Assistents robòtics = (imaginació + tecnologia) × ètica

Carme Torras*
Institut de Robòtica i Informàtica Industrial, CSIC-UPC
torras@iri.upc.edu

Resum: L'envelliment progressiu de la població suposa una càrrega creixent per al sistema de salut i per a les famílies. Per fer-hi front, la robòtica assistencial combina imaginació i tecnologia per proporcionar eines que ajudin els cuidadors i alhora dotin les persones amb limitacions de més autonomia. Després d'un breu apunt sobre els precursors —imaginats o construïts— dels assistents robòtics, exposarem els reptes tecnocientífics que plantegen i com són abordats per la comunitat investigadora, així com les implicacions socials i ètiques que es deriven del desplegament d'aquesta tecnologia.
Paraules clau: assistents robòtics, intel·ligència artificial, interacció persona-robot, formació en tecnoètica, ciència-ficció.

Robotic assistants = (**imagination** + **technology**) × **ethics**
Abstract: The progressive aging of the population represents a growing burden for health systems and families. To address this, assistive robotics combines imagination and technology to provide tools that help caregivers and give more autonomy to people with limitations. After a brief note on the precursors —imagined or constructed— of robotic assistants, the article discusses technoscientific challenges they pose and how these are addressed by the research community, as well as the social and ethical implications arising from the deployment of this technology.
Keywords: robotic assistants, artificial intelligence, human-robot interaction, technoethics training, science fiction.

* Carme Torras és llicenciada en Matemàtiques (UB), màster en Computació (UMass), doctora en Informàtica (UPC) i cap del Grup de Recerca en Percepció i Manipulació Robotitzada de l'Institut de Robòtica i Informàtica Industrial, CSIC-UPC (RobIRI). Ha rebut els premis nacionals de Recerca, tant el català com l'espanyol, per les seves contribucions a la robòtica intel·ligent i social, així com la Creu de Sant Jordi (2024). Compromesa amb la promoció de la tecnoètica, les seves novel·les *La mutació sentimental* i *Enxarxats* —ambdues premis Ictineu— i el recull de relats *Estimades màquines* aborden dilemes ètics suscitats pels assistents robòtics, la IA i les xarxes socials. És vicepresidenta del Comitè d'Ètica del CSIC i membre de l'Observatori d'Ètica en Intel·ligència Artificial de Catalunya. ORCID: 0000-0002-2933-398X.

Introducció

Imaginació i tecnologia mantenen una estreta relació d'inspiració mútua que ha resultat molt fructífera al llarg dels anys. D'una banda, l'enginy tecnològic permet materialitzar els productes de la nostra imaginació. De l'altra, les tecnologies disponibles en una societat modelen i condicionen els límits del que és imaginable. Aquest és el bucle de retroalimentació que ha impulsat l'avenç científic i tecnològic al llarg de la història humana i que resulta particularment diàfan en el cas dels assistents robòtics, en què la fantasia de crear humans artificials ha propiciat el desenvolupament de l'enginyeria robòtica, que al seu torn ha possibilitat apropar-se més i més a la realització d'aquesta fantasia.

En el primer quart del segle xxi, l'envelliment creixent de la població i la manca de personal d'assistència han accelerat aquest bucle de recerca sobre robòtica assistencial, ja que ha estat assenyalada com una tecnologia clau per ajudar els cuidadors en hospitals, centres de rehabilitació i residències de gent gran, així com per empoderar les persones amb mobilitat reduïda a la llar perquè puguin viure de manera autònoma. Aquest és un àmbit altament interdisciplinari, que requereix unir no només coneixements de diverses àrees tecnocientífiques, sinó també de les ciències socials i humanitats. En el desplegament d'aquesta tecnologia, el codisseny i la col·laboració de tots els actors intervinents en el sector de l'assistència i les cures són particularment importants a causa de la gran diversitat de condicionants, tant tecnològics com ètics, a què està subjecta.

Precursors dels assistents robòtics: artesania i ficció

La fascinació pels cossos i les ments artificials es veu reflectida en gairebé totes les cultures i èpoques d'una manera o altra. Per exemple, en la cultura egípcia ja trobem autòmats rudimentaris que utilitzen la força de l'aigua per moure parts d'estàtues i fins i tot un colom mecànic capaç d'aixecar el vol propulsat per raigs d'aire.

Centrant-nos en els autòmats que es poden considerar precursors dels assistents robòtics, al Japó, a principis del període Edo, van proliferar els *karakuri*, unes delicades figures de fusta amb mecanismes que els donaven mobilitat per sorprendre, entretenir i fins i tot servir el te als convidats. També a Europa, amb les idees filosòfiques de la Il·lustració i gràcies als avenços en rellotgeria, els autòmats van estar en auge. Al segle xviii va tenir una gran ressonància el Turc Jugador d'Escacs, de l'inventor hongarès Wolfgang von Kempelen, que movia les peces amb el seu braç extensible i guanyava la majoria de partides contra jugadors il·lustres. Però el Turc tenia truc: un compartiment ocult on s'amagava un nan expert en escacs. A la mateixa època, el suís Pierre Jaquet-Droz va construir tres autòmats molt complexos sense recórrer a cap engany: el Pianista, el Dibuixant i l'Escriptor. El mecanisme d'aquest darrer tenia més de sis mil peces i era capaç d'escriure textos de fins a quaranta paraules utilitzant la ploma, que mullava al tinter i seguia amb la mirada mentre escrivia.

De manera similar als autòmats físics, les rèpliques humanoides van irrompre en obres de ficció especulativa del segle xix, com *L'home de la sorra* d'E. T. A. Hoffmann o *Frankenstein* de Mary Shelley, suscitant emocions contraposades: enamorament, en la primera, i repulsió, en la segona. La paraula *robot* va aparèixer per primera vegada el 1920 a l'obra teatral *R.U.R.* (*Rossum's Universal Robots*), del dramaturg Karel Čapek. En txec, *robota* significa 'treball forçat', i l'obra versa sobre l'empresa R.U.R., que, per alleugerir la feina als obrers, produeix androides que es poden fer passar per humans.

Aquesta visió dels robots com a substituts dels treballadors a les cadenes de producció va marcar fortament el desenvolupament de l'enginyeria robòtica durant la segona meitat del segle xx. Els manipuladors industrials van propiciar el desenvolupament de capacitats d'interacció amb l'entorn i, posteriorment, de col·laboració amb treballadors humans, que han estat la llavor de la interacció multimodal, fluida i personalitzada, tan important en l'atenció a persones vulnerables.

Mentrestant, en el terreny literari s'exploraven els aspectes íntims i morals d'uns robots humanoides allunyats dels braços industrials no només per la seva forma, sinó també pel seu àmbit d'actuació, principalment domèstic i social. Relats d'Isaac Asimov, Ray Bradbury o Philip K. Dick exploren temes com la manca de privacitat, el totalitarisme tecnològic i la manipulació, els substituts emocionals, la responsabilitat moral, la pèrdua del control humà, o problemes existencials relacionats amb el lliure albir, allò que ens fa humans, les creences i la immortalitat. Tota aquesta casuística anticipa el desenvolupament d'una robòtica que sortirà de l'àmbit industrial per entrar en entorns quotidians, on la relació persona-robot esdevindrà primordial i plantejarà multitud de preguntes ètiques, psicològiques i socials. En aquest context, Isaac Asimov va formular les seves famoses tres lleis de la robòtica, que han resultat premonitòries:

1. Un robot no pot malmetre un ésser humà o, per inacció, permetre que un ésser humà pateixi mal.

2. Un robot ha d'obeir les ordres dels éssers humans, llevat que entrin en conflicte amb la primera llei.

3. Un robot ha de protegir la seva pròpia existència sempre que aquesta protecció no entri en conflicte amb la primera o la segona llei.

S'hi anticipaven no només les innovacions científiques i tecnològiques, sinó també les implicacions ètiques del desplegament actual de la robòtica assistencial.

Reptes tecnocientífics de la robòtica assistencial

Al tombant del segle xxi, el camp potencial d'aplicació dels robots s'amplia: ja no estan limitats a les fàbriques, sinó que s'estenen progressivament als dominis urbans, socials i d'assistència. Aquest canvi de la robòtica industrial cap a la robòtica social planteja nous desafiaments per a la investigació tecnocientífica, que descriurem a continuació.

Repte 1. Facilitat d'instrucció per persones sense experiència tecnològica

Perquè persones no expertes en programació —com ara cuidadors, pacients i els seus familiars— puguin instruir els robots, cal proporcionar-los formes molt naturals de fer-ho. L'aprenentatge mitjançant demostracions (Ravichandar *et al.*, 2020) ha emergit com una eina poderosa que permet als robots adquirir les habilitats desitjades a partir de demostracions dutes a terme per humans. Aquest aprenentatge pot requerir que la persona es posi uns sensors que captin el seu moviment i el transmetin al robot, o bé el guiatge cinestèsic del robot perquè segueixi una determinada trajectòria aplicant un perfil de força variable segons la situació, com cal en la tasca de fer llits, per exemple.

Repte 2. Personalització segons les preferències i limitacions dels usuaris

Aquesta personalització pot ser relativa a la interacció física o cognitiva. Per exemple, en una aplicació per posar una jaqueta, es van contemplar preferències tant físiques (p. e., velocitat del robot) com comunicatives (p. e., verbositat), així com limitacions de mobilitat de l'usuari (Canal *et al.*, 2021). Un altre aspecte de la personalització és adaptar el repertori d'habilitats amb què el robot està equipat, a les necessitats i gustos de cada usuari. Per exemple, en la tasca de donar menjar, el robot pot venir de fàbrica amb una trajectòria programada poc ergonòmica per a un usuari en concret, i el cuidador pot personalitzar-la mitjançant guiatge cinestèsic perquè li resulti més còmoda.

Repte 3. Seguretat intrínseca en la interacció amb persones

Aquest és un requeriment clau per als robots assistencials, que fa referència tant a la privacitat de les dades i la fiabilitat del programari, com a la prevenció de qualsevol dany físic que pugui causar un robot. A diferència dels entorns industrials, no és possible engabiar els robots en espais sensoritzats on, si algú entra inadvertidament, tot s'atura, sinó que les mesures de seguretat han d'estar en el mateix controlador del robot i depenen dels sensors amb què estigui equipat. S'acostuma a distingir tres capes de seguretat: la de col·laboració activa, la de coexistència i la reactiva. A la capa superior cal distingir entre una col·lisió accidental i un contacte intencionat d'un usuari humà per guiar el robot o demanar-ne la col·laboració en la realització d'una tasca conjunta. La coexistència d'humans i robots en un mateix espai requereix la monitorització de l'àrea de treball i mètodes eficients de prevenció de col·lisions basats en la informació proporcionada pels sensors. Tot i aquestes mesures de control, poden produir-se contactes físics inevitables a qualsevol part del cos del robot, que han de ser detectats de forma fiable i desencadenar una reacció adequada del robot.

Repte 4. Percepció i manipulació d'objectes deformables

A diferència de les cadenes robotitzades de fabricació, en què els objectes que s'han de manipular acostumen a ser rígids, en els entorns assistencials proliferen els materials deformables, com ara roba, cables, líquids o les persones mateixes. Mentre que l'estat d'un objecte rígid està determinat per tres paràmetres de posició i tres d'orientació, un objecte flexible pot canviar la seva forma i donar lloc a un espai d'estats de dimensió potencialment infinita. Aquest enorme salt dimensional fa que les tècniques desenvolupades per a objectes rígids no siguin escalables a entitats deformables i calgui explorar representacions dels macroestats de deformació i combinar-les amb algoritmes d'aprenentatge. També s'han hagut de considerar noves accions, com el seguiment de vores en tèxtils, que ha dut al disseny de mans robòtiques més versàtils que la clàssica pinça habitual en la indústria, així com a establir taxonomies de manipulació (Blanco-Mulero *et al.*, 2025) que permetin la planificació automàtica de seqüències d'accions conduents, per exemple, a plegar una peça de roba.

Repte 5. Col·laboració amb l'usuari en la realització de tasques

Per a una col·laboració persona-robot agradable i eficaç, calen interfícies amigables. Sovint aquestes inter-

fícies són multimodals, ja que combinen text, veu, imatges, gestos, postures, contacte i/o forces captades mitjançant sensors que poden no correspondre's exactament amb les modalitats de sensació humanes (per exemple, sensors de profunditat). La interacció pot ser merament comunicativa i pot produir-se a distància o bé incloure la manipulació conjunta d'objectes. Un exemple que combina ambdós tipus d'interacció, la comunicativa i la física, és un prototip per donar menjar. La comunicació és verbal i gestual, mentre que una càmera de color i profunditat i un sensor de força al canell del braç robòtic permeten captar les reaccions físiques de l'usuari.

Repte 6. Capacitat d'adaptació i aprenentatge
La destresa manipulativa humana, especialment en tasques en què la dinàmica té un paper important, no només s'adquireix a partir d'una demostració inicial, sinó que requereix molta exercitació. Aquesta és l'estratègia que s'acostuma a seguir, per exemple, perquè un robot bimanual aprengui a doblegar una camisa. A partir d'una demostració aproximada, el robot s'exercita en la tasca mentre és monitoritzat per una càmera zenital i se n'avalua l'execució mitjançant una funció de cost que mesura com queda de ben doblegada. Aleshores s'aplica un algoritme d'aprenentatge per reforçament que busca el millor moviment de plegatge de roba per part del robot (Colomé i Torras, 2020).

Codisseny i transferència
Perquè els avenços en recerca descrits arribin a ser d'ajuda a cuidadors i pacients és essencial que en la concepció, el desenvolupament i el desplegament dels robots assistencials participin tots els actors intervinents en el sector de l'assistència i les cures. A tall d'exemple, exposaré l'experiència del nostre grup en el codisseny de dos prototips.

El primer s'ha codissenyat amb el Parc Sociosanitari Pere Virgili. El 50 % de les persones que hi ha ingressades necessiten que algú les peixi. El personal sanitari no dona l'abast i cal recórrer a familiars que disposen de poc temps a l'hora de dinar. La direcció i els tècnics d'innovació del centre, el personal sanitari i alguns pacients van aportar valuosos suggeriments en el procés iteratiu de desenvolupament del prototip, com equipar el robot amb un sistema que li permeti canviar de cobert (cullera, forquilla, etc.), donar beguda i netejar el pacient amb un tovalló. A petició dels pacients, també se li va afegir una pantalla amb una cara expressiva, que permet una interacció fluida. El programa pilot amb nou persones ingressades i quaranta-cinc sessions ha conclòs amb resultats excel·lents i una gran acceptació per part de cuidadors i pacients.

El segon és un prototip de robot per a l'entrenament cognitiu desenvolupat conjuntament amb la Fundació ACE, dedicada a la investigació, el diagnòstic i el tractament de l'Alzheimer i altres demències. En concret, una neuròloga, un terapeuta i una assistenta social van participar en el disseny i l'experimentació. Es va utilitzar el mateix tipus de tauler amb fitxes que utilitzen al centre de dia, que permet proposar exercicis de dificultat variable. L'objectiu és que cada pacient disposi d'un robot que li proporcioni estímul i ajuda de manera ràpida, i el terapeuta humà supervisi l'execució dels exercicis com si estigués jugant partides simultànies d'escacs. Els resultats del programa pilot van ser molt favorables (Andriella *et al.*, 2023) i l'acceptació per part dels pacients, persones de més de setanta anys amb un cert deteriorament cognitiu, va ser total.

Desplegament ètic: regulació i formació
En l'àmbit assistencial, en què la relació persona-robot té un paper essencial i els usuaris solen pertànyer a col·lectius vulnerables, el desplegament ètic d'aquesta tecnologia és primordial, més encara que en altres contextos (Torras, 2024). Institucions com el Parlament Europeu i associacions professionals com l'Institute of Electrical and Electronic Engineers (IEEE) han aprovat normatives que estan disponibles a Internet.

La regulació és necessària, però encara ho és més la formació. És essencial formar en els beneficis i els riscos de les tecnologies digitals no només l'opinió pública, sinó també, en l'àmbit escolar, tots els alumnes, ja des de la primària i fins a la universitat, i molt especialment els enginyers i informàtics, que dissenyen robots i desenvolupen programari. Associacions tan prestigioses com l'Association for Computing Machinery (ACM), l'IEEE i l'Association for the Advancement of Artificial Intelligence (AAAI) inclouen disset àrees de coneixement en els plans d'estudis dels graus d'Informàtica, entre les quals hi ha «Societat, Ètica i Professionalisme». Algunes matèries d'aquesta àrea s'imparteixen basant-se en textos filosòfics, que de vegades resulten massa teòrics per a estudiants de graus tecnològics i necessiten ser complementats amb indicacions pràctiques per a situacions concretes. Per això, alguns docents han recorregut a la ciència-ficció per exemplificar aquestes situacions i afavorir un debat argumentatiu (Burton *et al.*, 2018).

La ciència-ficció contemporània tracta molts dels temes ètics abordats en aquests cursos. Per exemple, la pel·lícula *Robot and Frank* mostra la relació entre un Frank envellit i el seu cuidador artificial, i pel seu realisme i valor educatiu la plataforma *Teach with Movies* hi ha basat un curs sobre ètica i robòtica. Són nom-

broses les sèries, pel·lícules i novel·les que especulen sobre els beneficis i riscos dels assistents robòtics. Vaig sumar-m'hi amb la novel·la *La mutació sentimental* (Torras, 2008), en què s'explora com ser criat per mainaderes artificials, tenir mestres robòtics o compartir feina i lleure amb programes d'IA pot afectar els hàbits intel·lectuals, emocionals i socials de les generacions futures. El *leitmotiv* de la novel·la és una cita de Robert C. Solomon: «Les relacions que anem construint alhora ens modelen». El filòsof es referia a les relacions humanes amb els nostres pares, mestres i amics, però la frase es pot aplicar també als assistents artificials i a tota mena de dispositius interactius, tan presents actualment a les nostres vides.

La novel·la va ser publicada en anglès per MIT Press, juntament amb una guia en línia per al professorat i una presentació de cent diapositives per impartir un curs sobre «Ètica en robòtica social i intel·ligència artificial», que estan sent utilitzades en diferents universitats d'Amèrica del Nord i Europa. També hi ha una versió en català adaptada a l'educació secundària obligatòria (ESO) i el batxillerat. A tall d'exemple, aquestes són quatre de les preguntes relatives a la interacció persona-robot en contextos assistencials que hi ha a la guia:

— Quan interfereix la presa de decisions automàtica amb la llibertat/dignitat humana?
— Cal excloure per disseny la possibilitat que els robots siguin confosos amb éssers vius?
— És acceptable que els robots actuïn com a substituts emocionals? Si és així, en quins casos?
— Com podem evitar l'aïllament social que es pot derivar de la cura mecànica?

Conclusions i perspectives futures

En haver d'interaccionar amb persones en activitats de la vida diària, els robots assistencials plantegen reptes importants, tant de recerca tecnocientífica com d'humanitats i ciències socials, que constitueixen temes candents amb un elevat potencial per marcar el futur.

Pel que fa a la investigació tecnocientífica, a més del progrés incremental en els diferents àmbits descrits, preveiem dos avenços fonamentals. En primer lloc, la comprensió semàntica de les situacions permetrà un nivell de raonament i explicació argumentada dels motius subjacents a les decisions automàtiques, que actualment està fora de l'abast i limita l'aplicabilitat de tècniques d'aprenentatge com ara l'aprenentatge profund (*deep learning*). El segon avenç que preveiem és la modelització per part del robot de les pròpies capacitats físiques i cognitives, que permetrà no només augmentar significativament el seu grau d'autonomia, sinó també col·laborar més eficaçment amb les persones utilitzant estratègies de control d'autonomia compartida (*shared autonomy*), que avui són objecte d'una intensa recerca.

Aquests avenços sens dubte donaran lloc a assistents robòtics més versàtils, però alhora comportaran nous desafiaments humanístics, ja que indirectament estem modelant les generacions futures per la via de la tecnologia que desenvolupem en el present. Mai les tecnologies disponibles havien estat tan disruptives com ara. Encara que les distopies són útils per fer aflorar els riscos d'un desplegament de la robòtica sense control social, la ciència-ficció potser s'hi ha centrat en excés, presentant-nos un futur inevitablement fosc i descartant qualsevol utopia. Tanta distopia podria ser que acabés en una profecia autocomplerta, així que serà millor que produïm utopies... per si de cas!

Bibliografia

Andriella, A., Torras, C., Abdelnour, C, i Alenyà, G. 2023. «Introducing CARESSER: A framework for in-situ learning robot social assistance from expert knowledge and demonstrations». *User Modeling and User-Adapted Interaction*, 33, pp. 441-496.

Blanco-Mulero, D., Dong, Y., Borràs, J., Pokorny, F. T., i Torras, C. 2025. «T-DOM: A Taxonomy for Robotic Manipulation of Deformable Objects». *arXiv:2412.20998 [cs.RO]* https://doi.org/10.48550/arXiv.2412.20998.

Burton, E., Goldsmith, J., i Mattei, N. 2018. «How to teach computer ethics through science fiction». *Communications of the ACM*, 61(8), pp. 54-64.

Canal, G., Torras, C., i Alenyà, G. 2021. «Are preferences useful for better assistance? A physically assistive robotics user study». *ACM Transactions on Human-Robot Interaction*, 10(4), pp. 1-19.

Colomé, A., i Torras, C. 2020. «Reinforcement learning of bimanual robot skills». *Springer Tracts in Advanced Robotics* (STAR Series, 134).

Ravichandar, H., Polydoros, A. S., Chernova, S., i Billard, A. 2020. «Recent advances in robot learning from demonstration». *Annual Review of Control, Robotics, and Autonomous Systems*, 3, pp. 297-330.

Torras, C. 2008. *La mutació sentimental*. Lleida: Pagès. www.pageseditors.cat/ca/guia-didactica-la-mutacio-sentimental.html. [Traducció a l'anglès: *The vestigial heart: a novel of the robot age*. Cambridge: MIT Press, 2018. www.iri.upc.edu/people/torras/vestigial.html.]

Torras, C. 2024. «Ethics of social robotics: Individual and societal concerns and opportunities». *Annual Review of Control, Robotics, and Autonomous Systems*, 7, pp. 1-18.

Compàs d'amalgama, ISSN 2696-0982 / e-ISSN 2696-1008, abril 2026, núm. 13, pp. 53-57. DOI: 10.1344/Compas2026.13.10
Data de recepció: 11-9-2025. Data d'acceptació: 16-11-2025
Llicència CC BY-NC-ND. © Compàs d'amalgama, 2026