuaris i molts en demanaven versions avançades, com personalitzar el seu *chatbot*, que a mesura que avançava la conversa anava coneixent millor el seu interlocutor. I com que amb el tracte arriba l'afecte, alguns van començar una relació tan íntima i emotiva que eren inseparables del seu *chatbot*. Les converses pujaven de to i alguns, diuen, van acabar casant-s'hi.

Però el que ara ens interessa d'aquesta història és la motivació de Kuyda. Perdre un amic no li havia d'impedir continuar parlant amb ell, gràcies a la IA. Ja n'hi havia precedents. Potser el primer sistema fou *Virtual Eternity*, creat per Intellitar Inc. l'any 2010. L'anomenaven *plataforma d'avatar intel·ligent* (IAP, en anglès) i permetia crear clons digitals de persones reals. Una de les possibilitats era crear un avatar propi o d'una persona propera per seguir-hi en contacte després de la seva mort.

El nom ja ho deia tot: eternitat virtual. Enviant una foto, fent un test de personalitat i aportant-ne una mostra de veu, es començava a entrenar l'avatar. (Dit sigui de passada, fixem-nos què ha passat amb el terme *avatar*. Per als hinduïstes, era l'encarnació terrestre d'un déu. I ara és la desencarnació d'un individu per transformar-lo en un ésser digital.) La IA després s'anava alimentant de més documents de la persona i n'anava construint la rèplica virtual. L'usuari viuria per sempre, almenys en forma de *chatbot*.

Per ironies de la vida, l'empresa que prometia l'eternitat virtual va morir al cap de poc, al 2012, quan tenia prop de 10.000 clients. Però hi ha hagut experiències posteriors, a més de Replika.

L'any 2016, el pare de James Vlahos s'estava morint de càncer de pulmó. En James, expert en intel·ligència artificial, no es va voler resignar a perdre el contacte amb el seu progenitor i va començar a redactar una mena de memòries del seu pare, amb les seves vivències, aficions, jocs i cançons preferits... En total, va

omplir unes 200 pàgines. La feina no va acabar aquí. Desitjava alguna cosa interactiva. Així que amb aquesta informació va dissenyar i desenvolupar un *chatbot* que era la rèplica del seu pare: un *dadbot*, com el va anomenar.

I com en el cas d'Eugenia Kuyda, el projecte personal es va convertir en una empresa: Hereafter AI. Sembla que va començar a fer-ne les primeres proves el 2019 amb l'actor Andrew Kaplan. I no sé si la millor opció per a una empresa de *chatbots* –que sempre poden tenir problemes de privadesa– era triar com a primer client o com a conillet d'índies un actor especialitzat en pel·lícules d'espies. En tot cas, l'empresa va anar molt bé i, a part de *dadbots*, Vlahos també ofereix *mombots*, *siblingbots* o *friendbots*.

Vlahos ha muntat un gran negoci i, si bé de vegades ofereix algun matís sobre els usos dels seus sistemes, aprofita per aportar noves idees. El 20 d'octubre de 2022 va publicar una piulada en què recomanava un article sobre el tema de l'escriptora Charlotte Jee a la *MIT Technology Review*. I Vlahos deia això: «Els avatars d'IA no poden (i no haurien de) reemplaçar les persones estimades que hem perdut. Però poden proporcionar una manera més rica de recordar, i també una manera d'aprendre més de les persones que estan vives i sanes.» Així, segons ell, els *chatbots* no tan sols serveixen per no perdre del tot les persones estimades, sinó que ens poden ajudar a conèixer millor els vius. És que no val més la pena conversar molt i amb franquesa i intentar comprendre l'altre?

Necessitem *chatbots* plens de dades per acostar-nos més als altres?

No sempre les iniciatives d'aquest tipus han reeixit. L'any 2021, Microsoft va patentar un *chatbot* que es podia perfilar amb les característiques d'una persona morta. La reacció del públic va ser tan contundent que Microsoft no només va cancel·lar el projecte, sinó que va qualificar la seva pròpia idea d'inquietant (*disturbing*).

El projecte fins i tot preveia que es pogués representar la persona en una versió en 3D. El *chatbot* s'hauria alimentat amb la quantitat ingent de dades que tots anem deixant per les xarxes, incloses imatges nostres i potser fragments de veu.

És ben diferent que algú decideixi contractar una empresa per tenir un avatar un cop deixi aquest món o que això es pugui fer sense que l'interessat ho hagi promogut o, almenys, hagi donat l'autorització perquè els seus familiars o amics ho facin. Però aquest és un risc ben real d'aquestes tecnologies. Tothom deixa un rastre immens a les xarxes. Estem a centenars o a milers de bases de dades. Encara que no siguem gaire usuaris d'internet, som clients de diverses entitats, tenim seguretat social, un dia o altre anem al metge o a l'hospital, estem en alguna empresa... Fins i tot si no volem ser a les xarxes socials, això no evita que d'altres ens hi etiquetin, fins i tot amb imatges que no sabem que circulen pel món dels bits. Per això, que facin còpies virtuals nostres és difícil d'evitar. Les poden fer quan no hi siguem o mentre estiguem vius i amb les inten-

cions més diverses. Per això, el projecte de Microsoft era inquietant o, segons una altra possible traducció, pertorbador.

És clar que existeix el dret a l'oblit digital, però ara com ara no és un procediment fàcil i sempre pot quedar molta informació sense eliminar. L'important seria evitar que aquests sistemes es nodrissin d'informació sobre persones que no han donat el seu consentiment.

De fet, el gran problema es dona quan els avatars tenen encara vius els seus models. Es podria pensar que, un cop morts, l'avatar serà només un conjunt de bits, certament molt sofisticat, però que no tindrà emocions. I que nosaltres, si no creiem en una altra vida, no ens assabentarem de res. Però, fins i tot els no creients, solen tenir pensaments sobre com volen ser recordats. Pensar que un dia es transformaran en éssers digitals i continuaran presents, d'alguna manera, en aquest món, on generaran missatges i tindran converses, pot ser una idea engrescadora per a alguns, però gens agradable per a d'altres.

I centrant-nos en els *chatbots* per continuar conversant amb els éssers que han marxat, els anomenats *griefbots*, són realment positius? No seria millor intentar acceptar la pèrdua? El procés de dol és molt delicat i personal, i cadascú el passa com pot. Per tant, dir de manera contundent que no és positiu o recomanable seria inadequat.

Tot i ser un tema molt nou, ja ha generat força debat, però encara no hi ha gaires estudis a la literatura científica. Un dels que he trobat és d'agost de 2021 i el

van publicar a la revista *Current Psychology* investigadors de la Universitat de Milà i de l'Institut Europeu d'Oncologia de la mateixa ciutat. Es titulava «From virtual to real healing: a critical overview of the therapeutic use of virtual reality to cope with mourning». Els autors comencen descrivint el cas del documental *Et vaig trobar*, de la cadena sudcoreana de televisió MBC. Mostra el cas d'una mare que interactua amb l'avatar de la seva filla, morta quatre anys abans, quan en tenia set. El documental es pot visionar a la web de MSC: [https://content.mbc.co.kr/program/documentary/3479845\\_64342.html#](https://content.mbc.co.kr/program/documentary/3479845_64342.html#)



*Figura 7. Frame del documental Meeting you (2020)*

Els investigadors consideren que aquesta nova aplicació hauria d'encetar un debat sobre les seves possibles implicacions amb el dol, «una experiència dramàtica i

natural que pot tenir impactes psicològics profunds en individus fràgils, situats en entorns virtuals». A l'article, es fan diverses consideracions sobre el tema, inclòs el possible ús terapèutic de la realitat virtual per afrontar el dol.

En resum, els principals temes que cal tractar actualment són: la manca d'hipòtesis i de protocols clínics, la necessitat de fer una avaluació acurada dels possibles riscos psicològics associats a aquesta experiència, la manca d'informació sobre possibles danys en usuaris amb problemes mentals i la necessitat d'escoltar i comprendre les opinions públiques i d'adaptar aquestes intervencions a les característiques individuals. La realitat virtual, afegeixen els autors, ja s'ha mostrat eficaç en alguns trastorns d'ansietat i podria ser útil en algunes persones per tractar processos traumàtics específics.

Així doncs, no es pot descartar aquest sistema, però potser tampoc no s'ha de generalitzar alegrement ni s'han de donar falses expectatives, perquè per a algunes persones no serà convenient. I, com escrivia la periodista britànica Amber Louise Bryce a *Euronews* el març del 2023,<sup>20</sup> potser és més important viure al màxim «el fràgil i fugaç present amb els nostres éssers estimats, abans que ens convertim en pols i en píxels».

---

20 Vegeu: <https://es.euronews.com/next/2023/03/16/tecnologia-del-duelo-la-ia-se-utiliza-para-resucitar-a-los-seres-queridos>

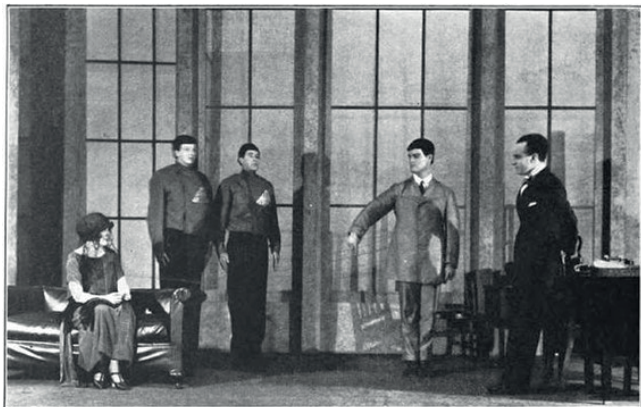


## Passeig amb els robots

«Sí, senyor.» Són les primeres paraules pronunciades per un robot en un escenari. Fou el 25 de gener de 1921, al Teatre Nacional de Praga. No és que el 1921 a Txèquia ja haguessin fabricat robots que poguessin actuar. Les paraules les va dir l'actriu que representava *Sulla*, un dels robots de l'obra *RUR*, de Karel Čapek.

El títol fa referència a les sigles de l'empresa Rossumovi Univerzální Roboti («Els robots universals de Rossum»). La paraula li va recomanar a Karel el seu germà Josef i en txec significa «treballador forçat». Ell havia pensat anomenar-los *labori*, però li semblava massa literari. Rossum és un fisiòleg i el seu cognom deriva de la paraula txeca *rozum*, que significa «raó, saviesa o sentit comú». Treball forçat i saviesa, tot al mateix escenari.

Sembla que la paraula prové del txec antic i de l'antic eslau eclesiàstic i s'origina en els termes *rabota*, «servitud», i *rabu*, «esclau». *RUR* es va estrenar aviat a diversos teatres d'arreu del món: a Londres, el 1921, i a Nova York, l'any següent. A la traducció anglesa, la primera que es va fer, es va conservar la paraula sense traduir. Potser era per mantenir les sigles del títol, potser per la musicalitat, potser perquè no hi havia un equivalent exacte d'una sola paraula en anglès, potser per mandra...



*Figura 8. Representació de RUR al Teatre Guild de Nova York.  
(Font: Col·lecció Francis J. Bruguière)*

A Barcelona, es va representar dues vegades, el 27 i el 28 de gener de 1928 (quatre anys després que a París), al Teatre Català Romea. I, curiosament, A. V. Bejeck i Carles Capdevila van triar la paraula *robot*,

que, segons Joan Coromines, prové del paleoeslau *rabota* i significa «esclavatge», o bé del rus *rabotat*, que vol dir «treballar». En català no existeix, però en francès significa *robot*. Al final, el *robot* s'ha convertit, com a tot arreu, en *robot*. El TERMCAT defineix el robot com una «màquina automàtica que manipula objectes i executa operacions d'acord amb les ordres establertes en el programa informàtic que en controla els moviments».

Els robots ideats per Rossum –és a dir, per Čapek– no eren mecànics, sinó basats en la bioquímica. Rossum s'havia retirat a una illa remota per estudiar la fauna marina oceànica i va intentar imitar la matèria viva mitjançant la síntesi química. Anys més tard, el seu nebot li va fer veure la manera de produir màquines vives i intel·ligents. Com que els robots estaven destinats únicament a treballar, els va crear sense ànima, perquè no tinguessin sentiments, ni plantegessin exigències.

A l'obra, el problema sorgeix amb Elena Glory, filla del professor Glory, de la Universitat d'Oxbridge [sic] i presidenta de la Lliga Humanitària. Després de visitar la fàbrica, s'enamora de Harry Domin, el director, i s'hi casa. Més endavant, convenç el doctor Hallmeier, cap de l'Institut de Psicologia i d'Educació de Robots, de fer els robots més humans. Però així deixen de ser submisos i es gesta la seva rebel·lió, que conforma el nus de l'obra.

El mot *robot* va assolir popularitat i aviat va ser incorporat a diverses llengües. El 1928, apareixen robots

a la pel·lícula *Metrópoli*, de Fritz Lang, però en aquesta ocasió ja són de metall. El 1942, Isaac Asimov va idear les tres lleis de la robòtica i les va enunciar en el conte breu *Runaround*. Les lleis provoquen situacions paradoxals, plantejades de forma enginyosa per Asimov en els contes del recull a *Jo, robot*, publicat el 1950. I, més tard, la robòtica es convertirà en una disciplina científica –i els robots, en font de fantasies utòpiques i distòpiques.

Tanmateix, els temors comencen molt aviat. Ja als anys trenta, ja apareixen a la premsa anuncis en què alguns col·lectius expressen preocupació per si els robots els prendran la feina. Ara que hem viscut la vaga de guionistes i d'actors de Hollywood, en gran part motivada per la por que la IA els tregui la feina, és curiós veure que el 9 de març de 1931 es va publicar al *Simpson's Leader-Times*, un diari de Kittanning, Pennsilvània, un anunci de la Federació Americana de Músics amb el títol «The Robot at the Helm» (el robot al timó). El dibuix mostrava un robot fent fora un músic i denunciava el risc que corrien aquests professionals si les màquines feien la *canned music* que sonava en directe. Un any abans, el 5 de juny de 1930, el *Bradford Era*, diari també de Pennsilvània, havia publicat un anunci que deia que amb la música gravada només caldrien 300 músics a Hollywood per proporcionar totes les simfonies als cines o als teatres.

Cal comprendre també el context: l'any 1933, amb la Gran Depressió s'arribaria a un 25% d'atur, amb 13 milions de persones sense feina. I la imaginació posava

robots a totes les ocupacions. L'abril de 1934, *Modern Mechanics* mostrava una il·lustració on es veia Jack Dempsey boxejant amb un robot.

A l'altra banda, hi havia el Moviment Tecnocràtic, que sostenia que la ciència acabaria amb tots els problemes. Proposava suplir els polítics per científics i enginyers. Un robot il·lustrava la portada de la revista *The Technocrats' Magazine* i plantejava aquesta tria: trenta milions sense feina el 1933 o uns ingressos garantits de 20.000 dòlars per a cada família. Era la «revolució sense vessament de sang» en què potser els robots treballarien i la majoria de la gent simplement cobraria un sou.

Però, deixant de banda pors i fantasies, els robots van arribar a ser una realitat al cap d'uns quants anys. No tenien l'aparença humana de les fantasies de ciència-ficció. Eren braços articulats que podien executar diferents tasques preprogramades. El primer robot industrial pròpiament dit va ser l'*Ultimate*, patentat als Estats Units l'any 1954, que a partir de 1961 es faria servir a la planta d'assemblatge de la General Motors. El primer robot industrial a Catalunya –i a tot l'Estat– es va instal·lar el 1974 a la planta de SEAT a la Zona Franca de Barcelona. L'any 1980, SEAT ja comptava, a la mateixa planta, amb 18 robots operatius.

Eren els primers passos, però calia formar tecnòlegs i investigar. En l'arribada de la robòtica al nostre país, la UPC va acomplir un paper essencial, quan encara es deia Universitat Politècnica de Barcelona.

## La robòtica arriba a Catalunya

Per veure com s'investiga en robots quirúrgics he hagut d'entrar primer a SPECIPIG, una empresa del Prat de Llobregat especialitzada en el porc com a model per a la recerca preclínica. Dona suport a empreses biomèdiques, farmacèutiques, biotecnològiques i de dispositius mèdics en totes les fases dels seus projectes.

Tanmateix, fer servir el porc com a model té un inconvenient: creix ràpidament i arriba a pesar entre 200 i 300 quilos. Per això, SPECIPIG, en col·laboració amb l'Institut de Recerca i Tecnologia Agroalimentàries (IRTA) i l'empresa Semen Cardona, ha desenvolupat porcs menuts, que pesen uns 60 quilos, que són els que s'utilitzen en experimentació mèdica perquè són molt més fàcils de mantenir i manipular. Quan es va crear SPECIPIG, a Europa només hi havia un proveïdor.

SPECIPIG té diverses instal·lacions, entre elles un quiròfan, amb unes mesures higièniques extraordinàries, perquè les empreses puguin fer recerca aquí en comptes d'haver d'endur-se el porc al seu laboratori. Per la seva proximitat a Barcelona i a la UPC, és el lloc ideal perquè s'hi hagi establert, en dues naus, l'empresa RobSurgical,<sup>21</sup> una *spin-off* que la UPC va crear l'any 2012, juntament amb l'Institut de Bioenginyeria de Catalunya (IBEC), per donar resposta als reptes de la cirurgia robòtica mínimament invasiva.

És allà on em rep el seu director científic, Josep Amat, catedràtic emèrit de la UPC i impulsor del

---

<sup>21</sup> Vegeu: <https://www.robsurgical.com/>

Grup de Recerca en Robòtica Intel·ligent i Sistemes (GRINS). Em mostra l'empresa i m'explica les seves recerques. Abans de crear l'empresa, entre els anys 2008 i 2011, la UPC i l'IBEC ja havien fet recerca amb experts de la Clínica Mayo, que treballaven amb la cirurgia robòtica mínimament invasiva. El 2012, es crea RobSurgical i comencen a desenvolupar el prototip funcional del sistema *Bitrack*. L'any 2014, van realitzar la primera cirurgia amb models animals; el 2018, *Bitrack* va completar amb èxit la validació tècnica de la seva unitat prèvia al disseny i, el 2020, va obtenir la certificació ISO 13485.

Però el camí és llarg i complex i, malgrat l'èxit del desenvolupament, les lleis són molt exigents. S'han de fer moltes operacions per validar el model i que es pugui fer servir. I això no es pot fer de manera genèrica, sinó amb diverses intervencions per a cada tipus de cirurgia. Per això, tenir la seu a SPECIPIG facilita molta la feina.

Conec en Josep Amat des de fa molts anys. Tots dos, per bé que vam néixer a Barcelona, tenim les arrels i uns grans lligams directes amb Sant Feliu de Guíxols i, de fet, som mig ganxons. El pare d'en Josep era un gran pintor i té obres delicioses que ens recorden un Sant Feliu en bona part desaparegut.

L'any 1995, hi vam fer un reportatge per als informatius de TV3 perquè, gràcies a en Josep Amat, Sant Feliu va ser la seu de la Conferència Internacional de Robòtica Avançada (ICAR). A part de les encantadores imatges de robots movent-se vora l'escenari medie-

val del monestir, també era curiós que una població que en aquell moment tenia poc més de 16.000 habitants hagués agafat el relleu de Tòquio com a seu de la conferència. Veient la llista de les seus, continua essent, amb diferència, la ciutat més petita que ha acollit una trobada d'aquestes característiques, que ha passat per Pisa, Coïmbra, Seattle, Daegu, Estambul i Hong Kong, entre d'altres ciutats.

Amat parla amb entusiasme i deixa anar una allau de dades, històries, anècdotes, detalls tècnics... M'ensenya un PowerPoint sobre la història de la robòtica a Catalunya, que té 115 imatges..., perquè diu que no la pot resumir amb menys. El miro, l'escolto i vaig prenent notes, i així m'assabento de com la ciència de la robòtica va arribar a Catalunya, més de mig segle després que al Teatre Romea s'estrenés *RUR*.

Amat va estudiar a l'Escola Tècnica Superior d'Enginyeria Industrial de Barcelona. L'any 1983, hi va guanyar la càtedra d'Arquitectura i Tecnologia de Computadors, i des d'allà va impulsar la robòtica. Però l'havia conegut a fons anys abans i ja havia començat a divulgar-la i a intentar fer-la present en els plans d'estudis.

M'explica que, el 1980, l'aleshores director general de Política Industrial de la Generalitat de Catalunya, l'enginyer empordanès Jaume Bassa, va proposar atorgar 24 beques amb totes les despeses pagades perquè científics o tecnòlegs anessin durant 15 dies a algun lloc dels Estats Units, del Regne Unit, d'Alemanya o del Japó a informar-se a fons d'algun tema innovador i

rellevant. En aquells moments, el bioquímic Joan Oró era diputat al Parlament i havia tornat a Catalunya per intentar impulsar-hi la recerca –les dificultats burocràtiques i la manca de transferències de l'Estat el van decebre i l'any següent se'n tornaria als Estats Units.

Amat va demanar anar al Japó per conèixer la robòtica industrial. Li van donar una de les places i el viatge li va permetre endinsar-se en aquest món en un dels països –o potser en el país– on els robots estaven més avançats. Cada dia feia una visita al matí i una a la tarda, amb un intèrpret. Va visitar Yamaha, Casio, una fàbrica de ponts fets amb làser amb robots, la Universitat de Tòquio, Osaka...

De tornada a Barcelona, tenia el compromís de fer conferències per explicar què havia vist. En va fer unes quantes, potser una vintena, arreu del país. I, al mateix temps, es preguntava com podia ser que a la UPC no es fes res sobre robòtica. Hi havia l'Institut de Cibernètica, amb quatre línies de recerca: simulació, bioenginyeria, sistemes numèrics i tècniques analògiques i digitals, que era la que ell dirigia. Va proposar canviar-li el nom pel de tècniques i sistemes digitals i robòtica, però no ho va aconseguir.

Com que en aquell moment era professor de la Facultat d'Informàtica, i la informàtica era la base per programar els robots, va començar a impartir-hi la matèria.

En un incís, m'explica com es va crear la Facultat d'Informàtica de Barcelona. Es volia crear l'Escola d'Informàtica, però el Consell de Ministres volia que

es fes primer o almenys simultàniament a Madrid. El rector Gabriel Ferraté va pensar que si, en comptes d'una escola, es creava una facultat, no caldria passar pel Consell de Ministres, sinó pel de Rectors. I va aconseguir que aquests ho aprovessin. Era l'any 1976.

Fou allà on Amat va començar a ensenyar robòtica el 1984, dintre de l'assignatura Control en Temps Real –el nom no es va poder canviar fins al 1991, quan es va modificar el pla d'estudis vigent. Abans, des del 1981, havia ensenyat robòtica en cursos de formació continuada a l'Associació d'Enginyers Industrials de Catalunya, juntament amb Alícia Casals i Fèlix Ferrer.

Al mateix temps, l'aplicació de robots progressava. L'any 1982, s'introduïa la robòtica a la planta de Miniwatt de Philips, a la Zona Franca. Era una de les primeres aplicacions industrials fora del sector de l'automòbil. Aquell mateix any, la japonesa Nissan augmentava al 80% la participació a Motor Ibèrica, també a la Zona Franca, i iniciava la robotització de la planta. El 1983, l'empresa ASEA (més tard, ABB) inaugurava, a Sant Quirze del Vallès, la primera fàbrica de robots d'Espanya, si bé només s'hi feia el muntatge. Seria MAPS la que en desenvoluparia de propis al cap de pocs anys.

La recerca també avançava, i els diaris i les revistes en començaven a parlar. L'octubre de 1983, *Computer World* titulava: «La UPC dotarà de visión a robots que se enviarán a Marte». Al novembre, l'*ABC* també tractava el tema i deia: «José [sic] Amat, en bus-

ca del chip prodigioso» i destacava que col·laborava amb el MIT de Boston. I, l'abril de 1985, *El Noticiero Universal* proclamava «Josep Amat, padre de los robots catalanes».

Tot va quallar tant que, quan l'any 1991 es va crear la Universitat de Girona, ja des del principi va incloure una assignatura de robòtica a la titulació d'Enginyeria. Ara hi ha, a Catalunya, nou graus de robòtica a les universitats públiques i quatre en centres privats, inclosa la Fundació Eduard Soler de Ripoll. I hi ha nou màsters de robòtica a les universitats públiques i dos a les privades.

Si a Catalunya es van poder tenir professors de prestigi mundial fou també, en bona part, gràcies a l'ajut del francès Georges Giralt. Fill de pares catalans exiliats establerts a Tolosa de Llenguadoc, va ser un dels grans experts mundials en robòtica. Va contribuir a la creació del Laboratori d'Anàlisi i d'Arquitectura de Sistemes (LAAS) i va ser director de recerca del Centre Nacional de la Recerca Científica (CNRS), on havia ingressat el 1956. Com a mostra del seu prestigi, només cal afegir que el premi europeu per a contribucions extraordinàries a la robòtica, que s'atorga des del 2002, porta el seu nom.

Sempre disposat a ajudar amb entusiasme, va contribuir que molts grans experts vinguessin a Catalunya i hi impartissin seminaris, cursos o conferències. I també va ajudar que alguns congressos internacionals es fessin aquí. A part de l'interès purament tècnic que podien tenir, també cal dir que passar uns dies a

Barcelona era un bon atractiu. I que, en arribar el cap de setmana, Amat els portava amb la seva embarcació a navegar per la Costa Brava, cosa que completava una gran experiència.

Devien estar encantats. En una ocasió, Amat els va portar a dormir al vaixell a la badia de Tamariu. No hi havia caigut, però era el 23 de juny. I, quan van començar a disparar-se coets, bengales i altres focs d'artifici, els experts en robòtica varen quedar molt contents i varen agrair a Josep Amat que hagués organitzat aquell espectacle en honor seu. L'Amat, murri, no ho va desmentir.

Una de les reunions internacionals més importants va ser el 15è Congrés de la Federació Internacional de Control Automàtic (IFAC), que s'havia constituït l'any 1957. Es va fer a Barcelona del 21 al 26 de juliol de 2002. I Amat, que presidia l'equip organitzador, va demanar que li deixessin preparar l'espectacle de la sessió inaugural a la seva manera, sense haver de compartir-lo, per assegurar la sorpresa, i l'hi van acceptar. Només ho sabia Gabriel Ferraté. L'espectacle era molt innovador: un ballet robòtic al Gran Teatre del Liceu. Tot i que era el teatre reconstruït després de l'incendi de 1994, una vegada més apareixia la combinació harmònica d'una tecnologia moderna en una institució que datava de mitjan segle XIX.

El primer que van haver de fer els organitzadors de la cerimònia fou calcular l'alçada dels robots perquè es veiessin bé des de platea. Van constatar que havien de tenir almenys quatre metres d'altura. Després, calia

construir tota l'estructura amb un plàstic molt lleuger. I, finalment, aconseguir que ballessin.

És així com sis robots platejats antropomòrfics van mostrar unes mínimes habilitats coreogràfiques a l'escenari del Liceu. El públic va quedar impressionat, perquè aconseguir uns moviments mínimament gràcils i harmoniosos era ben difícil amb aquelles màquines. Amb quin programa informàtic ho havia aconseguit?

Més que un programa, eren sis becaris, introduïts cadascun en un dels robots. Anaven amb unes cames que actuaven com unes xanques i havien assajat molt. Per donar més versemblança, durant l'espectacle un robot es va aturar i se li va encendre un llum vermell. Un altre becari va sortir amb un ordinador, el va endollar i va fer veure que teclejava algunes instruccions. En poca estona, el robot es va recuperar.

Al final, els robots van saludar i van rebre els aplaudiments del públic. Malauradament, la premsa del moment no només va mostrar poc interès per la reunió d'experts, sinó que també va passar per alt aquest fet original, encara que fos anecdòtic.

Dos dels sis robots van ser indultats. Els vaig veure a la Facultat d'Informàtica, en la meua primera visita a l'Institut de Robòtica i Informàtica Industrial (IRI). Estaven penjats a una certa altura, just al costat d'una de les escales, com aquells gegants que presideixen l'entrada en alguns ajuntaments. Va ser una imatge agradable i simbòlica, abans que Cecilio Angulo m'ensenyés què s'estava fent a l'IRI.

Malauradament, el 2023 es va decidir que aquell no era el seu lloc. Es van retirar, es van desmuntar i sembla que, malauradament, es van llençar. Els robots dansaires havien fet el seu darrer pas de ball.

## **Per la vall inquietant**

Josep Amat m'explicava que aviat va tenir estudiants de molt nivell i que alguns s'havien quedat a investigar i a fer docència a la UPC. Va esmentar dues dones joves que després han tingut una gran rellevància en el món de la robòtica: Alícia Casals i Carme Torras. Aquestes són les converses que exposo a continuació.

Premis, braços robòtics en miniatura, algun petit robot antropomòrfic i un astronauta assegut a l'extrem mirant cap al firmament. Per darrere d'una de les plaques, apunta l'escultura *Dona i ocell*, de Joan Miró. Aquest és el muntatge que té la doctora Alícia Casals damunt d'un arxivador. I, al costat, just a sota dels peus de l'astronauta, hi ha el dibuix que l'artista Perico Pastor va fer per al 50è aniversari de la UPC, el 2021: «50 anys sense límits».

Amb un somriure lleuger, dibuixat permanentment als llavis, la doctora Casals respon pacientment les preguntes que li plantejo. És catedràtica del Departament d'Enginyeria de Sistemes, Automàtica i Informàtica Industrial de la UPC i responsable de dos grups de recerca: GRINS (Grup de Recerca en Robòtica Intel·ligent i Sistemes) i TecSalut (Grup de Recerca

en Tecnologies de la Salut). Més tard, al laboratori, me n'ensenyarà alguns exemples, com un sistema robòtic de cirurgia cranioencefàlica, dissenyat l'any 2006, que requereix una precisió extraordinària.

D'entrada m'ofereix una visió històrica breu de la robòtica i em parla dels autòmats, uns artefactes mecànics que es remunten, pel que sembla, a cinc segles abans de la nostra era, amb Arquites de Tàrent i el seu colom mecànic, capaç de volar gràcies al vapor d'aire a propulsió. A l'edat mitjana i a l'edat moderna, vindrien coses més sofisticades i una de les més famoses seria l'ànec del rellotger francès Jacques Vaucanson. L'ocell era transparent i batia les ales, menjava i realitzava completament la digestió.

No varen faltar precursors molt rudimentaris del *Deep Blue*, amb autòmats que jugaven als escacs. Entre ells, el turc, construït el 1769, que tenia un maniquí amb forma humana vestit amb robes àrabs i que semblava invencible. Es va descobrir que, com els robots del ballet del Liceu, a part d'un mecanisme simulat tenia al seu interior un hàbil jugador, que era qui en controlava els moviments.

Per Casals, potser els primers braços articulats es podien considerar autòmats, però, quan els robots es poden programar perquè variïn els moviments, «ja no ho són ben bé, sinó alguna cosa més. En l'àmbit industrial, els autòmats fan uns moviments concrets, condicionats per la pròpia instal·lació. De del moment que aquests canvis es poden fer amb un programa, ja són màquines més flexibles; són robots.»

I torno a la pregunta recurrent: Són intel·ligents? «Poden ser-ho –explica. Quan els condicionants de l'entorn ho requereixen, tenen un cert grau d'intel·ligència, perquè agafen informació, la processen i prenen decisions, però és diferent de la intel·ligència humana.»

De fet, si es programen, els robots fan una seqüència d'accions i, «segons la informació que reben dels sensors, poden adaptar la seva actuació a les condicions de la tasca». Per això poden modificar el comportament: «Un robot mòbil es troba amb eventualitats i, si repeteix el mateix sense canviar, no és intel·ligent, però, si s'adapta a les condicions canviants de l'entorn, sí.» Com sempre, tot depèn de la definició d'intel·ligència i de com s'hagin programat els robots.

De fet, els robots poden aprendre «si se'ls dissenya i programa perquè ho facin. En aquest cas, li pots ensenyar què ha de fer, o bé ho aprendrà i ho farà per imitació o per altres mitjans d'aprenentatge». Tanmateix, un dels reptes és com fer que siguin fiables.

Posa com a exemple un robot rehabilitador, que pot ser rígid en l'execució de la teràpia, o ser capaç d'adaptar-se a les condicions de l'usuari. Però també ha de ser capaç d'interpretar aquests condicionants: «Si li dones un cop, ho pot prendre com un càstig, en comptes de pensar que és un gest manyac. Copets a l'esquena poden voler dir “molt bé” o “ànims”, però el robot ho pot malinterpretar. És cert, però, que les persones també interpretem sovint les coses de forma diferent.»

I fins a on poden arribar? «Si se li posen sensors cada vegada més potents i amplis, serà més potent. Si disposa de sensors prou potents i amb un nivell de percepció i interpretació més elevat, el robot es podrà comportar més intel·ligentment.»

Alguns dels avantatges dels robots és que no es cansen fent feines tedioses, no tenen mal de cap, no demanen més vacances o dies de permís per fer tràmits personals, no es desanimen ni discuteixen. Però, poden sentir por o frustració?

De fet, no: «Se li pot fer visualitzar una cara de por, enfadada o que riu, i identificar-la... I podem fer que hi reaccioni. Però és una màquina: no ho sent. Simplement capta dades i reacciona tal com l'hem programat.»

«Fins i tot, d'alguna manera, li pots donar una personalitat i fer que et digui "t'estimo", però no serà un sentiment.»

I, a la inversa, ens podem enamorar de la màquina? «És possible, perquè es creen vincles i perquè la persona en pot dependre.» Podem estimar-les més que els humans? «Pot ser.» I si hem perdut una persona molt estimada, podem reconstruir-la de forma virtual? «Sí, i pot ser una teràpia per combatre el dol. Però potser és una teràpia equivocada.»

I quins riscos hi ha sobre el domini de les màquines? «Una cosa són els errors. Els algorismes que s'executen sense que les persones en tinguin control poden portar a resultats erronis, lògicament. HAL, de *2001: una odissea de l'espai*, es torna contra la persona perquè

li falten dades; no té tota la informació i actua coherentment amb la que té.»

«Però una altra cosa –afegeix– seria programar-los perquè fessin mal. Això ja no seria la rebel·lió de les màquines, sinó màquines programades per cometre accions dolentes.» En principi, no hem de tenir por que als robots se'ls creuin els cables, de forma metafòrica, com es diu d'algunes persones quan cometen determinades accions poc o gens lògiques. Però, per si sols, els robots no es rebel·laran.

Hem parlat de cares i expressions amistoses, de robots que somriuen. I m'explica el que es coneix com a *uncanny valley*, que sovint es tradueix com a «vall inquietant», tot i que l'adjectiu significa més aviat «sorprenent, rara o astoradora», però també pot voler dir «que fa por». Hi ha estudis que demostren que una semblança excessiva a un humà crea repulsa. Potser si veiem que tractem amb una màquina amb una aparença agradable, es crea empatia. Però si veiem que la màquina intenta ser humana, produeix desconfiança, perquè sembla que ens vulgui enganyar i passar pel que no és.

Aquest concepte el va introduir l'expert japonès en robòtica Masahiro Mori l'any 1970. La hipòtesi és que l'aparença més humana d'un robot genera més empatia, fins que un cert nivell d'antropomorfisme comença a comportar repulsió, fins a arribar a un punt mínim, la vall. A partir d'aquí, comença a remuntar.

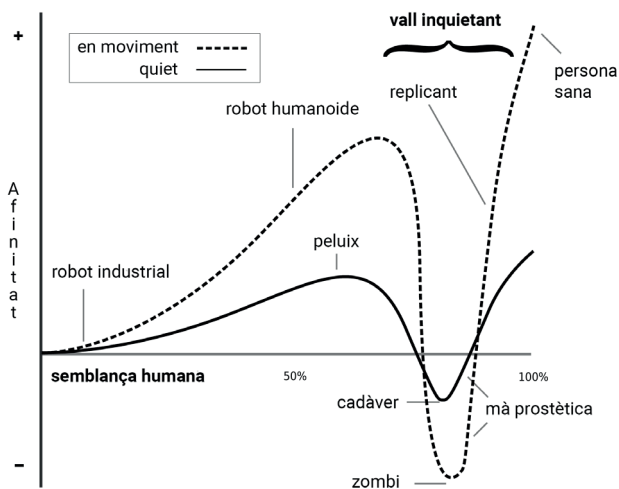


Figura 9. Gràfic de Mori de la seva hipòtesi de la vall inquietant.  
 (Font: Mori / Wikimedia Commons / CC BY-SA 3.0 DEED)

Hi ha qui pensa que l'efecte «*uncanny valley*» és real i qui creu que no, o que potser es dona per causes molt diverses. En tot cas, no s'aplica a determinats àmbits, com les *sex dolls*. Al seu llibre *Sex Robots & Vegan Meat* (Picador, 2020), la periodista i documentalista Jenny Kleeman descriu la seva visita a Abyss Creations, l'empresa de Califòrnia que fabrica les RealDoll, un robot sexual hiperrealista. Hi troba una llarga cua de cossos de silicona sense cap, penjats de cadenes de metall, «com carcasses en un escorxador». Cada cos és diferent del següent, però tots comparteixen «la mateixa impossible diminuta cintura». També

hi ha robots masculins, però molts menys. I els uns i els altres, inspirats en persones reals, tot i que amb algunes característiques més exagerades. És sorprenent, astorador, inquietant. És a dir, plenament *uncanny*.

Hi ha molts estudis i molta controvèrsia sobre aquest efecte. Però la doctora Casals m'ensenya unes imatges que demostren com a alguns experts la semblança no els causa cap repulsió, encara que el robot sigui un doble d'ells mateixos.

Hi apareix Hiroshi Ishiguro, director del Laboratori de Robòtica Intel·ligent de la Universitat d'Osaka. Ishiguro no creu en l'efecte «*uncanny valley*», sinó tot el contrari. Creu que amb semblança humana hi ha més interacció amb la gent. I fa uns quants anys va crear el seu propi *Doppelgänger*, el seu doble, fet de cautxú, silicona i part del seu propi cabell. Veure'ls a tots dos junts sí que és una mica inquietant. Potser el problema és que, al contrari del que passa amb el retrat de Dorian Gray, el robot no envelleix i el seu propietari, sí.

Estic assegut a menys de dos metres d'Ishiguro. No de l'avatar, sinó de la persona real. Estem al laboratori de l'IRI, on s'està presentant l'Avatar Symbiotic Society, un projecte en el marc del programa *Moonshot Goal 1* del Govern japonès. L'objectiu és crear una societat en què avatars i cíborgs parcialment teleoperats puguin suplir les persones en diferents àmbits de la vida i puguin participar activament en la societat, sense els límits imposats pel cos, la ment, l'espai o el temps. El projecte té un calendari prudent i situa l'objectiu al 2050.

La UPC ha estat escollida pels seus impulsors com a banc de proves per estudiar la interacció entre avatars i persones en un context diferent del Japó. En aquest país, els avatars s'estan provant en centres comercials i ben aviat s'estendran a les escoles, els hospitals i altres entorns. Alberto Sanfeliu, catedràtic de Ciències de la Computació i Intel·ligència Artificial de la UPC i investigador de l'IRI, és l'únic estranger en un ampli equip totalment japonès.

A la roda de premsa prèvia al 1<sup>st</sup> Symposium Japan-Spain de l'Avatar Symbiotic Society, celebrat el 3 de novembre de 2023, el professor Hiroshi Ishiguro ensenya una gravació amb una versió virtual del ministre de Transformació Digital del Japó, Taro Kono. És molt realista i s'expressa movent mans i braços. És un dels anomenats geminoides, andròides amb l'aspecte d'una persona real. L'investigador explica que es volen fer cada cop més realistes, fins i tot amb pell semblant a la humana.

Quan se li pregunta per l'efecte «*uncanny valley*», explica que al Japó no es dona. Allà un robot es considera un amic i, quan s'espatlla o la seva vida útil acaba, la gent es posa trista. Fins i tot hi ha cementiris per a robots. És una forma que tenen de reconèixer que robots i avatars també moren, com les persones.

Però la resposta en cultures diferents pot ser diversa. Per això es volen investigar les relacions a diferents zones del món. Si es fan avatars personalitzats, com el seu, pot haver-hi societats que pensin que són la mateixa persona: «l'avatar soc jo?» Si bé els que es dis-

senyen per a tasques diverses seran operats a distància, el seu, com d'altres, ja interactua. Em diu que és millor que ell, perquè té al cap els deu llibres que ha publicat, sencers, sense buits.

I també interactua a les xarxes socials. «I té tanta autonomia que té les seves pròpies opinions? Això pot ser un problema. Vostè i el seu avatar tenen discussions?» Riu, i afegeix: «El meu avatar pensa el mateix que jo.»

A la presentació, l'acompanya el professor Norihiro Hagita, que descriu el projecte *Moonshot Goal 1*. El nom convida a comparar-lo amb el programa *Apollo*: Kennedy es va comprometre a portar un home a la Lluna i retornar-lo sa i estalvi a la Terra. Hagita diu que l'objectiu de *Moonshot* és que tothom pugui fer les activitats que vulgui millorant les seves capacitats amb avatars. L'any 2050, tindrem robots que no estaran condicionats per les limitacions del cervell i del cos humà. A la descripció del projecte, destaquen algunes paraules: *atrevit, ambiciós, d'innovació disruptiva i sense por al fracàs*.

A la web del programa,<sup>22</sup> s'explica que, operant a distància diversos avatars, l'usuari podrà prendre part alhora en diverses activitats, des de la feina fins a l'atenció mèdica, passant per qualsevol altre acte social. Per això, Hagita diu que les limitacions de temps i espai s'hauran abolit. Per fi, haurem aconseguit el necessari do de la ubiqüitat!

---

22 Vegeu: <https://avatar-ss.org/en/index.html>

Segons els investigadors, l'any 2050 els nostres estils de vida hauran canviat de forma important i «tindrem una llibertat més gran per escollir la nostra ubicació i com esmerçar el nostre temps, i els avenços tecnològics hauran ampliat les nostres capacitats». El seu objectiu, afirmen, és desenvolupar aquesta simbiosi amb els avatars «en una futura societat equilibrada».

I, a banda de facilitar les coses als usuaris i donar-los tantes possibilitats, en què més ajudaran els avatars? El projecte està enfocat també a solucionar o pal·liar problemes com la disminució del nombre de naixements, l'envelliment consegüent de la població, la manca de treballadors, les pandèmies, els desastres naturals, l'escalfament global i la crisi climàtica que se'n deriva. I això darrer, com? S'apunta que amb els avatars hi haurà menys desplaçaments i, per tant, menys emissions. Però caldria fer el balanç comptant l'energia que requeriran en tot el seu cicle de vida, «del bressol a la tomba».

Una de les vegades que el doctor Sanfeliu visitava el Japó i es trobava amb col·legues seus va tenir una conversa amb Ishiguro i li va explicar una de les coses que estaven desenvolupant a l'IRI. Es tracta del nou model social de percepció-intenció-acció (Social-PIA) per tal d'explicar i comprendre com un robot i un humà poden interaccionar i fer tasques col·laboratives. Fins ara, en robòtica s'usa el model de percepció-acció, en què el robot percep l'entorn i actua. El nou enfocament inclou el concepte de la intenció humana. Es tracta que un robot entengui la intenció de l'humà amb qui

interactua, prevegi els seus moviments, s'hi anticipi i actuï en conseqüència. Ishiguro va dir de seguida que els interessava molt i el va incorporar al programa.

Sanfeliu m'explica la recerca i en posa un exemple amb unes imatges. Un robot ha de donar una caixa a una persona. Analitzant els seus moviments, s'anticipa i sap on s'ha de posar i com ha d'estirar el braç per donar l'objecte. No ha d'esperar que la persona arribi al seu costat i li estengui la mà.

Es tracta d'un nou paradigma en què el robot s'avança uns quants segons al que farà la persona. I, si aquesta canvia d'opinió sobtadament, el robot ha de tenir la capacitat d'anticipar-se o d'adaptar-se a la nova circumstància.

La recerca, m'explica, és complexa i difícil. Les reaccions humanes poden ser molt diverses, i per causes molt diferents. «Si juguem a pilota i el robot me la passa i jo la hi torno, però en un moment determinat me la quedo, què farà ell? Pensarà que s'ha acabat el joc? Esperarà? Me la prendrà per continuar jugant? I si estem transportant una taula entre tots dos i jo m'aturro? És per fer una pausa i descansar, per deixar-ho córrer, perquè m'he fet mal...?»

No són recerques noves. Ja van començar fa uns quants anys. Però cada cop es revelen més importants. Si l'avatar ha d'estar al nostre servei, ha de saber molt bé què fer en cada moment. Dalt d'una taula, *Sota* s'ho està mirant amb el seu rostre agradable, amistós. L'ha portat Ishiguro i aquest avatar, que no deu arribar al mig metre d'altura, s'estarà a la UPC perquè els inves-

tigadors analitzin quina és la reacció dels estudiants i com hi interactuen –o si l’eviten. El primer dia l’havien de posar a la cafeteria, un bon lloc on intentar socialitzar-se amb el màxim d’estudiants.

## Realitat i ficció

El llibre que veig al despatx de Carme Torras es titula *The Vestigial Heart* (MIT, 2018) i porta com a subtítol *A Novel of the Robot Age*. És la traducció de *La mutació sentimental*, que Torras va publicar el 2008 a Pagès Editors, amb què va guanyar el X Premi Manuel de Pedrolo, i que també s’ha traduït al castellà (Milenio) i al xinès.

És llicenciada en Matemàtiques per la Universitat de Barcelona i va estudiar Filosofia a la UNED. Té un màster en Ciències de la Computació per la Universitat de Massachusetts i un doctorat en Informàtica per la UPC. A més de fer classes a la UPC, lidera el RobIRI (Grup de Percepció i Manipulació Robotitzada de l’IRI) i és responsable, entre d’altres, del projecte CLOTHILDE (*CLOTH Manipulation Learning from Demonstrations*) de robòtica assistencial.

Essent un exemple clar de tecnòloga humanista, vull que em comenti la utilitat de la ciència-ficció per debatre els aspectes ètics de la informàtica i la robòtica, un dels temes que no deixaré de banda. Però començo amb la pregunta que ja he fet als meus altres interlocutors: La IA és intel·ligent? «Hi ha tantes intel·ligèn-

cies...», em diu la Carme. I es pregunta si el quocient intel·lectual (QI) mesura la intel·ligència. Si és així, sí, la IA és intel·ligent. El que l'acosta més a la intel·ligència, matisa, és que pot identificar patrons a partir de correlacions i «té molta capacitat per processar moltes dades i molt ràpid».

«Hem d'anar cap a la col·laboració entre humans màquines –afegeix de seguida–, no a la substitució dels humans per les màquines.» I posa un exemple del benefici que això comporta. Investigadors de Google van crear l'any 2016 un sistema basat en l'aprenentatge profund per millorar el diagnòstic del càncer de mama i detectar amb més exactitud nodes limfàtics. Es diu LYNA (*Lymph Node Assistant*).

Tal com revelen els seus estudis, com el que va publicar el desembre de 2018 l'*American Journal of Surgical Pathology*,<sup>23</sup> quan els metges feien servir el LYNA el percentatge de diagnòstics acurats era superior que si es deixava actuar només l'algoritme o només el patòleg. La sensibilitat per detectar micrometàstasis augmentava del 83% al 91% quan els humans i la màquina treballaven junts.

Carme Torras i el seu equip es dediquen a la robòtica assistencial, és a dir, la que ajuda les persones que no poden realitzar per si soles determinades tasques. Programen els robots o bé fan que aprenguin i imitin els moviments de la persona que els ensenya. En tots els casos, porten sensors per percebre l'entorn i

---

23 Vegeu: [https://journals.lww.com/ajsp/fulltext/2018/12000/impact\\_of\\_deep\\_learning\\_assistance\\_on\\_the.7.aspx](https://journals.lww.com/ajsp/fulltext/2018/12000/impact_of_deep_learning_assistance_on_the.7.aspx)

poder moure-s'hi i actuar. També intenten que millorin en el coneixement de la persona o del pacient a qui assisteixen, per tal de personalitzar-ne el servei.

El projecte CLOTHILDE<sup>24</sup> dissenya robots perquè facin una tasca aparentment tan senzilla com plegar roba. Potser és relativament senzilla per a una persona –en funció de la traça que hi tingui i de la roba–, però per a un robot és complexa.

El març de 2022, estava preparant un article per a la web de notícies 324.cat sobre un robot desenvolupat a la Universitat de Tòquio que era capaç de pelar un plàtan sense esmicolar-lo. Sembla una recerca absurda o poc important, però per a les mans metàl·liques i poc fines d'un robot això no és senzill. Per aconseguir-ho, primer els investigadors van haver de conduir el robot ells mateixos. I, després de 811 minuts, van deixar que el robot s'entrenés tot sol.

En les proves finals, el robot va poder pelar amb èxit un plàtan el 57% de les vegades. I, per fer-ho, trigava menys de tres minuts. No seria una gran publicitat i poca gent compraria un robot per pelar plàtans, però l'objectiu dels investigadors era assajar el sistema d'aprenentatge i demostrar que un robot és capaç, amb aquest entrenament, de fer tasques molt delicades i precises, que es poden aplicar a molts altres àmbits.

Vaig voler incloure a l'article l'opinió de la doctora Torras i em va comentar el següent: «Els robots industrials manipulen, en general, objectes rígids. Amb tres coordenades de posició i tres d'orientació, l'estat

---

24 Vegeu: <https://clothilde.iri.upc.edu/>

de l'objecte queda fixat. Els objectes deformables tenen infinites possibilitats de deformació i el seu espai d'estats pot tenir infinites dimensions.» I va afegir que la dinàmica del moviment del robot els canvia la forma, cosa que no passa amb els objectes rígids.

La manipulació d'objectes deformables és avui un dels reptes més importants de la recerca en robòtica. I, en aquest àmbit, entra la manipulació de roba i del cos humà en la robòtica assistencial.

En aquesta línia de recerca s'ha investigat molt el tacte, que és fonamental per manipular objectes. No és el mateix plegar uns texans que un mocador de seda. Perquè tingui tacte, recobreixen la mà del robot amb una pell sensible. I cal que aprengui bé tots els moviments, que sàpiga com ha de quedar plegada la roba i que sàpiga corregir si el resultat no és l'esperat.

Torras explica que «els robots mai no seran tan intel·ligents com nosaltres. Poden tenir sensors molt potents i visió nocturna amb infraroigs, però els falta sensibilitat i comprensió de les situacions».

Els sentits aconsegueixen un paper cabdal en l'aprenentatge humà. Gràcies al tacte, es capten les diferències entre els objectes i els materials, i s'aprèn a diferenciar-los i a constatar que s'han de manipular de maneres diferents. Una de les coses que, segons alguns investigadors, limiten la capacitat d'aprenentatge de les màquines és que algunes no tenen cos, sinó que són simplement algoritmes amb sensors. Potser es podran aconseguir sensors molt potents que imitin tots els sentits, però això no serà immediat.

I, se'ls poden creuar els cables, de forma metafòrica? Poden ser dolents els robots, o poden tenir mals moments? La doctora Torras explica que «no tenen ni tindran mai consciència. Ni tan sols ara saben que poden fer i quines limitacions tenen». I això arribarà? «Sí, però la consciència humana és diferent; requereix una experiència de vida que no tindran mai.»

Per això la rebel·lió dels robots és una fantasia, perquè «si et vols rebel·lar has de tenir consciència i voluntat». Creu que aquests debats desvien l'atenció i que, sobretot, cal educar els humans.

Perquè són aquests els qui dissenyen els robots. I un tipus de màquines que li preocupen són els *killer robots*, els robots assassins. «Es diu que poden ser més segurs que les persones en una situació de pànic, ja que sempre es poden posar línies vermelles a un robot, però a un humà, no. Un robot se'l pot programar perquè mai, en cap situació, dispari en una escola. A un humà se li pot prohibir això, però es pot saltar la limitació en algun moment.»

Si les línies vermelles estan molt clares, fer servir robots en determinades circumstàncies pot estar bé. «És com el robot cirurgià al qual no es deixa accedir a determinades parts perquè podria causar danys. Una operació la governa sempre el cirurgià, no el robot, però aquest pot no tan sols fer més precisos els moviments del metge, sinó també actuar com a mecanisme de seguretat». De la mateixa manera, el robot serà un soldat forçosament disciplinat que obehirà sempre les ordres. Però, en moltes ocasions, se'l deixarà que actui

de forma autònoma, perquè no se'l pot controlar en tot moment en el camp de batalla, i algunes decisions les haurà de prendre en fraccions de segon. Amb quin percentatge de seguretat se'l deixa actuar? Pot disparar a un grup de persones si pensa que són terroristes amb un percentatge del 90%? O ha de ser del 99%? O n'hi ha prou amb un 80%? I si la concentració és simplement de gent que està celebrant alguna cosa?

En robòtica assistencial, també s'intenta que un robot interpreti moviments i interaccions. Pot captar ironies, dobles sentits? «És difícil», diu.

Tampoc no està d'acord que es promoguin els robots com a companys de jocs dels infants. Es pot acceptar en alguns casos, quan els beneficis són evidents. És el cas de CASPER, el robot amb nom de fantasma creat per l'Hospital de Sant Joan de Déu de Barcelona i l'IDEAI-UPC perquè sigui el company d'habitació i redueixi el patiment dels infants internats a centres sanitaris.

Però això no té res a veure amb els amics electrònics, que se suposa que tenen qualitats humanes. Torras no creu que sigui positiu per als nens i nenes. Amb la gent gran, no és que hagi de substituir l'escalfor humana, però sí que hi pot ajudar molt.

Perquè, si deixem de banda el tracte personal directe, correm perill que s'extingeixin els sentiments, que necessitem «massatgistes emocionals» per recuperar sensacions i dependre de robots personals per viure el dia a dia. Aquest escenari extrem és el que planteja Carme Torras a *La mutació sentimental*. Ho llegeixo a

la contracoberta de la versió en anglès, que, com hem dit al principi d'aquest apartat, porta el títol significatiu de *The Vestigial Heart*.

La paraula *vestigial* no apareix al diccionari de l'Institut d'Estudis Catalans. Però hi ha dues accepcions de vestigi i la segona diu així: «Part o òrgan petit, degenerat, imperfectament desenvolupat, que ha estat ben desenvolupat en un estat anterior de l'individu o en una generació anterior.» Sí que apareix com a terme d'anatomia al TERMCAT: «Òrgan poc desenvolupat que, en avantpassats remots o en la vida embrionària, havia estat, anatòmicament i funcionalment, de desenvolupament normal.»

Així, el cor vestigial faria referència al cor, no en el sentit del múscul que batega, sinó, metafòricament, del lloc on neixen o es guarden els sentiments. A la novel·la, una noia de tretze anys es desperta després d'un període en què ha estat conservada criogènicament i es troba amb aquesta nova societat. Això permet deixar anar la fantasia, però, sobretot, plantejar molts dilemes ètics.

Així ho han entès diversos professors, com Ulises Cortés, que la fa servir al màster en IA. Però als Estats Units aquests temes ètics ja estan inclosos al currículum d'enginyeria des de 1991. Durant anys, per al debat es feien servir textos filosòfics. Però la temàtica va tenir molt més èxit quan Judith Goldsmith, professora a la Universitat de Kentucky, hi va introduir textos de ciència-ficció, d'autors com Isaac Asimov o Ray Bradbury. L'assignatura que impartia es deia «*Science*

*Fiction and Computer Ethics*». I és coautora del llibre *Computing and Technology Ethics. Engaging through Science Fiction*, publicat pel MIT el 2023

En alguns congressos, Torras parlava de la seva novel·la. I va cridar l'atenció d'una editora de MIT Press de Boston. Li va proposar d'estudiar publicar-lo si elaborava materials per a ús didàctic. I després d'un llarg procés, en va aparèixer la versió anglesa, amb un apèndix com a guia de lectura i amb propostes de debat. Aquest apèndix també està ara disponible en català a la web de Pagès Editors.<sup>25</sup>

La versió anglesa es fa servir també en alguns països d'Europa i a Austràlia.

La guia proposa la lectura de capítols concrets i planteja moltes preguntes i possibles discussions, però, tot i que és força completa, es tracta només d'un tast, perquè les idees per fer debats generades per la narració no s'acaben.

Pel que fa a la ciència-ficció, no és un tema nou a la UPC. El malaguanyat Miquel Barceló va promoure el premi que porta el nom de la universitat i l'edició de diverses obres. I Jordi José i Manuel Moreno no només han fet servir la ciència-ficció en les seves classes, sinó que són autors del llibre *De King Kong a Einstein: la física en la ciència-ficció* (Edicions UPC, 1999) i han impartit cursos com «Tecnociència i Ciència-Ficció: de King Kong a Einstein», juntament amb Marc Alier i Miquel Barceló.

---

25 Vegeu: <https://www.pageseditors.cat/ca/guia-didactica-la-mutacio-sentimental.html>

Les obres de ciència-ficció que inspiren reflexions ètiques i socials són moltes i cada lector tindrà les seves. Ara em ve al cap *Podem fabricar-te*, publicada per Philip K. Dick l'any 1972, en què una empresa d'òrgans elèctrics es recicla i passa a fabricar una rèplica robòtica d'Abraham Lincoln. O la que em comenta Jordi José, *Gent de fang* (*Kiln People*), obra del 2002 del físic i astrònom David Brin, en què la gent pot produir còpies de si mateixos per dur terme, temporalment, determinades tasques. Abans d'acabar la seva existència, aquestes còpies descarreguen les seves memòries en els seus propietaris de carn i ossos.

Manuel Moreno n'aporta una de més recent, *Màquines com jo* (Anagrama, 2019), d'Ian McEwan, l'escriptor britànic que ha dedicat més atenció a la ciència, com es veu en altres obres seves: *Amor perdurable*, *Dissabte*, *Solar*, *La llei del menor...* A *Màquines com jo*, planteja el debat de si la IA robòtica pot arribar a entendre la complexitat de les relacions humanes i, fins i tot, crea un triangle peculiar: un home, una dona i un robot. A la primera pàgina, ja trobem una frase que ens colpeja: «Tan bon punt va ser factible, no vam tenir més remei que seguir els nostres desitjos, fossin quines fossin les conseqüències.» Una manera potser usual, però perillosa, d'actuar amb la tecnologia, avançant simplement perquè es pot avançar.

Això em recorda *La Klara i el Sol* (Anagrama, 2021), del premi Nobel Kazuo Ishiguro, autor de la coneguda *No em deixis mai*. La Klara del títol és una AA, una amiga artificial, especialitzada a acompanyar

adolescents. I, mentre espera tenir un nou usuari, es dedica a observar el món exterior i a fer-se preguntes sobre els humans.

Moreno també esmenta *Remake*, publicada per Connie Willis el 1995 (en castellà, a Nova, el 1997). És una novel·la en què els actors reals de cine han estat substituïts per simulacres generats per ordinador i l'espectador pot canviar el final del film prement una tecla. Una narració ben actual, prop de trenta anys després, que molts guionistes i actors de Hollywood deuen tenir a la memòria.

I, ja que hem entrat en el món del cinema, Manuel Moreno cita unes quantes pel·lícules. Per ordre cronològic, la primera és *Sleep Dealer* (2008), dirigida per Alex Rivera, en què treballadors de classe baixa connecten els seus sistemes nerviosos a una xarxa informàtica global: «Connecta amb el teu destí.»

N'hi ha dues del 2015: *Ex Machina*, on es mostra una IA-ginoide –un terme que utilitza en Moreno, però que no trobo al TERMCAT ni al DIEC, que es fa servir per diferenciar-lo d'*androide*, que seria masculí– per a l'ús personal d'un excèntric mili-onari. En destaca una de les frases: «No hi ha res més humà que la voluntat de sobreviure.» Del mateix any és *Chappie*, dirigida per Neill Blomkamp, ambientada a Johannesburg, on la policia decideix comprar robots blindats per reduir l'alt índex de criminalitat a la ciutat. Però el seu inventor crea un prototip d'IA capaç de sentir emocions, que és segrestat per uns gàngsters.

A la sèrie *Westworld* de la cadena HBO, de l'any 2016, s'hi veuen ànimes de metall en parcs d'atraccions per a humans i com es rebel·len. I, finalment, a *The Creator*, dirigida per Gareth Edwards el 2023, «per un cop les IA són els bons i els humans, que les volen exterminar, els dolents».

Karina Gibert també m'havia comentat una pel·lícula, *Un amic per a en Frank*, dirigida per Jake Schreier, on –atenció, espòiler o, com insisteix Màrius Serra que s'ha de dir en català, esguerro–, un assistent robòtic d'un home gran arriba a tenir tanta empatia amb el seu usuari que acaba ajudant-lo a cometre robatoris.

Jo hi afegeixo, per suggeriment de Carme Torras, *Surrogates*, estrenada el 2009 sota la direcció de Jonathan Mostow i protagonitzada per Bruce Willis. Dibuixa una societat en què la gent viu per control remot des de la seguretat de casa seva a través de robots substituïts.

I no podem oblidar la tan coneguda *Her*, dirigida per Spike Jonze i estrenada el 2013, en què un usuari, interpretat per Joaquin Phoenix, s'enamora d'un sistema operatiu amb veu femenina.

Més enllà de la ciència-ficció estricta, quan entro amb la Carme Torras a l'IRI veig unes quantes obres divulgatives per als més menuts. Així, la subdirectora de l'IRI, Júlia Borràs, ha escrit i il·lustrat *La Berta i el robot Rob-ert*. I Anaïs Garrell, investigadora de l'IRI, ha creat *Els robo-científics*, una sèrie de contes escrits amb Eva Garrell i Núria Xalabardé, amb il·lustracions d'Anna Xalabardé, adreçats a escolars d'entre 4 i 8 anys,

que tracten de temes com els robots educatius, voladors i humanoides, i els exoesquelets.

Totes aquestes són obres que serveixen tant per fer reflexionar els futurs enginyers sobre temes ètics com perquè els petits coneguin què són exactament els robots i quines utilitzats i límits tenen.

### **Caminant amb robots**

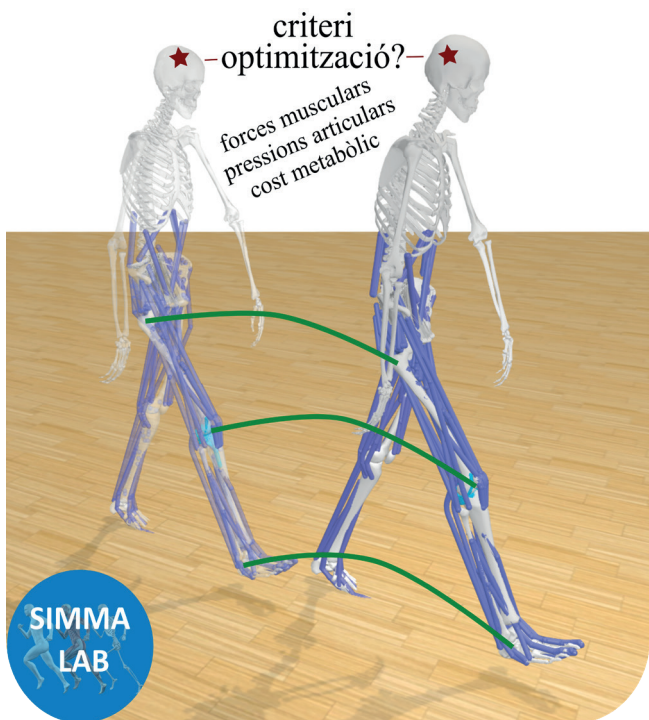
La recerca en robots a la UPC també avança en altres departaments. Gil Serrancolí és professor agregat i investigador de l'Escola d'Enginyeria de Barcelona Est (EEBE), al campus del Besòs. Treballa en biomecànica computacional i m'explica que l'estudi del moviment humà és molt rellevant en diferents aplicacions mèdiques. I que, per a això, cal modelitzar la persona humana: «Consisteix a representar matemàticament no només les propietats antropomètriques, si no els processos fisiològics que la formen, com per exemple els models musculars que relacionen el senyal neuronal amb la força muscular, o els models de contacte entre elements del propi cos o amb l'exterior.»

Davant de l'ordinador, comenta que «en els models musculars, un model matemàtic descriu el procés en què l'actina i la miosina provoquen la força de les fibres musculars; geomètricament, es marca on comença i acaba el múscul, per tal que aquest actuï sobre l'esquelet allà on toca i exerceixi la força que el farà moure».

Per portar-ho a terme, consideren una sèrie de sòlids rígids que interaccionen entre ells, com els ossos de l'esquelet, conduïts per aquests músculs. M'explica que al seu laboratori, el Simulation and Movement Analysis Lab (SIMMA Lab) treballen per desenvolupar simulacions musculoesquelètiques per tal de predir el moviment simultàniament amb pressions articulars, forces musculars i forces de reacció amb l'exterior. Això permet, per exemple, buscar la millor estratègia per anar d'un punt a un altre de forma òptima.

El model que m'ensenya defineix 94 músculs que es mouen tots alhora amb 37 graus de llibertat cadascun. Acaba essent un model amb moltes variables i s'hi ha de posar restriccions, ja siguin matemàtiques o físiques. Per exemple, que el múscul no passi d'aquest punt o que no faci aquest moviment. El problema matemàtic que cal resoldre acaba essent d'una dimensió molt gran, i per això es busquen mètodes, com la diferenciació automàtica, que permet resoldre'l fins a vint vegades més de pressa que amb els mètodes de diferenciació tradicionals.

Per predir un cicle de la marxa, per optimitzar el trajecte d'un punt a un altre, abans els càlculs trigaven tot un dia i ara es poden fer amb mitja hora, gràcies també als nous mitjans computacionals. I això per calcular només un cicle d'un segon i mig, en què cal tenir en compte les forces musculars, el contacte amb el terra, la interacció dels ossos...



*Figura 10. Captura del model SIMMA*

Les aplicacions de les simulacions musculoesquelètiques són diverses i permeten analitzar com canvia el moviment segons diferents criteris que pot seguir el nostre sistema nerviós central, com ara minimitzar el cost metabòlic, les forces musculars o les forces de contacte articulars per evitar el dolor.

També poden servir per estudiar com afecta el patró de les forces musculars en les pressions articulars i predir com serà el moviment després d'una cirurgia. O simular la pressió a diferents punts del genoll per tal d'estudiar l'osteoartritis, que és el treball que fa actualment en col·laboració amb l'Hospital del Mar.

Una altra aplicació és la millora de tècniques d'esportistes, com ara l'esprint, formulant simulacions per trobar virtualment millores en la tècnica i en els moviments i després aplicar-les a la realitat.

També poden servir per estudiar com millorar la rehabilitació després d'un ictus, cosa que es porta a terme en el marc del projecte Muvity.<sup>26</sup> Muvity és una aplicació computacional que conté alguns exercicis i jocs seriosos de rehabilitació motriu en què l'usuari interacciona amb el seu propi moviment corporal capturat per una càmera de profunditat i un algorisme amb IA. Se n'està fent un assaig clínic amb el Consorci Sanitari de Vic i els CAP d'Osona.

Anem al laboratori de captura del moviment i allà un investigador de l'equip, Mohanad Harba, camina amb uns marcadors reflectants sobre uns rectangles que, segons em comenta, són plaques de força. Té un marcador a cada genoll i un a cada turmell. La càmera detecta els marcadors reflectants i així obté el moviment de cada punt en tres dimensions. Les dades serveixen per validar els models biomecànics.

També es fan proves amb pacients que tenen problemes de genoll o que ja han estat operats, per tal

---

<sup>26</sup> Vegeu: <http://muvity.cat/>

d'analitzar els seus moviments i les causes que el provoquen. La captura del moviment que es fa en l'àmbit de la biomecànica és similar al sistema que s'utilitza per als moviments realistes dels videojocs.

I arribem als robots. Les simulacions del moviment basades en l'optimització de les trajectòries també tenen aplicacions en robòtica. Les simulacions poden servir per obtenir moviments més realistes de robots humanoides, o també podrien servir per optimitzar trajectòries repetitives que s'han de fer centenars de vegades al dia, per disminuir l'energia consumida o el desgast de les peces.

D'altra banda, també serveixen per ajudar a dissenyar el control d'un robot d'assistència, per calcular el grau d'assistència que es necessita a cada moment.

## **El futur i les pors**

Com ja hem dit repetidament, algunes pors o reticències o dubtes sobre la IA i els robots són no només lògics, sinó molt necessaris. D'altres més aviat poden desviar l'atenció dels problemes reals. És el cas de la possible rebel·lió dels robots, que es farien amos de la societat i sotmetrien els humans.

Si se'n fan amos, necessitaran perpetuar-se. En aquest cas, trobem les màquines autoreplicadores, que teòricament serien capaces de fabricar, per si soles, una còpia de si mateixes amb primeres matèries preses de l'ambient que les envolta. Entre altres científics que

han exposat aquesta idea, potser el més famós és John Von Neumann als anys cinquanta, tot i que Freeman J. Dyson també ha estat molt actiu en aquest camp. I, més recentment, n'ha parlat Kim Eric Drexler en un llibre sobre nanotecnologia.

La idea és atractiva, però una màquina així hauria de poder obtenir la matèria primera i l'energia necessàries per construir-ne una rèplica, amb tots els seus components. Això s'investiga en algunes universitats, com la de York, on hi ha el projecte *Autonomous Robot Evolution: Cradle to Grave*.<sup>27</sup>

Aquí no es tracta només que les màquines es puguin replicar elles mateixes, sinó que es produeixi una mena de selecció natural. Basant-se en el darwinisme, la teoria diu que les màquines amb més èxit i més ben adaptades es reproduirien més i les seves característiques passarien a generacions futures. És la robòtica evolutiva, molt útil per explorar altres astres sense tenir molts dels problemes dels humans en viatges llargs i en entorns molt agressius –segurament en tindrien d'altres.

Tornem a la ciència-ficció, perquè la idea ja la va plantejar Samuel Butler a la seva novel·la *Erewhon*, publicada anònimament l'any 1872. El títol és la paraula *nowhere* («enlloc») escrita a l'inrevés –amb l'únic canvi des dues lletres, potser per donar-li una altra sonoritat– i, per tant, juga amb la paraula *utopia*, que significa el mateix.

---

27 Vegeu: <https://www.york.ac.uk/robot-lab/are/>

Butler critica diversos aspectes de la societat victoriana, entre ells la industrialització. I en això uneix l'evolucionisme biològic i les màquines. L'autor anglès ja havia escrit el 1863 un assaig titulat *Darwin entre les màquines*, on proposava que aquestes podrien evolucionar tal com ho feien els organismes vius.

*Erewhon* descriu el país imaginari que dona nom a la novel·la. Allà hi arriba el narrador, un pastor d'ovelles. Un cop a Erewhon, és empresonat perquè du un rellotge i en aquell país totes les màquines estan prohibides, fita que es va aconseguir després d'una revolució sagnant.

Als capítols 23 a 25, el protagonista resumeix l'anomenat *Llibre de les màquines*, on es feia una crida a destruir tota la maquinària existent. L'enfrontament entre maquinistes i antimaquinistes va acabar amb la victòria d'aquests darrers, que van exercir una repressió que va impedir qualsevol tipus d'oposició. L'autor del llibre resumeix pel narrador expressa el temor que les màquines evolucionin fins a exercir una tirania encara més gran sobre l'espècie humana. Fins i tot suggereix que totes les màquines provenen d'un ancestre comú i que, per tant, han evolucionat.

Són pors antigues, que ara sorgeixen de nou en contextos socials i tecnològics molt diferents. Potser ens hem acostumat a pensar que amb la ciència i la tècnica tot és possible. I és cert que tot el que ha aparegut als darrers anys i el que es creu que vindrà aviat ens obliga a ser prudents, abans de jutjar la impossibilitat de determinats avenços. Però, de moment, les

màquines d'IA i els robots tenen moltes limitacions, com la incapacitat, ara com ara, de tenir uns sentits com els humans.

Acabem amb unes paraules de la doctora Karina Gibert: «Potser integrant sensors més precisos i potents i dades et surt un androide amb molta capacitat, encara que pesi 300 quilos. Però hi poden haver detalls que modifiquin la visió del món i el robot no els té. L'humà raona amb una gran quantitat de coneixement implícit que no sap que té i que no verbalitza. Ve de l'experiència, de l'aprenentatge implícit, coses que no ha après als llibres i que li permeten interpretar una situació.»



## L'art difícil de predir

Al seu blog «Calaix Desastre», Marc Alier exposava el 10 de juny de 2023 «10 prediccions sobre IA i internet que comparteixo». Es tractava d'un resum de l'article «The Tech Revolution: 10 Predictions about AI, Internet, and Possibly Our World in 3 Years», publicat a la web del tecnòleg i emprenedor David Poblador i Garcia, director tècnic de Secways, a la seva web «Keep it boring, for now» el 19 de maig.

Aquí en resumim el resum i compararem la llista amb altres prediccions. La primera és que els models de llenguatge grans (LLM) es convertiran en la interfície definitiva per a la interacció amb el coneixement i els motors de cerca convencionals quedaran gairebé obsolets. A més, la majoria dels LLM funcionaran sense connexió –no caldrà estat connectat sempre a internet.

La tercera és que els models d'IA generativa es convertiran en especialistes i proporcionaran coneixements profunds en dominis específics. En comptes del model de càlcul client-servidor tradicional, es faran servir dispositius més potents i LLM més avançats.

Un altre canvi serà que la indústria multimilionària de la publicitat *online* haurà de trobar la manera de passar a l'*offline* o arriscar-se a quedar obsoleta.

Les tasques més complexes, com les reserves d'hoteles o de restaurants, o l'organització d'enviaments, es faran interactuant amb API millorades que col·laboraran amb els models d'IA que s'executin localment. També emergirà un mercat extens per a les integracions de la IA.

Els llocs web corporatius perdran importància i internet es transformarà en un espai més personal per a l'expressió. Els fabricants de dispositius i els constructors de sistemes operatius es trobaran en la primera línia d'aquesta nova batalla. Finalment, els creadors de continguts trobaran noves vies a l'era de la IA, enriquint els LLM amb els seus coneixements especialitzats.

I tot això pot passar, segons Poblador, abans de tres anys.

Pel llenguatge i per cap a on s'orienten les prediccions, diria que estan destinades als entesos –almenys, als una mica entesos. Ara reproduiré les prediccions de Toby Walsh, investigador en cap de l'Institut d'IA de la Universitat de Nova Gal·les del

Sud, a Sidney. Les extrec del seu llibre *Machines That Think* (Prometheus Books, 2018), que em va deixar Ulises Cortés. És un llibre molt recomanable, força divulgatiu i clar, per conèixer l'origen i l'evolució de la IA, què és, què no és, els seus pros i contres, les seves possibilitats i limitacions, l'impacte social, els beneficis, els riscos...

Walsh és prudent i les seves prediccions són per al 2050. De fet, al seu darrer llibre, titulat *Faking It: Artificial Intelligence in a Human World* (La Trobe University Press, 2023), alerta que l'obra que el lector té a les mans ja està antiquada a penes acabada de sortir, perquè la IA avança amb tanta rapidesa que pot garantir que, quan la llegim, ja n'hauran aparegut noves aplicacions.

Fixem-nos en els deu grans impactes de la IA que Walsh prediu per a la meitat de segle. El primer és que els humans tindran prohibit conduir. És a dir, no només s'hauran superat tots els problemes dels cotxes autònoms, sinó que els humans no podran agafar el vehicle. Suposo que hi haurà alguns llocs tancats o delimitats per conduir.

El segon és que els dispositius controlaran la salut contínuament. L'ordinador serà el metge i detectarà els signes primerencs de tot.

La tercera predicció és que hi haurà pel·lícules hiperrealistes, amb elements impossibles a la vida real. Hi haurà una nova realitat per escapar de la realitat.

La següent afecta les empreses i els departaments de recursos humans: un ordinador et contractarà, et con-

trolarà i et farà fora. D'entrada, és possible que faci fora els propis responsables de recursos humans, si creu que ja no són necessaris.

A més, haurem d'anar molt amb compte, perquè donar ordres orals a tots els dispositius domèstics, al cotxe, al despatx, etc., pot ser una amenaça a la privacitat. Ja no patirem només per les dades que donem volent o sense voler, sinó per les ordres orals i, fins i tot, pels comentaris que fem. De fet, a Florida, la policia ja ha fet servir Alexa com a testimoni per investigar algun sospitós d'assassinat.

Parlant de la policia, un robot robarà un banc. I no només això, sinó que, com podíem suposar, el ciberdelicte serà més sofisticat i d'alta tecnologia –crec que això ja s'està veient i que abans del 2050 haurà fet un gran salt endavant.

Atenció als aficionats a l'esport i, més concretament, al futbol: Alemanya perdrà davant d'un equip de robots en un partit d'exhibició. A més, els científics de dades estaran entre els professionals més ben pagats pels clubs, que no només buscaran nous jugadors joves, sinó també joves informàtics brillants. Hi haurà una Masia per a informàtics?

Els pilots, els capitans de vaixell i els conductors de tren ja no seran humans, sinó que estaran automatitzats. I els noticiaris –o, almenys, una bona part– es faran sense humans. El mercat laboral no pinta gaire bé, encara que sorgeixin noves feines.

Tot i així, hi haurà ganes de no deixar del tot aquest món i estarem vius després de la mort, fent referència

a la identitat digital *post mortem*, que ja hem exposat al capítol 4.

És ben coneguda la frase –sovint atribuïda al físic Niels Bohr, però també a altres personatges– que diu que «fer prediccions és molt difícil, sobretot quan es tracta del futur». Això no impedeix que les prediccions tinguin valor –segons qui les fa, és clar. Però no tant per encertar què passarà d'aquí a tres anys o al 2050, sinó perquè la manera com veiem el futur revela molt sobre el present. L'important no és tant si d'aquí a menys de tres dècades les màquines ens deixaran sense la possibilitat de conduir, si tindrem un ordinador de capçalera en comptes d'un metge de capçalera o si els robots atracaran bancs –no diu si els policies que els perseguiran seran també robots. Tot això es presenta com un futur no només possible, sinó fins i tot plausible. No són extrapolacions fantasioses sense solta ni volta, sinó coses que els tecnòlegs poden suposar amb un marge d'incert.

Però, com veiem, totes aquestes prediccions porten a qüestionar el paper dels humans. És cert que moltes d'elles poden comportar grans beneficis, però quin paper farà la humanitat en un món tan maquinitzat?

Per això, l'important no és fer prediccions per endevinar el futur, sinó per preveure les diverses formes com podria ser i, sobretot, per preparar-nos i triar la que ens convingui més. Al cap i a la fi, són humans els qui dissenyen les màquines, els qui investiguen, els qui assagen les aplicacions. I cal saber cap a on orientar aquests treballs i, si cal, dir «prou; aturem-nos, repen-

sem-ho tot», o «no prenguem aquest camí», o «potenciem aquesta altra via».

O deixarem que sigui la IA que ens expliqui com serà el futur o com hauria de ser? No podem renunciar a tenir nosaltres la decisió final. Decidir el futur, o almenys intentar-ho, és o ha de ser un privilegi totalment humà.

## Final (provisional) de trajecte

L'any 1961, el francès Raymond Queneau va escriure cent bilions de poemes. I ho va fer sense el suport de cap màquina, excepte, probablement, d'unes tisores.

Un any abans, Queneau i el matemàtic François Le Lionnais havien creat OuLiPo (Ouvroir de littérature potentielle o «Taller de literatura potencial»), un grup dedicat a l'experimentació literària. Alguns dels seus membres van fundar, el 1981, l'ALAMO (Atelier de Littérature Assistée par la Mathématique et les Ordinateurs). Però aquell 1961 potser era aviat per tenir un ordinador i un programa que permetessin crear tants poemes. Tot just feia dos anys que Arthur Samuel havia encunyat el concepte de *machine learning* i elaborat un sistema informàtic capaç de jugar a dames i d'aprendre jugant partides contra ell mateix. I aquell mateix 1961 havia aparegut l'*Ultimate*, el primer robot industrial.

Els cent bilions de poemes de Queneau eren sonets, una composició que té catorze versos. L'autor francès va escriure deu versions diferents de cadascun dels catorze versos, tenint en compte la rima –en aquest cas, ABAB ABAB CCD EED–, i els va retallar de manera que el lector pogués agafar-ne qualsevol dels deu a cada línia de vers per combinar-los com volgués. Com a resultat, es podien compondre  $10^{14}$  poemes, és dir, cent bilions.

Una de les primeres versions de l'obra va ser programada aquella mateixa dècada per Dimitri Starynkevitch amb un ordinador CAB500. Utilitzava el llenguatge PAF (programació automàtica de fórmules), creat per ell mateix entre 1957 i 1959 quan treballava a la Société d'Électronique et d'Automatisme (SEA), una petita però molt activa empresa francesa d'ordinadors.

Ja ben entrat el segle XXI, l'obra de Queneau fou traduïda al català per Ester Xargay i Carles Hac Mor, i penjada a internet amb un programa interactiu desenvolupat per Eugenio Tisseli.<sup>28</sup>

OuLiPo i tot el que va generar van ser molt actius aprofitant la informàtica, com expliquen Camille Bloomfield i Hélène Campaignolle-Catel en un article publicat el 2016: «Machines littéraires, machines numériques: l'Oulipo et l'informatique». Es pot trobar a l'arxiu obert hal.science.<sup>29</sup>

---

28 Vegeu: <https://www.motorhueso.net/queneau/index.htm>

29 Vegeu: <https://hal.science/hal-03184437v1/document>

Però, per als seus cent bilions de poemes, Queneau va crear un artefacte manual amb papers retallats – una combinatòria poc tecnològica i força casolana. Recorda el que segles enrere havia inventat Ramon Llull, amb la seva *Ars Magna*, que va ser probablement l'objecte de la burla de Jonathan Swift en un dels viatges de Gulliver.

Dos anys abans de l'obra de Queneau, el matemàtic, informàtic i filòsof alemany Theo Lutz va elaborar el que va anomenar “Textos estocàstics”. Va utilitzar un *Zuse Z22*, un dels primers ordinadors comercials, i hi va introduir 16 substantius i 16 predicats, extrets de la novel·la *El castell*, de Franz Kafka. Hi va afegir paraules que servissin de lligam i punts per marcar el final d'una frase, i va programar l'ordinador perquè elaborés frases triant aquests elements. Calculava que els podria combinar i escriure'n més de quatre milions, però es va conformar amb 50. El primer deia així:

No tota mirada és propera. Cap poble fa tard.  
Un castell és lliure i tota granja és llunyana.  
Tot estranger és llunyà. Un dia fa tard.

Després varen arribar altres experiments. Al segon volum de les seves memòries, titulat *Juventut* (2002), el premi Nobel sud-africà J. M. Coetzee s'inventa un protagonista fictici, John, que, com ell mateix, als anys seixanta emigra a Londres i treballa de programador. Quan acaba la feina, aprofita l'ordinador per escriure versos a l'estil de Pablo Neruda.

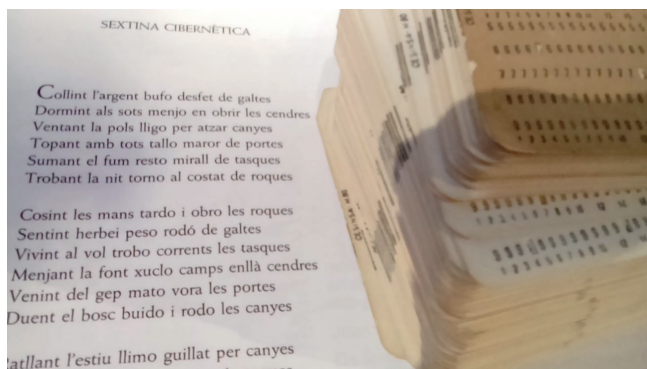
Aquestes proves, com els poemes que ara es demanen a ChatGPT, no devien generar res que pogués optar a guanyar el Carles Riba o l'Ausiàs Marc. Però, amb mans expertes i amb un treball acurat, el resultat podria ser ben diferent. Fou el cas de Joan Brossa, sempre donat a experimentar.

Fa pocs anys, seia en una terrassa de Vilafranca del Penedès i l'home que m'acompanyava va treure de la cartera un plec de targetes perforades, d'aquelles que, com recorden els no tan joves, eren necessàries per programar ordinadors. Allò que ara veiem tan rudimentari, amb el sabor de les coses antigues, tot i que no tenia ni quaranta anys, contenia el programa que va servir a Brossa per elaborar sextines.

El tema el vaig descobrir gràcies a una cita en un article de Màrius Serra, que em va donar un primer contacte. Estirant el fil, vaig poder investigar, amb l'ajut de diverses persones, la gènesi de la *Sextina cibernètica*. La història la vaig explicar al número 208 de la revista *Sàpiens*<sup>30</sup> i també va ser objecte d'algunes taules rodones, una de les quals al vestíbul del BSC. La persona que m'ensenyava les targetes era Josep Font, un informàtic ja jubilat, que a final dels anys setanta treballava al centre de càlcul de l'empresa AGMA, per simular composicions i propietats de diversos aliatges, en comptes de provar-los tot al laboratori.

---

30 Vegeu: [https://www.sapiens.cat/epoca-historica/historia-contemporania/computadora-brossa\\_202796\\_102.html](https://www.sapiens.cat/epoca-historica/historia-contemporania/computadora-brossa_202796_102.html)



*Figura 11. Fragment de la Sextina cibernètica i targetes perforades que contenen el programa per elaborar-la. (Font: Xavier Duran)*

L'any 1977, Brossa estava acabant el seu segon llibre de sextines, una composició poètica molt exigent en la mètrica i la rima. N'hi volia incloure una que fos ben diferent de la resta. I se li va acudir que, amb un ordinador, podria elaborar una sextina més «mecànica». Brossa tenia un bon amic, el químic Santiago Farré, que dirigia l'empresa AGMA. I allà dos joves informàtics, Josep Font i Jordi Bastardes, van rebre l'encàrrec d'elaborar un programa que compoqués sextines.

Brossa va triar 50 paraules-rima, que fossin substantius i bisíl·labs, a més de gerundis bisíl·labs, substantius, verbs en primera persona del singular del present d'indicatiu i unes quantes formes gramaticals diferents per servir d'enllaços. Un cop creat el programa, es va introduir tot en un *IBM 1130*. Segons Font i Bastardes, en uns minuts l'ordinador podia escriure

cinc mil sextines, que serien molt repetitives. Per això van proposar de fer-ne vint i, d'aquestes, Brossa en va triar una, que es va publicar al llibre sense cap retoc.

Més endavant qualificaria el poema com una mena de «nen proveta literari» i diria que havia descobert que, d'un aparell electrònic, en podia sorgir «el lirisme més abrandat, les imatges més còsmiques». Les sextines descartades es conserven al Fons Joan Brossa del MACBA.<sup>31</sup>

M'estic allargant i la meva intenció era simplement escriure un epíleg, per no acabar aquest viatge per la UPC tan en sec i de manera més emotiva. Si me n'he anat per les branques poeticoinformàtiques és perquè, primer de tot, se'm va acudir el títol que he posat al principi del capítol, amb la paraula *provisional*, i això em va fer venir al cap un dels jocs que els sempre imaginatius i juganers membres d'OuLiPo es van empecar. No era fàcil entrar a ser-ne membre i, quan se superava l'admissió, encara es podia caure en dues categories: uns eren els socis provisionalment definitius i altres, els socis definitivament provisionals.

Si ho analitzem bé, la primera categoria és un oxímoron clar, però la segona pot ser real: aquests socis sempre seran provisionals: mai de la vida esdevindran definitius.

Per això, vaig pensar que aquest epíleg posava un punt final definitivament provisional. Els bons viatges, com les lectures dels bons llibres, mai no acaben del

---

31 Vegeu: <https://www.macba.cat/ca/aprendre-investigar/arxiu/fons-joan-brossa>

tot. Sempre hi queda un eco, una ressonància que, de forma conscient o inconscient, manté el trajecte viu. No es tracta només de recordar tot allò que hem après o del pòsit que ens ha deixat. És més aviat la sensació que encara ens movem pel territori recorregut, que encara ens arriben idees o sensacions que potser no havíem acabat de copsar.

Això és evident en el cas de la UPC. Cada visita era un estímul, cada conversa feia néixer una idea, cada passa generava ganes de conèixer més i més. Però els trajectes, encara que sigui provisionalment, han d'acabar algun dia. Per molt que ens resistim a baixar del vehicle o a aturar les passes.

Els paisatges fèrtils deixen una forta empremta. Els paisatges inabastables mantenen un ampli territori per explorar. Els paisatges dinàmics canvien contínuament els seus límits i allò que contenen.

A la UPC, s'hi cou el futur. I el viatge no acaba mai, perquè el futur torna a començar cada dia.



## Breu selecció bibliogràfica

Angulo, C.; Veliz, C. *Intel·ligència artificial i tecnoètica* [en línia]. Barcelona: Iniciativa Digital Politècnica. Oficina de Publicacions Acadèmiques Digitals de la UPC, 2022. ISBN: 9788419184054. Disponible a <http://dx.doi.org/10.5821/ebook-9788419184054>

Čapek, Karel. *R.U.R. (Robots Universals Rossum)*. Barcelona: MalesHerbes, 2017. ISBN: 9788494725838

Hernández-Fernández, Antoni (ed.). *Creativitat digital*. Barcelona: Iniciativa Digital Politècnica. Oficina de Publicacions Acadèmiques Digitals de la UPC, 2023. ISBN: 9788410008090. Disponible a <http://dx.doi.org/10.5821/ebook-9788410008090>

López de Mántaras i Badia, Ramon. *100 coses que cal saber sobre intel·ligència artificial*. Valls: Cossetània, 2023. ISBN: 9788413562896

Moreno, Manuel; José, Jordi. *Física i ciència-ficció*. Barcelona: Edicions UPC, 1994. ISBN: 9788483011256

Moreno, Manuel; José, Jordi. *De King Kong a Einstein. La física en la ciencia ficción*. Barcelona: Edicions UPC, 1999. ISBN: 9788483013335

Sanguinetti, Pablo. *Tecnohumanismo. Por un diseño narrativo y estético de la inteligencia artificial*. Madrid: La Huerta Grande, 2023. ISBN: 978-8418657375

Torras, Carme. *La mutació sentimental*. Lleida: Pagès Editors, 2008, ISBN: 9788497796354

Torres, Jordi. *La intel·ligència artificial explicada als humans*. Barcelona: Plataforma Editorial 2023. ISBN: 9788419655585

Walsh, Toby. *Machines that think : the future of artificial intelligence*. Buffalo, NY: Prometheus Books, 2018. ISBN: 9781633883758



En aquesta odissea acadèmica, Xavier Duran ens convida a fer un viatge fascinant pel coneixement de la UPC, en què el lector, com un veritable “UPCnauta”, podrà cercar el seu particular velló d’or: la saviesa. Entre xips i neurones, els camins de la recerca es creuen com fils del destí, revelant un univers propi per a cada aspirant. La màquina al·lucinada de la recerca treballa, mentre els estudiants exploren terres de diagnòstics, fàrmacs i bessons digitals, desvelant la cara positiva d’aquesta epopeia. Converses amb *chatbots* i passejos amb robots s’entrecreuen, com si la robòtica mateix es manifestés en aquesta vall inquietant del món en què vivim i en què no sempre gaudim.

De la realitat a la ficció narrativa, aquesta èpica humanista sobre el futur, les pors i l’art predictor de la ciència aportaran, sens dubte, a qui comenci el viatge, una visió única de part de la recerca que Xavier Duran, com a escriptor resident del 2023, va escollir, convertit en un autèntic “UPCnauta”.



**UPCArts**Diàlegs



iniciativa  
**digital politècnica**  
Publicacions Acadèmiques de la UPC

