

nodes

El butlletí de l'ACIA

Any 2023. Número 59
PVP: 5 euros




La pandèmia de COVID-19 i la intel·ligència artificial





Edita:

Associació Catalana d'Intel·ligència Artificial (ACIA)

 www.acia.cat

 @acia_cat

 @AssocCatIA

 @acia_cat_rep

Consell editorial:

Vicent Costa (IIIA-CSIC), coord.

Josep Puyol Gruart (IIIA-CSIC)

Cecilio Angulo (IRI-UPC)

Ulises Cortés (UPC)

Emilia López-Iñesta (UV)

Lledó Museros (UJI)

Eloi Puertas (UB)

Marco Schorlemmer (IIIA-CSIC)

Marc Torrens (ESADE-URL)

Revisió ortogràfica:

Núria Altés

Dipòsit legal i ISSN:

GI 1598/2008

ISSN 2171-5602

ISSN 2014-5020 (internet)

Copyright:

Tots els autors identificats a cada article retenen els drets d'autor dels seus treballs. L'ACIA no es fa responsable de les opinions expressades pels seus socis o col·laboradors.

Consell Rector de l'ACIA:

Cecilio Angulo (president)

Lledó Museros (vicepresidenta)

Zoe Falomir (secretària)

Josep Puyol Gruart (tresorer)

Vocals: Francisco Grimaldo

Beatriz López

Jordi Nin

Eloi Puertas

Martí Sánchez-Fibla

Aida Valls

Disseny original:

Toni Casals

Maquetació:

Núria Pardinilla Gutiérrez

Sumari

Editorial	pàgina 3
Ciència ficció	pàgina 5
Estudi cas-control <i>Jordi Font-Agustí</i>	pàgina 5
Filosofia i pensament	pàgina 11
Baixos ètics a IA aplicada al triatge a les UCIs i la traçabilitat de contagis a través de les apps de rastreig <i>Angel Puyol</i>	pàgina 11
Societat	pàgina 15
La naturalesa sociotècnica de la IA: ètica, política i tecnologia <i>Núria Vallès-Peris</i>	pàgina 15
La pandèmia de COVID-19 i la intel·ligència artificial	pàgina 26
Quan la primera gran pandèmia del segle XXI va tornar a posar a prova la societat de la informació en què vivim <i>Hugo A. Ferrer García, Ismael Sanz</i>	pàgina 26
El fracàs de Radar COVID <i>Vicent Costa</i>	pàgina 34
Què en pensen? <i>Josep Puyol Gruart</i>	pàgina 40
Premis Marc Esteve Vivanco 2021 (Emran Saleh Ali Ali)	pàgina 44
2020 (Gerard Canal)	pàgina 47
2019 (Juan Manuel Acevedo Valle)	pàgina 50

Ens retrobem

Benvolgudes sòcies,
Benvolguts socis,

Fa més de tres anys des del darrer número de *NODES*, el butlletí de l'ACIA, i entremig, com bé sabeu, la crisi sanitària provocada per la COVID-19 bastí un escenari desolador que, a més d'apagar l'enrenou plaent de la quotidianitat, ens donà una nova perspectiva dels assumptes més vitals. La intel·ligència artificial no va restar al marge de la situació, i va contribuir en la lluita per una normalitat que, a dies d'ara, sembla arribar per fi. Així doncs, aquest número de *NODES* el dediquem, com era d'esperar, al paper de la intel·ligència artificial en el sotrac de la COVID-19, i el concebem com una tornada a la regularitat de la publicació del nostre butlletí.

Aquest número segueix la línia marcada pel seu precedent, de manera que el seu contingut es divideix en tres parts. D'una banda, pel que fa a les seccions fixes del butlletí, hem comptat amb contribucions per als apartats de «Ciència-ficció», «Filosofia i pensament» i «Societat». Així, en memòria de Miquel Barceló, el relat *Estudi cas-control*, signat per Jordi Font-Agustí, així com una nota introductòria sobre l'origen del text que ens ha regalat el mateix autor, obren aquest número 59 de *NODES*. En nom del consell editorial, agraiïc a Martí Domínguez, editor de la revista *Mètode*, i als editors del llibre *Científics Lletraferits (Monografies Mètode)*, Jordi de Manuel i Salvador Macip, el permís per reproduir el relat. Tot seguit, un article d'Angel Puyol sobre detecció i eliminació de biaixos ètics en algorismes de triatge i localització per a la COVID-19 estrena la secció de «Filosofia i pensament»; mentre que Núria Vallès-Peris inaugura una altra secció, la de «Societat», amb una anàlisi dels debats sobre els efectes socials i les controvèrsies ètiques de les tecnologies rere els processos d'automatització.

D'altra banda, són tres els textos que componen la discussió sobre el tema monogràfic d'aquest número, tot i que val a dir que l'article de la secció «Filosofia i pensament» bé podria incloure-s'hi en aquest apartat. En primer lloc, Hugo A. Ferrer García i Ismael Sanz presenten un xatbot inclusiu, guardonat amb el Premi ACIA pel bé social, sobre la COVID-19, i argumenten que la pandèmia ha estat un punt d'inflexió pel que fa a la informació digital. El número continua amb una reflexió, a càrrec de Vicent Costa, arran de la fi de l'aplicació per a mòbils Radar COVID. Per acabar amb aquesta part, Josep Puyol Gruart

Editorial

analitza i comenta els resultats de l'enquesta dissenyada per l'ACIA per tal de conèixer l'estat d'opinió sobre la temàtica central del número.

Finalment, tres articles de divulgació, corresponents als treballs guardonats amb el Premi Marc Esteva Vivanco a la millor tesi doctoral en Intel·ligència Artificial dels anys 2019, 2020 i 2021, clouen aquest número. En particular, Emran Saleh Ali Ali, Gerard Canal i Juan Manuel Acevedo Valle són els guanyadors, respectivament, dels anys 2021, 2020 i 2019.

En nom del consell editorial, done les gràcies a les autores i els autors de les contribucions que formen aquest número 59 de *NODES*, i aprofite per a animar-vos a enviar propostes per als propers números del butlletí.

Vicent Costa

L'origen del relat *Estudi cas-control*

Coïncidia amb en Miquel Barceló als actes de la Societat Catalana de Ciència Ficció i Fantasia, i vam estar junts en un grup de treball sobre tecnociència i literatura de l'institut de Tecnoètica que patrocinava la Fundació Epon al tombant del segle. Fa uns dotze o tretze anys, vaig escriure un relat que tenia el mateix Miquel Barceló com a protagonista. Ell el volia publicar a la secció del suplement Ciberp@ís del diari El País que aleshores dirigia, però no hi va ser a temps perquè la van suprimir. Vaig acabar publicant el relat al volum Científics lletraferits que va editar la revista Mètode de la Universitat de València el 2014. L'editor i els coordinadors del llibre autoritzen gustosament la reproducció del relat en homenatge a Miquel Barceló.

Jordi Font-Agustí
Badalona,
desembre 2022



nodes

Estudi cas-control

Jordi Font-Agustí

Per a en Miquel Barceló

En Marcel va entrar al seu despatx de l'Escola d'Enginyeria Informàtica, va remugar un «bon dia» amb la son posant sordina a la veu i va engegar l'ordinador. Era el primer gest que feia cada matí; el següent era desbrossar la pila de correspondència postal. En rebia tanta que els seus companys de departament li havien assignat la safata de dalt. Aquell matí, però, mentre premia el botó del seu pecé, en lloc d'estar per la safata, va buscar amb la mirada el cafè que omplia el despatx amb la seva aroma i va agrair mentalment al seu col·lega Tomàs que mantingués l'hàbit d'arribar d'hora al departament i engegar la cafetera. Encara en tenien una de cafè americà; la d'alta pressió havia estat vetada per motius mediambientals. Va respondre a les preguntes dels col·legues sobre l'acte i el sopar de la nit anterior i, mentre deixava la caçadora en el penjador que tenia la seva foto enganxada al damunt, va pensar que l'avís sonor de «Té correu nou» trigava més de l'habitual. De seguida es digué que això era previsible l'endemà de la festa literària del premi UPC de ciència-ficció.

La Nora, la professora visitant, se li va acostar amb una tassa de cafè a cada mà.

—Jo també he dormit quatre hores només —li digué fent-li un somriure entre còmplice i compassiu.

Ja estava gairebé mig curs amb ells i era manifest que havia aportat un agradable i càlid ambient d'ordre a aquell departament de savis bojos, distrets, atabalats i amb ganes de fer més coses que les abastables. Hi contribuïa la seva veu greu i càlida, la seva afabilitat i el seu accent sud-americà dolç i pausat. De fet, ella estava assignada al Departament d'Interfícies, que era just al costat, però com que no tenien taules lliures, li n'havien assignada una en el de Llenguatges i Metallenguatges.

Mentre acceptava el cafè de les seves mans grans i amb les ungles lacades de vermell, en Marcel pensà que no podia negar que se sentia cada vegada més atret per la seva sensualitat de mare terra encarnada en una còrpora amb bona ossada i un rostre amb un encant particular a pesar o potser a causa d'una manifesta asimetria.

—Ja no tinc edat per a segons què —va dir tornant-li el somriure.

—No diguis això —va respondre la Nora—. La lluna d'aquesta nit no era gaire propícia per al descans.

Aquella dona que tothom considerava brillant i que ell mateix reconeixia que l'atreïa, li despertava sentiments ambivalents. D'una banda, pensava que la seva competència professional era fora de dubte: l'havia impressionat el seu projecte de desenvolupament d'interfícies

que era capaç de fer que criatures amb paràlisi cerebral greu activessin mecanismes a través d'un pecé per mitjà d'ones cerebrals amb una riquesa de matisos que, fins aleshores, ningú havia aconseguit, però, per altra banda, veia amb desdeny la seva dedicació a l'astrologia i considerava una lamentable xacra que apliqués els seus coneixements informàtics a escriure programes per a traçar cartes astrals i altres col·lonades que ell ni coneixia ni tenia gens d'interès a conèixer. Tanmateix, havia d'acceptar que la Nora no era de les persones que es fan pesades amb les seves dèries: el primer cop que li havia preguntat pel seu signe zodiacal, ell li n'havia etzibada una de fresca i mai més li havia tornat a treure el tema; sí que li semblava veure, però, un somriure burleta dirigit a ell quan en parlava amb en Tomàs i la Rosa, l'associada que havia posat les fotos als penjadors; no els tenia per crèduls, però s'hi estaven tornant o almenys ho feien veure davant la Nora. Recordava que, el mateix matí que li va clavar el moc astral, ella va tenir notícia de la convocatòria del premi de ciència-ficció i, oblidant el desdeny rebut, li va fer explicar com anava, li va mostrar el seu entusiasme amb la idea que una universitat politècnica convoqués un premi literari i es va oferir a col·laborar-hi. La seva aportació al petit grup que mantenia el certamen havia estat estimulante com una onada d'aigua tropical en una cala ombrívola de la Costa Brava.

—Doncs a tu no se't nota, això de la lluna —va dir-li després de fer un glop de cafè, mentre es preguntava si potser s'havia d'atribuir a les seves activitats esotèriques el fet que, malgrat els seus quaranta anys repicats, no hi hagués en ella cap rastre de la gresca de la nit anterior.

Van intercanviar comentaris sobre el mal pas en què s'havia ficat el conferenciant en l'acte de lliurament de premis i com de bé que s'ho havien passat a la sobretaula del sopar i també després, quan dugueren els convidats de copes.

—Va estar molt bé —va dir en Marcel en to de cloenda.

—Molt —va dir ella i, amb un deix que simulava mandra, va afegir: —, però ara, millor que em posi a treballar.

Mentre la Nora tornava a la seva taula, ell contemplà amb discreció el moviment harmoniós del seu darrere.

En Marcel s'assegué davant de l'ordinador per despatxar el correu electrònic. Una concursant afirmava tenir proves concloents que la seva novel·la *Cibernauts antiianquis* no havia guanyat perquè l'havien vetat des de la xarxa Echelon d'espionatge de les comunicacions. Tancà els ulls per controlar els efectes de la ressaca, féu un glop de cafè i, en mirar de nou la pantalla, veié que el correu de la

reclamant jeia a la paperera. Estava segur que ell no havia clicat la icona d'eliminar el correu. Havia tornat a passar. Recordava haver llegit que la sensació d'haver viscut ja una determinada situació estava associada al cansament i a la necessitat del cervell de deturar per un moment els processos mentals, però, malgrat tot, estava segur que era la segona o tercera vegada que un correu electrònic se li eliminava només amb el desig de fer-ho.

Com tothom, en Marcel havia tingut més d'una vegada la sensació que l'ordinador era sensible als seus estats d'ànim; més exactament, als seus moments de tensió. Feia unes setmanes, durant el sopar de jubilació d'un col·lega, per a delectança dels comensals, ell i en Tomàs havien començat una dissertació bufa sobre el tema. «És universal l'evidència que, quan hom està treballant, a la intuïció sobtada que cal guardar de seguida el que s'està fent, segueix inexorablement que l'ordinador es pengi abans de poder-ho fer». Algú hi afegí que com més delicat i difícil de reproduir és el que es té a la pantalla, més curt és el lapse de temps entre la intuïció i la penjada. La Rosa digué que li feia vergonya confessar-ho, però que des que ja no li calia anar amb els tampons a la bossa, trobava que ja no hi havia cap setmana del mes en què l'ordinador li fes el burro. Tot seguit, tres membres del Departament d'Interfícies juraren i perjuraren que havien tingut una becària que, cada vegada que entrava al despatx, sobretot si aquell dia estava de males puces, es desbaratava la wi-fi. La Nora també havia assistit al sopar, però per molt que en Tomàs la punxà perquè hi digués la seva, ella es va limitar a dir que, ves, que mai se sap, que les ones cerebrals són tan electromagnètiques com les que pot emetre o rebre un ordinador, però que el més probable era que la becària els fes perdre l'oremus, que els homes sou com sou.

Es tornà a mirar el correu que havia anat tot sol a la paperera: es resistia a creure que respongués a la realitat la percepció que l'ordinador se li penjava més de l'habitual quan estava nerviós, o que dues o tres vegades un correu se n'havia anat a la paperera només pensant-ho; la seva ment racional li deia que l'única explicació possible era que la mateixa tensió li havia fet fer falses maniobres o que la ressaca li havia fet oblidar d'immediat que havia clicat la icona de la paperera. Seguí amb la safata d'entrada: quatre diaris, tres ràdios i dues cadenes de televisió li demanaven la seva opinió sobre l'obra guanyadora, la vicerectora li enviava una convocatòria de reunió, la biblioteca li reclamava tres llibres, Administració li demanava que enviés amb urgència les dades completes del conferenciant per pagar-li la minuta i el viatge.

Prengué un altre glop de cafè i, en tornar a la pantalla, va veure que s'havia obert el full de càlcul on tenia les dades del jurat, dels guanyadors i del conferenciant. Es va tocar la barba i va mirar al seu voltant: en Tomàs parlava per telèfon mirant per la finestra, la Rosa

estava absorta en la seva tauleta digital, la pissarreta, que en deia ella, i la Nora estava trastejant amb el seu ordinador i un cap de maniquí al qual havia posat una cinta al front amb uns sensors.

Mentre copiava les dades del conferenciant i enviava el correu a Administració, recordà la xerrada de l'acte de lliurament del premi. «Per a nosaltres, els llatins», havia dit el conferenciant, «un ordinador és masculí i una impressora és femenina perquè les nostres llengües tenen flexió de gènere; en canvi, per a un anglès, l'assignació es fa segons les característiques de l'aparell i utilitzant, gairebé sempre, prejudicis que ara es consideren políticament incorrectes i que es transmeten de generació en generació: de ben petits, els nens consideren les excavadores masculines, perquè tenen força, i els gladiols, femenins, perquè són delicats. Algunes assignacions depenen de com es comporti l'aparell: per a un ciutadà de Califòrnia, per exemple, el seu cotxe serà femení si té un comportament capriciós i masculí si és segur i potent. Tanmateix, és general que els ordinadors siguin considerats femenins. Pels Estats Units corre un acudit: l'ordinador és mascle o femella? No ho sé. Doncs és femella, perquè, malgrat que el comanda l'home, té els seus propis criteris i fa el que li dóna la gana». El comentari havia fet remoure la sala i s'havia sentit alguna veu femenina queixant-se. «Els homes i les dones», havia continuat de seguida l'orador per sortir del mal pas, «vivim en una simbiosi tal amb els ordinadors que ja comencem a parlar de les nostres funcions cerebrals i dels nostres òrgans com si es tractés de components d'ordinadors. A més, i això no ens passa amb cap altra màquina, tendim a projectar els nostres sentiments sobre els ordinadors de la mateixa manera que ho fem amb els animals de companyia».

En aquell moment, en Marcel no sabia dir si el seu ordinador era mascle o femella, però sí que tenia la sensació que l'estava obeint com si fos un gosset falder. Mirà de reüll al seu voltant, va veure que ningú l'observava, es va ajustar les ulleres i, encarant-se a la pantalla sense parpellejar, ordenà mentalment al full de càlcul que es tanqués. El marc del full de càlcul va perdre el color i es va obrir una finestreta nova: «Voleu guardar els canvis fets a Premis_CF.ods?». Pensà que no. Just ho acabava de fer, que li pujà al rostre una onada de vergonya i a la meitat de la rutina de tancar-se l'aplicació, l'ordinador va parpellejar, va proferir un reneç de color blau i es va penjar.

En Marcel es va tocar la barba, es mirà el rellotge, va tancar l'ordinador, s'acabà el cafè, va recollir la cartera per anar cap a l'aulari i, mentre agafava la caçadora, s'acomiadà genèricament dels seus col·legues. Abans de sortir, es va mirar la Nora i va sentir una efervescència de sexualitat d'una intensitat més alta que cap de les que aquella dona li havia provocat fins aleshores. «Almenys», pensà mentre tancava la porta amb suavitat darrere seu, «la meva bogeria no em fa preferir els ordinadors a les dones».

Ciència ficció

Quan en Marcel hagué marxat, la Nora es va aixecar, es va servir una nova tassa de cafè, va tornar davant de l'ordinador i va obrir la seva aplicació de correu electrònic. Va fer un glop, va clicar la icona de missatge nou i va escriure: «Si encara necessitàvem més clàries que el camí de la bioquímica pot competir amb el dels implants, aquest matí n'he obtingudes unes de molt bones. El conillet d'Índies involuntari ha vingut avui cansat i havent dormit poc. He posat al seu cafè la dosi estàndard de l'amplificador de l'activitat electromagnètica de les cèl·lules piramidals del còrtex cerebral. S'han produït dues interaccions involuntàries reeixides (correu electrònic i full de càlcul) i, per fi, ;-) dues interaccions voluntàries, totes les quals via ones μ . La segona interacció voluntària ha estat interrompuda quan la sensació de ridícul del subjecte ha provocat l'emissió d'ones Θ . Potser hauríem de reprogramar el receptor d'ones que hem instal·lat en el seu ordinador per tal que les filtri segons els paràmetres que tu mateix vas suggerir. Si m'envies un pegat, jo mateixa l'instal·laré a l'ordinador del conillet d'Índies. No he detectat cap indicati que l'amplificador de l'activitat electromagnètica de les cèl·lules piramidals tingui com a efecte secundari l'augment de la libido, com afirmen alguns experimentadors voluntaris».

Aquest relat es va publicar originalment al llibre:

Jordi de Manuel, Salvador Macip (2014). Científics lletraferits. Monografia Mètode, 7. Universitat de València.

El relat s'ha reproduït amb el consentiment del seu autor, Jordi Font-Agustí, l'editor de la revista Mètode, Martí Domínguez, i els editors del llibre, Jordi de Manuel i Salvador Macip.

Biaixos ètics a IA aplicada al triatge a les UCIs i la traçabilitat de contagis a través de les apps de rastreig

Angel Puyol

Entre l'octubre del 2020 i l'octubre del 2022, es va crear un grup de recerca sobre Detecció i eliminació de biaixos en algorismes de triatge i localització per a la COVID-19, finançat per la Fundació BBVA, del qual he estat investigador principal. L'objectiu principal del projecte ha consistit a explorar els biaixos que poden contenir els algorismes usats en els sistemes d'intel·ligència artificial (IA) tant per al triatge a les unitats de cures intensives (UCIs) per decidir quins pacients de Covid-19 han de quedar fora en cas de racionament, com els biaixos continguts en els algorismes basats en la geolocalització a través d'aplicacions mòbils per al control dels contagis.

En aquests dos anys, s'ha començat revisant de manera sistemàtica la literatura científica existent fent servir la metodologia d'una scoping review. Els resultats s'han publicat a Delgado, J.; de Manuel, A.; Parra, I.; Moyano, C.; Roda, J.; Guersenzvaig, A.; Ausín, T.; Creu, M.; Puyol, A., i Casacuberta, D. "Bias in algorithms of AI systems developed for COVID-19: A scoping review." *J Bioeth Inq.* 2022. Posteriorment, es van analitzar els diferents algorismes èticament rellevants que s'han descrit actualment associats al triatge a les UCI i a l'anàlisi de contactes de COVID-19 en una comunitat. Els resultats s'han publicat en diversos articles, entre els quals destaca Casacuberta, D.; Guersenzvaig, A., i Moyano, C. "Justificatory explanations in machine learning: for increased transparency through documenting how key concepts drive and underpin design and engineering decisions". *AI & Soc.* 2022; i s'han presentat a congressos internacionals a Hannover, Basilea, Berlín i Oegstgeest (Holanda).

L'anàlisi de la literatura ha mostrat que les principals fonts de biaixos identificats als sistemes de triatge i les aplicacions de monitorització per a COVID-19 estan principalment relacionats amb la variabilitat de les fonts de dades, la manca d'estandardització, la recopilació inadequada de dades i les desigualtats en la implantació tecnològica. A més, durant aquesta fase d'anàlisi s'han identificat problemes ètics relacionats amb la privadesa, el consentiment i la manca de regulació en les aplicacions mòbils. Els biaixos relacionats amb les disparitats de salut i els determinants socials de la salut no són els temes principals dels estudis a la literatura científica, i només s'inclouen tangencialment o indirectament, particularment en

"L'anàlisi de la literatura ha mostrat que les principals fonts de biaixos identificats als sistemes de triatge i les aplicacions de monitorització per a COVID-19 estan principalment relacionats amb la variabilitat de les fonts de dades, la manca d'estandardització, la recopilació inadequada de dades i les desigualtats en la implantació tecnològica."

modulés

Filosofia i pensament

“Durant el transcurs d’aquesta part de la investigació, s’ha desenvolupat tant un glossari, en el qual es recullen termes bàsics, com una biblioteca virtual que recopila la literatura més rellevant sobre el tema, de manera que faciliti l’accés tant a la terminologia específica relacionada amb la IA com amb el triatge tant per al públic en general com per a les persones interessades a investigar aquests temes.”

les revisions narratives de les aplicacions de rastreig i monitorització. La nostra investigació ha demostrat, però, que aquest tipus de biaix és més rellevant del que es creu i hauria de ser tingut més en compte en les anàlisis de biaixos a IA en salut a partir d’ara, també a IA aplicada a diagnòstic mèdic i prioritització de recursos sanitaris. Durant el transcurs d’aquesta part de la investigació, s’ha desenvolupat tant un glossari, en el qual es recullen termes bàsics, com una biblioteca virtual que recopila la literatura més rellevant sobre el tema, de manera que faciliti l’accés tant a la terminologia específica relacionada amb la IA com amb el triatge tant per al públic en general com per a les persones interessades a investigar aquests temes. Aquests recursos estan disponibles a la pàgina web creada per al projecte: www.airesearchsesgos-covid19.com.

A partir d’aquests resultats, s’han dut a terme una sèrie d’articles sobre l’impacte ètic de la IA i s’han analitzat les estratègies de mitigació dels biaixos. Alguns d’aquests articles estan actualment en revisió en revistes com ara *Public Health Ethics*, *AI & Society* i *Global Bioethics*, a més de revistes nacionals i altres de divulgació. Tot plegat ha donat lloc al desenvolupament d’una guia d’autoavaluació ètica per a organitzacions i professionals de l’àmbit de la salut. Aquest document constitueix una recomanació sobre els aspectes ètics que s’haurien de tenir en compte en el disseny i l’ús de sistemes de decisió automatitzada aplicats a la salut. A més d’una justificació teòrica, s’hi pot trobar adjunta una guia pràctica consistent en 35 preguntes binàries especialment dedicada a la justícia algorísmica per poder avaluar sistemes d’IA al llarg de tot el cicle de vida. Les respostes es renderitzen en un velocímetre que indica el nombre de respostes afirmatives amb un color, de manera que visualment es pugui obtenir un marcador gràfic de l’estat del sistema en termes ètics. Aquest sistema d’autoavaluació ètica ha estat implementat a la pàgina web del projecte, on, a més, es poden consultar tots els resultats de la investigació. Aquest document ha estat sotmès a una avaluació per un grup d’experts i expertes en ètica, filosofia de la ciència, epidemiologia, responsables sanitaris d’UCI i intel·ligència artificial de tot el territori espanyol. A aquest grup d’experts i expertes se li va enviar la documentació prèviament i vam celebrar una jornada de treball comú a Barcelona, on van poder aportar observacions i suggeriments de millora que van ser recollits per l’equip investigador.

El projecte vol tenir perllongament en tres àmbits diferents. A partir de l’elaboració de la guia d’autoavaluació ètica, una de les noves línies de recerca consistiria a poder testear-ne l’aplicació en un entorn real. L’equip investigador estaria interessat a poder comprovar l’aplicabilitat de l’eina tant en una nova fase d’avaluació dins d’una institució de recerca com en una institució sanitària, de la qual podríem extreure conclusions importants que donarien sentit a la seva aplicació a major

escala. D'altra banda, el desenvolupament d'un model interactiu de la guia d'autoavaluació permetria posar en pràctica els principis ètics i passar-los d'un pla teòric a un de pràctic. Així mateix, ens permetria crear un registre estadístic de dades d'ús. És important avaluar de manera estadística tant l'ús (nombre d'usuaris, satisfacció amb el producte, etc.) com els resultats de l'eina d'autoavaluació ètica que hem dissenyat segons els usuaris que anem tenint. Aquest conjunt de dades obtingut, si es replica a gran escala, podria suposar un avanç important i valuós per poder tenir una visió global de la situació ètica del desenvolupament de sistemes d'intel·ligència artificial en l'àmbit estatal.

Una segona línia de treball consistiria a elaborar un pla de formació adreçat a dissenyadors i aplicadors d'intel·ligència artificial en l'àmbit de la salut, amb l'objectiu d'augmentar la comprensió i la sensibilització envers els problemes ètics de la IA en salut reflectits a la guia d'avaluació ètica que hem elaborat. El pla que proposaríem constaria de tres pilars principals:

1) Accions de formació amb la finalitat que els equips que la realitzin puguin treure-li el màxim profit possible a la utilització de l'autoavaluació. Aquestes sessions podrien realitzar-se en centres de salut, hospitals, centres de recerca i qualsevol tipus d'institució relacionada amb la pràctica i la recerca al sector medicosanitari, així com en centres de disseny d'IA. Anirien destinades tant a personal de l'àmbit sanitari (gestors/es hospitalaris, membres de comitès d'ètica, personal mèdic, etc.) com a programadors d'algorismes, i buscarien expandir els coneixements i generar reflexió al voltant dels temes abordats a la guia.

2) La realització d'un curs de "formació per a facilitadors". Els facilitadors són els encarregats d'assistir a equips (que, en el futur, seran molt nombrosos) en la realització de l'autoavaluació (tant a partir de la nostra guia com d'altres). Aquest curs oferiria als facilitadors un coneixement més avançat, de manera que durant la realització de l'autoavaluació poguessin ajudar a la reflexió realitzada al si dels equips aportant nous coneixements i contribuint a desenvolupar-ne la sensibilitat ètica.

3) El desenvolupament d'un manual per a facilitadors que estaria estructurat al voltant de les preguntes de la guia d'autoavaluació ètica i que contindria dinàmiques docents, exercicis i documentació extra de suport. Aquest manual estaria organitzat amb l'estructura d'un kit d'autoaprenentatge, amb activitats individuals i en grup per posar en pràctica les idees desenvolupades en el document d'autoavaluació. Aquest tipus de manuals aplicats a la IA en l'àmbit de la salut no existeix en aquests moments i el seu desenvolupament seria completament innovador.

nodes

Filosofia i pensament

“La tercera línia de recerca que sorgeix arran d’aquest projecte consisteix a discutir èticament i filosòfica l’impacte ambiental de la intel·ligència artificial.”

La tercera línia de recerca que sorgeix arran d’aquest projecte consisteix a discutir èticament i filosòfica l’impacte ambiental de la intel·ligència artificial. Els sistemes d’aprenentatge automàtic engloben una sèrie de processos amb una empremta ecològica que no és neutra: des de la fase d’extracció de materials per construir els equips informàtics fins a la fase de processament de les dades, passant per altres fases com la de transport o la d’emmagatzematge i conservació de dades, es generen una sèrie d’emissions de gasos amb efecte d’hivernacle, de despesa energètica i de recursos, i, fins i tot, de destrucció d’hàbitats naturals. Arran de les investigacions dutes a terme en el marc d’aquest projecte sobre els biaixos de la intel·ligència artificial utilitzada per mitigar els danys a la salut clínica i pública produïts per la COVID-19, convé preguntar-se si és ètic fer ús de l’aprenentatge automàtic en medicina, però ara sumant la consideració pels seus efectes perjudicials al medi ambient.

La naturalesa sociotècnica de la IA: ètica, política i tecnologia

Núria Vallès-Peris

Amb el ràpid desenvolupament dels processos d'automatització, especialment amb la intel·ligència artificial (IA) i la robòtica i la seva progressiva aplicació en múltiples àmbits de la nostra vida, han proliferat també els debats sobre els efectes socials i les controvèrsies ètiques d'aquestes tecnologies. Qüestions com la llibertat, la responsabilitat, la privacitat, el control social, els biaixos, la desaparició de determinats perfils professionals o la transformació d'algunes tasques, entre molts altres temes, acompanyen la IA.

Ara bé, tot i aquest interès i debats, la majoria de sistemes d'IA i les tecnologies relacionades s'estan aplicant amb una supervisió mínima, amb escassa reflexió crítica, poca responsabilitat política i una deficient avaluació i seguiment del seu impacte (Coeckelbergh, 2020). Els qui acostumen a desenvolupar aquests sistemes són empreses privades multimilionàries que dominen el relat (i el mercat) i que alimenten el bombo publicitari, establint un marc de discussió poc creïble que, alhora, segresta els usos i el desenvolupament de la IA, i cancel·la l'anàlisi crítica sobre els seus efectes.

Quina és la ideologia dels fundadors de les grans multinacionals que dominen el mercat de la IA (Geburu, 2022)? Fins a quin punt no és una hipocresia i un engany parlar d'una IA beneficiosa i ètica si es fa servir mà d'obra precària i no es tenen en compte els costos mediambientals per desenvolupar alguns sistemes d'IA (Crawford, 2021)? És acceptable desenvolupar sistemes amb web scraping, sense preocupar-se de la parcialitat, el consentiment, els drets d'autor o els discursos d'odi (van Rooij, 2022)? Les dinàmiques que apunten aquest tipus de preguntes no fan referència a les eines i tecnologies de manera aïllada, sinó que tenen a veure amb l'entorn social, polític i econòmic en què es desenvolupen els sistemes d'IA. En l'escenari que s'obre des d'aquest plantejament, en aquest article es planteja una aproximació a l'estudi dels efectes socials i les controvèrsies ètiques de la IA tenint en compte tot l'entramat de relacions que es generen al seu voltant, és a dir, el sistema sociotècnic de la IA.

Amb aquests objectius l'article s'organitza de la manera següent: al primer apartat es planteja la problemàtica sobre problematitzar la proliferació de nombrosos marcs i principis ètics al voltant de la IA; seguidament es presenta l'aproximació dels estudis de la ciència i la tecnologia com a resposta a algunes de les problemàtiques plantejades, una perspectiva que vincula l'ètica de les tecnologies amb

“En l'escenari que s'obre des d'aquest plantejament, en aquest article es planteja una aproximació a l'estudi dels efectes socials i les controvèrsies ètiques de la IA tenint en compte tot l'entramat de relacions que es generen al seu voltant, és a dir, el sistema sociotècnic de la IA.”

modulés

la política; a l'apartat sobre 'Els efectes socials de la IA' es presenten les principals línies de recerca des d'aquesta aproximació i, finalment, a l'apartat sobre 'Els processos d'innovació i disseny' es desenvolupa una proposta d'acció política per a una IA al servei del bé comú, des de la comprensió de la IA com un sistema sociotècnic.

1. El problema de la proliferació dels principis ètics en la IA

Acompanyant la narrativa que es mou entre la utopia i la distopia al voltant d'una suposada societat futura automatitzada, s'han multiplicat les iniciatives que busquen establir principis ètics que serveixin per orientar l'adopció d'una IA socialment beneficiosa (com els *Asilomar AI Principles*; la *Montreal Declaration for Responsible AI*; la *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems* de la IEEE; la *Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems*; o *The Partnership on AI*, entre molts altres, tant amb vocació generalista com aplicats a àmbits específics - Floridi & Cowls, 2019). Ara bé, aquesta proliferació de marcs i principis ètics comporta una sèrie de problemàtiques, de les quals en destaquen tres:

“La tendència a l'especulació i la fantasia: en aquest escenari de múltiples debats sobre la IA hi ha una important tendència a l'especulació, la qual cosa dificulta articular propostes i recomanacions ètiques que puguin ajudar en les pràctiques quotidianes i les controvèrsies particulars.”

(1) La tendència a l'especulació i la fantasia: en aquest escenari de múltiples debats sobre la IA hi ha una important tendència a l'especulació, la qual cosa dificulta articular propostes i recomanacions ètiques que puguin ajudar en les pràctiques quotidianes i les controvèrsies particulars. Tot i que els avenços i aplicacions de la IA són indiscutibles, també és cert que el seu desenvolupament ha anat acompanyat d'una narrativa de solucionisme tecnològic (Morozov, 2016) o, el que en el món anglosaxó es coneix com a snake oil (Naranayan, 2019). L'snake oil al nostre context es podria traduir com una mena de 'bàlsam de tigre', un unguent miraculós que no se sap ben bé com funciona però que ho cura tot. El mateix succeeix amb certs discursos sobre la IA, que afirmen que aquesta tecnologia és la solució als problemes de pobresa, de l'escalfament del planeta, al col·lapse dels sistemes sanitaris, la crisi de les cures, etc. Els debats futuristes poden ser útils per identificar possibles riscos i delimitar escenaris futurs desitjables, però formen part d'una perspectiva hipotètica, no d'una controvèrsia actual. Quan aquest enfocament es trasllada al debat ètic planteja problemes importants, el principal dels quals és el determinisme tecnològic. Si es dona la impressió errònia que les discussions són sobre tecnologies actuals en lloc de potencialitats hipotètiques, es presenta un tipus de desenvolupament tecnològic imparabile front al qual la ciutadania només pot adaptar-s'hi, siguin quins siguin els seus riscos i independentment de les necessitats de gran part de la població.

(2) La mercantilització dels principis: tenint en compte la multiplicitat de propostes, tot i les bones intencions que es puguin tenir des de l'àmbit de la recerca, la indústria o l'Administració pública, allò

que té a veure amb tenir en compte criteris ètics en la IA s'ha convertit en una tasca confusa i aclaparadora. Diferents organitzacions, tant públiques com privades, competeixen per fer valer la seva proposta per damunt de les altres, i les parts interessades poden "comprar" la més atractiva. L'ètica es converteix així en un 'mercat de principis' (Floridi & Cowls, 2019). Integrar un determinat discurs o respondre una checklist que faci referència a qüestions ètiques es pot acabar convertint en un valor de canvi, en una forma de blanquejar problemes greus en l'aplicació de la IA o una manera de respondre als criteris de les entitats financeres, sense integrar mecanismes que en garanteixin el seu bon ús i els efectes socialment desitjables.

(3) La monopolització de la discussió: l'acapament del debat des d'una determinada manera d'entendre l'ètica basada en principis i propostes normatives (una perspectiva necessària però no suficient), que entén l'ètica fonamentalment en fer judicis morals (Sætra & Danaher, 2022), deixa en una posició marginal altres aproximacions a la discussió sobre les tecnologies per al bé comú. L'interès per desenvolupar una IA ètica i socialment responsable no és nou, i el debat al voltant d'aquestes qüestions és gairebé tan antic com l'establiment de la IA com a camp d'investigació acadèmica als anys 50. Ara bé, igual que la recerca i les aplicacions de qualsevol tecnologia es van transformant i diversificant, també ho fan les preguntes i aproximacions a les tecnologies des d'una perspectiva ètica i social. En aquest sentit, hi ha una llarga i consolidada tradició en la recerca acadèmica, en la política científica i en el disseny tecnològic que ofereix perspectives complementàries a la dels principis i les guies ètiques.

2. Els estudis de la ciència i la tecnologia: la política per altres mitjans

Al costat de les moltes potencialitats i aplicacions de la IA, els reptes que planteja a nivell social i polític són especialment complexos. En aquest article es presenta una aproximació a l'ètica i la IA complementària a aquella que es basa en l'establiment de principis i normes que, en lloc de partir de l'estreta vinculació entre ètica i moralitat, entén l'ètica com a vinculada a la política. En cap cas es tracta de substituir la necessària regulació i establiment de principis, sinó d'ampliar el marc d'anàlisi i reflexió integrant la idea de la IA com un sistema sociotècnic.

Aquesta manera d'entendre l'ètica de les tecnologies té una llarga trajectòria en els estudis de la ciència i la tecnologia, els coneguts com a STS (de l'anglès Science and Technology Studies). Els STS són un àmbit de coneixement que es va començar a articular a finals dels anys 60 i principis dels 70 del segle XX, moment en què la imatge de la ciència i la tecnologia que s'havia forjat durant la Il·lustració va entrar en un període de profunda crisi. La idea tradicional

“Amb referència a la noció de l'ètica de les tecnologies com a política, l'enfocament STS es basa en dos conceptes clau: la política dels artefactes i la democratització de les tecnologies.”

de l'autonomia i la neutralitat política del coneixement científic i el desenvolupament tecnològic va començar a ser qüestionada. Es van posar de manifest algunes de les conseqüències mediambientals més negatives del desenvolupament tecnocientífic i els riscos que suposen per a la salut humana; van aparèixer el moviment ecologista, el moviment anti-nuclear, el moviment feminista i, en general, nombrosos grups i associacions que, des de la cultura i l'activisme, qüestionaven explícitament els vincles estrets entre la investigació tecnocientífica, l'estament militar i la indústria. Paral·lelament al món acadèmic, des de la filosofia, la història de la ciència i la sociologia del coneixement, es van començar a desenvolupar aproximacions alternatives a aquelles defensades pels filòsofs racionalistes o neopositivistes. En l'àmbit del coneixement tecnològic aquestes perspectives alternatives construeixen una concepció de la tecnologia o els artefactes que desafia la nítida distinció entre les qüestions epistemològiques i ètiques, entre ciència i política, entre allò racional i allò mític, o entre natura i societat.

La idea central d'aquestes propostes rau en la idea que, en la construcció del món comú, societat, ciència, tecnologia i natura es construeixen mútuament entre si a través de diversos processos. Les societats contemporànies entomen rellevants reptes tecnocientífics (relacionats amb la IA, la generació d'energia o els organismes modificats genèricament, entre molts d'altres), uns reptes que no poden desvincular-se dels processos polítics, socials, culturals o econòmics. Amb referència a la noció de l'ètica de les tecnologies com a política, l'enfocament STS es basa en dos conceptes clau: la política dels artefactes i la democratització de les tecnologies.

La qüestió de la política dels artefactes parteix de la idea que les tecnologies no són bones ni dolentes, però tampoc són neutres (Kranzberg, 1986). Els artefactes tecnològics porten inscrits una sèrie de valors, relacions de poder, interessos econòmics, etc. (Winner, 1983). Per tant, les tecnologies són mediadores actives que contribueixen a la formació de les percepcions i la interpretació que fem de la realitat: són moral materialitzada, en tant que ofereixen propostes materialment inscrites sobre com actuar (Verbeek, 2008). Ara bé, el disseny de tecnologies no determina la societat ni les mediacions que configura, perquè aquestes mediacions estan conformades per factors que depenen tant de l'artefacte, com del context en què s'utilitza aquest artefacte (Feenberg, 2010).

L'altre concepte clau, sobre la democratització de les tecnologies, fa referència a la necessitat d'entendre els criteris ètics més enllà dels marcs jurídics i els catàlegs de recomanacions, que regulen l'ús de tecnologies ja dissenyades. El disseny de tecnologies en un món democràtic requereix de compromís polític i de canvis institucionals, que garanteixin que la producció de tecnologies responen a les necessitats

i desitjos de les persones, no de les grans corporacions internacionals. No és possible parlar de societats democràtiques sense una política democràtica de disseny tecnològic (Feenberg, 2010). No es tracta de votar com ha de ser un algoritme, sinó d'integrar mecanismes al llarg del procés de disseny d'un sistema d'IA, per conèixer les necessitats de la ciutadania i valorar els potencials efectes d'aquella aplicació amb els diversos actors directament concernits.

Ambdues idees estan totalment imbricades i configuren el marc d'estudi de l'ètica-política de la IA des dels STS. Ara bé, si ens fixem més en la primera, la política dels artefactes, emergeixen una sèrie de temàtiques i enfocaments propis de la recerca sobre els efectes de la IA; mentre que, si ens fixem més en la segona, la democratització de les tecnologies, el focus se situarà en l'estudi dels processos d'innovació i disseny. En els apartats següents es desenvolupen cadascuna d'aquestes qüestions, entenent, però, que la distinció entre els efectes socials de la IA i les formes d'innovació i disseny és purament analítica, ja que el disseny, la implementació i els efectes de les tecnologies no són processos lineals ni responen a una lògica causa-efecte (Callon et al., 2009).

3. Els efectes socials de la IA

L'abast de les transformacions i implicacions relacionades amb l'automatització que suposa la IA i la robòtica és múltiple i divers i, per tant, qualsevol revisió dels seus efectes i controvèrsies no pot pretendre ser exhaustiu. Tenint en compte aquesta limitació es poden analitzar els efectes de la IA en relació amb tres fases diferents del desenvolupament o aplicació d'aquestes tecnologies: (1) el procés de producció dels sistemes, (2) la gestió de les dades i desenvolupament d'algoritmes i, (3) finalment, l'elaboració de prediccions i judicis automatitzats.

3.1. Producció de sistemes

Actualment un petit nombre de poderoses corporacions tecnològiques despleguen aplicacions de la IA a escala planetària i els seus sistemes són considerats com a comparables, o inclús superiors, a la intel·ligència humana. Les formes d'aprenentatge automàtic han substituït els sistemes simbòlics i, juntament amb un relat dominat per les grans corporacions internacionals, s'han anat esvaint també les discussions i l'escepticisme en l'àmbit computacional sobre la fantasia de pensar que una màquina pot ser intel·ligent o que la intel·ligència és un procés racional, conscient i descontextualitzat.

Les xarxes globals de computació integren un entramat d'empreses tecnològiques amb una capacitat de producció i intervenció social inimaginable. Aquestes xarxes se sostenen a partir d'una intensa política extractiva de minerals i treball (Crawford, 2021).

La demanda del sector tecnològic de terres rares, petroli i carbó és enorme. La competència per produir models cada cop més ràpids i eficients, de processament del llenguatge i visió per computació, requereix de gran quantitat d'energia i suposa una major petjada de carboni. Juntament amb l'energia i els recursos naturals, la IA està feta de treball humà i, majoritàriament, d'explotació laboral. Els discursos al voltant de la IA i la digitalització, en general, han tendit a obviar el treball humà involucrat en la producció, la implementació i el manteniment dels sistemes. La lectura de les retòriques al voltant del programari, els dispositius i els entorns intel·ligents posa de manifest l'elusió d'aquest treball. Freqüentment les visions de futur que s'ofereixen, amb promeses imminents i respostes pràctiques a necessitats complexes, representen un somni utòpic molt problemàtic: una infraestructura perfecta i invisible, que uneix la promesa de màquines intel·ligents amb les necessitats de l'economia de serveis (Suchman, 2007).

“Més enllà dels greus problemes que suposa el capitalisme de la vigilància, el procés d'estandardització de dades també planteja profunds debats metodològics i epistemològics.”

3.2. Gestió de dades i desenvolupament d'algoritmes

Qüestions com la privacitat, la confidencialitat i la propietat de les dades han estat intensament discutides (Zuboff, 2019), perquè, malgrat existir nombroses recomanacions i marcs de regulació, els diversos contextos en què s'utilitza la IA fan difícil per als usuaris i els propis legisladors fer un seguiment de l'itinerari d'aquestes dades o dels seus possibles usos maliciosos: amb finalitats lucratives, per a la manipulació política o el control social.

Més enllà dels greus problemes que suposa el capitalisme de la vigilància, el procés d'estandardització de dades també planteja profunds debats metodològics i epistemològics. L'estandardització que necessita l'estadística, la ciència per excel·lència associada a qualsevol forma de gestió de grans volums de dades, suposa l'adopció de models teòrics sovint descartats o almenys en discussió des d'altres disciplines, que s'utilitzen acríticament en els sistemes estadístics per la seva utilitat per a la simplificació i codificació de la realitat social.

A partir de dades estandarditzades els algorismes estableixen patrons d'ordenació i previsió. No obstant això, la vida social està profundament travessada per eixos de desigualtat, per la qual cosa els algorismes reproduiran aquests patrons desiguals (per motius de gènere, classe social, grup ètnic, etc.); el que es coneix com a biaix o injustícia algorítmica, que vindria a ser un reflex de les injustícies socials i, sovint, la seva exacerbació. Pot ser que les dades estiguin esbiaixades i hi hagi una sobrerrepresentació de determinats grups que alimenten l'algorisme, una qüestió molt controvertida quan s'apliquen sistemes d'IA en contextos diferents dels que s'han pres les dades. Però, fins i tot quan les dades són representatives, els sistemes d'IA també estigma-

titzen i poden provocar una doble discriminació, no perquè l'algorisme per se sigui injust o discriminatori, sinó perquè reproduïx les formes d'organització social disponibles, profundament injustes i polaritzades (de Sousa Santos, 2016).

3.3. Elaboració de prediccions i judicis automatitzats

Els algoritmes són cada cop més precisos com a mecanisme per emetre judicis automatitzats, en temes com la detecció de correu brossa, el material protegit per drets d'autor, la detecció de discursos d'odi o la recomanació de continguts. Els judicis automatitzats són consubstancials a la lògica algorítmica, que opera a través d'una sèrie de proposicions. Alhora, una de les recomanacions més consensuades, quan s'aplica la IA per emetre judicis i determinar el curs de l'acció, és que l'humà sempre ha de prendre l'última decisió. És la idea de 'the human in the loop'.

No obstant això, la idea de la darrera decisió en mans de l'humà i no de la màquina és un plantejament que està constantment en tensió. D'una banda, perquè l'algorisme comporta l'elaboració de petits múltiples judicis i decisions automatitzades i, de l'altra, perquè, a mesura que es desenvolupen les tècniques i les tecnologies associades a la IA, augmenta també l'autonomia d'aquests sistemes, és a dir, la seva capacitat per gestionar un volum més gran de dades, establir més judicis i actuar en relació amb aquests judicis. Com es posa de manifest en els debats sobre els sistemes d'armes autònoms, per exemple, dir que l'última decisió ha de ser humana és una entelèquia, perquè els humans no podem gestionar el volum de dades i la velocitat dels judicis automatitzats que operen al sistema (Sparrow, 2007).

D'altra banda, la qüestió de l'autonomia va lligada a la responsabilitat. Enfront d'una autonomia creixent d'aquestes tecnologies, qui és responsable dels perjudicis o beneficis de l'aplicació de sistemes que possibiliten la delegació de l'acció i la decisió d'acord amb judicis automatitzats? A més d'aquest tipus de reflexions sobre l'autonomia i la responsabilitat de la IA, quan els algoritmes fan funcions de reconeixement facial, predicció del llenguatge o detecció d'objectes, les dades que utilitzen es converteixen en infraestructura i el significat específic o el context d'aquestes dades s'assumeix com a irrellevant. La cadena temporal de judicis i decisions d'aquestes infraestructures superen l'espai i el temps de la producció de les dades originals, i quan els sistemes d'IA s'utilitzen per fer prediccions (en relació amb la delinqüència, el rendiment a la feina, les malalties, etc.) o prendre decisions, el resultat final d'aquests judicis estarà completament descontextualitzat i realitzat segons l'acumulació d'una sèrie de criteris desconeguts i/o no explicats.

modies

4. Processos d'innovació i disseny

Una línia d'aportacions rellevants des dels STS gira al voltant de la necessitat de potenciar la democratització de la ciència i la tecnologia, garantint així que la innovació respongui a les necessitats i valors de la societat, mitjançant la utilització de metodologies per facilitar la participació pública (Felt, 2018). En l'àmbit de la IA aquesta preocupació ha donat fruit al desenvolupament d'una línia de treball que, a diferència de l'estudi dels efectes socials de la IA, té una voluntat propositiva, acompanyant la recerca i la innovació de propostes i actuacions específiques. Des d'aquesta perspectiva s'associa una IA ètica al desenvolupament de polítiques d'innovació responsables i; al desenvolupament d'experiències de disseny de la IA que integrin diverses formes de participació ciutadana i de compromís de l'Administració pública.

"L'RRI parteix de la idea que les consideracions al voltant de les implicacions ètiques o socials s'han d'integrar durant les fases inicials de disseny tecnològic, no a posteriori durant la seva implementació."

4.1. La Recerca i la Innovació Responsables

Diversos dels debats i les preocupacions sobre les implicacions ètiques i socials de la IA i la necessitat d'integrar la participació ciutadana al disseny de sistemes d'IA, estan avui representats en la política científica i de desenvolupament tecnològic sota el paradigma de la Recerca i la Innovació Responsables (RRI, del terme en anglès Responsible Research and Innovation). L'RRI parteix de la idea que les consideracions al voltant de les implicacions ètiques o socials s'han d'integrar durant les fases inicials de disseny tecnològic, no a posteriori durant la seva implementació. Aquesta qüestió no és només responsabilitat dels científics o tecnòlegs, al contrari, sinó que és necessari desplegar una sèrie de mecanismes polítics i institucionals, en què hi participin diversos actors i àmbits de coneixement.

L'RRI té els seus antecedents en una línia de la política científica que es va iniciar a finals de la dècada de 1980, amb el Projecte Genoma Humà. Des dels seus inicis, preocupats per com la nova informació genètica es podria utilitzar i com es podien protegir els individus i la societat dels riscos pel seu mal ús, el Projecte Genoma Humà va establir un grup de treball per tenir en compte i fer recomanacions sobre les implicacions ètiques, legals i socials de la recerca. Des d'aquell moment aquestes qüestions es van agregar sota les sigles ELSI (Ethical, Legal and Social Issues) als EUA o ELSA (Ethical, Legal and Social Aspects) a Europa, tot i que aquesta diferenciació territorial s'ha anat difuminant durant els últims anys. Ara bé, el marc ELSI va rebre algunes crítiques importants, la principal de les quals és que aquest tendeix a bifurcar la realitat, entenent com a dues entitats diferenciades la ciència i la societat (Argudo-Portal & Domènech, 2020).

Davant d'aquest tipus de crítiques la RRI presenta una alternativa a la tendència de respondre als impactes de la ciència i la tecnologia a posteriori i proposa introduir les dimensions socials,

ètiques i polítiques en el propi procés d'innovació tecnològica des de les fases inicials de disseny. RRI és un concepte paraigua, la intenció principal del qual és evitar conseqüències no desitjades de la ciència i la tecnologia i donar suport proactivament a noves maneres de desenvolupar innovacions d'acord amb les necessitats i esperances dels ciutadans (Owen et al., 2013). Es parteix de la idea que, juntament amb els beneficis de les innovacions i tecnologies, aquests sovint plantegen preocupacions ètiques o tenen efectes secundaris nocius. Actualment, aquesta perspectiva s'ha adoptat com a política científica de les principals agències finançadores dels EUA i la Unió Europea, que consideren que la recerca i la innovació s'han d'alinèar amb els reptes de la societat actuals, considerant la participació pública, l'educació i la participació ciutadana com a principis rectoros de l'RRI per garantir models tecnocientífics democràtics.

4.2. Democràcia, participació i interdisciplinarietat

Per entendre l'impacte social de la IA cal adonar-se que aquesta és més que la suma dels seus components de programari. Tal com s'ha explicat a l'apartat sobre la política dels artefactes, els efectes de la IA inclouen el context social en què es desenvolupen, s'utilitzen i modifiquen, amb la seva varietat de grups d'interès, institucions, cultures, normes i contextos. Per això es diu que la IA és un sistema sociotècnic, és a dir, que quan es consideren els efectes i la governança de la IA, o l'artefacte que incorpora aquesta tecnologia, el component tècnic no es pot separar de l'àmbit econòmic, polític, social o cultural en què aquest s'emmarca. Aquest sistema inclou persones i organitzacions amb rols molt diferents (dissenyador, fabricant, usuari, Administració pública, ecosistema d'innovació, legislació, etc.), les seves interaccions i els procediments i processos que organitzen aquestes interaccions (Dignum, 2021).

L'ètica de la IA no és, com alguns poden afirmar, una manera de donar a les màquines algun tipus de responsabilitat per les seves accions i decisions (Dignum, 2021). Per contra, l'ètica de la IA requereix d'una política científica i de desenvolupament tecnològic, que assumeixi la responsabilitat de vetllar perquè la ciutadania pugui confiar en les formes de desenvolupament de la IA i les seves aplicacions. Des d'aquesta responsabilitat política, les directrius, els principis i les estratègies han de tenir en compte la visió sociotècnica de la IA. No és l'artefacte o l'aplicació d'IA el que és ètic, fiable o responsable, és la xarxa col·lectiva de relacions en què participa la IA on cal situar la preocupació perquè aquell entramat sociotècnic garanteixi el bé comú i respongui a les necessitats de benestar individual i col·lectiu, en un context democràtic (Vallès-Peris & Domènech, 2021).

Per tant, el desenvolupament de la IA de manera ètica no és una qüestió que només tingui a veure amb una sèrie de programaris i

“Per entendre l'impacte social de la IA cal adonar-se que aquesta és més que la suma dels seus components de programari.”

la regulació de les seves aplicacions, sinó que té a veure amb el reconeixement de la seva naturalesa sociotècnica i la seva imbricació amb diversos processos que no són exclusivament tecnològics, sinó que són també econòmics, polítics o culturals. El desenvolupament d'estratègies democràtiques són clau des d'aquesta perspectiva, tant en la definició dels objectius socials i polítics pels que es vol desenvolupar la IA, com per identificar i donar resposta a les principals preocupacions de la ciutadania al voltant de la introducció d'aplicacions específiques. Davant d'aquesta necessitat es poden utilitzar diversos mecanismes per integrar el coneixement, opinions i valoracions de múltiples actors (desenvolupadors, usuaris, responsables polítics, etc.). Però no n'hi ha prou amb introduir mecanismes participatius, perquè no tots els col·lectius socials són iguals, ni viuen en les mateixes condicions, ni participen de la mateixa manera, ni les incerteses i les desigualtats es reparteixen per igual; per la qual cosa, els diferents grups no són iguals en la seva capacitat per imposar la seva lògica i participar en les societats democràtiques.

Les limitacions dels processos participatius per abordar aquestes qüestions són ben conegudes; per la qual cosa, per poder desenvolupar una IA ètica i al servei del bé comú, és necessari desplegar una sèrie d'estratègies complementàries, que tinguin en compte (Vallès-Peris & Domènech, 2020, 2021):

- El desenvolupament de formes alternatives d'inclusió i representació dels diversos actors concernits en la introducció de sistemes d'IA en contextos determinats, a través de la creació de fòrums híbrids, estratègies de ciència ciutadana, mecanismes de consulta de la ciutadania, etc., però també de la interlocució amb diversos moviments socials i la incorporació de les prioritats i necessitats de col·lectius d'activistes.
- La utilització de metodologies i formes de coneixement de les ciències socials, per tal d'analitzar de manera sistemàtica els efectes socials de les tecnologies i aplicacions actuals, així com també per indagar en les pors, esperances i incerteses dels diversos col·lectius sobre els potencials desenvolupaments tecnocientífics futurs, entre ells també del col·lectiu de dissenyadors i tecnòlegs dedicats al desenvolupament de la IA (actors clau per tal de valorar els riscos i dimensionar les funcionalitats reals i les potencialitats futures).
- El lideratge de l'Administració pública en l'establiment de prioritats i línies d'innovació, desenvolupament i aplicació de la IA, que integri mecanismes de seguiment i monitorització constant. En les societats democràtiques l'Administració pública és la responsable de vetllar pel benestar de la ciutadania i garantir la justícia social i l'equitat, criteris que cal avaluar i, si és necessari, aplicar mesures per corregir i modificar, una qüestió especialment

cabdal quan s'utilitzen els sistemes d'IA en àmbits especialment rellevants per a la vida de les persones, com la salut, l'educació, els serveis socials, l'habitatge o la justícia.

Referències bibliogràfiques

- **Argudo-Portal, V., & Domènech, M.** (2020). The reconfiguration of biobanks in Europe under the BBMRI-ERIC framework: towards global sharing nodes? *Life Sciences, Society and Policy*, 16(9), 1–15.
- **Callon, M., Lascoumes, P., & Barthe, Y.** (2009). *Acting in An Uncertain World : An Essay on Technical Democracy*. MIT Press.
- **Coeckelbergh, M.** (2020). *AI ethics*. MIT Press.
- **Crawford, K.** (2021). *Atlas of AI*. Yale University Press.
- **de Sousa Santos, B.** (2016). La incertidumbre: entre el miedo y la esperanza. In *La difícil democracia. Una mirada desde la periferia europea* (pp. 89–95). Akal.
- **Dignum, V.** (2021). The role and challenges of education for responsible ai. *London Review of Education*, 19(1), 1–11.
- **Feenberg, A.** (2010). *Between Reason and Experience*. Essays in Technology and Modernity. The MIT Press.
- **Felt, U.** (2018). Responsible research and innovation. *Routledge Handbook of Genomics, Health and Society*, July, 108–116.
- **Floridi, L., & Cowls, J.** (2019). A Unified Framework of Five Principles for AI in Society. *Harvard Data Science Review*, 1–15.
- **Kranzberg, M.** (1986). Technology and History: “Kranzberg’s Laws.” *Technology and Culture*, 27(3), 544.
- **Morozov, E.** (2016). *La locura del solucionismo tecnológico*. Katz Editores.
- **Narayanan, A.** (2019). How to recognize AI snake oil. *Arthur Miller Lecture on Science and Ethics*.
- **Sætra, H. S., & Danaher, J.** (2022). To Each Technology Its Own Ethics: The Problem of Ethical Proliferation. *Philosophy and Technology*, 35(4), 1–26.
- **Sparrow, R.** (2007). Killer robots. *Journal of Applied Philosophy*, 24(1), 62–77.
- **Suchman, L.** (2007). *Human–Machine Reconfigurations. Plans and Situated Actions*. 2nd Edition. Cambridge University.
- **Vallès-Peris, N., & Domènech, M.** (2020). Robots para los cuidados. La ética de la acción medida frente a la incertidumbre. *Cuadernos de Bioética*, 31(101), 87–100.
- **Vallès-Peris, N., & Domènech, M.** (2021). Caring in the in-between: a proposal to introduce responsible AI and robotics to healthcare. *AI & Society*.
- **Verbeek, P.-P.** (2008). Morality in Design. Design Ethics and the Morality of Technological Artifacts. In P. E. Vermaas (Ed.), *Philosophy and Design* (pp. 91–103). Springer.
- **Winner, L.** (1983). Do Artifacts Have Politics? In D. MacKenzie & et al (Eds.), *The Social Shaping of Technology* (pp. 55–81). Open University Press.
- **Zuboff, S.** (2019). Surveillance Capitalism and the Challenge of Collective Action. *New Labor Forum*, 28(1), 10–29.

nodes

La pandèmia de COVID-19 i la intel·ligència artificial

Quan la primera gran pandèmia del segle XXI va tornar a posar a prova la societat de la informació en què vivim

Hugo A. Ferrer García
Ismael Sanz

Premi ACIA pel Bé Social 2021

SemanticBots SL

Per conèixer per què la pandèmia va posar a prova la societat de la informació, hauríem de començar parlant a què fa referència aquest terme, un procés d'evolució profunda de la vida i les interseccions entre persones, governs, facultats i organitzacions per l'ús intensiu de les tecnologies de la informació i la comunicació (TIC), que faciliten la creació, distribució i manipulació de la informació i exerceixen un paper essencial en les activitats socials, culturals i econòmiques.

Fa pràcticament cent anys ja vam viure un episodi de pandèmia amb la grip espanyola, malaltia que, del 1918 al 1919, va contagiar i que s'estima que va provocar la mort de 40 milions de persones al món. Una pandèmia que va rebre l'apel·latiu de grip espanyola perquè la pandèmia va ocupar una major atenció de la premsa a Espanya que a la resta d'Europa, ja que Espanya no estava involucrada a la Primera Guerra Mundial i, per tant, no es va censurar la informació sobre la malaltia. En conseqüència, ja estem veient com a Espanya, durant aquesta primera pandèmia, ja existia, en certa manera, un preludi de la gestió informativa per mantenir al corrent i salvaguardar la salut de la població, tot i que encara es basa en premsa escrita o butlletins del govern. Aquests mètodes de comunicació per a la població eren els més estesos, ja que altres mitjans com la ràdio o la televisió no

estarien disponibles per al públic en general fins anys més tard.

A la premsa escrita i els butlletins tenien i tenen els seus propis problemes a l'hora d'assolir la població, com són la mateixa bretxa educativa, l'accessibilitat dels continguts informatius per a persones amb discapacitat i l'accés als mitjans de comunicació en funció de la localització de les persones i de l'escalabilitat del procés de producció de les impremtes.

Imatge del Butlletí Oficial Extraordinari de la província de Burgos publicat l'any 1918 amb motiu de la grip espanyola. (Enllaç:<https://www.rtve.es/noticias/20200507/imagen-gripe-espanola-1918-archivoburgos/2013540.shtml>)



La pandèmia de COVID-19 i la intel·ligència artificial

Des d'aquesta darrera pandèmia hi va haver molts canvis tecnològics i socials, entre els quals l'ús d'Internet per accedir a la informació. A finals dels anys 90 és quan comencen a incrementar-se els serveis de banda ampla a les llars espanyoles, i s'inicia així una època de boom de les comunicacions digitals i de l'anomenada tercera revolució industrial. Una revolució que comporta una transformació de les relacions socioeconòmiques de la nostra societat, en la qual un ciutadà o una ciutadana ha d'estar connectat a una xarxa per rebre informació, adquirir productes o serveis globalitzats i relacionar-se amb les administracions públiques. Els governs de tot el món encoratgen la ciutadania a fer els passos per convertir-se en ciutadà i ciutadanes digitals, i les empreses i els organismes a transformar-se digitalment.

Tota transformació és positiva, mentre se'n faci partícip tota la societat i s'habilitin les mesures oportunes, siguin tecnològiques, econòmiques o educatives, per abordar possibles problemes que puguin deixar una part de la població aliena a les millores en la seva qualitat de vida generades per aquesta transformació. En aquest cas, la societat digital comporta qüestions que les administracions públiques han de tractar, com és l'anomenada bretxa digital, que afecta 3 eixos principals: la manca d'accessibilitat als continguts digitals, la manca d'una educació digital per a l'ús de les noves tecnologies i l'accés a internet, sigui per mancances del senyal o per l'accés a aquest a causa de manca de recursos.

Per posar remei al problema de l'accessibilitat digital a la informació, els col·lectius de persones amb discapacitat han treballat juntament amb els governs per elaborar lleis que assegurin que els mitjans s'adaptin perquè puguin rebre la informació digital, a través de les pàgines web, complint una sèrie de característiques o criteris de desenvolupament. Lamentablement, malgrat que cada cop hi ha més conscienciació sobre aquesta problemàtica, seguim amb empreses i organismes que encara no han aconseguit adaptar-se a aquestes lleis i incloure criteris en els seus desenvolupaments relacionats amb l'accessibilitat.

Per què la pandèmia del coronavirus ha estat un punt d'inflexió quant a informació digital?

I de sobte, a aquesta societat digital que pretenem ser i que encara està en construcció la colpejà la pandèmia del coronavirus el 2020. Una pandèmia que era la primera en què la població comptava amb informació digital, diferent de la pandèmia de la grip espanyola del 1918, encara que semblant davant el desconeixement de la societat de com actuar per protegir-se i controlar el virus.

"Tota transformació és positiva, mentre se'n faci partícip tota la societat i s'habilitin les mesures oportunes, siguin tecnològiques, econòmiques o educatives, per abordar possibles problemes que puguin deixar una part de la població aliena a les millores en la seva qualitat de vida generades per aquesta transformació."

nodes

La pandèmia de COVID-19 i la intel·ligència artificial



El govern i la població no tenien un manual de suport per a aquesta situació d'emergència, no hi havia cap protocol previ i no estàvem preparats per la manca d'informació a què ens enfrontàvem. Les notícies sobre la naturalesa del virus, les restriccions, l'estat de cada comunitat i els protocols de seguretat eren llançats a través de diversos mitjans, i els més utilitzats pel govern i per les autoritats sanitàries van ser la televisió i els canals digitals, com les pàgines web. Part d'aquesta informació patia molts canvis i la societat depenia plenament d'estar informats contínuament per estar al corrent dels canvis ràpids que anaven esdevenint amb referència a la pandèmia. A més, aquest cop ens trobem amb un problema nou, la desinformació generada per les fake news llançades a les xarxes socials, principalment, i la capacitat de la població per verificar tota la informació que rebia.

Aquests problemes relacionats amb l'accés a la informació digital afectaven molt més les persones amb discapacitats, per la manca d'accessibilitat i usabilitat de la informació als mitjans de comunicació digital, com ara els webs dels organismes públics i empreses. Comprendre la qüestió de la bretxa digital facilita la gestió de les dificultats i permet generar propostes de solucions viables que permetin resoldre-les. Conèixer de primera mà les dificultats associades a la manca d'accessibilitat i usabilitat de la informació va fer que l'equip de SemanticBots col·laborés amb l'Administració i amb diverses associacions de persones amb discapacitat, desenvolupant un projecte de xatbot inclusiu relacionat amb la Covid utilitzant la plataforma Sayobo, basada en intel·ligència artificial.

La pandèmia de COVID-19 i la intel·ligència artificial

Per què neix Sayobo i com es posa en marxa el projecte del xatbot de la Covid inclusiu?

SemanticBots som un spin-off de la Universitat Jaume I, cofundada per investigadors del grup de recerca de Bases del Co-neixement Temporal (TKBG), experts en intel·ligència artificial i big data. L'empresa de desenvolupament de plataformes converses es va fundar el 2017, i el 2018 vam participar en el projecte «Cap a una innovació responsable: Els parcs científics com a motor de canvi», desenvolupat per la Unitat de Cultura Científica i de la Innovació del Servei de Comunicació de la Universitat Jaume I, en què es buscava promoure la Recerca i Innovació Responsables (RRI, de l'anglès Responsible Research and Innovation) a través de la participació i el diàleg amb els diferents grups d'interès perquè els processos d'innovació i els resultats siguin sostenibles, desitjables i acceptables.

Com moltes altres empreses i equips de recerca, desco-neixíem la importància d'aplicar l'RRI al desenvolupament dels nostres dissenys i productes, però poder participar en aquest projecte ens va permetre adonar-nos que en tots els processos calia plantejar-se i preguntar a la societat com afectava positivament o negativament la nostra feina.



Tenint en compte aquestes premisses, vam començar a analitzar a quins grups de la societat podia millorar la qualitat de vida l'ús o l'accés a eines d'intel·ligència artificial com les que estàvem dissenyant. La nostra conclusió va ser que a tothom, ja que tothom necessita accedir fàcilment a la informació digital, però encara més a aquells grups

models

La pandèmia de COVID-19 i la intel·ligència artificial

“Tenint en compte aquestes premisses, vam començar a analitzar a quins grups de la societat podia millorar la qualitat de vida l'ús o l'accés a eines d'intel·ligència artificial com les que estàvem dissenyant.”

“Els reptes tècnics a resoldre per a aconseguir una interfície conversacional accessible per a persones amb discapacitats visuals van ser majors dels esperats.”

que ja eren vulnerables a la digitalització. Això ens va portar a focalitzar la nostra feina en les persones amb discapacitats visuals: més de 253 milions de persones tenen discapacitat visual al món.

L'equip de l'empresa va decidir començar una investigació sobre quines solucions i eines estaven a la disposició de les persones amb discapacitats visuals quan busquen informació o fan gestions a través dels canals digitals d'empreses i institucions. Durant aquesta investigació es va poder conèixer de primera mà els problemes a què s'enfronten diàriament les persones amb discapacitat visual gràcies a la col·laboració i els coneixements aportats tant per professors de tiflotecnologia d'ONCE Castelló i València, com de persones amb discapacitats visuals de Castelló i Gandia. Durant aquesta col·laboració, i tenint en compte els canvis suggerits pels professors i usuaris, es va redissenyar la plataforma conversacional inicial a una de nova amb les adaptacions pertinents, a la qual vam anomenar Sayobo. Aquest desenvolupament d'una plataforma conversacional inclusiva es va interpretar a finals del 2019 i el març del 2020, coincidint amb la pandèmia, vam aplicar aquest coneixement al desenvolupament d'una eina d'informació digital per a la societat.

Els reptes tècnics a resoldre per a aconseguir una interfície conversacional accessible per a persones amb discapacitats visuals van ser majors dels esperats. Molts d'aquests reptes van tenir a veure amb la falta de maduresa de les tecnologies accessibles al web (hi ha diferències notables entre sistemes operatius i navegadors a l'hora de fer accessibles interfícies molt dinàmiques com la d'un xatbot), i la variabilitat a l'hora d'interaccionar amb els distints lectors de pantalla que usen els usuaris amb discapacitats visuals en el seu dia a dia. També hi va haver qüestions més subtils, com proporcionar alternatives a respostes basades en informació visual: imatges i vídeos.

La iniciativa de crear un xatbot inclusiu sobre la Covid-19 per part de SemanticBots es va dur a terme després de conèixer la Direcció General de Lluita contra la Bretxa Digital, que pertany a la Conselleria d'Innovació, Universitats, Ciència i Societat Digital de la Generalitat Valenciana i la problemàtica existent en alguns col·lectius per poder accedir a la informació digital relacionada amb la situació excepcional, creada per la crisi sanitària.

Des de l'empresa es va accedir a participar activament en el desenvolupament del xatbot inclusiu sobre la Covid-19 i la seva cessió de manera gratuïta. L'objectiu de tots els col·laboradors del projecte no era altre que assegurar que la informació digital sobre la pandèmia arribarà fàcilment a tota la societat, incloent-hi les persones amb discapacitats i d'edat avançada. Sayobo va començar a incorporar-se a les pàgines web oficials de diverses associacions de persones amb discapacitat de la Comunitat Valenciana, com per exemple CERMI, COCEMFE i FESORD. Aquest sistema inclusiu es podia instal·lar de

La pandèmia de COVID-19 i la intel·ligència artificial

manera gratuïta a qualsevol pàgina web d'organismes, associacions i empreses per millorar-ne l'accessibilitat i facilitar la informació digital a tots els usuaris.

El xatbot era també accessible en múltiples plataformes i dispositius des de les pàgines web on estava instal·lat i permetia atendre tots els usuaris alhora, tant en castellà com en valencià, que ajudava i complementava els canals habituals de comunicació oficial. L'eina es va convertir en un suport informatiu més, sobretot quan estar ben informats era vital davant de la crisi sanitària.

Com es va desenvolupar el sistema del xatbot de la Covid inclusiu?

La tasca que va fer l'equip de l'spin-off va tenir diversos vessants. D'una banda, es va crear la base de coneixement inicial, que es va implementar per oferir a la societat un compendi de les principals consultes sobre la Covid. Aquest coneixement partia d'una tasca prèvia de documentació respecte a qüestions sobre símptomes, mesures d'higiene, protocols d'actuació, mesures laborals i drets socials, així com altres mesures derivades de la situació d'alarma provocada pel coronavirus. Les preguntes i les respostes procedien de fonts oficials nacionals, autonòmiques i internacionals, a més d'altres organismes com l'Organització Mundial de la Salut (OMS), articles científics i informació digital contrastada.

“L'eina es va convertir en un suport informatiu més, sobretot quan estar ben informats era vital davant de la crisi sanitària.”



El hàndicap de l'equip en aquest cas era la necessitat de mantenir actualitzat diàriament els canvis que succeïen respecte a les respostes relacionades amb la pandèmia i les mesures de seguretat. Tots recordem per exemple els canvis que es van produir en les diferents fases i pel que fa a l'ús de mascaretes. Per això, una de les activitats fonamentals de l'equip va ser estar informat constantment per fer els canvis oportuns, i adaptar la informació en tot moment. En aquesta primera fase, la base de coneixement inicial el sistema d'intel·ligència artificial era capaç de reconèixer i donar resposta a 400 consultes.

models

La pandèmia de COVID-19 i la intel·ligència artificial

“Les associacions són molt conscients del problema de la digitalització, de l’accessibilitat i de la usabilitat de la informació, i, per tant, entenen la tasca social i tecnològica conjunta que estàvem duent a terme, i això és determinant per a tots en la implicació en el projecte.”

Durant el procés de documentació i de concepció del xatbot vam advertir que el govern i les institucions estaven centrant molt els documents i la informació que elaboraven a la societat en conjunt, però quedaven sense cobertura qüestions que podrien afectar les persones amb discapacitat. Per aquest motiu, a l’hora de fer evolucionar la base de consultes del xatbot, cal comptar amb el suport d’algunes de les principals associacions de persones amb discapacitat de la Comunitat Valenciana i del país, CERMI, COCEMFE, FESORD i ONCE. Aquestes associacions, a més de donar-nos suport en la resolució de noves consultes associades a la discapacitat que arribaven a través del xatbot, també ens van aportar documentació per incloure en la base de coneixement perquè la informació sobre seguretat i salut davant del coronavirus estigués més fàcilment disponible per als usuaris. Les associacions són molt conscients del problema de la digitalització, de l’accessibilitat i de la usabilitat de la informació, i, per tant, entenen la tasca social i tecnològica conjunta que estàvem duent a terme, i això és determinant per a tots en la implicació en el projecte.

Tota la base de coneixement del xatbot va anar evolucionant. Se’n van monitorar contínuament les interaccions amb els usuaris per tal d’afegir respostes a noves qüestions que no s’havien previst, i per a incloure informació extreta de noves publicacions que realitzaven les distintes administracions, que va donar una cobertura total a 540 qüestions l’any 2022.

Quin és el present i futur de SemanticBots i Sayobo?

Els nous desenvolupaments de xatbots que hem portat a cap des del 2020 han estat implementats mitjançant l’ús de la plataforma Sayobo, incloent-hi sistemes d’intel·ligència artificial a Smart Cities, destinacions turístiques intel·ligents (DTI), serveis de les administracions públiques, educació i empreses. Cada vegada són més les persones que formen part dels organismes públics i privats els qui prenen consciència dels reptes en el camp digital i en les possibles solucions per aconseguir assolir aquests reptes, entre els quals es té en compte la bretxa digital. A aquesta creixent consciència han influït els plans estratègics de molts organismes per abordar els Objectius de Desenvolupament Sostenible de l’Agenda 2030 de l’Organització de les Nacions Unides, que és una oportunitat més perquè la nostra societat avanci perquè la qualitat de vida millori per a tots i totes, i no deixar ningú enrere.

Des de SemanticBots som també conscients que el nostre treball de recerca i desenvolupament, centrat en la intel·ligència artificial com a eina habilitadora, permetrà una digitalització més inclusiva. Aquesta tasca facilitarà que encarem, amb la tecnologia i la innovació, solucions que redueixin les desigualtats socials, millorin l’educació i ens permetin crear Smart Cities o ciutats més inclusives.

La pandèmia de COVID-19 i la intel·ligència artificial

Des de l'equip veiem que la intel·ligència artificial associada a l'accessibilitat és un camp on encara queden moltes coses per fer, i on podem aportar tot el nostre coneixement per transformar el futur de la nostra societat digital. Per crear o dissenyar el món del futur, cal comptar amb les persones en aquesta transformació, i entendre els problemes i les necessitats; si no ho fem, estarem cometent els mateixos errors en què vam caure en el passat com a societat.



nodes

El fracàs de Radar COVID

Vicent Costa

“La situació de crisi generalitzada, causada pel SARS-CoV-2, ha estat el sotrac de la nostra època.”

El nou d'octubre de l'any 2022, la mascletà del migdia celebrada a la Plaça de l'Ajuntament de València culminava les celebracions relatives a la Diada Nacional del País Valencià. Aquell mateix dia, Radar COVID, l'aplicació per a mòbils dissenyada per la Secretaria d'Estat de Digitalització i Intel·ligència Artificial, cessà la seua activitat. Ara bé, cal dir que ho va fer lluny dels focs d'artifici i les maneres valencianes. Dos-cents vint caràcters d'una piulada del compte oficial de l'aplicació anunciaven la fi de l'operativitat de Radar COVID, i agraïen les persones que havien confiat en ella. Això sí, el missatge es llegia amb l'eco de qui constata una derrota quan ja s'ha calat foc a les naus i no queda ningú que faça atenció a la mar; i així ho van constatar les iròniques respostes de les usuàries i el to emprat per la premsa. Al remat, i deixant de banda la tragèdia, anunciar la cloenda d'una festa que no ha venut cap entrada té un punt de còmic.

En aquest treball, presentem una reflexió breu arran de la fi de Radar COVID; i ho fem amb la intenció de plantejar preguntes obertes i obrir nous debats, abans no hàgem de tornar a fer ús d'una aplicació semblant.

Breu introducció a Radar COVID

La situació de crisi generalitzada, causada pel SARS-CoV-2, ha estat el sotrac de la nostra època. L'apocalipsi tantes vegades augurada arribà, encara que val a dir que no esperàvem mesurar l'enemiga en dècimes de micra. La ciència i la tecnologia (i també la intel·ligència artificial en particular) han tingut, i encara tenen, un paper central en aquesta nova realitat, incloent-s'hi igualment les mesures sanitàries per fer front al virus. Radar COVID ha estat un exemple d'aquesta participació en la lluita. En efecte, de manera semblant a altres sistemes presentats en països com Alemanya o Suïssa, aquesta aplicació permetia el que comunament s'anomena rastreig de contactes, això és, un registre de les persones amb les quals ens hem creuat. Així, la informació recollida a partir del rastreig s'utilitzava per avaluar el risc de contagi de les usuàries. D'aquesta manera, una usuària rebia un avís de contacte si havia compartit espai durant el temps estipulat amb una persona que haguera resultat infectada. A més, Radar COVID donava la possibilitat d'informar de manera anònima un diagnòstic positiu, i fins i tot oferia l'opció de comunicar l'exposició de manera anònima a les persones amb qui la usuària havia estat en contacte. L'aplicació, per cert, pressuposava una màxima en aquests temps de dependència tecnològica, això és, la de no deixar-nos el mòbil quan eixim de casa.

La pandèmia de COVID-19 i la intel·ligència artificial

Radar COVID es llençà el vint-i-nou de juny de l'any 2020, mitjançant Google Play, i immediatament després es feu una prova pilot a La Gomera. El seu ús més general començà dos mesos després, i Catalunya fou el darrer territori en activar aquesta aplicació, el vint-i-set d'octubre del mateix any. Dos anys i poc després del seu llançament, tal com ja hem mencionat al principi d'aquest escrit, l'aplicació deixà d'estar operativa. Tenint en compte les dades oficials de participació recollides, es confirma que Radar COVID ha estat, en termes de participació i rellevància, un fracàs, car l'aplicació notificà poc més de l'1% dels contagis que es van produir durant el seu període d'activitat.



Què fem amb la privadesa?

Quan l'aplicació per a mòbils Radar COVID aterrà, la cultura occidental teixia, molt a poc a poc i amb alts i baixos, un procés de reconciliació amb la privadesa. D'una banda, la història ja havia il·lustrat els perills de l'ús abusiu de les dades. A tall d'exemple, podríem pensar en l'ús de les dades per part del règim nazi durant la Segona Guerra Mundial. Aquest desafortunat fet històric, per cert, el presenta de forma magnífica Carissa Véliz al llibre que signa amb Cecilio Angulo, Intel·ligència artificial i tecnoètica (Angulo & Véliz, 2022, pp.69—74). De l'altra, la filosofia continental, especialment la del segle passat, havia argumentat que la informació és indissoluble amb el poder, de manera que calia tenir-ne cura de no deixar-la en mans de qualsevol. I si bé els primers anys del segle XXI aturaren aquest procés (principalment, a causa del furor per la generalització de la xarxa i del capitalisme de la vigilància impulsat pel terrorisme en territori occidental), es podria dir que, a principis de l'any 2020, la societat estava «reaprenent el valor de la privacitat a cop de males experiències» (Angulo & Véliz, 2022, p.31).

D'una revolada, però, el món va canviar: un enemic letal arribà i no només calia actuar, sinó que també s'havia de donar la imatge que la presa de decisions no s'aturava. D'entre la llista d'accions per portar a terme, la idea d'utilitzar la tecnologia per rastrejar els contagis semblava una opció lògica i, a més a més, es pressuposava la seua eficàcia. Amb tot, restava la qüestió, gens trivial, sobre com integrar, en el context d'una ciutadania cada vegada més conscient amb la privacitat, el discurs a favor d'una cessió de dades tan evident.

A parer nostre, dos eixos vertebraren el discurs institucional. Per un costat, el valor del sacrifici que suposava l'acceptació del rastreig de contactes es presentà com un afer vital, quasi per causa major, perquè cedir una part del dret a la privacitat era una manera de contribuir a guanyar la batalla contra l'enemic invisible. Per l'altre, l'ús de l'aplicació Radar COVID es va plantejar com una part de

La pandèmia de COVID-19 i la intel·ligència artificial

l'excepcionalitat de l'època, la qual, tal com ha d'ocórrer amb tota excepcionalitat, se suspendria en un futur pròxim. Ara bé, arran de la fi de l'operativitat de l'aplicació, podem concloure que l'estratègia no fou encertada.

En primer lloc, sembla que la ciutadania trobà prou ingenu haver de creure que els governs (i de passada les grans empreses) no es veurien temptats de fer ús de dades com les que recollia l'aplicació. De fet, no calia erigir escenaris imaginaris o pensar en els períodes més obscurs del segle passat, puix que hi havia successos recents gens falaguers: des d'espionatge selectiu a persones del món de la política fins a d'altres amb més abast, com per exemple el cas de la xarxa geosocial de cites Grindr (Senabre & Costa, 2021). En efecte, des de fa anys, els usuaris d'aquesta xarxa disposen de l'opció d'indicar el seu estat serològic (això és, l'usuari té la possibilitat d'indicar si és portador del VIH) i el temps transcorregut des de la darrera anàlisi. L'abril de l'any 2018, l'organització noruega sense ànim de lucre SINTEF informà que els paquets de dades que Grindr cedia a tercers podien contenir l'estat serològic dels usuaris (en realitat, aquests havien acceptat els termes i les condicions de l'ús de l'aplicació, els quals informaven sobre aquesta possibilitat de cessió).

En aquest sentit, els discursos institucionals sobre l'aplicació Radar COVID afirmaven que el flux de dades era anònim i segur; el mes de setembre de l'any de llançament, però, l'Agència Espanyola de Protecció de Dades sancionà la Secretaria d'Estat de Digitalització i Intel·ligència Artificial perquè l'aplicació incomplia vuit normes del Reglament general de protecció de dades. Al capdavall, les sospites i preocupacions sobre les amenaces a la privadesa no van ser tan infundades com es podia creure en un primer moment. Altrament, pareix que va haver-hi una fallida de perspectiva. La intel·ligència artificial manté una relació constitutiva amb la societat en què es desenvolupa, i aquesta relació s'estén als enginys que la disciplina dissenya. En particular, la societat on aterrava Radar COVID, prou conscienciada

amb la importància de la privacitat, requeria una veritable aposta per la transparència de l'aplicació i un discurs d'acord amb aquest compromís. Previsiblement, deu anys enrere no hauria calgut insistir en aquest aspecte, això és, el de la privadesa; l'any 2020, emperò, la ciutadania necessitava més garanties.

En segon lloc, Radar COVID potser s'entenia com un pas més enllà d'un ús abusiu i eventual, i ressemblava a un control orwel·lià reservat fins aleshores a aquelles persones que, bé per haver comés un crim o bé per córrer un risc, accepten un cert grau

“En primer lloc, sembla que la ciutadania trobà prou ingenu haver de creure que els governs (i de passada les grans empreses) no es veurien temptats de fer ús de dades com les que recollia l'aplicació.”

“La intel·ligència artificial manté una relació constitutiva amb la societat en què es desenvolupa, i aquesta relació s'estén als enginys que la disciplina dissenya.”



La pandèmia de COVID-19 i la intel·ligència artificial

de geolocalització. Malgrat que l'excelsitud i la gravetat de la situació eren palmàries i pareixien justificar qualsevol mesura, a la fi, l'impacte i l'abast de l'aplicació per a mòbils foren nimis, i la singularitat de les circumstàncies no fou un argument suficient per a oblidar aquestes preocupacions de la ciutadania. Si ens preguntem els motius d'aquesta insuficiència, podríem destacar, d'una banda, que l'eficàcia de mesures com Radar COVID no estava garantida. El vint-i-quatre de març de l'any 2020, Carissa Véliz escrivia en El País:

«És necessari cedir la nostra privadesa a una app per frenar el coronavirus? No és evident. Les mesures més efectives no semblen passar per cap app. (...) L'opció de recórrer a apps no només és més invasiva des del punt de vista de la privadesa, sinó també molt menys precisa i efectiva [que la quarantena].» (Angulo & Véliz, 2022, pp. 40—41).

En efecte, sense evidència sobre l'efectivitat de l'ús de l'aplicació, justificar i, sobretot, motivar la cessió es tornava complicat, especialment en el context d'una societat civil coneixedora de la importància del dret a la privacitat. Tenint en compte els resultats finals sobre l'ús de Radar COVID, pareix que aquesta va ser una idea compartida per gran part de la població.

D'altra banda, cal accentuar la manca de claredat respecte a la durada d'aquesta mesura excepcional. L'horitzó d'una era autoritària en la qual el poder polític tindria accés a dades com les que recull l'aplicació Radar COVID requeria preguntes com les plantejades per Cecilio Angulo al llibre que hem esmentat anteriorment:

«Un cop instaurada la política del Gran Germà, de la col·laboració necessària, sempre "per un bé comú", qui desfarà el camí que ja s'ha fet? Qui m'assegura la desconexió del joc?» (Angulo & Véliz, 2022, pp. 36—37).

Al remat, sacrificar o limitar drets, tot i que temporalment, és un passeig per un estímul; i no conèixer la durada de la caminada és una incertesa amb flaïre de tempesta. Perquè els drets, com la confiança, es guanyen de mica en mica i es perden en un batre d'ulls.

Com a cloenda, interdisciplinarietat i esforç col·lectiu

Sovint sentim a parlar de la importància d'integrar les humanitats i les ciències socials en el desenvolupament de la intel·ligència artificial, i el cas de l'aplicació per a mòbils Radar COVID n'és un bon exemple, d'aquesta necessitat. Tal com hem argumentat, alguns dels problemes relacionats amb la privacitat reuneixen factors que pot tenir a veure amb els aspectes tècnics del disseny de l'aplicació. Al capdavall, el fet que l'anunci de la fi de Radar COVID mencionara la confiança no va estar una qüestió d'atzar. Acabem, doncs, aquesta reflexió amb dues propostes breus de cara a futures aplicacions semblants a la de Radar COVID.

“Perquè els drets, com la confiança, es guanyen de mica en mica i es perden en un batre d'ulls.”

La pandèmia de COVID-19 i la intel·ligència artificial

Per un costat, Radar COVID requeria una cooperació majoritària. Es tractava d'una aplicació de caràcter massiu, i l'objectiu d'evitar contagis només es podia donar si el seu ús no era marginal. Així doncs, comprometre la ciutadania no era, evidentment, un projecte senzill, ja que implicava tenir en compte, entre d'altres, la diversitat social i les motivacions, les dificultats i les circumstàncies personals de cada ciutadana. De fet, hi ha infinituds d'exemples de serveis públics que han necessitat estudis interdisciplinaris (incloent-hi especialment les ciències socials i l'antropologia) per tal d'augmentar el seu ús col·lectiu. I així, amb freqüència, la solució en aquests casos passa per canvis que tenen poca relació amb el disseny de la infraestructura. Per tant, i en vista dels resultats de Radar COVID, seria adient un estudi interdisciplinari per tal d'esclarir els punts febles i les raons del fracàs pel que fa al compromís general amb el seu ús.

“Perquè, comptat i debatut, potser la lliçó principal que podem treure d'aquest sotrac és justament el valor de no deixar ningú enrere.”

Per l'altre costat, en relació amb el disseny d'aplicacions com Radar COVID, és menester recordar que s'ha de tenir una cura especial de les discriminacions que es podien donar amb el seu ús, especialment aquelles relacionades amb l'accessibilitat: «aquesta infraestructura digital [Radar COVID] no sols s'ha de dissenyar per a tota la ciutadania, sinó que ha de ser accessible per a tots tipus d'usuaris» (Senabre & Costa, 2021, p.73). Al remat, la perspectiva ètica també ha d'integrar-se en el mateix disseny de l'aplicació i la inclusió és un valor bàsic i innegociable. En aquest sentit, defensem que la noció de bretxa digital, la qual ha assolit matisos i perspectives substancials els darrers vint anys, és un bon punt de partida per tal de definir un marc que guie un disseny inclusiu d'aplicacions d'aquest tipus. Així, a priori, en aquest cas el concepte teòric centraria l'atenció en dos dels col·lectius que més han patit les conseqüències de la crisi sanitària, això és, les persones grans i les persones amb discapacitat. Pel que fa al darrer, s'hi va haver d'afegir canvis per millorar l'accessibilitat d'aquest col·lectiu (això sí, els canvis arribaren quasi un any després del llançament de l'aplicació); mentre que les persones grans necessitaren l'ajuda del personal de les farmàcies per fer ús de l'aplicació. En un futur, seria convenient replantejar la integració de la inclusió en el disseny d'aquestes aplicacions i fer-ne un gir que, en comptes de concebre un disseny estàndard (per a usuàries estàndards) i d'afegir després els canvis pertinents per als altres col·lectius, reconduísca el disseny i el centre, des del principi, en les usuàries amb necessitats especials (ben segur que, amb aquesta estratègia, la resta de les usuàries s'hi adaptarà amb facilitat). Perquè, comptat i debatut, potser la lliçó principal que podem treure d'aquest sotrac és justament el valor de no deixar ningú enrere.

La pandèmia de COVID-19 i la intel·ligència artificial

Referències bibliogràfiques

- **Angulo, C.; Véliz, C.** Intel·ligència artificial i tecnoètica. Iniciativa Digital Politècnica. Oficina de Publicacions Acadèmiques Digitals de la UPC, 2022.
- **Senabre, E.; Costa, V.** Intel·ligència artificial. Com els algorismes condicionen les nostres vides. València: Sembra Llibres, 2021 (segona edició).

nodes

La pandèmia de COVID-19 i la intel·ligència artificial

Què en pensen?

Josep Puyol Gruart

"L'enquesta consistia en quatre assercions i s'havia de triar entre cinc graus d'acord i, opcionalment, deixar un comentari."

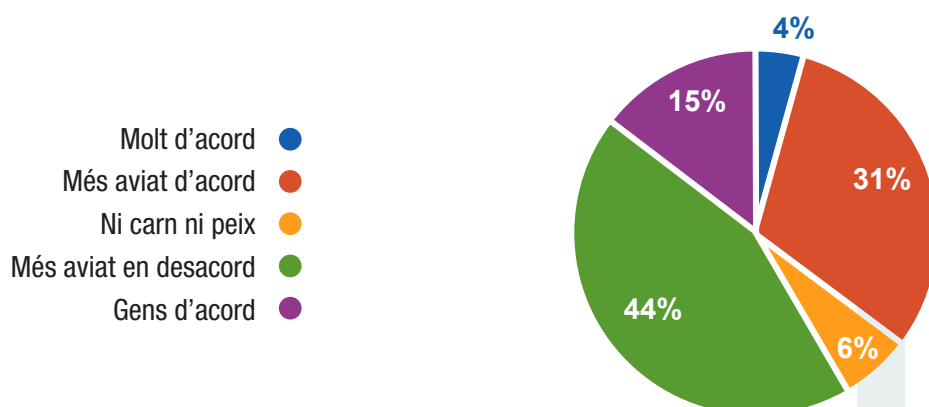
A l'Associació Catalana d'Intel·ligència Artificial (ACIA) vam dissenyar una petita enquesta informal, sense pretensió científica, per conèixer l'estat d'opinió sobre alguns temes que relacionen la IA amb la recent epidèmia de COVID-19, tema principal d'aquest número 59. Encara que no ens podem deslliurar de relacionar-ho també amb la qualitat democràtica (veure NODES58).

Com que ja va patir per la nostra insistència perquè omplís aquesta petita enquesta, i tal com vam prometre, en publiquem els resultats. Gràcies per la vostra participació.

L'enquesta consistia en quatre assercions i s'havia de triar entre cinc graus d'acord i, opcionalment, deixar un comentari. També opcionalment es podia fer un comentari general sobre el tema d'IA i COVID-19.

Tenim 48 enquestes contestades i aproximadament el 35% hi heu fet comentaris, en total 79. L'anàlisi de l'enquesta tampoc té cap pretensió especial i ens limitarem a comentar els resultats i emfatitzar algunes de les vostres opinions que ens semblen rellevants, sense arribar a conclusions. Pels més agosarats, els resultats (anònims) complets els podeu trobar en un fitxer CSV a: <https://www.acia.cat/EnquestaNODES-2023.csv>

1 La IA està contribuint d'alguna manera al descrèdit científic. A causa de la pandèmia, molta gent veu IA pertot: més control de la població, dubtes sobre les motivacions de les campanyes de vacunació, més invasió de la intimitat, etc. Tot això només fa que fer créixer l'escepticisme científic.

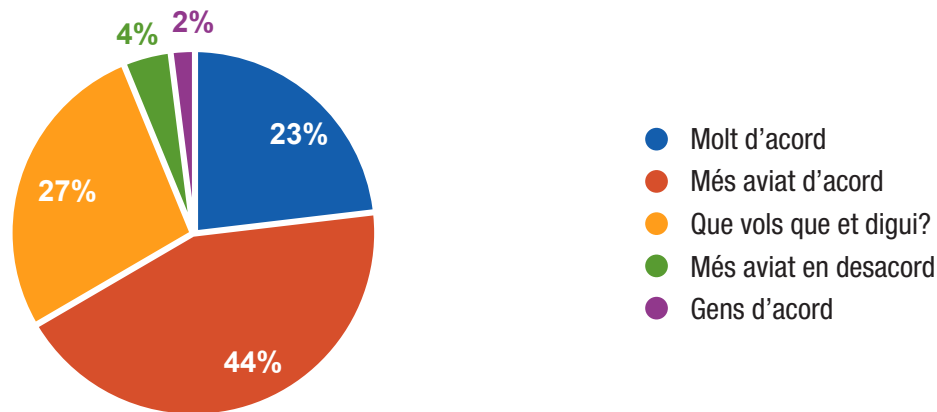


El 59% dels enquestats no estan d'acord que la IA contribueix al descrèdit científic, i un 15% d'aquests fan èmfasi que no hi estan gens d'acord. Un 35% sí que hi estan d'acord, i un 4%, clarament.

La pandèmia de COVID-19 i la intel·ligència artificial

Mirant els 18 comentaris, ens trobem que els que estan més en contra de l'enunciat ens deixen frases com ara [... *les promeses de la IA ens fan pensar en un poder infinit que ens resoldrà els problemes de la humanitat en una forma de transhumanisme...*], que la IA és un [... *concepte «calaix», i que moltes vegades els mitjans de comunicació «venen» com a IA coses que no tenen perquè ser-ho ...*] o que [... *la gent percep la IA com una eina molt potent i versàtil ...*]. Els més favorables ens diuen coses com que per a [... *bona part de l'opinió pública IA sigui sinònim de «big brother»*], que [... *la reflexió crítica de la IA sovint s'acompanya d'un escepticisme científic que desvia l'atenció...*] o que [... *La gent té reticències perquè els discursos al voltant de la IA són o massa triomfalistes o massa conspiranoics. No ajuden ni els mitjans de comunicació ni el relat que la IA (o la digitalització) ho pot solucionar tot.*]

2 Totes les informacions sobre la COVID-19 han fet sorgir campanyes més o menys orquestrades, o de vegades ingènues, de desinformació. Aquesta experiència portarà un avenç espectacular en la lluita contra la desinformació utilitzant tècniques d'IA.

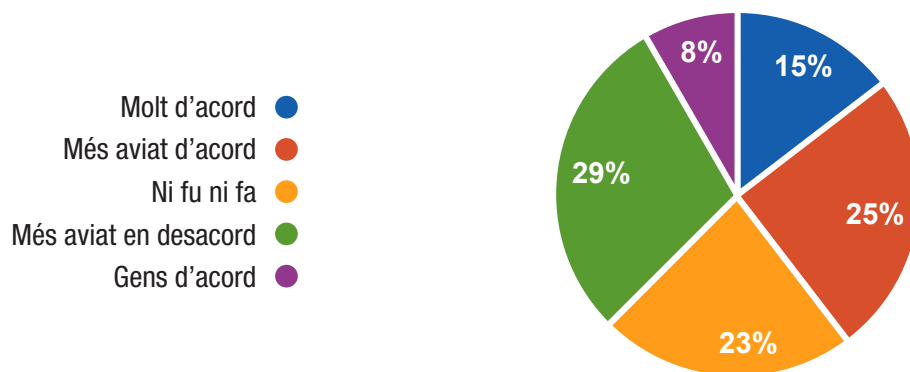


Els nostres enquestats són molt optimistes quant a la lluita contra la desinformació per part de la IA. Un 67% ho veuen factible i entre aquests, un 23% hi està molt d'acord. Hi ha molt indecisos, un 27%, i molt pocs que ho veuen inviable, un 2%.

Pel que fa als 16 comentaris, són interessants els dels que no estan en absolut d'acord, per exemple, un diu que [*La IA ajudarà a la desinformació, o a combatre-la ... Són les xarxes i la falta de regulació que no posa fre a unes organitzacions ... que no tenen escrúpols.*] o [*... ha quedat clar en aquest període la utilització Orwel·liana del terme «desinformació» per combatre la llibertat d'expressió quan aquesta qüestiona els interessos de les autoritats...*]. Tot i que un que hi està més aviat d'acord diu que [... *massa tecno-optimisme hi veig...*], curiosament els que hi estan totalment d'acord no fan cap comentari.

La pandèmia de COVID-19 i la intel·ligència artificial

3 Alguns governs han utilitzat l'excusa de la pandèmia per fer un control i una vigilància de la població a uns nivells mai sospitats. Aquest és un moviment imparable i molt perjudicial per a la democràcia que perjudicarà els científics.



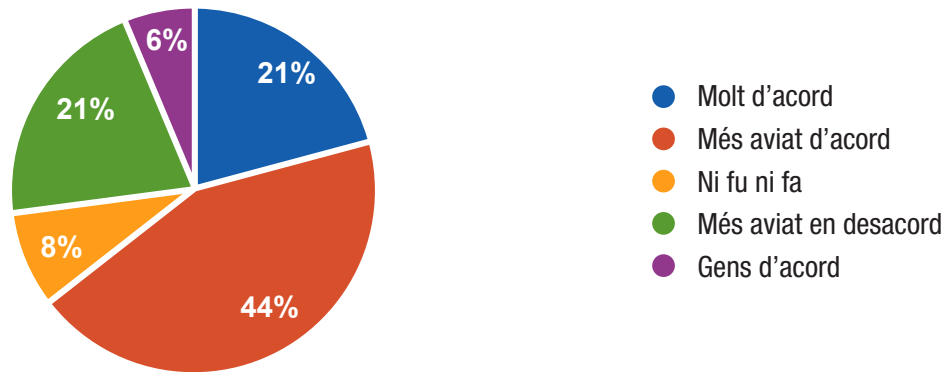
En aquesta pregunta els resultats són molt dispersos, potser degut al fet que l'afirmació no era prou clara. Un 23% no en tenen una idea clara, els que hi estan mitjanament d'acord o en desacord, molt similar (25% i 29% respectivament). Si hi sumem els més radicals, ens queda un 40% d'acord i un 37%, no.

Dels 21 comentaris podem veure que, independentment de la resposta, molts insisteixen que el perjudici pot ser per la democràcia i per amplis sectors de la població, i que no necessàriament implicaria només els científics. Malgrat no estar gens d'acord amb la frase, algú assegura que alguns moviments [... demostren el poder negatiu de la desinformació, la desigualtat i les dretes antidemocràtiques.]. Entre els que hi estan bastant d'acord, també es comenta que [L'abast del control que possibiliten els avenços en IA és un assumpte d'urgència, a causa de la seua perillositat, i difícil d'aturar si tenim en compte les dinàmiques dels sistemes democràtics actuals.]

"Dels 21 comentaris podem veure que, independentment de la resposta, molts insisteixen que el perjudici pot ser per la democràcia i per amplis sectors de la població, i que no necessàriament implicaria només els científics."

La pandèmia de COVID-19 i la intel·ligència artificial

4 S'han produït avenços espectaculars en la indústria farmacèutica gràcies a la IA. Això farà que la població reconegui la importància incontestable de la ciència per a l'avenç de la humanitat.



Aquí les respostes dels enquestats són més clares. Un 65% és favorable a l'afirmació i només un 27% s'hi mostra en desacord.

Els 18 comentaris són molt diversos, encara que molts comenten que no tenen clar que la població en general relacioni la IA amb la indústria farmacèutica. Un diu que *[Falta feina de comunicació i divulgació per al públic general, perquè siguin conscients d'aquests avenços]*. Surten entre els comentaris termes com "crítica a la ciència", "ceguesa habitual", "conspiranoics", "antivacunes", etc. També ens trobem comentaris apassionats com que *[... la «ciència» és utilitzada per les autoritats de la mateixa manera que a l'edat mitjana s'utilitzava la «religió», com una forma acrítica de fer callar les veus que qüestionen les seves accions.]* o un que es posa *[... en dubte l'avenç de la humanitat en el panorama actual d'auge de la ultradreta i de crisi mediambiental.]*

"Els 18 comentaris són molt diversos, encara que molts comenten que no tenen clar que la població en general relacioni la IA amb la indústria farmacèutica."

Comentaris generals

Respecte dels comentaris sobre la relació de la IA i la COVID-19, només en tenim set, variats. Els més sincers són un que diu que aquesta relació [no la veig, la veritat] i una esmena a la totalitat [No entenc a què respon o quin objectiu té aquesta enquesta]. Altres sí que hi veuen una relació, com un que diu que a causa de la COVID-19 [... la recerca sobre educació a través de plataformes digitals s'ha disparat...] o que ha sigut [... una bona eina per assistir en moments claus de la pandèmia (sobretot en l'àmbit de la logística als hospitals i UCIs)...]

Gràcies pel vostre interès!

nodes

Development of Machine Learning Techniques for Diabetic Retinopathy Risk Estimation

Emran Saleh Ali Ali

Directors/es: Aida Valls, Antonio Moreno, Pedro Romero
Universitat Rovira i Virgili. Departament d'Enginyeria Informàtica i Matemàtiques

“L’objectiu principal d’aquesta tesi doctoral és la construcció d’un sistema de suport a la decisió clínica (Clinical Decision Support System, CDSS) que utilitzi només dades del registre electrònic de salut, com ara dades fisiològiques i d’anàlitziques.”

La retinopatia diabètica (RD) és una malaltia crònica. És una de les principals complicacions de la diabetis, que afecta la vista i pot arribar a produir ceguesa. La pateixen al voltant d’un 10% dels diabètics. Per aquest motiu, les persones diabètiques han de fer-se periòdicament proves específiques per tal de detectar signes de desenvolupament de la retinopatia en una fase inicial. Aquesta prova consisteix a fer una fotografia del fons de l’ull amb un retinògraf. Posteriorment, un especialista ha d’anàlitzar la imatge i buscar signes de lesions. Aquest cribratge precoç disminueix el risc de pèrdua de visió, ja que permet iniciar el tractament quan la malaltia es pot controlar. Ara bé, el nombre de pacients diabètics està tenint un creixement molt ràpid, de manera que es fa difícil fer-los aquestes proves anualment, ja que consumeix massa recursos, tant de personal com de temps. Tenint en compte la baixa incidència de la malaltia (10%), aquest sistema de detecció no és eficient.

L’objectiu principal d’aquesta tesi doctoral és la construcció d’un sistema de suport a la decisió clínica (Clinical Decision Support System, CDSS) que utilitzi només dades del registre electrònic de salut, com ara dades fisiològiques i d’anàlitziques. El sistema ha d’ajudar a estimar el risc de desenvolupar RD. Aquesta eina la podran utilitzar els metges de capçalera per poder determinar si cal derivar el pacient a la prova del retinògraf o no. La tesi forma part d’un projecte de recerca finançat pel Instituto de Investigación Carlos III (amb fons FEDER) i hi participen el grup de recerca ITAKA (Intelligent Technologies for Advanced Knowledge Acquisition) de la Universitat Rovira i Virgili i el



grup d’oftalmologia de l’Hospital Universitari Sant Joan de Reus, a través de l’Institut d’Investigació Sanitària Pere Virgili.

Durant la tesi s’ha disposat d’un conjunt de dades de més de 3000 pacients, classificats segons si tenen retinopatia diabètica o no (classificació binària). En aquesta tesi doctoral s’han estudiat diversos mètodes d’aprenentatge automàtic supervisat per construir

un CDSS basat en regles lingüístiques difuses. El coneixement expressat en aquest tipus de regles facilita que el metge sàpiga quines combinacions de les condicions són les que poden provocar el risc de desenvolupar RD.

Per començar s'han usat algorismes d'aprenentatge per construir un arbre de decisió difús (Fuzzy Decision Tree, FDT). S'ha partit d'algorismes existents basats en ID3 per distingir les dues classes. Donat que a vegades les regles difuses no diferencien suficientment les dues classes, la tesi proposa una millora que permet etiquetar pacients com a "no classificats" (unknown), per evitar així cometre errors de classificació (falsos positius o falsos negatius). Els resultats obtinguts donen una especificitat i una sensibilitat del 80%. Malgrat ser millors que altres mètodes de classificació, no són prou bons per ser utilitzats en un entorn real. Cal destacar que el model difús ens permet determinar el grau de pertinença de cada persona a la categoria de DR, que és 0 lliure de risc, i 1 amb risc de DR. Els valors intermedis ens indiquen el nivell de confiança en el risc de desenvolupar DR, i amb això el sistema calcula la data de la propera visita per tornar a mesurar el risc, entre 6 mesos i 2 anys.

"Donat que a vegades les regles difuses no diferencien suficientment les dues classes, la tesi proposa una millora que permet etiquetar pacients com a "no classificats" (unknown), per evitar així cometre errors de classificació (falsos positius o falsos negatius)."

The screenshot shows the RETIPROGRAM interface. On the left, there is a form for patient information with fields for NHC, Nom, Sexe (Dona), Edat (63 anys), EVOL (14), TTM (Se suministra insulina), HbA1c (7.3), HTAR (Hipertensió mal controlada), CKDEPI (99), MA (772), and BMI (30). A 'Calcular' button is at the bottom of the form. On the right, a blue box displays the risk assessment results: '¿EXISTEIX RISC? sí', 'SEGÜENT VISITA: Necessita revisió amb un Oftalmòleg', and 'CERTESA: 38%'. Below this is a horizontal bar chart with a green segment and a red segment. Two buttons are present: 'El resultat es correcte' (green) and 'El càlcul sembla ser incorrecte' (red). At the bottom of the blue box, there is an 'Observacions' field and a 'Desar les dades' button. The version number 'versió 1.1' is visible at the bottom center.

Figura 1. Interfície gràfica del RETIPROGRAM a l'Hospital Universitari Sant Joan de Reus

Per millorar els resultats anteriors, la tesi estudia la utilització de múltiples arbres usant la tècnica dels Fuzzy Random Forests (FRF). S'utilitzen entre 100 i 300 arbres per fer la classificació. Una etapa molt rellevant és la de la fusió dels resultats de cada arbre. A la tesi es proposen diverses tècniques d'agregació. Per millorar la decisió

Premi Marc Esteva Vivanco 2021

final dels nostres models, es proposen tres mesures difuses que permeten ponderar les regles i fer una agregació mitjançant integrals de Choquet i Sugeno. La definició d'aquestes mesures difuses es basa en els valors de confiança de les regles. En particular, una n'és una mesura difusa que aprofita la jerarquia dels arbres i els valors dels diferents nodes. Amb aquests mètodes s'ha aconseguit millorar la qualitat de la classificació i obtenir una especificitat del 82% i una sensibilitat del 83%.

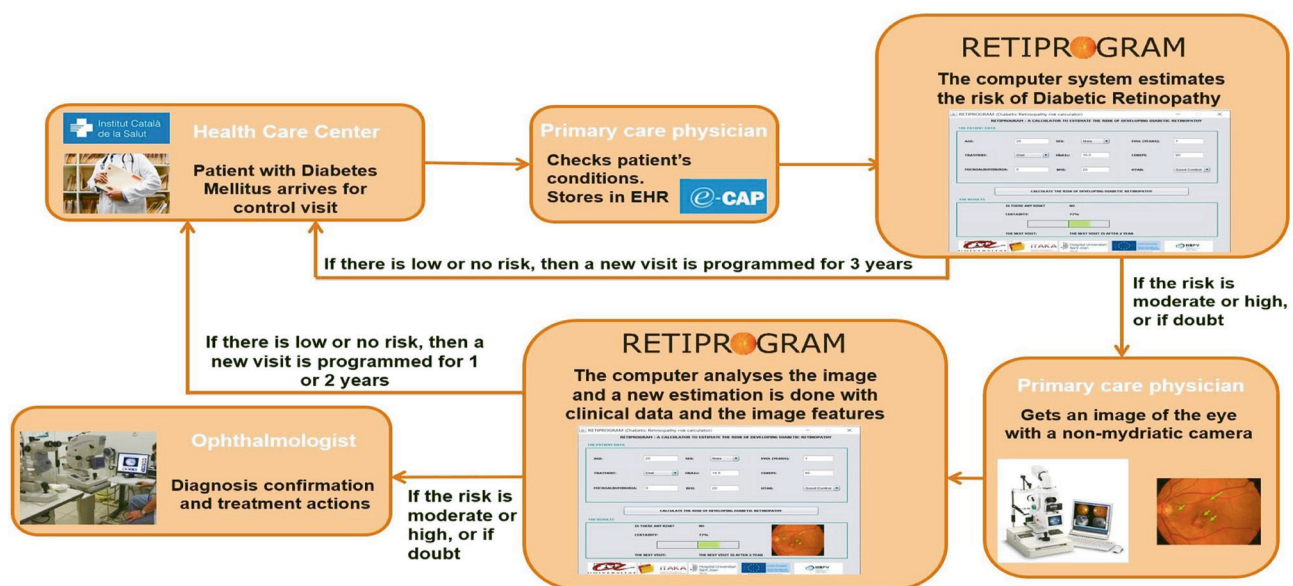


Figura 2. Ús del RETIPROGRAM en el cribatge de Retinopatia diabètica.

El resultat final de la recerca feta ha donat lloc a un programari anomenat RETIPROGRAM. Aquest software està llicenciat i és explotat per part de la companyia Up2Smart. Actualment, es troba en fase de prova pilot a l'Hospital Universitari Sant Joan de Reus, on s'està utilitzant per l'equip d'oftalmologia. El recull de noves dades permetrà fer una validació més exhaustiva del sistema. Esperem que en un futur es pugui instal·lar en centres d'assistència primària de tot Catalunya per ser usat pels metges de capçalera per fer l'avaluació preventiva i el cribatge de la Retinopatia Diabètica.

El projecte de recerca inclou també una segona fase en què la imatge capturada pel retinògraf s'analitza de forma automàtica usant visió per computador i les característiques observades es podran integrar en el RETIPROGRAM i utilitzar-les per afinar la predicció.

Adapting robot behavior to user preferences in assistive scenarios

Introducció i motivació

La robòtica assistencial pot jugar aviat un paper essencial en el benestar de les persones. Aquest tipus de robots poden allargar la independència de la gent gran, de persones amb mobilitat reduïda, i també d'aquells que necessiten ajuda per fer algunes de les tasques més bàsiques i necessàries per viure. Aquesta extensió de la independència té el potencial de millorar el benestar psicològic dels usuaris, que tenen menys sentiments de càrrega i culpa, atès que el robot serveix com a eina per habilitar aquests usuaris a acomplir tasques del dia a dia. Aquests robots assistencials poden ser socials, en què el robot ajuda per mitjà d'interaccions socials; o físics, en què el robot ajuda la persona en tasques que impliquen contacte entre la persona i el robot.

Malgrat això, hi ha múltiples formes diferents d'assistir una persona per realitzar tasques del dia a dia. Els cuidadors (humans), per exemple, solen tenir en compte les preferències i necessitats de la persona dependent per tal de fer-los la tasca més agradable. Per tant, aquests robots assistencials també haurien d'adaptar el seu comportament a les necessitats, capacitats i preferències específiques de cada individu, en comptes de fer que sigui la persona qui s'adapti al robot. En aquest cas, veiem els robots com a eines intel·ligents i no pas com a substituïts del cuidador humà.

El principal objectiu de la tesi és, doncs, avançar en el camp de la personalització dels robots autònoms i de l'adaptació del seu comportament, fent èmfasi en les tasques assistencials físiques. Per aconseguir-ho, hem desenvolupat mètodes per adaptar el comportament del robot de forma autònoma, de manera que coincideixi amb les expectatives de la persona assistida.

Els mètodes desenvolupats a la tesi mostren l'habilitat del robot per adaptar-se a les preferències de l'usuari. Els apliquem principalment a tres tasques assistencials: donar menjar, posar una sabata i posar una jaqueta, tal com es mostra a la Figura 1.

Recerca desenvolupada

La tesi comença definint les bases per a la personalització de robots amb un framework de personalització per a robots assistencials, que consisteix de tres fases: programació a fàbrica, adaptació a casa i execució adaptada. Aquesta metodologia permet adaptar el comportament del robot a usuaris específics mentre es duu a terme

Gerard Canal

Directors/es:
Guillem Alenyà, Carme Torras
Institut de Robòtica i Informàtica
Industrial (IRI), CSIC-UPC

"El principal objectiu de la tesi és, doncs, avançar en el camp de la personalització dels robots autònoms i de l'adaptació del seu comportament, fent èmfasi en les tasques assistencials físiques."



Premi Marc Esteva Vivanco 2020



Figura 1. Tasques assistencials físiques. Se'n poden trobar vídeos www.iri.upc.edu/thesis/show/103

la tasca assistencial, i es pot iterar tantes vegades com calgui. A la tesi, exemplifiquem el mètode mitjançant la tasca de péixer, on mostrem com un usuari no expert pot modificar el comportament del robot per tal que s'aturi davant la boca en casos on l'usuari pot agafar el menjar per ell mateix.

Després proposem una definició conceptual de les preferències per a robots assistencials en diferents tasques. Aquestes definicions, que són més abstractes que les adaptacions a baix nivell, permeten desenvolupar un comportament més complex. Les preferències es classifiquen en una taxonomia on considerem aquelles preferències que poden millorar la capacitat de decisió del robot i ajudar-lo a escollir aquelles accions que són més agradables per a l'usuari.

"Les contribucions de la tesi mostren la importància d'aquest tipus d'adaptació per a millorar l'execució de la tasca assistencial i el gran impacte social que la robòtica assistencial pot arribar a tenir."

Aquestes definicions són després utilitzades per crear algorismes robustos que permeten adaptar el comportament del robot mitjançant tècniques de planificació en Intel·ligència Artificial (IA). En aquest context, es defineixen les capacitats del robot per mitjà de les accions que aquest pot desenvolupar. El planificador s'encarrega de trobar un pla, que és una seqüència d'accions que compleixen la tasca. Suggestim afegir preferències en aquest entorn, on associem a cada acció un cost proporcional a les preferències (així doncs, les accions que no compleixen les preferències són més «cares»). Amb aquesta definició, hem desenvolupat mètodes per adaptar les accions a l'usuari de manera que, quan l'usuari canvia el comportament, el model de planificació també ho fa.

Finalment, com que obtenir les preferències de l'usuari és complicat, també proposem un mètode per suggerir valors de preferències als usuaris, tenint en compte les preferències ja conegudes. Aquest mètode analitza un subespai dels possibles plans d'acció per tal de trobar els valors de preferències que fan que el cost total del pla sigui el menor possible, i es proposen a l'usuari.

La tesi conclou amb un estudi d'usuaris -amb qui es duen a terme les tres tasques- on es mostra la utilitat dels mètodes proposats, com els usuaris poden identificar quan el robot usa les seves preferències i com això incrementa el seu benestar.

Premi Marc Esteva Vivanco 2020

Conclusions

La tesi se centra en la integració de preferències d'usuari per a modificar el comportament del robot en tasques assistencials físiques. Les contribucions de la tesi mostren la importància d'aquest tipus d'adaptació per a millorar l'execució de la tasca assistencial i el gran impacte social que la robòtica assistencial pot arribar a tenir.

nodes

Sensorimotor exploration: constraint awareness and social reinforcement in early vocal development*

Juan Manuel Acevedo Valle

Director: Cecilio Angulo
Universitat Politècnica de Catalunya.
Departament d'Enginyeria de Sistemes,
Automàtica i Informàtica Industrial

** El text presentat a continuació és una traducció del resum en anglés lliurat per l'autor*

"La nostra recerca es va centrar a entendre el paper de la consciència de la restricció motriu i els comportaments imitatius durant el balbuceig vocal."

Malgrat una visió optimista sobre el futur de la robòtica, els robots encara no han impregnat la nostra quotidianitat, i tampoc no s'ha resolt una sèrie de problemes perquè aquests puguin comportar-se de manera eficient en entorns oberts i incerts. Així, d'entre els camps més complexos associats a la robòtica, es troben la visió per ordinador, la navegació, el control del moviment i la interacció humà-robot.

L'evidència procedent de les ciències biològiques suggereix que l'adquisició de coneixement està condicionada tant per l'entorn en què es desenvolupa un agent intel·ligent com pel mateix cos de l'agent, el qual se sotmet a processos de desenvolupament que configuren allò que pot aprendre i quan ho pot fer. Per exemple, com a humans, el nostre cos i la nostra ment canvien al llarg de les nostres vides. Mitjançant robots i simuladors, la robòtica del desenvolupament pretén comprendre els processos clau descrits per la psicologia del desenvolupament per tal de permetre l'adquisició progressiva d'una interacció eficient d'un agent artificial amb l'entorn, incloent-hi els seus semblants. L'ús de la robòtica per adquirir coneixement sobre el desenvolupament primerenc dels nadons pot aportar avantatges a diferents camps de la ciència. Per exemple, entendre el balbuceig i altres comportaments rítmics observats en els nadons ens podria ajudar a comprendre mecanismes cognitius dels nadons que encara no entenem, així com a construir robots capaços d'aprendre a controlar els seus cossos mecànics de la mateixa manera que ho fan els nadons.

La nostra recerca es va centrar a entendre el paper de la consciència de la restricció motriu i els comportaments imitatius durant el balbuceig vocal. El balbuceig vocal és una etapa inicial de l'adquisició de la parla en la qual els nadons aprenen a controlar el seu tracte vocal per produir els sons desitjats. Per aconseguir-ho, un infant ha d'aprendre a controlar el seu sistema articulador i construir un mapa sensoriomotor de gestos articuladors i sons de parla. Durant l'època del balbuceig, els nadons exploren el seu propi cos per aprendre a controlar el seu tracte vocal i produir els sons previstos. Per reproduir aquest mecanisme amb un agent artificial, vam començar per integrar un simulador de tracte vocal amb una arquitectura curiosity-driven exploration. Els estudis anteriors que utilitzaven arquitectures similars no van reconèixer la rellevància de les limitacions en l'àmbit sensorial, motor, cognitiu i social.

Nosaltres, tot i la necessitat de tenir en compte tota mena de limitacions, ens vam centrar a les limitacions motrius i socials. Tal com suggereixen els estudis de desenvolupament, abans d'aprendre un model sensoriomotor, els infants han de descobrir primer les limitacions imposades pel seu sistema motor (se suposa que la informació tàctil és important per assolir aquesta fita de desenvolupament).

Vam argumentar que la informació tàctil proporciona una manera de fer front a la redundància en el model sensoriomotor adquirit durant l'exploració, per exemple, descobrint restriccions motores. Per descobrir aquestes limitacions, l'agent està dotat d'un mecanisme per obtenir informació tàctil, i aquest mecanisme genera un senyal que indica si s'ha arribat a una configuració motora conflictiva. En el cas del tracte vocal artificial, aquesta informació indicaria que l'agent s'està mossegant la llengua o el tracte vocal està bloquejat i, per tant, no es pot produir la parla.

Finalment, vam plantejar la hipòtesi que l'exploració sensoriomotora no només està impulsada pel progrés de la competència i el descobriment de limitacions, sinó també per la rellevància dels objectius sensorials amb finalitats de socialització. Els treballs científics que estudien el desenvolupament vocal primerenc han demostrat que la parla ambiental dona forma a l'espai sensorial explorat durant el balbuceig. De manera més general, s'ha demostrat que els comportaments imitatius són crucials per al desenvolupament primerenc dels nadons, ja que restringeixen l'espai de cerca durant l'exploració sensoriomotora. Així, per tenir en compte les interaccions socials durant el desenvolupament, vam incloure un expert en unitats sensorials rellevants per a la comunicació. Les interaccions amb aquest instructor expert es produeixen quan l'enunciat d'un aprenent és «prou» semblant a un enunciat rellevant per a la comunicació. En aquest cas, l'instructor reformula l'enunciat com la unitat sensorial rellevant per a la comunicació. Quan l'aprenent percep una acció de l'expert, immediatament intenta imitar-lo. En els nostres experiments vam considerar un instructor expert en la producció de vocals alemanyes.

En resum, en els nostres experiments, un agent artificial explora el seu sistema oïda-tracte vocal escollint objectius auditius, aprenent a produir sons previstos i evitant configuracions motores no desitjades, i imita una expressió produïda per un instructor. De manera coherent amb un gran nombre d'estudis de psicologia del desenvolupament, els nostres experiments van suggerir que l'exploració sensoriomotora es pot enriquir amb diferents modalitats perceptives i socials. En el cas del desenvolupament vocal primerenc, vam observar una bona evolució de la complexitat dels gestos vocals mitjançant vocals alemanyes, que és necessària per a l'aparició del llenguatge parlat en els nadons.

Enllaç a la tesi: <https://upcommons.upc.edu/bitstream/handle/2117/168572/TJMAV1de1.pdf;sequence=1>.

"Vam argumentar que la informació tàctil proporciona una manera de fer front a la redundància en el model sensoriomotor adquirit durant l'exploració, per exemple, descobrint restriccions motores."

"De manera coherent amb un gran nombre d'estudis de psicologia del desenvolupament, els nostres experiments van suggerir que l'exploració sensoriomotora es pot enriquir amb diferents modalitats perceptives i socials."

nodes

Amb el suport de:

