

connexions



Ètica i justícia en la Intel·ligència Artificial, des d'una perspectiva de gènere

vista prèvia >

La Intel·ligència Artificial diàriament desenvolupa noves potencialitats i genera oportunitats que poc abans semblaven impossibles. Però com tot avenç tecnològic en la història, aquest corre el risc de provocar l'aparició de noves desigualtats, discriminacions i injustícies. Cal que la Intel·ligència Artificial sigui socialment intel·ligent, i que serveixi tothom de forma equitativa. Per fer-ho cal superar els biaixos de gènere que estan marcant actualment el seu camí. Que és el nostre com a societat.



Anna Grau

Consultora especialitzada en *smart cities*, transformació digital i mobilitat del futur
Co-fundadora i presidenta del projecte Young IT Girl

@AnnaGrauGal

El sector tecnològic és, sense cap mena de dubte, el paraigua que acull la transformació de la societat digital que som, i que cada cop més serem, en la forma com ens relacionem, en la que prenem decisions i en la que creem noves idees i construïm projectes empresarials. Alhora, aquesta forma de vida cada vegada suscita més canvis de paradigma en les persones.

Un exemple clar d'aquesta disrupció digital és la Intel·ligència Artificial, IA d'ara en endavant. La IA és una tecnologia, basada en la ciència de la computació, que busca afegir la conducta humana en l'automatització de les activitats i la presa de decisions, amb la finalitat d'assistir i atendre a les persones. La IA percep el seu entorn en què es mou, pensa, aprèn i actua d'acord i, en conseqüència, amb allò que l'envolta. Per aconseguir aquest efecte, la IA compta amb el disseny d'algoritmes que raonen i prediuen de forma automatitzada, a partir a la informació que han captat al seu voltant.

En la revolució digital actual, la IA és més present que mai, sense adonar-nos-en, en accions quotidianes com

ara buscar la millor ruta per anar a un lloc, demanar menjar a domicili des d'una *app* de mòbil, rebre suggeriments i recomanacions en una plataforma de música, o veure la nostra pel·lícula preferida. Els algoritmes formen part de com vivim, són a tot arreu i ens ajuden a ser més eficients amb el nostre temps. No obstant, com en tot avenç tecnològic, hi ha diversos aspectes d'aquests algoritmes que cal redirigir o mitigar pel seu impacte, no sempre positiu, cap a la societat. Un d'aquests impactes negatius pot ser la reproducció d'injustícies o discriminacions que es viuen al món no digital, com per exemple la discriminació de gènere.

Els prejudicis de la societat plasmats a la lògica de la IA

Podem entendre la IA com el procés d'aprenentatge d'una nadó des de la seva naixença. Totes dues neixen amb una inspiració per part dels creadors o progenitors, creixen progressivament gràcies als aprenentatges o entrenaments que se'ls hi dona i s'inspiren en base a l'entorn social i referències que tenen a l'abast. Dit

en altres paraules, creixen i aprenen en base a allò que perceben. Ambdues són un full en blanc principalment per dues raons: per una banda, són com una esponja del llenguatge i imatges dels que estan envoltats per determinar la seva forma de ser o la seva lògica de presa de decisions —els algoritmes. I, d'altra banda, ambdues necessiten prendre com a referència una mostra neutra representativa de la realitat per no mostrar cert biaix en les seves decisions o forma de desenvolupar-se.

Per què aquest símil? Mentre una noia no identifiqui en allò que percep una referència d'una dona líder d'un projecte empresarial en el sector tecnològic, no hi haurà una presència femenina en aquest món. De la mateixa manera, mentre una solució d'IA sigui creada només per homes i, per tant, no inclogui una mostra representativa de la realitat en diversitat de gènere, aquesta solució tindrà un biaix en la seva gestió i presa de decisions.

A mode d'exemple il·lustratiu, moltes empreses compten amb sistemes de *machine learning* per a la selecció de personal. Si aquests sistemes només

Només un 30% dels llocs de treball en el sector de les tecnologies de la informació i la comunicació a Catalunya els ocupen dones

prenen dades històriques dels seus treballadors, el més probable és que la política de selecció premii el talent masculí respecte al femení. De la mateixa manera, una nadó que no se li dona un casc d'astronauta o no té un referent femení proper, no creixerà amb la idea i confiança de poder ser-ne una.

Així mateix, mentre que hi hagi uns estereotips i cultura global en què els estudis científics i tècnics no són un entorn per les dones, la confiança traslladada en les noies que vulguin emprendre aquests camins serà nul·la o molt poca. Altra vegada, referent a la IA, mentre els seus serveis de computació no estiguin creats per equips diversos en gènere, els seus algorismes representaran una lògica esbiaixada en gènere.

Totes les persones tenim prejudicis pel que hem viscut, però pensem quin impacte pot tenir el fet que una tecnologia com la IA, la qual està present més del que ens pensem en les nostres vides i que dominarà el món, tingui discriminació. Això podria perpetuar errors al futur que ara no voldríem. No pot ser, en conseqüència, que la IA no tingui en compte

al 50% de la població que són les dones. O, tornant al símil entre la IA i una nadó, no podem ser o somiar en convertir-nos en algú que no hem conegut mai. De la mateixa manera, créixer sota un parany tecnològic no representatiu de la societat que vivim ens fa prendre decisions inequívocatives.

Invisibilitat femenina en el món tecnològic

Només un 30% dels llocs de treball en el sector de les tecnologies de la informació i la comunicació a Catalunya els ocupen dones.¹ No es tracta d'una segmentació de gèneres per especialitzacions, sinó d'un sector que ja s'ha impregnat, i encara ho farà en major mesura, en tots els aspectes de la nostra vida. Per què és important augmentar la presència femenina en el món digital? Segons l'informe de novembre de 2018 de l'European Center for the Development of Vocational

Training,² es preveu que un 84,5% de l'oferta ocupacional serà d'àmbit tecnològic al 2025 i la baixa presència femenina al sector implica un enorme risc d'exclusió laboral per raons de gènere en un futur molt proper. És a dir, és necessari, en primer lloc, que tothom disposi d'unes competències digitals primordials i, en segon lloc, que la societat digital i de la informació que s'està consolidant generi igualtat d'oportunitats. Històricament, la dona no ha estat representada com una referència en la ciència i tecnologia. Per demostrar aquest fet només caldria agafar un llibre d'educació primària i procurar d'identificar quantes científiques o tecnòlogues hi apareixen. En aquesta línia, podem trobar a la Marie Curie (1867-1934) i, pot ser, com a molt, alguna més. Algú preguntarà si aquest fet passa perquè en aquella època no hi havia dones en aquest sector? No, no és per aquesta raó. Sempre hi ha hagut científiques, però han estat invisibilitzades.

1 INSTITUT CATALÀ DE LES DONES, «Dones en les Tecnologies de la Informació i la comunicació. Dossier estadístic».

2 CEDEFOP, «European Center for the Development of Vocational Training».

La IA aporta molts beneficis, però ara per ara encara té un llarg recorregut perquè representi a tothom i, per tant, tingui un efecte a tota la humanitat, sense desviació desfavorable per a persones o col·lectius

Alguns exemples ineludibles són l'Ada Lovelace (1815-1852), la primera dona programadora de la història que al 1815 va participar en la màquina analítica de Babbage,³ l'antecedent a l'ordinador modern. Un altre nom imprescindible és el de Hedy Lamarr (1914-2000), una estrella del Hollywood clàssic i inventora de l'espectre eixamplat per salt de freqüències, creat pel radioguiatge dels torpedes americans durant la Segona Guerra Mundial, un mètode de posicionament per satèl·lit utilitzat pel GPS i que va ser el precursor del que ara anomenem WiFi. O per exemple, Mary Anderson (1866-1953), inventora del braç mecànic que incorporen els cotxes que ara coneixem com netejaparrisa, i que al 1916 ja incorporaven tots els automòbils. I moltes més. Un llistat interminable de coneixement, invenció i saviesa que no ha estat

3 La màquina analítica de Babbage és de naturalesa mecànica la qual podia ser programada per l'usuari per executar un repertori d'instruccions en l'ordre desitjat. Aquesta inclou la majoria de les parts lògiques d'un ordinador actual i capacitat d'emmagatzemar 1.000 números de 50 dígits cada un d'ells. Charles Babbage (1791-1871) va ser un dels pares de la informàtica, però la seva concepció de la màquina fou a les primeres dècades del segle XIX, i la tecnologia disponible aquella època no li va permetre desenvolupar el projecte.

present a la nostra societat ni a l'ensenyament i que, per tant, resulta difícil d'encoratjar a combatre l'esclatxa de gènere avui existent a la indústria, la ciència i la tecnològica. Sense l'existència d'aquests casos, i molts d'altres referents femenins en el sector, la IA no hauria arribat al que coneixem avui. Un dels pares de la IA Marvin Lee Minsky (1927-2016), estava convençut que la IA salvaria la humanitat. Considerava que els avantatges i beneficis que podria aportar esdevindrien la clau per resoldre les desigualtats i les problemàtiques actuals. És evident que la IA aporta molts beneficis i, cada vegada, més capes de valor i millores en moltes indústries i als nostres usos quotidians, però cal dir que, ara per ara, encara té un llarg recorregut perquè representi a tothom i, per tant, tingui un efecte a tota la humanitat, sense desviació desfavorable per a persones o col·lectius.

Per entendre aquest esdevenir de la IA, a més del desig de Lee, cal fer esment a l'afirmació de Cristina Aranda (1976), cofundadora de MujeresTech; «qui domini la Intel·ligència Artificial, dominarà el món». Així, davant

d'aquesta realitat, cal debatre i indagar en si aquesta tecnologia és ètica, justa i facilitadora per la societat en la seva globalitat.

Discriminació amagada dins la IA

Segons el *Gender Social Norms Index* (GSNI), índex establert per les Nacions Unides que mesura com els pensaments socials obstrueixen la igualtat de gènere en àrees com la política, l'ensenyament, assenyala que el 90% dels homes i dones mostren un biaix contra les dones, en una mostra de dades relativa a 75 països i cobrint més del 80% de la població mundial.⁴ D'altra banda, segons l'article «The Ultimate List of Cognitive Biases: Why Humans Make Irrational Decisions»,⁵ s'enumeren fins a un total 49 biaixos cognitius negatius cap a les dones. Nogensmenys, de biaixos, n'hi ha fins a un centenar, entre ells culturals i cognitius. Per tant, aquests biaixos formen part del pòsit cultural de l'entorn.

4 UNITED NATIONS, «Almost 90% of Men/Women Globally Are Biased Against Women».

5 HOCHMA, «The Ultimate List of Cognitive Biases: Why Humans Make Irrational Decisions».

Una pregunta derivada de la presència tan majoritàriament masculina en la IA és si aquest avenç tecnològic, idealment tan decisiu per la societat, acaba tenint efectes desfavorables per a les dones

Així, els models i sistemes tecnològics que creem i entrenem, també en la IA, són un reflex de nosaltres mateixos. El biaix de gènere en la IA és una realitat. Per exemple, només cal esmentar que el 78% dels professionals en IA són homes,⁶ fet que provoca que la discriminació impregni els algoritmes d'aquesta tecnologia. Una altra pregunta que sorgeix d'aquesta presència tan majoritàriament masculina en la IA, és si aquest avenç tecnològic, idealment tan decisiu per a la societat, acaba tenint efectes desfavorables per a les dones. El 2014, Amazon va estrenar un algoritme per reclutar nous treballadors per als seus magatzems.⁷ L'eina puntuava d'una a cinc estrelles els millors candidats. Tot semblava ideal: la IA estalviaria hores al departament de recursos humans. Però, un any més tard, la multinacional es va adonar que en els llocs tècnics, com el de desenvolupador de programari, no s'havia contractat cap dona. Podria ser que no hi hagués cap candidata amb aptituds per a aquella feina? L'exemple ja és un clàssic dels errors que ha

6 WORLD ECONOMIC FORUM, «Global Gender Gap Report: Assessing Gender Gaps in Artificial Intelligence».

7 BAEZA-YATES, «¿Por qué la inteligencia artificial discrimina a las mujeres?».

provocat recentment la IA. Les dades massives que van servir per alimentar l'algoritme de selecció de personal es van basar en currículums rebuts en la dècada anterior, en què la majoria dels programadors eren homes. Quan el sistema automàtic detectava la paraula «dona» —o un sinònim—, directament penalitzava el currículum posant-li menys puntuació.

Un altre exemple d'inclinació esbiaixada va passar el 2016. Investigadors de Microsoft Research i de la Boston University van fer servir una col·lecció massiva de notícies de Google per entrenar algorismes sobre els estereotips femení i masculí que sortien a la premsa. Els resultats donaven que els homes eren programadors informàtics i les dones mestresses de casa. Ells eren metges i elles, infermeres. I té la seva lògica, perquè als EUA la majoria dels periodistes que redactaven les notícies sobre les quals es va basar l'entrenament de l'algoritme eren homes. Per tant, Google només reflectia el biaix de gènere existent en la realitat.

Altres exemples escandalosos derivats de la subrepresentació dels

rols femenins al món tecnològic són els algoritmes bancaris que atorguen límits de crèdits més elevats als homes, així com les plataformes d'adquisició laboral que ignoren les dones i les minories, com s'apunta en diversos estudis que denuncien que a diverses plataformes de professionals com ara LinkedIn i Viadeo, havent-hi homes i dones amb qualificacions similars, en els primers resultats del cercador només apareixen homes.⁸ Aquests són alguns dels exemples de molts encara per identificar, en què en l'ús de la IA s'ha discriminat les dones. Tal com explica la investigadora Elisabet Golobardes: «L'algoritme com a tal no té cap biaix. Són les dades que s'introdueixen i l'objectiu pel qual ha estat dissenyat el que discrimina».⁹

L'ètica de la IA

Avui la IA està en boca de tothom, i resulta imprescindible per a les persones en moltes aplicacions. Tant el

8 ZEHLIKE (et al.), «FA*IR: A Fair Top-k Ranking Algorithm».

9 APDCAT, «Inteligència Artificial. Decisions Automatitzades a Catalunya».

Els usuaris estan menys exposats a punts de vista diferents i queden aïllats intel·lectualment en el seu propi biaix informatiu

que és el *machine learning* [aprenentatge automàtic] —com, en particular, el *deep learning*¹⁰ [aprenentatge profund]. Cal que pensem, com a societat, què té d'ètic el que fem servir de la IA? És un debat i una preocupació compartida entre les institucions i empreses responsables del disseny dels seus algoritmes, però també cal que estigui en el punt de mira de la ciutadania com a principal usuària d'aquesta tecnologia. Quan pensem en els avenços tecnològics, ens venen a la ment diverses reflexions de com es reaccionarà davant un problema de caire ètic. Un exemple d'això podria ser el vehicle autònom i l'ètica davant d'una possible incidència o resposta immediata. Pensant, doncs, en la IA, què està bé o malament i què representa un biaix o una decisió objectiva?

En aquest sentit, s'han desenvolupat diversos documents com ara la guia ètica de la Comissió Europea per a

una IA fiable,¹¹ la Declaració de Barcelona¹² o l'Estratègia IA per a Catalunya¹³ —presentada l'any 2019 i que incorpora la proposta de creació d'un Observatori Ètic—; alguns exemples que responen a la preocupació per assegurar els drets bàsics de la ciutadania davant d'aquest avenç tecnològic. A més, aprofundint en la IA, un fet que succeeix i que nodreix la desinformació o conflicte ètic és l'anomenat «filtre bombolla», fenomen encunyat pel ciberactivista Eli Pariser (1980) per descriure la selecció personalitzada de la informació que rep cada individu. Aquest filtre té la conseqüència que introdueix l'usuari a una bombolla adaptada perquè aquest es trobi còmode, però amb la derivada de quedar aïllat d'altres visions de la realitat. En altres paraules, aquest fet fa que els usuaris estiguin menys exposats a punts de vista diferent i quedin aïllats intel·lectualment en el seu propi biaix informatiu. Un bon exemple d'aquest

efecte de la IA és que si t'agraden les sèries policíacques, només et suggereix sèries dins d'aquest mateix gènere, sense explorar noves opcions.

En aquest context, la percepció de la informació abundant i l'elevat volum de dades generades a l'entorn digital es dilueix. La confecció dels algoritmes sovint impedeixen pensar fora de la caixa i, en conseqüència, conèixer altres formes de pensar i idees, enriquir-se amb el coneixement del món i, fins i tot, estimular la nostra creativitat. I és una entrada al sectarisme.

Principals biaixos de gènere a la IA

Els principals biaixos de gènere identificats a la lògica i l'aplicació de la IA són principalment quatre.

En primer lloc, la IA no afavoreix allò desconegut i, per tant, quan els històrics de dades preses com a referència no són representatives de la societat o, dit d'una altra manera, no inclouen les dones, aquestes pateixen una discriminació en els elements decisors. És a dir, aquelles casuístiques en què hi hagi un conjunt de dades utilitzades en l'entre-

10 El *deep learning*, en anglès, com a variant de la IA és l'aprenentatge profund. És a dir, un conjunt d'algoritmes d'aprenentatge automàtic, però que intenta modelar abstraccions d'alt nivell en dades utilitzant arquitectures computacionals que permeten transformacions no lineals múltiples i iteratives de dades expressades. En altres paraules, aquest aprenentatge permet reconeixements més avançats d'aglomerats de dades complexes de desxifrar automàticament.

11 EUROPEAN COMMISSION, «EU artificial intelligence ethics checklist ready for testing as new policy recommendations are published».

12 ARTIFICIAL INTELLIGENCE RESEARCH CENTRE, «Barcelona Declaration for the proper development and usage of artificial intelligence in Europe».

13 GENERALITAT DE CATALUNYA, «Estratègia IA per a Catalunya, 2019».

L'especialització de la IA constitueix un cas d'especial singularitat, ja que hi ha únicament una dona per cada deu homes cursant aquests estudis a Catalunya

nament que siguin incompletes o esbiaixades. Per exemple, si només el 10% d'una mostra són dones, a l'hora d'aplicar el model d'entrenament de *machine learning*¹⁴ és més probable que davant una dona produeixi una freqüència d'error major. En un món tecnològic on la presència femenina és minoria, aquesta distinció en les lògiques succeeix molt més del que voldríem. Val la pena dir que aquest biaix no només fa témer pel present pel que fa a la poca presència d'avui al món professional, sinó que també enfosqueix les previsions de futur degut a la manca de tria de carreres tècniques per part de noies, fet que retroalimenta el biaix. Referent a aquest aspecte, diversos estudis apunten l'existència d'estereotips de gènere en les carreres STEM,¹⁵ ja que només hi opten un 26% de les noies que accedeixen a estudis universitaris, davant d'un 40% dels nois. Curiosament, a aquesta dada cal afegir que el 60% de les persones

graduades a les universitats catalanes són dones.¹⁶ L'especialització de la IA constitueix un cas d'especial singularitat, ja que hi ha únicament una dona per cada deu homes cursant aquests estudis a Catalunya.¹⁷

En segon lloc, la IA plasma i és un digne reflex dels responsables codificadors dels serveis d'aprenentatge, entorn absolutament masculinitzat. Segons Reuters¹⁸ el percentatge de dones ocupades en posicions tècniques en les majors companyies de *machine learning* és només del 20%.¹⁹ La formació del model de *machine learning* compta amb un aprenentatge supervisat, és a dir, que les dades de la formació s'etiqueten per ensenyar i entrenar al model com es pot comportar. Atès que tothom té prejudicis i biaixos, tant conscients com inconscients, aquests es representen involuntàriament a

l'etiquetatge i formació de la IA. Segons un estudi d'investigació de WIRED, l'eina de cerca i buscador de la IA de Google llista fins a 641 treballadors en *machine intelligence*, d'entre els quals només el 10% són dones. D'altra banda, el llistat resultant a través del cercador de *Pages* conclou un total de 115 persones, només el 15% d'aquests són dones.²⁰ La IA és un sector emergent en ple creixement. La demanda de professions tecnològiques incrementarà un 14% anual fins al 2025 a Europa i la IA encapçala aquest moviment.²¹ Així, la IA necessita molts més professionals dels que ara mateix l'acadèmia genera. Actualment hi ha un dèficit professional important, cosa que determina pràcticament la plena ocupació del sector.²²

En tercer lloc, les característiques i tècniques de modelatge. La creació d'aquests serveis de computació en identificació de la veu, detecta de

14 *Machine learning* o aprenentatge automàtic fa referència a la capacitat d'una màquina per aprendre mitjançant l'adaptació de certs algorismes de la seva programació respecte a certa entrada de dades al seu sistema.

15 L'acrònim STEM agrupa les àrees de coneixement de la Ciència, la Tecnologia, l'Enginyeria i les Matemàtiques, en anglès.

16 GENERALITAT DE CATALUNYA, «Pla STEMcat d'impuls de les vocacions científiques, tecnològiques, en enginyeria i en matemàtiques».

17 GENERALITAT DE CATALUNYA, «Estratègia IA per a Catalunya, 2019».

18 DASTIN, «Amazon scraps secret AI recruiting tool that showed bias against women»; GOMEZ, «Women in Artificial Intelligence: mitigating the gender bias».

19 WEF, «Global Gender Gap Report: Assessing Gender Gaps in Artificial Intelligence».

20 SIMONITE, «AI Is the Future—But Where Are the Women?».

21 RANDSTAD, «El sector industrial en España necesitará 3,5 millones de profesionales especializados en 2025».

22 Ibidem.

La demanda de professions tecnològiques incrementarà un 14% anual fins al 2025 a Europa, i la IA encapçala aquest moviment

forma més acurada les veus amb cordes vocals més fortes i molt menys aquelles més febles. És, altra vegada, una casualitat?

Google compta amb un reconeixement de veu amb presència comprovada de biaix de gènere. En aquest estudi es comprova que el reconeixement funciona correctament per més de 1.500 paraules d'entre 50 diferents accents de múltiples vídeos. A més, garantint una mostra igualitària de presència d'homes i dones per la identificació de cada dialecte. Els resultats van concloure que menys de la meitat de les paraules de les veus femenines —un 47%—, estaven captades correctament. Per contra, en el cas dels homes, la xifra de captura exacta ascendia al 60%.²³

Això genera impactes reals en la vida de les persones, en àmbits de la vida quotidiana, i d'altres que estan per arribar, com ara el reconeixement de veu en els vehicles o bé qualsevol estudi de programari dictat mèdic, on les dones, de forma objectiva, ho tenen més difícil.

23 TATMAN, «Google's speech recognition has a gender bias».

En quart lloc, hi ha un biaix sexista. Sense anar més lluny, totes les veus de les assistències virtuals de les empreses tecnològiques mundialment més reconegudes, ja sigui el Google Assistant, la Siri d'Apple, la Cortana de Microsoft o bé l'Alexa d'Amazon són de dones. D'altra banda, tres d'entre aquestes quatre, a més, tenen nom femení. Per no esmentar que, per un costat, Microsoft va donar nom a Cortana per un personatge femení del videojoc *Halo*, la qual és una dona sensual i desvestida. Per l'altre costat, Apple fixà el nom de Siri que significa, en noruec, «dona preciosa que et porta a la victòria».

Aquest quatre biaixos poden ser els més destacats, però n'hi ha molts més. Per exemple, al 2018, s'estima que es van vendre 100 milions d'assistències intel·ligents de veu aproximadament i, addicionalment, segons un estudi de Gartner, es preveu que al 2020 la mitjana de les persones mantindrà més conversa amb una teleassistència, ja sigui pels *smartphones*, tabletas o d'altres assistències virtuals, que amb la seva pròpia parella.²⁴

24 LEVY, «Gartner's Top 10 Strategic Predictions for 2017 and Beyond: Surviving the Storm Winds of

Conclusions

Els models i algorismes de la IA són un reflex de la societat que, inevitablement, té prejudicis i estereotips. Amb la finalitat d'eradicar els biaixos de gènere en la IA, en què s'ha de treballar?

Cal treballar per tenir una major presència femenina en la creació de continguts digitals i de la IA, tot assegurant que els seus serveis de comunicació estaran desenvolupats amb dones i pel benestar de les dones. Alhora, cal garantir que les dades que nodreixen qualsevol algoritme de IA siguin representatives. És a dir, d'homes i dones. Així, convindria impulsar programes d'autoconsciència del biaix als homes creadors d'aquests algorismes que dominaran el món.²⁵

Addicionalment, cal crear programes de sensibilització i conscienciació sobre la rellevància de la diversitat en equips programadors de plataformes de IA per mitigar qualsevol discrimi-

Digital Disruption».

25 TYN MAGAZINE, «Cómo terminar con el sesgo de género en los algoritmos de inteligencia artificial».

Cal crear programes de sensibilització i conscienciació sobre la rellevància de la diversitat en equips programadors de plataformes de IA per mitigar qualsevol discriminació inconscient

nació inconscient. Així mateix, és necessari formar a la societat, empreses, institucions i centres educatius per combatre desigualtats i trencar amb els estigmes.

Els darrers anys hi ha hagut un esforç molt gran a Catalunya, amb l'aparició de diferents entitats, iniciatives i esdeveniments, per tal d'inspirar i promoció vocacions tecnològiques a les noies en edats primerenques, amb el fi de combatre l'esclatxa de gènere digital. És un pas essencial, però cal començar a integrar programes de formació i sensibilització en l'ús just i inclusiu de la IA. N'és un exemple l'entitat de les Young IT Girls²⁶ que, en el marc del seu objectiu global d'inspirar vocacions tècniques i científiques a les noies tot actuant a les escoles —ja sigui de forma presencial o a distància—, ha començat a incloure en les seves activitats la introducció a la IA i, així, crear consciència de la seva importància a les generacions futures.

Així, és necessari millorar els coneixements en la IA dels educadors i la

seva didàctica per generar millors aprenentatges i interès per les carreres tècniques, que alhora ocupen a una de les especialitzacions amb major nombre de llocs de treball de present, i sobretot, de futur. També cal incrementar-ne la motivació mitjançant tècniques i plataformes de *e-learning*²⁷ que les facin més atractives i accessibles a tothom.

Calen esforços socials i econòmics per redirigir els models d'IA existents, allunyant-los d'històrics de dades i patrons poc representatius. És a dir, cal una neteja i escombrada dels biaixos, que normalment són més difícils d'eliminar que els propis estigmes de les persones. Per aquesta finalitat s'ha de pensar en una inversió d'investigació tecnològica orientada a tal finalitat. En definitiva, és primordial que la IA d'avui sigui reflex dels valors de la població pels quals està creada. ■

²⁶ Entitat sense ànim de lucre que busca trencar amb l'esclatxa digital de gènere, inspirant vocacions tecnològiques a les noies en educació primària i secundària: <www.youngitgirls.org>.

²⁷ L'*e-learning* significa aprenentatge electrònic en anglès, com a possibilitat de l'educació en remot a través de l'entorn digital.

■ Bibliografia

APDCAT. «Intel·ligència Artificial. Decisions Automatitzades a Catalunya» [en línia]. Disponible a: <www.apdcat.gencat.cat>.

ARTIFICIAL INTELLIGENCE RESEARCH CENTRE. «Barcelona Declaration for the proper development and usage of artificial intelligence in Europe» [en línia]. Declaració duta a terme el 8 de març del 2017. Disponible a: <www.iiia.csic.es>.

BAEZA-YATES, Ricardo. «¿Por qué la inteligencia artificial discrimina a las mujeres?» [en línia]. A *Medium*, de 22 d'octubre del 2019. Disponible a: <www.medium.com>.

CEDEFOP. «European Center for the Development of Vocational Training» [en línia]. Novembre 2018. Disponible a: <www.cedefop.europa.eu>.

DASTIN, Jeffrey. «Amazon scraps secret AI recruiting tool that showed bias against women» [en línia]. A *Reuters*, de l'11 d'octubre del 2018. Disponible a: <www.reuters.com>.

EUROPEAN COMMISSION. «EU artificial intelligence ethics checklist ready for testing as new policy recommendations are published» [en línia]. Disponible a: <www.ec.europa.eu>.

GENERALITAT DE CATALUNYA. «Estratègia IA per a Catalunya, 2019» [en línia]. Disponible a: <www.participa.gencat.cat>.

GENERALITAT DE CATALUNYA. «Pla STEMcat d'impuls de les vocacions científiques,

tecnològiques, en enginyeria i en matemàtiques» [en línia]. Disponible a: <www.smartcatalonia.gencat.cat>.

GOMEZ, Emilia. «Women in Artificial Intelligence: mitigating the gender bias» [en línia]. Disponible a: <www.ec.europa.eu>.

HOCHMA, Tomer. «The Ultimate List of Cognitive Biases: Why Humans Make Irrational Decisions» [en línia]. Disponible a: <www.humanhow.com>.

INSTITUT CATALÀ DE LES DONES. «Dones en les Tecnologies de la Informació i la Comunicació. Dossier estadístic» [en línia]. De febrer de 2020. Disponible a: <www.dones.gencat>.

LEVY, Heather P. «Gartner's Top 10 Strategic Predictions for 2017 and Beyond: Surviving the Storm Winds of Digital Disruption» [en línia]. A *Gartner*, de 18 octubre 2016. Disponible a: <www.gartner.com>.

RANDSTAD. «El sector industrial en España necesitará 3,5 millones de profesionales especializados en 2025» [en línia]. A *Randstad*, del 21 de març de 2016. Disponible a: <www.randstad.es>.

SIMONITE, Tom. «AI Is the Future—But Where Are the Women?» [en línia]. A *WIRED*, del 17 d'agost de 2018. Disponible a: <www.wired.com>.

TATMAN, Rachael. «Google's speech recognition has a gender bias» [en línia]. A *Making Noise & Hearing Things*, del 12 de juliol de 2016. Disponible a: <www.makingnoiseandhearingthings.com>.

TYN MAGAZINE. «Cómo terminar con el sesgo de género en los algoritmos de inteligencia artificial» [en línia]. A *TyN Magazine*, de 14 de desembre de 2019. Disponible a: <www.tynmagazine.com>.

UNITED NATIONS. «Almost 90% of Men/ Women Globally Are Biased Against Women» [en línia]. De 5 de març de 2020. Disponible a: <www.undp.org>.

WORLD ECONOMIC FORUM. «Global Gender Gap Report: Assessing Gender Gaps in Artificial Intelligence» [en línia]. Disponible a: <www.reports.weforum.org>.

ZEHLIKE, Mieke (et al.). «FA*IR: A Fair Top-k Ranking Algorithm» [en línia]. Disponible a: <www.arxiv.org>.