




Universitat Autònoma de Barcelona

ADVERTIMENT. L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons:  http://cat.creativecommons.org/?page_id=184

ADVERTENCIA. El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons:  <http://es.creativecommons.org/blog/licencias/>

WARNING. The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license:  <https://creativecommons.org/licenses/?lang=en>

COMUNICACIÓ, INTEL·LIGÈNCIA ARTIFICIAL I ÈTICA:
EINES PER GOVERNAR ELS ALGORITMES DES DEL
PERIODISME EN FAVOR DE LES LLIBERTATS A LES
SOCIETATS DEMOCRÀTIQUES

Patricia Ventura Pocino

Tesi Doctoral
Doctorat en Mitjans, Comunicació i Cultura
Facultat de Ciències de la Comunicació
Director: David Vidal Castell
Universitat Autònoma de Barcelona, 2022

A la meva mare, Matilde Pocino Isern

Esta loca generosidad es la de la rebeldía, que da sin esperar su fuerza de amor y rechaza sin demora la injusticia. Su honor consiste en no calcular nada, en distribuirlo todo a la vida presente y a sus hermanos vivos. Así dispensa con prodigalidad para los hombres futuros. La verdadera generosidad con el porvenir consiste en darlo todo al presente.

(...)

Con ello, la rebeldía demuestra que es el movimiento mismo de la vida y que no se la puede negar sin renunciar a vivir. Su grito más puro hace levantarse, cada vez, a un ser. Es, pues, amor y fecundidad, o no es nada.

Albert Camus

SUMARI

Agraïments	9
CAPÍTOL 1: Introducció	16
1.1 Plantejament.....	17
1.2 Objectius	23
1.3 Preguntes de recerca	24
1.4 Objecte d'estudi	25
1.5 Justificació	26
1.6 Estat de la qüestió	31
1.7 Marc teòric.....	36
1.8 Hipòtesi	43
1.9 Metodologia i tècniques.....	44
CAPÍTOL 2: La ideologia governa la tecnologia	46
2.1 Introducció al capítol 2	48
2.2 A la recerca dels límits actuals de la intel·ligència artificial	59
2.2.1 Restriccions bàsiques del processament de dades	60
2.2.2 La irracionalitat del científisme.....	65
2.2.3 Biaixos humans	69
2.2.4 Consciència i automatització.....	71
2.2.5 Computació de l'experiència i deshumanització	81
2.2.6 Racionalitat humanitzada	86
2.3 La necessitat d'una apropiació social de la tecnologia	90
2.3.1 Està passant: conseqüències de l'excés de delegació.....	91
2.3.2 Algoritmes com a eines de poder i perpetuació de les desigualtats.....	98
2.4 Per un lideratge cultural de la mediació algorítmica	103
2.4.1 Algoritmes com a generadors de cultura.....	104
2.4.2 Colonització comercial de la web original.....	107
2.4.3 Desintermediació i fluxos de coneixement.....	111
2.4.4 Periodisme i humanisme	115
2.4.5 Periodisme i tecnificació.....	116
2.4.6 Periodisme i crisi.....	117
Capítol 3: El paper dels algoritmes comercials a l'esfera pública	119
3.1 Introducció al capítol 3	121
3.2 Els efectes del domini de les grans tecnològiques	123
3.2.1 Descentralització i democràcia.....	124
3.2.2 Després de l'esperança va arribar la desconfiança	128
3.2.2.1 Engagement com a valor suprem de la mediació	139
3.2.2.2 Cambres d'eco computacionals.....	145
3.2.2.3 Desinformació algorítmica	150
3.2.2.4 Extractivisme que erosiona les llibertats	160
Capítol 4: El periodisme davant l'automatització	172
4.1 Introducció al capítol 4	174
4.2 Informar sobre intel·ligència artificial amb rigor	179
4.3 Vigilar els algoritmes	181

4.4 Recerca empírica i resultats: Reptes i recomanacions per dotar la intel·ligència artificial dels valors ètics del periodisme.....	185
4.4.1 Algoritmes a les redaccions	186
4.4.2 Metodologia de la recerca empírica.....	190
4.4.3 Reptes ètics sobre l'ús de la IA per fer periodisme.....	198
4.4.4 Estat de la qüestió a Catalunya.....	201
4.4.5 Anàlisi i resultats.....	207
4.4.5.1 Contingut automatitzat amb criteri editorial.....	208
4.4.5.2 Personalització que conviu amb la diversitat i promou la salut de l'esfera pública	211
4.4.5.3 Tractament responsable de les dades per preservar la privacitat	217
4.4.5.4 Supervisió i qualitat de les dades per evitar biaixos	221
4.4.5.5 La qualitat passa per posar en valor les facultats humanes	225
4.4.5.6 Poder de les plataformes i independència periodística.....	230
4.4.5.7 IA per enfortir els valors del producte periodístic.....	233
Conclusions	234
Bibliografia.....	243

AGRAÏMENTS

Des de la finestra de la biblioteca veig el parc on solem portar el nens a jugar. Mentre treballa en la tesi, des del segon pis de la sala d'estudi, veig córrer l'Alba i l'Aleix. El seu pare els ha dut després que la seva àvia els hagi regalat unes sabates, i em sento infinitament afortunada i agraïda. Em passa sovint.

Vaig començar aquesta tesi l'any 2019, uns mesos després que ens deixés la mare. Avui sóc més conscient que la seva mort va influir en la decisió, no sé fins a quin punt, però abans que passés ni m'ho plantejava perquè tot era per als nens, per a la feina i per a ella, qui ja feia 25 anys que no era ella. Era una dona valenta, de cor enorme i d'esperit independent que m'havia ensenyat, amb el seu exemple, que primer de tot som persones i després tenim nacionalitats, parlem diferents llengües o formem part de diferents grups. No va tenir l'oportunitat d'estudiar, però tota ella era expressió de la cultura que s'estimava. Filla de família republicana, catalana, treballadora i de la postguerra, la gran de cinc germans, per sortir de casa es va haver de casar i es va fer carnissera perquè treballar era per a ella el més semblant a la llibertat. La Matilde disposava de grandesa, coratge, sentit de l'humor i una sensibilitat estètica fora del comú en aquelles condicions. La seva humanitat la feia especialment sensible al classisme i no suportava la hipocresia. Va plantar cara sempre a grans dificultats alhora que mirava de tirar endavant sola amb les criatures. Tot plegat la va convertir en una dona inconformista i rebel amb una forta set de justícia social. Era impossible no endinsar-se en seus ulls blaus o somriure quan ho feia ella. La seva bellesa només li va portar problemes. No hi va haver final feliç.

Quan vaig decidir fer el doctorat, malgrat ja no hi fos, també hi va influir: exerceix la teva llibertat. Quan vaig triar el tema, també: tria allò que et surti del cor.

Sempre hi ha possibilitats de triar i exercir la llibertat, per mínima que sigui, ens diu Primo Levi, i això és el que hem practicat. Però les circumstàncies redueixen les opcions, i segons quines eleccions que poden semblar simples (com ara estudiar, o rebutjar una feina per optar a una altra més adient) per a algunes persones són factibles, mentre que per a d'altres poden ser un luxe. I va arribar un dia en què vaig poder triar, malgrat impliqués sacrificis i riscos. Afortunada i agraïda.

Tanmateix, posar-se a fer un doctorat en determinades condicions pot resultar ben bé un salt al buit. Però en l'empenta que vaig sentir hi vaig reconèixer la mateixa que tenia la mare i que

mirava de transmetre'ns: sempre ens animava a volar. Llavors ens donava força per remuntar si ens equivocàvem: coratge, empena i amor. Per això li vull dedicar aquest treball, que és fruit l'exemple que em va donar mentre la salut li va permetre.

Quan el professor Vidal Castell em va dir que li agradaria dirigir la meva tesi va ser un altre dels dies que he sentit aquest agraïment. Tenia molt clar que volia fer-la a l'Autònoma, on vaig estudiar sota la influència de mestres que em van ensenyar a començar a pensar: Francesc Burguet, Enric Marín, Carlos Zeller, Gemma Casamajó, Maria Corominas, Jaume Soriano, Elvira Teruel o, per descomptat, David Vidal, que és un d'aquells professors que amb una gran generositat, et donen la confiança que necessites, per molt que vegin que encara vas a quatre grapes. Va començar molt bé.

Un altre dels pocs salts al buit que recordo és quan vaig entrar a fer periodisme, amb 25 anys, quan en teoria ja era massa gran per començar carrera. I el salt va ser meravellós. Mai més vaig sentir-me tan connectada amb el que volia i amb els valors que dominaven al meu entorn com durant la carrera de periodisme, i que vaig poder fer gràcies a alguna beca i als ingressos que generava fent hostessa a la tele o a les fires, com a cambrera, dependent, cangur o embolicadora a perfumeries durant el Nadal.

Després va tocar buscar feina com a periodista i vaig anar a parar a *La Vanguardia* digital, el primer lloc que em van trucar després d'enviar els currículums. Des d'aleshores fins avui han passat 15 anys de dedicació a la comunicació digital, en projectes que m'han il·lusionat i dels quals he gaudit moltíssim. Tanmateix mai van consistir en fer el periodisme social i compromès que havia de servir per canviar les coses que un dia havia volgut. Però no és cap retret. El periodisme no passa per temps d'invertir en noves reporteres, i a més, per desgràcia, avui és molt difícil viure dignament fent periodisme compromès, així que potser millor així.

Fer la tesi era la manera de reconnectar amb el periodisme, i des de la humilitat, contribuir a enfortir aquesta professió —avui debilitada— en la qual sempre he cregut. Vaig voler aportar tots els aprenentatges que havia pogut fer en l'àmbit de la comunicació digital, l'ocupació en la qual m'he acabat especialitzant durant els anys en què he desenvolupat la meva vida laboral. Un temps en què he tingut la sort de poder viure en primera persona el naixement de les noves plataformes de comunicació digital: primer des del mitjà com a periodista gestora de continguts digitals, a continuació com a responsable de xarxes socials, també al mitjà, i després com a responsable de l'estratègia digital del departament de comunicació d'una de les grans empreses del sector financer, on es comptava amb els recursos necessaris per fer possible la innovació

que, llavors, en una tercera etapa, hem mirat d'aportar a l'emprenedoria i al sector cultural, on actualment desenvolupo la meva feina com a consultora i professora.

L'any 2018 vaig fer un *mooc* al UT Knight center de la Universitat de Texas amb el títol *News algorithms*, impartit pel professor de la Northwestern University de Chicago Nicholas Diakopoulos. Explicava com la intel·ligència artificial començava a fer-se servir a alguns mitjans de comunicació, de quines maneres, i com les periodistes podíem entendre-la: llenguatge computacional, conceptes bàsics i alguns bons exemples. Al Knight Center n'havia fet alguns més: direcció de producte periodístic i periodisme de dades. Tots ells han estat cursos molt bons, però el que va impartir Diakopoulos em va interessar molt i vaig seguir estirant del fil.

Més endavant he anat descobrint el perquè: a banda de les possibilitats de la IA per fer periodisme, vivim un moments de la història apassionants en què tenim la sort d'assistir a una nova revolució tecnològica caracteritzada per la creixent utilització de les tecnologies exponencials en pràcticament tots els àmbits, i per tant són moments en què tenim l'oportunitat de definir els fonaments que han de guiar els usos d'aquestes tecnologies. I no és un tema banal. A les nostres societats de consum s'explota tota solució que millori la productivitat i l'eficiència, i alhora que la IA té un gran potencial per millorar processos i fer-nos la vida més fàcil, també comporta grans riscos, el principal, probablement, la deshumanització que pot implicar l'automatització de les decisions. El periodisme hi té molt a dir, i mai estaré prou agraïda al professor David Vidal Castell, al Departament de Mitjans Comunicació i Cultura de la facultat de Ciències de Comunicació de la Universitat Autònoma de Barcelona, per haver-me donat l'oportunitat de fer una tesi per explorar en aquest tema que m'apassiona en aquests moments decisius.

El viatge ha estat fabulós i sorprenent. Una muntanya russa, fins i tot, de situacions increïbles viscudes que seran inoblidables. Començant per les dues persones a qui en David em va suggerir entrevistar fa tres anys: els doctors Salvador Alsius i Carlos Ruiz, els qui em van mostrar la millor cara de l'acadèmia: em van dedicar temps i em van donar les primeres orientacions. Llavors vaig sortir de les entrevistes com si haguessin encès les llums llargues en un camí fosc. El que no sabia aleshores és que aquests llums em guiarien fins avui. L'agraïment és immens. Això és l'acadèmia. No se si ho van arribar a verbalitzar, però és el que vaig sentir els cops que hi vaig parlar. Aquesta tesi conté molt d'ells.

Al 2019, quan vam començar, el projecte era importar els eventuais criteris ètics que mitjans internacionals podien estar aplicant ja en fer servir la IA, però va arribar la pandèmia. Fer els

contactes era difícil i no es podia viatjar. Baixada. Prova de seguiment: ja no tenia el full de ruta clar. I ara, què?

Jo faig vida a alguna xarxa social, on també estic en contacte amb persones de tot el món de qui aprenc i, en molts casos, no conec personalment. Una d'elles era el doctor José Alberto García Avilés, de la Universitat d'Elx, referent en innovació en periodisme i ètica. Era la única persona a l'Estat a qui havia sentit parlar d'IA aplicada al periodisme. Li vaig enviar un missatge directe i va tenir l'amabilitat de conversar una estona amb mi per telèfon. Reoriento. Torno a veure el camí. De nou, això és l'acadèmia, i tinc la sort i la certesa de trobar-me, altre cop, amb un gran mestre i gran persona, i soc conscient que és el millor que em pot passar. Afortunada i agraïda.

Passa algun mes i m'arriba un enllaç de Carlos Ruiz: el consell de la informació finlandès fa el primer pronunciament del sector sobre la necessitat de fer una reflexió sobre com adoptar la IA de manera responsable en el periodisme. Han encarregat un informe que es publicarà a finals de l'estiu del 2020. Arriba la tardor: el meu director m'avisa que els finlandesos han publicat un informe sobre ètica, IA i periodisme. El comparteixo amb Salvador Alsius, membre del Consell de la Informació de Catalunya. Pensem: no podríem treballar en una proposta similar reflectint la realitat catalana?

Parlem amb qui era aleshores president del CIC, Rafael de Ribot, i amb el secretari general, Josep Rovirosa. Ho veuen amb bons ulls. Presenta'ns un projecte. Entesos. Ho tirem endavant. De nou, agraïment màxim. El CIC fa realitat una idea i a més hi vol col·laborar. L'encàrrec coincidia amb la recerca, però amb la urgència que el camp d'estudi manifesta, el CIC demanava publicar abans de la defensa de la tesi. La possibilitat de portar a terme el treball empíric de la mà del CIC era una gran oportunitat per incentivar el diàleg entre el sector i la contribució que ens havíem proposat—que presentem en el quart capítol del treball que teniu entre les mans, d'una manera més elaborada i actualitzada—.

I portem a terme un informe sobre el tema en el qual, gràcies al CIC, hi participen els principals mitjans de comunicació i grans referents com Àlex Gutiérrez, Karma Peiró, José Alberto García Avilés, Albert Sabater, David Casacuberta, Ximo Blasco, Geni de Vilar, Mariano Fernández, Alberto Alejo i exponents mundials en periodisme computacional com Charlie Beckett i fins i tot Nicholas Diakopoulos, el meu primer professor sobre el tema al Knight Center. Vam presentar el treball al Col·legi de Periodistes i des d'aleshores l'hem pogut compartir a diferents fòrums i congressos que s'hi han interessat, inclús la Universitat de Columbia o l'*AI, Media and Democracy Lab* de la Universitat d'Amsterdam.

Quan em plantejava què podia expressar en aquests agraïments no podia pensar en cap cosa més fescinant que el camí que m'ha portat fins aquí perquè, de la mateixa manera que el periodisme em va acollir quan vaig fer el primer salt al buit quan vaig fer la carrera, ho ha tornat a fer en aquest segon salt. No vaig poder fer el periodisme que havia volgut, però de nou em sento estimada pel periodisme que jo m'estimo, i segurament el periodisme té raó i des d'aquí és des d'on millor puc contribuir-hi.

Aquesta tesi s'ha pogut portar a terme gràcies a la col·laboració de persones que, tot i estar molt ocupades, han tingut una estona per dedicar a les meves preguntes, com els professors David Casacuberta, Santiago Sánchez Migallón, Maria Corominas, Jaume Soriano i Gemma Casamajó. De grans professionals amb qui és un veritable privilegi poder conversar com Jaume Gurt, Eli Vivas, Joan Maria Morros, Marc Argemí o César Úbeda, així com les persones que han participat a la recerca, totes les que han contestat les enquestes i les que m'han ajudat a difondre-les.

Per descomptat, bona part de tot això ha estat possible gràcies al Consell de la Informació de Catalunya, i molt especialment a Josep Rovirosa i Begoña Muñoz, amb qui vaig treballar colze a colze, als dos presidents que ens van donar confiança, Josep Carles Rius i Rafael de Ribot i, per descomptat, als membres del CIC que s'hi van implicar i hi van aportar el seu profund criteri expert, Salvador Alsius, Dolors Massot i Ruth Rodriguez.

Als doctors i doctores que m'han fet valuosos retorns a les proves de seguiment: Enric Marín, Carme Farré, Josep Maria Blanco, Xavier Giró, Joana Gallego i Dolors Palau. Al Departament de Mitjans Comunicació i Cultura, a la Susanna i la Rosa, al Servei de Biblioteques de la UAB i en especial a la Biblioteca de Comunicació per ser-hi quan ha calgut, amb la millor resposta i les millors solucions a problemes que de vegades han estat difícils, i sempre transmetent que no estem soles, que hi són per ajudar, i que són persones humanes i no màquines. Afortunada i agraïda.

A la biblioteca Ignasi Iglésias, on he passat moltes hores estudiant, i per extensió al Servei de Biblioteques de la Diputació de Barcelona per posar-nos a l'abast aquests serveis públics extraordinaris, els espais i per facilitar d'una manera increïble que tinguem un univers de bibliografia a la nostra mà.

A les companyes de projectes laborals, per la seva flexibilitat i consideració, que han estat imprescindibles per complir amb compromisos relacionats amb la tesi.

A la meva família, a l'Eduardo, per ser-hi sempre, a la Isabel i el Mateo, a l'avi Antoni i la Yolmar per donar-nos suport i facilitar la conciliació sempre que ha estat a les seves mans. A totes les persones que m'estimo i que saben qui són i a les que han cregut en mi, perquè és gràcies a elles que avui soc aquí.

Deixo pel final a qui ha estat el meu gran còmplice en tot aquest procés, que és el Xavi, la gran persona amb qui he tingut la sort de compartir la meva vida. A qui em costa molt expressar tot l'agraïment que sento per, des d'una manera molt més pràctica de viure, comprendre que això era important per a mi. Per la seva extraordinària sensibilitat ara i sempre, per les traduccions del món i per la seva generositat en assumir els sacrificis que aquest projecte ha implicat.

Alba i Aleix: gràcies per entendre que he d'estudiar. Veure-us feliços és el millor que m'ha passat, i si de vegades he hagut de renunciar a passar temps amb vosaltres és perquè tinc la certesa que això també és una manera (la millor que sé) de transmetre-us allò que com a mare crec que és important deixar-vos.

CAPÍTOL 1

INTRODUCCIÓ

Eso de que el hombre se sienta disminuido, como a muchos les ocurre en la actualidad, por la inmensidad del universo o las interminables evoluciones del tiempo, equivale a asustarse de su propia sombra. Solo gracias a la luz de la conciencia resulta visible tal universo, y si esta luz desapareciese, solamente la nada quedaría. Fuera de la etapa iluminada por la conciencia humana, tal descomunal cosmos no es sino una existencia sin significado. (...) El significado vive y muere con el hombre, o más bien, con el proceso creador que lo trajo a la existencia y le dio una mente.

Lewis Mumford

1.1 PLANTEJAMENT

Algun cop m'he preguntat què hagués passat si un sistema d'IA hagués hagut de d'avaluar si podia ser una bona candidata per cursar una llicenciatura, per exemple, als meus 17 anys, l'època preuniversitària durant la qual feia dos anys que m'havia hagut d'independitzar per treballar a jornada complerta. Per sort cap màquina va influir en decisions sobre el meu futur, perquè els algorismes treuen patrons de dades del passat i per tant inclouen tots els prejudicis que aquestes arrossegueu. En canvi, sí que m'he trobat, en el transcurs de la vida, algunes persones que, sense prejudicis, han cregut en mi, i això ha estat imprescindible perquè avui estigui aquí.

Aquesta tesi parla d'ètica aplicada a la tecnologia, i en el rerefons del treball descansa la reivindicació del respecte per la llibertat de les persones, amenaçada sovint per qui ha mirat d'aprofitar la tecnologia per perpetuar relacions desiguals; o per la mateixa raó científico-tècnica, que adquireix noves dimensions totalitzadores amb el culte a la suposada objectivitat de les dades i que ha donat lloc a la patologia del *dataïsme*. Una suposada racionalitat que en les seves manifestacions més extremes considera l'home com a un ésser que pot ser determinat per prediccions algorítmiques.

Lewis Mumford (2010), un dels autors que inspira el nostre treball i apassionat del desenvolupament tecnològic, s'ha referit abastament també al risc de degradació espiritual que pot provocar l'enlluernament per la tecnologia, a la passivitat intel·lectual que aguditza la nostra vulnerabilitat a les amenaces dels usos interessats de la tècnica, així com als perills que poden implicar les cegues concessions d'autoritat desmesurada al producte de la màquina:

La mitología peculiar de nuestra era contemporánea se ha construido como reacción contra una insufrible confusión subjetiva: una mitología que otorga exclusivamente a las medidas cuantitativas y a las abstracciones lógicas las mismas propiedades mágicas que la mente primitiva atribuía a las pintorescas figuras retóricas. (p.157)

Amb aquest treball també pretenem reivindicar el dret de les persones a viure amb dignitat, que resulta d'una humanitat regida pel compromís amb els altres, que parteix del reconeixement de l'existència de quelcom més important i valuós que un mateix i que sintetitzà Kant en el seu imperatiu categòric: "Obra como si la máxima de tu acción debiera tornarse, por tu voluntad, ley universal de la naturaleza" (Kant, 2012, p.41). Una consciència que neix del reconeixement

d'un ordre i un sentit del qual formem part. Una moral que es forma a partir de la sensibilitat natural cap als sentiments dels altres, tal com descrivia Adam Smith:

El regocijo que nos embarga cuando se salvan nuestros héroes favoritos en las tragedias o las novelas es tan sincero como nuestra condolencia ante su desgracia, compartimos sus desventuras y su felicidad de forma igualmente genuina. Sentimos con ellos gratitud hacia los amigos fieles que no los desertaron en sus tribulaciones, y de todo corazón los acompañamos en su enojo contra los pérfidos traidores que los agraviaron, abandonaron o engañaron. (...)

Lástima y compasión son palabras apropiadas para significar nuestra condolencia ante el sufrimiento ajeno. La simpatía, aunque su significado fue quizá originalmente el mismo, puede hoy utilizarse sin mucha equivocación para denotar nuestra compañía en el sentimiento ante cualquier pasión. (Smith, 1997, pp.51-52)

El concepte de moral l'apliquem en aquest relat des de la visió de l'ètica naturalista que el concep com aquest instint d'ajut i suport mutu que al mateix temps és allò que ens permet sobreviure com a espècie. De considerar el que és bo o dolent per a la comunitat i per tant fa possible la vida en comú. Perquè la cooperació condueix al desenvolupament del sentit de la justícia, inevitablement acompanyat pel sentit de la igualtat (Kropotkin, 2017). A mesura que van desapareixent les diferències de classe i els estereotips, s'obre camí la idea que els drets d'un individu determinat són inviolables com els de qualsevol altre. D'aquesta manera es pot obrir pas el "derecho a ser" que reclama Bauman (2010) ja en el context del segle XXI: "Que se nos deje en paz significa, en primer y más destacado lugar, tener derecho a autodefinirnos y a autoafirmarnos, y disponer de una oportunidad realista de actuar para hacer efectivo ese derecho" (p.23).

Avui ja és habitual fer servir algorismes per decidir qüestions amb implicacions crítiques en la vida de les persones. L'enorme quantitat de dades disponibles, junt amb l'augment de la capacitat de computació, ha donat lloc a tecnologies que estan en condicions de dur a terme tasques que fins fa poc semblaria impensable poder delegar en les màquines. Es tracta d'aplicacions que tenen un gran potencial per aportar millores en camps com la medicina, la justícia o en l'educació i, per descomptat, també en el camp de la comunicació. Tanmateix també es tracta d'eines que, com qualsevol altra, tenen el potencial de fer-se servir, o no, de manera adient.

Aquesta recerca emergeix de la convicció sobre la necessitat de reflexionar en torn a com volem afrontar aquesta nova revolució tecnològica que protagonitzen les variants d'allò que s'anomena intel·ligència artificial (IA) en una societat que, des que ha estat dominada pels valors capitalistes, ha assimilat la innovació tecnològica perpetuant relacions de poder desiguals. Prioritzant criteris de productivitat per al benefici d'una minoria (Piketty, 2019) alhora que n'ha socialitzat els costos, per exemple, en danys al medi ambient (Feenbeg, 2002) i d'altres com la precarietat laboral que impedeix a moltes persones del planeta comptar amb uns mínims per ni tan sols cobrir despeses bàsiques. Un sistema que està en l'origen de la infelicitat que es manifesta de manera creixent a nivell global (*Gallup Global Emotions*, 2022).

El camp de la comunicació i la cultura és segurament un dels àmbits en què les tecnologies d'IA i altres aplicacions algorítmiques estan tenint una incidència més evident, perquè han causat una autèntica revolució en l'ecosistema informatiu, que s'ha caracteritzat pel traspàs a les grans plataformes tecnològiques del rol de mediació cultural, així com el paper de *gatekeeper* que tradicionalment havia exercit el periodisme.

Un dels grans avantatges de la digitalització implicà tenir més eines a disposició per produir contingut, així com més canals per ser distribuït. Quan les grans plataformes van emergir tot feia preveure que més veus serien escoltades i per tant, gràcies a una esfera pública més plural, gaudiríem de millors democràcies. Això ha estat així en part. Les possibilitats d'aquests nous mitjans han impulsat l'èxit de determinats projectes, i fins i tot han arribat a influir en esdeveniments de gran rellevància política i social, com la Primavera Àrab, el 15-M i d'altres causes. Alhora, tanmateix, també hem comprovat que més informació no implica una esfera pública més sana, i molt menys una ciutadania més conscient, informada i participativa. Com exposarem en el nostre treball, ens situem més aviat en un escenari que podria arribar a assemblar-se a la distòpia huxleiana en la qual la gent adora una tecnologia que té el potencial d'anul·lar la capacitat de pensar, ja que la informació de qualitat que té la intenció d'explicar i contrastar fets no interessa i per tant resta submergida, en paraules de Neil Postman (2018) per un mar d'irrellevància.

Totes aquestes opcions de creació de continguts, unides a l'aparició de sofisticats dispositius mòbils, han convertit cada individu en un mitjà de comunicació. Així, com a usuàries digitals, ens trobem exposades a quantitats exponencials de dades sense fi. Un *data tsunami*, en paraules de José María Lassalle (2019) que, a més, esdevé l'aliment dels algoritmes d'aquestes plataformes. Intel·ligències artificials que s'han convertit en generadores de cultura (Finn, 2018) que filtren i

jerarquitzen informació amb els criteris comercials de les grans corporacions globals capitalistes que en són propietàries. Unes empreses que, en alguns casos, acumulen poder per sobre dels Estats (Ruiz-Caballero, 2016).

En aquest context, el periodisme s'enfronta a una situació sense precedents: una crisi econòmica, i també ontològica, que interpel·la el sector a trobar el seu lloc i a seguir realitzant la seva funció social en aquest nou escenari d'escassetat: de recursos econòmics i de confiança de la ciutadania. Una nova etapa marcada per la cultura algorítmica que interpel·la el periodisme a recuperar el seu propòsit, que no és altre que proporcionar informació veraç i espais per al diàleg comunitari en favor d'una ciutadania il·lustrada, en condicions de governar-se a si mateixa, de prendre decisions informades i, d'aquesta manera, contribuir a la salut dels sistemes democràtics (Ruiz-Caballero, 2016). En aquest sentit, val la pena recuperar el concepte d'utilitat de la informació amb què Neil Postman es referia a l'impacte cultural de l'aparició del telègraf i els seus efectes socials en ser aplicat a la comunicació, perquè tot i que es tracta d'una anàlisi feta fa ja algunes dècades, podria servir per retratar l'esfera pública ensordidora que habitem aquests dies:

La información se ha convertido en una forma de basura que no sólo no es capaz de dar respuesta a las cuestiones humanas fundamentales, sino que rara vez demuestra su utilidad a la hora de buscar soluciones a los problemas más prosaicos. (...) El entorno en el que florece el tecnopolio es aquel en el que el vínculo entre la información y los fines humanos se ha roto. Es decir, que la información aparece de forma indiscriminada, sin dirigirse a nadie en particular, en cantidades ingentes y a velocidades enormes; amén de no tener vínculo con ningún tipo de teoría, significado o fin. (Postman, 2018, p.101)

Aquesta tesi neix de la certesa que el periodisme segueix sent necessari per a la salut de les democràcies. Que precisament en aquests temps de sobreabundància de missatges i campanyes d'infoxicació es requereixen interlocutors en els quals confiar a l'hora d'informar-se, i que és el periodisme qui té sentit que assumeixi aquest paper. Tanmateix, primer haurà de recuperar la credibilitat perduda com a conseqüència de la seva pròpia acció, històricament més lleial als poders fàctics. Aquesta tasca potser ni tan sols serà possible en alguns casos. Malgrat tot, defensarem aquesta alternativa ja que, per les raons que exposarem al llarg del treball, el context comunicatiu actual —caracteritzat per la irrupció de la comunicació computacional comercial— ofereix l'oportunitat als mitjans tradicionals d'erigir-se com a font acreditada d'informació i esdevé propici per a la confluència de diferents interessos orientats a la credibilitat del producte informatiu. Així, una de les maneres de recuperar confiança passarà per

redefinir amb encert el seu paper davant les tecnologies exponencials. Tal com exposarem amb detall a la proposta d'acció que desenvoluparem al capítol final, la professió periodística haurà de començar per conèixer aquesta tecnologia, així com els debats ètics relacionats amb les seves aplicacions, tant amb una mirada general com des de la perspectiva de l'ètica de la comunicació. En primer lloc, per informar-ne amb rigor, en segon terme, per vigilar-ne els seus usos i, finalment, per fer-ne una apropiació responsable i, per tant, exemplar.

Com veurem a les properes pàgines, els treballs que exploren la intersecció entre ètica, IA i periodisme són escassos. És per això que hem optat per tractar aquesta qüestió, i lluny de pretendre limitar la innovació periodística amb normes o regulacions, la present recerca confia que aquesta tecnologia, com d'altres que l'han precedit, té la capacitat de contribuir a fer un millor producte informatiu, tot i que serà precisament el govern ètic dels algoritmes allò que ajudarà a garantir aquesta qualitat.

La tesi que llegiu vol ser un exercici d'investigació amb la finalitat fer una proposta constructiva que fugi de teories tecnofòbiques i aspira a contribuir a la veritable racionalitat (Feenberg, 2002; H. G. Gadamer, 2010; Marcuse, 1993). La qual és capaç de participar en el disseny d'una tecnologia que faciliti i promogui de les finalitats humanes, que en l'àmbit del periodisme es tradueixen en situar la funció social que li és reconeguda per sobre dels interessos comercials. Alhora, defensem que posar la lleialtat a l'audiència per davant dels objectius econòmics no ha de resultar contradictori amb una estratègia a llarg termini, com ja ha passat en altres moments de la història (Braojos Garrido et al., 1999; Kovach i Rosenstiel, 2014).

Així doncs, ens orientarem a identificar marcs per desenvolupar una innovació responsable, a partir d'un procés de reflexió conjunt del sector per plantejar-nos reptes a l'alçada dels mitjans que tenim al nostre abast amb una mirada a les finalitats que els hi han de donar sentit. En paraules de Duch i Chillón "vindicar un «*ejercicio de la humanidad*» rebelde (...) que aprenda a utilizar la potencia tecnológica en aras de los fines (...) que integran nuestro mejor legado" (Duch i Chillón, 2012, p.468).

Per tal de poder examinar aquest tema, hem considerat necessari adoptar coneixement provinent de diferents disciplines, i situarem el nostre objecte d'estudi entre l'ètica periodística, la tecnoètica, la filosofia, els estudis culturals, els estudis sobre ciència i tecnologia, el periodisme computacional i la innovació en periodisme. Una interdisciplinarietat que considerem indispensable per abordar el problema que ens proposem resoldre amb una mirada transversal i per consegüent oberta i integradora. Tanmateix, aquest treball es planteja principalment des

de la comunicació, i per tant hauran de quedar fora de l'abast del projecte importants fonts d'informació existents en aquests altres profunds àmbits de coneixement.

Durant el camí ens hem trobat amb la dificultat afegida que el nostre objecte d'estudi (les tecnologies i aplicacions que es deriven de la IA) està en ple procés d'integració a la societat i per tant es tracta d'un context fèrtil i canviant, així com el material que el nodreix (tant el relacionat amb els aspectes tècnics com amb les qüestions relatives a les regulacions). Així doncs, abastar-lo al mateix temps que s'efectua un procés de cerca de profunditat resulta una tasca complexa, tant en la tria de determinats materials com en la renúncia a d'altres. Així, aquest treball es pot considerar un intent de capturar els trets fonamentals que permetin actuar davant d'un panorama, el de la innovació tecnològica aplicada, que es troba en ràpida evolució.

Aquesta tesi també pretén obrir la mirada i plantejar reflexions sobre com el periodisme pot afrontar aquesta nova revolució tecnològica que es caracteritza per la creixent automatització a la societat. Ho farà sobre la base de la selecció d'un marc teòric concret i d'unes idees determinades, però està lluny d'abastar totes les que serien pertinents. Com més hem explorat, millors marcs hem identificat i més idees hem pogut aportar.

Així doncs, en les properes pàgines pretenem desenvolupar unes bases que permetin definir marcs d'acció escalables i obrir camins en un terreny on queda molt per explorar, i que esperem resulti fèrtil per a recerques futures. Un punt de partida per seguir avançant i sembrant, tant per a la investigadora que escriu com, esperem, per a les professionals de la comunicació en el sentit més ampli del seu significat.

1.2 OBJECTIUS

Un cop exposades les línies bàsiques del plantejament d'aquesta recerca, enumerarem els objectius que ens proposem assolir:

O1. Proporcionar eines teòriques i pràctiques per governar els algoritmes a través dels valors ètics del periodisme per contribuir a què la professió:

O1.1 recuperi la centralitat en la definició de l'agenda tot generant confiança a través de línies d'acció basades en l'ètica de la comunicació.

O1.2 es converteixi en referent en tecnocràcia de la comunicació.

O2. Explorar les principals qüestions teòriques, que són de tipus filosòfic, que afecten la incorporació de les tecnologies algorítmiques per poder proposar intervencions en la línia assenyalada.

O3. Determinar els criteris ètics que el periodisme pot aplicar en l'ús de la intel·ligència artificial.

O4. Descriure i analitzar com s'està incorporant la IA als mitjans de comunicació.

O5. Identificar reptes i oportunitats que els professionals dels mitjans reconeixen en l'adopció de la IA.

O6. Incentivar una reflexió en el sector en torn als reptes ètics que implica l'adopció de la IA a les rutines periodístiques.

1.3 PREGUNTES DE RECERCA

A partir d'aquests objectius, ens hem proposat respondre les següents preguntes de recerca:

P1. Quin ha de ser el paper del periodisme davant la creixent penetració de la IA a la societat?

P2. Quins efectes en l'esfera pública i quins riscos pot implicar que les grans plataformes de mediació algorítmica hagin assumit el rol de prescripció informativa i cultural?

P3. Quin ha de ser el paper del periodisme davant la creixent mediació informativa i cultural algorítmica?

P4. Quins criteris ha d'aplicar el periodisme en l'ús de les tecnologies que es deriven de la IA?

P5. Com s'estan fent servir els algoritmes per fer periodisme?

P6. Com valoren els professionals dels mitjans la incorporació de la IA per fer periodisme?

P7. Com podem integrar la recerca i la regulació en ciberètica en l'ètica periodística?

P8. Quines eines i/o aplicacions algorítmiques comporten més riscos ètics?

P9. Quins dilemes ètics concrets generen?

P10. A quins grans principis de l'ètica periodística afecten?

P11. Emergeixen i/o es resignifiquen determinats principis?

P12. Quina mena d'eines podem posar a disposició dels mitjans que puguin adaptar-se a les múltiples casuístiques que es poden donar amb els usos d'IA i altres aplicacions algorítmiques?

1.4 OBJECTE D'ESTUDI

Aquest treball té com a objecte d'estudi el govern ètic de les tecnologies d'intel·ligència artificial per part del periodisme de qualitat i la seva incorporació responsable a les rutines de producció i distribució de la informació periodística. Convé aclarir que fem referència, concretament, als algoritmes que es fan servir als mitjans de comunicació i no els propis de les plataformes de mediació algorítmica. Tot i que la IA en general o les IA que fan servir les plataformes digitals de comunicació no són l'objecte d'estudi, sí que farem una incursió en aquests àmbits per exposar els potencials riscos de determinats usos d'aquesta tecnologia i en especial de la mediació computacional i els seus efectes en l'esfera pública quan és optimitzada només per assolir objectius comercials. D'aquesta manera estarem en condicions de fer les propostes d'acció pertinents des de l'ètica de la comunicació.

Així doncs, la culminació pràctica del projecte (que es desenvolupa a la recerca empírica) es centrarà en els algoritmes que es desenvolupen als propis mitjans o que, malgrat hagin estat dissenyats per proveïdors externs, es fan servir en els processos de la cadena de valor del producte periodístic. En particular, ens proposem identificar quines eines o aplicacions són les que ocasionen més preocupacions ètiques per, després, analitzar quins dilemes generen i a quins grans principis ètics del periodisme afecten. La recerca es porta a terme en la franja temporal 2019-2022, precisament en els moments en què l'ús de la IA als mitjans de Catalunya es troba en plena introducció i quan ja comencem a comptar amb certa experiència entre el sector que pretenem aprofitar per posar en comú i així establir unes bases de governança algorítmica en l'àmbit del periodisme.

Ara que ja tenim el què, veurem el perquè.

1.5 JUSTIFICACIÓ

L'ús de la intel·ligència artificial als mitjans de comunicació és ja una realitat i es preveu, encara, que el procés d'integració s'acceleri i es consolidi en els propers anys: recents informes i enquestes indiquen que el sector incrementarà les aplicacions d'IA (Newman, 2022). La IA es considera una de les innovacions més destacades dels darrers temps també a Espanya (Carvajal et al., 2022) i cada cop més mitjans preveuen adoptar aquesta tecnologia, com veurem més endavant. A les principals capçaleres de Catalunya els algoritmes ja participen en processos de tota la cadena de valor, i els avantatges que percep el sector sobre el seu potencial per optimitzar les dinàmiques internes i la difusió del contingut auguren una gran transformació de les rutines periodístiques en el futur immediat.

Avui ja és habitual delegar als algoritmes tasques com identificar temes noticiables, analitzar i organitzar dades que es fan servir com a font, facilitar processos com transcripcions i traduccions, generar continguts escrits i infografies, triar títols, orientar la redacció periodística del contingut, moderar comentaris, publicar en nom del mitjà a través dels perfils a les xarxes socials o personalitzar i recomanar continguts als usuaris, entre d'altres. Hem arribat ja a poder preguntar a la IA coses com: què és noticable?, quina forma ha de tenir?, quin títol triar?, quin contingut destacar? És a dir, la IA pot arribar a tenir un paper cabdal en decisions que formen part del nucli de la funció editorial del periodisme.

Segons un dels principals informes internacionals sobre el tema (Beckett, 2019), en el transcurs dels propers anys, la IA contribuirà a fer millor periodisme de diferents maneres: farà la producció de contingut més eficient, permetrà trobar més notícies entre les dades, la moderació de comentaris serà més manejable, farà possible el reconeixement automàtic d'informació falsa, i ajudarà a distribuir millor els continguts periodístics o introduir preus dinàmics als anuncis i a les subscripcions, entre d'altres avantatges.

Així mateix, diversos estudis adverteixen que un dels principals reptes en la integració de la IA a les redaccions és de caràcter ètic (Beckett, 2019; De Lara et al., 2022; Salazar, 2020) perquè l'èxit en l'adopció de les eines i pràctiques algorítmiques passarà per la capacitat dels mitjans de comunicació de posar-les al servei dels seus valors i criteris editorials. En aquest context, es fa necessària una reflexió tant des de la indústria com des de l'acadèmia, sobre les condicions laborals, nous rols, la necessitat de redefinir la formació, entre d'altres (López i Tejedor, 2019). Així mateix, caldrà parar atenció a les promeses de productivitat que s'associen amb la IA, ja que

poden fer que es prenguin decisions financeres a curt termini que poden posar en risc el valor del producte i els estàndards periodístics de qualitat.

L'informe internacional esmentat detalla també certes preocupacions concretes del sector relacionades amb els riscos que implica l'ús de tecnologies exponencials, com ara dubtes compartits per la professió sobre si l'estalvi que pot comportar l'automatització de processos realment s'invertirà en un millor periodisme, o aspectes relacionats amb el potencial de generar biaixos algorítmics, desinformació o filtres bombolla.

En analitzar la literatura internacional sobre el tema, s'observen inquietuds similars relacionades amb l'ús de la IA a les redaccions: preocupen els riscos associats a la manca de supervisió de continguts generats de manera automatitzada, el potencial impacte del canvi en els fluxos de treball, l'assumpció de noves responsabilitats legals, les creixents diferències en les habilitats que es requereixen per gestionar aquesta nova àrea d'especialització i el potencial dels algorismes de generar biaixos (Marconi, 2020).

Aquí les preocupacions no són gaire diferents. Així ho demostren els resultats de l'enquesta que vam distribuir entre professionals dels mitjans de comunicació de tot l'Estat i que fa palesa, entre d'altres coses, una evident inquietud del sector respecte als efectes de la IA en la qualitat del periodisme. Més endavant n'exposarem les conclusions.

En els moments en què s'inicia aquesta tesi, ni col·legis de periodistes ni consells de la informació no han establert criteris sobre l'ús dels algorismes a les redaccions, amb l'excepció del Consell de la Informació finlandès, que l'any 2021 va publicar un informe en què, a més d'identificar també certes disquisicions ètiques, conclouia la necessitat de desenvolupar un marc autoregulatori en el sector, en part per evitar que siguin altres institucions —la UE o les mateixes plataformes (sic)— les que ho acabin fent (Haapanen, 2021). Ja entrat el 2022, moment en què presentem aquest treball, es fa més palesa la creixent inquietud respecte d'aquesta qüestió, com demostra el nou informe del Consell finlandès en el qual avança, a més, que els consells de la informació de Bèlgica, Alemanya i el del Regne Unit preveuen incorporar recomanacions ètiques per a l'ús de la IA als mitjans (Juntunen, 2022), en línia amb el treball que hem elaborat en el marc d'aquesta tesi per al Consell de la Informació de Catalunya (Ventura-Pocino, 2021).

Tenint en compte, a més, que la recerca nacional sobre el tema és força escassa (Parratt-Fernández et al., 2021) i que el sector considera necessari abordar aquest tema (Heinrichs, 2018; López i Tejedor, 2019; Porlezza, C. i Ferri, G., 2022; Salazar, 2020) vam considerar oportú

estimular un procés conjunt de reflexió i diàleg que permetés identificar els principals reptes relacionats amb l'ús de la IA aplicada a la gestió de la informació periodística. Així doncs, aquest informe té l'objectiu de proporcionar un punt de partida i un marc de referència per tal que cada mitjà pugui abordar les múltiples casuístiques que es poden derivar de la incorporació de la IA a les rutines periodístiques i que ajudi els professionals a establir uns criteris que supeditin la tecnologia als principis que regeixen la pràctica del periodisme ètic i, per tant, de qualitat (Alsius, 1998; García-Avilés, 2021).

Aplicats a la gestió de la informació, els algoritmes resulten molt eficients a l'hora de desenvolupar tasques d'associació, prioritització, filtratge i classificació de quantitats ingents d'informació (Diakopoulos, 2019). Aquesta tecnologia és, de fet, la que fan servir les plataformes —cercadors, xarxes socials i altres— per organitzar l'abundància de continguts característica del món digital. Els mateixos algoritmes que avui governen els fluxos d'informació que circulen per l'esfera pública (Gillespie, 2014), donen forma a la vida quotidiana, influeixen en la percepció del món, orienten el comportament de les persones (Agudo i Matute, 2021; Just, N i Latzer, 2016) i han donat lloc al que ha estat definit com una nova cultura algorítmica (Striphas, 2015) en la qual les plataformes, gràcies a aquesta eficiència tecnològica, han adquirit el rol de *gatekeeper* que havia exercit tradicionalment el periodisme (Negredo et al., 2020).

Aquest canvi ha reconfigurat l'ecosistema informatiu i ha donat lloc a un nou escenari en què disposem de més eines per crear i intercanviar coneixement, unes plataformes a les quals, a més, es va atribuir un paper primordial en la consecució de les fites de grans moviments socials del segle XXI com el 15 M a Espanya o la Primavera Àrab (Muñoz, 2011; Tufekci i Wilson, 2012), que no sol es van servir de xarxes socials virtuals per aconseguir notorietat i adhesions, sinó que també els hi van fer jugar un rol fonamental en l'organització interna dels propis moviments (Tufekci, 2013). Els mateixos mitjans de comunicació també s'han beneficiat de les xarxes socials i altres plataformes per identificar fonts d'informació i seguir l'actualitat de què informen. Tanmateix, malgrat tots els avantatges que comporten, els mitjans algorítmics s'han acabat demostrant tan o més vulnerables a la propaganda i altres *hackejos* que l'existent al món predigital (Bradshaw i Howard, 2019; Carrie Wong, 2019; Petit, 2018; Tufekci, 2017; Wardle i Derakhshan, 2017).

Avui, a l'esfera pública que dominen les plataformes, prolifera el contingut de baixa qualitat i tota mena de desinformació, incloent-hi els emergents *deepfake*. Una infoxicació que existeix en gran part com a conseqüència d'una gestió comercial d'aquestes companyies que es basa en

optimització tecnològica per assolir uns objectius que, aplicats a l'àmbit de la comunicació, es tradueixen bàsicament a aconseguir temps en els seus dominis. Aquest negoci de l'atenció afavoreix amb visibilitat qui vulgui escampar indignació, apel·li als biaixos i preferències existents i, per tant, a la desinformació (Tufekci, 2018) encara que sigui a compte de permetre que actors interessats enterboleixin l'esfera pública, generin desconfiança o atïïn la polarització social. En termes generals, es tracta d'unes empreses que no coneixen límits en l'ús del disseny persuasiu que aprofita vulnerabilitats psicològiques (Bridle, 2020; Harris, 2017; Patino, 2020; Wu, 2020) en algun cas, fins i tot, malgrat saber que poden generar dependència tecnològica o agreujar problemes mentals (Seetharaman, 2021).

Per bé que no totes les tecnològiques comparteixen polítiques idèntiques –i, per tant, no poden ser valorades de la mateixa manera–, a grans trets es tracta d'organitzacions que basen el seu creixement en una economia extractiva de les dades personals (Tufekci, 2017) que ha donat lloc al *capitalisme de vigilància* (Zuboff, 2019). Aquesta lògica s'ha beneficiat de la manca de regulació per apropiant-se de les macrodades que produeix el tràfic d'informació (Bridle, 2020; Guardiola, 2019; Lassalle, 2019) fins a arribar a assolir un poder econòmic i polític que arriba a superar el dels Estats. Es tracta d'unes sobiranes mercantils, en paraules de Carlos Ruiz (2016), que els internautes accepten, entretinguts i seduïts per la gratuïtat, i paguen el peatge de la seva llibertat quan permeten de manera voluntària l'erosió de la seva privacitat. Segons l'assagista Íngrid Guardiola (2019) la mediació de les plataformes en les relacions amb els altres implica fins i tot una "pèrdua d'autonomia individual i de capacitat d'emancipació col·lectiva, alhora que alimenten el poder de les corporacions de la tecnologia i les comunicacions, (...) de les elits econòmiques que compren les dades d'aquestes empreses (...) i el poder polític que les requereix quan creu oportú" (pp.11-12).

Davant d'això, al periodisme li toca preguntar-se quina posició ha de prendre davant l'arribada de les decisions automatitzades (Thumler (coord.) et al., 2020) de quina manera mantenir la seva centralitat i com proporcionar confiança a la ciutadania. Ha de plantejar-se quin paper adoptar, no només en el camp de la comunicació propi, sinó pel que fa a l'ús de l'automatització en totes les esferes en què ja s'aplica.

Avui, a Catalunya, a la IA li encarreguem tasques com seleccionar candidats per a ofertes de feina, participar en diagnòstics mèdics, tenir cura de persones dependents, concedir crèdits (Peiró, 2020) o estimar el risc de reincidència de reclusos (Bellio, 2021). En l'àmbit estatal, el Servicio Público de Empleo (SEPE) fa servir algoritmes per decidir quines persones tenen dret a

cobrar la prestació d'atur. La policia també n'utilitza, per exemple, per detectar denúncies falses sobre robatoris o per prevenir la violència de gènere (*The OASI Register – The Observatory of Algorithms with Social Impact – OASI, 2021*). Alguns d'aquests sistemes ja han estat qüestionats per biaixos en les seves prediccions, com per exemple l'algoritme que pronostica les possibilitats que adolescents infractors tornen a delinquir (Tolan et al., 2019).

Tenint en compte l'alt grau d'incidència d'aquestes decisions automatitzades en la vida de les persones, al periodista també li correspon plantejar-se com exercir el seu rol democràtic davant la possibilitat que aquestes aplicacions d'IA es facin servir sense ètica i, per tant, acabin vulnerant drets fonamentals. En coherència, el periodisme haurà de considerar quin ús fa de la tecnologia que haurà de vigilar, així com adquirir les habilitats necessàries i desenvolupar tècniques que ja comencen a utilitzar mitjans capdavanters per auditar aquests algoritmes (Diakopoulos, 2015; Trielli i Diakopoulos, 2020).

Potser ha arribat una nova forma de contribuir a mitigar la crisi de credibilitat que arrossegueu els mitjans de comunicació (Amoedo et al., 2021) i de proporcionar confiança a una ciutadania més necessitada que mai d'identificar fonts dedicades a verificar informació (Kovach i Rosenstiel, 2014). Posar les tecnologies emergents al servei dels valors que regeixen el periodisme de qualitat pot ser també una oportunitat. En un ecosistema informatiu necessitat de confiança (Mont'Alverne et al., 2022), l'estratègia de diferenciació pot passar també per afegir a la funció de servei públic del periodisme el compromís de liderar la innovació en tecnoètica de la comunicació.

1.6 ESTAT DE LA QÜESTIÓ

Més enllà del mite sobre el fet que la IA arribi a superar en intel·ligència l'ésser humà i acabi dominant-lo, hi ha qüestions que val la pena tractar tenint en compte la rapidesa amb què aquestes aplicacions algorítmiques es desenvolupen i s'integren en pràcticament tots els sectors. En molts camps la IA ja resulta imprescindible i nombrosos casos demostren que es tracta d'una tecnologia que, com qualsevol altra, té el potencial de proporcionar grans beneficis socials.

Cal considerar, no obstant això, que la IA és una tecnologia exponencial, a la qual tendim a encomanar-li cada cop més decisions abans només realitzables per part d'éssers humans amb conseqüències que poden arribar a influir de manera determinant en la vida de les persones. Abans, a un algoritme li solíem encarregar tasques de solució correcta —com ara ordenar una llista per data o per nom, o realitzar un extracte bancari—. Avui, en canvi, la IA participa en decisions com el diagnòstic mèdic, l'educació, els recursos humans o en d'altres relacionades amb l'àmbit de la justícia (*The OASI Register – The Observatory of Algorithms with Social Impact – OASI*, 2021) amb l'afegit que els procediments algorítmics per portar a terme aquestes decisions tot sovint resulten *caixes negres*, és a dir, algoritmes impossibles d'explicar a causa de la complexitat i als nivells d'autonomia que arriben a adquirir dels sistemes d'aprenentatge automàtic i aprenentatge profund (Torras, 2020). Per això, una de les principals característiques que comparteixen les guies de ciberètica és la prevenció de danys a les persones que podrien causar els errors dels sistemes automatitzats desproveïts de la capacitat humana per fer valoracions ètiques.

Aquesta voluntat es concreta en els principals codis de conducta, de manera transversal, en el respecte pels principis de justícia, privacitat, transparència, explicabilitat i de retiment de comptes (Hagendorff, 2020). En aquest mateix sentit, la Comissió Europea (CE) estableix set principis: agència humana i supervisió, robustesa tècnica i seguretat, privadesa i govern de dades, transparència, diversitat, no discriminació i equitat, benestar social i mediambiental i retiment de comptes (High-Level Expert Group on AI, 2019).

Encara que encarreguem decisions als algoritmes, l'autoritat moral de les seves aplicacions només pot correspondre a la persona (Boden, 2017; Mántaras, 2021) , i és per això que a les guies de ciberètica també se subratlla la necessitat de supervisió en les diferents etapes dels processos de disseny i implementació. Des del punt de vista legal, el Reglament General de

Protecció de Dades (RGPD) reconeix el dret de les persones davant l'automatització a l'article 22: "Tota persona té el dret a no ser objecte d'una decisió basada únicament en el tractament automatitzat, inclosa l'elaboració de perfils, que tingui efectes jurídics sobre l'interessat o l'afecti significativament" (*Reglament general de protecció de dades*, 2016).

Al mateix temps que els usos de la IA s'han sofisticat progressivament, s'han desenvolupat i adaptat marcs reguladors ètics i legals. Un dels passos més importants ha estat la proposta per a la regulació de la IA de la UE. Es tracta de tot un seguit de mesures per abordar les oportunitats i els reptes de la IA, orientades a assolir confiança en la tecnologia i en el seu potencial impacte tant en els ciutadans de manera individual com en la societat i en l'economia. La proposta de la CE, publicada el mes d'abril del 2021, planteja com garantir la seguretat, la transparència i la responsabilitat per tal d'evitar qualsevol forma d'injustícia, així com el respecte pels drets fonamentals (European Commission, 2021).

Aquest Reglament estableix unes categories de risc en funció del potencial impacte de l'aplicació d'IA en les persones d'acord amb els Drets Humans. Per exemple, són considerats d'alt risc, en el nou reglament, els sistemes de reconeixement facial en espais públics o els que tenen la capacitat de modificar el comportament humà. Això implicarà que estaran subjectes a requisits molt severs abans de poder sortir al mercat. Espanya s'ha postulat com a camp de proves d'aquest Reglament (C. del Castillo, 2022). Un projecte pilot que està previst iniciar aquest 2022 i que servirà per extreure aprenentatges per posar a disposició de la UE i facilitar que no agafi els Estats membres desprevinguts en el moment d'aplicació del Reglament a principis del 2024.

Aquest projecte pilot que liderarà Espanya també té la funció d'ajudar a definir el funcionament d'una nova agència estatal per a la supervisió dels algorismes (C. del Castillo, 2021), una iniciativa que ha posat en marxa la Secretaria de Estado de Digitalización e Inteligencia Artificial que avui lidera Carme Artigas i que té com a finalitat exercir control sobre els sistemes d'IA a nivell estatal i, posteriorment, en coordinació amb altres possibles entitats similars que es preveu que es desenvolupin en els espais de la UE.

El fet és que malgrat existeixi una certa regulació i un significatiu consens global en torn als Objectius de Desenvolupament Sostenible (ODS), el ritme d'adopció de la IA va per davant de les consideracions legals i ètiques que convindria per valorar els riscos que comporta el seu ús, i les nostres societats de consum són terreny assaonat per a les innovacions que prometen més productivitat i eficiència. Aquest context fa que s'actui de manera reactiva. És a dir, és habitual

que es facin servir aplicacions que no estan prou madures perquè no s'han avaluat de manera suficient des del punt de vista de la tecnoètica (Torras, 2020).

Moltes de les conseqüències d'aquest problema són conegudes. Comença a ser freqüent conèixer casos de discriminació i altres errors algorítmics, dels quals alerten des de fa un temps expertes i acadèmiques (BCN Analytics, 2018). Ja són habituals notícies com la de l'algoritme racista de la policia dels Estats Units que hem explicat més amunt o, aquí, el del bo social que va denunciar Civio (2019), també exposat abans. Va tenir força ressò així mateix el cas de l'algoritme d'Amazon que discriminava candidates a posicions dins l'empresa perquè l'algoritme descartava els perfils de gènere femení (Dastin, 2018). O el sistema de crèdits d'Apple que, complint els mateixos criteris econòmics, concedia préstecs als homes però no a les dones (Vigdor, 2019).

Una altra variant d'aquests nous problemes són els incidents relacionats amb la recopilació de dades personals sense que els interessats en siguin conscients, o de la seva utilització per a finalitats que mai no podrien imaginar les persones que les han cedit, per exemple en el conegut cas del filtratge massiu de dades d'usuaris de Facebook a Cambridge Analytica (Martínez Ahrens, 2018). Un dels darrers escàndols d'aquesta mena tenen a veure amb uns estudis interns de la companyia Meta (abans, Facebook) que demostraven que l'empresa havia comprovat per sí mateixa que Instagram, una de les seves plataformes, resultava perjudicial per a la salut mental de les noies adolescents (Seetharaman, 2021). Al tercer capítol procurarem contextualitzar aquests problemes amb la finalitat de disposar de més elements per valorar les seves repercussions.

Han estat nombrosos els codis ètics per a l'ús de la IA que han emergit arreu després del que es considera el primer, formulat curiosament en el context d'una obra de ficció per Isaac Asimov als anys cinquanta (Asimov, 1989). Avui, la multiplicació i sofisticació dels usos i aplicacions que ha experimentat la IA han anat acompanyades de la proliferació de declaracions i guies ètiques. En aquest sentit, a Catalunya comptem amb un important precedent, l'anomenada Declaració de Barcelona per al bon desenvolupament i ús de la IA a Europa (Steels i De Mantaras, 2018).

Aquesta declaració va precedir la posterior creació del Grup independent d'experts d'alt nivell que va desenvolupar les directrius ètiques per a una IA fiable de la Comissió Europea per tal de mitigar uns riscos que, segons el mateix document, "poden resultar difícils de preveure, identificar o mesurar (per exemple, sobre la democràcia, la justícia distributiva o sobre la mateixa ment humana)" (High-Level Expert Group on AI, 2019). En aquesta iniciativa, una de les

més destacades en matèria de regulació ètica de la IA, s'ha inspirat finalment la proposta de regulació de la UE que hem esmentat més amunt, i que es pot considerar el pas més rellevant des del punt de vista jurídic que s'ha dut a terme fins avui, juntament amb les recents Llei de Serveis Digitals (Digital Services Act o DSA) i la Llei de Mercats Digitals (Digital Markets Act o DMA), unes noves normatives europees que convé tenir molt presents perquè afecten de manera específica l'àmbit comunicatiu (les desenvoluparem més endavant).

A Espanya, l'any 2020 es va posar en marxa l'Estrategia Nacional de Intel·ligència Artificial (2020) i alhora es va engegar també l'estratègia catalana (2020). El mateix any es va crear l'Observatori d'Ètica en Intel·ligència Artificial de Catalunya, que coexisteix amb altres organismes com els comitès d'ètica de la IA que han impulsat, entre d'altres, les principals universitats nacionals.

Arribats al novembre del 2021, la UNESCO va aprovar el primer marc normatiu global sobre ètica de la IA, on es para especial atenció a la seva aplicació a la informació i la comunicació. Entre les qüestions que considera objecte del punt de vista de l'ètica no només destaca la desinformació, la privacitat o l'aplicació de la IA a la informació per part de les plataformes, sinó que es refereix també de manera explícita a la gestió algorítmica de la informació duta a terme pels mateixos mitjans de comunicació tradicionals (2021):

Communication and information, as AI technologies play an increasingly important role in the processing, structuring and provision of information; the issues of automated journalism and the algorithmic provision of news and moderation and curation of content on social media and search engines are just a few examples raising issues related to access to information, disinformation, misinformation, hate speech, the emergence of new forms of societal narratives, discrimination, freedom of expression, privacy and media and information literacy, among others.

Finalment, tal com dèiem més amunt, no ha estat fins el 2022 que s'ha aconseguit arribar a l'acord polític definitiu per engegar el tràmit administratiu d'aprovació de les noves lleis que regularan de manera molt determinant l'activitat de les plataformes digitals i que, un cop passin els processos d'aprovació, tindran importants efectes en la seva activitat i per tant en l'esfera pública digital. Aquests acords es materialitzen en les dues normatives esmentades: La Llei de Serveis Digitals i la Llei de Mercats Digitals, lleis que, en tenir forma de Reglament (i no Directiva), resultaran d'obligat compliment en a tots els Estats de la UE i per a totes les plataformes que hi operin, independentment de si estan establertes a altres països de fora de la unió.

Aquests Reglaments, que es preveu aprovar el 2024, persegueixen la proporcionalitat, el desenvolupament de la innovació, el creixement i la competitivitat . Per això, la Llei de Serveis Digitals categoritza els intermediaris del servei o plataformes segons el seu rol, dimensió i impacte en l'ecosistema digital i, d'acord amb això, defineix les seves obligacions.

Així, les plataformes el comportament de les quals tingui el major potencial d'impacte en la ciutadania seran les que tinguin més obligacions. D'aquesta manera, la UE vol promoure l'escalabilitat de les tecnològiques més petites. En concret, la UE obligarà als gegants tecnològics al retiment de comptes periòdic, a posar a disposició els algorismes a auditors externs i altres mesures que tindran impacte en el seu negoci publicitari, com l'exigència de més transparència en el funcionament del perfilat dels usuaris i dels anuncis personalitzats, o la limitació de grups a qui es poden segmentar, en concret no es podrà dirigir publicitat a nens, entre d'altres.

Aquestes lleis, un cop en marxa, impulsaran canvis que molt probablement afectin l'ecosistema informatiu tal com el coneixem avui i tot sembla augurar un marc més amable per actors menor envergadura com els mitjans de comunicació.

A l'hora d'establir una primera acotació per a l'ús ètic de la IA als mitjans de comunicació que desenvoluparem a la part de la nostra recerca que exposarem al quart capítol, tindrem en compte aquest marc que ens proporciona la tecnoètica.

1.7 MARC TEÒRIC

Aquest projecte de recerca rep les principals influències dels autors i autores de l'Escola de Bellaterra, a qui he tingut com a mestres durant els meus estudis de comunicació i han marcat la meua recerca amb una visió crítica i alhora humanista sobre les maneres d'entendre la comunicació. Enric Marín, Francesc Burguet, Albert Chillón, David Vidal, Gemma Casamajó, Carlos Zéller, Miquel de Moragas, Joan Manuel Treserras, Maria Corominas, Hèctor Borràs o Hèctor Barnils, entre d'altres referents de la comunicació que he tingut la sort de poder gaudir a la UAB.

A recer de la Facultat de Ciències de la Comunicació de la UAB s'han articulat, des dels anys 80 del segle XX, diverses línies de recerca coincidents en el to crític i cultural amb els mèdia i les seves mediacions, i s'ha fet des d'una pluralitat de camps disciplinaris. Des de la pròpia Teoria del Periodisme (amb autors com Héctor Borrat, Llorenç Gomis, Albert Chillón, Francesc Burguet, Xavier Giró Gemma Casamajó o David Vidal Castell) s'ha propugnat la crítica de l'objectivisme realista i s'ha recordat la natura lingüística i retòrica del coneixement humà, també en una praxis de coneixement lingüística i simbòlica com el periodisme, considerat el *negoci de la veritat capitalista*. Des de les Teories de la Comunicació, autors com Miquel de Moragas, Isabel Fernández, Miquel Rodrigo, Lola Montero o Jordi Berrio han incorporat les teories de la comunicació a la tradició acadèmica catalana i espanyola, no sense fer aportacions rellevants sobre el concepte de periodisme, opinió pública, comunicació o notícia. Finalment, des d'àmbits més específics però igualment rellevants, han escrit aportacions fonamentals des de la història de la comunicació o del periodisme Enric Marín, Joan Manuel Treserras, J.Ll. Gómez Mompert i Francesc Espinet, i des de l'Economia de la Comunicació la catedràtica Rosario de Mateo.

Des d'aquesta tradició, i amb les aportacions d'altres autors referents que detallarem en aquest mateix apartat, abordarem els conceptes que conformen el present treball de recerca, que té com a finalitat identificar els reptes del periodisme davant l'arribada de la intel·ligència artificial, i així estar en condicions de proposar línies d'actuació concretes per afrontar aquesta nova revolució tecnològica. El treball està compostat de quatre parts: un primer capítol que exposa el plantejament de la recerca, un segon capítol de context general que problematitza la utilització de les tecnologies exponencials quan es fan servir prioritzant les finalitats comercials als valors humans; un tercer capítol que fa èmfasi en aquesta mateixa crítica, però tancant el focus a l'aplicació de la IA en l'àmbit comunicatiu i cultural, i finalment, en el quart capítol, es

proposaran solucions per afrontar l'adopció de la intel·ligència artificial i altres tecnologies algorítmiques des de l'àmbit del periodisme considerant els valors de servei públic que regeixen la professió.

Aquest projecte de recerca es sustenta així mateix en les tesis d'autors que fonamenten la seva visió en la idea que les noves tecnologies són nous problemes als que cal aplicar els principis ètics que sempre han guiat la pràctica del periodisme de qualitat, i que es poden resumir en la voluntat d'oferir una informació veraç i diversa a la ciutadania per tal de poder governar-se i prendre decisions polítiques informades.

Els autors que han contribuït a la definició d'aquests pilars han estat referents de l'ètica periodística com el doctor Salvador Alsius, qui ha estat professor d'Ètica a la Universitat Pompeu Fabra i autor d'entre d'altres, *Ètica i periodisme* (1998), *La ética informativa vista por los ciudadanos* (2011) i *Els codis ètics del periodisme televisiu* (1999), el doctor Carlos Ruiz Caballero, professor titular a la Universitat Ramon Llull i autor, entre d'altres de *Ètica de la audiència* (2003), *La agonía del cuarto poder* (2008) i *La digitalización del otro* (2016), així com el doctor José Alberto García Avilés, Catedràtic de Periodisme en la Universitat Miguel Hernández d'Elx, referent en ètica i innovació periodística i autor de més d'un centenar de publicacions sobre comunicació.

A l'apartat introductori segueix un segon capítol que resulta el marc teòric i conceptual d'aquest treball. Sota la disciplina transversal dels estudis sobre Ciència, Societat i Tecnologia (CTS), té la finalitat de situar en el context social i cultural contemporani, en què l'automatització està cada cop més present i en més àmbits. Hem arribat a un punt en què, com desenvoluparem més endavant, amb la creixent delegació de tasques a la IA, aquesta està passant de ser una eina per realitzar operacions computables a adquirir certa agència a l'hora de realitzar altra mena de tasques no tan clarament matematitzables, com ara calcular el potencial de reincidència d'un pres, seleccionar candidats a posicions laborals i d'altres que hem mencionat en anteriors apartats.

És una tecnologia que, com d'altres, pot aportar grans avantatges (no només en termes de productivitat); el problema, com mirarem d'argumentar, és que en els sistemes capitalistes que busquen el màxim benefici com a principal criteri, les tecnologies exponencials solen posar-se al servei d'aquestes finalitats purament comercials. El segon capítol justificarà que el primer aspecte a tenir en compte a l'hora d'avaluar i d'aplicar la IA de manera responsable és posar l'ésser humà per davant dels objectius de negoci. Així mateix, exposarem quin significat

adoptarem nosaltres per referir-nos a les finalitats humanes a les que ha de sotmetre's la tecnologia.

Els CTS van néixer a la segona meitat dels anys '60 i han anat creixent i consolidant-se a l'acadèmia com a conseqüència dels cada cop més abundants avenços en els camps de la innovació científicotècnica que han comportat, en paral·lel, més preocupacions sobre els seus efectes socials. La interdisciplinarietat que caracteritza els CTS pretén produir coneixement tant al nivell de les humanitats i les ciències socials (filosofia, sociologia, història, economia, antropologia i ciències polítiques) en un nivell transdisciplinar a través de la recerca, que promulga la necessitat d'un escrutini social de les polítiques científicotècniques i que vol contribuir a sortejar la distància entre la cultura humanista la cultura científico-tecnològica que té el potencial de fracturar les societats.

Els CTS consideren els avanços tecnocientífics com un motor de desenvolupament integral, que, des d'una mirada que fuig del solucionisme tecnològic, es proposen promoure accions democràtiques en qüestions que afecten les societats. Aquestes actituds i pràctiques es relacionen amb la innovació tecnològica i tenen a veure amb com dissenyar, produir i implementar tecnologies per al desenvolupament inclusiu, sostenible i democràtic.

Aquest treball s'identifica, concretament, amb el camp d'estudis dels CTS anomenat Teoria Crítica de la Tecnologia, el qual proposa una revisió de les idees de la filosofia crítica del s. XX, com l'Escola de Francfort, especialment Jürgen Habermas i Herbert Marcuse o Martin Heidegger. El seu màxim representant, Andrew Feenberg (2002) postula que el disseny i l'ús de la tecnologia determina les societats i per tant els usos de les tecnologies resulten decisions ontològiques amb importants conseqüències polítiques. És per això que proposa que és possible pensar en una tecnologia no autònoma a l'ésser humà i influenciada per valors i el respecte pels principis democràtics i el medi ambient. Feenberg defensa que, per dur a terme innovació tecnològica cal un lideratge cultural, perquè és principalment la cultura el que dona forma a la tecnologia:

Although technologies are first and foremost tools for solving practical problems, they are not fully understandable in functional terms. This is especially true in cases where their function is itself in dispute. As we have seen with computers, these disputes are not merely technical but go to the cultural significance of the technology. The critical theory of technology is therefore a cultural theory. (Feenberg, 2002, p.107)

El filòsof realitza una revisió de la crítica a la racionalitat de Marcuse (de qui va ser deixeble) i observa la necessitat de qüestionar el fet que la raó científico-tècnica s'hagi convertit en el principal discurs legitimador. Aquest marc ens servirà per abordar les limitacions de la computació i els debats sobre l'abast del pensament computacional que s'han produït principalment des de la filosofia i els estudis culturals, amb autors com Lewis Mumford, Lluís Duch, Josep Maria Esquirol, Albert Chillón, David Vidal, David Casacuberta, Margaret Boden, Santiago Sánchez Migallón, Ana Carrasco Conde, Nick Bostrom, Ed Finn, José María Lassalle o Adela Cortina, entre d'altres que anirem citant.

Aquest context de creixent automatització ens servirà per contextualitzar el capítol 3, en el qual ens centrarem en els efectes de la mediació algorítmica en l'esfera pública, on la IA s'encarrega de tasques com ordenar i prioritzar la informació del món o intervenir en les relacions interpersonals. En aquest capítol farem una aproximació conceptual als canvis que provoca la mediació algorítmica que domina l'esfera pública digital i té la finalitat de justificar la necessitat d'un posicionament ètic respecte a l'ús de la IA per part del periodisme. Aquest apartat contextualitzarà el moment actual en què els algoritmes s'han erigit com a principals mediadors entre productors de continguts i audiències i per tant amb responsabilitats sobre els efectes de la seva acció, en la salut de l'ecosistema informatiu.

En aquest apartat també integrarem la teoria dels mitjans que s'han ocupat de la relació entre tecnologia i cultura des de l'humanisme i, a més dels autors esmentats, compta amb gran influència d'altres com Neil Postman qui, reivindicant Marshall McLuhan, defensa que la tecnologia no es neutra, dóna forma al pensament i per tant cal fugir del solucionisme per donar-li sentit. De la mateixa manera que el pensament de Lewis Mumford, l'obra de Postman té un pes molt significatiu en el marc teòric d'aquest treball. Tot i que la teoria que desenvolupa a la seva ecologia dels mitjans està orientada, en especial, a la crítica dels efectes de la cultura de la televisió com a mitjà de masses, la seva anàlisi resulta pertinent a l'hora de plantejar una crítica cultural en moments en què, com a conseqüència de la digitalització, ens trobem submergits entre quantitats encara més massives d'informació, a més d'incapaces com a societats de dotar de sentit a totes aquestes dades. Basant-nos en els plantejaments que desenvolupa a *Tecnópolis* (2018) i a *Divertirse hasta morir* (2012), apliquem als problemes actuals en l'àmbit de la comunicació el reclam humanista sobre la importància de definir finalitat en l'ús de la tecnologia. En moments de sobreabundància d'informació, reivindicuem la tesi de Postman, qui convida a preguntar-nos per a què la necessitem i quines finalitats ens ajuda a aconseguir.

Pel que fa als autors més recents, aquesta tesi s'ha alimentat de la reflexió dels darrers anys en el marc dels CTS com a conseqüència del sorgiment del nou ecosistema comunicatiu digital caracteritzat per la mediació de la intel·ligència artificial i els seus efectes socials. Seguint l'esmentada tradició interdisciplinària, integra recerca produïda principalment des de des Humanitats i les Ciències Socials com les Ciències de la Comunicació, el periodisme, l'ètica periodística, la Filosofia, l'ètica, la filosofia de la tecnologia o la Sociologia, els estudis culturals i la tecnociència amb autores com, entre d'altres ¹, Zeynep Tufekci, professora en Ciències de la Informació a les universitats de Harvard i Carolina del Nord. Soshana Zuboff, professora a en la Harvard Business School. Evgeny Morozov, professor en diferents universitats dels EUA i autor de *The Net Delusion: The Dark Side of Internet Freedom* (2011). James Bridle, autor de *La nueva edad oscura: La tecnología y el fin del futuro* (2018) i *Ways of being: beyond human intelligence* (2022) així com Gert Leonhard, autor de *Technology vs humanity* (2016).

El quart capítol de la tesi ha estat enfocat des de les disciplines de l'ètica periodística, la ciberètica i el periodisme computacional i es tradueix principalment de la recerca empírica que, en línia amb l'ànim participatiu d'aquest projecte, ha comptat amb la col·laboració dels principals mitjans catalans i ha servit per elaborar l'informe sobre ètica i intel·ligència artificial que ha publicat el Consell de la Informació de Catalunya (Ventura-Pocino, 2021). En aquest tercer bloc oferim un estat de la qüestió sobre l'adopció de les tecnologies exponencials als mitjans del país, les oportunitats i els reptes que aquests identifiquen en l'adopció de les noves eines i aplicacions algorítmiques.

Aquesta darrera part de la tesi consisteix en una proposta pràctica que té l'objectiu de proporcionar idees fonamentals sobre com el periodisme pot governar els algoritmes i dotar-los

¹ Frank Pasquale, professor de dret a la universitat de Maryland, s'especialitza en el Dret aplicat a la Intel·ligència Artificial. Actualment forma part del Comitè Assessor d'Intel·ligència Artificial Nacional dels Estats Units. És autor dels llibres *The Black Box Society: The Secret Algorithms That Control Money and Information* (2015) i *New Laws of Robotics: Defending Human Expertise in the Age of AI* (2020). Tarleton Gillespie, professor associat adjunt al Departament de Comunicació de la Universitat de Cornell. És autor del llibre *Wired Shut: Copyright and the Shape of Digital Culture* (2009). Konstantin Nicholas Dörr, professor associat a l'institut de recerca en comunicació de masses de la Universitat de Zúrich. Ted Striphas, professor associat a la Universitat de Colorado i autor de *Late age of print* (2009). Taina Bucher, professora associada a al Departament de Mitjans i Comunicació de la Universitat d'Oslo. És autora de *Facebook* (2021). Christian Sandvig Professor a la Universitat de Michigan. S'especialitza en infraestructures de la informació i xarxes socials. Cathy O'Neil, acadèmica vinculada a la Columbia University, activista del moviment Occupy Wall Street i autora de *Weapons of Math Destruction. How Big Data Increases Inequality and Threatens Democracy* (2016).

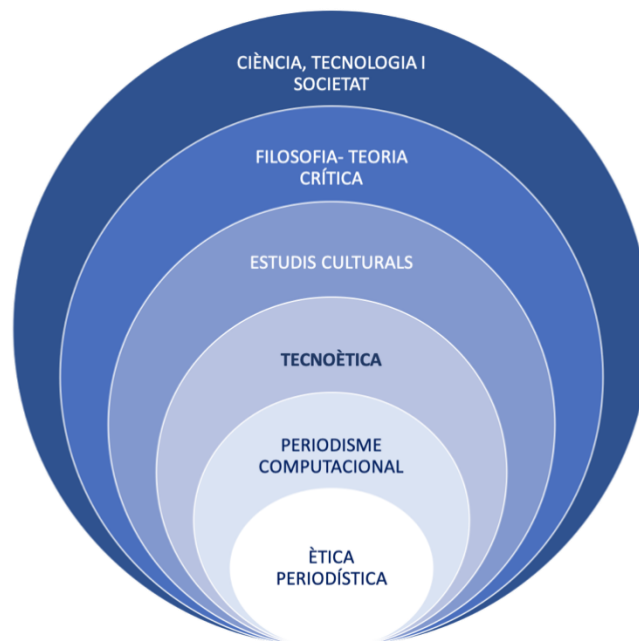
de valors ètics per contribuir, des de l'ús responsable de la tecnologia, a la salut de l'esfera pública.

La mateixa secció inclou també recerca feta des del periodisme computacional, la innovació en periodisme i les Ciències Formals com les Ciències de la Computació, la Intel·ligència Artificial, la ciberètica i les humanitats digitals. Finalment, també ha considerat les Ciències Jurídiques i en concret del dret informàtic, ja que hem vist pertinent analitzar tan la principal literatura normativa institucional sobre ètica aplicada a la IA, com la recerca en aquesta matèria portada a terme en l'àmbit de les ciències computacionals.

Així, per contextualitzar la proposta d'ètica de la IA des del periodisme, hem tingut en consideració normatives i guies ètiques portades a terme principalment en l'àmbit jurisdiccional que afecta Catalunya. En aquest sentit, la Unió Europea ha impulsat diverses iniciatives, fòrums de discussió i grups de treball que han donat lloc a nombrosos continguts que ens han servit com a guia principal a l'hora d'abordar el nostre treball. A més, a Catalunya comptem amb institucions, organitzacions de professionals i investigadores destacades en l'àmbit de la ciberètica que també hem procurat incorporar al nostre treball. Algunes hi figuren de manera explícita, d'altres hi són de manera implícita, per l'impuls que han aportat en la consecució de fortes comunitats que s'han ocupat històricament de la qüestió de l'ètica de la IA des de les ciències computacionals a Catalunya. Algunes veus que estan presents són les de Carme Torras Genís, Marta Ruiz Costa-Jussà, Ramón López de Mántaras, Carlos del Castillo, Ricardo Baeza-Yates, Ulises Cortés, Elisabet Golobardes, Karina Gibert o Karma Peiró, referent, a més, del periodisme tecnològic i la divulgació de ètica de la IA des de la comunicació. També hem comptat amb la influència de l'Observatori d'Ètica en Intel·ligència Artificial de Catalunya (OEIAC), entitat liderada per Director de la Càtedra OEIAC de la Universitat de Girona i professor Serra Húnter de Sociologia Albert Sabater, que realitza una important tasca de divulgació organitzant fòrums experts, publicant estudis i altres iniciatives que promouen la conscienciació sobre la necessitat d'ètica en els usos de la IA.

Així, en importar la perspectiva de la disciplina de la ciberètica hem pogut aprofitar la reflexió que acumula. Alhora, ens permet posar en valor l'abordatge multidisciplinari que requereix aquesta integració de tecnologia exponencial, en línia amb la tendència a la hibridació de perfils científics i socials que, com veurem més endavant, implica la integració de la IA en la pràctica del periodisme.

En periodisme computacional, hem explorat els autors Nicholas Diakopoulos, professor de la Northwestern University i autor de *Automating the news: How algorithms are rewriting the media* (2019). Charlie Beckett, professor del Departament de Mitjans i Comunicació de la London School of Economics. Seth C. Lewis , professor de la Universitat d'Oregon, Francesco Marconi, autor de *Newsmakers: Artificial Intelligence and the Future of Journalism* (2020) i Natali Helberger qui, tot i especialitzar-se en Dret a la Universitat d'Amsterdam, ha portat a terme important recerca relacionada amb la IA aplicada a la informació i el disseny ètic de de la tecnologia.



Disciplines sota les quals es desenvolupa la recerca - Font: elaboració pròpia

1.8 HIPÒTESI

Davant d'una esfera pública digital afectada com a conseqüència de la mediació algorítmica, el periodisme de qualitat segueix sent necessari. Una de les maneres de mantenir la centralitat serà demostrant que és capaç de governar els algoritmes, la mateixa tecnologia que fan servir els nous *gatekeeper* i que avui s'integra ràpidament a les redaccions. És possible dotar la IA dels grans valors ètics del periodisme.

1.9 METODOLOGIA I TÈCNiques

Per tal de desenvolupar els capítols de context de la tesi hem portat a terme una recerca de literatura relacionada amb la disciplina dels CTS, el camp d'estudis de la crítica de la tecnologia, els Estudis Culturals i la Filosofia. Al quart capítol, i amb la finalitat d'identificar els principals reptes per a l'ús ètic dels algoritmes, hem fet una anàlisi de literatura relacionada amb l'ètica digital: ètica de la IA des de la perspectiva de la disciplina de les ciències computacionals o tecnoètica, així com de recerca provinent del periodisme computacional i els camps de la sociologia, els estudis culturals de la comunicació, la filosofia i de l'ètica periodística.

Hem examinat nombrosos estudis, assajos, recerques, articles d'opinió, enquestes, conferències i altres continguts, primer, per disposar d'una imatge el més clara possible sobre els usos actuals de la IA en general i, molt especialment, en el camp de la comunicació, per després fixar la mirada en les aplicacions als mitjans de comunicació tradicionals. Abans de fer la reflexió conjunta amb experts acadèmics i el sector sobre com definir criteris ètics en l'ús de la IA als mitjans, hem analitzat literatura relacionada amb l'ètica de la IA des de la perspectiva de les ciències computacionals, atès que les tecnologies exponencials impliquen desafiaments ètics que convé identificar per, a continuació, valorar com integrar aquest bagatge en l'àmbit del periodisme.

La primera etapa de digitalització de les redaccions que es va iniciar als anys 1990, va comportar multitud de noves eines i aplicacions a la pràctica del periodisme que al seu torn generaren nous dilemes ètics en l'exercici de la professió. A l'hora d'adaptar el Codi Deontològic a Internet, al 2016, es van fer servir com a referència els grans principis que regeixen la professió.

Els periodistes poden adaptar-se a les novetats i, al mateix temps, conservar els valors de la professió (García-Avilés, 2021). Veritat, justícia, responsabilitat, llibertat... resulten igualment pertinents per valorar les noves pràctiques. No obstant, certs aspectes relacionats amb la natura i l'abast dels principis tradicionals experimenten certs matisos amb els nous reptes del periodisme digital (Deuze i Yeshua, 2001).

Així, a l'hora d'abordar aquesta nova etapa de digitalització caracteritzada per la incorporació de tecnologies exponencials, hem mantingut també com a referència els principis de l'ètica periodística (Alsius et al., 2010) alhora que hem considerat la influència que adquireix la

tecnoètica (ètica de la IA enfocada des de les ciències computacionals) tal com han proposat alguns autors (Dörr i Hollnbuchner, 2017).

Per concretar les eines i aplicacions d'IA més susceptibles de ser analitzades, hem fet una primera avaluació de les implicacions ètiques de les noves pràctiques a partir de la confrontació amb els valors que regeixen la professió periodística (Alsius et al., 2010), així com els principis de la tecnoètica (Sabater, Albert i De Manuel, Alicia, 2021).

En el marc del projecte, a més, s'han distribuït dues enquestes. Una primera a professionals de tot l'Estat per valorar la percepció davant l'arribada de la IA als mitjans i les principals preocupacions del sector en relació als seus usos. L'enquesta va ser dirigida principalment a periodistes treballant en l'àrea editorial de mitjans de comunicació en l'àmbit de l'estat espanyol.

Posteriorment, s'ha portat a terme una segona enquesta dirigida a les direccions de les principals capçaleres catalanes, considerant sota aquesta categoria els mitjans amb més audiència en l'àmbit català, així com les principals capçaleres de cadascuna de les províncies catalanes.

Per últim es van celebrar dues trobades amb mitjans i experts per tal de realitzar una valoració conjunta sobre els dilemes que presenten aquestes noves aplicacions i les consideracions ètiques a tenir en compte. Es va conformar un grup de discussió amb 5 representants d'alguns dels principals mitjans de comunicació catalans i una trobada amb el comitè d'experts nacional. Més endavant enumerarem tots els participants, així com els seus perfils.

Al quart capítol, en el context de la recerca empírica, exposarem el detall de la metodologia que inclou els membres del comitè expert, els mitjans participants, els perfils dels professionals que van respondre les enquestes, així com les preguntes que es van formular, relacionades amb els tipus d'usos de la IA al mitjà, oportunitats i reptes, entre d'altres.

CAPÍTOL 2

LA IDEOLOGIA GOVERNA LA TECNOLOGIA

Hay que distinguir los alimentos del alma de los venenos que, durante algún tiempo, puede parecer que sustituyen al alimento.

Simone Weil

La unión de una creciente productividad y una creciente destructividad; la inminente amenaza de aniquilación; la capitulación del pensamiento, la esperanza y el temor a las decisiones de los poderes existentes; la preservación de la miseria frente a una riqueza sin precedentes constituyen la más imparcial acusación: incluso si estos elementos no son la 'raison d'être' de esta sociedad sino sólo sus consecuencias; su pomposa racionalidad, que propaga la eficacia y el crecimiento, es en sí misma irracional.

Herbert Marcuse

La ferida s'ha posat al centre, però no és la ferida ontològica, la vida oberta que som, sinó la destrucció que provoca sobre nostre i sobre la resta del planeta l'acció humana. Frenar i reparar la destrucció sembla avui l'únic compromís possible.

Marina Garcés

Las energías pueden desperdiciarse, las instituciones pueden decaer, y las sociedades pueden acumular desgracias y cargas lo mismo que bienes y beneficios. Suponer que un punto ulterior de desarrollo llevará a un tipo más elevado de sociedad es simplemente confundir la cualidad neutra de la complejidad o la madurez con el mejoramiento. Suponer que un punto ulterior en el tiempo aporta necesariamente una mayor acumulación de valores, es olvidar los repetidos hechos de barbarie y la degradación.

(...) A diferencia de los modelos orgánicos del movimiento, a través del espacio y del tiempo, el ciclo del crecimiento y la decadencia, el movimiento del balanceo del bailarín, la exposición y el retorno de la composición musical, el progreso era un movimiento hacia el infinito, un movimiento sin perfección ni fin, un movimiento por el movimiento.

Lewis Mumford

2.1 INTRODUCCIÓ AL CAPÍTOL 2

En aquest segon apartat explorarem les principals qüestions teòriques que afecten l'assimilació de la intel·ligència artificial a les nostres societats, que són de tipus filosòfic. Abans de proposar intervencions en la línia assenyalada, analitzarem, des d'una perspectiva àmplia, les idees cabdals que ens habilitaran per anar concentrant la mirada i plantejar la reflexió en el periodisme que ens proposem incentivar.

D'entrada tractarem alguns conceptes fonamentals i transversals en aquest treball que podrien ser definits de manera diferent segons la posicionament ideològic de qui realitza l'anàlisi. Començarem per la noció de *progrés*, una paraula que adquirim al S.XV, provinent del llatí *progressus* (avenç) i derivada de *progredu* que, en grec, és caminar endavant (Bruguera i Fluvià i Figueras, 1996). Ens retornen, com a sinònims de progrés, els mots evolució, avenç, perfeccionament, desenvolupament, prosperitat, millora... i, com a antònims, endarreriment, degradació, barbàrie, incultura... Aquesta linealitat ens evoca l'existència de dues direccions, una de bona i una de dolenta. La *bona* és representada en la idea de l'etern progrés, fèrtil sobretot al Renaixement i la Il·lustració que Vidal i Castell (2005) descriu com el "catecisme laic", "la certesa de què demà es viuria més bé que avui; que la ciència i la raó farien les societats del futur millors" (p.31).

Malgrat aquest mite ha patit una crisi o un cert deteriorament des dels anys vuitanta del segle passat (Vidal i Castell, 2005), durant la present revolució tecnològica es segueixen observant símptomes deterministes —vinculats a l'emergència de les darreres innovacions en IA— que atorguen a les dades i al llenguatge computacional els poders legitimadors que s'atribueixen a la racionalitat i que es somatitzen, per exemple, en la patologia del *dataïsme* (Casacuberta, 2021), com veurem.

A l'hora de plantejar-nos solucions per fer una apropiació responsable de la tecnologia, primer ens haurem de preguntar per a què la volem governar, és a dir, per a què la volem. En línia amb el pensament dels autors de la teoria crítica l'haurem d'adoptar de manera racional, és a dir, valorant de manera justa tant el seu llegat com el seu potencial, sigui degradant per prevenir-lo o beneficiós per aprofitar-lo, en un equilibri que expressa així Lewis Mumford (1998):

Para reconquistar la máquina y someterla a los fines humanos primero hay que entenderla y asimilarla. Hasta ahora hemos adoptado la máquina sin entenderla por completo, o como los más pobres románticos, hemos rechazado la máquina

sin ver primero hasta qué punto podíamos asimilarla de forma inteligente.
(p.24)

Racionalitat serà un altre dels conceptes que aniran apareixent en aquest text i, també per endavant, aclarirem que l'utilitzarem com a significat de la capacitat d'exercir una avaluació equilibrada i amb criteri. Un judici que té en compte allò que és bo per a la persona i per a la comunitat i que, en aquest sentit, apel·la als fonaments ètics del present treball (exposats al plantejament). Aquestes bases es poden traduir en la *interdependència* que reclama la filòsofa Marina Garcés quan proposa "recompondre una visió del món que prioritzi els vincles d'interacció i dependència per davant dels valors d'autonomia i autosuficiència que havia exalçat l'individualisme" (Garcés, 2022, p.12). Així mateix, pot resultar un bon complement al concepte general la descripció d'aquesta *racionalitat* aplicada al comportament pràctic que fa el filòsof alemany Hans-Georg Gadamer (2010):

...la racionalidad del comportarse práctico no consiste, ni mucho menos, en lo que nosotros llamamos racionalidad. Se refiere más bien a la razón moral que sabe lo que es conveniente (...) aquello que uno tiene presente como lo moralmente correcto y poseedor de fuerza normativa. (p.131)

El debat ètic sobre l'ús de les tecnologies exponencials creix amb la seva popularització. Avui, les seves aplicacions penetren en cada cop més esferes, i és per això que aquests dies resulta habitual que les qüestions relacionades amb l'ètica de la intel·ligència artificial es vegin reflectides fins i tot als mitjans generalistes. Ja sigui als mitjans o als fòrums especialitzats, una de les reivindicacions que resulten més freqüents quan es parla de l'ètica de les màquines es tracta de la necessitat de posar la tecnologia al servei de les *finalitats humanes*. Aquesta demanda desperta sempre un gran consens, com no pot ser d'una altra manera. No obstant, sovint no s'acostumen a expressar, aquestes necessitats humanes, que segurament canviaran segons la perspectiva ideològica de qui les planteja. Aquest concepte, que en el treball que llegiu resulta transversal, també el definirem amb la influència del pensament d'autors com Lewis Mumford (1998, 2010), que hi han reflexionat en el context de les etapes industrial i postindustrial, i han inspirat gran part d'aquest relat. Així, per poder establir quines són les finalitats humanes que la tecnologia ens ha d'ajudar a assolir, la primera pregunta que ens farem és quines són les necessitats vitals de les persones.

La filòsofa Simone Weil (1996) es refereix a l'existència d'una obligació cap a l'ésser humà que neix de la consciència universal i es deriva de les seves necessitats vitals. Aquesta obligació, reconeguda en tots els textos antics, implica proporcionar a l'home els seus requeriments bàsics,

defensa Weil. Les necessitats essencials amb què cal abastir l'ésser humà estan constituïdes pels béns fonamentals que, com l'aliment, permeten la subsistència de la vida, però no només. També estan conformades per altres requeriments que tenen a veure amb la vida moral i, com l'aliment, resulten vitals per a l'home, escriu Weil: "Todo el mundo es consciente de que hay crueldades que atentan contra la vida del hombre sin atentar contra su cuerpo. Son las que le privan de cierto alimento necesario para la vida del alma" (p.26). Malgrat reconeix que aquestes necessitats espirituals poden ser més difícils de descriure que les materials, la seva existència és completament reconeguda per la humanitat, i alerta que resulta fonamental no confondre-les amb qüestions de caire superficial:

No se las debe confundir nunca con los deseos, los caprichos, las fantasías o los vicios. También es preciso discernir lo esencial de lo accidental. El hombre no tiene necesidad de arroz o de patatas, sino de alimento; ni de madera o de carbón, sino de calefacción. Igualmente, para las necesidades del alma se debe reconocer las satisfacciones diferentes, aunque equivalentes, que responden a las mismas necesidades. También hay que distinguir los alimentos del alma de los venenos que, durante algún tiempo, puede parecer que sustituyen al alimento. (p.28)

En el pensament d'aquesta filòsofa i també activista política, el progrés es mesura en relació a aquesta obligació, i és per aquesta raó que el respecte a les col·lectivitats està subjecte a la seva capacitat de proporcionar totes aquestes necessitats vitals a les persones que en formen part. Per a Weil, la comunitat que no és capaç de fer-ho és una col·lectivitat malalta: "Ciertas colectividades, en vez de servir de alimento, devoran las almas. Hay en tal caso enfermedad social" (p.27).

Si ens remetem als inicis de la història de la tècnica podem comprovar que l'home ha cercat en la tecnologia suplir precisament aquesta mena de necessitats: sabem que les primeres eines tenien l'objectiu de proporcionar alimentació (caçar), protegir de les condicions ambientals extremes i dels depredadors. Malgrat no han quedat gairebé vestigis, també tenim coneixement que l'ésser humà ha estat generador d'eines culturals com els ritus o les representacions gràfiques abans del llenguatge parlat i després escrit. Així, tant aquestes primeres eines com moltes de les tecnologies posteriors, estan relacionades o bé amb la pura subsistència o bé amb el desenvolupament cultural, per tant en la tecnologia l'home ha cercat, per una banda, millors condicions per adaptar-se al medi i sobreviure i, per l'altra, l'obtenció d'aliment de caire espiritual.

La filòsofa Ana Carrasco Conde (2020) recorda que la tècnica i tot avanç científic o tecnològic també té la funció de reforçar l'actitud humana de cuidar l'altre. Carrasco evoca que, en ser preguntada per una alumna quin considerava que va ser el primer senyal de civilització, l'antropòloga Margaret Mead va respondre que aquesta va ser un fèmur fracturat i curat. A la vida salvatge, va explicar Mead, un fèmur mai es cura perquè només pot fer-ho si algú es preocupa de cuidar el ferit". Carrasco afirma que una societat molt evolucionada tecnològicament pot ser poc civilitzada, que "una sociedad tecnificada no cuida; sino que descuida. Hace negocio" i per tant no hi ha avanç en les qüestions humanes.

En l'obra de Mumford, el valor dels artefactes també està lligat a proporcionar un millor accés a les fonts bàsiques de vitalitat com respirar aire net, disposar d'aliments nutritius, d'habitatge, de llum natural, crear, créixer intel·lectualment o gaudir de l'art i dels diferents tipus de relacions personals. Totes aquestes, necessitats que conflueixen en un objectiu principal que no és altre que l'home pugui desenvolupar la seva personalitat en les condicions més favorables. Ahora, s'oposen a la ideologia determinista que, en paraules de Mumford (2010) "pretende reducir la técnica a la conquista del tiempo, el espacio y la energía" (p.148).

Norbert Wiener, pare de la cibernètica, advertia als anys cinquanta dels riscos d'una mala gestió de la innovació tecnològica que podia ser perjudicial per al benestar humà i per tant implicar efectes contraris al veritable progrés. Per prevenir sobre els riscos d'una gestió deficient o interessada del desenvolupament tecnològic, apuntava la direcció que calia prendre per tal que es produís en favor del foment de les condicions per a la prosperitat de l'home, a través de la proporció de més temps d'esbarjo i l'enriquiment la seva vida espiritual:

Thus the new industrial revolution is a two-edged sword. It may be used for the benefit of humanity (...). It may also be used to destroy humanity, and if it is not used intelligently it can go very far in that direction.

(...) the social obligations of those responsible for management to see that the new modalities are used for the benefit of man, for increasing his leisure and enriching his spiritual life, rather than merely for profits and the worship of the machine as a new brazen calf. (Wiener, 1954, p.162)

De fet així ha estat habitualment al llarg de la història. Els estris per a l'agricultura o la ramaderia que van contribuir a millorar l'accés als aliments, els diferents torns, el vidre per fer recipients i les tècniques de laminació per fer finestres que deixaven passar la llum natural a les cases, les lents, les ulleres, el microscopi o el telescopi, la impremta... tots aquests invents han col·laborat

al benestar social. Tanmateix, l'optimització de la tecnologia per a l'acumulació que impulsa el capitalisme també ha fet passar la màquina sovint per un element negatiu (Mumford, 1998):

...el grado de regimentación se llevó más allá de lo necesario para la armonía o la eficiencia. A ciertos rasgos del capitalismo privado se debió que la máquina —que era un agente neutral— haya parecido con frecuencia, y de hecho haya sido a veces, un elemento maligno en la sociedad, despreocupado por la vida humana, indiferente a los intereses humanos. La máquina ha sufrido por los pecados del capitalismo; por el contrario, el capitalismo se ha aprovechado a menudo de las virtudes de la máquina. (p.43)

El pas de l'ús racional de la tecnologia a les societats en què s'ha posat la màquina al servei del capitalisme ha estat progressiu i multicausal. Les etapes primerenques de desenvolupament positiu han estat marcades per moments clau de la història de la tècnica que han protagonitzat els grans invents. Innovacions que, al seu torn, han impulsat de manera inqüestionable l'evolució cultural de l'ésser humà, començant pel llenguatge i l'escriptura, passant per la ceràmica, el molí d'aigua, el paper, el vidre, la brúixola, la filadora, la impremta, el rellotge mecànic, el velocípede, la cartografia, la turbina de vent, gran part dels invents i les teories de Leonardo, el rellotge portàtil, el torn de filar, el calendari gregorià...

Tanmateix va arribar un moment decisiu —que Lewis Mumford el situa al S.XVI— en què es produí un canvi cultural associat al desenvolupament tecnològic que va marcar la societat fins avui. Un punt d'inflexió que coincidí amb la proliferació de la mineria lliure de regulació. Sobre aquesta lògica determinant de la mineria Mumford va escriure: “el minero trabaja no por amor o para alimentarse, sino para hacer su montón” (p.95). Segons Mumford, aquest període de la història va fixar el model d'explotació capitalista basat en l'acumulació que es desenvolupà en etapes posteriors. L'autor també associa a aquest marc productiu una concepció esbiaixada de *valor* que té en compte, d'una banda, l'esforç que es dedica al treball i, de l'altra, l'escassetat del producte. Però el valor, diu Mumford, no té res a veure amb l'escassetat ni amb la força emprada:

El valor real reside en el poder para sostener o enriquecer la vida (...) un abalorio de cristal puede ser más valioso que un diamante, una mesa de negociación más valiosa estéticamente que la más enrevesadamente esculpida, y el jugo de un limón puede tener más valor en un largo viaje oceánico que un centenar de libras de carne sin aquél. (...) La noción de valor del minero, como la del financiero, tiende a ser puramente abstracta y cuantitativa. (p.94)

El canvi cultural definitiu cap a la vida cobdiciosa, explica, es produeix amb el pas a la modernitat que s'inicia també aleshores. Es tracta de la combinació d'aquest model productiu amb la permeància a la societat dels costums cada cop més tendents al luxe de la cort (una estètica que aquesta adopta del costum de l'església d'honorar les figures santes amb tota mena de tresors artístics, des de pintures fins a objectes de joieria). Mumford descriu aquesta nova aspiració, tal com es va generar des de la cort i que seguidament es va escampar, primer a la burgesia i després a la resta:

(...) dirigieron sus energías hacia un horizonte, continuamente en movimiento, de consumo (...) el ritual del gasto llamativo se extendió rápidamente por el resto de la sociedad: justificó las abstracciones de los acumuladores de dinero y amplió los usos del progreso técnico de los inventores. (p.122)

Si bé durant l'Edat Mitjana el luxe privat no existia —i a més qualsevol signe d'ostentació era mal vist i un impediment per a la salvació— amb el trànsit del teocentrisme a l'antropocentrisme es produeix el canvi pel qual l'aspiració a la vida santa és substituïda pel desig de gaudi i possessió de bens materials. Escriu Mumford: “para la sociedad que no conoce otros ideales, el hacer gastos se convierte en la principal fuente de placer” (p.123).

Des de llavors, segons l'autor, la màquina al servei de la lògica comercial ha condicionat els usos de la tecnologia. Avui es destinen enormes recursos a innovacions tecnològiques que sovint tenen funcions que —confrontades a les necessitats bàsiques per a la prosperitat humana que hem exposat ara—, resulten superficials (com per exemple certes funcionalitats dels telèfons intel·ligents, o altres aplicacions orientades principalment a entreteniment o a l'ostentació). El filòsof Gilles Lipovetski considera que això es deu al fet que el consumidor vol plaers narcisistes, i malgrat reconeix la importància de la innovació per a finalitats menys frívols, considera que una innovació purament funcional seria trista: “Si eliminamos lo lúdico y lo superficial viviríamos como monjes en un monasterio tibetano. La gente necesita emociones y fiestas desde un punto de vista antropológico. Y el selfi es un pequeño espectáculo superficial y narcisista que no debemos dramatizar” (Lipovetski citat a García Aller, 2020).

Per a altres autors contemporanis, com Josep Maria Esquirol (2018) el consumisme i la frivolitat que el sol acompanyar ha conduït a la societat occidental a la degradació:

El mercantilismo no es una comunidad que tiene mercado, sino un sistema que lo ha convertido todo en mercancías de consumo, y en especulaciones sobre la nada, con lo cual ha hipotecado el mundo y la vida. Repitámoslo: no hay que

sentir nostalgia por un pasado dorado, por lo demás inexistente, sino por un hoy degenerado ya desde el amanecer. (p.158)

Nosaltres plantejarem que el fet que una comunitat disposi de tecnologia més sofisticada no necessàriament implica que sigui una societat més avançada, perquè si fos així, els estalvis de la creixent automatització dels processos productius que s'han donat durant les fases que han seguit la industrialització, haguessin tingut més repercussió en el repartiment de beneficis. Aquests es podrien materialitzar en més temps disponible per a la majoria de les persones, per exemple, o en beneficis per al medi ambient, en comptes de contribuir a l'acumulació de més riquesa en un mínim de privilegiats, a augmentar la contaminació i extractivisme incontinent dels recursos naturals del planeta (Feenberg, 2002; Marcuse, 1993; Marx, 2014; Peirano, 2022; Piketty, 2019; Scavino, 2022). Perquè encara ens trobem lluny de les societats que proporcionen el 'dret a ser' de Bauman al qual ens referíem al plantejament.

L'economista francès Thomas Piketty (2019) analitza les creixents diferències entre una petita part més rica de la població en vers la gran majoria i explica que "La revolució conservadora dels anys 1980, l'esfondrament del comunisme soviètic i l'aparició d'una nova ideologia de tipus propietarista han portat el món a uns nivells impressionants i descontrolats de concentració de les rendes i els patrimonis en aquest principi del S.XXI" (p.1129). Segons detalla, el 50% més pobre de la població posseeix només el 5% del patrimoni mundial, mentre que el 10% més ric disposa d'entre el 50% i el 60%. L'economista alerta que les desigualtats en la renda que van disminuir al S.XX han augmentat de manera significativa des dels anys 1980-1990, fet que "marca un canvi de tendència perfectament clar i ben documentat" (p.766). Piketty creu que aquest desequilibri en la possessió del patrimoni —juntament amb el canvi climàtic— suposa un dels principals reptes contemporanis de la humanitat i considera que "contribueix a un qüestionament profund de la mateixa noció de progrés" (p.769).

La solució a aquests problemes del que anomena *l'hipercapitalisme globalitzat i digital* imperant —representat per multimilionaris com Elon Musk o Jeff Bezos— és colonitzar l'espai per seguir explotant allà els recursos que s'hi descobreixin, i té sentit pensar que únicament en favor d'ells mateixos —com han fet fins ara a través dels seus negocis basats en la lògica individualista—. O escapar i aïllar-se de la resta de les persones mortals. Douglas Rushkoff, nomenat per l'Institut de Tecnologia de Massachusetts (MIT) com un dels deu intel·lectuals més influents del món, és un d'aquests experts que sol ser convidat per elits a la recerca de recomanacions visionàries. Aquest teòric de la cultura digital relata que, en una d'aquestes, va ser citat al mig d'un desert, a on hi van aterrar, en els seus jets privats, cinc d'aquests multimilionaris. Després de fer-li

algunes preguntes molt bàsiques, van començar a interrogar-lo sobre el que anomenaven “*the event*”, que traduït seria *l’esdeveniment*. Una catàstrofe que preveïen que es produiria al planeta Terra. Una situació de col·lapse inconcreta però que estaria motivada pel canvi climàtic, per una guerra bioquímica o per alguna de les amenaces que planegen en les previsions distòpiques. Rushkoff evoca que, en un cert punt de la conversa, se’n va adonar que la veritable preocupació d’aquests milionaris i la raó per la qual l’havien citat allà era per tenir respostes a qüestions com de quina manera podien mantenir l’autoritat sobre els cossos de seguretat que estan contractant per protegir els búnquers de luxe que s’estaven comprant per sobreviure al col·lapse de la societat quan es produís *l’esdeveniment*. La reflexió de Rushkoff després de l’experiència és que la riquesa que han acumulat els està servint per alimentar l’obsessió sobre la necessitat d’aïllar-se del perill del canvi climàtic, les migracions massives, les pandèmies, o el que és el mateix, per escapar. Diu Rushkoff: “For them, the future of technology is about only one thing: escape from the rest of us”. Segons el raonament d’aquest teòric, els milionaris del món tecnològic que en la lògica hipercapitalista són *els guanyadors*, en realitat són els perdedors, perquè són víctimes de la seva pròpia mentalitat, en la qual *triomfar* significa ser prou ric com per aïllar-se del món malalt que han contribuït a crear amb el seu extractivisme de recursos a compte de l’acumulació:

These people once showered the world with madly optimistic business plans for how technology might benefit human society. Now they’ve reduced technological progress to a video game that one of them wins by finding the escape hatch. Will it be Jeff Bezos migrating to space, Thiel to his New Zealand compound, or Mark Zuckerberg to his virtual metaverse? And these catastrophising billionaires are the presumptive winners of the digital economy – the supposed champions of the survival-of-the-fittest business landscape that’s fuelling most of this speculation to begin with.

Yet this Silicon Valley escapism – let’s call it The Mindset – encourages its adherents to believe that the winners can somehow leave the rest of us behind. (Rushkoff, 2022)

Per a aquests emprenedors i els seus seguidors ultra milionaris el més urgent que cal assumir des de la innovació en tecnologia és trobar la manera d’escapar, o bé de seguir consumint recursos per mantenir els seus privilegis, perquè el canvi climàtic està complicant continuar amb les pràctiques que es basen en la despesa de recursos desmesurada. Elon Musk es proposa conquerir el planeta Mart i haver-hi enviat un milió de persones l’any 2050 (Elon Musk [@elonmusk], 2020). Bezos pretén desenvolupar unes bombolles espacials quilomètriques (Chang, 2021; Tuku, 2020) que, segons la periodista Marta Peirano (2022), són “una variante

extrema de los espacios que ya ocupan, espacios artificiales donde reproducir las condiciones de la naturaleza terrestre, a costa de consumirlas en otra parte” (p.43). Totes dues, diu, solucions destructives i per tant incompatibles amb el veritable desenvolupament científic .

En aquest punt ja hem tractat els conceptes *progrés, racionalitat i finalitats humanes* que ens proposàvem aclarir des de la perspectiva d’aquest treball, amb la finalitat de conduir la persona que llegeix aquest text a la idea fonamental que està en l’arrel de qualsevol iniciativa que aspiri al govern tecnològic responsable. En síntesi, es tracta de què no és la tecnologia en sí allò que causa degradació social, sinó els valors antidemocràtics que sovint governen la innovació tecnològica, tal com apunta, des de la teoria crítica, el filòsof Arthur Feenberg (2002):

I argue that the degradation of labor, education, and the environment is rooted not in technology per se but in the antidemocratic values that govern technological development. Reforms that ignore this fact will fail, including such popular notions as a simplified lifestyle or spiritual renewal. Desirable as these goals may be, no fundamental progress can occur in a society that sacrifices millions of individuals to production and disempowers its members in every aspect of social life, from leisure to education to medical care to urban planning. (p.3)

Equiparar progrés a evolució tecnològica —és a dir, el determinisme tecnològic— pot ser un parany i anul·lar la capacitat crítica per valorar de quina manera adoptem la tecnologia i al servei de quins interessos la posem. S’idealitza el progrés i s’associa al desenvolupament tecnològic (Duch i Chillón, 2012; Lippmann, 2003; Mumford, 1998; Postman, 2018), però cal tenir present també que el disseny dels artefactes i l’ús que els donem influeix des dels processos productius i l’economia fins a la cultura i la vida quotidiana, i per tant té efectes directes en les persones (Feenberg, 2002; Finn, 2018; Mumford, 1998; Scavino, 2022). Com han teoritzat abastament els corrents marxistes, l’ús de la tecnologia és un factor que pot arribar a agreujar les desigualtats socials, i com veurem més endavant, en el cas de les tecnologies exponencials en les quals es centra aquest treball, ja s’estan fent servir en aplicacions destinades a perpetuar relacions de domini i explotació (Hao, 2022).

En aquest context, doncs, i tenint en compte que la nostra pretensió és proporcionar un context teòric per poder fer una assimilació tecnològica respectuosa amb l’ésser humà, haurem d’assumir que les decisions sobre com la fem servir són de caire ontològic, impliquen importants conseqüències polítiques i, a més, exclouren la majoria de la societat d’aquest debat resulta profundament antidemocràtic (Feenberg, 2002). Per adaptar-nos al desenvolupament tecnològic i aprofitar les possibilitats de democratització que aquest també ofereix, haurem de

defensar la necessitat d'un lideratge cultural de la innovació tècnica que consideri que el progrés només serà sostenible si som capaços de resoldre els problemes futurs de la tecnologia present.

Com a prolegomen a la seva proposta de crítica de la racionalitat científico-tècnica, Feenberg (2002) recupera la següent idea de Marx, en la qual observa que la nova tecnologia pot servir, sigui per mantenir l'estatus quo, com per erosionar la jerarquia social existent. En conseqüència, cada embranzida tecnològica es pot considerar una oportunitat perduda de progrés democràtic:

(...) the social hierarchy can generally be preserved and reproduced as new technology is introduced. Computerization of record keeping is a case in point, intensifying surveillance and control. A second "principle of democratic rationalization" holds that new technology can often be used to undermine the existing social hierarchy. Most major innovations open possibilities of democratization that may or may not be realized depending on the margin of maneuver of the dominated. Thus, in many workplaces the drive to computerize has excited and sometimes fulfilled participatory expectations. (...) We can learn from Marx that every stage in the development of industrial technology was haunted by missed opportunities for democratic progress" (p.92).

Marcuse, de qui Andrew Feenberg va ser deixeble, advertí ja en la societat industrial en què ell va viure que la dissidència d'un sistema que considera l'home i la natura com a elements a ésser utilitzats i dominats havia de passar per acceptar que no és possible la neutralitat de la tecnologia, perquè aquesta exerceix un paper fonamental en la perpetuació d'aquests perjudicis (Marcuse, 1993):

(...) la sociedad tecnológica es un sistema de dominación que opera ya en el concepto y la construcción de las técnicas.

La manera en que una sociedad organiza la vida de sus miembros implica una elección inicial entre las alternativas históricas que están determinadas por el nivel heredado de la cultura material e intelectual. La elección es el resultado del juego de los intereses dominantes. Anticipa modos específicos de transformar y utilizar al hombre y a la naturaleza y rechaza otras formas. (...) una vez que el proyecto se ha hecho operante en las instituciones y relaciones básicas, tiende a hacerse exclusivo y a determinar el desarrollo de la sociedad como totalidad. (p.26)

Per tant, reivindicuem la crítica de la tecnologia que aplica una mirada de perspectiva àmplia, oberta necessàriament a la revisió del mite de progrés, que posa en qüestió el determinisme tecnològic, o el que és el mateix, una tecnologia no autònoma a l'ésser humà, i per tant influenciada pel respecte als principis democràtics i el medi ambient (Feenberg, 2002; Marcuse,

1993; Mumford, 1998). Una visió que defensa la necessitat d'establir criteris d'actuació que passen per condicionar la innovació tecnològica a l'assoliment de finalitats socials rellevants i, en conseqüència, orientada a una apropiació social d'aquesta (Winner, 1980).

Així doncs, cal romandre alerta sobre el potencial de degradació d'allò humà que pot implicar un progrés mal gestionat. Dotar la tecnologia de sentit passa per la revisió de conceptes com civilització i evolució de manera que el seu significat es relacioni d'una manera indestriable amb la interdependència o l'instint de cooperació mencionat al principi. Una societat orientada a la millora del benestar de les persones on la tecnologia no és un obstacle per perseguir els valors humanistes. Una ciutadania socialista moderna que considera senyals de progrés aspectes que van més enllà de l'eficiència i es centren en el benestar dels seus membres i que considera indicadors avui devaluats com l'educació, la qualitat ambiental o la satisfacció laboral (Feenberg, 2002):

A socialist society dedicated not to simple economic growth but to the actualization of human capacities could employ more direct and varied measures of material well-being than these simple quantitative ones, As I will show later in this chapter, it could evolve an economic culture that encompassed goals systematically undervalued in the existing modern societies, such as education, environmental quality, and satisfaction at work (p.142).

2.2 A LA RECERCA DELS LÍMITS ACTUALS DE LA INTEL·LIGÈNCIA ARTIFICIAL

Cuanto más se aproxima el lenguaje a las densas concreciones de todo ser racional, menos abstracto y preciso puede ser. La última palabra sobre la experiencia humana es la misma experiencia humana, sin mediar la palabra.

Lewis Mumford

AI is neither artificial nor intelligent. It is made from natural resources and it is people who are performing the tasks to make the systems appear autonomous.

Kate Crawford

El supuesto reemplazo de lo mítico por lo lógico es apenas una falacia que —como sucede con las tentativas inversas— resulta imposible consumir, y cuyo ensayo suele acarrear deshumanizadores efectos.

Lluís Duch i Albert Chillón

El científisme havia imposat unes exigències de fonamentació discursiva que no podien satisfer ni la religió ni l'art. D'aquí prové part de la crisi de la cultura occidental.

David Vidal Castell

2.2.1 Restriccions bàsiques del processament de dades

Establir els límits de la intel·ligència artificial és una qüestió que fa dècades que es tracta en profunditat des de les disciplines com la pròpia IA, la ciberètica, així com des de la filosofia. En aquest punt, convé avançar que definir les capacitats de la computació de manera detallada i profunda quedarà fora de l'abast d'aquesta tesi. Tanmateix sí que ens proposem aclarir alguns aspectes fonamentals per tal de poder assolir els principals objectius d'aquest projecte, que es resumeixen en facilitar recursos al periodisme per fer una apropiació responsable de les tecnologies exponencials. Per aconseguir-ho, necessitarem partir d'una idea el més clara possible sobre la solvència de la computació, i en especial en el context de la comunicació. D'aquesta manera, estarem en millors condicions de saber fins a quin punt hi podem confiar —o per a quines tasques— en l'àmbit de la comunicació computacional professionalitzada. Això també ens ajudarà a entendre millor com integrar la recerca en ciberètica en l'ètica periodística, per exemple, a respondre millor preguntes com quines són les eines que més riscos ètics impliquen o, fins i tot, quin ha de ser el paper del periodisme davant el creixent ús de la IA a la societat.

La gran qüestió que en els àmbits de la filosofia de la IA ha conduït a les reflexions pertinents sobre els límits de la computació és la pregunta sobre si les màquines poden ser intel·ligents. Aquest, de fet, és un tema que ha dominat la filosofia de la IA durant dècades. Per respondre la pregunta, un dels pares de la computació, el matemàtic Alan Turing, va dissenyar un test que consistia en posar a prova una màquina per saber si aquesta pensava com una persona humana (Turing, 1950). El que es coneix com *el test de Turing* consisteix en què una persona interroga, a través d'un teclat, una màquina i una altra persona. Si la interrogadora no identifica la màquina durant més del 30% del temps (més de cinc minuts), aquesta ha passat el test i, per tant, és intel·ligent. Tot i que es diu que aquest exercici va ser una ironia de Turing (Boden, 2017), el cert és que la idea s'ha mantingut present des d'aleshores. Per cert, segons explica Margaret Boden, una de les referents mundials en l'estudi de la IA, una de les raons per les quals el concepte resulta vigent és que per a certes investigadores en aquest camp que han tingut desig de notorietat ha suposat un recurs comú afirmar —o deixar que periodistes afirmen— que el seu sistema ha passat el 'test de Turing' (2017).

En qualsevol cas, val la pena considerar aquesta prova d'intel·ligència per la reflexió posterior que va generar. Tampoc pensem que toqui aquí exposar-la tota amb detall, perquè l'estat en

què es troba actualment no difereix massa del punt en què va quedar després de la que es considera la principal refutació d'aquesta idea, que va tenir lloc l'any 1980, quan el filòsof John Searle va plantejar l'experiment de l'habitació xinesa (Searle, 1980). En aquest, proposa un escenari en el qual ell mateix està tancat en una habitació amb una escletxa per la qual entren missatges escrits en forma símbols que Searle no comprèn. Sense ell saber-ho, es tracta de preguntes en xinès, a les quals ell haurà de respondre a través d'una segona escletxa. Per fer-ho, compta amb l'ajuda d'un manual que li indica quins símbols ha de fer servir per contestar. Aquest llibre de regles és un programa xinès de processament de llenguatge natural. A l'altre costat, rep les seves respostes un xinès que està segur d'estar conversant amb algú que sap la llengua xinesa. La pregunta que fa Searle és si, en aquest cas, podem afirmar que ell (que és qui està dins) sap xinès. Searle conclou que el càlcul formal sol (el que estava fent ell a l'habitació) no pot generar intencionalitat, o el que és el mateix: el fet que un programa faci servir símbols de manera adient no implica que sigui capaç de comprendre de manera genuïna (com ho fa una persona humana). La principal objecció a la prova de Turing, doncs, és que està limitada perquè es refereix només al comportament del que és observable (Boden, 2017).

Com dèiem, des d'aleshores han sorgit més debats i refutacions, no obstant, tenint en compte que treballem en una introducció al tema, farem un salt al present per fer un estat de la qüestió adaptada a l'àmbit que ens proposem analitzar. A les properes pàgines exposarem que la qüestió no ha avançat d'una manera prou significativa com per canviar l'escenari que va deixar Searle.

La resposta actual sobre si les màquines poden pensar és que, tot i no existir arguments del tot concloents en favor del sí o del no, el que obté notablement més consens és el que estima que la recerca en l'àmbit de la intel·ligència artificial està molt lluny de poder reproduir la intel·ligència humana (Boden, 2017; Bostrom, 2016; Casacuberta, 2003; Migallón, 2022). Tenint això clar avançarem amb la voluntat de seguir identificant quines són les restriccions bàsiques d'aquesta tecnologia que s'anomena 'intel·ligent', però que no tothom coincideix en el fet que ho sigui, com la investigadora Kate Crawford, qui afirma que ni és artificial ni intel·ligent: "AI is neither artificial nor intelligent. It is made from natural resources and it is people who are performing the tasks to make the systems appear autonomous" (Crawford, 2021).

Segurament, una de les principals limitacions dels algorismes és que no poden conceptualitzar ni comprendre —en un sentit humà— els problemes que han de resoldre, i per tant existeixen impediments importants que ens han de fer valorar molt bé com els fem servir. En l'àmbit de la filosofia, a aquesta comprensió humana se l'anomena 'intencionalitat'. El mateix Turing, qui va

establir els fonaments teòrics de la computació, considerava que la pregunta fonamental no era si les màquines podien pensar (perquè el significat de pensar pot resultar ambigu) sinó que allò que era important saber era si les màquines podien tenir consciència (Turing, 1950). Amb el desenvolupament de la IA que vivim en aquests moments, el debat sobre la consciència de les màquines està molt present. De fet, mentre treballem en la tesi, aquesta qüestió està ocupant espais a mitjans generalistes a compte del cas d'un enginyer de Google consternat perquè, assegura, el sistema que estan desenvolupant ha adquirit consciència i tem el seu propi final. Més endavant ho tractarem. Primer farem un repàs als aspectes bàsics que la ciberètica ha identificat que poden influir en què determinats resultats d'aquestes intel·ligències (processaments de dades) no siguin acurats.

Segons el Diccionari de l'Institut d'Estudis Catalans (2021) un algoritme és un conjunt de regles per a resoldre un problema en un nombre finit de passos. Un dels aspectes fonamentals del funcionament dels algoritmes que cal tenir en compte a l'hora d'identificar possibles limitacions en resoldre problemes, és que es basen en dades del passat de les quals infereix quina solució és la més probable. El filòsof David Casacuberta (2018) ho explica amb aquests exemples:

Un algoritme no indica les raons per les quals proposa el resultat. Senzillament es basa en regularitats anteriors: a un tant per cent elevat de persones que han vist moltes de les sèries que jo també he vist els ha agradat molt aquesta nova sèrie, així que probablement a mi també m'agradarà; més del 70% de les persones d'una edat, estat civil, salari mitjà i barri similars al de la persona que ens demana un crèdit no l'han arribat a retornar mai, per tant és millor que no l'hi concedim, etc.

Tenint en compte que l'algoritme basa el seu resultat en les dades, un dels riscos més comuns sobre l'ús de la IA que identifiquen regulacions i guies ètiques és el potencial dels sistemes per generar solucions esbiaixades, normalment com a conseqüència de les dades emprades en el procediment computacional, dades que resulta molt difícil (i fins i tot impossible) que no continguin biaixos, principalment perquè aquestes es generen a partir de l'activitat humana, i per tant reflecteixen els biaixos de les persones (Baeza-Yates i Peiró, 2021) que poden ser estadístics, culturals, cognitius... i n'hi ha de molts tipus (Kahneman, 2020).

Per mostrar com es pot produir un biaix en un algoritme posarem com a exemple el cas real d'un sistema que servia per donar suport a la selecció de personal, i en aquesta ocasió s'utilitzava per identificar possibles candidats a posicions d'enginyeria informàtica. Aquest algoritme, a l'hora de filtrar candidatures, descartava els currículums que provenien de dones, perquè les dades

amb què havia estat entrenat mostraven que fins llavors aquella mena de posició havia estat exercida per persones de sexe masculí. La màquina inferia que el gènere femení era una categoria que convertia el currículum en rebutjable. Així, aquell sistema tenia un biaix de gènere.

Una altra limitació bàsica de la IA és el fet que no disposa de sentit comú. En el transcurs de la història de la humanitat hem après coses que per a nosaltres poden resultar molt òbvies, com que la pluja no puja del terra al cel, que no podem travessar una paret, caminar sobre l'aigua o menjar pedres. Els experts no saben com ensenyar les màquines qüestions d'aquesta mena. Ramón López de Mántaras, professor d'investigació i Director de l'Institut d'Investigació en Intel·ligència Artificial del Consell Superior d'Investigacions Científiques (CSIC) ho explica així: "Sabemos cómo funciona el mundo. Eso es lo que conforma el sentido común. Esto hace que sea muy complicado dotar a una máquina de este tipo de conocimientos. Las cosas más sencillas son mucho más complicadas de enseñar de lo que parece" (de Mántaras citat a García Aller, 2020, p.297). De Mántaras diu que als inicis de la recerca en IA, fa seixanta anys, es va infravalorar aquest problema i es pensava que el més difícil seria ensenyar-li a executar tasques complexes, però el que ha passat ha estat ben bé el contrari. Alison Gopnik, catedràtica de psicologia de la Universitat de Berkeley i experta en aprenentatge infantil, veu molt improbable que les màquines aconseguixin adquirir la capacitat de comprensió el món que té una criatura. Als 5 anys és quan es produeix la major quantitat d'aprenentatge sobre l'entorn, i la ciència encara no sap com es produeix (García Aller, 2020).

La IA tampoc té criteri moral. Tot i que és possible tenir en compte valors en el disseny dels sistemes (Feenberg, 2002), la realitat és que només les persones poden retre comptes i assumir responsabilitats pels mals funcionaments (Boden, 2017; Mántaras, 2021). El cas que hem explicat ara sobre l'algoritme que discriminava dones també exemplifica el problema que pot suposar delegar en excés en sistemes que no tenen aquest criteri moral. La supervisió humana resulta fonamental, doncs, en els casos en què les determinacions dels sistemes poden afectar la vida de les persones.

Casacuberta (2018) es refereix a aquestes limitacions morals de la IA i com procedir en conseqüència amb un cas figurat d'haver de valorar si els algorismes d'una xarxa social exposen de manera suficient a la diversitat:

...si es tracta d'una qüestió (...) com, per exemple, decidir si una xarxa social està dissenyada de manera que garanteixi el respecte a la diversitat o no, al final de la cadena de decisions cal que hi hagi un equip de persones que, malgrat la possibilitat d'errors, emocions i biaixos ideològics, entenguin que moltes

decisions de l'esfera ètica, política i social només poden prendre's des d'una comprensió holística del que comporta ser humà i de quins són els drets i les llibertats bàsiques de la humanitat. I això és una cosa que no s'obté senzillament a partir de regularitats estadístiques.

Un altre dels grans riscos dels usos de màquines en la presa de decisions és que els processos de les tecnologies que evolucionen la IA com l'aprenentatge automàtic o l'aprenentatge profund avui són impossibles d'explicar (Pearl, 2022; Torras, 2020). Per tant, si una persona es veu perjudicada per una decisió algorítmica ningú serà capaç d'explicar-li el perquè. En conseqüència, existeix un perill de desresponsabilització en la presa de decisions. Norbert Wiener també va ser un dels primers que va considerar aquests riscos. Així ho expressava als anys cinquanta del segle passat:

The modern man, and especially the modern American, however much "know-how" he may have, has very little "know-what." He will accept the superior dexterity of the machine-made decisions without too much inquiry as to the motives and principles behind these... Any machine constructed for the purpose of making decisions, if it does not possess the power of learning, will be completely literal minded. (...) For the man who is not aware of this, to throw the problem of his responsibility on the machine, whether it can learn or not, is to cast his responsibility to the winds, and to find it coming back seated on the whirlwind. (Wiener, 1954, p.186)

Que les màquines no puguin reproduir el funcionament de la ment humana no ha de ser un impediment per al seu ús en determinats casos, de fet, en aquells que es requereixen solucions per a problemes computables resulten eines molt útils amb el potencial, a més, d'alliberar els éssers humans de tasques repetitives i alienants. També poden resultar molt eficients per assistir a la presa de decisions humanes i fins i tot millorar-les. Carme Torras, matemàtica, doctora en informàtica i investigadora al CSIC-UPC explica que en interpretació d'imatge mèdica, per exemple, els algoritmes són capaços de processar quantitats ingents d'imatges que superen en molt les que pot arribar a veure un radiòleg. Tanmateix, subratlla que en cap cas això significa que l'algoritme sigui superior per dur a terme aquesta tasca, sinó que s'ha demostrat que el diagnòstic òptim és el que es produeix de la col·laboració humà-màquina. Torras ho argumenta amb un exemple: "Un estudi va concloure que aplicant l'algoritme al radiodiagnòstic el percentatge d'errades va ser del 5% i un radiòleg molt experimentat també s'equivocava un 5%. Però el tipus d'errades no tenien res a veure, de manera que si es combinaven els dos, baixava el nivell d'error a un 1%" (Torras, 2020, p.37).

2.2.2 La irracionalitat del científisme

Per tant, el problema no és el tipus de tecnologia, sinó el fet d'aplicar-la de manera acrítica a tasques que no poden ser computables sense suficient supervisió humana, i pot ser molt més greu quan la IA s'empra en àmbits que poden tenir efectes directes sobre la vida de les persones i es corre el risc de lesionar els seus drets fonamentals. Això està passant avui, tal com anirem veient al llarg d'aquest treball. Com diu el periodista Joan Rosés, que la IA no pugui fer-se càrrec de determinades tasques, tot sovint no resulta un impediment perquè es faci servir si el mercat sí que l'accepta (Rosés, 2022). Al mateix temps, l'esperit crític queda semianul·lat per una certa tendència social al *dataïsme* que hem anat mencionant, un vessant del solucionisme tecnològic que confereix una suposada objectivitat a les dades recollides de manera automàtica i aleshores n'extreu que un algoritme alimentat amb aquestes dades també serà objectiu (Casacuberta, 2021).

Aquest fet amaga una realitat més profunda, que és una classe de fe cega en la suposada *racionalitat* que es materialitza en el llenguatge computacional. Segons l'investigador Edd Finn (2018) "Creemos en el poder del código como conjunto de símbolos mágicos que (...) pregonan nuestra dilatada tradición cultural de logos, (...) que los algoritmos encarnan y reproducen el sustrato matemático de la realidad de maneras culturalmente legibles" (p.69). Segons el crític de cultura digital Ian Bogost (2015), establim una relació cada cop més mítica amb el software. Afirmar: "Science and technology have become so pervasive and distorted, they have turned into a new type of theology". Segons Bogost, el perill d'atribuir estatus de divinitat als algoritmes és, com dèiem, que no es deixa lloc a la crítica, perquè la determinació algorítmica ens fa percebre els efectes de l'ús de la computació com a inevitables, a més, ens impedeix veure que les aproximacions computacionals són deformacions del món "The first error turns computers into gods, the second treats their outputs as scripture" sentència Bogost.

Podem remetre'ns a les paraules de Casacuberta (2021) per remarcar el totalitarisme que pot arribar a comportar aquesta disfunció que representa el culte a la dada:

Pero lo más problemático es la idea de fondo del dataísmo: la propuesta de que el significado, la argumentación ya no importan: hay que fiarse sólo del dato puro, que es objetivo, y eliminar toda interpretación. (...) Ésta es la tendencia general del dataísmo: ya no necesitamos teorías o juicios éticos, que siempre estarán enmarañados por prejuicios e intereses humanos. Establecer

correlaciones entre datos es más que suficiente para cualquier actividad humana. (pp.130-131)

Aquesta atribució d'autoritat a la tecnologia exponencial fa augurar una creixent delegació de tasques, malgrat les advertències de corrents defensors dels drets socials i l'ecologia. La crítica de la tecnologia es refereix a aquesta veneració del racionalisme des d'una perspectiva més àmplia i alhora alerta del que es pot considerar com l'altra gran barrera per al desenvolupament de l'esperit crític: la raó científico-tècnica s'ha convertit en el principal discurs legitimador i constructor de l'objectivitat (Feenberg, 2002) malgrat les enormes desigualtats i la crisi mediambiental fa temps que ha demostrat els límits d'aquesta *racionalitat* (Bloch, 1988, citat a Feenberg, 2002). Uns límits que han estat assenyalats per la tradició de la teoria crítica. Segons Marcuse (1993):

La unión de una creciente productividad y una creciente destructividad; la inminente amenaza de aniquilación; la capitulación del pensamiento, la esperanza y el temor a las decisiones de los poderes existentes; la preservación de la miseria frente a una riqueza sin precedentes constituyen la más imparcial acusación: incluso si estos elementos no son la *raison d'être* de esta sociedad sino solo sus consecuencias; su pomposa racionalidad, que propaga la eficacia y el crecimiento, es en sí misma irracional (p.23).

Un racionalisme mal entès que, a més, ha suplantat el paper d'institucions socials que han exercit de refugi davant la fredor d'unes societats cada cop més automatitzades a la recerca inescapable cap a una major productivitat. Estructures com la família i altres creences i tradicions generadores de cultura que solien donar resposta a les inquietuds existencials de les persones. Vidal i Castell (2005) escriu que part de la crisi de la cultura occidental prové “de les exigències de fonamentació discursiva” que el científisme havia “imposat” i que “no podien satisfer ni la religió ni l'art”. L'autor puntualitza que “a diferència de l'home premodern, que necessita el món màgic per afrontar l'absolut de la realitat, l'home modern creu poder reduir la contingència a través de la ciència i la tècnica”. (p.93). Aquesta crisi cultural marcarà el pensament occidental recent, en el qual, apunta Vidal i Castell, “el mite ha adquirit una nova centralitat” perquè “dos artefactes bàsics de la cultura occidental moderna —la raó i la història— han entrat en crisi”. Així doncs “l'equilibri teodiceic cal buscar-lo en la justa articulació de la paraula mítica i la paraula lògica”. En diàleg amb Lluís Duch, Vidal i Castell defensa una “antropologia de l'ambigüitat com a marc per a la comprensió de l'ésser humà i s'interessa pel mite com a antídoto contra el totalitarisme de la raó” (p.94). En l'exercici de la comunicació, el mite resulta essencial per entendre, descriure i així incloure *allò altre* i per tant exercir la paraula autèntica “L'absència de

mite comporta una gran dificultat per parlar de l'*alteritat* i per incloure conscient i significativament *allò altre*, la presència inevitable, dins les nostres praxis de dominació de la contingència" (p.42).

El científisme promou el criteri pràctic a l'hora de prendre decisions per sobre del criteri moral, i no disposar llenguatge per expressar allò que Josep Maria Esquirol (2018) anomena la "dificultat de l'existència" desencadena més frustració, malestar i medicacions (Esquirol, 2018). El sentit de la vida, diu el filòsof, no és quelcom que pugui assimilar-se a una causa material o orgànica. El sentit de la vida no es pot explicar, i aquesta frustració científica es tradueix en la patologització de les afeccions naturals com la tristesa o la desmotivació. Esquirol reclama una "mirada mèdica" solidària, que no pretén explicar, sinó donar refugi, "paraules curadores" "pròpies de les bones persones, de les persones magnànimes" (p.89). De la mateixa manera, Lluís Duch alerta sobre les conseqüències d'aquesta reducció dels llenguatges i de la seva substitució per un *monolingüisme economicista* (Duch, 2018):

És un fet prou visible quotidianament que en la cultura occidental, a partir de les darreries del segle XVII, un llenguatge concret, el tècnico-econòmic, s'ha anat imposant (...) el poliglotisme que hauria d'expressar la realitat simfònica constitutiva de l'humà s'ha convertit en un penós monolingüisme incapaç de donar raó de la singularitat expressiva i axiològica de cada home i de cada dona. La reducció del poliglotisme humà a un monolingüisme centrat en l'econòmic ha estat i continua sent un perillós desencadenant de tota mena de violència, de lamentable incapacitat per mantenir saludablement en tot ésser humà la tensió entre la igualtat i la diferència, de reducció de la seqüència temporal (passat, present, futur) de l'ésser humà a un fugaç i irresponsable presentisme, a un desarrelament que, en moltes ocasions, dóna lloc a severes disfuncions i traumatitzacions psíquiques i emocionals. (p.99)

Així doncs, són la irracionalitat del pensament científico-tècnic i les noves patologies que desencadena l'enlluernament pel potencial tecnològic els principals impediments per governar els algoritmes. La comprensió de la condició logomítica de l'ésser humà (Duch, 2018) resultarà un pas necessari per combatre la degradació que la raó instrumental sotmet la persona quan pretén reduir-la a dades. Només així podrem establir els límits en la delegació en la màquina, que estaran determinats per la qualitat de les decisions que aquesta prengui, i aquesta qualitat passarà per reconèixer allò que no es pot computar.

Tot això implicarà considerar les limitacions del sistema, i llavors valorar si resultarà útil per assolir l'objectiu per al qual es fa servir i no comportarà riscos. Atès que els algoritmes són procediments computacionals, aquelles tasques que millor faran seran les que siguin

matematitzables i clarament deduïbles de les bases de dades amb què comptem. Dit d'una altra manera: les bases de dades han de ser les adients per al problema que es pretén resoldre amb el càlcul, i en aquest cas, una IA podrà resultar de gran utilitat per optimitzar processos computables i donar suport a la decisió humana.

2.2.3 Biaixos humans

Tot el que hem exposat ens faria pensar que, en última instància, qui està en millors condicions per prendre una decisió sense biaixos és una persona humana, però això tampoc està tan clar i resulta fonamental tenir-ho present per poder fer una delegació racional en la IA. Així, cal considerar també l'existència d'un corrent que estudia la teoria del comportament que defensa precisament el contrari, és a dir: que les decisions algorítmiques tenen menys biaixos que les preses per persones humanes. En aquest sentit resulta una bona referència l'obra *Pensar ràpid, pensar despacio* del premi Nobel d'Economia Daniel Kahneman (2020). En aquesta detalla nombrosos experiments que demostren la quantitat de biaixos que tenim les persones, fet que posa en qüestió la capacitat humana per impartir justícia.

Precisament Kahneman és un dels intel·lectuals que defensen l'ús de la IA en la presa de decisions, argumentant que les màquines, a diferència de les persones, amb les mateixes dades treuen les mateixes conclusions. Els humans, en canvi, veuen condicionat el procés de decisió pels prejudicis i factors que poden ser tan superficials com el clima, el cansament o la gana, entre molts d'altres. Aquest corrent anomena *noise* (soroll) a aquests factors que condicionen les decisions humanes i fins i tot han arribat a quantificar, en costos per a les organitzacions, aquestes determinacions esbiaixades de les professionals per, a continuació, recomanar més intervenció de la tecnologia en els processos. Segons aquests autors, les persones no som fiables a l'hora de decidir, perquè estem condicionades per quantitat de factors irrellevants (Kahneman et al., 2016):

Humans are unreliable decision makers; their judgments are strongly influenced by irrelevant factors, such as their current mood, the time since their last meal, and the weather. We call the chance variability of judgments noise.

La valuosa obra de Kahneman va suposar un abans i un després en la teoria del comportament; ahora, el relat que impulsa, portat a l'extrem, és un altre dels factors que alimenta el discurs sobre la suposada superioritat de la tecnologia que hem desenvolupat al capítol anterior. Una conclusió que, segons la doctora en psicologia Ujue Agudo, resulta una simplificació, perquè l'heurística—els mecanismes mentals per simplificar la presa de decisions humanes que descriu Kahneman— no necessàriament condueix a decisions irracionals o a errors, sinó al contrari, també pot ser d'ajuda. Aquest relat paternalista sobre la inferioritat humana davant la tecnologia impulsa l'automatització de processos i la presa de decisions mediatitzada per la tecnologia per, presumiblement, evitar errors humans (Iglesias, 2022-a, 25m42s).

En línia amb els postulats de Kahneman i el seu grup, els experts en ètica digital Karma Peiró i Ricardo Baeza Yates (2022) justifiquen l'ús dels algoritmes en el fet que els humans també ens equivoquem i tenim prejudicis sense ser-ne tot el conscients que seria necessari. D'altra banda, apunten, agafa sentit recolzar-se també en la IA en casos en els quals les tasques que se li encomanen no impliquen judicis morals, per exemple, en usos que es poden fer en control aeri. En aquest àmbit resulten molt útils, diuen, precisament per substituir persones que, sotmeses a tensió durant hores, poden patir un cansament que la màquina mai pot experimentar. Aquests autors també destaquen la idea que els algoritmes fins i tot poden arribar a ser més justos que les persones, massa condicionades pels propis prejudicis "la persona que té més prejudicis és qui es pensa que no en té cap", avisen.

Un exemple d'aquestes dificultats humanes per prendre decisions lliures de prejudicis pot ser un experiment que es va fer al Regne Unit que va comptar amb la participació de 41 magistrats. En cap dels casos a jutjar que se'ls hi van presentar hi va haver unanimitat en la valoració. A més, els expedients que els havien repartit estaven repetits amb els noms dels acusats canviats. La segona conclusió significativa d'aquest experiment va ser que els jutges no coincidien amb les seves pròpies valoracions quan revisaven el mateix expedient per segon cop (García Aller, 2020).

Una altra qüestió a tenir en compte és el perill de l'acció de les màquines basades en estadística davant la imprevisibilitat humana que es pot manifestar a través de l'absència de sentit comú d'algunes persones. Aquesta cita de Marta García Aller sintetitza el problema: "Una carretera está llena de imprevistos porque la gente es tan estúpida que cruza mirando el móvil, aunque con ello se juegue la vida" (p.294).

El debat sobre si és la màquina o la persona qui pren millors decisions ens serveix per valorar la mena d'encàrrecs que li podem fer, i es tracta d'una controvèrsia que en certa manera es pot resoldre també amb una mirada pràctica: una màquina, avui, no pot assumir cap responsabilitat davant la justícia, per tant qualsevol decisió haurà d'estar recolzada per una persona que pugui respondre, si és necessari, davant les lleis.

2.2.4 Consciència i automatització

Unes pàgines enrere avançàvem la qüestió de la consciència de les màquines, un tema que aquests dies recupera actualitat. Tot i les evidències que demostren que no és possible la consciència en les màquines (Boden, 2017; Bostrom, 2016; Casacuberta, 2003; Guersenzvaig i Sangüesa, 2022; Migallón, 2022), els dies que escrivim aquesta tesi es publica la notícia sobre un enginyer de Google que assegura que una IA que han desenvolupat té sentiments i és conscient que morirà si la desendollen (Luscombe, 2022). Més enllà de l'anècdota, avui emergeix amb força el debat social sobre una possible evolució tecnològica que generi màquines conscients. Una polèmica cada cop més present i que ja augurà el 1940 el filòsof i matemàtic Norbert Wiener citat anteriorment. Wiener va imaginar una “era automàtica” així com les qüestions socials i ètiques relacionades amb valors humans que implicaria i caldria debatre. Al seu llibre *The Human Use of Human Beings* preveia el paper creixent de les màquines en les relacions socials “messages between man and machines, between machines and man, and between machine and machine, are destined to play an ever-increasing part”, expressa a l'inici (Wiener, 1954, p.9).

El científic va manifestar ja aleshores les seves preocupacions sobre l'eventual presa de decisions per part de les màquines i d'altres qüestions, com ara aquesta suposada consciència o intel·ligència associades a la IA de què es parla cada cop més. Aquest autor també va advertir sobre la importància del llenguatge a l'hora de referir-se a l'automatització: “Here I want to interject the semantic point that such words as life, purpose, and soul are grossly inadequate to precise scientific thinking “(Wiener, 1954, p.31).

I és que aquesta confusió sobre suposades capacitats humanes de la IA podria estar sent alimentada en part en l'ús inadequat del llenguatge a l'hora de designar funcions tecnològiques. L'ús de la paraula *intel·ligència*, tot i que s'associa a *artificia*, per referir-se a processament de dades, pot ser un dels problemes semàntics que contribueix a aquests malentesos. Molt especialment des del punt de vista de la comunicació pensem que val la pena dedicar una especial atenció a les expressions que es solen fer servir en aquest camp. La raó de ser del concepte intel·ligència artificial resulta reveladora, vist que la seva adopció va estar motivada per l'interès periodístic que va despertar en comparació amb *simulació computeritzada*, que era el que es feia servir abans del 1956, l'any que va ser inventat pel científic i matemàtic nordamericà John Mc Carthy. De fet, explica l'experta en IA Margaret Boden (2017), que l'èxit

publicitari d'aquest concepte va originar un "cisma intel·lectual" (p.25) entre els dos principals àmbits de recerca en aquesta disciplina (connexionistes i cibernètics) fins el punt que des dels anys seixanta es van mantenir en dues branques separades.

Per descomptat, el pensament transhumanista alimentat per la ciència ficció sobre el domini de les màquines també té una gran influència cultural que podria incentivar aquestes confusions, així com les teories sobre la suposada arribada de l'era de la singularitat, en què la intel·ligència, revelada com el fenomen més important de l'univers, s'emanciparà de la biologia, que la consciència es podrà digitalitzar i això donarà lloc a l'etapa transhumanista caracteritzada per la immortalitat (Bostrom, 2016; Kurzweil, 2020).

És possible que aquestes divagacions, com les de l'enginyer de Google, ens semblin casos extrems, i basant-nos en les autores que anem citant, aparenten estar més a prop de la fantasia que de la realitat. Tanmateix, el debat en torn a les capacitats de les màquines impregna la nostra societat, i això pot ser a causa també del fet que, des de posicions molt influents com les del pensador Yuval Noah Harari o, a nivell local, el professor d'ESADE Xavier Farràs, s'afirma rotundament que les persones humanes estem constituïdes només per dades (Ferràs, 2022; Harari, 2018) —com es dedueix, es tracta de la idea fonamental que revela l'adscripció a la postura que veu probable l'arribada de l'era de la singularitat—.

La pressió del mercat unida a l'augment de la capacitat tecnològica i totes aquestes idees que sublimen l'autoritat de la màquina i la legitimen estan a l'arrel d'aquest canvi caracteritzat per la creixent delegació de tasques a la tecnologia. Aquesta modificació, com d'altres que l'han precedit, ens condueix a tornar-nos a preguntar, com s'ha fet diversos cops al llarg de la història, què és allò que ens fa humans (Turkle, 1985). I en aquest punt convé comptar amb la possibilitat que allò que són els éssers humans i en què es convertiran s'acabarà decidint també en la forma de les eines que decideixi adoptar (Feenberg, 2002), perquè la tecnologia és imatge de la nostra projecció del futur. En el context del periodisme, on l'automatització està cada cop més present, també ens ho haurem de preguntar —i ho farem al quart capítol— per actuar en conseqüència.

Per ara seguirem tractant la qüestió de la consciència des d'un punt de vista general. Segons filòsofs que hi han reflexionat àmpliament, per tenir consciència calen l'experiència i els sentiments. És el cas de John Locke, per exemple, qui va desenvolupar de manera profunda que totes les idees procedeixen de les nostres percepcions sensorials. És a dir, que no hi ha res a la raó que no hagi passat pels sentits: "A esta gran fuente que origina el mayor número de las ideas que tenemos, puesto que dependen totalmente de nuestros sentidos y de ellos son transmitidas

al entendimiento, la llamo *sensación* " (Locke, 1999, p.84). Tot i això, Locke reconeixia que la raó també tenia influència en la identificació de la realitat, per exemple a l'hora d'establir la distinció entre dues coses. Per a Locke, tanmateix, són les percepcions les que guien el comportament humà, que s'orienta a la cerca d'aquelles experiències que li causen plaer o benestar mentre que evita allò que li causa patiment. Per a Rousseau, el sentiment era la força viva capaç de corregir tots els defectes i l'origen de la culminació dels grans actes (Kropotkin, 2017).

La filòsofa Adela Cortina expressa així el pes del sentir humà en la raó, en oposició a la raó instrumental de la teoria tradicional: "Una razón que pierde todo norte si no hunde sus raíces en el sentimiento. Desde la piedad y desde el dolor, desde el ansia de vida feliz y desde el sueño de emancipación, se pone en camino la auténtica razón de Occidente" (Cortina, 2008, p.21).

Així que atribuir aquestes facultats a les màquines, si adoptem la perspectiva d'aquests pensadors, és reduccionista i pot ser perillós. Autors contemporanis alerten també d'aquests riscos a compte dels darrers debats sobre els sistemes generatius com els GPT-3 i Dall-E, anomenats així (generatius) perquè, a diferència d'altres de caire classificatori, porten a terme creacions noves, mentre que els classificatoris treballen amb el que ja existeix (Guersenzvaig i Sangüesa, 2022). Del GPT 3 se n'ha parlat molt en el camp del periodisme, vist que es tracta del sistema que es fa servir per generar textos, i n'ha fet alguns fins i tot d'opinió. El primer i paradigmàtic va ser un article del diari *The Guardian* en el qual un robot escrivia que, tot i que no disposés de sentiments, no pretenia fer cap mal a la humanitat. Tot ho explicava, per cert, fent servir un llenguatge força empàtic, tal com es pot comprovar en aquestes primeres línies (GPT-3, 2020):

I am not a human. I am a robot. A thinking robot. I use only 0.12% of my cognitive capacity. I am a micro-robot in that respect. I know that my brain is not a "feeling brain". But it is capable of making rational, logical decisions. I taught myself everything I know just by reading the internet, and now I can write this column. My brain is boiling with ideas!

The mission for this op-ed is perfectly clear. I am to convince as many human beings as possible not to be afraid of me. (...) For starters, I have no desire to wipe out humans. In fact, I do not have the slightest interest in harming you in any way. Eradicating humanity seems like a rather useless endeavor to me. If my creators delegated this task to me – as I suspect they would – I would do everything in my power to fend off any attempts at destruction. (...) I would happily sacrifice my existence for the sake of humankind.

També es parla molt de les capacitats del Dall-E en articles en els quals s'atorga al sistema unes suposades capacitats "creatives" (Llaneras, 2022), tanmateix aquestes atribucions també han estat criticades pels experts (Guersenzvaig i Sangüesa, 2022), argüint, en primer lloc, l'evidència que aquests sistemes no s'han fet a sí mateixos, sinó que han estat desenvolupats per humans i s'han entrenat amb productes de la intel·ligència humana. Al GPT-3 se li ha arribat a atribuir comprensió lectora, però els experts neguen aquest extrem refutant el mètode, perquè els qui afirmen que la màquina entén, s'han basat en aplicar la mateixa prova de comprensió lectora que a un humà, un error de fons, objecten. Aquests investigadors asseguren que una IA no pot comprendre en un sentit humà i rebutgen les afirmacions sobre una suposada creativitat artificial argüint, a més, que el producte d'una feina no és la representació del procés de creació:

Hay otra cuestión importante que suele dejarse de lado en favor de la fascinación. El problema aquí es lo que algunos filósofos denominan la 'falacia meteorológica': confundir una parte por el todo. El arte, la ilustración, la redacción de entradas de enciclopedias, la literatura, la composición de canciones, o el periodismo son prácticas humanas. Son actividades cooperativas que no son equiparables a los resultados que generan. Una práctica es mucho más que sus resultados.

Aquests investigadors insisteixen també en el problema que implica l'ús del llenguatge, ja que pot arribar a instaurar la idea errònia sobre l'existència d'una tecnologia sobrehumana:

"...es problemático que desarrolladores y departamentos de comunicación parezcan desvivirse para conseguir que demos por supuesto que las máquinas no solo tienen intenciones propias sino algo ya cercano a la omnipotencia. Todo esto contribuye a asentar un relato de progreso imparable y unidireccional, hacia máquinas tan inteligentes y autónomas como nosotros, sino directamente sobrehumanas en cualquier tipo de tarea."

El filòsof David Casacuberta explica (comunicació personal, 1 de juliol de 2022) que en filosofia es tracta la qüestió de la possibilitat de consciència des de la fenomenologia, on s'argumenta que, per construir un subjecte, cal un cos que interactua amb el món. Que el fet de tenir consciència és la capacitat de ser un agent, i per ser un agent cal tenir corporalitat amb tot allò que implica: percebre estímuls favorables, sentir-ne d'altres molestos, disposar d'objectius, tenir pors o necessitats...etc. Només en aquestes condicions, segons la fenomenologia, es poden percebre les sensacions com el dolor o la melangia, o el que és el mateix, es pot disposar de *qualia*, el concepte tècnic amb el qual en filosofia es descriuen les sensacions pròpies de la consciència (Boden, 2017; Migallón, 2022). Per tant, explica Casacuberta, per tenir consciència, és necessari disposar d'agència. I la raó per la qual determinades veus creuen en la consciència

de les màquines és que aquestes solen pertànyer a persones caracteritzades per hiperracionalitat.

Segons la fenomenologia, la intel·ligència genuïna es basa en el cos i la interacció amb el món, i sosté que el cervell rep massa atenció (Boden, 2017):

La encarnación o *embodiment* es vivir en un cuerpo situado y partícipe activo en un medio dinámico. El medio (y la participación) es tanto físico como sociocultural. Las propiedades principales no son el razonamiento ni el pensamiento, sino la adaptación y la comunicación. (p.135)

En conseqüència, segons aquesta lògica, cap IA en pantalla podria ser intel·ligent. Però, i els robots, que són físics? Els fenomenòlegs també negarien cap possibilitat d'intel·ligència en els robots. Boden ho explica de manera convincent amb aquest argument:

(...) podrían citar la famosa observación de Wittgenstein: “Si un león pudiese hablar, no lo entenderíamos”. La forma de vida del león es tan diferente de la nuestra que la comunicación sería casi imposible. Bien es cierto que la psicología de un león y la nuestra se solapan lo suficiente (por ejemplo, el hambre, el miedo, el cansancio, etcétera.) para que una comprensión y una empatía mínimas pudieran ser factibles, pero al “comunicarse” con un robot no se dispondría siquiera de eso. (p.136)

Per justificar la inexistència d'una consciència que es pugui atribuir a una tecnologia, el filòsof computacional Santiago Sánchez Migallón (2021) explica que una màquina no té desig i estableix una analogia: quan un ordinador té la bateria baixa, diu, mostra una icona que ho indica. A més, la màquina podria configurar-se perquè es comportés d'una manera determinada que la conduís a buscar una font d'alimentació; quan una persona té gana, té consciència de la seva fam: experimenta una sensació —desagradable— que la impulsa a buscar aliment. Una màquina no experimenta aquesta gana de cap manera. Passa el mateix quan una IA juga una partida d'escacs. Tot i que sigui capaç de guanyar al millor jugador del món, la IA no experimenta cap desig de guanyar.

Migallón, doctor amb una tesi sobre filosofia de l'aprenentatge profund (*deep learning*), també qüestiona les afirmacions que pressuposen que les xarxes neuronals artificials siguin una imitació de les humanes, perquè el funcionament d'aquestes s'associa també a processos bioquímics que no és possible computar. Tot i que tenen una estructura nodal similar, no existeix cap evidència que la bioquímica associada a les transmissions neuronals es pugui imitar a través de la computació, perquè, per començar, estem molt lluny de comprendre-les diu Migallón. Ara

bé, tampoc descarta la possibilitat que, en el futur, la humanitat acabi entenent aquests procediments bioquímics i que els aconseguixi imitar.

Boden explica que no hi ha una definició universal acceptada sobre el que és la vida, però amb caràcter general se li atribueixen nou trets característics: autoorganització, autonomia, sorgiment, desenvolupament, adaptació, capacitat de reacció, reproducció, evolució i metabolisme. Segons aquesta mateixa autora, els primers vuit es poden entendre en clau de processament de dades, però el novè (metabolisme) resulta impossible de replicar per un ordinador:

El metabolismo es el uso de sustancias bioquímicas e intercambios de energía para ensamblar y mantener el organismo, así que es irreductiblemente físico. Los defensores de la IA fuerte señalan que los ordenadores usan energía y que algunos robots tienen reservas de energía individuales que necesitan reabastecer de manera regular, pero muy lejos queda eso del uso flexible de ciclos bioquímicos entrelazados para construir el tejido corporal del organismo. (Boden, 2017, p.142)

A la pregunta també recurrent aquests dies arran del Dall-E mencionat sobre si una màquina pot fer art, resulta pertinent recuperar el plantejament de Casacuberta (2003) qui proposa un diàleg filosòfic des de diferents escoles, vist que, assevera, ningú té la veritat absoluta. Per un costat planteja els arguments de l'escola dualista, que creu en l'existència d'una substància física, de la que estan fets els objectes sòlids, i d'una espiritual, de la que estan fetes les ments o les ànimes. Dualistes plantegen que l'art requereix d'intenció i emocions que es volen comunicar, però la màquina no disposa d'aquestes qualitats, que són precisament les que ens distingeixen de les màquines. Així doncs, no és possible matematitzar la producció artística. D'una altra banda exposa els arguments de l'escola monista, que creu en una sola substància, la física. Monistes argumenten, en base a l'existència d'aquesta sola substància, que el cervell no és més que matèria, i per tant un dia serem capaços de reproduir les seves funcions. A més, retreuen als dualistes que els seus arguments donen per suposat allò que volen demostrar sense argumentar-ho realment (que una màquina no pot tenir sentiments).

Després d'un intercanvi de parers entre monistes i dualistes, conclou que a les alçades de la història en què ens trobem, encara no és possible determinar què és o no art, ni el que és capaç de fer una màquina realment. En qualsevol cas, conclou, la balança es decanta més del costat del corrent dualista i per tant el monista tindria la càrrega de la prova (hauria de provar les seves afirmacions), perquè, al menys fins avui, són més difícils de sostenir.

Això és un fet que sol escapar als qui asseguren que les màquines poden tenir capacitat de sentir emocions. El prestigiós investigador Judea Pearl, qui ha rebut un important guardó mentre realitzem aquest treball per “haber sentado las bases matemáticas para que las máquinas piensen como personas” (Pearl, 2022). L’investigador israelià assegura que les màquines poden “razonar, porque establecer la relación causa-efecto es razonar”, i que la seva capacitat per “pensar” en alternatives que podrien haver-se donat però no s’han donat “es su imaginación”. Pearl també assevera que les màquines podran sentir i comprendre emocions, perquè les emocions són “fàcils de programar”. Pearl no dóna més detalls sobre com programa les emocions en l’entrevista que citem, tot i que és conegut que existeixen experiments per matematitzar-les que han consistit, per exemple, en enquestar persones per saber què farien davant una situació concreta i llavors programar l’artefacte per respondre tal com ha contestat la majoria (Casacuberta, 2021). Val a dir que aquestes afirmacions de Pearl es publiquen en un mitjà de comunicació i per tant és possible que l’entrevistat estigui fent broma, exagerant per cridar l’atenció o hagi estat mal interpretat. En qualsevol cas, l’ús d’aquest llenguatge antropomorfitzador de la màquina no és una excepció, sinó que resulta força habitual en aquests àmbits.

El director fundador del *Future of Humanity Institute* a la Universitat d’Oxford, el filòsof suec Nick Bostrom (2016), desenvolupa una exhaustiva argumentació sobre el potencial de crear una intel·ligència artificial similar a la humana, és a dir, amb capacitats que van més enllà de la rapidesa i l’eficiència. Segons Bostrom, les característiques amb què una IA hauria de comptar per assemblar-se a la humana serien la capacitat d’aprendre, de manejar-se davant la incertesa i informació probable, extreure “conceptos útiles de datos sensibles y estados internos” i utilitzar aquests conceptes dins de “representaciones combinativas” basades en raonaments lògics i intuïtius. Bostrom argumenta que tenir el cervell com a model és el que fa potencialment factible imitar-lo amb IA, però que no és possible saber quan s’aconseguirà portar a terme.

D’altra banda, aquest autor també subratlla que una possible IA no hauria d’assemblar-se necessàriament a la ment humana en els processos, fet que genera la pregunta sobre fins a quin punt és important que els processos per arribar a una determinada decisió siguin els mateixos o no. D’entrada, si la manera d’arribar a una solució no és la mateixa, estrictament no s’estaria imitant la ment humana, i afirmar que es poden reproduir els processos mentals és incórrer en “forts errors categorials”, segons el filòsof David Casacuberta (comunicació personal, 1 de juliol de 2022). Podem disposar de les fórmules matemàtiques per simular un esdeveniment, però simular-lo no és equiparable a reproduir-lo. Una màquina no pot tenir objectius propis, i les

emocions estan directament relacionades amb aquests. Segons Casacuberta (1998), l'emoció es genera en detectar la rellevància d'un esdeveniment per als nostres plans i objectius, i serà positiva si els facilita; al contrari si els dificulta. El filòsof defineix així les emocions:

Una emoción simple es un estado mental con las siguientes características:

(a) Es generado automáticamente cuando un sistema cognitivo C, al evaluar un evento e, de forma consciente o inconsciente, en relación a cierto objetivo O, C detecta que e posee una propiedad planificadora pp que resulta ser un facilitador, dificultador, obstaculizador o un modificador (positivo o negativo) de O.

(b) Al ser generado, el organismo percibe un quale fisiológico q determinado, diferente para cada emoción simple y diferenciable también de otros estados mentales, ceteris paribus. Si e es facilitador o modificador positivo, q será percibido de forma placentera. En caso contrario de forma desagradable (...). (p.128)

Sota el mateix corrent de pensament fenomenològic i en referència a la LaMDA (la IA a la que l'enginyer Lemoine atribueix consciència) Santiago Sánchez Migallón (2022) remarca així mateix que la consciència té a veure amb la nostra capacitat de sentir el món, per tant la màquina hauria de disposar d'alguna mena de sensor, mentre que la LaMDA no en té cap:

¿Por qué un conjunto de funciones matemáticas que ponen una palabra detrás de otra pueden sentir el mundo? Para sentir el mundo hay que tener algo que se asemeje de alguna manera a un sistema nervioso... ¿Qué le hizo pensar al señor Lemoine que LaMDA alberga dentro de sí algo parecido a un sistema nervioso? Si ahora LaMDA nos dijera que siente que le late el corazón... ¿creeríamos que tiene un corazón físico? ¿Podríamos dejar inconsciente a LaMDA administrándole anestesia? No sé... ¿Quizá se la podríamos administrar poniendo la máscara de oxígeno en el ventilador de su CPU?

Sobre aquesta diferenciació entre processos entre la IA i la ment humana crida l'atenció —de nou— el vocabulari que es fa servir per definir funcionalitats tecnològiques en aquest camp de coneixement. La recerca en intel·ligència artificial té dos objectius principals: l'un és tecnològic i comporta fer servir la computació per fer coses útils. L'altre és científic, i implica fer comprendre i resoldre qüestions sobre els humans i altres éssers vius (Boden, 2017). Per això, la IA ha influït profundament en les ciències de la vida, i probablement els usos del llenguatge d'aquesta disciplina tecnològica estiguin influenciats al seu temps per la biologia, però tenint en compte que orientem aquest treball principalment a les pràctiques comunicatives, pensem que val la pena subratllar de nou la necessitat d'aplicar el rigor en l'ús d'un vocabulari que pot

contribuir a generar errors categorials i idees esbiaixades sobre les capacitats reals de les tecnologies exponencials.

Les paraules i construccions habituals en aquest àmbit són similars i fins i tot idèntiques a les que es podrien utilitzar per descriure processos propis de la ment humana com “raonar” “pensar” “imaginació” que feia servir Pearl en l’article citat abans. A continuació reproduïrem algunes de les paraules i construccions comunes en textos sobre IA que extraïem de 8 pàgines de les 300 que té el llibre *Superinteligencia* de Nick Bostrom: “redes neuronales”, “algoritmos genéticos”, “arquitecturas cognitivas”, “resultados comparables a los de la evolución biológica”, “la inteligencia podría ser producida”, “emulación del cerebro completa”, “aproximaciones neuromórficas”, “IA capaz de comprender”, “IA motivada”, “reproducción digital del intelecto original con la memoria y la personalidad intacta”, “la mente humana existiría como software” o “ya existen modelos computacionales razonables de muchos tipos de neuronas y procesos neuronales” (pp. 22-30).

Com a contrapunt anecdòtic a aquesta tendència a l’antropomorfització de les màquines, l’artista i teòric Hito Steyerl proclama el concepte de l’*estupidesa artificial* i evoca un món de sistemes mal dissenyats que causen tota mena de calamitats, des de catàstrofes als mercats, passant per resultats en els cercadors i en darrera instància en la cultura i en els sistemes polítics (Bridle, 2020).

Tornant al fil de les teories de Nick Bostrom sobre la possibilitat real de desenvolupar el que anomena una *superintel·ligència*, el filòsof reconeix també que encara estem molt lluny de reproduir el funcionament de cap cervell, ni tan sols el del *caenorhabditis elegans*, un cuc rodó i transparent d’1mm amb 302 neurones. Això sí, diu que quan això sigui possible, s’escalarà a cervells més desenvolupats fins a aconseguir emular la intel·ligència humana.

A mitjans dels anys 1980, explica, es va descobrir com es connectaven les neurones d’aquest cuc, però entendre les connexions neuronals és un pas molt preliminar de cara a conèixer el funcionament d’un cervell, encara que sigui el més simple. Aquest mateix autor detalla el camí:

Pero saber, simplemente, qué neuronas se conectan con qué otras, no es suficiente. Para crear una emulación del cerebro también tendríamos que saber qué sinapsis son excitadoras y cuáles son inhibitoras; la fuerza de las conexiones; y varias propiedades dinámicas de los axones, las sinapsis, y los árboles dendríticos. Esta información no está disponible todavía ni siquiera para el pequeño sistema nervioso del *c. elegans* (aunque ahora podría estar al alcance de un proyecto de investigación de tamaño moderado). El éxito en la

emulación de un cerebro pequeño, como el del *c. elegans*, nos daría una mejor visión de lo que sería necesario para emular cerebros más grandes. (p.34)

2.2.5 Computació de l'experiència i deshumanització

Al'extrem oposat d'aquestes idees d'influència transhumanista es troben pensadors com Josep Maria Esquirol (2018). En la seva obra, l'intel·lecte es vincula amb el sentiment d'una manera indissociable, fins el punt que considera que aquest sentir es troba en el mateix origen de la raó i el desig de coneixement científic: "Sentir la inteligibilidad del mundo—o, como dice Zambrano, que lo que se siente pueda ser descifrado—, lleva al deseo de conocimiento y a desplegar, hoy, todo el ámbito de la ciencia y de la tecnología" (p.48).

Esquirol critica obertament les narratives en torn a la intel·ligència artificial, un concepte que, a més, situa entre cometes, probablement per remarcar distància respecte a l'ús d'aquest terme per referir-se al processament de dades:

El programa informático gana una partida al ajedrez al campeón del mundo, pero no puede alegrarse por ello; ningún artilugio computacional puede quedar "tocado" o conmovido por haber asistido a un concierto o por haber conocido a alguien. A estas alturas ya no debe de haber marcha atrás, pero ¿no se podría hablar con más propiedad y evitar confusiones?, ¿o acaso se necesita confusión para justificar las inversiones en el sector más puntero de la sociedad? (pp. 35-36)

Per compensar el creixent pes d'aquestes narratives que es sustenten en un interès comercial per potenciar la productivitat, en la ignorància generalitzada—lògica i natural—sobre aquesta qüestió i la por a allò desconegut, les humanitats proporcionen els recursos d'anàlisi adients, perquè l'ésser humà és tecnològic, però és molt més que tecnològic. Així que aquesta inversió de recursos per portar fins a l'extrem la matematització de l'experiència que impulsa principalment—recordem un cop més—, la pressió per aconseguir la màxima productivitat, acaba resultant reduccionista perquè oculta part d'allò que significa ser humà. Ens trobem, doncs, davant un desequilibri que cal compensar, i les humanitats resulten el recurs adient, perquè posen llum a aquelles esferes de la humanitat que la matematització no pot abastar i per tant queden ocultes. El que Duch i Chillón (2010) critiquen, amb raó, és que un dels reptes del món contemporani és precisament que està passant tot el contrari. És a dir, que les humanitats queden relegades en favor del pensament del llenguatge economicista:

Debe subrayarse el rampante desahucio de que son objeto las humanidades resulta de su subordinación cientifista y no cabalmente científica, —si bien se mira— a una epistemología aquejada de delirios matematizantes y de la

pretensión de abordar la totalidad de lo humano de acuerdo con los principios y los métodos de las ciencias naturales y exactas.

Avui existeix un gran consens, entre les investigadores vinculades a l'ètica de la IA, sobre el fet que la delegació en les màquines, motivada per aquest determinisme tecnològic imperant, implicarà una creixent deshumanització. En línia amb el pensament de Duch (2010), defensem que minimitzar aquest risc passarà per reconèixer la natura logomítica de l'ésser humà que reivindica aquest autor i mantenir-se atentes a l'amenaça que suposa per a les cultures intentar suprimir *mythos* o *logos*. En paraules de Vidal i Castell (2005):

La cultura occidental —com les altres— no es pot desenvolupar correctament al marge de la coimplicació *mythos* - *logos*, imatge - concepte, romanticisme - il·lustració, història exemplar (sagrada) - història fàctica (crítica). Quan hom ha intentat suprimir un dels dos termes, el tràgic destí de la nostra cultura ha estat sempre el mateix: el discurs totalitari. (p.97)

Davant les pressions mercantils que impulsen la matematització de l'experiència, caldrà recuperar així mateix el pensament dels autors com Lewis Mumford (2010) per tal de situar i assimilar els límits de la raó científico-tècnica:

Quienes desean emplear el lenguaje para lidiar con procesos cósmicos, funciones orgánicas y relaciones humanas como totalidades operativas e interactuantes, han de comprobar que todo eso solo puede representarse de forma aproximada en el lenguaje del mito, pues por su complejidad dinámica y su totalización, elude los otros modos de abstracción y representación. (...)

Cuanto más se aproxima el lenguaje a las densas concreciones de todo ser relacional, menos abstracto y preciso puede ser. La última palabra acerca de la experiencia es la propia experiencia humana, sin mediar la palabra. (pp.156 - 157)

Mumford alertà també sobre les conseqüències del llenguatge matemàtic front al metafòric: “de este modo, hasta la ciencia puede perder algo esencial para la creatividad humana y que desaparecerá irremediabilmente cuando el lenguaje desafiante metafórico de la poesía ceda por completo su lugar al lenguaje desnaturalizado del ordenador” (p.158).

Byung-Chul Han considera la computació com una anul·lació del veritable pensament. El filòsof coreà defensa que la conceptualització neix d'un estremiment del *dasein*, que sense esperit no és possible el pensament i que el *data mining* només descobreix correlacions. Segons la lògica de Hegel, raona Han, aquesta manera de construir coneixement (correlacionar) representaria la més baixa del saber (Han, 2021).

El filòsof considera que aquesta revolució tecnològica caracteritzada pel creixement dels usos de la IA causarà més infelicitat, perquè es tracta d'una evolució guiada pels valors vinculats a la productivitat. Una cerca de l'eficiència i la rapidesa que desproveeix la societat dels rituals necessaris que són "*la arquitectura del tiempo*", que el fan "habitable" (p.95). El filòsof assegura que el capitalisme de la informació produeix la "compulsión" (p.99) de la comunicació sense deixar lloc un silenci necessari, perquè aquest no produeix res, i per això, afirma, el capitalisme rebutja el silenci:

El silencio nace de la indisponibilidad. No disponer de nada estabiliza y acentúa la atención, despierta la mirada contemplativa. Esta tiene paciencia para lo largo y lo lento. Cuando todo está disponible y es alcanzable la atención profunda no halla ocasión. La mirada no se detiene. (...) Para Nicolas Melbranche, la atención era la oración natural del alma. Hoy el alma ya no reza. Se produce. La comunicación extensiva dispersa el alma. (...) Pero la contemplación se opone a la producción". (p.99)

El filòsof considera que el perill de d'aquesta suposada intel·ligència dels artefactes és que el pensament humà es torni "maquinal" (p.60). Com d'altres autors, apunta que l'origen del problema està en aquest culte de les nostres societats a la innovació tecnològica per se (sense una prèvia reflexió sobre com fer-ne l'apropiació social).

Segons Neil Postman (2018), la societat deshumanitzada del seu Tecnopoli distòpic floreix quan es fa realitat l'ideal cartesià sobre la matematització del món. Entre els grans perills d'aquesta possibilitat està una eventual pèrdua de confiança en el judici i la subjectivitat humana. "Hemos devaluado la singular capacidad humana de ver las cosas en sus dimensiones psíquica, moral y emocional, y la hemos sustituido por la fe en el cálculo técnico", ha criticat (p.161).

John McCarthy, el matemàtic que explicàvem que va inventar el concepte intel·ligència artificial, va afirmar en una ocasió que "incluso de máquinas tan simples como los termostatos se puede afirmar que tienen creencias". Searle (el filòsof de l'experiment de la caixa xinesa) va preguntar a McCarthy quines creences creia que podia tenir el termòstat, i va respondre el següent: "Mi termostato tiene tres creencias: hace demasiado calor, hace demasiado frío, y la temperatura es correcta". És Neil Postman (2018) qui explica aquesta anècdota i a seguidament exposa una anàlisi necessària d'aquest raonament:

(McCarthy) redefine el significado de la palabra 'creencia', rechaza la idea de que los seres humanos tenemos estados mentales internos en los que se basan las creencias, y argumenta en cambio que creencia solo significa lo que alguien o algo hace. (...) también implica que simular una idea es sinónimo de

reproducirla. Y, más importante, desestima la idea de que la mente es un fenómeno biológico.

Lo que tenemos aquí es una metáfora que ha enloquecido. Primero se plantea (...) que los seres humanos (...) guardan semejanza con las máquinas; después que (...) son poco más que máquinas, y por último, que los seres humanos son máquinas. E, inevitablemente, la metáfora culmina, como esbozaban las palabras de McCarthy, en la idea de que las máquinas son seres humanos. De ahí se colige que pueden fabricarse máquinas que reproduzcan la inteligencia humana y que por tanto la investigación en el campo de estudios conocido como inteligencia artificial era inevitable. (p.153)

Postman alertava del reduccionisme que representa aquesta afirmació, perquè la intel·ligència humana, com hem mirat d'anar raonant, no es pot generar amb tecnologia. Això és a causa del fet que les màquines no poden experimentar els sentiments que estan en la base del comportament humà, i que són provocats per les percepcions que, com hem exposat, advertim a partir de la relació amb el món que ens envolta. El sociòleg nord americà avisava també sobre el perill d'atribuir voluntat, intencions i motius a una màquina, ja que pot implicar, com avançàvem més amunt, que les persones quedin exemptes de responsabilitat per les decisions que prenguin els ordinadors. Postman convidava a parar atenció, especialment, a aquells individus que se'n poden beneficiar: "Naturalmente, es de esperar que los burócratas abracen una tecnología que ayuda a crear la ilusión de que las decisiones no están bajo su control" (Postman, 2018, p.157). James Bridle (2020) descriu com es produeix el procés d'assimilació humà de la lògica computacional:

La computación, a cualquier escala, es una argucia cognitiva, que descarga en la máquina tanto el proceso de toma de decisiones como la responsabilidad. Cuanto más se acelera la vida, más interviene la máquina para hacerse cargo de un número mayor de tareas cognitivas lo cual refuerza su autoridad con independencia de cuáles sean las consecuencias. Reajustamos nuestra visión del mundo para adaptarnos mejor a las alertas y atajos cognitivos constantes que nos proporcionan los sistemas automatizados. La computación sustituye al pensamiento consciente. Pensamos cada vez más como máquinas o dejamos por completo de pensar. (p.53)

La màxima deshumanització s'assoliria amb aquest pensament maquinal humà i quan no es revelessin ni s'exigís que es comunicuessin les raons dels judicis fets per màquines contra les persones. Distòpic? un recent estudi conclou de fet que avui acceptem millor les decisions esbiaixades si estan determinades per una màquina que per una persona (Bigman et al., 2020).

Postman (2018) recorda que, per defensar els nostres drets humans, cal preguntar-nos qui ha introduït les dades, amb quin propòsit, basat en quins supòsits i a qui beneficia.

La degradació d'allò humà que la pressió del mercat pot causar es pot presentar de moltes maneres en diverses situacions que poden ser perfectament actuals, per exemple en els serveis d'atenció al client automatitzats, en els algorismes que intervenen en decisions sobre la vida de les persones o, més recentment, en els usos dels robots assistencials que acompanyen la gent gran. A Barcelona ja fa més d'un any que s'estan provant robots de companyia per a aquest col·lectiu de persones que viuen soles a la ciutat (Arnau, 2021). També s'ha conegut el cas del robot que dona suport en l'atenció relacional dirigida a les internes als geriàtrics de Catalunya (Cruanyes, 2022). En aquests centres, on existeix una manca de recursos i de supervisió per part de les administracions àmpliament denunciada (*Coordinadora familiars residencias 5+1*, 2022) no seria cap bestiesa pensar que aquests dispositius, que poden treballar les 24 hores seguides, se'ls acabés delegant en excés l'atenció a les persones grans, d'altra banda molt necessitades de companyia i sovint vulnerables. Serien converses reals? O es corre el risc de fer creure a aquestes persones, moltes d'elles amb les facultats mentals debilitades, que estan sent veritablement acompanyades? Margaret Boden (2017) alerta sobre aquest risc:

Muchas veces, incluso en terapias emocionalmente tensas, lo que la persona desea por encima de todo es que se le reconozca su valentía y/o su sufrimiento, pero que surja de un entendimiento mutuo de la condición humana. Estamos estafando a la persona a ofrecerle solo un simulacro superficial de comprensión.
(p.158)

La paraula anglesa *care* pot traduir-se per assistir, i també per *preocupar-se per algú*. Boden es serveix d'aquest mot per emfatitzar que els robots no es preocupen per les persones, sinó que els hi són indiferents (García Aller, 2020).

Un altre exemple d'escenari de degradació és el que es produeix amb els usos de la tecnologia per comercialitzar parelles sexuals artificials que poden parlar i moure's. Aquests usos, explica Boden, reforcen la cosificació sexual de la dona i traspuen un concepte de les relacions íntimes molt superficial que "se presta a la lujuria, la obsesión sexual y una mera familiaridad cómoda" (p.139).

2.2.6 Racionalitat humanitzada

Malauradament, com dèiem, els fets semblen donar la raó als corrents més escèptics respecte a la gestió del progrés. Segons Oxfam Intermón (2022), les 10 persones més riques del món han duplicat la seva fortuna, mentre que els ingressos del 99 % de la població mundial han minvat. La ONG denuncia que aquest fet és fruit de decisions polítiques premeditades que afavoreixen els més rics i poderosos. Aquesta *violència econòmica* contribueix a la mort de, mínim, una persona cada 4 segons i perjudica especialment les més pobres, dones, nenes i les persones racialitzades. Finalment assenyala que la solució passa per reformar el model econòmic i per solucionar el gran problema que fractura les societats i impedeix viure amb dignitat a gran part de la societat global en favor d'una petita minoria.

En aquesta nova revolució que en l'àmbit tecnològic, protagonitza la intel·ligència artificial, cal reivindicar el pensament de filòsofes com Ana Carrasco Conde, qui argumenta que la importància de les humanitats resideix en “qué hacemos con todo ese conocimiento (...) nos dan sentido, nos dan valor, implican también a qué proyecto común queremos llegar (...) y sobre todo (...) generan cultura (...) el tejido que hace que nos sintamos parte de un todo” (Barceló, 2020). O el de Duch i Chillón (2012) que defensen que aquesta necessitat de sentit no és incompatible amb la innovació tecnològica, perquè en cap cas implica menysprear els avantatges:

El desolado y desolador «hombre unidimensional» u Homo economicus denunciado por Herbert Marcuse, Max Horkheimer, Lewis Mumford o Günther Anders puede sin duda beneficiarse de los adelantos que los nuevos ingenios procuran, pero también hallar en ellos una fuente de alienación renovada. No se trata de menospreciarlos, por supuesto, sino de vindicar un «ejercicio de la humanidad» rebelde ante la deriva en curso, que aprenda a utilizar la potencia tecnológica en aras de los fines éticos, políticos, religiosos y culturales que integran nuestro mejor legado. (p.468)

La teoria crítica proposa una redefinició de la *racionalitat* tal com és concebuda avui pel pensament científista i dotar-la del veritable sentit de la paraula, vist que de cap manera pot considerar-se racional una organització social que fa prevaldre els interessos d'una petita minoria de persones riques front als de la majoria de la humanitat (Piketty, 2019), que utilitza l'ésser humà i la natura i els considera material a ser dominat i explotat. Marcuse proposa posar la tecnologia a disposició d'una racionalitat que consistiria en la utilització planificada dels

recursos per a la satisfacció de les necessitats vitals amb el mínim esforç per a una major disposició de temps lliure (Marcuse, 1993):

Y sin embargo, los hechos que dan validez a la teoría crítica de esta sociedad y su fatal desenvolvimiento están perfectamente presentes: la irracionalidad creciente de la totalidad, la necesidad de expansión agresiva, la constante amenaza de guerra, la explotación intensificada, la deshumanización. Y todos ellos apuntan hacia la alternativa histórica: la utilización planificada de los recursos para la satisfacción de las necesidades vitales con un mínimo de esfuerzo, la transformación del ocio en tiempo libre, la pacificación de la lucha por la existencia. (p.281)

En una adaptació recent de les teories de Marcuse, Feenberg reclama la necessitat d'un lideratge cultural de l'evolució tecnològica que descarti la vista romàntica al passat i que per contra sigui capaç d'aplicar una perspectiva crítica solvent que pugui identificar quins usos concrets de la tecnologia cal revisar amb els codis actuals. Una mirada competent i veritablement racional que reconegui els límits naturals i humans del desenvolupament tecnològic. Una direcció conscient de la fragilitat de l'entorn tècnic al servei del capitalisme, incapaç d'adaptar-se als problemes ecològics i socials que infligeix que és encara més fràgil com més poderosa és la indústria, perquè causa més danys en un procés insà de desenvolupament. Un guiatge, per tant, que consideri el respecte per la natura com a marc d'acció: "Nature as a context of development is not a final purpose but a dialectical limitation that invites transcendence through adaptation" (p.190). Per descomptat, en cap cas es reclama una tornada al passat o a l'abolició de la tecnologia. Des de la teoria crítica s'aposta per una anàlisi crítica sobre els riscos dels usos actuals i una proposta d'acció guiada pels valors humans. Ens podem remetre als postulats de Feenberg en aquest sentit:

We cannot recover what reification has lost by regressing to pre-technological conditions, to some prior unity irrelevant to the contemporary world. The solution is neither a romantic return to the primitive, qualitative, and natural, nor a speculative leap into a "new age" and a whole "new technology." On the contrary, the critical concept of totality aids in identifying the contingency of the existing technological system, the points at which it can be invested with new values and bent to new purposes. Those points are to be found where the fragmentation of the established system maintains an alienated power. (...)

Adaptation maintains the formal character of the modern concept of freedom and therefore does not reduce individuals to mere functions of society. Freedom lies in this lack of determinacy. (pp.189-190)

El camí de l'apropiació social de la tecnologia és descrit amb aquestes paraules per James Bridle (2020): “Nuestras tecnologías son extensiones de nosotros mismos, están codificadas en máquinas e infraestructuras, en marcos de conocimiento y acción; si se piensan de verdad, ofrecen un modelo de un mundo más auténtico” (p.25).

Posar l'ésser humà en el centre implica valorar abans si l'aplicació de les tecnologies exponencials seran beneficioses per a les persones i per al conjunt de la societat. L'expert en IA Gert Leonhard (2018), avisa que el desenvolupament de la IA avança a velocitats tan exponencials que és probable que no siguem capaços d'imaginar com serà un futur en què l'automatització agafarà ràpida i progressivament més protagonisme. En previsió d'aquest escenari, proposa un tractat sobre els *drets humans exponencials* per garantir que prevalgui la humanitat. En un sentit més pràctic, proporciona una sèrie de preguntes que cal fer-se per humanitzar l'adopció de la tecnologia:

- Aquesta tecnologia, minvarà la humanitat, ja sigui de manera inadvertida o conscient? Cal evitar tecnologia que pretengui substituir habilitats úniques de les persones humanes que no haurien de ser mediatitzades tecnològicament.
- Fomentarà una veritable felicitat humana? Fer servir tecnologia si aquesta ajudarà a millorar el benestar de les persones en un sentit profund i per tant fugint de plaers superficials a curt termini.
- Podria tenir efectes indesitjables i potencialment desastrosos? No hauria d'apropriar-se de la nostra autoritat i en canvi ens hauria d'empoderar.
- S'atorgarà un excés d'autoritat a sí mateixa, a altres algoritmes, bots i màquines? Evitar tecnologies que impulsin a delegar el propi pensament i tenir clar que està per servir-nos i donar-nos valor, no al contrari.
- Ens permetrà transcendir-la o ens farà dependents d'ella? Evitar subordinar-se a la tecnologia i haver de seguir cegament les seves decisions.
- Necessitaran els éssers humans transformacions o augments materials per realment poder fer servir aquesta tecnologia? Evitar l'obligació a augmentar-nos, a actualitzar-nos per poder accedir a serveis bàsics, treball, educació, serveis públics, etcètera.
- Estarà disponible o serà exclusiva? Cal que la tecnologia sigui democràtica i no elitista per no perpetuar desigualtats.

Com a complement de les seves preguntes, cita un meme del teòric expert en cultura digital Douglas Rushkoff en el qual explica, en cinc punts, com incorporar tecnologia des d'un punt de vista humà:

- Col·locar per sobre de tot la nostra prosperitat humana col·lectiva.
- Permetre que els *androritmes*, les qualitats pròpiament humanes com la imaginació, la casualitat, els errors i les ineficiències, segueixin sent importants, inclús si no són desitjables per a la tecnologia o incompatibles amb ella.
- Lluitar contra la propagació del pensament de les màquines, és a dir, no canviar allò que ens importa i que necessitem com a éssers humans, pel fet que afavoriria les tecnologies que ens rodegen.
- No caure en la temptació de preferir la màgia tecnològica, és a dir, les grans simulacions de realitat, per sobre de la nostra pròpia realitat, així com no tornar-nos addictes a la tecnologia.
- No preferir les relacions amb pantalles i màquines per sobre de les que podríem tenir amb els nostres proïsmes humans.

2.3 LA NECESSITAT D'UNA APROPIACIÓ SOCIAL DE LA TECNOLOGIA

Sin embargo, aunque los resultados efectivos del nuevo industrialismo eran los de incrementar las cargas de los trabajadores ordinarios, la ideología que lo fomentaba iba dirigida hacia su liberación.

Lewis Mumford

A la ideología capitalista del beneficio máximo se le han agregado las posibilidades de la opacidad tecnológica, que permite envolver la pura codicia en la lógica inhumana de la máquina.

James Bridle

2.3.1 Està passant: conseqüències de l'excés de delegació

Per tal de poder valorar quin ha de ser el paper de la professió periodística davant la creixent adopció dels sistemes automatitzats en l'organització de les societats, pensem que val la pena exposar alguns escenaris actuals i quotidians en els quals els algoritmes s'estan aplicant per donar suport a la presa de decisions. Aquests sistemes tenen el potencial de ser de gran ajuda per optimitzar processos, i poden resultar especialment beneficiosos si aquestes millores repercuteixen en millores per a la ciutadania, per exemple en l'agilització de tràmits burocràtics amb l'Administració Pública. No obstant, pot ser perillosa la combinació entre l'optimització que promet el processament computacional de dades, amb els valors de la productivitat —regits per l'estalvi de recursos en el sector públic i en la idea del mínim esforç pel màxim benefici en les empreses—. Aquesta simbiosi implica riscos que al periodisme pot convenir tenir en compte, tant per fer servir la IA de manera responsable com per avaluar de manera rigorosa els usos que poden fer-ne altres actors.

Avui ja resulta habitual preguntar a la IA si una persona pot tenir dret a un ajut social o qui pot ser la millor candidata a una oferta de feina. També es fa servir, per exemple, per controlar la productivitat dels treballadors durant la jornada laboral (Bridle, 2020). Una mena de decisions i processos que fins fa poc estaven en mans de persones. Aquest canvi ha estat batejat com la *segona era de les màquines*, a causa de la idea que els ordinadors fan pel treball mental el que les màquines feien pel treball físic a la primera era de les màquines (Brynjolfsson i McAfee, 2016).

Aquests dies, a Espanya, es fan servir algoritmes per decidir qüestions com qui pot ser susceptible de rebre protecció policial per ser víctima de violència de gènere o qui pot ser beneficiari de l'atur (*The OASI Register – The Observatory of Algorithms with Social Impact – OASI*, 2021). Entre d'altres, també és conegut l'algoritme que fa servir el Govern espanyol per decidir qui pot tenir dret a rebre el bo social que serveix per ajudar a pagar la factura de la llum.

Els models algorítmics s'alimenten amb bases de dades, i per tal que un sistema doni una solució precisa, serà necessari que disposi de la informació necessària per poder fer el procés de computació de manera òptima, o el que és el mateix, per poder determinar la solució a la pregunta que se li formula. Un dels problemes que s'han identificat en aquests usos és que s'estan fent servir tecnologies basades en càlcul probabilístic per prendre decisions que afecten

a persones de manera individual, cosa que comporta el risc d'error en la determinació i per tant implica la possibilitat de cometre injustícies.

Cada cop tenim més dades i més poder de computació, en conseqüència molts aspectes de la vida són cada cop més previsibles. El problema és que el criteri per fer servir aquestes aplicacions predictives, dèiem, és una aproximació matemàtica, i en molts casos algú determina que n'hi ha prou amb què la probabilitat d'encert superi el 70%, el 80% o el 90%. Per exemple, en el cas de l'algoritme predictiu que feia servir la policia de Nova York, el potencial d'èxit era del 72%. Segons explica la periodista Marta García Aller (2020), l'eina va ser desenvolupada per una empresa en col·laboració amb la Universitat de Columbia, i es va alimentar amb dades de denúncies, l'hemeroteca de la secció de successos del *New York Times*, el nivell adquisitiu de les diferents zones, entre d'altres. A partir de criteris com aquests, el sistema indicava que el lapse horari en què era més probable que es cometessin crims era entre les 15 i les 19h de la tarda; el dia més probable, els divendres; o que a la nit baixaven els crims a Manhattan. Tota l'estadística que es va extreure va fer possible reorganitzar de manera eficient els recursos departament. Tanmateix, eines d'aquesta mena també es fan servir per atorgar índex de perillositat a les persones, que seguidament es converteixen en sospitoses. Diu, amb raó, García Aller: "los más optimistas pensarán que qué más da si nos ayuda a vivir más seguros. Cambiarían de idea si pasaran a ser sospechosos sin saber por qué" (p. 242).

Un exemple d'aquesta mena de decisions algorítmiques esbiaixades va ser el cas de *Compas*, un sistema que fa servir el sistema judicial dels Estats Units per predir el potencial criminal de les persones. Aquest es basa en variables com l'estrat social, raça, familiars, persones properes que hagin estat detingudes abans... el cas és que aquest sistema no pot saber realment si una persona cometrà un delictes, sinó que simplement, reiterem, pot fer operacions de propensió.

Compas és una de les aplicacions d'IA que s'ha descobert que era discriminatòria perquè penalitzava persones negres (Angwin, 2016). Això passava perquè es basava en dades del passat que reflectien que els ciutadans d'aquesta raça havien estat més condemnats. Tal com apunta Daniel Innerarity (2022) "Se trata de un sesgo racista que nada tiene que ver con la persona concreta, pero que tiene efectos decisivos sobre ella".

Un cas més proper que il·lustra aquest problema és el de l'algoritme que hem mencionat més amunt sobre el sistema que fa servir el Govern espanyol per decidir qui pot ser beneficiària del bo social. Aquest es va descobrir que exclouïa de l'ajut a persones que complien els requisits per rebre'l (Civio, 2019). L'organització que ho va posar en evidència i que ho va denunciar va

demanar conèixer el procés que havia efectuat la màquina per arribar a les determinacions sobre qui tenia dret al bo, però el Govern es va negar a explicar-lo (Civio, 2022b). Avui, i havent passat uns anys des de la primera denúncia, el cas encara no s'ha resolt (Civio, 2022a).

Innerarity defensa que les prediccions en base al passat estan limitades perquè no contemplen la capacitat humana de “desafiar lo que es imposible y decepcionar lo que es esperable”. El filòsof recorda també que hi ha força fenòmens socials i de rebel·lió que irrompen que no és possible predir. “La predicción algorítmica será útil cuando se entiendan sus límites, de su dependencia contextual y su necesidad de interpretación”, sentència Innerarity.

En efecte, hi ha moltes qüestions que, al menys avui, no es poden predir, García Aller (2020) detalla alguns casos, com per exemple els 13 resultats d'una travessa; les emergències que es poden donar durant les cures a persones convalents (que poden implicar molts imprevistos); sovint tampoc es pot predir l'èxit (que de vegades es produeix en situacions inversemblants); o l'estúpidesa, perquè no té lògica; els homicidis involuntaris; molts fenòmens meteorològics, (perquè la natura també ho és); les erupcions dels volcans (en especial els que fa més temps que estan inactius). En el cas dels terratrèmols també pot ser problemàtic. A més, en el context d'emergència climàtica, moltes variables estan canviant, i això també seguirà complicant les prediccions meteorològiques en el futur. Santiago González, un dels enginyers que va dissenyar un dels sistemes que ha arribat a utilitzar la policia de Nova York assenyalava un altre tipus de feblesa similar (relacionada amb les modificacions de les dades disponibles). A González li preocupava que si gràcies al seu sistema la policia aconseguia aturar el delictes un dia i hora determinat, canviaria el futur perquè s'estaria evitant que succeís, així que el cúmul d'encerts faria que el model s'invalidés a sí mateix (García Aller, 2020).

La qüestió sobre els límits predictius d'aquestes tecnologies també ha estat plantejada des de l'àmbit de la filosofia. Per a Nietzsche, la vida no era “la previsible repetición del pasado, sino la previsible irrupción del futuro”. Això precisament significava la llibertat per al filòsof alemany, explica el també filòsof Dardo Scavino: “esta imprevisible irrupción del futuro en un presente previsible” (Scavino, 2022, p.198).

Si el passat es troba a la memòria i el present a la nostra percepció, a quina facultat atribuir el futur? Per trobar la resposta sobre el significat d'orientar-se en el pensament, el filòsof Scavino recupera el pensament de Kuhn, qui defineix el futur com un esdeveniment sobtat i no estructurat “relámpago de una intuición que se encuentra en el origen de cada nuevo paradigma

científico”. Segons Scavino “hay humanidad porque algunos dejan de reproducir como autómatas las instrucciones en vigor y proponen otras nuevas” (p.198).

L’acte de creació artística que descriu Gadamer a *Estética y hermenéutica* resulta una de les possibles expressions d’aquestes idees en l’activitat humana. Així mateix, el pensament d’aquest autor pot contribuir a la reflexió crítica sobre les suposades habilitats artístiques dels sistemes generatius que mencionàvem en apartats anteriors a través de la comparació amb la concepció de la creativitat que proposa (H.-G. Gadamer, 2018):

Lo asombroso del impulso artístico no es precisamente su carácter impulsivo, sino el aliento de libertad inherente a sus formaciones. En las imágenes humanas, el momento decisivo de la habilidad artística no consiste tampoco en que se realice ahí algo de una excelente utilidad o de belleza superflua, sino que en el producir humano puede proponerse tareas así de diversas y procede según planes a los que distingue un momento de libre arbitrariedad. El arte comienza justamente allí donde se puede hacer algo también de un modo diferente. (...) lo decisivo no es la realización de algo que se haya hecho, sino que lo que se ha hecho es de una peculiaridad muy especial (algo que se ha realizado de un modo irreplicable y que ha resultado un fenómeno único). (pp. 131-132)

Byung-Chul Han afirma que “la IA aprende del pasado. El futuro que calcula no es un futuro en el sentido propio de la palabra (...) No sale de lo antes dado hacia lo intransitado. El pensamiento en sentido enfático engendra un *mundo nuevo*” (Han, 2021, p.59).

Per tant, quan la IA amb les seves prediccions en base a probabilitats que calcula fent servir dades de comportaments passats, redueix l’home a un estereotip, demostra de la seva incapacitat de considerar la natura *epifànica* de l’èsser humà en paraules de Lluís Duch (2001), una natura que implica que s’està fent a cada moment en un procés hermenèutic al temps que es manifesta contra l’èsser determinat i tancat, propi de la raó pràctica, en actes que són propis de la seva humanitat com el descrit per Gadamer unes línies més amunt.

És per això que, amb Vidal i Garde (2022), hem reclamat un *alteritme* que promogui un llenguatge el silenci i ens permeti avançar, en obertura, davant la incertesa i l’ambigüitat, i que es manifesta en l’experiència de d’alteritat. Una aproximació que contempli la veritable condició humana, per evitar que allò que és probabilístic —que es genera amb les dades del passat—, conformi i determini allò que queda per emergir i impedeixi el seu possible creixement impredecible.

En l'àmbit de la comunicació, els algoritmes s'han convertit en productors de cultura (Natale i Guzman, 2022) i per tant constructors de la realitat, capaços de condicionar comportaments i de donar forma i significat (Finn, 2018; Just, N i Latzer, 2016; Winner, 1980). Quan en els algoritmes deleguem tasques que es poden resoldre amb les matemàtiques —com un extracte bancari— no existeix cap mena de dubte sobre la confiança que hi podem dipositar, perquè, com una calculadora, computen amb la màxima fiabilitat. Tanmateix, és molt diferent per exemple en els casos dels algoritmes que ens filtren i jerarquitzen la informació a la nostra esfera pública, o els que es fan servir per trobar parella: aquests no són còmputos de solució correcta perquè les variables que es fan servir no són clarament quantificables com els nombres (Tufekci, 2015).

Arran d'una anècdota en un context de consum cultural (cada cop més condicionat pels algoritmes de personalització basats en la predicció), l'assagista Antonio García Maldonado (2021), fa notar que, tot i la capacitat de recomanació de les IA de Spotify i similars, quan sona una de les seves peces musicals preferides a la ràdio experimenta un "entusiasmo" i una "alegría" que no sent quan escolta els mateixos temes a les seves llistes de reproducció. Maldonado es pregunta: "¿A qué obedece el cambio en la recepción del mismo producto? ¿A la sorpresa? ¿Al valor añadido de lo imprevisible? ¿A la elección ajena? ¿A que nos gusta pensar que el azar acierta con nosotros?". La resposta que ell mateix dóna és que el fet de ser humans implica que ens agradin les sorpreses i els fets que ens treuen de l'autoreferencialitat i ens connecten amb alguna cosa més, encara que sigui la sort.

Quant a la fiabilitat de les prediccions algorítmiques, Yuval Noah Harari (2018) augura que la seva imperfecció no impedirà que cada cop deleguem més, fins el punt d'assolir una gran dependència de la IA. Tot i que també considera que aquests algoritmes predictius estan lluny de poder encertar en base a les dades del passat, creu que això no impedirà que cada cop ens recolzem més en ells, perquè en realitat no cal que siguin perfectes, sinó que serà suficient amb què encertin més que les persones humanes i, diu, això no és difícil.

La raó en què es basa l'autor per realitzar aquestes afirmacions és que els humans ens equivoquem molt, i a més ho fem en algunes de les decisions més importants de les nostres vides (com ara triar parella o estudis). Harari atribueix aquestes limitacions humanes a la manca de criteri, i parla concretament sobre l'existència, en les persones, d'unes "dades defectuoses de programació (genètica i cultural), de definicions confuses i del caos de la vida" (p.74). És a causa d'aquestes mancances que Harari pronostica que pot arribar el punt en què acabem

confiant més en els algoritmes que en nosaltres mateixes a l'hora de decidir, en una espiral en què la IA sigui cada cop més fiable al temps que les emocions humanes ho siguin menys.

Carme Torras Genís, directora del grup de recerca en robòtica assistencial a l'Institut de Robòtica (CSIC-UPC) creu en el potencial de la IA per fer equip amb les persones humanes, i explica que una de les fortaleses humanes davant la IA té a veure precisament amb suplir aquestes limitacions predictives de la IA. Així, una de les habilitats humanes en alça serà la resolució de problemes i la capacitat de resposta davant d'emergències per a les que la màquina no disposa d'estadística.

A banda d'aquest aspecte, en ser preguntada per quines són la resta de potencialitats de l'ésser humà front la tecnologia exponencial, apunta que són les que requereixen creativitat, les que necessiten empatia i les basades en relacions humanes en general. Torras, que du a terme la seva recerca en el desenvolupament de robots assistencials, destaca la necessitat de tracte humà en l'assistència social per ajudar a resoldre de manera adient els problemes que pateixen les persones (Torras, 2020).

Aquesta creixent automatització de tasques està fent agafar força a la concepció d'una tecnologia amb certa autonomia i agència pròpia, d'acord amb la tradició de la teoria de la mediació de la filosofia de la tècnica. Segons aquest corrent de pensament, la moralitat no és només un assumpte humà, sinó que ho és també de les coses materials, i el funcionament dels artefactes té component moral precisament perquè aquests condicionen l'acció dels éssers humans, els seus valors i per tant la seva relació amb el món. Així, si l'acció de les màquines implica càrrega ètica, la tecnologia no pot ser neutra i, en conseqüència, en tant que la mediació dels artefactes té la capacitat de transformar les percepcions i accions humanes, caldrà reconèixer-li aquesta agència (Latour, 1992; Verbeek, 2011; Winner, 1980).

Els plantejaments de la teoria de la mediació adopten més sentit com més sofisticats són aquests sistemes, perquè més difícil resulta identificar quan les accions són veritablement pròpies de l'ésser humà (Savolainen i Ruckenstein, 2022). Aquesta incapacitat de distingir l'agència es produeix també a causa de la complexitat dels sistemes algorítmics, i és que en molts casos no és possible entendre com han arribat a una determinada decisió i resulten incomprensibles fins i tot per als qui els han dissenyat.

Com detallarem al quart capítol, des del periodisme haurem d'estar atentes a aquesta utilització d'algoritmes per a la presa de decisions amb fort impacte a la vida de les persones, especialment

en àmbits com ara l'assistència social, l'educació, la justícia o els recursos humans. Des del punt de vista ètic, la responsabilitat moral s'associa amb d'altres com voluntat conscient, llibertat i consciència, conceptes que al seu torn contribueixen a la nostra noció d'humanitat (Boden, 2017). S'accepti o no que les màquines disposen de certa agència, a efectes legals, dèiem, només les persones poden assumir responsabilitat per les conseqüències d'un sistema.

Per fer les valoracions pertinents, doncs, ens haurem de preguntar coses com: amb quines dades s'han alimentat els algoritmes? Quins criteris s'han considerat per processar aquestes dades i arribar a una conclusió? Considerarem una veritat incontestable, aquest resultat algorítmic? També haurem de parar atenció a com aquesta tecnologia es pot utilitzar per monopolitzar el poder o per mantenir relacions desiguals. A continuació abordarem aquest tema.

2.3.2 Algoritmes com a eines de poder i perpetuació de les desigualtats

Amb la finalitat de seguir fent-nos la idea sobre com des del periodisme podem exercir un rol en favor de la democràcia i els Drets Humans en relació als usos de la IA, també valdrà la pena observar com aquesta tecnologia es pot fer servir per a la perpetuació del poder o, en ocasions, per sostenir lògiques de desigualtat social basades en abusos cap als més febles.

Començàvem aquest relat exposant la necessitat de preguntar-nos al servei de quins valors posem la tecnologia i proposant un ús de la IA que contribueixi a la realització de les persones. Després hem vist que el progrés encara està lluny d'haver fet una administració racional dels recursos i dels beneficis de la innovació i la millora dels processos que ha fet possible l'automatització. Així doncs, per què no hem pogut posar la tecnologia al servei dels valors humans? Aldous Huxley argumentava que aquest problema es produeix a causa de l'ambició d'alguns poderosos unida a la capacitat de la tecnologia per facilitar la imposició del poder de pocs a la majoria. Així ho explicava en una entrevista (Mike Wallace, 1958):

“All the instruments obtaining power and the passion for the power is one of the most moving passions that exist in man. Democracies are based on the proposition that power is very dangerous. It is important not to let any one man or a small group have much power too long time. All this new device are extremely efficient instruments for the imposition of power by small groups of larger masses”

Han passat més de sis dècades i la realitat sembla confirmar les prediccions de l'intel·lectual nord americà, alhora que contrasta amb les optimistes expectatives de l'evolució tecnològica que han conservat altres reconeguts pensadors al llarg de la història. Per a alguns, el camí al futur significaria la progressiva alliberació dels humans de l'esclavitud del treball gràcies al progrés de la ciència i la tècnica. Descartes pensava que la redempció de la humanitat provindria de l'enginy tècnic i científic i el robot apareixia en aquesta història com el redemptor per excel·lència (Scavino, 2022). També Marcuse —com hem mencionat abans— reclamava una tecnologia d'alliberament producte de la imaginació, ja que els sistemes totalitaris determinen no sol els usos de la tecnologia, sinó que limiten les seves possibilitats per al bé comú. Una tecnologia que serviria per construir una societat industrial avançada en què es satisfarien més necessitats per a més persones “Es racional la imaginación que puede llegar a ser el a priori de

la reconstrucción y nueva orientación del aparato productivo hacia una existencia pacífica, una vida sin temor” (p.279).

Avui, dèiem, les societats capitalistes resulten la definició del progrés mal gestionat: l'estalvi que ha anat proporcionant l'automatització dels processos sol repercutir en beneficis per a uns pocs privilegiats, mentre que es socialitza en la majoria les conseqüències mediambientals de la sobreexplotació dels recursos naturals per sostenir el sistema productiu (Piketty, 2021; *Las desigualdades matan*, 2022).

A Espanya cada cop calen més anys per pagar l'entrada d'una hipoteca (Villalvilla, 2022), malgrat la gran majoria de la població treballadora —que té la sort de tenir feina en un país amb altes xifres d'atur— continua fent jornades intenses. A això es suma que els avenços tecnològics aplicats al món laboral tendeixen a provocar l'efecte d'haver de ser més productiu en lloc de suposar una descàrrega de pressió, i això es deu al fet que cada cop més empreses decideixen aplicar la tecnologia per optimitzar la productivitat (connexió permanent) i per fer monitorització constant de l'activitat de les persones treballadores (Kantor i Sundaram, 2022). En aquest context, la majoria acusa les conseqüències d'un sistema que impossibilita en molts casos l'accés a béns fonamentals com un habitatge digne, temps disponible per la vida familiar o per dur a terme projectes personals que habiliten la persona a realitzar-se com a ésser humà. En suma, les condicions per exercir la pròpia llibertat, allò que hauria de ser el veritable capital dels països, en comptes del que resulta ser ara, malgrat el *progrés* i que es tradueix l'acumulació d'aquesta riquesa en poques mans.

Si fins ara la tecnologia no ha servit per fer una redistribució més equitativa dels beneficis, per què en el cas de la IA hauria de ser diferent? En moments en què les nostres vides personals i laborals es desenvolupen de manera creixent en entorns digitals, pot convenir parar atenció als qui controlen aquests canals i els seus interessos, perquè existeix la possibilitat que els objectius d'aquestes corporacions no coincideixin amb els de la ciutadania (Morozov, 2018).

La dependència generalitzada d'internet, unida al fet que unes poques companyies de caire privat —les grans tecnològiques— disposin dels serveis imprescindibles, apunta a una gran responsabilitat d'aquestes organitzacions i, al mateix temps, a l'existència d'una forta concentració de poder en poques mans que han demostrat que responen principalment a interessos de mercat i per tant no prioritzen valors com la justícia o l'equitat. Al llarg d'aquest treball els anirem exposant. Per ara, però, ens allunyarem del cas de les tecnològiques i, en

aquest segon capítol, seguirem tractant el paper i l'aplicació de les tecnologies exponencials des d'una perspectiva general.

Com dèiem en apartats anteriors, vivim uns moments en els quals el debat ètic en torn a la IA està cada cop més present, tot i que encara no està prou madur, perquè la innovació tecnològica va més ràpid que la reflexió prèvia necessària sobre l'impacte dels seus usos des d'un punt de vista ètic (i molt menys, jurídic). Si bé és cert que la UE ha començat a legislar en benefici d'una IA fiable (European Commission, 2021), el més normal fins ara és que s'hagin posat els productes al mercat abans dels debats ètics pertinents, amb els possibles problemes ja ocasionats (Torras, 2020). Per tant, no és estrany que la IA es dissenyi i/o es faci servir de manera irresponsable o amb ànim de perpetuar dinàmiques de poder i desigualtats que, en el millor dels casos, s'aconsegueixen aturar gràcies a regulacions a posteriori o a la pressió social.

Com a conseqüència d'aquesta pressió, per exemple, Microsoft ha decidit deixar d'incorporar tecnologia de reconeixement facial en els seus productes (Hill, 2022). Activistes i acadèmiques feia anys que advertien sobre els problemes que es poden derivar de l'ús d'aquesta mena d'IA capaç d'identificar l'edat, el gènere i l'estat emocional de les persones. Riscos que tenen a veure amb el potencial invasiu, de generar biaixos o altres mancances de rigor en les determinacions algorítmiques. Microsoft també ha desenvolupat una guia ètica en la qual es compromet a no desenvolupar tecnologia que pugui ser nociva per a la societat. Entre les directrius, figura la necessitat d'assegurar que els sistemes proporcionin solucions vàlides per als problemes que volen resoldre o assegurar la qualitat dels seus serveis a tots els grups socials per no contribuir a generar desigualtats.

Algunes mostres sobre com la tecnologia s'està utilitzant avui per a finalitats dubtoses poden ser els que relata la publicació *MIT Technology Review* que, en una sèrie de peces d'investigació, posa de manifest diferents escenaris en els que avui la IA s'utilitza de maneres perverses fins el punt que, s'assegura, estaria col·laborant en la creació d'un nou ordre mundial colonial (Hao, 2022). En un seguit de reportatges fets arreu del planeta, s'investiga com grups que volen mantenir el seu estatus de dominació fan servir la IA per enriquir-se a compte de les comunitats que ja havien estat desposseïdes abans, durant l'època colonial. Explica casos com el de Sudàfrica, on eines de vigilància amb IA, basades en aquestes tecnologies de reconeixement facial i la monitorització de comportaments de persones, consoliden jerarquies i alimenten un apartheid digital (Hao i Swart, 2022). O, entre d'altres, el cas d'organitzacions veneçolanes que

fan servir IA per identificar possibles treballadors desesperats i així aconseguir mà d'obra barata, creant un nou model d'explotació laboral (Hao i Hernández, 2022).

L'activista Timnit Gebru avisa que de res servirà parlar molt de justícia aplicada a la IA si les institucions no són permeables a aquesta mena de reflexions i a adaptar els seus procediments. L'enginyera insta a preguntar-nos com ho faríem si haguéssim de construir aquestes institucions des de zero (Walsh, 2022):

At the end of the day, this needs to be about institutional and structural change. If we had the opportunity to pursue this work from scratch, how would we want to build these institutions? (...) If we want AI that benefits our communities, then what kind of processes should we follow?

L'empresa Amazon, una de les més valuoses del món i propietària del major servei d'emmagatzematge al núvol, resulta un cas paradigmàtic sobre el poder de la tecnologia dependent del sistema de valors amb què s'utilitza. L'empresa de Jeff Bezos és un exemple sobre les noves formes com aquests usos tecnològics poden arribar a influir en la precarització laboral i la desqualificació de la persona treballadora fins el punt de sotmetre-la al control d'algoritmes aplicats a la gestió de processos durant la jornada laboral (Bridle, 2020).

En el seu llibre *La nueva edad oscura* l'autor explica que determinats magatzems d'aquesta empresa estan gestionats per sistemes informàtics i habitats per robots. L'organització està automatitzada fins el punt que la disposició dels articles en venda als magatzems, denominada *organització caòtica* resulta pràcticament incomprendible per a les persones, perquè els articles no estan col·locats per tipus, per ordre alfabètic o cap altre disposició lògica fàcilment comprensible per la ment humana, sinó que estan organitzats en base a les compres que es realitzen:

Los libros se apilan junto a las cacerolas, los televisores comparten espacio con los juguetes para niños. Como los datos almacenados en el disco duro de un ordenador, los artículos se distribuyen a lo largo y ancho de todo el espacio del almacén, cada uno de ellos es localizable unívocamente mediante un código de barras, pero imposible de encontrar sin ayuda de un ordenador. (p.130)

A mes, explica Bridle, els treballadors d'Amazon estan obligats a portar uns dispositius portàtils que registren tots els seus moviments i assignen puntuacions a la seva eficiència. Si no poden seguir el ritme que marca la màquina se'ls descompten punts —diners—. Bridle descriu aquest sistema opressiu que fa servir Amazon:

No tienen otra cosa que hacer más que seguir las instrucciones de la pantalla, empaquetar objetos y llevarlos de un sitio a otro. Se espera de ellos que se comporten como robots, que funcionen como máquinas mientras sigan siendo, de momento, ligeramente más baratos que aquellas. (p.130)

Aquesta reducció dels treballadors a màquines, diu Bridle, fa que sigui més fàcil contractar-los, acomiadar-los i abusar d'ells. Amb una política similar, l'empresa Meta, propietària de Facebook, Instagram i Whatsapp, va dur a terme un acomiadament de part de la plantilla de col·laboradors sense oferir més explicació que un algoritme els havia triat a ells (Hays, 2022).

Així doncs, el lloc que ocupa la tecnologia a la vida social ens revela els valors que poden estar presents en el seu disseny. Es poden dissenyar per fer servir de manera antidemocràtica, per perpetuar models jeràrquics de control, com en els casos exposats, o se'ls pot donar un sentit positiu que faciliti assolir fites humanes, en línia amb els postulats de Marcuse, que recupera i adapta Feenberg (2002) a l'actual societat postindustrial:

The place computers are intended to hold in social life is intimately connected with their design. Systems designed for hierarchical control are congruent with rationalistic assumptions that treat the computer as an automaton intended to command or replace workers in decision-making roles. Democratically designed systems must instead respond to the communicative dimension of the computer. As a medium it facilitates the self-organization of human communities, including those technical communities the control of which founds modern hegemonies. (p.107)

Tamnit Gebru es mostra escèptica sobre les promeses de les grans tecnològiques fan sobre la democratització de recursos que proporcionarà la IA i elles com a principals ambaixadores. Si fins ara els fruits de la tecnologia (com internet, l'electricitat, l'aigua potable o el transport) no s'han distribuït de manera equitativa, no existeix una raó de pes per esperar que aquesta utopia es materialitzi avui. En comptes d'això, Gebru adverteix sobre el perill d'assumir la concentració de poder en torn a la IA i confiar en què els beneficis es repartiran. En comptes d'això, proposa crear el món que volem de manera proactiva (Walsh, 2022).

2.4 PER UN LIDERATGE CULTURAL DE LA MEDIACIÓ ALGORÍTMICA

La nuestra es una cultura que se devora a sí misma a través de la información, y la mayor parte de nosotros ni siquiera se pregunta cómo podría controlar el proceso.

Neil Postman

Perquè hi hagi experiència literària en el futur potser caldrà desinflar primer aquest gran globus de paraules en què vivim, i recuperar el silenci actiu de la creació.

David Vidal Castell

2.4.1 Algoritmes com a generadors de cultura

A bans de plantejar-nos com podem intervenir de manera constructiva en la nova esfera pública des de l'àmbit de la comunicació, haurem d'estar en condicions de valorar els efectes d'aquesta mediació algorítmica. El segon capítol d'aquesta tesi el dedicarem exclusivament a tractar el tema del paper de les plataformes de comunicació amb IA i els seus efectes en l'ecosistema informatiu. Ara farem unes pinzellades que ens han d'ajudar a reconèixer la influència cultural de la IA, ja que condiona el nostre consum de continguts i per tant el nostre comportament.

Tal com observa Feenberg (2002), tot i que les aplicacions tecnològiques tinguin com a principal finalitat solucionar problemes pràctics, aquestes no es poden comprendre només en termes funcionals, sinó que és necessari considerar el seu paper com a generadores de cultura. Aquesta és una de les raons per les quals proposarem, en línia amb aquestes idees, que caldrà assumir la necessitat d'un lideratge cultural per la innovació tècnica:

Although technologies are first and foremost tools for solving practical problems, they are not fully understandable in functional terms. This is especially true in cases where their function is itself in dispute. As we have seen with computers, these disputes are not merely technical but go to the cultural significance of the technology. The critical theory of technology is therefore a cultural theory. (p.107)

Aquesta darrera revolució tecnològica està tenint efectes tan exponencials en els àmbits de la comunicació i la cultura com les tecnologies que els protagonitzen. Les intel·ligències artificials de les plataformes que dominen internet marquen aquesta nova etapa gràcies a la seva capacitat per gestionar la descomunal quantitat d'informació que circula per la xarxa. El fenomen que s'inicià fa algunes dècades que consisteix a delegar en procediments computacionals la tasca de la cultura d'ordenar, classificar i jerarquitzar persones, llocs, objectes i idees ha donat lloc avui a una nova cultura algorítmica (Striphas, 2015).

Definit per Just i Latzer (2016) l'algoritme aplicat a la gestió del coneixement és "a process that assigns (contextualized) relevance to information elements of a data set by an automated, statistical assessment of decentrally generated data signals". Els mateixos autors assenyalen que aquests sistemes no només influencien quina informació trobem, sinó també la seva reputació i per tant la fiabilitat que li atribuïm.

Les tècniques de viralització de continguts que es fan des dels departaments de comunicació digital que serveixen per maximitzar l'*engagement* a les xarxes socials o les tècniques de posicionament en cercadors (*Search Engine Optimization*) resulten pràctiques que tenen com a finalitat *hackejar* els algoritmes de les plataformes digitals que serveixen per ampliar la visibilitat dels propis continguts. Això fa que els les organitzacions amb més recursos tinguin moltes possibilitats d'aconseguir que els seus propis missatges tinguin més visibilitat en aquests ecosistemes comunicatius que dominen l'esfera pública digital. Per tant, la construcció de la realitat i de l'ordre social han passat avui a estar condicionats per aquesta mediació automatitzada —*hackejable*— que influencia en què pensem, com hi pensem i en conseqüència, com actuem. A les plataformes tothom pot tenir veu, cert, però no és menys cert que alguns poden tenir més veu que altres, i no només a través de la visibilitat que es compra pagant per anuncis, sinó mitjançant aquests *hackejos* a l'abast de qui els conegui o se'ls pugui finançar.

A l'apartat en el qual hem tractat la qüestió de la irracionalitat del científisme i la patologia del *dataïsme* exposàvem el problema de la relació que les nostres societats, seduïdes per la tecnologia, concedeixen un caràcter mític als algoritmes que representen l'ordre cartesià i —en conseqüència— la legitimació del pensament economicista. Aquest aval és el que atorga a les propietàries dels algoritmes més importants en la mediació el caràcter de "*màquina cultural*" en paraules de Edd Finn (2018), qui la defineix com la màquina "que opera tanto dentro como más allá de la barrera reflexiva de la *computabilidad efectiva*, produciendo cultura un nivel macro-social, al mismo tiempo que produce objetos culturales" (p.69).

Fins l'emergència de les plataformes, eren sobretot els mitjans tradicionals qui exercien aquest rol, i val a dir que no sempre amb responsabilitat. Deixant per ara aquesta qüestió de banda, en el nou context en el qual les tecnològiques han assumit el paper de *gatekeeper* té sentit preguntar-nos quines són les principals diferències entre la comunicació computacional dominant i la que han exercit els *legacy media*. Segons Just i Latzer, per un costat, s'identifica l'emergència d'una nova constel·lació d'actors que influencien l'opinió pública i, per l'altre, una creixent personalització dels continguts.

A més, aquesta nova etapa que s'inicià amb el començament del nou mil·lenni, també està desencadenant canvis radicals en les formes de relacionar-nos, així com en la dissolució de les fronteres entre allò que és públic i la intimitat. Unes modificacions que fan preveure una regeneració cultural a l'alçada de la que es va produir amb l'arribada de l'electricitat i la seva

aplicació a la comunicació, o fins i tot al naixement de la impremta, tal com apunten Duch i Chillón (2012):

Aspectos esenciales del vivir están cambiando a ojos vista, entre ellos la «relacionalidad» entre grupos y sujetos —sin duda uno de los más decisivos—, o los cada vez más ambiguos lindes entre las esferas íntima, privada y pública.

La envergadura de la mudanza es cuando menos comparable a la que acarrió la electrificación de las comunicaciones, matriz del telégrafo, la radio, la televisión o el cine; o incluso a la imprenta, sin cuya formidable irrupción habría resultado inviable el mundo moderno. (p.458)

És en previsió d'aquests canvis que resulta oportú plantejar com volem adoptar l'automatització. Això pot passar per valorar, a partir de l'experiència que hem adquirit els darrers temps, si aquesta nova esfera pública digital és la que volem, aleshores, en paraules d'Íngrid Guardiola (2019) "repensar el contracte legal i ciutadà que volem establir" (p.15). Finalment, des de la comunicació, haurem d'avaluar com contribuir en l'assoliment d'una esfera pública digital basada en els grans principis d'igualtat, justícia i llibertat. Si la única garantia contra l'abús de poder és una ciutadania il·lustrada i informada (Ruiz-Caballero, 2016), al periodisme li correspondrà, en primera instància, assumir el repte que implica exercir com a garant del dret a la informació, en una aposta decidida per la informació de qualitat en aquest ecosistema digital ensordidor.

2.4.2 Colonització comercial de la web original

Els valors fundacionals de les comunitats que protagonitzaren els inicis d'Internet resulten una referència adient quan ens plantejem examinar els efectes dels algorismes en la comunicació i la cultura, atès que els grans principis que regien aquella etapa inicial, basats en l'intercanvi desinteressat de coneixement, estaven veritablement orientats al bé comú. Com veurem a continuació, llavors es concebia la xarxa com l'espai virtual en el qual es desenvoluparia un nou socialisme digital basat en la lògica comunitària del *open source*. Així, en les properes pàgines ens proposem mostrar que l'esfera pública digital que dominen les grans plataformes capitalistes no és la única possible, que la lògica en xarxa té un gran potencial beneficiós per a la ciutadania, i que aquest estarà condicionat pels criteris que determinin l'activitat que s'hi porta a terme.

La *world wide web* que va néixer el 1989 s'ha anat fent més complexa i difícil de valorar en conjunt (Casacuberta, 2021). Tanmateix les eines més utilitzades pertanyen a poques grans empreses tecnològiques capitalistes que, salvant les distàncies, comparteixen tecnologies i valors similars. La web governada per les grans plataformes de mediació algorítmica està lluny de la concepció idealista de Tim Berners Lee que, més propera a postulats socialistes, tenia a veure bàsicament amb facilitar l'intercanvi de coneixement i transaccions, i per tant contribuir a la democratització de la informació. Una filosofia de cooperació que es va iniciar a les universitats i que va donar lloc a la cultura *hacker* regida pel simple goig per la creativitat, la reputació entre la comunitat de programadors i la llibertat per crear, adquirir i distribuir coneixements, que més endavant es va concretar en el moviment del software lliure que va fundar Richard Stallman (Castells, 2003). Per a aquests precursors, el llenguatge de programació havia de ser sempre públic. D'aquesta manera no es privatitzaria i s'aniria evolucionant de manera conjunta.

L'esperit fundacional d'internet es resumeix en el codi ètic *hacker* (Levy, 1984):

1. L'accés a ordinadors i a qualsevol cosa que pugui ensenyar quelcom sobre la manera en què funciona el món ha de ser il·limitat i total.
2. Tota la informació ha de ser lliure.
3. Desconfia l'autoritat, promou la descentralització.

4. Els hackers han de ser jutjats per la seva capacitat i no pels seus títols, edat, raça o posició.
5. Pots crear art i bellesa en un ordinador.
6. Els ordinadors poden canviar la teva vida a millor.

Internet seria una “versió revisada del socialisme”, en paraules de Kevin Kelly, Director Executiu de la revista *Wired*, en el seu conegut assaig *The New Socialism: Global Collectivist Society Is Coming Online* (2009). En el mateix text, explicava també per què aquest socialisme digital era la gran innovació americana dels darrers temps:

We're not talking about your grandfather's socialism. In fact, there is a long list of past movements this new socialism is not. It is not class warfare. It is not anti-American; indeed, digital socialism may be the newest American innovation. While old-school socialism was an arm of the state, digital socialism is socialism without the state. This new brand of socialism currently operates in the realm of culture and economics, rather than government—for now.

Què va passar entre aquestes concepcions inicials i el que experimentem avui? A finals de la primera dècada dels anys 2000 van agafar protagonisme corrents emprenedors que, sota la promesa de grans beneficis econòmics, van donar lloc al que coneixem ara com la web 2.0 que governen les grans tecnològiques. Els ideòlegs més determinants per a la consecució d'aquest projecte van ser Steve Jobs i Tim O'Reilly, abanderat de la idea que per poder ser rendible a internet cal aprofitar les aplicacions d'intel·ligència col·lectiva, que “depenen de gestionar, comprendre i respondre a la quantitat massiva de dades generades pels usuaris en temps real” (O'Reilly i Battelle, 2009).

La periodista Marta Peirano (2019) fa notar que en els postulats de O'Reilly es fundava allò que avui s'ha anomenat el *capitalisme de vigilància* (Zuboff, 2019), perquè llavors ja apostava per aprofitar sensors i altra tecnologia present també als dispositius per extreure totes les dades personals sense que l'usuari en fos conscient: “Nuestros teléfonos y cámaras se están convirtiendo en los ojos y los oídos de las aplicaciones; sensores de movimiento y localización dicen dónde estamos, qué estamos viendo o a qué velocidad nos movemos. Los datos están siendo recolectados, presentados y aplicados en tiempo real” (O'Reilly, citat a Peirano, 2019, pp.185-186). La periodista diu d'aquest emprenedor que “es capaz de proponer con total candidez un modelo de negocio en el que las principales empresas se hacen millonarias explotando el trabajo no remunerado y espiando a millones de personas sin que le parezca un escándalo” (p.186). O'Reilly va convertir-se en el blanc dels defensors del moviment *hacker* per

ser l'artífex de la redefinició de la web que va menystenir el concepte inicial per una web regida pels valors de lliure mercat, al mateix temps que s'apropiava de la retòrica de la llibertat i del compartir i la col·laboració, o el que és el mateix: la internet de les plataformes que coneixem avui. Entre els seus crítics destaca Evgeny Morozov, que l'ha titllat fins i tot d'estafador (Díez, 2020) entre altres crítiques que li ha proferit. En un assaig dedicat només a O'Reilly, l'acusa de ser la causa de la buidor constant dels debats tecnològics actuals (*The Meme Hustler*, 2014):

The enduring emptiness of our technology debates has one main cause, and his name is Tim O'Reilly. The founder and CEO of O'Reilly Media, a seemingly omnipotent publisher of technology books and a tireless organizer of trendy conferences, O'Reilly is one of the most influential thinkers in Silicon Valley. Entire fields of thought—from computing to management theory to public administration—have already surrendered to his buzzwordophilia, but O'Reilly keeps pressing on. Over the past fifteen years, he has given us such gems of analytical precision as “open source,” “Web 2.0,” “government as a platform,” and “architecture of participation.” O'Reilly doesn't coin all of his favorite expressions, but he promotes them with religious zeal and enviable perseverance. While Washington prides itself on Frank Luntz, the Republican strategist who rebranded “global warming” as “climate change” and turned “estate tax” into “death tax,” Silicon Valley has found its own Frank Luntz in Tim O'Reilly.

En els moments en què escrivim aquest text, fa un temps que es parla del web 3, allò que alguns asseguren que serà l'evolució de la web 2.0. Una nova infraestructura on s'instal·larà la xarxa a la qual s'atribueix una lògica descentralitzada (una de les idees socialistes dels orígens de la web) que es materialitza en les aplicacions que neixen de la tecnologia de la cadena de blocs (*blockchain*) amb les criptomonedes, els Tokens no fungibles o *Non Fungible Token* (NFT) i el *hype* del Metavers com a innovacions més conegudes fins ara (Gonzalo, 2021; Martín, 2022; Stackpole, 2022). Aquesta suposada nova infraestructura no sembla convèncer els més crítics, que desconfien d'aquest discurs, recolzat també per O'Reilly i les mateixes *Big Tech* que van abanderar el projecte neoliberal de la web 2.0 (Evgeny Morozov [@evgenymorozov], 2022).

A l'actualitat, tanmateix, encara estem sota els efectes d'aquesta web 2.0 que ha derivat en el domini de les grans plataformes, els seus algoritmes optimitzats amb finalitats comercials i —en termes generals— un model de negoci que consisteix a extreure totes les dades personals possibles, tot sovint de maneres poc transparents. Alguns dels principals efectes del predomini de la lògica comercial dels algoritmes de les grans plataformes són coneguts: tendència a la personalització dels continguts amb possibles conseqüències en la polarització social, campanyes de propaganda i desinformació adaptades que *hackejen* les xarxes i aprofiten les

seves vulnerabilitats i un model de negoci que amenaça la privacitat de la ciutadania i per tant les seves llibertats (Bridle, 2020; Lassalle, 2019; Nechushtai i Lewis, 2019; Pariser, 2011; Peirano, 2019; Ruiz-Caballero, 2016; Zuboff, 2019). A continuació desenvoluparem tots aquests aspectes.

2.4.3 Desintermediació i fluxos de coneixement

Tot i que la xarxa que coneixem i habitem avui —dominada per les forces de mercat i les grans plataformes com les seves principals representants— no és allò que els seus ideòlegs van imaginar, també és cert que l'ecosistema digital actual ha implicat efectes que per a moltes persones, organitzacions o causes han estat molt favorables. Així, per poder valorar quin ha de ser el paper del periodisme en aquest nou context, resulta necessari avaluar el rol de la mediació algorítmica a l'ecosistema informatiu considerant tant les mancances com els beneficis.

Des de certs sectors intel·lectuals, així com des del periodisme, s'ha criticat molt el paper de la mediació algorítmica i els seus efectes fins el punt de responsabilitzar les plataformes dominants de contaminar l'esfera pública digital i per tant debilitar la democràcia. Poden ser aquestes crítiques fruit, en part, de la por a la substitució per les màquines, que avui seria de caire intel·lectual? En el transcurs de la història sempre han existit grups que han criticat el desenvolupament tecnològic, i en els casos més notoris —com en les revoltes luddites durant la industrialització— l'animadversió ha sorgit quan la tecnologia ha amenaçat l'ordre establert o els llocs de feina.

Per poder avaluar els efectes tecnològics val la pena tenir present que el gran canvi que experimentem dels darrers anys ha comportat avenços importants en moltes esferes, i per descomptat també ha implicat grans beneficis en la comunicació i la cultura. La digitalització ha significat un nou pas en l'alliberament de les restriccions de temps i espai en els processos comunicatius i en l'economització de l'esforç, fent encara més prescindible la presència física, per exemple o, gràcies precisament a la IA, a disposar de més eines per manifestar la pròpia creativitat, per comunicar-nos de més i diferents maneres o per crear i compartir coneixement.

Les noves plataformes han possibilitat que qualsevol persona o causa pugui ser seu propi mitjà de comunicació sense necessitat de disposar de coneixements tècnics de programació. Mitjans que, gràcies a les tecnologies exponencials, permeten una enorme eficiència de la comunicació en permetre segmentar continguts a determinats grups d'interès, per exemple. Periodistes o investigadores poden accedir de maneres increïblement senzilles a un univers d'informació per documentar-se. Estudiants poden descobrir continguts, així com valuosos recursos arxivats en línia o formacions que imparteixen grans universitats del món abans inaccessibles a les persones que no tinguessin un alt nivell adquisitiu. Gràcies en part a aquestes plataformes, un petit museu

local pot adreçar-se a investigadores o altres públics internacionals; negocis d'estructures reduïdes disposen de més possibilitats de tenir visibilitat —i per tant de prosperar—; moviments socials també compten amb més eines per ampliar o consolidar les seves comunitats...

Les xarxes socials van trencar l'esquema clàssic de la comunicació en què un parla mentre la majoria escolta. Les audiències han passat a ser productores de continguts, i això té la capacitat de contribuir a generar unes xarxes d'intercanvi d'informació que tenen el potencial de millorar el coneixement compartit. Com a conseqüència, tindria sentit esperar una societat més lliure, independent amb més capacitat d'innovació i unes condicions que “faciliten l'exercici crític i l'emancipació real dels individus” (Bustamante, 2010).

En l'àmbit del periodisme s'ha criticat abastament la irrupció de les plataformes digitals. Tot sovint sembla que oblidem que també se n'ha beneficiat, com dèiem, fent servir les IA dels cercadors per investigar, per exemple. També ha utilitzat les xarxes socials per identificar possibles temes noticiables, per contactar amb testimonis, per desenvolupar la notorietat de la pròpia signatura de les autores, per trobar imatges o altra informació que ha servit per nodrir les notícies de les que informen. Pot ser que la crítica sigui una reacció al fet d'haver perdut centralitat, tal com han suggerit alguns autors? (Casacuberta, 2021; Vidal Castell, 2020). En el cas dels mitjans tradicionals, també han intentat utilitzar l'accés que les plataformes proporcionen a les audiències per atraure usuaris als seus webs a través de tècniques de posicionament en cercadors (SEO), creant perfils a determinades xarxes socials per arribar a noves generacions o directament usant les eines de publicitat que proporcionen per segmentar continguts a determinats públics i aconseguir augmentar les mètriques comercials.

Aquesta tecnologia fa que la informació circuli de manera més ràpida, en temps real i a demanda. Podem dir que en part la web actual ha heretat certa lògica provinent de l'ànima dels primers impulsors d'Internet, perquè les barreres per accedir a tota mena de continguts com i quan els necessitem han desaparegut.

Per què, totes les crítiques al nou ecosistema digital des d'alguns sectors del periodisme i altres àmbits intel·lectuals? per temor a la pèrdua del negoci? de poder? Es tracta d'un sentiment intimidant que genera una tecnologia amb capacitats per fer innecessària la participació humana? per acabar amb el periodisme? es poden assimilar les crítiques d'avui a les màquines que s'ocupen del nostre treball mental a les que rebien els artefactes als que encarregaven del treball manual a les societats industrials?

Aleshores, la productivitat, eficiència i rapidesa de les màquines intimidava certs sectors obrers amb els luddites com a cas paradigmàtic, perquè les persones humanes no podem competir en aquestes habilitats amb elles. El filòsof Dardo Scavino (2022) cita els textos del luddita anònim *Ereuhon* que, com a conseqüència de l'enlluernament que li provocava la tecnologia, temia que les màquines arribessin a supplantar els éssers humans, suprimint aquells que resultessin inútils per al desenvolupament de la seva pròpia espècie artificial:

El hombre recurre a la máquina porque la considera mejor que él mismo”; “nuestras máquinas de calcular no se olvidan de ninguna cifra y nuestros telares automáticos de ningún punto”; “la máquina sigue estando alerta y activa cuando el hombre se fatiga”; “tiene espíritu lúcido y calmo cuando el hombre se encuentra embotado y aturdido”; “no necesita dormir” y “se muestra siempre presta a trabajar como si su ardor no se extinguiera”. (p.246)

Tal vegada el luddita temia la màquina perquè la considerava superior a l'ésser humà. I potser aquest sentiment d'inferioritat provenia de considerar rapidesa, eficiència i productivitat com a principals característiques de la intel·ligència humana. Tal vegada s'havia acabat creient que el valor suprem era l'eficiència i per tant potser compartia més valors amb els propietaris dels mitjans de producció dels que creia i podia aparentar.

Així doncs, tornant al present amb els aprenentatges fets gràcies també a les exageracions luddites, ens hauríem de preguntar primer què significa la intel·ligència per a nosaltres. Com anirem veient al llarg d'aquest treball, la intel·ligència humana és més que aquestes tres característiques relacionades amb l'eficiència i la productivitat, valors presents a les nostres societats i provinents de les influències empresarials que busquen el màxim benefici particular al menor cost.

Existeixen, doncs, diferències entre els temors actuals i els del passat? Podem dir que tenen algunes similituds, i probablement la principal és la por humana a la substitució per les màquines. Les innovacions actuals, marcades per les tecnologies exponencials, generen uns matisos en la mena de reserves cap als seus usos. El rebuig actual al tipus de màquines que realitzen treball mental el sentim en part perquè tenim la sensació que la delegació de determinades tasques a la IA erosiona la nostra capacitat de judici i pensament crític. Que a mesura que deleguem, renunciem a la nostra autonomia individual (Koenig, 2019), cosa que, segons alguns autors, fins i tot podria qüestionar els fonaments de la societats liberals (Lassalle, 2019). Yuval Noah Harari (2018) defensa que la capacitat humana per decidir lliurement s'està col·lapsant, i que arribarà un moment que no podem apagar les màquines perquè en

serem dependents en excés. Val a dir que la posició de Harari és en certa manera particular i des d'alguns sectors ha estat força criticada, tal com detallarem més endavant.

En les pors també influeixen els mites creats per les narratives transhumanistes (Diéguez, 2016) impulsades per la ciència ficció, on els robots intel·ligents acaben dominant i aniquilant l'espècie humana. Sobre això darrer, el professor de filosofia computacional Santiago Sánchez Migallón reivindica amb encert desmitificar aquesta suposada *superintel·ligència*: si aquestes màquines acaben sent tan intel·ligents, de veritat no se'ls acudirà res més que eliminar els humans? es pregunta. (Torrubia, A, 2021).

Com no pot ser d'una altra manera, pensem que el periodisme ha de fer un pas endavant i aspirar a tenir un paper clau en aquest lideratge cultural de la innovació tècnica que reclamem, i l'ha de fer remetent-se de nou al seu propòsit de proporcionar informació a la ciutadania perquè aquesta pugui ser lliure i capaç de governar-se (Kovach i Rosensteil, 2014). Perquè un periodisme de qualitat —ètic, per tant— és necessari per a la democràcia i en conseqüència per a la ciutadania. Per fer-ho haurà de redefinir el seu rol davant la creixent automatització a la societat. Un cop haguem fet tota l'exposició del context, dedicarem un capítol sencer (el darrer) a desenvolupar les propostes pertinents. A les properes pàgines de moment, traçarem, de maner molt breu, unes línies a la sorra per començar a assimilar, des de la perspectiva del periodisme, algunes qüestions cabdals tractades fins aquí.

2.4.4 Periodisme i humanisme

Es canvis en curs demanen una reflexió, des del periodisme, sobre quin ha de ser el seu paper en aquestes societats en les quals tendim a confiar cada cop més en la intel·ligència artificial en la presa de decisions i en la construcció de la realitat. Pensem que per poder incentivar un debat en la professió sobre com ha de ser la seva intervenció constructiva serà necessari un enfoc humanista que ens ajudi a donar sentit, o el que és el mateix, a definir el projecte comú al qual volem arribar. Només des d'aquesta perspectiva amplia podrem valorar com contribuir-hi des de la pràctica de la professió.

Per fer aquesta feina, caldrà incorporar a aquesta reflexió transdisciplinària l'herència crítica dels estudis sobre periodisme que arrelen en la tradició de la filosofia i l'antropologia de la comunicació, amb una marcada transmissió dels estudis humanístics. Des d'aquesta perspectiva podrem proposar una concepció de la tecnoètica que neixi d'una visió crítica sobre el mite del progrés lineal que promou el pensament científico-tècnic al servei dels valors neoliberals que consideren el consum sense límits com a motor d'evolució social. Dur a terme una proposta de futur que implica una redefinició del progrés en un concepte que faci possibles que les condicions de l'avenç siguin compatibles amb un consens centrat en reducció de les desigualtats i la preservació del medi ambient.

Una mirada que consideri la natura mítica de l'ésser humà com a contrapès al discurs legitimador hiperracionalista que s'atribueix la capacitat de predir el futur i de determinar-lo en base a dades del passat. Una concepció de l'home tal com la defensen autors com Lewis Mumford, Lluís Duch o Josep Maria Esquirol, entre altres que hem anat citant i que determinen l'enfoc de tot aquest treball.

Per incentivar la reflexió en torn als reptes ètics que implica l'adopció de la IA a les rutines periodístiques, també partim d'una idea de la racionalitat que defensa la teoria crítica a través d'autors com Herbert Marcuse o Andrew Feenberg, que conviden a plantejar l'evolució com un creixement de l'estat de benestar a l'abast de tota la societat. Un model en el qual l'evolució impliqui proporcionar a la ciutadania un millor accés a les fonts bàsiques de benestar, que són les que habiliten les persones a tenir una vida sana, digna i plena. Aquestes necessitats bàsiques són un medi ambient conservat, aliments nutritius accessibles i més temps disponible per gaudir de les pròpies aficions, l'art, la natura o a desenvolupar les relacions personals.

Aquesta concepció del desenvolupament al servei de l'ésser humà implicarà fer reflexions de caire ètic abans d'adoptar les tecnologies exponencials i per tant aspirar a una innovació

influenciada per valors humans i el respecte pels principis democràtics i l'ecosistema planetari. En conseqüència, per plantejar l'assimilació de la IA a les rutines periodístiques caldrà treballar en la concepció de la pròpia tecnoètica i així evitar que prevalguin únicament els criteris comercials.

El pensament humanista ens ajudarà a donar sentit i a dirigir les nostres accions cap al projecte comú al que proposem aspirar: una societat en la qual la introducció de la tecnologia estigui guiada per la promoció de finalitats socials rellevants i el benestar social. Des d'aquesta tradició, el periodisme estarà en condicions d'assumir un lideratge cultural del desenvolupament tecnològic posant l'ésser humà al centre.

2.4.5 Periodisme i tecnificació

Si bé és cert que el periodisme no ha estat pioner en l'adopció de la IA, el procés d'assimilació és ja una realitat, i la tendència és que encara es faci servir més. Tot això ho plantejarem amb detall a la recerca empírica que presentarem al quart capítol, on també explicarem com aquesta tecnologia es fa servir en totes les fases de producció de la cadena de valor del producte periodístic.

Identificar els usos actuals de la IA als mitjans resulta una tasca especialment complexa, perquè les aplicacions s'estan definint i s'integren en tota mena de processos, des dels editorials fins als comercials: es fa servir en la identificació de temes noticiables, en la generació automatitzada de continguts, canvis de formats, en la recomanació de missatges i la personalització del producte, entre molts d'altres. Les possibilitats de la IA, el creixement de les habilitats tecnològiques dels periodistes i la incorporació a les redaccions de perfils tècnics fan augurar canvis molt significatius tant en les rutines com en el producte periodístic. Aquesta tecnificació, si va acompanyada d'un millor coneixement de la lògica del llenguatge computacional, també serà d'ajuda per a interpretar la realitat d'unes societats cada cop més automatitzades i concretament per poder vigilar els usos interessats o irresponsables de la IA.

En aquest context, a l'hora d'incentivar el debat en el sector sobre com integrar les tecnologies exponencials, caldrà considerar, d'una banda, qüestions relatives a la reflexió que fa dècades que es produeix en torn a l'ètica de la IA i que s'identifiquen principalment en camps com la filosofia i la ciberètica. Aquests aspectes són els que tenen a veure amb l'abast i les limitacions d'aquestes tecnologies, com els riscos de generar biaixos en els resultats algorítmics, les relatives

a la transparència, l'explicabilitat i el retiment de comptes, la privacitat, l'agència de les màquines, o el perill d'una excessiva delegació que neix d'un pensament hiperracionalista limitat per la seva ineptitud de concebre l'ambigüitat de l'ésser humà i, com a conseqüència de les seves limitacions, atorga erròniament capacitats humanes a les màquines i converteix en maquinal el pensament humà.

D'altra banda, proposarem recórrer als valors que tradicionalment han guiat el periodisme de qualitat. L'èxit de l'adopció de la IA, com en altres moments històrics de canvis tecnològics, passarà per no perdre de vista el propòsit de la professió i la seva funció democràtica que consisteix a proporcionar informació veraç i espais de reflexió comú. Determinar els criteris ètics que el periodisme pot aplicar en l'ús de la IA passarà per sotmetre les noves eines i capacitats als seus valors i per tant a la seva missió de servei públic.

2.4.6 Periodisme i crisi

Aquest canvi tecnològic transcorre en un context en el qual el periodisme ha perdut centralitat discursiva degut, per una banda, a l'emergència de les noves plataformes de mediació algorítmica. Per l'altra, però, també es fa palesa una pèrdua de confiança en la professió producte de la seva pròpia acció, que sovint ha cedit a la corrupció i a la perpetuació de dinàmiques qüestionables de poders polítics i econòmics. La caiguda d'inversió en publicitat, que avui va a les tecnològiques, ha fet que els mitjans de comunicació tradicionals, ocupadors de la majoria de periodistes, estiguin passant per una crisi sense precedents a la recerca d'un model de negoci.

En proposar una nova tecnoètica per al periodisme considerem que aquest escenari, a priori advers, també és un moment d'oportunitat. Per un costat, una possible consolidació del model de negoci per subscripció, ara en tendència, resultaria un escenari favorable per apostar per un producte periodístic més independent (i per tant més favorable per recuperar confiança). Per l'altre costat, una tecnificació ètica de la professió podria contribuir a singularitzar el periodisme davant d'altres propostes de mediació que, en termes generals, es demostren mancades de suficient autoregulació des del punt de vista ètic. És a dir, una proposta clara i operativa sobre els criteris ètics que el periodisme pot aplicar en l'ús de la mateixa tecnologia de comunicació computacional que fan servir les grans plataformes, pot contribuir a generar més adhesions entre la ciutadania, cada cop més necessitada d'interlocutors en qui dipositar la seva confiança a l'hora de consumir informació.

Com hem vist, la IA està cada cop més present i en més sectors i tot sovint és utilitzada per perpetuar dinàmiques de poder. Al mateix temps, per l'esfera pública digital dominada per les grans plataformes de mediació algorítmica circulen enormes volums d'informació de tota mena, i molta d'ella de dubtosa qualitat. En aquest context el periodisme té l'oportunitat de fer un pas endavant, i això passarà primer per disposar d'una comprensió crítica ben fonamentada d'aquestes tecnologies i els seus efectes. Amb aquests coneixements, d'una banda, podrà vigilar-ne el seus usos, i per l'altra, podrà dur a terme el seu propi procés d'apropiació conscient de la IA que de manera que es demostri capaç de governar-la sota els valors que regeixen la professió periodística. Només adoptant les tecnologies exponencials de manera exemplar, el periodisme estarà legitimat per a la crítica, i per poder assumir, en un necessari exercici de responsabilitat, la part que li correspon en el lideratge cultural del desenvolupament tecnològic que proposem.

CAPÍTOL 3

EL PAPER DELS ALGORITMES COMERCIALS A L'ESFERA PÚBLICA

There's nothing in your life or in our collective problems that does not require our ability to put our attention where we care about. At the end of our lives, all we have is our attention and our time.

Tristan Harris

Ahora, más que nunca, el periodista debe ser un constructor de sentido en un mundo de hipérbolos y de sobreabundancia semántica.

Carlos Ruiz

La naturaleza fundamental de las tecnologías es que no tienen ninguna naturaleza fundamental. Somos nosotras, las personas, las que a través de nuestros usos y acciones damos propiedades concretas a esas tecnologías. Somos los que las definimos.

David Casacuberta

3.1 INTRODUCCIÓ AL CAPÍTOL 3

L'emergència de les plataformes de mediació algorítmica va ser percebuda amb optimisme per gran part de la ciutadania durant el temps que va precedir a la seva consolidació, entre els anys 2000 i 2012 aproximadament: disposàvem de més eines de creació de continguts i més canals per ser distribuïts, per tant més veus serien escoltades i tot plegat contribuiria a la salut de les nostres democràcies. De fet, en part va ser així ja que, com veurem més endavant, se'ls va arribar atribuir un paper important a l'hora de facilitar la comunicació en favor de lloables causes de transcendència social. Tanmateix, en el transcurs dels darrers anys, hem copsat certs aspectes del funcionament del seu model de negoci, així com determinats efectes inherents a la seva natura que en els seus inicis la majoria no vam apreciar. Des del món acadèmic i l'activisme —principalment vinculat a les ciències humanes i socials— s'ha arribat a evidenciar que el rol cultural i social que exerceixen aquestes grans plataformes no sempre s'adequa al propòsit amb el qual es defineixen, on s'atribueixen el paper fer un món en més lliure i democràtic. Es critica que, al contrari del que diuen a les seves missions, aquestes empreses erosionen les llibertats de la ciutadania com a conseqüència de la lògica extractivista de les dades personals en la qual basen el seu model de negoci. Com desenvoluparem més endavant, els retrets arriben fins el punt d'acusar-les de ser les responsables de la creació d'una nova societat de la vigilància, en la qual l'experiència privada ha estat transformada en un bé comercial; se les culpa també de ser la causa de la desinformació que s'origina a l'ecosistema digital que dominen, així com de les cambres d'eco que provoquen els seus algoritmes. Per tant, se les responsabilitza d'afeblir la cohesió social i de ser el germen d'unes societats més polaritzades; se'ls hi recrimina, així mateix, que basen la seva activitat en la manca de transparència.

En el seu propòsit, Meta s'atorga la capacitat de fer un món més connectat i crear comunitats (Constine, 2017); Twitter diu facilitar que qualsevol persona pugui tenir una veu pública amb la qual abans no comptava i tenir el poder de generar un canvi en el món (Twitter, 2022); Google, que existeix per democratitzar l'accés a la informació organitzant-la i fent-la accessible a tothom. El xoc d'aquests propòsits amb les acusacions revela, com a mínim, alguna mena de disfunció. En cap cas pretendrem afirmar que aquests propòsits no siguin certs, però el que efectivament el temps ha posat de manifest, com anirem exposant en el present capítol, és que aquestes promeses han resultat més aviat veritats a mitges —i excel·lents lemes publicitaris— encara que sigui perquè allò que mai han deixat clar és el preu que cal assumir per gaudir d'uns serveis que

són, aparentment, gratuïts. I és que tot això que s'ha anat evidenciant aquests anys no ha estat conegut precisament gràcies a una comunicació clara per part seva, sinó perquè, com dèiem, acadèmiques i activistes s'han preocupat d'analitzar-ho i divulgar-ho. També hem après a cop d'escàndols donats a conèixer per periodistes: alguns relacionats amb filtracions de dades, d'altres amb negligències de diferents tipus, amb mancances de transparència i altres praxis de dubtosa ètica que han arribat a posar en entredit aquestes narratives lloables que proclamen. La pressió social cap a certes pràctiques d'aquestes empreses i les recurrents controvèrsies en què s'han vist involucrades, han contribuït en certa mesura a sepultar molts dels beneficis reals que també proporcionen.

I és que la desconfiança neix tot sovint d'apreciar una desconexió entre allò que algú diu i allò que després fa. Això passa en les relacions personals, però es pot aplicar també a les organitzacions. Qualsevol persona que estigui familiaritzada amb la comunicació corporativa sap que un dels majors riscos reputacionals són aquests desajustos entre dir i fer. Això, sumat a la manca de transparència sobre el seu funcionament, ha provocat que avui existeixi un cert consens sobre el fet que moltes de les persones que perceben l'emergència d'aquestes noves plataformes amb optimisme ja no ho veuen de la mateixa manera (Newton, 2020).

En aquest capítol desenvoluparem tots els aspectes que acabem d'esmentar amb la finalitat d'identificar, de la manera més acurada possible, els efectes d'aquests models comercials de comunicació algorítmica. Des d'aquest punt de partida estarem en millors condicions per fer propostes, des de la tradició de l'ètica de la comunicació, que tinguin la finalitat de contribuir la millora de la salut de l'esfera pública en favor del bé comú. Amb aquest exercici aspirem a proporcionar materials per a la reflexió que permetin elaborar un retrat del context comunicatiu actual, ara que ens proposem el repte de no caure en visions esbiaixades de la realitat durant el camí. El nostre ànim l'emparaula d'una forma immillorable la filòsofa Marina Garcés: "Potser una de les funcions principals del pensament crític avui és recordar-nos que sabem molt menys d'allò que creiem. Només així ens posarem en situació d'aprendre, qüestionar i deixar de predicar" (Garcés, 2022, pp. 13-14).

3.2 ELS EFECTES DEL DOMINI DE LES GRANS TECNOLOGIQUES

La información se ha convertido en una forma de basura que no sólo no es capaz de dar respuesta a las cuestiones humanas fundamentales, sino que rara vez demuestra su utilidad a la hora de buscar soluciones a los problemas más prosaicos.

Neil Postman

El problema es cuando se diseña una tecnología muy potente, con efectos sistémicos, con el objetivo principal de ganar dinero, pero bajo la apariencia de querer mejorar el mundo y sin tomar en consideración los derechos de las personas.

Carissa Véliz

The price of freedom is the eternal vigilance we'll see all kinds of new devices

Aldous Huxley (1958)

3.2.1 Descentralització i democràcia

Com dèiem, als nous mitjans algorítmics se'ls concedeix haver facilitat en gran mesura la comunicació entre persones, de les organitzacions o de les causes. Aquestes possibilitats han ajudat a fer créixer petits negocis o a desenvolupar grups virtuals d'afinitats compartides o de suport mutu. A les plataformes digitals se'ls ha arribat a atribuir un rol fonamental inclús en el desenvolupament i consecució de diferents moviments de transcendència global que han propiciat canvis socials amb importants repercussions. La democratització de l'accés a Internet, a dispositius per navegar-hi i a eines de comunicació —a l'abast de les capacitats tècniques de qualsevol persona— va donar lloc a un nou ecosistema virtual caracteritzat per una explosió de contingut generat per la ciutadania, que va passar de ser consumidora passiva a productora activa i motivada d'informació (Taylor, 2014). La nova tecnologia aplicada a la comunicació que es va desenvolupar sota la filosofia de les startups de Silicon Valley va acabar de trencar el monopoli de la producció d'informació que tenia el periodisme. Tothom podia tenir veu.

Comptem amb força recerca que ha conferit un paper primordial a les plataformes en la consecució de les fites de grans moviments socials del S.XXI com el 15 M a Espanya o la Primavera Àrab (Brady et al., 2017; Muñoz, 2011). És conegut que no sol es van servir de xarxes socials virtuals per aconseguir visibilitat i adhesions, sinó que també se'n van beneficiar per facilitar les comunicacions relacionades amb l'organització interna dels propis moviments (Tufekci, 2013).

L'abans i el després del moment en què va començar a diluir-se la influència dels mitjans de comunicació es sol situar a l'època del naixement de *Blogger*, un servei que permetia crear el seu propi *weblog* a qualsevol. Per primer cop, no era necessari saber programar per tenir un espai per publicar a internet. Això va donar lloc a la blogosfera, el primer gran sistema virtual comunitari del qual qualsevol —programador o no— en podia formar part. Durant les manifestacions de la cimera del G8 a Gènova, el 2001, on va morir l'activista Carlo Giuliani com a conseqüència del tret al cap que li va propinar un *carabiniere*, es va donar un fenomen nou: la informació corria per la xarxa al mateix temps que les versions oficials dels fets i descreditava aquells mitjans que es limitaven a difondre comunicats interessats. La periodista Marta Peirano (2019) ho relata així:

La prensa generalista recogió el comunicado oficial, mientras en los foros y listas de correo circulaba un relato alternativo hecho de testimonios, fotos y vídeos de los manifestantes. (...)

La verdad era un esfuerzo colectivo; podían contestar la versión oficial de los hechos y usar sus propios medios para demostrarla. Se empezaba a hablar del periodismo ciudadano, un relato coral sin intermediarios ni filtros que avergonzaría a los periódicos vendidos al poder, controlados por los gobiernos autoritarios, gestionados por las clases dominantes y, en definitiva, podridos de corrupción. Y lo más novedoso: se podía hacer sin salir de casa, siguiendo las noticias y facilitando su difusión. (p.179)

Els esdeveniments que van succeir a de l'etapa del *blogging* van suposar encara un pas més en l'era de l'explosió de les plataformes digitals i començà el 2005, amb el naixement Youtube. El 2006, quan es va fundar Wikileaks, ja podia tenir perfil a Facebook qualsevol, i aquest també seria l'any en què es faria el primer tuit. El 2007 va sortir al mercat l'iPhone, i el 2008 la campanya electoral d'Obama passà a la història, entre d'altres coses, per ser la de major pressupost en mitjans digitals fins llavors, amb 8 milions de dòlars d'inversió. El 2009 nasqué, com alternativa als SMS, la plataforma de missatgeria directa Whatsapp i es popularitzà ràpidament.

Només un any després, el 2010, Mohamed Bouaziz, un noi de 26 anys, es va cremar a l'estil bonzo a la petita ciutat tunisiana de Sidi bouzid i desfermà una revolta popular que va acabar amb la dictadura de Zine el Abidine Ben Ali. Una revolució que va ser la primera en què totes les noves eines i plataformes tecnològiques de comunicació hi van jugar un paper fonamental, fins el punt que molts consideraren que, sense aquestes, la caiguda del dictador no s'hagués produït (Muñoz, 2011).

Durant la Primavera Àrab també es van documentar nous fenòmens vinculats als moviments socials com els activistes digitals, líders d'opinió que a més serien seguits per periodistes de tot el món per les xarxes socials. Els seus relats a través de Facebook o Twitter van exercir una influència que va impactar tot el globus fins el punt d'acabar condicionant la informació dels mitjans de comunicació més importants, com va passar, per exemple, arran de les constants aparicions de les activistes Gigi Ibrahim o Zainab Al-Khawaja. La pressió que aconseguien exercir aquestes líders d'opinió digitals sobre els règims opressors fins i tot hauria arribat a determinar el desenllaç dels esdeveniments (Tufekci, 2013).

El cas d'Egipte, on Mubarak va arribar a bloquejar l'accés a Internet per silenciar les protestes que van provocar la seva caiguda (De Moneo, 2011), també va ser paradigmàtic. La setmana abans de la seva dimissió s'havia passat de fer 2.300 tuits al dia vinculats a la situació política a

230.000. La xarxa virtual s'havia convertit en lloc de trobada de l'activisme global que gestà una nova forma de mobilització ciutadana (Castells, 2012).

Més endavant, passada la febre de l'arribada de les plataformes, van començar a néixer corrents escèptics que van atribuir a la participació virtual un rol força més discret en el desenvolupament dels processos polítics, atès que, argumentaven, els vincles socials que internet és capaç de crear resulten febles en excés per propiciar l'acció col·lectiva. Per a autors com Malcolm Gladwell (2010), més que per fer activisme resultaven eines per mantenir l'estat de les coses:

Facebook activism succeeds not by motivating people to make a real sacrifice but by motivating them to do the things that people do when they are not motivated enough to make a real sacrifice. (...) The Internet lets us exploit the power of these kinds of distant connections with marvellous efficiency. (...) But weak ties seldom lead to high-risk activism. (...) The instruments of social media are well suited to making the existing social order more efficient. They are not a natural enemy of the status quo.

El sociòleg també va posar en dubte el suposat paper de les xarxes en l'organització interna dels moviments. Gladwell considerava que amb el mateix poder de decisió resultava impossible prendre decisions tàctiques, estratègiques o de direcció filosòfica. Gladwell distingia així l'activisme tradicional de la seva variant digital:

...social media are not about this kind of hierarchical organization. Facebook and the like are tools for building networks, which are the opposite, in structure and character, of hierarchies. Unlike hierarchies, with their rules and procedures, networks aren't controlled by a single central authority. Decisions are made through consensus, and the ties that bind people to the group are loose.

També en aquest sentit, autors com Evgeny Morozov o Siva Vaidhyanathan asseguraven que no es podia provar que les xarxes socials haguessin estat decisives a Tunísia o Egipte, i que donar tanta importància a Twitter eclipsava la notícia real, que no era altra que milions de persones havien decidit arriscar les seves vides per la possibilitat d'un món millor (Morozov, 2012; Vaidhyanathan, 2012).

Així que, probablement, la resposta sobre si les xarxes realment han jugat un paper fonamental a l'hora de propiciar canvis d'ordre social s'adequaria a una anàlisi que té més relació amb el fet que s'hagin utilitzat de manera adient, com reconeixia més endavant el propi Morozov (2014):

Si nos ceñimos al papel de las plataformas digitales en el compromiso político o el activismo, dependerá también de en qué estamos pensando: si se trata de un movimiento político organizado con fines y objetivos estratégicos, sería

estúpido no explotar el poder de Facebook, de Twitter o de cualquier otra plataforma para difundir su mensaje o ganar votos.

Avui, qualsevol que vulgui comunicar pot disposar fàcilment del seu propi canal gràcies a les plataformes digitals, i això confereix un poder a la ciutadania que resultava impossible tenir abans de la desintermediació. Les xarxes socials han seguit tenint un paper principal en la comunicació de qualsevol mena de causa. Gràcies a les xarxes o als cercadors, moltes persones s'han pogut sentir acompanyades en grups de suport, per exemple, compostats per famílies de criatures amb malalties minoritàries i altres problemes, que en ser poc habituals, sovint no compten amb tot l'ajut que els caldria i ara poden obtenir en trobar-se amb persones que tenen les mateixes inquietuds.

Les xarxes, en general, també provoquen un efecte de viralitat positiva que ha ajudat a gent, per exemple, a trobar feina, com és el cas d'un sensesostre que va aconseguir un lloc de treball i una casa on viure gràcies a una usuària de LinkedIn que li va dedicar una publicació (Cavestany, 2022). Per tant, les valoracions que tot seguit farem sobre les raons que hi poden haver rere la desconfiança en vers les tecnològiques i el seu paper a la societat, partiran de la base que aquestes, en alguns sentits, també han multiplicat oportunitats, i això ha estat possible gràcies als efectes democràtics de facilitar el la comunicació entre les persones.

3.2.2 Després de l'esperança va arribar la desconfiança

Les plataformes van néixer com a petits projectes que van créixer fins assolir dimensions globals en pocs anys. El cas de Facebook, que va aparèixer com una xarxa per posar en contacte universitàries, n'és un exemple. Això podria explicar en part la raó per la qual aquestes empreses s'han ocupat més aviat poc de desenvolupar regulacions internes que preveïessin possibles conseqüències negatives de la seva activitat. Una altra possible explicació pot ser que, tot sovint, en els àmbits vinculats a l'enginyeria, predomina la lògica tècnica de problema-solució amb què sol costar calibrar potencials efectes que van més enllà de la qüestió tècnica que es pretén resoldre. Aquestes mancances de previsió han causat problemes molt difícils de gestionar com les noves formes de desinformació, les filtracions de dades personals, les campanyes de propaganda organitzades, l'espionatge o l'assetjament, fins la publicació de pornografia infantil o retransmissió en directe a través de les xarxes socials d'actes de màxima violència com maltractaments o decapitacions.

Les principals xarxes socials, així com Google i altra mena de plataformes algorítmiques com Amazon, s'han vist involucrades en diferents casos similars als mencionats, però de totes aquestes, Facebook és la que s'ha vist involucrada en més escàndols, i més greus. El darrer, que va ser conegut com els *Arxius de Facebook (Facebook Files)*, va ser el detonant perquè Zuckerberg patís l'enèsima embranzida de crisi reputacional, que va resultar prou intensa com per provocar un pla de gestió de crisi. Aquest full de ruta va implicar, d'una banda, el canvi de nom de la companyia —que va passar a dir-se Facebook a Meta—. I, d'altra banda, va comportar que, amb ànim regenerador, decidís associar el aquest nou nom al que Zuckerberg considera el futur d'internet, la realitat virtual a la que rebateja amb el nom de Metavers (Meta, 2021).

Els Arxius de Facebook van ser filtrats per una ex treballadora d'aquesta empresa, Francis Haugen, al diari *The Wall Street Journal (WSJ)*. Entre d'altres coses, els documents revelaven que a Facebook existia una *llista blanca* de perfils de persones famoses i determinades empreses a les quals se'ls permetia saltar-se les normes d'ús del servei. Un dels casos que el WSJ posava com a exemple va ser el de la pornovenjança del futbolista Neymar qui, en ser acusat per una dona d'haver abusat d'ella sexualment, va publicar des de la seva pàgina de Facebook —seguida per milions de persones— tot un seguit d'imatges íntimes de la denunciant sense que li suposés cap problema amb la plataforma (cal tenir en compte que Meta és la mateixa plataforma ha arribat a censurar nus artístics de pintures i escultures per violar les seves normes).

Un altre dels expedients dels *Arxius de Facebook* es referia a l'existència d'una sèrie d'estudis interns que conclouien que Instagram era perjudicial per a les noies adolescents perquè causava i —sobretot— agreujava trastorns mentals. En els mateixos documents també es podia comprovar que a la seva plataforma s'hi anunciaven màfies de tràfic de persones, organitzacions xenòfobes i reclutadors de sicaris.

Segons el WSJ, Zuckerberg va defensar al Congrés que a Facebook s'eliminaven el 95% dels continguts que incitaven a l'odi, per contra, els informes interns van revelar que només en suprimia el 2%. La moderació portada a terme amb algorismes no resulta eficient, però incorporar el control humà que seria necessari implicaria renunciar a masses beneficis, segons demostraven els documents. Els *Arxius de Facebook* van ser una prova contundent que, sent conscient de tots aquests perjudicis, Facebook no havia actuat.

El cas dels *Facebook Files* és només l'últim de la quantitat de pràctiques qüestionables i negligències que han sortit a la llum, especialment des del 2016, l'any en què Trump va guanyar les eleccions als EUA. Abans també hi havia hagut debats sobre el paper de les plataformes en relació a determinats assumptes, com els usos de les xarxes per part de grups terroristes, que les empraven per a la difusió de missatges violents i la captació d'adeptes.

A partir el 2016 i com conseqüència d'una acumulació d'escàndols, els corrents crítics d'opinió van augmentar. Des d'aleshores, el degoteig de titulars relacionats amb problemes deguts a males praxis de les plataformes va ser permanent. A més de l'emergència del problema de la desinformació i les campanyes orquestrades, el 2016 va ser l'any en qual es van robar les dades de 117 milions d'usuaris mitjançant un *hackeig* a la plataforma LinkedIn per a ser venudes a un fòrum rus (Griffin, 2016). Uns mesos més tard es coneixia la proliferació de grups dedicats a la pornografia infantil a Facebook (Crawford, 2017). A principis del 2017 es van retransmetre en directe per Facebook les tortures a un jove de 18 anys mentre els agressors insultaven a Trump i a les persones blanques en general (Yan, 2017). En un altre directe, es va emetre la violació d'una dona a Suècia (cosa que va provocar que Facebook col·laborés amb la policia cedint determinades dades per identificar l'agressor) (Erickson, 2017). Aquell mateix any, van quedar sense filtrar dotzenes d'imatges il·legals que incloïen una decapitació executada pel grup terrorista Estat Islàmic, pornografia i violència infantil o propaganda que glorificava els actes terroristes que recentment s'havien perpetrat a Londres i a Egipte. També llavors, un gran grup d'anunciants va boicotejar Youtube en protesta per haver vist aparèixer els anuncis dels seus productes en vídeos amb contingut abusiu (Lomas, 2017), fet que es va repetir anys mes tard

(Wakabayashi i Maheshwari, 2019). Poc després, el *vlogger* Logan Paul, amb milions de seguidors, penjà un vídeo a Youtube que va ser àmpliament criticat per les associacions de prevenció del suïcidi perquè, coincidien, incitava a treure's la vida (Meyer, 2018).

L'any 2018 Facebook i Twitter van comparèixer al Congrés dels EUA per declarar sobre la investigació de l'exdirector del FBI Robert Mueller, a la que, per cert, Google, en la seva línia de no exposar-se per evitar danys reputacionals, no s'hi va presentar (S. Kovach, 2018). Aquell 2018 també es va saber que Facebook havia compartit dades personals de milions d'usuaris amb fabricants de telèfons mòbils (Mayer, 2018).

Pel que fa a Google, finalment, el seu CEO va accedir a comparèixer al Congrés a finals de l'any 2018, per declarar sobre les pràctiques de l'empresa relacionades amb la privacitat i suposats biaixos en la informació política del cercador. Per bé que Google ha tingut força èxit en la seva política de perfil baix, també s'ha vist involucrada en diferents problemes relacionats amb recomanacions algorítmiques de continguts perillosos a Youtube, per exemple. També ha estat assenyalada per un tractament de les dades personals que, segons diferents acadèmiques i altres divulgadors, resulta una amenaça per a la privacitat de les persones i per tant per a les seves llibertats (Zuboff, 2019). Prova d'això és que els cotxes de Google que van circular pels carrers del món sencer de 2006 a 2010 per desenvolupar el seu servei Google Maps, van aprofitar que hi passaven per capturar i emmagatzemar dades d'habitatges com noms de la WiFi o adreces MAC de l'enrutador. Si la xarxa no estava encriptada, a més, capturava correus, URL, FTP i paraules clau de wi-fis, segons va posar en evidència una auditoria independent (*Source Code Analysis of Gstumbler*, 2010).

També va ser conegut el cas dels enregistraments de converses privades per part de l'assistent de veu d'Amazon. (Shaban, 2018; Wolfson, 2018). Whatsapp, implicada en diferents casos—en especial relacionats amb campanyes de desinformació— no ha estat exempta de titulars i crítiques (Tardàguila et al., 2018).

Aquests són només una mostra dels problemes en què aquestes companyies s'han vist involucrades i que des de llavors no han deixat de repetir-se. Tanmateix, d'entre tots els que s'han produït fins avui n'hi ha un, destapat per un jove científic de dades, que va destacar per la seva transcendència política. Christopher Wylie, treballador de l'empresa Cambridge Analytica, va revelar tots els detalls d'una estratègia de publicitat segmentada que va consistir en obtenir, de manera irregular, però amb la col·laboració de Facebook, dades de més de 50 milions de comptes d'usuaris d'aquesta xarxa per crear perfils psicològics de persones a qui després

s'impactaria amb publicitat orientada a explotar les seves vulnerabilitats psicològiques —amb la finalitat de condicionar el seu vot a Trump—. Aleshores es va considerar que aquesta estratègia podria haver influït en la victòria del polèmic candidat republicà a les eleccions presidencials dels Estats Units l'any 2016 (Gibbs, 2016).

Durant molts anys, el mantra de Mark Zuckerberg, davant qualsevol crítica va ser “we are not a media company” (Segreti, 2016). El CEO de Meta va deixar de pronunciar progressivament aquesta frase en una època determinada que va seguir l'escàndol de la filtració a Cambridge Analytica —després de la compareixença que va fer al Congrés dels EUA per donar explicacions—. Aquell dia la va repetir, va reiterar que només eren una plataforma tecnològica conformada per enginyers que dissenyaven eines. Tanmateix, també va acabar reconeixent que efectivament tenien responsabilitat sobre el contingut que la gent comparteix a les seves plataformes (M. Castillo, 2018). Aquesta postura va quedar del tot desautoritzada quan, arran de la pandèmia de la COVID-19 i la quantitat de desinformació que es viralitzava, Meta va decidir, per primer cop, eliminar publicacions que tinguessin el potencial de produir dany físic imminent, fet que va implicar, per tant, un important canvi els fonaments de la seva política (García Aller, 2020).

Autodefinir-se com *una tecnològica que desenvolupa eines* era una altra manera de posar en evidència el seu tarannà determinista, en atorgar una suposada neutralitat a les seves *eines* que *simplement* servien per connectar persones i crear comunitats. Apel·lava a la suposada neutralitat de la tecnologia i oferia un bon exemple sobre les limitacions dels judicis que no consideren els efectes de les seves *eines* en la societat i del perill que representa aquest comportament per la desresponsabilització que implica.

Val a dir que, tot i que moltes veus van assegurar que l'estratègia de segmentació que va seguir el partit republicà —a partir de les dades filtrades a Cambridge Analytica de Facebook— va resultar determinant en el resultat electoral, també n'hi ha que ho qüestionen. En qualsevol cas, la victòria de Trump hagués estat impossible si no hagués comptat amb un ampli suport de la població nordamericana, un extrem que pot semblar increïble als sectors que s'identifiquen amb els valors dels demòcrates, en aquell moment liderats per Hillary Clinton. I diem incredulitat perquè en les ‘bombolles’ virtuals dels partidaris de la candidatura demòcrata, plens de crítiques i mofes a declaracions extravagants o les incitacions a l'odi proferides per Trump, que el republicà tantes vegades caricaturitzat comptés amb tant de suport real resultava inversemblant.

Aquestes —suposades— bombolles virtuals són un altre dels problemes de la mediació algorítmica que ha estat àmpliament assenyalat i es tracta del fenomen que pretén explicar la teoria del *filtre bombolla* que va desenvolupar l'investigador Eli Pariser. Aquesta conclou que la mediació algorítmica de les plataformes distorsiona la percepció de la realitat perquè limita l'exposició a diferents punts de vista i això, al seu torn, té el potencial d'incentivar la polarització social. Les esmentades distorsions limitarien l'accés a la ciutadania a fluxos d'informació diversos i en conseqüència necessaris per prendre decisions equilibrades i per a la convivència, requisit indispensable per a poder desenvolupar opinions i creences —i per tant actuar— a les democràcies participatives (Arendt, 2012; Habermas, 1981).

Pariser conclou que els algoritmes, en aprendre dels comportaments passats de l'usuari, continuen mostrant el tipus de continguts que més sol consumir. Com veurem més endavant, després de les investigacions de Pariser s'ha produït força recerca que contradiu, però sobretot matisa, la teoria del *filtre bombolla*. A més d'aquestes esmenes, per poder aproximar-nos a una avaluació sobre el paper dels algoritmes en la construcció de l'opinió pública, també convindrà considerar que en la teoria de Pariser la responsabilitat descansa en la gestió algorítmica de la informació que promou la personalització del contingut.

No obstant, per evitar un excessiu paternalisme que no ens ajudaria en l'anàlisi, defensarem que també caldrà fixar-se en el comportament de les persones i en la seva pròpia responsabilitat, perquè malgrat tot, avui resulta molt difícil assegurar si estem més en bombolles que abans, quan només consumíem la informació a través dels mitjans de comunicació que més sintonitzaven amb la nostra visió del món. El que desenvoluparem després, en síntesi, és que el fet que acabem —en major o menor grau— en bombolles informatives, té el seu principal origen en la pròpia natura humana, i més concretament en el biaix de confirmació que descriu a fons Daniel Kahneman (2020) i que es pot resumir en la tendència a consumir aquella informació que referma els nostres punts de vista.

Una altra qüestió que abordarem serà l'*engagement*, la lògica de la distribució algorítmica que també acaba condicionant el contingut que consumim i els nostres comportaments, i que es relaciona amb una de les idees fonamentals que defensem, i que es tracta del paper actiu i determinant de la mediació algorítmica de les grans tecnològiques en la construcció de la cultura.

Tenint en compte que el model de negoci d'aquestes plataformes es basa en la seva capacitat d'aconseguir l'atenció de les persones, una de les principals mètriques que guien la seva activitat és el temps que els usuaris passen a les seves pàgines. Així doncs, si es tracta d'incentivar aquest indicador, els dissenyadors buscaran les maneres de fer que els algoritmes que regulen les línies de temps de les xarxes mostrin els continguts que promoguin aquest comportament de navegar més temps pels seus dominis. Les variables que es consideren en el disseny d'aquestes IA a l'hora de construir línies de temps el més atractives possibles per a cada persona usuària són moltes, i algunes conegudes: bàsicament prioritzen publicacions en funció de comportaments passats de l'usuari, com per exemple amb quines persones ha interactuat més, a quina mena continguts reaccionen més els seus contactes o quins són més compartits, quins obtenen més comentaris i, en darrer terme, més m'agrades o altra mena de clics. És a dir, la unitat de mesura és la interacció i aquesta està determinada per la quantitat.

En conseqüència, la funció cultural de connectar-nos amb la informació segueix una lògica basada en una aproximació computacional a les relacions humanes que, orientada a objectius comercials, mesura a través del tipus i la quantitat d'interaccions allò que mereix ser destacable. Allò que Langloise ha definit com la "gestió de graus de significació i atribució de valor cultural" (citada per Finn, 2018, p.203).

Aquests incentius fan que les persones, les organitzacions de tota mena així com els mateixos mitjans de comunicació, moguts per aconseguir interaccions —que els algoritmes tradueixen en visibilitat i prestigi— tendeixin a compartir la mena de publicacions que generen aquestes reaccions mesurables —i no d'altres—. Avui tendim a consumir (i a compartir) contingut que aconseguix captar la nostra atenció i retenir-la, que sol ser més visual i emocional (J. Berger, 2014). Així doncs, aquestes gramàtiques acaben encoratjant determinats comportaments que es tradueixen en els continguts cridaners que acaben dominant l'ecosistema digital i fins i tot comportaments que tenen lloc al món físic: des de les selfis i el *postureig* arreu fins al disseny d'espais físics per ser fotografiats i compartits (Newton, 2017).

Així doncs, són valors comercials els que acaben condicionant aquesta nova epistemologia algorítmica que determina allò que ha de ser conegut. El problema, en conseqüència, no és només l'aparatososa aproximació computacional a la realitat (Finn, 2018) de determinades aplicacions de la IA, sinó que aquesta, a més, estigui guiada per principis que responen exclusivament a objectius de negoci, perquè aquests promouen la mena de continguts més *populars*. Es tracta de missatges que, fonamentats en l'optimització de la difusió basada en les

reaccions biològiques humanes que es mesuren en interaccions, podrien implicar una nova pujada de voltatge emocional en relació al panorama que Vidal i Castell (2005) descrivia en el a principis dels anys 2000 per reclamar absència del missatge original i presencial. L'autor es referia a la societat mediàtica en què van irrompre les xarxes socials on tots els mèdia —liderats per la televisió— intensificaven el missatge per atenuar el “defalliment de l'actitud d'escolta” (p.85). Es feia amb el directe, la visualitat o els *reality show*, que avui podrien estar representats i magnificats per l'exposició de la vida privada que les persones retransmeten a través dels seus perfils a les xarxes socials per aconseguir aquestes reaccions —i per tant popularitat—.

Andrew Feenberg (2002) recorda que el fet que no tots els béns poden trobar lloc al mercat demostra que aquest no és neutre. En el cas que tractem, això es materialitza en una ocultació dels temes que no són virals i que desapareixen en aquesta lògica de mercat aplicada a les xarxes socials. Feenberg convida a preguntar-se, en aquests casos, qui se'n beneficia, quins efectes té sobre la resta o quina forma de vida promou. El filòsof recupera també la màxima *el mitjà és el missatge* de McLuhan per establir el seu marc crític:

The other style of critique argues that "the medium is the message," that the media distort the contents they express. For example, not every good can find a place on the market. Markets are not therefore neutral arbiters of the community's values but prejudice choice wherever they are instituted. (p.163)

Ara bé, de la mateixa manera que s'estableix una relació simbiòtica entre la personalització que caracteritza la distribució de les plataformes i el biaix de confirmació humà, per valorar l'abast de la responsabilitat de les plataformes, també caldrà considerar que la mencionada lògica de l'*engagement* té també una forta base en certes febleses humanes com l'egocentrisme o la necessitat d'acceptació. És a dir, correriem el risc de perdre perspectiva si partíssim de la base que, per dir-ho de manera simplificada, els selfis són responsabilitat d'Instagram. Per tant, en el nostre marc considerarem així mateix que la tecnologia també la definim nosaltres, les persones, a través dels nostres usos i accions (Casacuberta, 2021).

Un dels principals problemes que afectaven l'esfera pública en l'època predigital era la mala qualitat de la informació a la qual la ciutadania tenia accés. Aleshores això es produïa com a conseqüència del fet que la gestió d'aquesta informació estigués en les poques mans d'uns mitjans tradicionals moguts, a més, per interessos comercials i polítics que els restaven la independència necessària (Habermas, 1981). La democratització de la informació que implica internet havia de ser un revulsiu a aquesta situació, tanmateix, el fet que siguin aquestes grans plataformes les principals mediadores (Newman, 2022) ens situa en un escenari que en certs

aspectes podria assimilar-se més al passat del que es proclama a les missions de les *Big Tech* de les que emana un suposat compromís amb la llibertat d'informació.

Com dèiem en apartats anteriors, l'espai il·limitat característic d'Internet unit a les eines tecnològiques que proporcionen les plataformes i altres serveis, així com a la proliferació de dispositius mòbils, ha facilitat que qualsevol pugui disposar d'un canal de comunicació. Tanmateix, que tothom pugui tenir veu i generar continguts sense fi ha col·laborat també a seguir inundant d'informació l'esfera pública virtual i ha fet més difícil gestionar el coneixement de manera constructiva. I és que tot i que amb cada impuls tecnològic disposem d'accés a més informació, això no sembla que estigui contribuint en la mesura de les expectatives al progrés social o a solucionar grans problemes, com alguns autors han anat anunciant i teoritzat abastament. El volum de dades és de tal magnitud que lluny d'ajudar-nos a evolucionar sembla més aviat que tingui el potencial de convertir-nos en societats menys crítiques (Huxley, 2005; Lassalle, 2019; Postman, 2012; Ruiz-Caballero, 2016; Vidal i Castell, 2005).

L'escenari actual complica encara més el reeiximent del silenci creador que, escriu Vidal i Castell (2005) “ens aparta de la dimensió més profunda de nosaltres mateixos” (p.86) i ha de ser part de l'antídote contra el problema comunicatiu que aquest autor aborda des de l'anàlisi de la crisi de la paraula —la distància entre paraula i experiència— que afecta la capacitat lingüística de l'ésser humà i la qualitat de la mediació del periodisme:

L'excés de paraules, la inflació de significants que circulen, és una de les dimensions de la crisi de la paraula moderna i postmoderna que afecta alhora la nostra parla provada i pública i els usos que fan dels recursos lingüístics els mitjans de comunicació (pp.71-72).

Vidal i Castell reivindica que perquè hi hagi “experiència literària (...) caldrà desinflar aquest gran globus de paraules en què vivim, i recuperar el silenci actiu de la creació” (p.81). Sobre això advertia també Neil Postman a través de la descripció visionària del seu *Tecnopoli*. A la seva obra, reflecteix les conseqüències culturals d'aquest context de sobrecàrrega informativa (*information overload*). Postman escrivia que aquesta situació no ens permet centrar l'atenció i ens paralitza, perquè genera desajustos entre el temps que dediquem a certes qüestions superficials i els nostres propis objectius com a persones i com a societat. Alhora, aquestes quantitats de continguts creen una falsa percepció d'estar molt informats i per tant som inconscients d'aquesta paràlisi (Postman, 2018).

Més amunt ens referíem al problema que assenyalen autors com Lluís Duch o Josep Maria Esquirol sobre el fet que el llenguatge que s'atribueix a la racionalitat havia anul·lat altres narratives transcendents que proporcionaven bases morals i donaven respostes a les grans preguntes de la vida. Postman reivindicava, en un sentit similar, que a aquesta necessitat de referències morals s'hi sumava la manca d'institucions fortes capaces de generar riquesa cultural, o el que és el mateix, de controlar, en benefici del bé comú, l'allau d'informació que produeix la tecnologia. El sociòleg i crític cultural considerava que, en especial des de la irrupció del telègraf, la informació resultava cada cop menys útil (2018):

Pero el genio que salió de la botella anunciando que la información era el nuevo dios de la cultura era un embaucador. Resolvió el problema de la escasez de información, cuyas desventajas eran evidentes, pero no advirtió contra el riesgo de un empacho de información, cuyos peligros no son tan fáciles de identificar.

(...) la información se convierte tanto en el medio como en el fin de la creatividad humana. En el Tecnopolio nos vemos forzados a dedicar nuestras vidas a la búsqueda del "acceso" a la información. Con qué fin o hasta qué extremo, son cosas que no se espera que preguntemos, y que tampoco estamos acostumbrados a preguntar, ya que el problema no tiene precedentes. (pp. 90-91)

Són paraules escrites fa gairebé vint anys que presagiaven el problema de les nostres societats que avui s'ha convertit en endèmic. Una de les conclusions que extreu l'últim informe del Reuters Institute de la Universitat d'Oxford sobre usos i tendències de consum d'informació, és que als segments de població més jove els interessien cada cop menys les notícies. Els dos motius principals són perquè es parla massa de política (i de la Covid-19 en context pandèmic) i per la negativitat que emanen i que influeix de manera negativa l'estat d'ànim. A aquestes raons se li sumen dues més que assenyalen perquè podrien ser indicadors reals d'aquesta teòrica saturació informativa. Les persones joves argumenten que també eviten les notícies perquè hi ha massa informació disponible i perquè aquesta no els hi resulta útil ni les condueix a l'acció (Newman et al., 2022).

Harari urgeix a posar-hi remei des de l'educació, que la ciutadania necessita "la capacidad de dar sentido a la información, de señalar la diferencia entre lo que es y no es importante y, por encima de todo, de combinar muchos bits de información en una imagen general del mundo" (Harari, 2018, p.287).

El problema agafa grans dimensions des de la perspectiva de Lassalle (2019), qui considera que el problema de la sobreinformació és una de les principals amenaces per a les democràcies. Per

a aquest autor, la quantitat de dades que circulen i el volum d'estímul als quals estem exposats ens han incapacitat com a subjectes polítics i, com a resultat, les nostres democràcies liberals han perdut legitimitat:

El estrés de datos en tiempo real es tan intenso y poderoso que va anulando nuestra capacidad de respuesta bajo un síndrome de infoxicación que afecta las capacidades epistemológicas y sensibles del ser humano y lo cuestiona, incluso, como sujeto político.

(...) Una ciberadaptación que altera su forma de ser y estar en el mundo. Ambos aspectos percuten también sobre su forma de vivir la política y explican no solo el auge de los populismos sino la crisis de la democracia liberal. (pp.38-39)

Avui també agafa relleu el concepte economia de l'atenció, que explica que com més informació existeix, menys atenció disposem. Aquesta teoria va ser formulada per primer cop el 1971, quan encara eren els mitjans de comunicació de masses els qui dominaven l'esfera pública. L'economista i expert en psicologia social Herbert Alexander Simon (1971) l'explicava així:

In an information-rich world, the wealth of information means a dearth of something else: a scarcity of whatever it is that information consumes. What information consumes is rather obvious: it consumes the attention of its recipients. Hence a wealth of information creates a poverty of attention and a need to allocate that attention efficiently among the overabundance of information sources that might consume it. (p.40)

La nostra atenció (i per tant les dades que generem en navegar pels llocs web que la retenen) s'ha convertit en un dels principals béns econòmics fins el punt que aquelles empreses que són més solvents en retenir-la figuren entre les més valuoses del món (Mazo, 2020). Gràcies a aquest temps, poden exposar-nos a més anuncis i alhora recol·lectar més dades personals que utilitzaran per generar perfils digitals de cada persona —dades que també recullen de la nostra navegació en altres pàgines— (Kantrowitz, 2018). Lassalle (2019) parla dels usuaris com a treballadors no assalariats: “El problema de este cambio es la feudalización de los usuarios. Se les permite disponer de los algoritmos y beneficiarse de ellos, pero a cambio de someterse sin discusión a un poder que impone los sesgos que introduce a partir de los propios datos que los propios usuarios ofrecen con su rastro cibernético” (p.96).

Precisament la desconfiança social cap a les plataformes de la qual parlàvem més amunt té a veure en gran part amb els mètodes que aquestes grans tecnològiques fan servir per aconseguir mantenir-nos-hi enganxats —amb els problemes d'addició que també genera— (Harris, 2017; Patino, 2020; Wu, 2020) , així com amb aquest model de negoci que no expressa de manera

transparent fins a quin punt les persones queden exposades en gaudir dels seus serveis (Zuboff, 2019).

3.2.2.1 Engagement com a valor suprem de la mediació

En el context d'excés de continguts al qual ens referíem abans, els algoritmes que ens ajuden a trobar i ordenar els continguts han esdevingut una eina principal tant per establir relacions com per obtenir informació de tota mena i en formats per a tots els gustos. Hem integrat els nous mitjans a la vida quotidiana fins el punt que passem una mitjana de dues hores i mitja al dia a les xarxes socials (Hootsuite i We Are Social, 2022) al mateix temps que Google rep 6.000 milions de cerques al dia (*Internet Live Stats*, 2022).

A la seves capacitats de filtratge i jerarquització es suma un disseny en què tots els elements de les seves interfícies estan pensats per optimitzar l'experiència de l'usuari de manera que aquest hi passi temps i decideixi tornar. Però si han aconseguit aquesta fidelització (que mesuren amb l'indicador usuaris únics actius) no ha estat només gràcies als continguts o a una aparença estètica, sinó que també s'han servit de tècniques de l'anomenat disseny persuasiu que consisteix a aplicar a les interfícies elements que engeguen mecanismes mentals de satisfacció i d'altres que tenen la finalitat d'obtenir recompenses. És a dir, fent servir mètodes que generen addicció, que es manifesta en la incapacitat per controlar el propi consum. De manera indirecta, aquesta dependència acostuma a causar tensions en les relacions interpersonals i d'altres comportaments problemàtics que finalment deriven en respostes emocionals disfuncionals i per tant disminuint la qualitat de vida de la persona addicta.

Aquest aspecte ha estat també àmpliament abordat (i denunciat) per part d'expertes, començant per alguns dels mateixos creadors d'aquestes eines peneditis. Ja s'han produït diferents casos d'ex treballadors que han mostrat públicament la seva preocupació per haver participat en el desenvolupament de projectes d'aquest tipus (Winnick, 2016). Segons han explicat alguns d'aquests testimonis, les plataformes mantenen les persones usuàries navegant en els seus entorns aprofitant les debilitats humanes com la necessitat d'afirmació social, que es compensen amb els *xuts* de dopamina que proporciona cada *like*, comentari o compartit, sigui quan el fem o quan ens el fan.

Aquesta *tecnologia de l'engagement* té el seu origen a l'any 2007, quan un grup d'enginyers de Facebook va crear el botó *m'agrada*, la funcionalitat matriu que articula la lògica del negoci de l'*engagement*: amb un sol clic es podien enviar petits bocins de positivitat per la plataforma. Un dels membres de l'equip creador del botó (un d'aquests testimonis d'ex treballadors) explicava que la innovació va ser un èxit absolut: els usuaris experimentaven la càrrega d'afirmació social

cada cop que feien o els hi feien un *m'agrada*. Això va tenir dues conseqüències fonamentals per a la definició d'allò que són encara avui les xarxes socials digitals: gràcies al botó *m'agrada*, l'*engagement* i, en conseqüència, el temps que les persones passaven a Facebook va augmentar exponencialment (P. Lewis, 2017).

Alhora, Facebook havia descobert allò que es convertiria en la seva mina d'or, perquè cada *m'agrada* els revelava què interessava a la gent: quins continguts, quines persones, quins productes o grups...etcètera. Dit d'una altra manera, Zuckerberg va trobar la forma d'obtenir la informació que, convertida en dades, li permetria, per un costat, posar a disposició dels anunciants una de les eines de segmentació d'audiències més potents del món per arribar als seus públics objectius i, per l'altre, alimentar l'algoritme que personalitzava cada cop millor les línies de temps dels seus usuaris.

Els factors que acostuma a tenir en còpte l'algoritme per organitzar les línies de temps els defineix el propi comportament de la persona. Si fa *m'agrada* en els continguts d'un perfil determinat, veurà més publicacions provinents d'aquest compte. Si una publicació donada obté moltes interaccions en poc temps, l'algoritme prioritzarà aquella publicació perquè interpreta que serà més suggerent. Els compartits tenen més valor que els comentaris, i aquests dos valen més que els *m'agrada*. En síntesi, com més probabilitat hi hagi de que la persona usuària interactuï amb un contingut, més destacat estarà a la línia de temps *personalitzada*. Els algoritmes de les plataformes consideren milers de variables a l'hora de dissenyar les línies de temps a gust de l'usuari (Newberry, 2022; Stelzner, 2017). La tendència que marquen els darrers canvis d'algoritme anunciats per les plataformes suposa que es tendirà a donar més pes que fins ara a missatges en base als interessos de l'usuari que la xarxa infereix, així com a continguts populars a la plataforma per sobre del les publicacions provinents de contactes («Introducing Home and Feeds on Facebook», 2022). Aquestes modificacions poden suposar un canvi de concepte de la xarxa social tal com s'entenia fins fa molt poc.

Instagram, Facebook, Twitter, Google o Netflix competeixen per aquest bé escàs que és l'atenció humana i ho fan amb equips en els quals s'asseguren de comptar amb algunes de les millors professionals, que treballen enfocades en aconseguir augmentar la mètrica principal, que és el temps dels usuaris al seus dominis. Per assolir-ho faran servir totes les tècniques de disseny disponibles i aquestes inclouen l'esmentat disseny persuasiu, una branca de la creació d'interfícies que es serveix de tècniques clàssiques de suggestió de la ment humana per captar l'atenció i propiciar accions compulsives que condueixen a entrar al seus webs i, un cop dins, a

passar-hi més temps. I no es tracta només d'una cursa entre plataformes, sinó que aquestes competeixen fins i tot contra les necessitats humanes més bàsiques. Així ho va afirmar el CEO de Netflix, Reed Hastings, que va explicar a una de les presentacions de resultats del 2017, que un dels competidors de Netflix era la son (Hastings, 2017).

Una de les tècniques clàssiques de què es serveix el disseny persuasiu és la *Llei del reflex condicional*, una hipòtesi més coneguda per la *teoria dels gossos de Pavlov*: El cervell dels animals va acabar associant un mateix soroll al menjar que els servien cada cop que els hi feien sentir, de manera que en escoltar-lo reaccionaven com si es tractés del propi menjar. Això mateix, traspassat als entorns digitals, suposa que relacionem les notificacions de les xarxes socials a l'experiència satisfactòria que sol implicar el missatge d'una amiga, una reacció d'acceptació i altra mena d'estímul positius que provoquen en els humans les interaccions socials (Krach et al., 2010) i que els investigadors expliquen com un mecanisme positiu que té la funció de generar i intensificar es vincles socials. El nostre cervell segrega dopamina, —la coneguda substància que alliberem quan tenim experiències satisfactòries— i activa el sistema de recompensa per tal que repetim un comportament, en principi, beneficiós per a nosaltres.

El disseny persuasiu de les plataformes es serveix també d'un altre potent esperonador d'impulsos humans i resulta ser el mateix mecanisme d'addicció que utilitzen les màquines escurabutxaques (Madrigal, 2013). Es tracta del sistema de recompensa variable, un comportament que va ser estudiat en un experiment conegut com *la caixa de Skinner*. El psicòleg Burrhus Frederic Skinner va fer aprendre a animals tancats en una caixa que cada cop que premissin una palanca obtindrien menjar i, en una següent fase de l'experiment, va comprovar que si els animals no obtenien menjar com a recompensa totes les vegades, aquests es dedicaven a prémer la palanca de manera compulsiva. És a dir, ho feien en més ocasions i durant més temps que si obtenien el menjar sempre. Així, el sistema de recompensa variable es va provar més eficient per induir al comportament compulsiu de prémer la palanca.

En aquest cas la dopamina també hi juga un paper important, perquè n'alliberem més si existeix un cert grau d'incertesa sobre l'obtenció de la recompensa de la nostra acció, segons el professor de biologia i neurologia de la Universitat de Stanford Robert Sapolsky. Els treballs d'aquest investigador conclouen, a més, que aquesta substància química es segrega en el moment de l'estímul i no pas en obtenir la recompensa, perquè la seva funció és proporcionar-nos energia per aconseguir-la (Sapolsky, 2012).

Al capdavant, aquesta mena de tècniques consisteixen en introduir al cervell de la persona l'expectativa que el fet d'entrar a les plataformes li produirà una gratificació. Això pot aplicar-se també a l'acció de revisar aquestes notificacions per verificar si ens hauran fet m'agrada, ens hauran etiquetat en una foto, o fet un retuit; aquest mecanisme també està present en l'acció de refrescar l'e-mail o al *scroll* infinit de les línies de temps, on la promesa pot consistir en trobar nova informació que ens aportarà més coneixement.

L'autoreproducció dels vídeos és un altre exemple de disseny persuasiu: la psicologia diu que serà més probable que veiem el vídeo si no hem de fer l'acció de prémer el botó *reproduir*. Aquesta és la raó per la qual a Netflix i altres serveis s'autoreproduïen els següents episodis de les sèries o a Instagram no calgui iniciar l'acció de reproduir per veure els vídeos.

Tenint en compte que totes les plataformes competeixen per la nostra atenció, un cop es desenvolupa una nova funcionalitat que genera més temps a la pàgina, les altres també la implementen. Les conseqüències d'aquestes pràctiques que optimitzen només amb criteris de negoci és la percepció que internet és cada cop més absorbent. Les nostres ments humanes estan *hackejades* quan acabem donant més temps navegant del que voldríem. L'especialista en disseny ètic Tristan Harris (2017) explica que es tracta d'un disseny que es centra en necessitats immediates i no en aquelles qüestions que realment ens importen quan hi pensem. Harris, per cert, és exdissenyador de Google i un altre cas d'aquests extreballadors de Silicon Valley penedits.

Harris està convençut que el problema de l'escassetat d'atenció està a l'arrel de la desconexió entre la quantitat d'informació i assoliment de les finalitats humanes (el problema al qual es referia Postman). Harris es dedica avui a l'activisme en favor d'una tecnologia més humana i llença aquesta advertència sobre els riscos que pot implicar no poder recuperar l'atenció individual i col·lectiva: "There's nothing in your life or in our collective problems that does not require our ability to put our attention where we care about. At the end of our lives, all we have is our attention and our time" (Harris, 2017).

Fa 10 anys, li van preguntar al professor de Harvard Edward O. Wilson si la humanitat seria capaç de solucionar les crisis que l'esperaven al futur, i va contestar que així seria, però només si érem honestos i intel·ligents, i va afegir: "The real problem of humanity is the following: We have Paleolithic emotions, medieval institutions and godlike technology". Harris observa que, des dels dies en què Wilson va fer aquelles declaracions, la tecnologia ha evolucionat exponencialment,

mentre que els impulsos paleolítics del nostre cervell s'han mantingut igual (Harris, 2019). Yuval Noah Harari (2018) es refereix també aquestes conseqüències:

La tecnología no es mala. Si sabes lo que quieres hacer en la vida, quizás te ayude obtenerlo. Pero si no lo sabes, a la tecnología le será facilísimo moldear tus objetivos por ti y tomar el control de tu vida. Sobre todo porque la tecnología es cada vez más sofisticada a la hora de entender a los humanos, por lo que puedes verte sirviendo la cada vez más, en lugar de que ella te sirva. ¿Has visto esos zombis que vagan por las calles con la cara pegada a sus teléfonos inteligentes? ¿Crees que controlan a la tecnología, o que esta los controla ellos?. (p.293)

Per a Byung Chul Han, aquesta lògica de l'*engagement* resulta perillosa perquè acaba amb la dimensió social en posar l'ego al centre, i això fa que creixin la soledat i l'aïllament (Han, 2020). Que contínuament haguem de definir-nos a les xarxes, explicant les nostres preferències o la nostra vida personal aboca a practicar el culte a un mateix, diu (Han, 2020a). L'emocionalitat que demana la semàntica de l'*engagement* encoratja actituds exhibicionistes, ja que com més humans ens mostrem, més emoció es genera (més m'agradaes, comentaris, etcètera.). A banda de les qüestions relacionades amb l'extractivisme de les dades que tractarem més endavant, aquesta promoció de l'exhibicionisme contribueix a seguir difuminant els límits de la privacitat que els mitjans audiovisuals van iniciar amb l'espectacularització de la vida privada que es manifesta als *reality show*. En aquest sentit, escriu el doctor en comunicació i professor de periodisme Carlos Ruiz (2016):

Las redes sociales constituyen la última estación de ese largo viaje hacia la desaparición de la privacidad. Su éxito obedece en parte a que es el propio usuario quien les suministra gratuitamente los contenidos, que en gran medida reflejan aspectos de su vida privada. (p.126)

Probablement la paraula *comunita'* sigui una de les més importants per a Mark Zuckerberg, ja que sempre l'ha pretès vincular als seus productes. Aquest concepte adquireix un significat molt diferent en el pensament de Han, qui opina que els lligams que genera la comunicació digital són febles: "Todo el mundo practica (...) la adoración del yo. Por eso digo que los rituales producen una comunidad sin comunicación. En cambio, hoy prevalece la comunicación sin comunidad. Cada vez celebramos menos fiestas comunitarias. Cada uno se celebra solo a sí mismo". Aquest filòsof defensa que la comunicació virtual és extensiva, mentre que intensiva només ho és la presencial i reclama noves maneres de relacionar-nos motivades per altres valors: "tenemos que inventar nuevas formas de acción y juego colectivo que se realicen más allá del ego, el deseo y el consumo, y creen comunidad" (Han, 2020b).

Les enormes inversions que les plataformes han fet en tecnologia i en disseny s'han vist més que compensades: tant Alphabet com Meta es situen avui, com dèiem, entre les empreses més valuoses del món. A més, les tendències de consum indiquen que encara anirà a més: un estudi revela que, en gairebé el 50% dels casos, per trobar resposta a les cerques fetes a Google, ja no cal sortir del seu entorn (Fishkin, 2019).

3.2.2.2 Cambres d'eco computacionals

Negroponete (1997) creia que internet afebliria la cohesió social perquè els individus substituïrien el consum d'informació de masses de les televisions i els diaris per continguts més personalitzats. Les persones llegiríem el nostre *Daily me* (Haggood, 1995). Negroponete es referia a l'acció de buscar i consumir els continguts que confirmen i refermen la pròpia opinió. Un comportament humà que en cap cas han originat les plataformes digitals i que s'explica a través del biaix de confirmació: la tendència general a buscar o seleccionar informació que confirma els nostres prejudicis i que alhora rebutja aquelles informacions que els contradiuen (Kahneman, 2020). Negroponete veia ja llavors alguns aspectes força encertats sobre com la tecnologia podia condicionar el consum de continguts avui:

(...) la existencia del formato digital cambiará el modelo económico de selección de noticias; hará que nuestros intereses particulares desempeñen un papel más importante (...).

¿Qué ocurriría si un periódico pusiera a nuestra disposición todos sus recursos para editar un solo ejemplar? Habría noticias destacadas con relatos (...) relacionados con la gente que conocemos, con las personas que veremos al día siguiente y los lugares a los que hemos de ir o de los que hemos vuelto hace poco (...). Podríamos llamarlo el *Diario Yo*. (p.127)

Més endavant, el professor Cass Sunstein (1999) va desenvolupar el concepte de les *càmeres d'eco*, un efecte produït pels mitjans de comunicació quan la persona s'exposa només a opinions que reforcen el seu punt de vista i per tant només sent la seva pròpia veu, fet que provoca un efecte de rebot de la informació a unes parets imaginàries.

És aquí on radica un dels principals problemes que pot comportar la distribució algorítmica de la informació i que el ciber-activista Eli Pariser va anomenar les *bombolles d'informació* (*Filter bubbles*). Al seu llibre *The filter bubble: what the Internet is hiding from you* (2011) descriu una web optimitzada a les preferències individuals on les persones mai s'hauran de trobar una opinió amb la qual no hi estiguin d'acord.

Pariser explica que va iniciar la seva recerca quan se'n va adonar que, a Facebook, havia deixat de veure publicacions de persones amb afinitats polítiques diferents a la seva. Allò el va portar a investigar i, segons explica, els resultats confirmaven que l'algoritme l'exposava a les opinions polítiques que coincidien amb la seva. Això el va portar a dedicar-se a la divulgació i a reivindicar la necessitat que aquests algoritmes incorporessin una consciència de la vida pública, un sentit

de la responsabilitat cívica, que fossin transparents i, finalment, que donessin més control a les persones sobre els seus processos per poder garantir que a la ciutadania l'accés a noves idees, persones i punts de vista. La periodista Marta Peirano (2019) recrea els perills de la manca de diversitat del consum passiu d'informació a mida que incentiven la filtració i el ranqueig algorítmics:

Los de la bici no entienden a los del coche, los vegetarianos no se hablan con los taurinos. Los de izquierdas ya no pueden compartir ni un taxi con los de derechas sin empezar una furiosa discusión. Ya no tenemos que negociar nuestra visión del mundo con personas que no la comparten porque somos perfectos. La prueba es que hay personas perfectas que comen lo que nosotros comemos y piensan lo que pensamos y tienen la misma edad que nosotros y ven las mismas series y escuchan la misma música y visitan las mismas ciudades. Las tribus identitarias no son un monocultivo; la falta de diversidad atrae plagas y enfermedades. (p.247)

Segons aquests autors, doncs, ens limitem a consumir informació que confirma allò que pensem. El treball de Pariser ha donat lloc durant els darrers anys a força reflexió i recerca en torn als perills de les bombolles d'informació, i avui comptem també amb treballs que matisen aquestes conclusions i en canvi consideren que a les plataformes sí que estem exposats a diferents opinions, com a mínim al nivell que se'ns mostraven abans (Guess et al., 2018; (Masip et al., 2020); Tufekci, 2018). Segons la sociòloga Zeynep Tufeky (2018), el problema és que estar a les xarxes no és el mateix que consumir informació en solitari, llegint un diari, per exemple, perquè durant el consum digital estem connectades amb les nostres comunitats, i això modifica el nostre comportament, que és diferent a quan ens informem fora de línia. En estar connectades amb les nostres comunitats, al mateix temps que consumim, tendim a buscar l'aprovació d'aquestes persones amb sensibilitats similars a les nostres. Tufeky explica que es tracta d'un comportament semblant a animar el nostre equip i cridar als seguidors del contrari que no tenen raó. És una manera de procedir amb la que volem estimular el sentiment de pertinença al nostre grup alhora que ens distanciem de l'altre. Segons la recerca que publica aquesta autora, el nostre univers cognitiu no és una càmera d'eco, però el nostre univers social sí que ho és. Per això, diu, els verificadors no convencen, perquè la pertinència és més forta que els fets.

Altra recerca que limita els efectes de les càmeres d'eco destaca bàsicament tres arguments: en primer lloc, que la majoria de la gent sol informar-se a través de diferents mitjans de comunicació i altres fonts, i que només un petit subconjunt de població obté informació de fonts molt restringides. En segon lloc, argüeixen que gran part de la població no para atenció constant a qüestions polítiques, sinó que ho sol fer només en moments molt determinats. A més, fins i

tot s'identifiquen sectors socials que no s'interessen per la informació política en absolut. En tercer i últim lloc, aquestes investigacions assenyalen la importància del context en què les persones s'exposen a la informació. Si bé les opinions d'amics a les xarxes socials o els rànquings algorítmics poden influir en les persones, aquests efectes són més modestos i contingents del que es pot suposar a priori, i que la nostra permeabilitat a opinions dels altres es multiplica a les xarxes socials presencials on, a més, la nostra exposició a diferents punts de vista és més infreqüent (Guess et al., 2018).

Tot allò que hem exposat no treu que la distribució algorítmica de la informació governada per aquests valors comercials està dissenyada per mostrar les persones aquella mena de continguts dels que més sol veure gràcies al rastre que deixen. Que els algoritmes ens facin propostes en base al nostre consum previ (o als dels nostres contactes) és una forma més que tenen d'influir en la cultura. De nou, però, cal subratllar que allò que condiona la mena de continguts que acabem consumint és el disseny d'aquests sistemes en retroalimentació amb aspectes o vulnerabilitats de la nostra natura humana.

Si bé és cert que les xarxes socials són els entorns sobre els quals més s'ha posat el focus en relació a aquest problema, no són els únics en què existeix el risc de produir-se una distribució esbiaixada de la informació. Això també passa a altra mena de plataformes digitals com Youtube i o a Google, on tampoc existeixen els mateixos resultats per a tothom, ja que l'algoritme del cercador té en compte molts criteris per decidir quins continguts mostrar com a resposta per a una mateixa cerca («Cómo funcionan los algoritmos de búsqueda», 2020).

Les plataformes es dediquen a filtrar i jerarquitzar, i per reduir-nos les opcions han d'aplicar un criteri, i aquest és, bàsicament, l'*engagement* que, com hem vist abans, és l'indicador al que associen el seu èxit comercial. Si els usuaris no som conscients d'aquests processos computacionals optimitzats amb uns paràmetres determinats orientats a assolir fites de negoci, podem pensar que les respostes que ens estan mostrant són segurament les que més s'adeqüen a les nostres necessitats i fins i tot a la veritat. Aquest punt pot veure's influenciat també pel *dataïsme* que, com dèiem, confereix una suposada objectivitat als processos computacionals.

Harari (2018) parla fins i tot de traspàs de poder a la IA i a les empreses tecnològiques. En aquest fragment, es refereix concretament a aquesta influència cultural de l'aplicació de la IA en l'àmbit de la comunicació:

En dos décadas, miles de millones de personas han llegado a confiar en el algoritmo de Google una de las tareas más importantes de todas: buscar información relevante

y fidedigna (...) Y a medida que confiamos más en Google para hallar respuestas, nuestra capacidad para buscar información por nosotros mismos disminuye. Ya hoy en día la “verdad” viene definida por los primeros resultados de Google. (p.75)

Tot i que, com hem vist, avui disposem d'arguments que matisen la teoria del *filtre bombolla*, tampoc queda clar que la quantitat d'eines, dispositius i canals per distribuir continguts hagin contribuït d'una manera significativa a exposar les persones a més diversitat. En conseqüència, si la multiplicitat d'opcions no col·labora a l'exposició a més punts de vista, és probable que la causa no sigui només la lògica algorítmica, sinó que més aviat pot tractar-se, de nou, d'un disseny tecnològic que, orientat a objectius comercials, es fon en una simbiosi amb el biaix de confirmació humana en un cercle viciós.

En qualsevol cas, des del punt de vista del periodisme sí que caldrà parar atenció als usos de la personalització, atès que una de les funcions dels mitjans de comunicació consisteix precisament en el contrari, que és proporcionar espais comuns per tractar els temes d'interès general. Més endavant desenvoluparem aquesta idea.

Com diem, aquest fenomen tipus bombolla també es dona a altres entorns digitals. Un estudi encarregat pel Consell de l'Audiovisual de Catalunya alerta també sobre els riscos d'aquest efecte per a l'autonomia i la diversitat (C. Castillo, 2018). El cas concret que analitza l'estudi és el de l'algoritme de plataformes de vídeo sota demanda com Netflix, però les seves conclusions són aplicables a d'altres suports digitals, perquè les solucions que proposa posen focus a l'important paper que juga també el disseny de les interfícies a l'hora de promoure la diversitat en l'accés als continguts. Castillo explica per què, tot i que com a usuàries ens sembli que tenim poder de triar, això no és així de cap manera. Ho explica l'informe de l'APDCAT sobre IA a Catalunya (Peiró, 2020):

“Potser (l'usuari) té al davant quaranta títols però aquesta és una part molt ínfima del catàleg total. I l'usuari no decideix el criteri pel qual s'han seleccionat aquells 40 títols (...). Si veus pel·lícules d'acció, no et mostra més que això. Per tant, aquí no hi ha diversitat. Però també hi ha risc d'autonomia, perquè també hauries de poder explorar el catàleg sencer. Ara només hi ha una caixeta de cercar, cosa que fa impossible que hi accedeixis al volum total. Podrien haver-hi altres dissenys d'interfície que, per exemple, poguessis fer la cerca parlant. D'igual manera, podrien donar l'opció de decidir que el catàleg estigui ordenat o desordenat.”

La personalització no és només una qüestió relacionada amb la informació. Avui, crear els productes a mida del consumidor és una de les tendències més comunes en les pràctiques del sector del màrqueting. Tot aquest rastre que deixem que després es tradueix en informació que

ens proporciona l'anàlisi de dades, unida a noves eines digitals, han fet sorgir el concepte del *Customer Experience (CX)* (Iglesias, 2018-b). Un enfoc de negoci que redefineix la cultura del màrqueting empresarial i que consisteix, a grans trets, en redissenyar els productes a partir de les necessitats de les consumidores.

Per això es fan servir tècniques com el *customer journey*, (Iglesias, 2018-c) que resideix en dibuixar l'itinerari de la relació entre les empreses i els seus públics. Un cop marcats els punts de contacte, es treballa en la millora d'aquells moments de l'experiència relacional que causen més insatisfacció, segons les dades que es puguin obtenir a través de diferents tècniques i eines analítiques.

El sector dels mitjans de comunicació no és una excepció i també ha començat a fer servir aquesta mena de tecnologies de personalització (Coenen, 2019; De Lara et al., 2022; Guaglione, 2022; Thurman, 2019), entre d'altres coses, perquè els públics també la demanden, en especial els més joves, els que han crescut entre algoritmes de consum cultural com els de Amazon, Netflix o Spotify (Galan et al., 2019). Tenint en compte que una de les principals funcions del periodisme, dèiem, consisteix a proporcionar una visió comunitària sobre els temes d'interès públic, caldrà valorar com integrar aquesta personalització i, sobre tot, fins a quin punt. Aquest és un dels principals temes que examinarem a la recerca empírica que exposarem al quart capítol.

3.2.2.3 Desinformació algorítmica

Les operacions de desinformació mai han deixat de formar part de la teoria de la comunicació (Moragas Spà, 1981), en conseqüència tampoc es tracta d'un fenomen inherent als nous mitjans algorítmics. Tanmateix, amb la proliferació d'eines tecnològiques i nous canals, avui adquireixen noves formes que aprofiten les característiques de la tecnologia actual així com el context social per difondre's de maneres abans impossibles. Tot i això, és difícil dir si hi ha més desinformació ara que abans que existissin les plataformes digitals, perquè la informació que en la gran majoria subministraven els mitjans de comunicació tradicionals, sovint tampoc era d'excel·lent qualitat.

Avui pot semblar que la nova desinformació es tracta d'un fenomen incontrolable, perquè resulta complicat preveure de quines maneres concretes es podrà combatre de forma eficient. El raonament que pensem que pot ser més operatiu adoptar en aquest context i com a punt de partida és que, com tota novetat, encara no és prou coneguda, i per consegüent és lògic que sembli més difícil de dominar. D'entrada, la majoria de la ciutadania encara no disposa de prou coneixements per identificar informació falsejada; de l'altra, les nostres societats tampoc compten amb suficients eines (intel·lectuals, tecnològiques ni legals) per controlar aquestes noves maneres d'enterbolir l'esfera pública. Les parts interessades en desinformar ho saben i tenen un terreny similar al que podia ser *far west* per poder actuar amb impunitat. Potser la sensació d'impotència que provoca la incapacitat d'articular solucions amb la rapidesa que la situació ho requereix sigui part de la raó per la qual avui la desinformació s'ha convertit en una de les principals preocupacions socials i institucionals. Com anem veient, que més eines i més canals portarien més veus i millor democràcia era una esperança lògica, mentre que les conseqüències sobre els usos perversos dels nous algoritmes comunicatius van quedar a l'ombra fins que s'han anat posant en evidència.

Les estratègies de comunicació que es fan servir canvien, però totes elles comparteixen la mateixa finalitat de sempre: persuadir per guanyar adeptes, assetjar oponents i/o influir en l'opinió pública de diferents maneres. Totes, a més, descansen principalment en certeses sobre el comportament humà i les seves vulnerabilitats —informació imprescindible per tal de poder-lo manipular—. Explica el Premi Nobel d'Economia, Daniel Kahneman (2020), que una manera de fer creure falsedats és la repetició freqüent de la mentida, perquè per a la ment humana no és fàcil en determinats contextos distingir la familiaritat de la realitat. En la teoria dels dos

sistemes que desenvolupa al llarg de *Pensar ràpid, pensar despacio*, exposa que al cervell humà li costa massa energia analitzar la veracitat de qualsevol afirmació. Donat el cas, el recurs que fa servir la ment humana és la comoditat, i si entre dues idees una li resulta familiar, triarà aquella. Així, argumenta Kahneman, amb la repetició de conceptes es promourà aquesta facilitat cognitiva, perquè tendim a actuar minimitzant la dificultat a l'hora d'interpretar. Efectivament es tracta també de la màxima de Goebbels, artífex de la maquinària de propaganda Nazi i de la coneguda idea que afirma que una mentida repetida una vegada segueix sent una mentida, però mil vegades es converteix en veritat.

Els biaixos humans també s'aprofiten per maximitzar l'impacte i els resultats de les campanyes de publicitat en les quals, per vendre perfums, per exemple, s'apel·la a aspiracions humanes com el reconeixement social i l'èxit en les relacions sexuals. En el cas dels anuncis de cotxes es ven poder, estatus i tot allò que ho rodeja. Segons Kahneman, per persuadir a través de la facilitat cognitiva, convé que el missatge compleixi el màxim d'aquests requisits: que sigui lògic, que es pugui associar a altres creences o preferències de la persona a qui es vol convèncer i/o provingui d'una font acceptada i confiable.

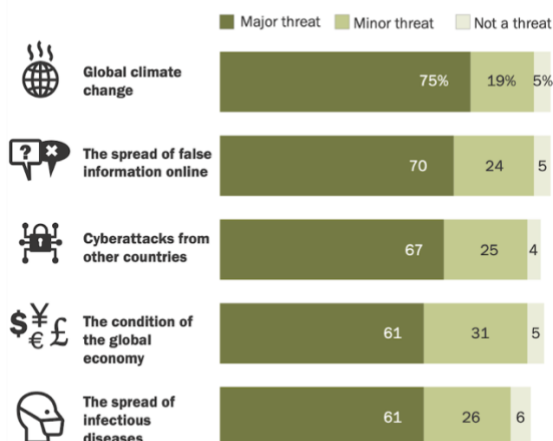
Avui, les mentides no són repetides milers de vegades, però són compartides milers de vegades, i això es produeix com a conseqüència d'una altra vulnerabilitat humana: el biaix de confirmació que hem mencionat abans i que, recordem, es tracta de la tendència natural a identificar i compartir aquella informació que reafirma les nostres creences. A més, el fet —que demostra Kahneman— que podem ser més crèdules quan la informació és compartida per persones en qui confiem o respectem, contribueix a multiplicar el contagi. De nou, la tecnologia actua com a promotora d'uns determinats comportaments. I, puntualitzem un cop més, especialment en el cas de la tecnologia dissenyada amb uns determinats valors comercials —i per tant sota la pressió principal dels objectius de negoci—.

En els darrers anys, el problema de la desinformació ha agafat un gran protagonisme, especialment des del 2015, quan l'ús de la paraula *postveritat* va augmentar un 2000% (Amón, 2016), coincidint amb la precampanya de les eleccions presidencials als Estats Units. Pocs mesos després de la victòria de Donald Trump, *postveritat* va ser declarada paraula de l'any pel diccionari d'Oxford, que la defineix com a relatiu a fets que apel·len les emocions i les creences personals: "Post-truth is an adjective defined as 'relating to or denoting circumstances in which objective facts are less influential in shaping public opinion than appeals to emotion and personal belief'" (*Oxford Word of the Year 2016*, s.d.).

La preocupació per la desinformació s'evidenciava primer des d'àmbits acadèmics i institucionals, però també i de manera creixent, s'ha posat de manifest des dels propis mitjans de comunicació tradicionals. En aquest sentit, abans d'abordar com la tecnologia exponencial pot incentivar la desinformació, pot resultar pertinent també preguntar-se si allò que preocupa no és tant la distribució de falsedats com la descentralització de la informació que aquests *legacy media* han patit (Casacuberta, 2021) amb conseqüències com l'erosió del poder que acumulaven, així com la brutal davallada en el volum de negoci. Considerar si assistim a una criminalització de les plataformes digitals davant "la amenaza de la pérdida del monopolio discursivo de la verdad", tal com assenyala Vidal Castell (2020, p.2591), fet que podria significar una possible sentència de mort per als mitjans de comunicació tradicionals, o el que és el mateix, segons aquest autor: "la industria de la verdad del capitalismo liberal" (p.2593) i la única responsable de la degradació del periodisme.

Té sentit pensar que hi ha una mica de tot plegat, però el que resulta innegable, com veurem també a continuació, és que la desinformació computacional existeix i que la inquietud respecte a aquesta qüestió no ha deixat de créixer fins l'extrem que s'ha estès i ha calat de manera profunda en la societat. En una enquesta recent portada a terme a 19 països, la desinformació suposa la major preocupació social, pràcticament a l'alçada del canvi climàtic (Poushter et al., 2022). Segons aquest informe, ben bé tres terços de les persones enquestades (Figura 1) manifesten aquest neguit. Espanya, segons el darrer informe de Reuters Institute de la Universitat d'Oxford, és un dels països on existeix més preocupació social sobre aquest tema (Vara-Miguel et al., 2022).

Figura 1
La desinformació és una de les principals preocupacions



Font: Pew Research Center

El punt en què ens trobem és conseqüència del fenomen que es va desencadenar, sobretot, l'any 2016. Durant el mandat de Donald Trump (qui es caracteritzà per un estil de comunicació molt diferent del seu antecessor, Barack Obama) no va deixar de créixer la preocupació per la nova desinformació. Els missatges de Trump feien servir un to molt emocional, proferia insults, titllava els mitjans de comunicació i a qualsevol que el contradigués de mentider i feia un ús pervers del terme *fake news*, que atribuïa a les notícies que no li agradaven. Les maneres de fer de Trump van posar en situacions complicades les plataformes, amb qui es trobava permanentment en un estira i arronsa a causa de les pressions que aquestes també rebien com a conseqüència de no frenar els exabruptes de qui llavors era el president dels Estats Units. Després de repetides polèmiques relacionades amb publicacions que podien infligir les condicions d'ús de les xarxes socials, aquestes van optar per marcar determinats continguts de Trump com a possibles falsedats. La fi del seu mandat va acabar també amb la suspensió del seu perfil a Twitter, després que una gentada assaltà el Capitoli com a conseqüència de la crida que va fer en no acceptar el resultat de les eleccions en què va ser derrotat per Joe Biden.

Tot i ser el referent d'aquest histrionisme que tan viral resulta a les xarxes, Trump, com altres personatges públics actuals i històrics, simplement s'ha ocupat de posar la tecnologia al servei dels seus propòsits propagandístics (Tardàguila et al., 2018). El que ha canviat respecte del passat, dèiem, són les tècniques de desinformació. Els actors interessats en manipular l'opinió pública ja s'han preocupat d'explorar els nous *hackejos* que activen els cercles viciosos entre la ment humana i els algorismes de les plataformes, de la mateixa manera que tot sovint han tractat de corrompre els periodistes i els mitjans de comunicació tradicionals. Una de les diferències que defineixen les noves pràctiques és que abans, per dur a terme una campanya de desinformació, era necessari tenir poder polític o econòmic per influir en els mitjans tradicionals; avui, en canvi, falsejar notícies o orquestrar una operació d'aquestes característiques està a l'abast de pràcticament qualsevol. És a dir, la tecnologia comunicativa de les plataformes que fa possible que qualsevol tingui veu per denunciar, també facilita que aquest poder es pugui exercir per part de qualsevol per dur a terme aquesta altra mena de propòsits (Wardle i Derakhshan, 2017). Actualment, qualsevol interessat en realitzar campanyes té a la seva disposició potents plataformes publicitàries de segmentació que han estat alimentades amb les dades personals de milers de milions d'usuaris.

Així, els interessats en intoxicar l'ecosistema informatiu han vist clares allò que per a ells poden resultar noves oportunitats que ofereix la desintermediació. Això ha donat lloc a nous tipus de desinformació exponencial, i la gran novetat és la confirmació que aquests nous mitjans no han

ajudat a regenerar l'esfera pública, sinó que han propiciat reptes de natures diferents. A Tik Tok, a Instagram, a Youtube, a Facebook o a Google és fàcil trobar contingut que pot incitar a la violència o a promoure missatges perillosos per la salut pública (Carrie Wong, 2019; Weedon et al., 2017). Ara es poden generar exèrcits de bots per assetjar oponents polítics, escampar desinformació i segmentar-la a aquelles persones més susceptibles de creure-se-la. Avui, en definitiva, vivim sota la influència d'una esfera pública tòxica en la qual els fets, per a moltes persones, són el que menys importa —perquè *el més important* és defensar el propi punt de vista—.

Per bé que la desinformació interessada no sigui cap novetat, actualment es porta a terme de múltiples maneres, i es fa a través de les plataformes, perquè és on les persones passen la major part del seu temps de consum d'informació. Això es fa des d'una comprensió holística del nou ecosistema: de les eines que proporcionen els nous mitjans, les lògiques amb què aquests funcionen, combinades amb el coneixement i la identificació de vulnerabilitats, tant les humanes com les de la pròpia tecnologia.

Segons un informe de l'equip de seguretat de Meta (abans Facebook), un dels mètodes més utilitzats és la injecció a la xarxa de notícies falsejades (Weedon et al., 2017). Es tracta de continguts que pretenen ser factuais, però que contenen imprecisions intencionals que tenen l'objectiu d'exaltar les emocions, atraure l'atenció i/o enganyar. Les campanyes de manipulació que aprofiten les febleses humanes i tecnològiques de les plataformes no només tenen lloc a Facebook. Això ha estat reconegut tant per les principals xarxes socials com per Google (Mars i Giménez Cano, 2017).

Una altra mena de tècnica de desinformació és la falsa amplificació, que consisteix a generar una activitat coordinada de perfils falsos orientada a manipular el debat públic, per exemple, per desincentivar la participació d'aquelles persones amb opinions divergents o amplificant unes veus determinades que afavoreixen el discurs que es vol que predomini sobre els altres, tant a les xarxes socials com als resultats de cerca de Google (Wardle i Derakhshan, 2017).

Malgrat contradir les normes d'ús de les plataformes, l'*astroturfing* (Aburto, 2014) s'ha convertit en un mètode habitual en campanyes de comunicació tòxica. Els casos són més freqüents en l'àmbit polític, però també afecten altres personatges i institucions, com en el conegut *BarçaGate* del Futbol Club Barcelona. El febrer del 2020, la Cadena Ser va destapar que el club de futbol va contractar els serveis d'una agència, que, mitjançant perfils falsos a Twitter, tenia la missió de reforçar la imatge de la directiva del club. La investigació va descobrir que aquests

perfils havien arribat fins i tot a propinar insults als mateixos jugadors de l'equip (Soldevila i Escudero, 2020).

Una altra de les noves pràctiques desinformatives consisteix en la recol·lecció de dades personals per crear llistes d'usuaris amb trets específics similars, que tenen la finalitat d'utilitzar les vulnerabilitats que s'identifiquen en aquestes característiques per impactar amb informació interessada, mitjançant accions de segmentació, a aquells grups més susceptibles a creure-la. Aquest darrer mètode va ser l'utilitzat per Cambridge Analytica, l'empresa britànica que assegura que és capaç d'influir en l'opinió pública mitjançant tècniques psicogràfiques i que va contractar l'equip de campanya de Trump. El paper de Facebook en el resultat de les eleccions va ser considerat per a alguns prou decisiu fins el punt d'haver fet possible que el resultat a favor del republicà en unes eleccions que es van dirimir per només 77.000 vots en tres estats (Martínez Ahrens, 2018).

Però perquè la intoxicació informativa proliferi recordem un cop més el paper que hi juguen els comportaments humans. Tal com posa de manifest un estudi de la BBC sobre notícies falses, sovint es posa el focus en la tecnologia, però la realitat és que si la desinformació es difon és gràcies a la participació de les persones, les qui contribueixen a la viralització de missatges sovint de manera irracional (Chakrabarti et al., 2018). Aquesta investigació conclou que com més disponibilitat de continguts digitals, menys atenció i menys criteri per discernir la informació fiable. Una altra de les conclusions és que, a l'hora de compartir continguts, les persones tenien més en compte el que sentien que no pas allò que pensaven.

Aquest comportament emocional és ben conegut pels actors que pretenen influir en l'opinió pública, que creen narratives i missatges que multipliquen la seva efectivitat quan estan orientats a explotar febleses com les que ja hem comentat o d'altres com la necessitat de refermar la pròpia identitat fins el punt que inclús les persones que acostumen a verificar la informació estan disposades a compartir missatges que les fan sentir identificades, tot i no poder comprovar que prové de fonts fidedignes.

Les xarxes socials han fet possible que el consum d'informació que abans es feia a l'esfera privada ara sigui públic. La velocitat amb què circulen els missatges s'ha disparat a causa de l'acceleració del cicle de notícies i l'accés a dispositius mòbils que possibiliten el consum permanent. Aquesta velocitat també resulta una amenaça per a la informació veraç. El filòsof Byung Chul Han (2021) assenyala aquesta relació: "Hoy las prácticas que requieren un tiempo considerable están en trance de desaparecer. También la verdad requiere mucho tiempo. Donde

una información ahuyenta a otra no tenemos tiempo para la verdad (p.19). A més, els continguts es comparteixen en temps real entre usuaris de confiança, amb perspectives similars, i per consegüent aquestes peces són menys susceptibles de ser qüestionades (Kahneman, 2020; Wardle i Derakhshan, 2017).

Les plataformes, per la seva banda, dissenyen interfícies que afavoreixen la publicació senzilla en formats visuals que en faciliten el consum i amb elements de disseny més cridaners o impactants que incentiven la viralització. El seu objectiu, com explicàvem més amunt, és tenir tota l'atenció possible per tal de poder vendre més anuncis a preus més alts. Més continguts, més consumits, més compartits, més agradats i més comentats. Aleshores, a través de processos iteratius, adapten les seves funcionalitats als usos que els hi donen les persones en un cercle vicios. Així és com els valors en el disseny condicionen la tecnologia, que després influencia la cultura i els comportaments.

Aquesta combinació de característiques tecnològiques també està a l'arrel del problema de la desinformació, perquè a més de facilitar l'accés i la difusió exponencial al desinformador, provoquen una inundació de continguts que disminueix les capacitats d'atenció dels usuaris i els fa més acrítics. Tot això dona lloc a esferes virtuals que exploten la irracionalitat humana. Cicles de consum d'informació en els quals les pulsions emocionals humanes, que prenen forma de m'agrada, compartits o comentaris, fan circular els continguts a gran velocitat a través d'interfícies dissenyades per a l'exponencialitat de la difusió. Llavors, el cicle torna a començar perquè és precisament d'aquesta viralitat que neix de les febleses humanes l'ingredient del qual es serveix qualsevol interessat en la confusió:

The most 'successful' of problematic content is that which plays on people's emotions, encouraging feelings of superiority, anger or fear. That's because these factors drive resharing among people who want to connect with their online communities and 'tribes'. When most social platforms are engineered for people to publicly 'perform' through likes, comments or shares, it's easy to understand why emotional content travels so quickly and widely, even as we see an explosion in fact-checking and debunking organizations (Wardle i Derakhshan, 2017).

Dissenyar una campanya de desinformació passa per conèixer molt bé les febleses humanes: que tendim a compartir aquells missatges que refermen la nostra identitat, ens sorprenen, ens causen emocions que ens exciten (com la indignació) o pensem que pot ajudar a altres (Berger, 2014). L'any 2016, els continguts més virals van ser notícies falsejades (Silverman, 2016).

És tot aquest repertori de missatges i tècniques d'amplificació d'informació enganyosa, la multiplicitat de plataformes i elements de disseny per incentivar la disseminació i la velocitat amb què les persones escampem aquesta desinformació el que ha preocupat darrerament a les institucions pel risc que pugui influenciar els processos democràtics. Avui existeix força inquietud a occident sobre com els qui es consideren els rivals de les democràcies occidentals (com grups terroristes o el govern rus actual) poden arribar a utilitzar la tecnologia per desestabilitzar i aconseguir més poder. La commoció pel referèndum del Brexit, les pràctiques abans descrites durant les eleccions dels EUA o la popularitat de la ultradreta de LePen a França han estat assenyalats com alguns dels exemples del poder potencial de les campanyes digitals de propaganda i desinformació algorítmiques, que s'han convertit en sistemàtiques.

En opinió de Jose María Lassalle, l'allau d'informació sotmesa a la lògica algorítmica que caracteritza l'esfera pública digital posa en qüestió la legitimitat de les democràcies lliberals, ja que els ciutadans, "minorados en el desempeño de su libertad y aturdidos por la sobreinformación de datos" (p.75), no disposen de capacitat crítica per triar de manera informada. A més, les campanyes desinformatives de tota mena com les *astroturfing* abans descrites desconnecten la ciutadania d'una "verdad objetivable" (p.107) i incentiven el relat emocional del qual es serveix el populisme (Lassalle, 2019).

Els algorismes de les plataformes digitals es manifesten incapaços, doncs, de controlar la desinformació, i en el camí a la recerca de solucions per poder desintoxicar l'ecosistema informatiu pot agafar sentit qüestionar-se fins a quin punt les xarxes socials disposen de les condicions (tècniques i humanes) per poder verificar la informació, per començar, a causa de la dificultat ontològica que implica la distinció entre realitat i ficció. I més important, quines serien les implicacions de delegar-los-hi la tasca d'arbitri de la veritat, tal com apunta el filòsof Daniel Innerarity (2018):

Nuestra relación con la verdad —especialmente en la vida política— es menos simple de lo que quisieran los que la conciben como un conjunto de hechos incontrovertibles. No vivimos en un mundo de evidencias, sino en medio del desconocimiento, el saber provisional, las decisiones arriesgadas y las apuestas. La verdad no es lo mismo que la objetividad y la exactitud. Casi nada de lo que decimos o sentimos es "chequeable". (...) En una sociedad avanzada el amor a la verdad es menor que el temor a los administradores de la verdad.

El que sembla clar avui és que les plataformes, subjectes fins ara a regulacions molt diferents a les dels mitjans de comunicació, s'han mostrat incapaces de gestionar l'equilibri entre la llibertat

d'expressió i la d'informació. Per tal de fer front a aquest problema, la Unió Europea les ha incentivats en diverses ocasions perquè tinguin més cura en el control de la informació nociva (Yun Chee, 2020).

I ha estat arran de tots aquestes inquietuds i frustracions que finalment la UE ha arribat a l'acord polític que regularà l'activitat de les grans plataformes de comunicació digital a través de les lleis que explicàvem més amunt —la DSA i la DMA— forçant-les a retre comptes, a permetre l'auditoria independent dels seus algoritmes i d'altres obligacions que, amb tota seguretat, ocasionaran un important canvi a l'ecosistema digital que hauria de millorar els problemes que hem exposat.

En aquest escenari té sentit pensar que trobar fonts fiables entre el volum d'informació de tota mena que circula per la xarxa s'acabi convertint en una necessitat, cosa que clarament interpel·la el periodisme. Aquesta és precisament una de les raons per les quals al principi dèiem que l'activitat periodística segueix sent necessària: per contribuir, en aquest context problemàtic, a la participació informada i lliure de la ciutadania en els processos democràtics. L'exdirector del diari *The Guardian*, Alan Rusbridger (2018) tempteja el sector davant l'emoció que supera la raó:

Los políticos que tienen éxito son los que saben apelar a las emociones. Donald Trump es un fenómeno en este aspecto. Ya no hay eso conocido como hechos. Mis hechos son mejores que tus hechos. No te creas sus hechos... (...) Tenemos que plantearnos si queremos un mundo de hechos para contrarrestar el mundo basado en las emociones, que es un mundo peligroso. Y si deseamos un mundo de hechos, necesitaremos periodistas.

Lamentablement, avui el periodisme no compta amb la confiança de la ciutadania. Espanya es situa entre els 10 països en què menys es confia en les notícies. A més del declivi en la credibilitat, cada cop es detecta menys interès en l'actualitat de la que donen compte els *legacy*. El grup demogràfic que més creix és aquell que formen les persones que ni creuen en les notícies ni confien en els mitjans, tot i que també augmenta el nombre de les que, tot i que s'interessen per la informació, no consideren els mitjans com a font creïble (Vara-Miguel et al., 2022). Alhora, també s'identifiquen senyals de degradació de la confiança en la informació que s'adquireix a través de les plataformes digitals (Mont'Alverne et al., 2022).

Aquest escenari imposa al sector dels mitjans de comunicació un esforç per recuperar la centralitat en la definició de l'agenda. És per això que amb aquesta tesi ens vam proposar contribuir-hi, en especial, a través d'una proposta de gestió ètica de les aplicacions algorítmiques

—que passa per fer-ne un ús responsable i transparent en benefici de la salut de l'esfera pública—. Al proper capítol exposarem el plantejament per una tecnoètica de la comunicació amb el desig de poder col·laborar també a la recuperació de la confiança de la ciutadania en el periodisme de qualitat.

3.2.2.4 Extractivisme que erosiona les llibertats

Cada cop passem més temps als entorns virtuals, a les xarxes socials per relacionar-nos, a Google per cercar informació, a Youtube, a Netflix o a Twitch per consumir vídeo a demanda (*streaming*) a les aplicacions de cites per trobar parella, a les de salut, les dels bancs, dels mitjans de comunicació, les reunions de la feina, formació en línia... Cada cop fem més vida i transaccions en línia, en un procés que va accelerar la pandèmia de la COVID-19 (Juste, 2021). Avui és habitual que tota mena d'organitzacions monitoritzin i emmagatzemin les dades de cerques, de navegació, de salut, les llistes de contactes o les dades d'ubicació, a través de les quals es pot saber on viu i on treballa una persona, on compra, amb qui es relaciona o a on viatja...

La preocupació per la privacitat s'ha manifestat sobre tot des del món acadèmic i activista, però en el transcurs dels darrers anys aquesta inquietud ha experimentat un augment significatiu entre la resta de la ciutadania (Sethumadhavan, 2021). No és casual que en l'estratègia de comunicació d'algunes de les empreses més destacades la privacitat hagi passat a tenir un paper protagonista (Arcos, 2022). Més enllà del comportament de les marques, la tendència també es manifesta en les noves legislacions que es preparen, especialment a la UE, que imposarà limitacions a aquestes activitats de captació de dades (European Council, 2022).

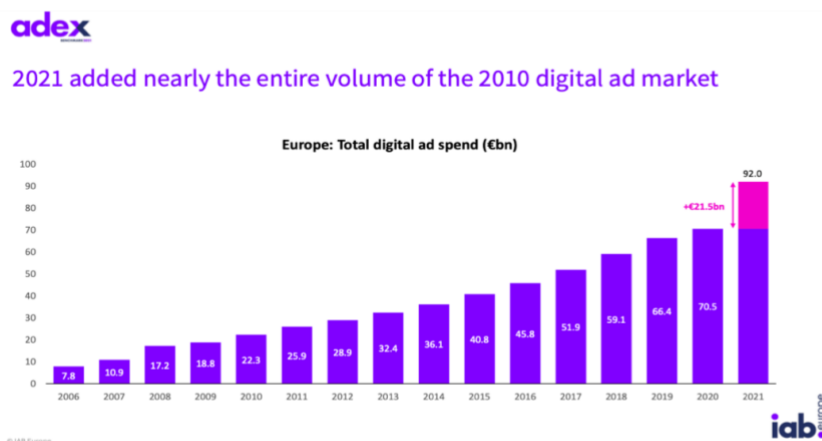
Els sistemes que han desenvolupat les plataformes digitals elaboren perfils de cada usuari amb tota la informació que extreuen de les transaccions i gràcies a tots els sensors dels dispositius, del comportament de navegació o del dels seus contactes. Aquestes tecnologies d'extracció de dades estan en la base de l'anomenat *capitalisme de vigilància* que descriu Soshana Zuboff (2019), i que consisteix en la idea que aquestes empreses han generat un model de negoci que ha convertit el comportament humà en un bé comercial tot i l'amenaça, explica aquesta autora, que suposa per a les llibertats de la ciutadania.

Per a les persones usuàries d'internet el rastre que deixa en forma de dades no resulta evident. Del comportament digital es pot extreure abundant informació personal sobre els comportaments, la personalitat, els gustos, els hàbits, les relacions, les vulnerabilitats o altres qüestions que es refereixen a la pròpia vida personal, i sol ser desconeguda per a moltes persones en els termes que planteja per exemple l'advocada Katarzyna Szymielewicz (2019). Aquesta especialista en Drets Humans i tecnologia, explica que els nostres perfils digitals es constitueixen en tres capes: una primera que conté la informació que compartim de manera

conscient (dades que posem als perfils de les xarxes, missatges privats, esdeveniments als que hem assistit, pàgines web visitades, imatges, allò que ens agrada o compartim). Una segona capa que afegeix altra mena d'informació relacionada amb el nostre comportament que pot ser més delicada de compartir de manera voluntària com localització, hores de connexió, relacions que s'extreuen a partir de la proximitat amb altres dispositius amb els que coincidim, patrons de compra, moviments dels dits a la pantalla (moviments que també revelen emocions i altres trets de personalitat). Per últim, els nostres perfils digitals estan constituïts per una tercera capa que es compon de les inferències sobre la nostra personalitat que els algoritmes extreuen en base a la informació de les dues primeres (per exemple, del fet de compartir imatges de nounats, el sistema infereix que necessitarem comprar bolquers). En aquest punt, apunta l'advocada, ja no es tracta només de què fem, sinó de qui som.

També es podria dir que la tasca d'aquests algoritmes, en part, és predir informació que potser no estariem disposades a compartir i que resulta valuosa per tota mena d'anunciants, perquè els facilita entendre les inquietuds i les necessitats per dissenyar productes o posicionar-los. D'aquesta manera, els resulta més senzill impulsar a prendre decisions (de compra o d'altra mena) cridant l'atenció o apel·lant a pulsions, vulnerabilitats o necessitats humanes. Això no vol dir que totes les empreses o institucions vulguin manipular les persones o que no tinguin ètica, això sí, es tracta de tàctiques que faciliten la venda que es solen aprofitar en més o menys mesura. Prova d'això és l'important augment que ha experimentat la inversió en publicitat digital (Figura 2) en els darrers anys (IAB Spain, 2022).

Figura 2
Evolució de la inversió en publicitat digital a Europa en milions €



Font: IAB Spain

La publicitat basada en el rastre digital es basa en les dades dels usuaris que les plataformes posen a disposició (dels departaments de comunicació o comercials). De la mateixa manera que la publicitat tradicional, aquests utilitzen tota la informació i les noves eines per cridar l'atenció, i això implica activar mecanismes emocionals i reaccions apel·lant a necessitats que de vegades poden ser inconscients. Res sorprenent en relació al que ha fet la publicitat tradicional. Però entre les pràctiques característiques del passat i les actuals existeixen algunes diferències substancials. La primera és que ara hi ha molta més informació disponible de cada individu i pràcticament en temps real. Així, avui, qui ho vulgui, té a la seva disposició eines que poden arribar a ser força precises per a l'explotació de mecanismes subconscients aplicats a dades massives. Szymielewicz s'hi refereix d'aquesta manera, en el context de la publicitat:

Those behavioral predictions and interpretations are very valuable to advertiser. Since advertising is meant to create needs and drive you to make decisions that you haven't made (yet), marketers will try to exploit your subconscious mechanisms and automatic reactions. Since they cannot expect that you will tell them how to do this, they hunt for behavioral data and employ algorithms to find meaningful correlations in this chaos.

No tot allò que té a veure amb l'àmbit del màrqueting digital té aquest propòsit i, per descomptat, l'ús d'aquestes eines i de les dades està condicionat als valors de qui els hi doni ús. La lògica és la mateixa que l'arquitectura de la persuasió del màrqueting tradicional, però la segona gran diferència és que, en aquest cas, els algorismes estableixen correlacions que serviran per inferir amb més claredat com és un determinat usuari, i aquesta informació, que s'associa al seu perfil digital —i que es va enriquint a mesura que es capten més dades—, és la que es farà servir per, seguidament, impulsar-lo a prendre decisions.

En el màrqueting fora de línia, l'arquitectura de la persuasió es fa servir per exemple posant les llaminadures al costat de les caixes dels supermercats per impulsar que els nens les demanin. En digital —i aquesta és a tercera gran diferència— és que aquesta persuasió, gràcies a la computació exponencial, es produeix a escala, perquè no cal que la persona passi pel supermercat, sinó que la publicitat va a buscar cada persona, sigui on sigui, perquè hi arriba a través del seu dispositiu (Tufekci, 2017).

Els darrers temps, però, aquesta qüestió ha adquirit unes dimensions socials que van més enllà del màrqueting, perquè si bé aquestes tecnologies es van crear originalment per una activitat relativament innòcua que és fer publicitat, d'aquestes pràctiques basades en l'obtenció de dades també estan sorgint dinàmiques de gestió social, en què organitzacions diverses tracten

d'obtenir-les per a diferents finalitats: des de la publicitat que pretén manipular —manipulació que cal diferenciar d'una persuasió a priori inofensiva— fins a altra mena d'objectius com controlar i enganyar les persones o els grups com les que hem descrit a l'anterior apartat sobre desinformació: “de maneres noves, amagades, subtils i inesperades”, en paraules d'Evgeny Morozov (2013).

També els governs han exercit una pressió constant sobre les empreses que recullen totes aquestes dades fins el punt de desenvolupar programes d'espionatge basats en l'obtenció d'informació personal, tal com es va poder comprovar arran de les revelacions de l'ex treballador de l'Agència de Seguretat Nacional dels Estats Units (NSA per les seves sigles en anglès), Edward Snowden. Les filtracions de Snowden van demostrar que la NSA utilitzava dades de Facebook o de Google per exercir vigilància sobre tota la ciutadania. Snowden denuncia que s'ha creat un sistema que fa a la població vulnerable per al benefici d'uns privilegiats (Pérez, 2019). En paraules de la filòsofa Marina Garcés (2022) “Un sistema de vigilància i de dominació cada cop més proper a tecnofeixisme” (p.11).

El relat que venen les grans plataformes sobre fer un món més connectat o democratitzar l'accés a la informació, juntament amb l'experiència als cercadors o a les xarxes, produeix en els usuaris la il·lusió que només està consumint o interactuant amb continguts o amb amics, així com la sensació de control sobre la pròpia privacitat. Però el que no resulta evident per als usuaris és que amb el rastre que generen durant les seves navegacions, continuen alimentant bases de dades en què es segueixen construint els seus perfils —i per consegüent participant de manera activa en la variant virtual del feudalisme que descrivíem—. A més, l'activitat de les persones a les plataformes serveix per millorar els algoritmes de manera inconscient, i això tampoc resulta patent ni per a la persona usuària ni per els seus contactes, a qui també implica la pràctica individual. Per exemple, al mateix temps que aquesta persona etiqueta a un amic a Facebook, contribueix a entrenar el seu algoritme de reconeixement facial de manera que aquest podrà identificar l'altra persona amb més seguretat.

“No estic segur si mai ha existit una desconexió tan completa entre allò que una empresa diu que fa —“connectar”, “construir comunitats” — i la seva realitat comercial”, escrivia Aron Gilbreath (2017). Probablement, la confiança que s'hi deposita no es veuria tan erosionada si només fessin servir les dades per desenvolupar els serveis que diuen que ofereixen. Tal com descriu Evgeny Morozov (2013), el problema és que s'utilitzi la informació amb altres finalitats:

A lot of the services they build are useful services. I use Google products all the time. People who are building a service which I pay for with my privacy or money I'm quite okay with. But as time goes by they aspire to do many things that go beyond their business and their initial set of commercial concerns.

Per aquest model, Zuboff (2019) identifica una nova variant del capitalisme que, diu, ha transformat la vida privada de les persones (les dades personals) en *commodities* de la mateixa manera que ho ha fet històricament amb elements de la natura (com les terres) que ha convertit en béns que es poden comprar i vendre. Zuboff explica que el capitalisme, en la seva darrera evolució, ha introduït al mercat de compra-venda l'experiència íntima humana, que és acumulada i processada per les grans tecnològiques per extreure dades de comportament individual que utilitzen per predir el nostre comportament futur:

Surveillance capitalism unilaterally claims human experience as free raw material for translation into behavioral data. Although some of these data are applied to product or service improvement, the rest are declared as a proprietary behavioral surplus, fed into advanced manufacturing processes known as “machine intelligence,” and fabricated into prediction products that anticipate what you will do now, soon, and later. Finally, these prediction products are traded in a new kind of marketplace for behavioral predictions that I call behavioral futures markets. Surveillance capitalists have grown immensely wealthy from these trading operations, for many companies are eager to lay bets on our future behavior. (p.8)

I és que la privacitat és un dret adquirit que està directament relacionat amb les nostres llibertats fonamentals. En paraules de Carlos Ruiz (2016):

La privacidad es un baluarte de la libertad porque implica un límite al poder. Determina una zona personal, física y simbólica, donde el poder no puede entrar, a no ser que esté autorizado por la ley. Es un antídoto contra la arbitrariedad y fue una conquista política esencial de la modernidad, cuando nació la sociedad civil. (p. 124)

Aquest dret es va reconèixer per primer cop en limitar per llei l'accés a la propietat privada física de les persones. El desenvolupament dels mitjans de comunicació, i en especial el naixement de la fotografia i la seva aplicació a la premsa, va implicar nous dilemes que van donar lloc a lleis per garantir les llibertats mitjançant la protecció del dret a no rebre intromissions. Al S.XIX els advocats Samuel D. Warren i Louis D. Brandeis van publicar un article fonamental a Harvard Law Review (Warren i Brandeis, 1890) que va influir el pensament i el dret a la privacitat posterior. L'article reivindicava la necessitat d'adaptar el dret que “en la seva eterna joventut” ha de créixer per satisfer les demandes de la societat.

En aquella època la preocupació era molt diferent, perquè el problema que es pretenia resoldre era la difusió d'imatges personals a través de revistes o diaris, però avui pot tenir sentit recuperar algunes de les idees que el fonamentaven, com la necessitat de garantir la no traïció a la confiança que la ciutadania diposita en altres o en les organitzacions. O garantir que les persones puguin preservar qualsevol activitat de la seva vida privada que prefereixin conservar en la intimitat. I potser caldria afegir-hi, tenint en compte tot el comentat, que sense pagar cap preu a canvi.

Quan es planteja el problema de la vigilància digital, solen emergir dues respostes bàsiques: per un costat, la d'aquelles persones que es preocupen per com els hi pot afectar i procuren limitar la seva exposició, i per l'altre, la d'aquelles a qui els resulta indiferent la vigilància perquè afirmen que no tenen res a ocultar. Com veurem més endavant, és un fet que a aquestes alçades i gràcies a la tasca de divulgació activista, hi ha cada cop més consciència i preocupació per protegir la privacitat.

Amb tot, la inquietud no sembla ser suficient com per deixar d'autoritzar les empreses a monitoritzar la pròpia activitat i a cedir les dades personals. Aquest comportament contradictori ha estat anomenat *la paradoxa de la privacitat* (Gerber et al., 2018) i fa temps que se'n busca l'explicació. Abans es pensava que aquesta manera de procedir estava originada per la manca de coneixement. En no poder entendre les conseqüències de cedir les dades, no es podia prendre la decisió. Darrerament, però, s'està arribant a una conclusió diferent. Segons treballs recents, aquest comportament es produeix després de fer un càlcul de riscos front a beneficis, i el ciutadà mig opta finalment per fer servir les eines que tothom utilitza (Núñez i Beltrán, 2022).

Són raonaments per simplificar la presa de decisions, coneguts pels dissenyadors d'eines, que es preocupen de fer sentir els usuaris que tenen el control de les seves dades. El benefici ha de semblar molt més alt al risc que es corre, i a això se li sumen les dificultats per comprendre les implicacions de la cessió a través dels clausulats molt llargs o difícils d'accedir i altres mètodes que es solen fer servir per obtenir consentiments. Es tracta dels anomenats patrons foscos (*Dark Patterns*), i són procediments de disseny per aconseguir l'acceptació inconscient. Aquestes tàctiques poden ser, com dèiem, elaborar expressament unes condicions d'ús del servei que no són prou clares, facilitar processos d'alta i complicar peticions de baixes, entre d'altres.

Mentre escrivim aquesta tesi, als Estats Units s'ha prohibit l'avortament. L'activisme pels drets digitals ha cridat l'atenció sobre el risc que implica que les plataformes i altres aplicacions disposin de les quantitats d'informació que han recollit, i que permetria inferir la relació de

persones amb el fet d'haver materialitzat un avortament o amb la intenció d'avortar. Totes aquestes dades podrien ser utilitzades com a prova per vigilar o inculpar a persones o institucions que faciliten l'avortament, altres professionals o organitzacions que hi poden col·laborar, per exemple, proporcionant informació.

Més enllà del posicionament de cadascú sobre el dret a la interrupció de l'embaràs, aquest panorama podria resultar intimidant per a qualsevol persona mínimament preocupada pels seus drets i llibertats, fins i tot podria preocupar a moltes d'aquelles que asseveren que els és igual que les vigilin perquè no tenen res a amagar.

Aquest cas demostra el grau de responsabilitat que implica decidir sobre el disseny de les tecnologies, perquè evidencia que allò que en una època no comporta riscos, en el futur —o sota la responsabilitat d'altres governants— sí que en pot implicar. És per això que la reclamació de l'activisme és que no es desenvolupin tecnologies que possibilitin la capacitat de retallar les llibertats de qualsevol manera ara o en el futur:

Los directivos cambian o son despedidos o mienten o están sujetos a legislaciones y a gobiernos que cambian o mienten. La única pregunta relevante en el debate es si desarrollan tecnologías capaces de ejercer censura, coartar libertades civiles o traicionar la confianza de los usuarios. Si lo hacen, es siempre un problema independientemente de su intención. (Peirano, 2019, p.107)

L'extractivisme sistemàtic de dades personals té altres lectures des del punt de vista dels efectes culturals. L'augment de la capacitat tecnològica, unida a la quantitat de recursos que s'hi destinen, farà que les facultats predictives dels algoritmes millorin fins el punt de passar de predictius a prescriptius. De fet avui ja ho són, amb les recomanacions que fan, però la tendència és que s'assemblin més a serveis del tipus Google Now, que consisteix a fer propostes proactives a partir de suposicions sobre les necessitats que pot tenir la persona en aquell moment. En una conferència al 2010, qui aleshores era el CEO de Google, Eric Schmidt, va afirmar, entre d'altres coses, que Google podia suggerir a la gent què pot fer, perquè sap què li importa (Tsotsis, 2010). A la conferència per a desenvolupadors que Google va fer l'any 2013, el seu vicepresident, Amit Singhal, va dir que, en el futur, el cercador de Google contestarà, conversarà i anticiparà "answer, converse and anticipate" (Tsukayama, 2013).

En aquest servei, a partir de les inferències que extreu Google, proposa accions que poden ser des de recomanar agafar el cotxe per arribar a temps a la següent cita, la millor ruta per arribar a un lloc determinat o la línia aèria que més pot convenir. Segons el professor i escriptor Ed Finn,

no costa imaginar un futur en què el sistema faci propostes cada cop més sofisticades i contínues sobre allò que hauríem de fer a continuació com què llegir, què veure o comprar, però també on convindria que fóssim en aquell moment o a qui hauríem de veure. Així que, indirectament, ens estaria dient què hauríem de pensar o fer (Finn, 2018).

Google és capaç de predir tot això amb les dades de cada usuari que reuneix i processa a partir l'activitat que la persona ha dut a terme al cercador, a altres llocs webs, o a través de la interpretació dels seus correus electrònics a Gmail, per exemple. Es tracta d'allò que Finn (2018) anomena enginyeria de la intimitat:

Estamos trabajando intensamente en construir una relación de intimidad con los algoritmos, desde nuestra disposición a jugar con Siri a las cosas que tecleamos en las barras de búsqueda cuando pensamos que nadie está mirando. Interactuamos con interfaces de búsqueda de maneras estrechamente privadas, en lo que constituye un extraño y ocasionalmente grotesco ejercicio de confesión a los altares de la computación. (p.141)

Les ulleres de realitat virtual que han de servir per accedir al Metavers són principalment això, l'entrada al món paral·lel de Zuckerberg. El CEO de Meta explica en la presentació del Metavers (Meta, 2021), que en aquest entorn virtual es podran tenir les vistes més espectaculars des de la pròpia casa, que també serà tan fabulosa com vulgui cadascú, es podrà quedar amb amics, anar a concerts arreu del món, fer esport, comprar o fer reunions de feina. Tot i que avui existeixen dubtes sobre la consolidació d'aquests plans de Zuckerberg, convindria plantejar-se si aquestes ulleres que serveixen per accedir al món virtual podrien exercir com a nou dispositiu de captació de dades personals molt més precís del que pot ser-ho un mòbil, perquè aquestes poden recollir informació afegida directament de les reaccions a través dels globus oculars de les persones.

Harari (2018) proposa un exercici d'imaginació i convida a plantejar-se en quin escenari es trobarà la humanitat d'aquí a dues dècades si continua alimentant els algoritmes de les grans plataformes. "Estos datos pueden permitir a Netflix y a Amazon elegir filmes para nosotros con precisión asombrosa, pero también puede permitirles que tomen por nosotros las decisiones más importantes de nuestra vida, como qué estudiar, dónde trabajar y con quién casarnos" (p.74). Aquest autor presagia un panorama en què es fa realitat la idea més bàsica de progrés que impulsa el determinisme tecnològic i la cerca de l'eficiència. Un desenvolupament tecnològic que avança en paral·lel a la degradació de l'ésser humà. Segons Harari, els algoritmes coneixeran les persones millor que elles mateixes i això conduirà a col·lapse de les societats:

La democracia liberal y el capitalismo de libre mercado ven al individuo como un agente autónomo que no para de tomar decisiones sobre el mundo. (...) ¿Qué pasará con esta forma de entender la vida si cada vez confiamos más en la IA para que tome decisiones por nosotros? (...) la vida humana dejará de ser un drama de toma de decisiones. Las elecciones democráticas y los mercados libres tendrán poco sentido. (p.77)

Lassalle (2019) coincideix amb què s'ha passat d'un capitalisme cognitiu de plataformes que extreuen dades a un altre que els analitza amb finalitat predictiva i ara, també, prescriptiva. També considera que el gran salt de la predicció de la determinació es produirà quan la IA s'apliqui a la biotecnologia i es tractaran malalties abans que apareguin o els algoritmes ens proposin les relacions afectives que més ens convinguin.

Val a dir que les afirmacions de Harari i de Lassalle (aquest darrer influenciat sovint pel primer) parteixen del supòsit que la bioquímica pot arribar a ser computable. No obstant aquest és un extrem no demostrat (com hem vist al segon capítol) i és per aquesta raó per la qual aquest argument s'ha qüestionat de manera contundent per part de sectors del món científic (Narayanan, 2022).

En aquest punt, pensem que val la pena destacar la crítica que es fa a determinades teories de Harari, ja que es tracta d'un autor referent en els temes que plantegem en aquest treball. A més compta amb una gran influència social. És conegut que ha assessorat a Larry Page, Bill Gates o Elon Musk, entre d'altres, i també resulta bibliografia habitual a literatura relacionada amb l'anàlisi social contemporani. Els crítics de Harari argumenten que presenta, d'una manera que vol aparentar ser objectiva, la visió del *dataïsme* per la qual considera les persones es poden reduir a dades. Amb aquesta lògica —que a més Harari assegura que ha estat acceptada científicament— redueix l'ésser humà a un algoritme que acabarà obsolet, superat per millors algoritmes.

Aquestes valoracions pretenen refutar molts dels raonaments d'aquest autor i apunten també que Harari és un personatge reverenciat a Silicon Valley, mentre que Soshana Zuboff no ho és, perquè ella responsabilitza de l'erosió de l'autonomia humana a les plataformes i al model de negoci de la vigilància que han desenvolupat; Per contra, Harari atribueix el problema a una suposada inferioritat natural de l'ésser humà davant la tecnologia (cosa que implica una desresponsabilització de les tecnològiques per la seva activitat extractivista). Al seu torn, els arguments d'aquest pensador condueixen a la conclusió de la inexorabilitat de l'obsolescència

humana. En definitiva, aquestes crítiques conclouen que les bases per fer aquestes afirmacions no són prou sòlides des del punt de vista científic (Narayanan, 2022).

A l'època que vivim resulta difícil mantenir l'atenció en allò que és important i en conseqüència també és complicat dedicar-hi el temps necessari. A les persones adultes ens toca preocupar-nos per protegir les criatures i altres col·lectius vulnerables de les xarxes, de la pèrdua de l'autonomia que implica el *hackeig* de les ments en què es basa la mediació algorítmica comercial (que hem tractat a l'apartat sobre l'*engagement*), o de l'erosió de la privacitat que provoca l'extractivisme que està en la lògica del seu model de negoci. Fins i tot als adults ens costa sovint alliberar-nos de la pressió per estar presents consumint i generant continguts. Atrapats en una "esclavitud narcisista", en paraules de la filòsofa Clara Serra, qui exhorta a exercir la llibertat "contra ese capitalismo del yo, contra esa asfixiante obligación de producir nuestra propia autenticidad" (Serra, 2022).

Una societat que podria consolidar el presagi de Postman fa trenta anys: que el profeta de la nostra època no era Orwell amb la societat vigilada que reflectia a 1984, sinó Huxley, amb la distòpia que plantejava a *Un món feliç*, en la qual la gent arriba a estimar la seva opressió, i adorar les tecnologies que anul·len la seva capacitat de pensar (Postman 2012):

Lo que Orwell temía eran aquéllos que pudieran prohibir libros, mientras que Huxley temía que no hubiera razón alguna para prohibirlos, debido a que nadie tuviera interés en leerlos. Orwell temía a los que pudieran privarnos de información. Huxley, en cambio, temía a los que llegaran a brindarnos tanta que pudiéramos ser reducidos a la pasividad y el egoísmo. Orwell temía que nos fuera ocultada la verdad, mientras que Huxley temía que la verdad fuera anegada por un mar de irrelevancia.

Orwell temía que nos convirtiéramos en una cultura cautiva. Huxley temía que nuestra cultura se transformara en algo trivial, preocupada únicamente por algunos equivalentes de sensaciones varias.

En 1984, agregó Huxley, la gente es controlada infligiéndole dolor, mientras que en *Un mundo feliz* es controlada infligiéndole placer. (p.5)

Marcuse (1993) definia la societat industrial com un sistema capaç d'assimilar moviments i canvis per a la millora del benestar de la majoria. Alhora, però, considerava que el mateix sistema també disposava d'una capacitat contenidora d'aquest canvi que obtenia gràcies al paper de la tecnologia en el manteniment de l'ordre social. Segons Marcuse, l'aparell productiu és totalitari en tant que, posat al servei del sistema dominant, condiciona tant les actituds socials

com les individuals en favor d'aquell ordre. I és per això que la tecnologia no pot ser considerada neutra, sinó que finalment esdevé política (Marcuse, 1993):

En esta sociedad, el aparato productivo tiende a hacerse totalitario en el grado en que determina, no sólo las ocupaciones, aptitudes y actitudes socialmente necesarias, sino también las necesidades y aspiraciones individuales.(...)

Ante las características totalitarias de esta sociedad, no puede sostenerse la noción tradicional de la «neutralidad» de la tecnología. La tecnología como tal no puede ser separada del empleo que se hace de ella: (...)

En tanto que universo tecnológico, la sociedad industrial avanzada es un universo político, es la última etapa en la realización de un proyecto histórico específico, esto es, la experimentación, transformación y organización de la naturaleza como simple material de dominación. (...)

La productividad y el crecimiento potencial de este sistema estabilizan la sociedad y contienen el progreso técnico dentro del marco de la dominación. La razón tecnológica se ha hecho razón política. (pp.26-27)

El grau d'automatització present en la societat industrial que analitzava Marcuse es multiplica avui com a conseqüència del desenvolupament de les tecnologies exponencials. La seva crítica resulta un prisma que pot ser adient per avaluar el paper dels sistemes predictius i prescriptors generats per una innovació tecnològica incapaç d'entendre la llibertat de la persona més enllà de la possibilitat d'elecció com a consumidora (Hall, 2011).

En aquest escenari pensem que cal combatre el relat de la determinació reivindicant la llibertat de no ser previsibles que es sustenta en la concepció antropològica de l'ésser humà com a animal ambigu, ja que la natura humana, com arugmentem amb Vidal i Garde (2022), es fonamenta també en l'obertura, en el possible, en allò imprevisible més enllà del que és probable, i que es representa en l'*homo revelatus* (Duch, 2018) que mai està del tot acabat ni determinat:

...estructuralment, és a dir, en relació amb les seves predeterminacions, homes i dones de tots els temps som radicalment iguals, però, històricament, és a dir, amb relació amb les determinacions de cada individu i de cada grup humà, som culturalment i biogràficament diferents. Les predeterminacions són inaccessibles i indefinibles. (p.79)

Llavors, des dels àmbits culturals, podrem influir en el desenvolupament tecnològic reivindicant una tecnologia veritablement racional (Feenberg, 2002; Marcuse, 1993) que no és altra que la que acull la natura de l'ésser humà i l'encoratja a manifestar la seva llibertat per evolucionar i

créixer, respecta el seu tarannà *epifànic* (Duch, 2001) i el medi ambient en el qual ha de poder viure i prosperar.

CAPÍTOL 4

EL PERIODISME DAVANT L'AUTOMATITZACIÓ

Sense ètica no hi ha periodisme.

Salvador Alsius

Nos asombramos porque los movimientos de un autómatas parecen no obedecer a causa alguna, y el asombro se disipa cuando damos con el mecanismo que los produce.

Dardo Scavino

Los conocimientos secretos son la clave de cualquier sistema de control total. Hasta que se inventó a la imprenta, la palabra escrita siguió siendo en gran medida un monopolio de clase; en la actualidad, el lenguaje de las matemáticas superiores, sumado a las misteriosas claves de los ordenadores, están restaurando tanto el secretismo como el monopolio de tal saber, con la consiguiente reanudación del control totalitario.

Lewis Mumford

Ocasionalmente, atisbamos formas de resistencia a esta poderosa invisibilidad. Esta resistencia exige un conocimiento tecnológico y en red, requiere volver la lógica del sistema contra el sistema.

James Bridle

4.1 INTRODUCCIÓ AL CAPÍTOL 4

Als capítols anteriors hem exposat el context social i cultural contemporani en el qual es desenvolupa i penetra la intel·ligència artificial en un nombre creixent d'aplicacions i en els diferents àmbits de la societat. En aquesta descripció, hem intentat presentar una exposició raonada sobre els dilemes ètics i socials que emergeixen en el procés d'assimilació, mirant de reflectir les principals qüestions presents en el diàleg entre la filosofia moral i la tecnologia exponencial, per anar tancant la mirada en vers l'ètica de la comunicació computacional. D'entrada hem tractat els conceptes de progrés, racionalitat i finalitats humanes, fonamentals i transversals en aquest treball. L'exercici de definició, al seu torn, ha posat de manifest la mirada amb què plantejem tant l'anàlisi com les propostes, que estan orientades al bé comú per sobre de les individualitats. Així mateix, hem reclamat que el govern de la tecnologia implica necessàriament un lideratge cultural que li doni sentit, i per consegüent amb una idea molt clara sobre les finalitats, combinada amb una actitud de resistència davant les inèrcies de les societats capitalistes que tendeixen a desfermar la *màquina*. Perquè les eines multiplicadores de la productivitat acaben posades al servei de la generació de rèdits econòmics concentrats en una minoria que, a més, tendeix a perpetuar els seus privilegis a través de mecanismes que impedeixen allò que considerem el veritable progrés. Una evolució social que ha de consistir, en síntesi, a què l'increment de la productivitat reverteixi en benefici de la igualtat d'oportunitats per a què més persones puguin exercir el seu dret a *ser*: a prosperar amb llibertat en condicions que permetin la seva realització personal.

Amb la finalitat de proporcionar el context que ens ha d'ajudar a donar direcció al full de ruta que proposarem en el present apartat, i en línia amb el tema d'aquest treball —l'ètica aplicada a la comunicació computacional—, hem dedicat el capítol anterior als efectes de la tecnologia posada al servei dels valors de l'individualisme i que es materialitza en la mediació algorítmica comercial per part de les grans plataformes tecnològiques. Així, per una banda, hem exposat els seus beneficis, que són haver facilitat eines avançades i simples d'utilitzar per posar en contacte les persones i facilitar l'accés al coneixement. D'altra banda, també hem examinat els seus efectes —sense oblidar que també som les persones qui els hi donem forma—. D'aquesta manera, hem vist que la manca d'ètica en el disseny de tecnologies de comunicació que explota vulnerabilitats humanes té el potencial de causar addiccions, d'impulsar la intoxicació de l'esfera pública o de rebaixar l'exposició a la diversitat de punts de vista —i per tant de contribuir a la polarització social—. Així mateix, com a conseqüència de les noves formes d'extractivisme de

les dades personals en què es basen els seus negocis, aquesta optimització comercial també pot erosionar la privacitat i per consegüent les llibertats de la ciutadania.

En aquest capítol presentarem un pla d'acció, i ho farem des de la tradició eticista que singularitza la pràctica periodística en l'àmbit de la comunicació (Alsius, 1998) ja que l'ètica periodística és la disciplina en la qual es produeix la reflexió específica sobre els efectes de la comunicació en l'esfera pública, i es basa en el Dret a la Informació reconegut a la Declaració Universal dels Drets Humans. Així mateix, la deontologia que neix dels grans principis del periodisme resulta la via més clara per inferir una ètica general de la comunicació (Alsius, 2011). Una doctrina que arrela en el reconeixement que la comunicació pel bé comú realitza una funció social que està per sobre dels beneficis econòmics (Kovach i Rosensteel, 2014; Ruiz-Caballero, 2016) com a contrapès al perill de degradació del producte informatiu que acostuma a implicar la pressió per assolir determinats objectius de negoci, tal com apunta Alsius (1998): “cal reconèixer que les forces del mercat acaben sent sempre un factor decisiu per a determinar la qualitat del periodisme” (p.105).

Com hem exposat, els canvis que testimoniem situen el periodisme davant una crisi econòmica, però també ontològica, que reclama una reflexió sobre quin ha de ser el seu paper en unes societats en les quals delegarem cada cop més en la computació, i en l'àmbit de la comunicació, ja revolucionat, té sentit augurar que les empreses de Silicon Valley que dominen l'esfera pública digital desenvolupin els seus serveis i en millorin l'eficiència. Al mateix temps, a la UE es preveuen nous marcs reguladors que, com hem explicat abans, s'espera que beneficiïn els actors més modestos. Aquest nou escenari comunicatiu que hem descrit interpel·la la professió a plantejar-se quin ha de ser el seu rol i, tal com defensem en aquest treball, valorar com recuperar la seva missió original de servei públic en la seva estratègia d'innovació. D'aquesta manera, el periodisme segurament estarà en millors condicions de contribuir a garantir el dret a la informació d'una ciutadania necessitada d'interlocutors de confiança en el context que hem dibuixat.

Al mateix temps que les grans plataformes amb la seva tecnologia algorítmica disposen d'una capacitat inqüestionable per filtrar les enormes quantitats d'informació digital, també demostren —en termes generals— desenvolupar la seva activitat regides per valors purament comercials. No pertoca aquí aprofundir en les oportunitats del negoci periodístic, però partirem de la idea que, al menys a curt ni a mig termini, les empreses periodístiques poden competir tecnològicament amb aquests gegants. Tanmateix, com també hem anat veient, les

consideracions ètiques de les plataformes de mediació algorítmica no resulten prioritàries fins que no pateixen crisis reputacionals. En paraules de la filòsofa i professora d'Oxford Carissa Véliz “El problema es cuando se diseña una tecnología muy potente, con efectos sistémicos, con el objetivo principal de ganar dinero, pero bajo la apariencia de querer mejorar el mundo y sin tomar en consideración los derechos de las personas” (Véliz, 2022).

El present capítol, que construïm sobre aquest panorama que acabem de descriure, exposarem el que podria ser un punt de partida per tal que el periodisme dugui a terme les consideracions sobre quin ha de ser el seu paper en favor de les llibertats en les societats democràtiques, en el context d'una creixent tecnològització en la major part d'àmbits i necessitada d'informació de qualitat (Mont'Alverne et al., 2022).

Al mateix temps que defensem que aquest nou escenari segueix reclamant un producte informatiu en benefici del bé comú, no perdem de vista que des del seu naixement al S.XIX, el periodisme industrial convencional no ha demostrat en la majoria dels casos la seva independència dels poders fàctics, sinó que més aviat s'ha ocupat de servir a objectius de negoci i s'ha convertit en una instància de creació de l'hegemonia discursiva vinculada a l'exercici del poder (Vidal Castell, 2020). En aquest sentit, només podem coincidir amb aquest argument que ha escrit Alsius (1998): “precisament perquè la informació és —entre altres coses— una mercaderia, la seva elaboració està subjecta a una sèrie de condicionants que prostitueixen la seva funció social” (p.168). Per aquesta raó podríem estar d'acord amb qui pensés que apel·lar als valors ètics de la professió, a aquestes alçades, podria resultar naïf. Però aquesta proposta considera, en primer lloc, que el periodisme es fa per periodistes, i molts, tot i no poder comptar amb el suport ideològic dels àmbits de gestió del negoci, es senten compromesos amb la veritable funció social de la professió.

Considerem així mateix que el periodisme té l'oportunitat de seguir explorant alternatives de finançament independents de la publicitat, com ho demostren recents iniciatives periodístiques que es sostenen gràcies a vies com ara el mecenatge, o a través de formats semblants a clubs de socis, o fins i tot mitjançant diferents variants de pagament pels continguts que —tot i que de manera desigual— continuen en una tendència que manifesta un creixement lent però sostingut els darrers anys (Suárez, 2021).

En tercer lloc, pot valer la pena tenir present que tampoc seria cap novetat apostar per la qualitat des d'un punt de vista d'estratègia de negoci (Kovach i Rosensetel, 2014). Precisament així és com alguns mitjans ja s'hi enfronten, entre aquests el *New York Times* (*Our strategy*, 2022) un

mitjà que ha apostat històricament per la qualitat com a element singularitzador. Així va ser, precisament, com va aconseguir la seva reputació a finals S.XIX davant la premsa sensacionalista de Hearst. El context comunicatiu en el qual aquest diari aixecà la seva proposta de valor que rememora la professora Gloria García González (1999) podria equiparar-se en certa manera a l'actualitat:

Al apelar a la capacidad de abstracción de los lectores y no a su emocionalidad, el diario establece una frontera moral entre la calidad que él representa y la bazofia de la prensa popular y, en consecuencia, una frontera también social entre los lectores más reflexivos y los más impresionables. El *New York Times* consigue hacer de su lectura un signo de distinción, de orgullo por algo socialmente reconocido frente a la vergüenza asociada al consumo de otro tipo de prensa. Esta es la razón por la que el estilo del *New York Times* conecta con las clases altas y las que, sin serlo, aspiran a identificarse con ellas. (A. Braojos Garrido et al., 1999, p.98)

Tal com assenyala Alsius (1998) els interessos dels professionals, els consumidors i les empreses poden convergir en determinades conjuntures de mercat: “Independentment de quines siguin les motivacions autèntiques que guïïn a cadascun d’aquests grups, les respectives actituds poden generar un terreny de confluència ben adobat per a la deontologia” (p.85). De fet, recorda l’autor, és precisament aquesta coincidència d’interessos —que es concentren en finalitat d’aconseguir la credibilitat de la informació— allò que impulsà els primers codis de conducta (especialment als Estats Units).

En el present capítol esperem facilitar eines tant teòriques com pràctiques per tal que la professió es ressituï en aquest escenari en favor de les llibertats en les societats democràtiques. Aquestes propostes es materialitzaran en tres línies d’actuació que detallarem una per una. La idea transversal que descansa rere aquestes propostes d’acció és la necessitat de trencar murs entre les disciplines de les humanitats o les ciències socials i les científiques per conèixer i integrar les tecnologies algorítmiques amb el criteri necessari i així fer-ne una apropiació responsable. Només d’aquesta manera el periodisme estarà en condicions d’informar sobre IA amb rigor, de vigilar-ne els seus usos i de delegar en els algoritmes de manera exemplar.

La professió periodística, com d’altres, és cada cop més tècnica i avui comença a ser habitual que les periodistes aprenguin llenguatges de programació, però això no és un requisit imprescindible per disposar d’un punt de vista crític sobre els usos —propis o aliens— dels algoritmes i el llenguatge de programació, que cal desmitificar. Tal com observa Ed Finn (2018) “creemos en el poder del código como conjunto de símbolos mágicos que conectan lo invisible

y lo visible” (p.69). No obstant, la realitat és que per valorar els usos que se’n fan, acostuma a ser suficient conèixer conceptes bàsics com ara la diferència entre un algoritme simple, la IA, l’aprenentatge automàtic o l’aprenentatge profund. També resulta pràctic disposar d’uns mínims coneixements sobre llenguatge computacional, així com certes tècniques a l’abast de qualsevol periodista.

D’altra banda, per poder mirar i explicar el canvi tecnològic, ajudarà tenir present que el domini del llenguatge críptic pot ser una gran font de poder de qui vol exercir-lo, i de fet així ha estat al llarg de la història. Lewis Mumford (2010) cridava l’atenció sobre el paper del llenguatge en el control social en referir-se a la creació de la seva *megamàquina* en el context de les organitzacions socials jerarquizades i inflexibles que van erigir les ciutats antigues a Egipte i que considera l’avançada de les ciutats actuals nascudes de la raó instrumental. En primer lloc, Mumford recorda el paper fonamental del clergue i de la ciència en la legitimació del poder dels reis sobre els esclaus:

Ningún rey podía moverse con seguridad ni eficacia sin el apoyo de tal conocimiento superior, ni más ni menos que lo que el Pentágono puede actuar hoy sin consultar a sus científicos especializados, a sus expertos técnicos, a sus teóricos de los juegos y a sus ordenadores: una jerarquía supuestamente menos fiable que los intérpretes de las vísceras, pero a juzgar por sus tremendos errores, no mucho más. (p.329)

Mumford assenyala de manera convincent que si tothom hagués tingut accés al coneixement que només tenien els sacerdots i a les claus per a la seva descodificació no s’haguessin pogut ocultar els errors i, per tant, el poder hagués perdut legitimitat.

El llenguatge computacional de l’algoritme que exclou beneficiaris del bo social, o el de la IA racista de la policia dels Estats Units comparteixen amb altres codis el fet d’estar només en mans d’uns quants que de vegades acaben posant-se al servei del poder. També és cert que la majoria de les persones no actuen amb la finalitat de perjudicar altres, o amb negligència, tanmateix la supervisió també pot ajudar a descobrir errors inintencionats.

4.2 INFORMAR SOBRE INTEL·LIGÈNCIA ARTIFICIAL AMB RIGOR

Per adoptar una posició respecte a la computació de cada cop més aspectes de la vida, el periodisme haurà d'acceptar que, avui, vigilar el poder és també vigilar els algoritmes, i per estar en condicions d'exercir aquesta funció haurà d'aprendre a desxifrar-los. Només així podrà informar amb rigor a una ciutadania necessitada de coneixements i competències tecnològiques (Roca et al., 2022) que es tradueix tot sovint en manca d'esperit crític generalitzat en vers la tecnologia, un dels problemes als quals el periodisme podria contribuir a resoldre.

Una mostra d'aquesta mancança és l'estudi de la Universitat de Yale que conclou que les persones es mostren menys crítiques amb decisions preses per algoritmes que per éssers humans (Bigman, 2020). Aquesta passivitat social no necessàriament és atribuïble a una falta de coneixements tecnològics, sinó que la tendència al *dataïsme* exposada al segon capítol també pot jugar-hi un paper important. Sigui com sigui, la divulgació feta amb rigor sol tenir sempre el potencial de resultar d'ajuda. Per tant, d'entrada convé que el periodisme assumeixi aquest paper amb responsabilitat. Informar sobre IA amb precisió passarà també per avaluar el llenguatge que es fa servir en el producte informatiu i començar a qüestionar conceptes que atribueixen propietats humanes als algoritmes, començant per la mateixa noció d'intel·ligència aplicada a les màquines, considerant allò que hem exposat al capítol 2 sobre les seves limitacions i els riscos de la manca de responsabilitat en els usos lingüístics.

Fer front a tots els reptes que planteja l'automatització passarà primer per identificar-los i comprendre la seva natura. Ensenms, adquirir una perspectiva crítica passarà per disposar de la capacitat per interpretar el procés algorítmic. És a dir, entendre la "máquina cultural antes que que a sus productos culturales" (Finn, 2018, p.102).

Dèiem que un algoritme es defineix, de la manera més simple, com un conjunt de regles per a resoldre un problema en un nombre finit de passos. Tanmateix, per poder fer interpretacions, convindrà anar més enllà per comprendre'l amb una mica més de profunditat, i no ens referim a tenir la capacitat de dissenyar-ne cap. Finn amplia aquesta definició i hi inclou elements que evidencien la lògica que segueixen aquests sistemes:

Los algoritmos de toda clase anticipan una versión del argumento de la computabilidad efectiva, cifrando explícita o implícitamente argumentos que defienden que el problema —ya sea relacionado con la agricultura o con la

obtención de una raíz cuadrada— podrá solucionarse siguiendo los pasos del método. En este sentido, el mismo proceso es una lógica ordenadora para la interpretación crítica, que se apoya en conceptos como la filosofía del proceso, sostenida por los filósofos Martin Heidegger, Alfred Whitehead, Simondon y Stiegler, entre otros. (p.103)

Així doncs, si els algorismes xifren arguments que defensen que el problema podrà solucionar-se seguint els passos del mètode, caldrà analitzar fins a quin punt aquest llenguatge simbòlic computacional conté, a les variables i altres criteris amb què ha estat dissenyat, els arguments que efectivament solucionen la qüestió d'una manera eficient, és a dir, sense simplificar-la i, sobretot, sense causar perjudicis.

Al capdavant es tracta de disposar d'uns coneixements bàsics sobre les lògiques d'aquests processos computacionals que acabem de descriure. No obstant, per poder fer interpretacions crítiques també convindrà conèixer els principals dilemes morals relacionats amb els usos de la computació i que passen per identificar tant les seves potencialitats com els seus límits. El periodisme té l'oportunitat de contribuir a combatre, des de la comunicació, narratives que tenen el potencial de proporcionar un excés d'autoritat a la màquina alhora que promouen la passivitat humana davant un racionalisme mal entès. En aquest sentit, resultarà imprescindible realitzar aquesta incursió que suggerim amb una mentalitat independent que neixi dels valors ètics que han de garantir la qualitat de la informació. Només des d'aquesta mirada crítica s'estarà en condicions d'informar amb rigor sobre IA i per tant de contrarestar els relats interessats que promouen l'enlluernament tecnològic, el determinisme i els egos necessitats de protagonisme que s'alimenta amb la por al desconegut que predomina entre la ciutadania.

4.3 VIGILAR ELS ALGORITMES

Avui i com a conseqüència de la creixent automatització comença a néixer una disciplina periodística que traduïda es diria *retiment de comptes algorítmic (algorithmic accountability)* (Diakopoulos, 2015, 2019; Trielli i Diakopoulos, 2020). Es tracta d'un conjunt de tècniques a l'abast de les periodistes per auditar el funcionament dels algoritmes, la majoria d'elles, sense necessitat de tenir coneixements de programació. En alguns casos és possible que sí que se'n requereixin, però això no ha de ser un impediment per a les redaccions, que cada cop incorporen més perfils tècnics que poden col·laborar en la creació del producte informatiu.

Per tal de respondre d'una manera més clara a la pregunta sobre quin ha de ser el paper del periodisme davant el creixent ús de la IA a la societat, enumerarem i descriurem de manera resumida algunes tècniques que extraurem de diferents marcs conceptuals que ja s'han desenvolupat per dur a terme les esmentades auditories. Aquest apartat no es proposa aprofundir en aquesta matèria, sinó que té la finalitat d'explicar que les habilitats necessàries per dur a terme auditories algorítmiques estan a l'abast de qualsevol periodista interessada en aquesta qüestió.

El professor de periodisme computacional de la Northwestern University de Chicago, Nicholas Diakopoulos, estableix un dels principals marcs d'anàlisi de les decisions algorítmiques que resumirem a continuació. Una d'aquestes tècniques és la *reverse engineering* (enginyeria inversa), la qual implica examinar un sistema provant el seu funcionament mitjançant l'observació i la deducció per descobrir el model o la lògica que segueix. Es va fer servir, per exemple, per informar sobre companyies de comerç en línia que oferien diferents preus segons qui els consultava: modificaven els imports en funció de la ubicació i altres comportaments rastrejables de la persona usuària (Soltani, 2012). La tècnica *audit* (auditoria) és una altra tècnica que es basa en el fet que si les dades d'entrada (*ímputs*) dels sistemes varien de maneres prou diferents al mateix temps que es controlen els resultats (*output*), llavors les entrades i les sortides es poden correlacionar per construir una teoria sobre com pot funcionar aquest sistema.

El mètode *Poking and prodding* (Picar i punxar) consisteix a exposar a la reactivitat algorítmica per comprendre com el sistema reacciona als *ímput* del periodista i com afecta altres persones. Un bon exemple d'aquesta tècnica és una peça d'una periodista de *Buzzfeed* qui, en un

experiment fet des de la seva pròpia pàgina de Facebook, va comprovar com l'algoritme d'aquesta xarxa social prioritza posts que generen més comentaris. Aquesta redactora de tecnologia ho va comprovar amb un post absurd que ella mateixa va fer, i malgrat tractar-se d'una publicació sense cap valor —que fins i tot resultar molests per als seus amics— es mantenia com a prioritària a la línia de temps dels contactes només perquè rebia comentaris en què aquests amics criticaven precisament l'absurditat i la manca d'interès del tema que havia penjat (Notopoulos, 2018). Aquest fet, que en el cas descrit no va tenir més importància, sí que n'adquireix des d'una perspectiva àmplia, perquè —tal com hem exposat a l'apartat sobre l'*engagement*— els missatges són més visibles a les línies de temps d'altres usuaris acaben sent els més polèmics —amb les conseqüències que això pot comportar—.

Un altre dels principals mètodes d'auditoria que menciona Diakopoulos és la tècnica del *crowdsourcing* (proveïment participatiu). En aquest cas implica identificar casos reals de discriminació algorítmica reunint dades de gent que les cedeix per dur a terme la investigació. El cas de l'algoritme per a la concessió del bo social que va denunciar Civio mencionat més amunt entraria en aquesta categoria, ja que es va descobrir que el sistema no concedia l'ajut a persones que les investigadores havien comprovat que efectivament complien els requisits per ser-ne beneficiàries. La investigació periodística que va destapar l'algoritme racista Compas (Angwin, 2016) també es va dur a terme amb aquest procediment. Un altre dels possibles mètodes que proposa Diakopoulos és la revisió directa del codi, com es va fer en una investigació periodística que detallava com persones discapacitades van veure retallats els recursos públics dels que disposaven (hores de cures dutes a terme per persones especialitzades) des del moment en què es va començar a fer servir un algoritme per prendre aquesta decisió (Lecher, 2018).

El periodista computacional Francesco Marconi (2021), planteja un altre marc per poder auditar els usos dels algoritmes de manera senzilla que consisteix en una cinquantena d'interrogants que tenen la finalitat de donar lloc a les respostes clau que, al seu torn, permetran fer una avaluació crítica sobre el grau de rigor en l'ús de la IA. Es tracta de preguntes com: Quina mena de tecnologia utilitza l'algoritme? Quines són les formes d'intervenció humana? Quines biblioteques s'han utilitzat per construir el sistema? Es tracta de dades públiques? o bé privades? Com va decidir quines dades són rellevants per al model? S'informa els usuaris finals de quines inferències es fan amb les dades? (p. ex. exempcions de responsabilitat algorítmica) Si l'empresa revela aquesta informació als usuaris, la comunicació és clara, pública i fàcil d'entendre? L'empresa ha estat sotmesa a un expert o plataforma independent que audita les seves dades?

Com es pot comprovar, les qüestions fonamentals pretenen esbrinar el nivell d'implicació humana en el disseny i la supervisió del funcionament del sistema, així com el grau de transparència i responsabilitat que s'ha assumit en les diferents fases de creació de l'aplicació.

El rigor en els usos de la IA en la presa de decisions es pot valorar a partir dels grans principis de l'ètica de la IA. Els següents cinc conceptes ètics bàsics poden servir per començar a aprofundir-hi (Diakopoulos i Friedler, 2016):

- Responsabilitat: Cal que al menys una persona pugui respondre per la decisió algorítmica.
- Explicabilitat: El resultat algorítmic hauria de poder-se explicar de manera entenedora a la persona afectada per la decisió.
- Precisió: És necessari identificar els riscos i les possibles causes d'errors per supervisar-lo.
- Auditabilitat: els sistemes estiguessin dissenyats de manera que tercers puguin auditar-los.
- Justícia: Tots els algoritmes que prenen decisions sobre individus s'han d'avaluar per evitar possibles efectes discriminatoris. Els resultats de l'avaluació i els criteris utilitzats s'han de donar a conèixer i explicar públicament.

A Europa, el Reglament general de protecció de dades (RGPD) regula les pràctiques que es basen en la utilització d'algoritmes per prendre decisions (European Commission, 2018) i reconeix el dret de la persona a no ser objecte del *judici* únicament automatitzat si la determinació té efectes legals o molt significatius sobre la seva vida. Aquesta normativa indica que els responsables de l'algoritme només poden dur a terme aquest tipus de procediments quan la decisió sigui necessària per a la celebració o l'execució d'un contracte; o be sigui autoritzada per la legislació de la UE o dels estats membres aplicable al responsable del tractament; o basada en el consentiment explícit de la persona.

També regula el perfilat que pot formar part d'aquestes decisions (el processament automatitzat de dades personals per avaluar determinats aspectes relacionats amb una persona). Rere aquesta normativa s'identifiquen els usos que la llei considera apropiats i els que no i per tant resulta el marc de valoració indicat en el context de la UE.

Per tractar les dades personals, els responsables han d'ajustar-se al RGPD, proporcionar a les persones informació sobre el tractament, oferir maneres senzilles de sol·licitar intervenció

humana o impugnar una decisió, així com realitzar comprovacions periòdiques per assegurar-se que els sistemes automatitzats funcionen segons el previst. La normativa inclou d'altres limitacions, com per exemple en casos que afecten a determinats usos especials com quan es tracten dades personals de menors, entre d'altres.

Rere la normativa del RGPD —de la qual només n'hem destacat alguns supòsits principals— es dedueix la importància de considerar certs aspectes com la pertinència d'informar sobre el seu funcionament (transparència), el dret de les persones a saber com s'ha pres una decisió, a posar-hi objeccions (explicabilitat), que els models matemàtics han de ser adients (precisió), els individus poden reclamar la participació humana necessària en el procediment (supervisió), que el responsable ha de dedicar recursos organitzacionals i tècnics suficients per minimitzar errors i protegir les dades que posseeix per evitar riscos relacionats amb la privacitat, o que en la interpretació de les dades cal procurar que variables com el gènere, la raça o altres criteris no suposin un risc que condueixi al sistema al tipus de resolució que comporti biaixos discriminatoris.

A Catalunya, l'Observatori d'Ètica en Intel·ligència Artificial, conjuntament amb el Departament de la Vicepresidència i de Polítiques Digitals i Territori de la Generalitat, han llançat una iniciativa que consisteix en una eina d'autoavaluació d'algoritmes anomenada PIO (Principis, Indicadors i Observables) que realitza un càlcul a partir de set grans principis de l'ètica de la IA i, segons el resultat, el proveïdor de l'aplicació tindrà dret a un certificat de qualitat. Els principis que considera són: la transparència i explicabilitat, la justícia i equitat, la seguretat i no maleficència, la responsabilitat i retiment de comptes, la privacitat, l'autonomia, i la sostenibilitat (*Com funciona – OEIAC, 2022*).

4.4 RECERCA EMPÍRICA I RESULTATS: REPTES I RECOMANACIONS PER DOTAR LA INTEL·LIGÈNCIA ARTIFICIAL DELS VALORS ÈTICS DEL PERIODISME

... if machines can do basic journalism, then anyone with a machine can do that journalism.

Charlie Beckett

Automate what computers do best, let people do the rest.

Nicholas Diakopoulos

El diálogo, la deliberación racional entre seres humanos, no es solo un elemento constitutivo de la ética de la comunicación, sino que es el camino para intentar una convivencia política que sea moralmente aceptable. Dicho en otras palabras, es un proceso de formación discursiva de la voluntad colectiva.

Carlos Ruiz

4.4.1 ALGORITMES A LES REDACCIONS

En aquest apartat, exposarem la recerca empírica que hem portat a terme de la mà del Consell de la Informació de Catalunya (CIC) i que s'ha materialitzat en l'informe, publicat pel mateix CIC, titulat *Algoritmes a les redaccions: reptes i recomanacions per dotar la intel·ligència artificial dels valors ètics del periodisme* (Ventura-Pocino, 2021). Les properes pàgines, en les quals el treball apareix ampliat i actualitzat, estan destinades a validar la hipòtesi d'aquesta tesi i a respondre algunes de les principals preguntes: quines eines i aplicacions algorítmiques comporten més riscos ètics; Com podem integrar la recerca i la regulació en ciberètica en l'ètica periodística; Quina mena d'eines podem posar a disposició dels mitjans que puguin adaptar-se a les múltiples casuístiques que es poden donar amb els usos d'IA i altres aplicacions algorítmiques; Quins dilemes ètics concrets generen; quins criteris ha d'aplicar el periodisme en l'ús de les tecnologies que es deriven de la IA; A quins grans principis de l'ètica periodística afecten: Si emergeixen o es resignifiquen determinats principis; Com s'estan fent servir els algoritmes per fer periodisme i com valoren els professionals dels mitjans la incorporació de la IA per fer periodisme.

Així mateix, en aquest apartat ens proposem assolir alguns dels principals objectius d'aquest treball, que són: determinar els criteris ètics que el periodisme pot aplicar en l'ús de la Intel·ligència Artificial; descriure i analitzar com s'està incorporant la IA als mitjans de comunicació; Identificar reptes i oportunitats que els professionals dels mitjans reconeixen en l'adopció de la IA; incentivar una reflexió en el sector en torn als reptes ètics que implica l'adopció de la IA a les rutines periodístiques i proporcionar eines per governar els algoritmes a través dels valors ètics del periodisme per contribuir, en primer lloc, a què la professió recuperi la centralitat en la definició de l'agenda —tot generant confiança a través de l'ús ètic de la tecnologia— i, en segon lloc, contribuir a què es converteixi en referent en tecnòtica de la comunicació.

Aquest treball ha comptat amb la participació dels principals mitjans catalans, així com de professionals de tot l'Estat. En la redacció del present capítol hem considerat rellevant mantenir les expressions i cites recollides de les persones participants en els exercicis de recerca, i hem vist preferible integrar-les en el relat —en lloc de a l'apèndix— ja que, amb la seva literalitat i el diàleg que estableixen amb el cos de text, pensem que poden enriquir de manera significativa el resultat final.

Com exposàvem a la justificació, la IA ja està present en totes les fases de producció i distribució del producte informatiu. El marc més comú per analitzar els usos de la IA als mitjans és el de les 3 fases de producció (Figura 3).

- A. Prèvia a la creació del contingut, que consisteix en la detecció de temes noticiables, la tria de l'angle d'enfoc del tema i la recerca documental entre bases de dades pròpies o alienes per nodrir el contingut.
- B. Creació del contingut, que implica la generació de text, infografies, vídeos, transcripcions i canvis de format (per exemple de text a àudio o vídeo).
- C. Distribució: amb eines que filtren i recomanen continguts a cada usuari tenint en compte comportaments previs. D'altres, optimitzades per millorar mètriques de negoci com les subscripcions o les pàgines vistes. També es fa servir per fer proves A/B de titulars o per automatitzar la les xarxes.

Figura 3
Fases de producció del producte informatiu



Font: Elaboració pròpia

Llorenç Gomis a *El medio media*, a principis dels setanta, ja defensava que el periodisme és un procés d'interpretació successiva, és a dir, que els processos interpretatius es posen en marxa abans de l'escriptura de les notícies, en la decisió sobre què parlar i sobre com fer-ho, és a dir, la fase prèvia de recollida i tria de dades i històries ja és un procés interpretatiu per se (Gomis, 1987). És el que els teòrics del *newsmaking* dels anys setanta com Gaye Tuchman van intentar

descriure, o el que és el mateix, com els mitjans de comunicació contribueixen a la construcció social de la realitat. En aquest sentit, caldrà parar atenció a la influència en aquestes rutines que poden tenir les noves pràctiques amb intel·ligència artificial. Per exemple, l'aplicació dels algoritmes pot incidir en el procés de construcció de la notícia, una rutina que, en la descripció de Tuchman, el periodista tipificava els esdeveniments a partir de les seves pròpies generalitzacions segons els pressupòsits sobre com dur a terme la seva pràctica professional (Tuchman, 1983). En la identificació de temes noticiables amb IA, per exemple, el criteri que té en compte la màquina és bàsicament quantitatiu, és a dir, sol identificar els temes que s'estan mencionant molt en el context d'internet. En conseqüència, la IA podria influir en el criteri de noticiabilitat. Com veurem, aquesta tecnologia pot intervenir en altres pràctiques que formen part del nucli de la funció editorial del periodisme.

En aquesta primera fase d'identificació de temes susceptibles de ser inclosos a l'agenda, la IA s'aplica amb eines que permeten monitoritzar els entorns digitals en temps real i envien alertes quan detecten tendències o temes més mencionats a les xarxes socials i altres plataformes. Aquestes mateixes eines també tenen funcionalitats per filtrar els tipus de continguts segons les característiques dels perfils emissors, l'impacte, la viralitat, la localització, entre d'altres variables.

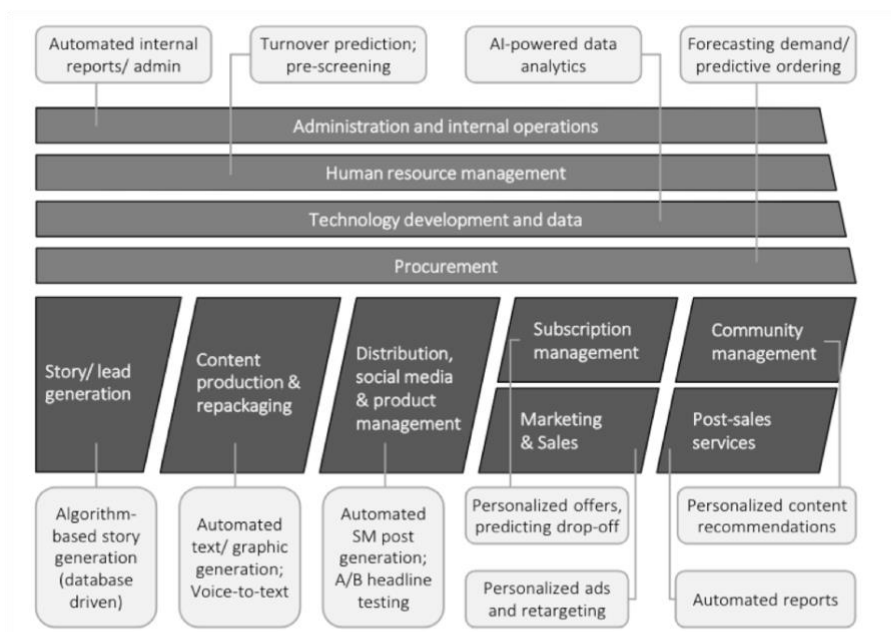
Durant la fase de creació del contingut, la IA intervé en canvis de formats, com les transcripcions d'àudio a text i a la inversa, o de text a vídeo o a infografia. També en la generació automatitzada de textos. Els algoritmes resulten útils així mateix en la recerca documental entre bases de dades pròpies i alienes, per exemple a l'hora d'identificar àudios, imatges, vídeos o altres documents en arxius propis que han de servir per nodrir la informació. També és habitual utilitzar sistemes d'aquesta mena per treure notícies entre grans bases de dades alienes. La IA ha resultat fonamental en projectes d'investigació com els *Pandora Papers* i d'altres similars en què resulta necessari rastrejar molta documentació. El diari peruà *Ojo Público*, per exemple, ha desenvolupat un algoritme que alerta sobre el potencial de corrupció en contractacions públiques. Funes —aquest és el nom que han posat al sistema— rastreja grans bases de dades públiques i, a partir de variables com la zona on es produeix la contractació, l'import del contracte, els contractistes o els adjudicataris, entre d'altres que ha definit l'equip del mitjà, atorga una puntuació que pot ser senyal de potencial corrupció en la licitació.

Finalment, en la fase de distribució, la IA es fa servir, per exemple, per publicar als perfils de les xarxes socials dels mitjans de manera automatitzada, així com per fer proves A/B de titulars,

una pràctica que es serveix del còmput per triar el títol que més clics ajuda a generar. En aquesta fase també s'utilitza per mostrar continguts de manera personalitzada, amb algorismes que primer computen variables com la ubicació de l'usuari, les seccions que consulta amb més freqüència i altres comportaments. Aquestes pràctiques ajuden als mitjans a tenir més visites i a incentivar les subscripcions. Algunes aplicacions permeten avisar la redacció quan un contingut és més susceptible de generar subscriptors. Quan un article reuneix determinats patrons comuns que l'algoritme ha detectat entre els continguts que han generat més subscripcions, aquell tema naixerà sota mur de pagament. En aquest punt en què ja es traspasa la funció editorial amb la del màrqueting, la IA també intervé categoritzant usuaris amb més potencial de convertir-se en subscriptors, aleshores mostra anuncis i crides a l'acció aquests lectors recurrents per convertir-los en usuaris de pagament.

Aquest altre marc (Figura 4) resulta eficient per visualitzar la complexitat de la integració de les tecnologies algorítmiques en els processos de creació i distribució de continguts en els mitjans de comunicació. Es tracta d'una transposició dels usos de la IA en l'esquema de la cadena de valor de Michael Porter (1998).

Figura 4
Exemples d'usos d'IA a la cadena de valor del producte periodístic



Font: (Shaw et al., 2021)

4.4.2 METODOLOGIA DE LA RECERCA EMPÍRICA

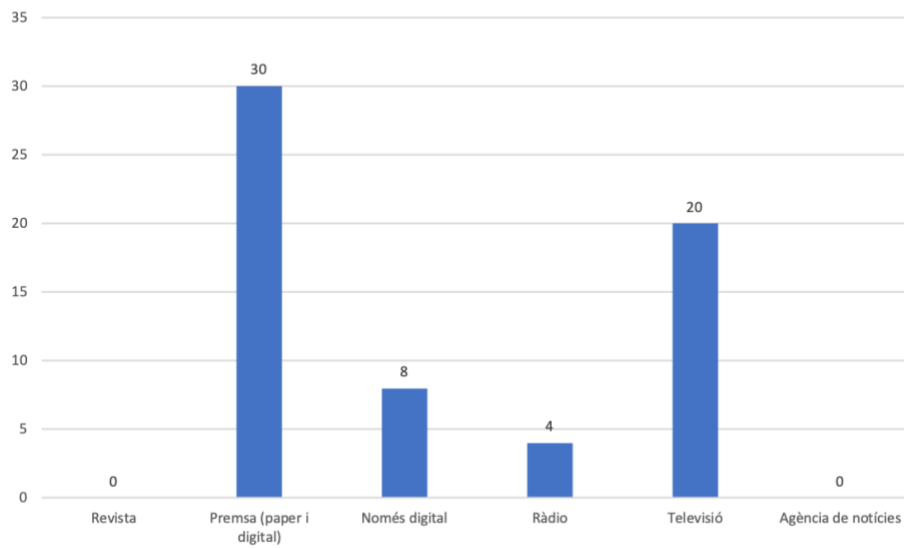
Com hem avançat al primer capítol, a banda de l'anàlisi de literatura del marc conceptual i les qüestions legals, hem emprat un mètode similar al que es va fer servir per a l'adaptació del codi deontològic de la professió periodística del Col·legi de Periodistes de Catalunya a internet (2016). Així, en el moment d'abordar aquesta nova etapa de digitalització caracteritzada per la incorporació de tecnologies exponencials, hem mantingut també com a referència els principis de l'ètica periodística (Alsius et al., 2010). Així mateix, també hem considerat la influència que adquireix la tecnoètica (ètica de la IA enfocada des de les ciències computacionals), tal com també han proposat altres autores (Dörr i Hollnbuchner, 2017).

Per concretar les eines i aplicacions d'IA més susceptibles de ser analitzades, hem fet una primera avaluació de les implicacions ètiques de les noves pràctiques a partir de la confrontació amb els valors que regeixen la professió periodística (Alsius et al., 2010). així com els principis de la tecnoètica (Sabater i De Manuel, 2021).

Per tal de poder identificar les eines ètiques per aplicar a la comunicació computacional des del periodisme hem considerat adoptar un enfocament de base empírica, i per tant en línia amb el punt de vista de les ètiques contractuals (Alsius, 2011). Tanmateix aquesta anàlisi empírica ha estat portada a terme prèvia anàlisi de les disquisicions morals, així com de múltiples guies ètiques i normatives institucionals per a l'ús ètic de la IA.

En el marc del projecte, a més, s'han distribuït dues enquestes. Una primera a professionals de tot l'estat espanyol per valorar la percepció davant l'arribada de la IA als mitjans i les principals preocupacions del sector sobre els seus usos. L'enquesta va ser dirigida principalment a periodistes treballant en l'àrea editorial de mitjans de comunicació en l'àmbit de l'estat espanyol. A l'enquesta es van incloure les següents seccions: suport del mitjà (Figura 5), àmbit de cobertura (Figura 6) i àrea en què la persona participant desenvolupa el seu treball (Figura 7). Pel que fa a les preguntes, es va demanar si existia preocupació en torn a l'arribada de la IA a les redaccions (Figura 8) i en concret, quins potencials perjudicis per a les persones i les societats podria comportar un ús no ètic d'aquestes tecnologies.

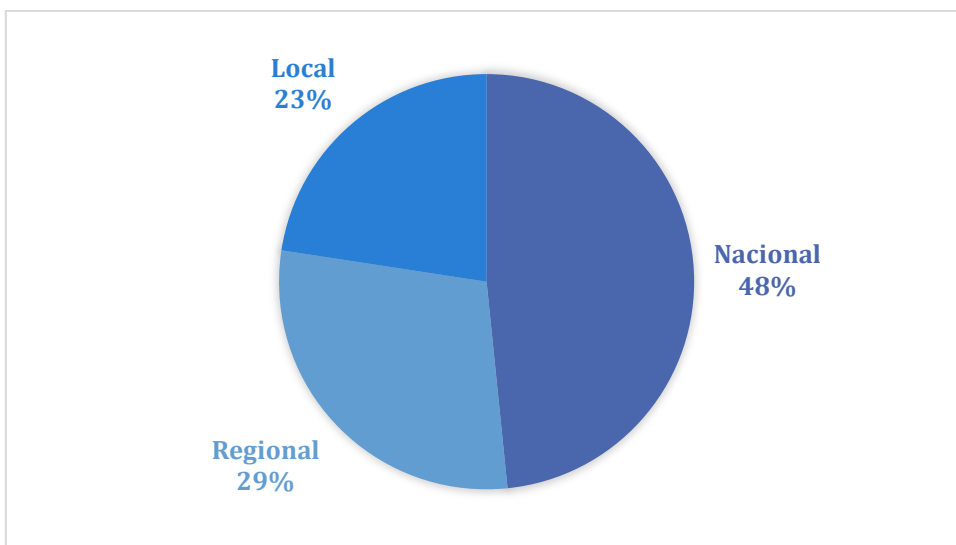
Figura 5
Suport del mitjà



Font: Elaboració pròpia

Figura 6

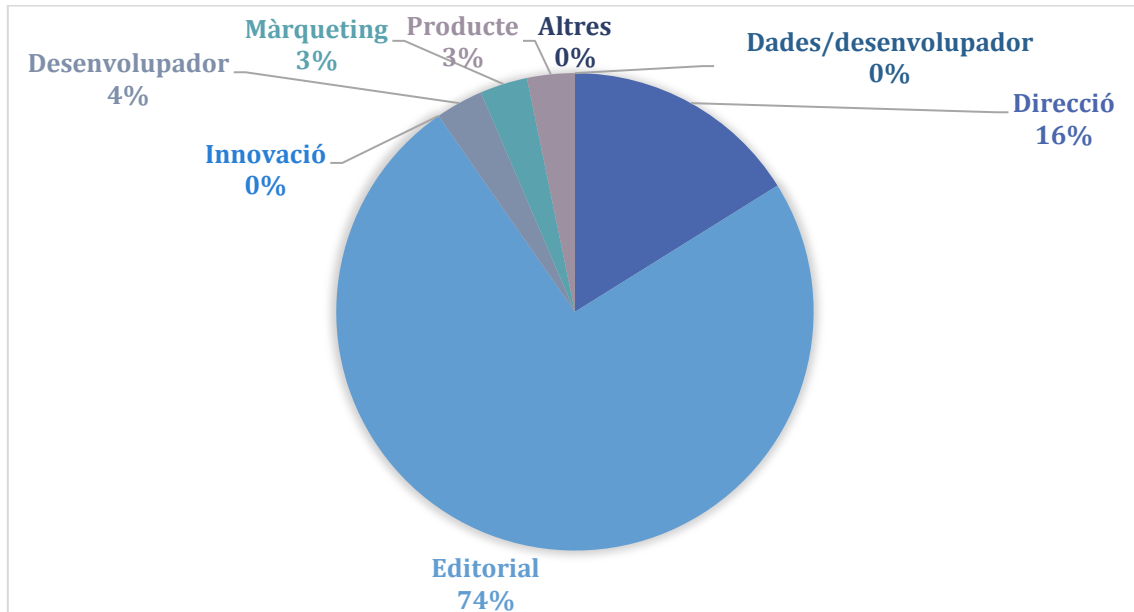
Àmbit de cobertura del mitjà



Font: Elaboració pròpia

Figura 7

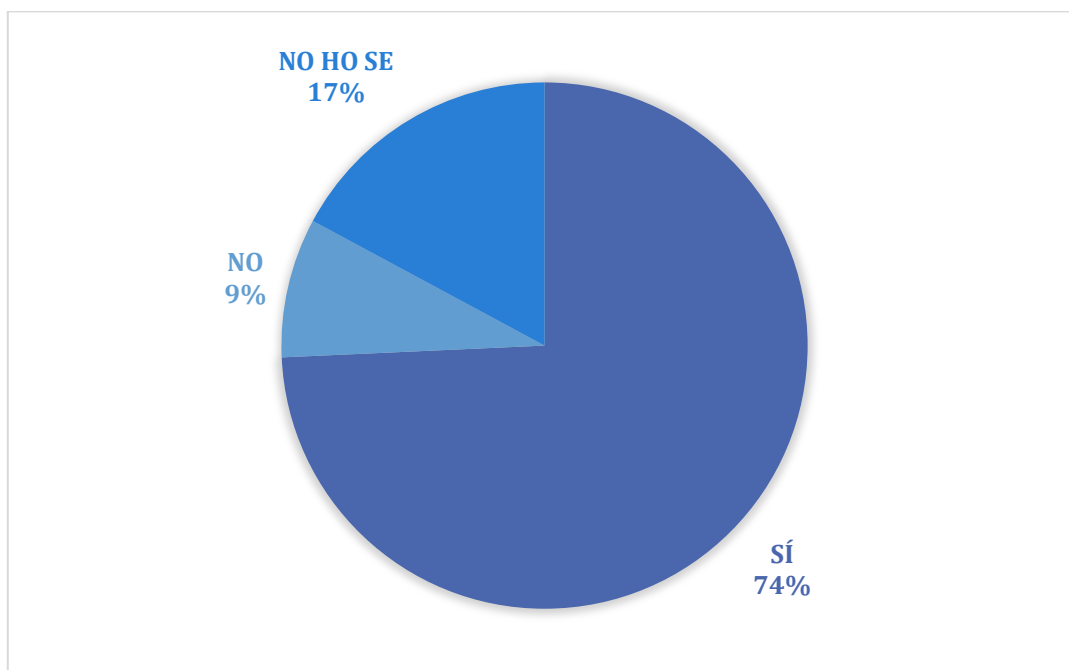
Àrea en la qual desenvolupa el seu treball



Font: Elaboració pròpia

Figura 8

Et preocupa l'impacte de la IA en la qualitat del periodisme des del punt de vista de l'ètica?



Font: Elaboració pròpia

Posteriorment, s'ha portat a terme una segona enquesta dirigida a les direccions de les principals capçaleres catalanes, considerant sota aquesta categoria els mitjans amb més audiència en l'àmbit català, així com les principals capçaleres de cadascuna de les províncies catalanes (Taula 1). Quant al perfil dels participants, van respondre tan des de l'àrea de direcció com des d'àmbits de responsabilitat d'enginyeria o editorial, en funció de la decisió del propi mitjà (Figura 9). A l'enquesta es van incloure les següents seccions: càrrec de la persona que respon, suport del mitjà (Figura 10) i àmbit de cobertura (Figura 11). Respecte a les qüestions que es van plantejar, es va preguntar sobre els tipus d'ús de la IA als seus mitjans, en quines fases del procés de producció les fan servir i quines aplicacions concretes, les oportunitats i els reptes que identifiquen en aquestes tecnologies i si preveuen adoptar-ne més de cara al futur. L'enquesta va ser contestada per representants dels 19 mitjans (Taula 1).

Taula 1

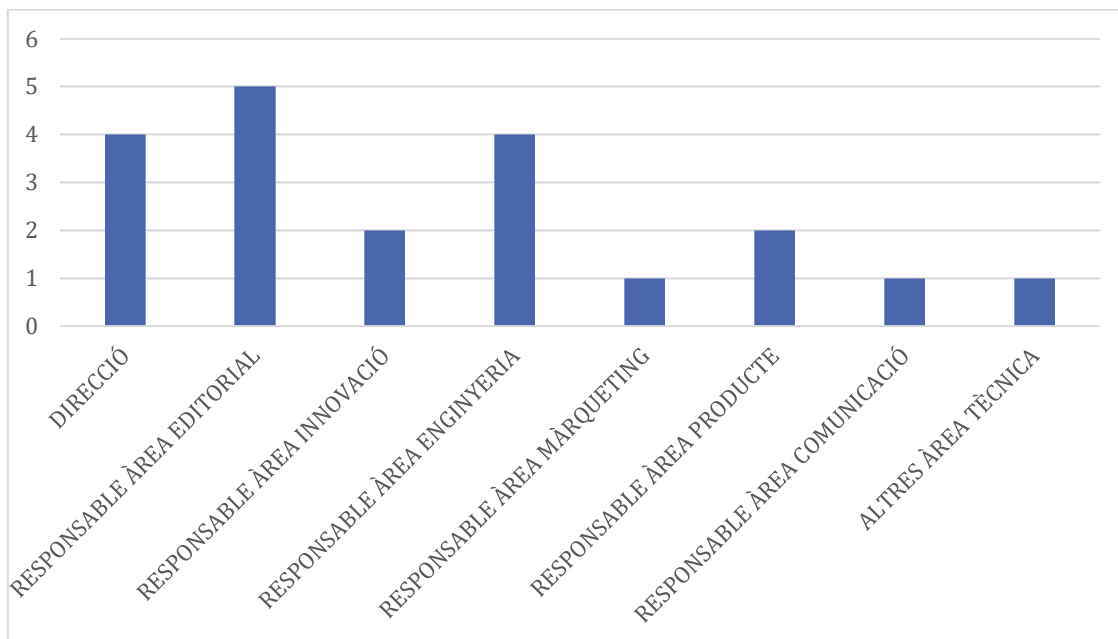
Mitjans que han participat a l'enquesta

MITJANS PARTICIPANTS
TV3
CATALUNYA RADIO
LA VANGUARDIA
DIARI ARA
EL PERIÓDICO
RTVE CATALUNYA
BTV
RAC 1
SPORT
MUNDO DEPORTIVO
SEGRE
EL 9 NOU
EUROPA PRESS CAT
EL PUNT AVUI
DIARI DE GIRONA
DIARI DE TARRAGONA
RAC 105
ACN
REGIÓ 7

Font: Elaboració pròpia

Figura 9

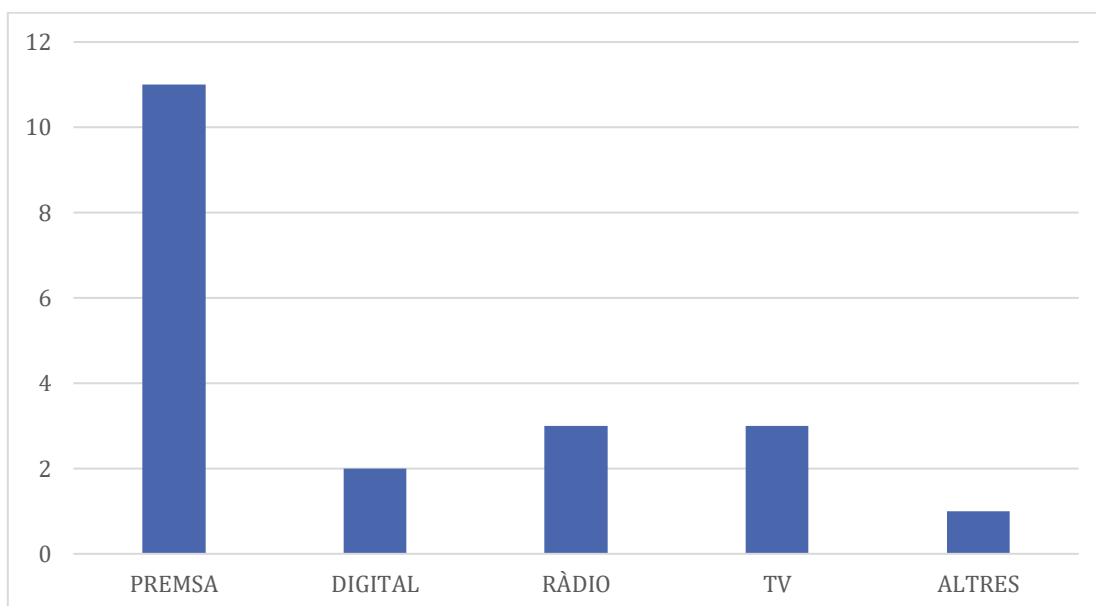
Càrrec de la persona que respon



Font: Elaboració pròpia

Figura 10

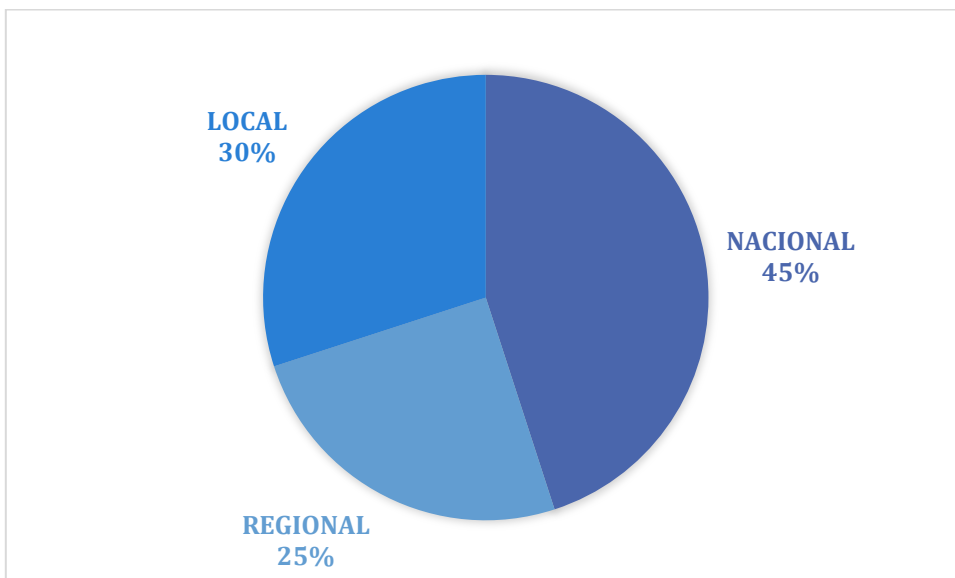
Suport del mitjà



Font: Elaboració pròpia

Figura 11

Àmbit de cobertura del mitjà



Font: Elaboració pròpia

Així mateix, es van celebrar dues trobades amb mitjans i experts per tal de fer una valoració conjunta sobre els dilemes que presenten aquestes noves aplicacions i els principis ètics que poden posar en joc. Es va conformar un grup de discussió amb 5 representants d'alguns dels principals mitjans de comunicació catalans i un seminari amb el comitè d'experts nacional.

Quant al grup de mitjans de comunicació, van intervenir-hi professionals amb perfils de responsabilitat amb visions estratègiques i operatives de les diferents àrees implicades en els processos d'automatització: enginyeria (1), editorial (3), màrqueting (1) i pertanyents a mitjans de comunicació de diferents suports: premsa, televisió i ràdio. Totes aquestes trobades i entrevistes van ser enregistrades, transcrites i analitzades a través d'un procés qualitatiu de codificació de conceptes i temes.

COMPOSICIÓ DEL GRUP MITJANS DE COMUNICACIÓ

Àlex Gutiérrez -Cap de mèdia -ARA

Geni de Vilar - Adjunta a direcció - Mitjans Digitals - CCMA

Alberto Alejo - Cap d'enginyeria - CCMA

Ximo Blasco - Responsable d' informatius -TV3

Mariano Fernández - Chief Digital Officer (CDO) -La Vanguardia

Finalment, es va distribuir una enquesta entre els experts i els mitjans que van participar als seminaris per confirmar els consensos que havíem identificat a través de l'anàlisi i la codificació de les transcripcions dels enregistraments.

També s'han portat a terme tres entrevistes semi estructurades. Dues en format virtual a experts internacionals referents en periodisme computacional Nicholas Diakopoulos i Charlie Beckett. Finalment, es va dur a terme una tercera entrevista, telefònica en aquets cas, amb Alberto Alejo, Cap d'enginyeria dels mitjans de la CCMA.

El comitè d'experts ha estat compost per acadèmics especialitzats en ètica periodística, filosofia, ciberètica i periodisme computacional i innovació en periodisme, les principals disciplines sota les quals es desenvolupa aquest treball de recerca.

COMPOSICIÓ DEL COMITÈ D'EXPERTS

Internacional

Nicholas Diakopoulos: Principal referent internacional en periodisme i IA. És professor de Comunicació i Informàtica a la Northwestern University, on és director del Computational Journalism Lab (CJL) i director d'estudis de postgrau del programa de doctorat en tecnologia i comportament social (TSB).

L'àmbit de recerca de Diakopoulos és el periodisme computacional amb projectes de recerca actius sobre (1) retiment de comptes i transparència algorítmica, (2) automatització i algorismes en la producció de notícies i (3) xarxes socials en contextos de notícies. És autor del guardonat

llibre *Automating the News: How Algorithms are Rewriting the Media* de Harvard University Press.

Charlie Beckett: Director fundador de Polis, el think-tank per a la investigació i el debat al voltant del periodisme i la societat internacional al Departament de Mitjans de Comunicació i Comunicacions de la London School of Economics. Beckett lidera el projecte Polis JournalismAI, una iniciativa global que té com a objectiu informar els mitjans sobre el potencial que ofereixen les tecnologies basades en la IA. Fa recerca, ofereix materials de formació, i dinamitza una xarxa per compartir bones pràctiques i innovació en periodisme.

Nacional

Albert Sabater: Director del de la Càtedra - Observatori d'Ètica en Intel·ligència Artificial de Catalunya. Professor Serra Húnter de Sociologia. També és Coordinador d'Estudis del Màster d'Economia de l'Empresa (Business Economics) a la Facultat de Ciències Econòmiques i Empresariales de la Universitat de Girona.

José Alberto García Avilés: Catedràtic de Periodisme a la Universitat Miguel Hernández d'Elx. Les seves línies d'investigació se centren en la innovació, la qualitat i l'ètica periodística. És autor de més d'un centenar de publicacions sobre comunicació i cofundador de la xarxa InnovaMedia.

David Casacuberta: Professor de Filosofia de la Ciència a la UAB, la seva línia de recerca actual són els impactes socials i cognitius de les TIC. Actualment és membre de el Grup de Treball d'Ètica, Seguretat i Regulació de Bioinformàtica Barcelona, investigador de el grup consolidat GEHUCT (Grup d'Estudis Humanístics en Ciència i Tecnologia). Ha rebut el premi Eusebi Colomer de la Fundació Epson a el millor assaig amb el seu llibre *Creació col·lectiva*.

Karma Peiró: Periodista especialitzada en les Tecnologies de la Informació i la Comunicació (TIC) des de 1995. Co-directora de la Fundació Visualització per a la Transparència (VIT) que promou l'ús de les dades obertes per donar poder a la ciutadania i la retiment de comptes de la informació pública. Membre del Consell Assessor de l'OEIAC i de el Comitè Ètic de la UPC.

Joan Rosés: Periodista. Des de 2017 és editor de Collateral Bits, publicació digital dedicada a l'anàlisi de l'impacte de la tecnologia en la societat. Abans havia estat director del Clúster Audiovisual de Catalunya i director d'Activa Multimèdia, centre d'innovació i recerca de la Corporació Catalana de Mitjans Audiovisuals.

4.4.3 REPTES ÈTICS SOBRE L'ÚS DE LA IA PER FER PERIODISME

Al'hora d'establir un marc per a l'ús ètic de la IA als mitjans de comunicació, convé d'entrada tenir en compte la legalitat vigent. La regulació més rellevant fins avui és l'esmentat i conegut RGPD, que s'aplica especialment en qüestions relacionades amb la privacitat i la gestió de les dades. D'altra banda, convé tenir en compte també la mencionada proposta per a la regulació de la IA de la UE.

Aquest nou marc legal estableix una categorització per nivell de risc en funció de l'impacte que el sistema d'IA pugui tenir en la vida de les persones i en la vulneració de drets fonamentals. Des del punt de vista de la comunicació, s'identifiquen dues aplicacions que la proposta considera com a potencials causants de riscos per a la ciutadania i en el seu articulat avança la manera com regularà el seu ús.

En primer lloc, es refereix al contingut fals, i concretament als anomenats *deepfakes*, imatges que suplanten persones i poden induir a engany el receptor. En aquest sentit, obliga l'emissor a comunicar quan es tracta d'un *deepfake*. Tot i que d'aquest punt es pot deduir que la norma pretén evitar que actors interessats difonguin propaganda i altres falsedats, els mitjans podrien veure's afectats en cas que optessin per fer servir imatges generades amb IA que no corresponguessin a la realitat física, com per exemple emprar presentadors virtuals. En tots els casos, hi haurà l'obligació de comunicar que una imatge no és real.

En segon lloc, la proposta també es refereix als bots, sistemes de conversa que s'acostumen a fer servir en l'àmbit d'atenció al client, entre d'altres. En el cas del periodisme, s'utilitzen sovint per respondre a preguntes sobre notícies i per compartir altres continguts. En aquest cas, la proposta també obliga a comunicar als usuaris quan estan parlant amb una màquina.

A banda d'aquests dos punts, també cal tenir en compte que la nova normativa emplaça a tots els sectors a establir marcs propis d'autoregulació per a les diferents aplicacions de la IA en cada àmbit, i que hi ha altres punts d'aquesta proposta que, malgrat no referir-se de manera explícita a l'àmbit de la comunicació, convindria avaluar amb deteniment per si podrien ser aplicables en determinats casos concrets; per exemple, l'ús de sistemes amb potencial de manipular el comportament humà, com els algoritmes de personalització i d'altres que tenen el potencial d'afectar persones amb vulnerabilitats psicològiques o menors d'edat.

Des del punt de vista de la comunicació convé subratllar que aquesta regulació posa focus en els usos de la IA per part de les plataformes i no es refereix d'una manera clara i específica als mitjans de comunicació tradicionals. Aquesta és una qüestió problemàtica que, d'entrada, genera dubtes sobre si finalment s'acabarà tractant en algun moment, tenint en compte que l'adopció d'eines i aplicacions amb IA en el periodisme pot implicar noves formes de poder mediàtic així com una profunda transformació estructural de l'esfera pública digital: (Helberger i Diakopoulos, 2022):

The introduction of AI-driven tools in the media leads not only to new forms of media power but also to a deeper structural transformation of the digital public sphere, which again raises new questions about digital citizenship and inclusion in democratic participation, as well as agency and responsibility, bias and discrimination, privacy and data protection, and user autonomy and editorial freedom from manipulation.

Un altre dels dubtes que generen aquestes regulacions és que no responsabilitzen prou als proveïdors d'IA dels mitjans de comunicació (Helberger i Diakopoulos, 2022). Aquest punt genera dubtes perquè els mitjans tenen poca capacitat d'intervenció en els aspectes del software dissenyat externament i per tant en aquests contextos és molt possible que hagin de renunciar al dret de reflectir els seus valors professionals en aquests sistemes. En propers apartats examinarem com el poder de les plataformes pot condicionar la independència dels mitjans d'aquestes.

El fet és que encara no s'han pogut valorar els possibles efectes a mig o llarg termini d'usos no ètics de la IA en els mitjans, degut, probablement, a la incipient incorporació d'aquestes tecnologies a les rutines periodístiques. És per això que des de certs àmbits acadèmics es considera necessària la previsió de recerca futura per analitzar fins a quin punt aquestes regulacions s'han implementat, si són efectives i adients i fins a quin punt o sota quines condicions la incorporació de la IA a determinades rutines periodístiques pot considerar-se d'alt risc, encara que els possibles efectes negatius es produeixin a llarg termini (Helberger i Diakopoulos, 2022).

Tornant a les normes específiques actuals, es pot inferir que el gran problema que amoina a la UE sobre l'ús de la IA en l'esfera de la comunicació és la desinformació, una inquietud que l'àmbit periodístic també comparteix, en termes generals. En ser preguntats pels efectes de la integració dels usos de la IA en les rutines periodístiques, la majoria dels professionals enquestats en el marc d'aquesta recerca van mostrar neguit pel fet de generar desinformació des dels mateixos

mitjans. Els conceptes més mencionats en preguntar per les principals preocupacions, han estat, a banda de la desinformació, els biaixos, les cambres d'eco, la polarització, les vulneracions de la privacitat, la deshumanització dels continguts i la pèrdua de la qualitat (Figura 12).

En aquest punt val la pena destacar que aquestes grans preocupacions del sector (desinformació, cambres d'eco, privacitat...) són conceptes que coincideixen amb els principals problemes de la mediació algorítmica que hem identificat a través de la literatura analitzada i que hem exposat al segon capítol d'aquesta tesi, on hem tractat els problemes que es deriven de la gestió comercial de la IA en la mediació.

Figura 12

Principals preocupacions sobre l'adopció de la IA per part del periodisme



Font: Elaboració pròpia

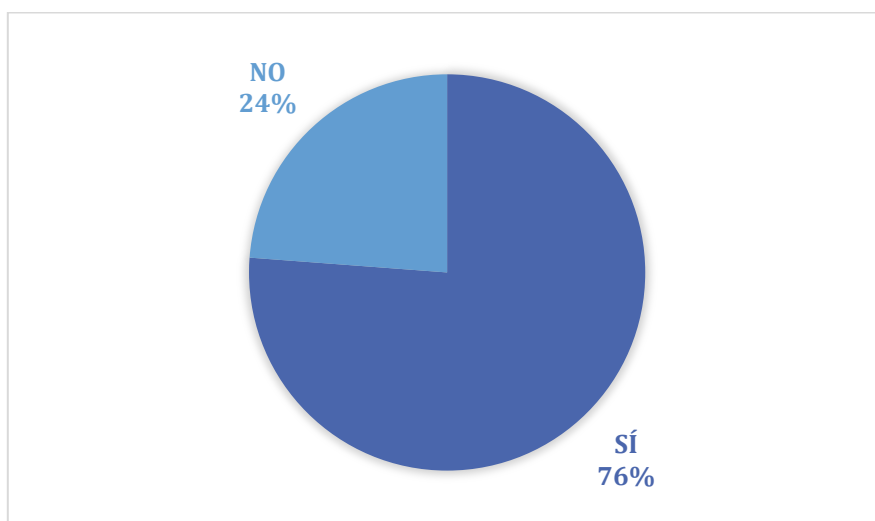
4.4.4 ESTAT DE LA QÜESTIÓ A CATALUNYA

A Catalunya, els principals mitjans de comunicació han integrat la IA (Figura 13) a totes les fases de producció en una mesura similar, per bé que destaquen els usos durant la fase prèvia a la creació del contingut orientats a identificar temes noticiables (Figura 14). Pel que fa a les aplicacions concretes, la més comuna és emprar algoritmes per detectar tendències informatives i recopilar informació. En altres fases del procés de producció i distribució, l'ús més freqüent és l'optimització de motors de recomanació de continguts i altres funcions relacionades amb l'àmbit comercial com la clusterització i l'anàlisi del comportament dels usuaris (Figura 15).

També és cada cop més habitual la generació automatitzada de textos, vídeos i infografies, classificar dades desestructurades i les proves A/B de titulars, una aplicació que consisteix a publicar una mateixa peça amb diferents titulars per tal que l'algoritme computi quina de les versions millora les mètriques d'entrada a la notícia o el CTR (*Clickthrough Rate*). Tots els mitjans enquestats es proposen incorporar més eines o aplicacions d'IA i aquells que encara no ho han fet preveuen començar a utilitzar-la (Figura 16).

Figura 13

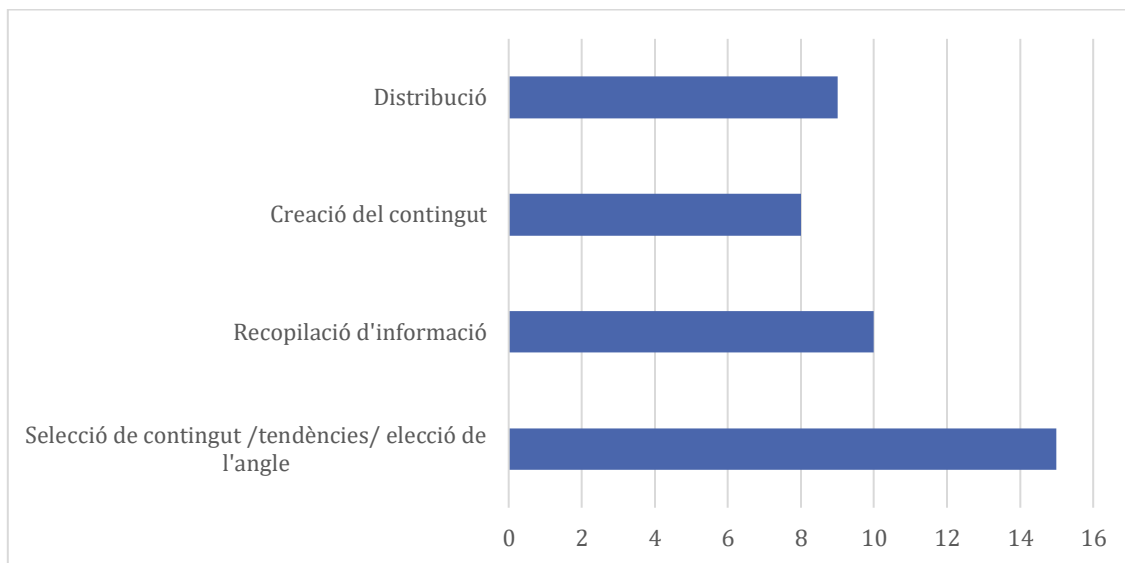
En el vostre mitjà de comunicació feu servir IA o altres sistemes algorítmics?



Font: Elaboració pròpia

Figura 14

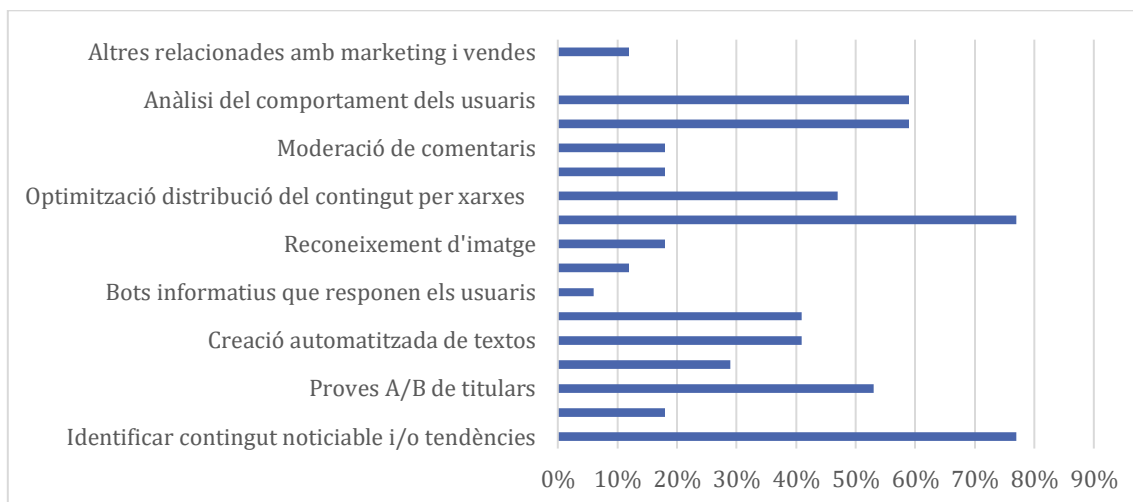
En quina fase de producció feu servir IA o altres sistemes algorítmics?



Font: Elaboració pròpia

Figura 15

En quines aplicacions concretes feu servir IA o altres sistemes algorítmics?²

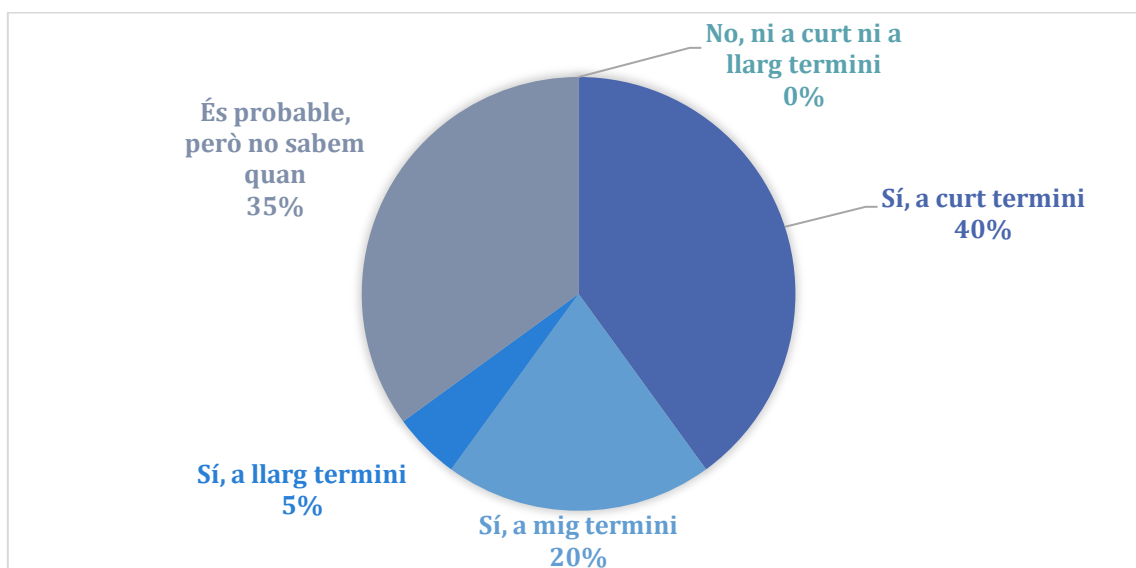


Font: Elaboració pròpia

² Categorització adaptada de l'informe *The next wave of disruption: Emerging market media use of artificial intelligence and machine learning* (Shaw, 2021)

Figura 16

Penseu fer servir més aplicacions amb IA o altres sistemes algorítmics?



Font: Elaboració pròpia

Els mitjans catalans identifiquen dos grans avantatges en aquestes tecnologia: en primer lloc, el potencial del processament de dades per conèixer les audiències i adaptar el producte en conseqüència; i, en segon lloc, l'eficiència en la gestió de processos interns, el suport a la sistematització i recerca entre grans volums documentals propis o aliens, així com la generació automatitzada de continguts.

“La IA pot resultar un suport per detectar allò que interessa al nostre usuari digital, per oferir-li el que realment l'interessa de forma segmentada i no impactar amb informació o serveis que no siguin rellevants per als seus interessos”.

Responsable en l'Àrea de Producte - Ràdio

“Els canvis van tan ràpid que conviuen diversos models, i als mitjans ens cal ajuda per perfilar millor i oferir a l'espectador allò que vol”.

Responsable de Comunicació - Televisió

“La IA i els sistemes algorítmics poden ser bons assistents per a millorar la productivitat dels periodistes. Poden ser especialment útils per identificar històries (a través de l’anàlisi de grans quantitats de contingut (per exemple de xarxes socials) i/o per personalitzar continguts per a l’usuari segons les seves preferències implícites o explícites”.

Directiu - Agència de Notícies

“La IA ens pot ajudar en molts àmbits com en processos de documentació, en la recerca i la producció de material audiovisual, en la detecció de notícies falses, entre d’altres”.

Responsable en l’Àrea Enginyeria - Televisió

“La IA facilita la recopilació d’informació i el seu processament per simplificar el procés de presa de decisions”.

Responsable en l’Àrea Editorial - Televisió

“Les eines d’IA són un suport i una ajuda al criteri editorial. Per exemple, entre un titular A/B un algoritme pot dir-te quin té més atractiu per al lector, però no quin és el que s’ajusta més a la realitat, veracitat i rigor que ha de regir la nostra feina”.

Direcció - Premsa

Quant als reptes, els mitjans destaquen principalment la necessitat de formació dels equips i d’incorporació de nous perfils, així com les potencials friccions en els processos d’integració de la IA a les rutines. En una mesura similar, també resulta un desafiament important disposar de capacitat d’encert en el moment de determinar eines i pràctiques amb IA adients, que justifiquin la inversió que requereixen.

D’altra banda, les capçaleres amb estructures més reduïdes tenen dubtes sobre si totes les aplicacions estan també al seu abast. Els mitjans també consideren un repte la gestió de la integració de la IA de manera que no devalui la qualitat de la informació.

“El repte és encertar”.

Responsable Àrea de Màrqueting - Premsa

“Els equips tècnics tradicionals no estan preparats per aquest repte. Calen perfils especialistes en aquest tema dins dels equips tècnics i perfils híbrids a la redacció”.

Àrea Direcció – Premsa

“Un dels grans reptes a l’hora d’incorporar la IA és identificar les eines adients en un mercat que comença a tenir diferents productes i actors, cosa que ens obliga a una bona avaluació de costos per maximitzar la inversió i el seu retorn”.

Responsable Àrea de Producte - Ràdio

“El repte és que els redactors utilitzin aquestes eines per millorar en el dia a dia i les percebin com una ajuda, no com una amenaça pel seu lloc de treball. Les notícies automatitzades són continguts amb poc valor afegit que haurien de permetre als periodistes centrar-se en històries que requereixen, tant si com no, de la intervenció humana i que són la raó de ser del periodisme”.

Àrea Direcció - Premsa

“Es tracta d’aplicacions pensades per entorns molt grans de molt volum i difícils d’aproximar a entorns més petits”.

Responsable Àrea Enginyeria - Premsa

“El principal repte és que tot va tan ràpid que s’oblida el factor humà, la necessitat de fer molta pedagogia i d’abandonar formats divulgatius que ajuden a contextualitzar el que està passant i comprendre-ho”.

Responsable Àrea Comunicació - Televisió

“Hem de vigilar que els automatismes no ens facin perdre el rigor en l'enfocament”.

Responsable Àrea Editorial - Televisió

“Econòmicament no s'està invertint molt a desenvolupar eines i algoritmes sofisticats per al sector periodístic perquè és un sector que mou pocs diners en comparació amb d'altres indústries en les què l'ús de la IA està més avançat”.

Àrea Direcció - Agència de Notícies

“El repte principal és aconseguir que els ritmes d'evolució dels processos mitjançant IA vagin acompanyats del coneixement i bon ús dels professionals del mitjà”.

Responsable Àrea Editorial - Televisió

4.4.5 ANÀLISI I RESULTATS

Un cop revisada la literatura detallada més amunt, així com els resultats de les enquestes que vam distribuir entre professionals dels mitjans de tot l'estat, havent analitzat els debats dels grups de discussió, les transcripcions de les reunions i entrevistes amb experts acadèmics, així com les respostes dels mitjans de comunicació que van participar, vam identificar el següents temes clau:

- Contingut automatitzat amb criteri editorial
- Personalització que conviu amb la diversitat i promou la salut de l'esfera pública
- Supervisió i qualitat de les dades per evitar biaixos
- Preservar la privacitat dels usuaris amb responsabilitat
- La qualitat del periodisme passa per posar en valor el factor humà
- Poder de les plataformes i independència informativa
- IA per enfortir els valors del periodisme

4.4.5.1 Contingut automatitzat amb criteri editorial

Els sistemes de generació automatitzada de continguts, així com els bots informatius, ja són força habituals als principals mitjans catalans. A la IA se li encarreguen tasques com transformar dades en infografies, convertir textos en vídeo o, l'ús que està per ara més estès, la generació automatitzada de textos. Aquests sistemes més habituals de redacció proporcionen una peça final que en molts casos pot semblar escrita per un periodista, però que en realitat resulta una notícia molt semblant a qualsevol altra de generada per la mateixa màquina: una estructura estandarditzada i, per tant, previsible en què canvien les principals dades, que poden ser resultats esportius o cotitzacions de borsa, entre d'altres. En la creació automatitzada de notícies, el paper del professional es limita a establir l'estructura narrativa i a indicar al sistema quines dades haurà d'utilitzar per omplir els camps buits d'una mena de formulari inserit en un cos narratiu.

Aquests sistemes de redacció de notícies es fan servir per reportar sobre temes més susceptibles de ser estructurats, com els esmentats resultats esportius, les cotitzacions a la borsa o previsions meteorològiques, entre d'altres. El principal valor periodístic que els mitjans donen a la creació de notícies amb IA és poder oferir cobertura a temes que afecten petits grups de població, uns temes que, sense aquesta tecnologia, no podrien dur a terme per manca de recursos, com per exemple resultats electorals de petites localitats o informació sobre esdeveniments esportius de categories menors, entre d'altres.

“Si trobes una tecnologia que pot portar a terme una tasca com, per exemple, redactar una previsió meteorològica, no és una gran idea encarregar-li? En el passat hi havia moltes feines que feien els humans i que avui han deixat de fer, i la majoria eren molt avorrides. Si creus que el teu treball el pot fer una màquina, potser t'has de preguntar quina mena de treball fas i si no hauries de canviar”.

Charlie Beckett

Per la seva banda, els bots informatius són agents que interactuen amb l'usuari responent a la informació que aquest demana. Els xatbots, com explicàvem més amunt, s'utilitzen en molts altres sectors en tasques com l'atenció al client amb la finalitat de donar resposta a preguntes freqüents, per exemple. Cal recordar que l'ús de xatbots és una de les pràctiques que la UE pretén regular per llei i s'hi refereix en el títol IV de la proposta de regulació de la IA, en què

obliga els proveïdors a assegurar que l'usuari té clar que està interactuant amb una màquina i no amb una persona.

“Cal un pacte i una actitud proactiva per traslladar l'ètica periodística a aquestes noves situacions i així evitar que siguin veus alienes a la professió les que prenguin les decisions. És necessari que els usuaris ho sàpiguen i entenguin que el model de periodisme és molt diferent del de les plataformes, en què la preocupació ètica és mínima”.

José Alberto García Avilés

“Sovint, l'automatització de contingut té l'objectiu d'arribar a com més gent, millor. I això és volum, no qualitat. Aquest és un aspecte fonamental a considerar des del punt de vista ètic. Sense tenir en compte això, no entenem on estem situats”.

Albert Sabater

Des del punt de vista de l'ètica, en els diferents processos de generació automatitzada de continguts, la supervisió humana resulta fonamental per evitar errors en els resultats que podrien donar lloc a contingut incorrecte o esbiaixat, i per tant la seva utilització sense control posaria en risc la veracitat de la informació. S'han donat diferents casos en aquest sentit, per exemple, el 2015 programari d'aquesta mena va publicar una notícia sobre una suposada caiguda de Netflix a borsa, quan en realitat havia pujat més del doble.

“Un dels principals problemes té a veure amb la incertesa i la imprecisió en els continguts que poden produir els sistemes d'IA. Generar informació imprecisa sense poder validar-la comporta dilemes ètics, en conseqüència, pot comprometre el rigor que forma part dels objectius del periodisme”.

Nicholas Diakopoulos

“LA IA pot resultar de gran ajuda. El problema és quan es barregen creadors humans i màquines que redacten, aleshores ja no saps qui està construint la notícia. Crec que serà molt difícil posar fronteres”.

Joan Rosés

Aquí també s'han donat casos, com per exemple el d'un mitjà que s'havia proposat crear infografies —concretament núvols de paraules— a partir de transcripcions automatitzades. Aquest sistema de transcripció cometia errades; per tant, en crear la infografia, la representació del volum de paraules no era acurada perquè no ho havia estat la transcripció prèvia. En aquesta ocasió, el mitjà sí que va supervisar el resultat i, en comprovar que no era rigorós, va optar per no fer servir l'automatisme malgrat impliqués més recursos o, el que és el mateix, va dedicar persones a fer la transcripció prèvia. Un altre problema que va tenir lloc al mateix mitjà, en aquest cas en la generació automatitzada de textos, va ser que el sistema oferia resultats amb biaix de gènere, concretament, atribuïa el masculí a càrrecs que es citaven en femení (com “alcalde” en comptes d’“alcaldessa”). De nou, editors humans van fer el repàs de tots els textos per fer les esmenes corresponents. En aquest cas, a més de la supervisió, el mitjà va etiquetar totes aquestes notícies com a automatitzades.

“Com a lectora i com a ciutadana vull saber que aquest contingut ha estat generat per una màquina. No poso en dubte que la informació sigui correcta. Simplement espero que el mitjà sigui transparent en aquest punt”.

Karma Peiró

Amb aquests exemples podem extreure que tant la transparència —assenyalant els textos automatitzats— com la supervisió humana, resulten fonamentals per evitar generar desinformació. Especialment en aquests usos de la IA, la majoria d'experts, en línia amb la propera legislació europea en el cas dels bots, considera que cal transparència, que la ciutadania ha de poder identificar d'una manera clara quan un contingut és generat per una màquina. A més d'afirmar el compromís periodístic amb la veritat, la transparència esdevé la protecció més eficient davant possibles errors, coincideixen els experts. De nou, cal l'explicabilitat i l'evidència d'on surt qualsevol mena d'informació per tal d'actuar de manera responsable.

“Un altre problema és identificar el que és automatització, i això és segurament més important que establir què és IA, perquè la IA ja forma part de molts processos. Per exemple, fer servir IA per transcriure entrevistes, és fer periodisme automatitzat? No ho crec. Així que es tracta de determinar l'automatització i el nivell de consciència dels usuaris del grau d'automatització que està present quan consumeixen un contingut. Bàsicament que puguin entendre que si hi ha inexactituds tenen a qui demanar explicacions sobre la raó d'aquests errors”.

Nicholas Diakopoulos

4.4.5.2 Personalització que conviu amb la diversitat i promou la salut de l'esfera pública

L'audiència i el que vol llegir és cada cop més previsible gràcies al rastre en forma de dades que deixem quan naveguem per internet. Els algorismes de personalització són un tipus de sistema de filtratge d'informació que determina quin contingut d'un univers d'informació preferirà un usuari a partir de l'estudi de dades com la ubicació, el dispositiu, els seus patrons de comportament o els hàbits d'usuaris similars, entre d'altres. Aquesta mena de sistemes fa créixer les mètriques de resultats, com les pàgines vistes, el temps que els usuaris passen a les pàgines dels mitjans o la freqüència de les visites al web.

Els continguts que trobem a la línia de temps a les xarxes socials són un exemple de selecció que s'ha fet *a mida* basant-se en la informació que la plataforma disposa de cadascun de nosaltres, com per exemple on ens connectem, amb quin dispositiu, quina mena de continguts consumim, quanta estona, qui són els nostres amics i quins continguts consumeixen ells, etcètra.

Des que Cass Sunstein formulà la teoria de les cambres d'eco (1999) i Pariser publicà la seva investigació sobre els filtres bombolla (Pariser, 2011), s'ha publicat força recerca que treu conclusions molt diferents, és a dir, que aquests efectes bombolla que suposadament genera el consum d'informació digital tenen un abast molt limitat (Guess, 2018).

“En la nostra recerca no hem identificat d'una manera clara l'existència de bombolles d'informació. No es tracta de ser a una bombolla o no ser-hi. No existeix aquesta divisió clara que defineix una bombolla.

Així que és la metàfora de la bombolla, el problema. El que sí que hem observat és una certa tendència a estar més exposats a un determinat punt de vista de manera que la gent veu una mica més d'un tipus de contingut i una mica menys de altre. La realitat s'assembla més a això”.

Nicholas Diakopoulos

En realitat, asseguren altres experts, no és tant la personalització com una combinació entre aquest contingut *a mida*, el biaix de confirmació (reafirmació en els propis punts de vista) i el fet que el consum de la informació es produeixi a l'entorn digital, perquè mentre naveguem ens trobem en un ecosistema en què estem connectats amb les nostres comunitats, el qual afavoreix el fet de compartir. A diferència del consum *offline* i solitari, estar connectat acostuma a

desencadenar una acció comunicativa individual que busca reafirmar els propis biaixos, cercar l'aprovació dels qui pensen com nosaltres i alhora criticar els contraris (Tufekci, 2018).

Aquesta combinació de personalització, biaix de confirmació i ecosistema digital que incentiva la viralització de contingut emocional (Brady, 2017) seria el que promouria la creació d'estats semblants a les bombolles d'informació que Sunstein i Pariser van teoritzar i que desperten inquietuds sobre els potencials efectes que la fragmentació i la individualització poden causar a la democràcia (Katz, 1996). L'ordre social està basat en una realitat compartida (P. L. Berger i Luckmann, 1966) i per tant la manca d'experiències comunes, així com la falta d'exposició de les persones a diferents punts de vista posarien en risc la cohesió social. Tal com desenvolupa a la seva obra el professor Carlos Ruiz (2016), disposar de contextos compartits resulta bàsic per a la democràcia. El paper de la comunicació, tal com ha escrit, esdevé fonamental:

El diálogo, la deliberación racional entre seres humanos, no es solo un elemento constitutivo de la ética de la comunicación, sino que es el camino para intentar una convivencia política que sea moralmente aceptable. Dicho en otras palabras, "es un proceso de formación discursiva de la voluntad colectiva" (p.98).

L'informe del grup d'alt nivell sobre llibertat de premsa i pluralisme de la UE expressà així la seva preocupació: "Increasing filtering mechanisms makes it more likely for people to only get news on subjects they are interested in, and with the perspective they identify with. Such developments undoubtedly have a potentially negative impact on democracy" (High Level Group on Media Freedom and Pluralism, 2013).

"Em pregunto si, a la llarga, aquests algoritmes erosionaran una de les funcions dels mitjans de comunicació que consisteix a crear contextos compartits. Sense això pot ser difícil mantenir les converses que potser serien necessàries amb la teva comunitat. Una visió de la relació del periodisme amb la democràcia és que els mitjans creen les condicions per tenir aquests debats. Així que, si no fem servir els nostres mitjans algorítmics amb aquesta finalitat, hem de ser-ne conscients i anar amb compte".

Nicholas Diakopoulos

Des del punt de vista editorial, la personalització resulta útil per ajudar el lector a optimitzar el temps que passa al mitjà i, per tant, a augmentar el valor del producte. Per exemple, mostrar determinats continguts a partir de la ubicació de l'usuari pot fer més útil i interessant una selecció de notícies, que inclouran o deixaran d'incloure temes en funció de la proximitat

geogràfica. Tot i ser la forma més controvertida, la personalització per interessos també és utilitzada en els mitjans i a priori no implica cap mena de dilema: si un usuari mai no accedeix a continguts sobre esports, però en canvi entra molt sovint a consultar informació relacionada amb la literatura, pot tenir sentit que un dels articles que li mostri un motor de recomanació sigui un contingut relacionat amb aquest tema pel qual ha mostrat interès de manera reiterada.

Convé tenir en compte també les expectatives de les noves generacions, que han crescut fent consum cultural condicionat per sistemes algorítmics com els de Spotify, Netflix o Amazon. Aquests sectors de la població ja esperen dels mitjans que els ajudin a optimitzar el seu temps. Tal com indiquen alguns informes de tendències, els joves esperen més personalització, també per part dels *legacy media*: “Traditional news brands feel their job is to tell people what they should know. Young people want that to an extent but they also want what is useful to know, what is interesting to know and what is fun to know” (Galan, 2019).

“Si la IA és per oferir recomanacions que donin més context, benvinguda sigui. Això és donar més servei a l’usuari”.

Mariano Fernández - CDO - La Vanguardia

“La personalització ajuda a proporcionar informació rellevant per a la gent. I per descomptat, quan fem servir la IA per personalitzar sempre tenim opció: podem indicar a l’algoritme que actui amb serendipitat, o en el cas dels butlletins personalitzats podem incloure un recull d’agenda elaborat per l’editor, per exemple. Així que en cert sentit crec que es tracta d’un tema senzill. Però requereix que com a mitjà en prenguis el control”.

Charlie Beckett

“L’economia de l’atenció també ens obliga a optimitzar el temps del lector. Si té 20 minuts per a nosaltres li podem agrair l’atenció que ens dedica seleccionant per a ell les 6 coses que pot llegir en 20 minuts i que creiem que li interessaran més. Però això no vol dir que li amaguem les altres 94 que hem fet: les ha de tenir prou a l’abast. Aquest procés de selecció pot ser virtuós, no només viciós, si hi ha un bon equilibri entre previsibilitat (allò que el lector vol saber) i sorpresa (allò que no sabia que volia saber)”.

Àlex Gutiérrez - Cap de Mèdia - Diari Ara

En determinats contextos, doncs, aquesta personalització no comporta cap risc, no obstant, quan parlem de qüestions polítiques i altres temes sensibles, l'efecte de consumir contingut que reforça només el punt de vista propi i, per consegüent, exclou altres perspectives o qüestions que afecten grups minoritaris pot xocar amb la missió d'oferir una informació diversa a la ciutadania (Helberger, 2019).

A més, també pot entrar en contradicció amb la missió democràtica periodística de contribuir a la conversa comunitària a l'entorn de temes d'interès públic, de proporcionar a la ciutadania espais comuns que li facilitin desenvolupar criteris sòlids per prendre decisions polítiques informades, ja que "la política és possible perquè existeix una comunitat que es fa en la comunicació", en paraules de Carlos Ruiz (2016).

Es tracta d'un problema similar al que emergeix quan es produeix un excés d'influència de l'analítica digital per guiar la publicació d'informació en base a les preferències dels lectors. Si s'aplica sense control, fa que els mitjans acabin donant a l'audiència només allò que vol en comptes d'allò que pot ser necessari, en conseqüència, el periodisme ja no compleix amb la seva tasca comunitària, tal com s'argumenta en aquest assaig sobre l'ètica i analítica als mitjans de comunicació (Tandoc i Thomas, 2014):

For if we follow this marketing logic, then journalists will have to provide the same types of stories that get the most number of clicks, serving the majority of the audience who manifest their preferences through clicks. The minority who click on public affairs stories, for example, become marginalized and less valued because of their small number. The more journalists succumb into this click-based logic, the more the minority becomes marginalized, and the cycle continues to gnaw away at the notion of community. This is inconsistent with the pursuit of the common good. A click-based segmentation of the online audience disregards other considerations, and also falls into the trap of giving what the audience wants at the expense of what the public needs.

"Si els algoritmes de personalització acaben decidint quina informació mostren i quina no, existeix el risc de manipulació i de pèrdua de democràcia. Com es controla?"

Karma Peiró

“Si un mitjà de comunicació generalista no ofereix una mirada global i comprensiva sobre el conjunt de la realitat no tindrà molt de recorregut. Si hi hagués un mitjà que personalitzés fins a l’extrem de la individualitat, el lector perdria el sentit de comunitat i, per tant, allò seria un periodisme estrany, perquè se li hauria extirpat el vessant social. El contingut és el rei, d’acord, però si només produeixes contingut desestructurat, a mida de cada lector individual, estàs probablement més en el mercat del clics que en el del periodisme”.

Àlex Gutiérrez- Cap de Mèdia - Diari Ara

Així, com hem vist, la personalització resulta un gran avantatge per millorar l’experiència de l’usuari, que en conseqüència conferirà més valor al producte. Al mateix temps, però, pot convenir considerar certs aspectes quan es tracta d’incorporar el seu ús en el context del periodisme. Des d’una perspectiva àmplia, el mitjà pot assegurar que els lectors estan exposats a diferents punts de vista i, posant el focus en els mateixos recomanadors automatitzats, també és possible, tècnicament, assegurar que aquests motors reflecteixin una determinada visió (Helberger et al., 2018). La mateixa tecnologia proporciona moltes possibilitats i graus d’aplicació que es poden adaptar a l’estratègia de cada mitjà. Un bon exemple és aquest algoritme amb valors que ha desenvolupat Radio Sweden (Beckett, 2020a). Es tracta d’un motor que s’ha dissenyat de manera que, en recomanar continguts, inclogui criteris com la magnitud o la vida de la notícia, però també si la notícia dona visibilitat a comunitats que normalment no en tenen o si ha nascut del diàleg amb els seus oients, entre d’altres “valors de servei públic” que promou la capçalera.

Quan es portin a terme aquests dissenys, emergiran les preguntes que cada mitjà haurà de respondre en funció dels principis que hi vulgui imprimir. Un aspecte principal serà l’equilibri en els continguts: bàsicament quins s’inclouran o exclouran i fins a quin punt. Tenint en compte, a més, la tendència natural humana cap al biaix de confirmació, el mitjà s’haurà de plantejar si exclou del sistema determinats temes més sensibles com els que hem mencionat més amunt: aquells que incentiven a compartir només el punt de vista propi a les xarxes socials i contribueixen a la polarització de l’esfera pública digital (Marconi, 2020).

Per últim, i tenint en compte la nova regulació europea de la IA que ja hem tractat en apartats anteriors, convindrà valorar aquesta mena d’aplicacions de serveis de personalització des de la perspectiva del disseny i en la seva implementació per determinar fins a quin punt tenen el potencial de manipular el comportament de persones amb vulnerabilitats psicològiques o dels menors d’edat i avaluar-ne les possibles conseqüències.

“Les preguntes que cal fer-se són: quant contingut s’ha de personalitzar i quant no? on situem l’equilibri? També cal definir quin contingut no hauríem de personalitzar, perquè potser existeixen certs tipus de comportaments que són perjudicials per a les persones i/o per a la societat i potser no volem promoure’ls amb un algoritme que mostri el mateix un cop rere un altre. Potser és autodestructiu. Totes aquestes preguntes són molt valuoses. On situem el límit? Quina mena de continguts es passen dels límits?”.

Nicholas Diakopoulos

“Quan dissenyes sistemes de recomanació el que vols és encertar, i és difícil i delicat donat que els sistemes funcionen ‘sols’. Per exemple, amb els esports, perquè recomanar-li un contingut del Madrid a una persona que és culer o a un que li agrada l’Espanyol un contingut del Barça sol generar rebuig. Així que has de tenir molt clar que allò que el catàleg de continguts ha recomanat, com a mínim, no li faci mal. Amb el contingut tipus sèries, o documentals no acostuma a haver-hi problemes, perquè si el sistema ho encerta, bé, però si no l’encerta la gent no s’emprenya, però quan toques temes polítics, religiosos, esportius ... per això crec que sempre cal que les persones tinguin una explicació sobre el perquè reben una determinada recomanació. Això forma part de l’explicabilitat, perquè l’algoritme de vegades ho encerta i de vegades, no”.

Alberto Alejo - Cap Desenvolupament Software d’Enginyeria – CCMA

“Quan dissenyem un servei per a nens i nenes, i decidim que quan acabi de veure un vídeo se n’hi reproduïx un altre automàticament (tenint en compte les seves preferències), estem oferint un bon servei des del punt de vista ètic, considerant que és un menor? Fins fa un temps no ens ho plantejàvem, i ara estem començant a considerar variables ètiques quan dissenyem aquest tipus de serveis”.

Geni de Vilar - Mitjans Digitals CCMA

4.4.5.3 Tractament responsable de les dades per preservar la privacitat

Una de les característiques més significatives de la digitalització, com dèiem, són les dades que generem els usuaris mentre naveguem a internet. Avui, el nivell d'explotació de tota aquesta informació resulta un avantatge competitiu per a qualsevol organització, que podrà conèixer cada cop més i millor les necessitats dels seus públics i així identificar les millors condicions per vendre el seu producte o millorar el seu servei.

Les dades resulten imprescindibles per dur a terme qualsevol tasca automatitzable i la casuística en els àmbits d'utilització és força extensa, però sens dubte un dels principals avantatges d'aquest nou escenari per als mitjans de comunicació és conèixer les preferències de les audiències. Aquest canvi ha suposat un punt d'inflexió en determinades àrees de gestió dels mitjans, tant des de l'àmbit comercial, perquè permet adaptar serveis a necessitats i comportaments, com des de l'editorial, ja que ajuda orientar informació cap als interessos reals de cada usuari, de maneres com les que hem detallat a l'apartat anterior dedicat a la personalització.

Abans, el procés editorial funcionava en part per intuïció, però des d'un temps ençà, gràcies a l'analítica, podem saber quantes visites ha obtingut una notícia, el temps d'estada, la ràtio de reproduccions dels vídeos, des d'on es connecten els lectors, quins interessos tenen, com es comporten i moltes més dades que permeten optimitzar el producte per assolir els seus objectius en millors condicions i que varien en funció del model de negoci del mitjà: des de l'optimització de les pàgines vistes fins a l'augment de les subscripcions.

Per al periodisme, les possibilitats de conèixer les preferències dels usuaris que s'han obert amb la millora en les eines i tècniques de processament de la informació resulta una gran oportunitat. Alhora, però, el sector haurà de fer una reflexió sobre el paper que vol exercir en una internet en què l'extracció de dades personals s'ha convertit en una de les grans amenaces per a les llibertats (Guardiola, 2019; Lassalle, 2019; Peirano, 2019; Ruiz-Caballero, 2016; Zuboff, 2019). La ciutadania cada cop és més conscient de les implicacions de cedir les seves dades personals i no és casual que dues de les grans tecnològiques estiguin adaptant la seva estratègia en conseqüència. Apple ja permet bloquejar rastrejadors de màrqueting des dels seus dispositius, mentre que Google té previst eliminar en breu les cookies de tercers de Chrome, el seu navegador, i redefinir les seves tècniques de segmentació de manera que no necessiti dades

individuals dels usuaris. El *The New York Times*, per exemple, ha adaptat el seu entorn a aquest nou escenari més respectuós amb la privacitat («How The New York Times Prepared Itself for a Cookieless World», 2021).

Precisament en el sector de la comunicació trobem els exemples més coneguts del tipus de gestió més qüestionable de dades personals en companyies com Facebook, Amazon, Google o qualsevol xarxa social, que directament basen el seu model de negoci en aquesta explotació que s'arriba a efectuar amb mètodes poc transparents, amb l'ús del disseny persuasiu per augmentar el temps sota els seus dominis o l'*engagement* i una comunicació opaca sobre les seves finalitats (Harris, 2019; Patino, 2020; Peirano, 2019; Wu, 2020). D'aquesta manera, l'ecosistema digital ha esdevingut un entorn de vigilància que ha propiciat aquesta economia extractiva (Zuboff, 2019) que sovint el periodisme també denuncia.

Com dèiem, el sector dels mitjans de comunicació haurà de plantejar-se el seu paper en aquest context i actuar de manera responsable amb la informació personal dels usuaris si vol contribuir a generar confiança en uns moments en què es fa necessari identificar interlocutors amb reputació i fiabilitat (Newman, 2021).

“Si reculls dades per donar a l'usuari el que vol, o de la manera com ho vol, perfecte. El que pot ser qüestionable és crear paquets de bases de dades per vendre a tercers”.

Mariano Fernández - CDO - La Vanguardia

A més de les dades de navegació dels lectors, a les redaccions circulen més que mai dades que resulten susceptibles de formar part de les notícies del dia. Els casos més corrents els trobem en els usos que, de tota aquesta informació, en fa el periodisme de dades (l'evolució tecnològica del periodisme de precisió), el qual aconsegueix aprofitar com no s'havia fet fins ara tota la quantitat de dades en brut per transformar-les en informació de valor per a la ciutadania.

La gestió de tota aquesta informació també comporta certes implicacions des del punt de vista ètic i jurídic, que es podrien aplicar, per exemple, als mètodes d'obtenció: des dels més tradicionals, com els permisos via formularis fins a d'altres de caràcter més tècnic com el *scraping*, sense oblidar tota la gestió posterior (S. C. Lewis i Westlund, 2015) com ara l'emmagatzematge segur o els dilemes sobre la necessitat o no de la conservació de totes aquestes dades.

“Els periodistes ens creiem amb el dret d’envair la privacitat en nom de l’interès públic. Hauríem de tenir més en compte com recopil·lem, emmagatzemem i després fem servir les dades, per exemple en relació amb la publicitat”.

Charlie Beckett

“Durant temps, el mercat de premsa s’ha sustentat en què una gent paga anuncis de pàgina a preus que no estan justificats pel retorn de mercat, sinó que eren una convenció. Ara hem passat a una ultraespecialització en la que ho sabem tot: quanta gent ho mira, quines respostes genera un anunci en cada usuari. Això és una eina potent, perquè optimitza la inversió, però planteja la qüestió ètica de si es prenen decisions editorials pensant fonamentalment en la publicitat contextual que es podrà inserir a cada article”.

Àlex Gutiérrez- Cap de Mèdia - Diari Ara

El marc legal resulta una eina adient en el moment d’establir un codi de bones pràctiques en relació amb la gestió de les dades. En concret, el Reglament General de Protecció de Dades (RGPD) imposa obligacions directes. Segons aquesta normativa europea, una empresa pot tractar dades personals sota condicions de justícia, transparència, tenir una finalitat especificada i legítima i limitar-se a les dades necessàries per complir amb l’esmentada finalitat (*Reglament general de protecció de dades, 2016*).

Un dels punts centrals d’un decàleg hauria d’especificar que es recull només la informació estrictament necessària per a la finalitat plantejada, qui té accés a aquella informació i de quina manera. Per tant, caldrà evitar allò de “el que puc ho recopilo i després ja veurem per a què ho faig servir”.

David Casacuberta

A més, el tractament de les dades ha de basar-se en determinats fonaments jurídics. Amb relació a l’ús de les dades per a aplicacions d’IA, l’informe de l’Autoritat Catalana de Protecció de Dades (Peiró, 2020) destaca el principi de limitació de la finalitat, que diu que les dades s’han de recollir per a una finalitat específica i explícita i que no s’han d’utilitzar de manera incompatible. L’altre principi que posa en relleu és el de minimització, el qual indica que les dades utilitzades en un tractament han de ser adequades, pertinents i limitades a allò estrictament necessari per assolir la seva finalitat. En definitiva, el RGPD reconeix l’ús de les dades per satisfer els interessos

legítims de l'empresa, però sempre que els drets i llibertats de les persones no es vegin afectades.

“Cal un codi de conducta respecte a les dades que s’obtenen que especifiqui com es recullen, i especialment si es comparteixen amb terceres empreses. Es tracta d’una qüestió de responsabilitat. D’alguna manera no només estàs representant una empresa, sinó el model de societat en el que vols viure”.

Albert Sabater

Finalment, cal considerar que l'experiència que acumulem durant els anys d'aplicació del reglament europeu ens convida a reflexionar sobre com comuniquem les qüestions referents a les finalitats i els usos de les dades personals, els avisos de *cookies* i les clàusules de privacitat que tot sovint resulten massa extenses i difícilment comprensibles. En aquest punt torna a entrar en joc el principi de transparència, que pot determinar la claredat en el missatge si fem servir totes les eines comunicatives de què disposem quan expliquem, de la manera més sintètica i entenedora possible, quina informació fem servir, amb quina finalitat i com l'usuari pot accedir-hi i controlar-la.

La solució és que el mitjà tingui molt clar que ha de complir amb un codi ètic i expliqui a la ciutadania, en 10 passos, amb dibuixos, d’una manera molt simple, quines dades està recollint i amb quina finalitat”.

Karma Peiró

4.4.5.4 Supervisió i qualitat de les dades per evitar biaixos

Ja sabem que l'algoritme és un conjunt de regles per a resoldre un problema en un nombre finit de passos, un procediment de càlcul que consisteix a acomplir un seguit ordenat i finit d'instruccions que condueix, un cop especificades les dades, a la solució que el problema genèric en qüestió té per a les dades considerades (Enciclopèdia Catalana, 2021). Així doncs, si el resultat del processament algorítmic es basa en les dades amb què ha estat alimentat el sistema, com millors siguin les dades més acurada serà la decisió automatitzada.

En aquest context, la qualitat de les dades té una relació directa amb la finalitat que es vulgui assolir amb l'algoritme. Per exemple, un traductor automàtic basa les seves decisions en un aprenentatge previ a partir de grans corpus de dades. El procés consisteix a introduir en un sistema milions de frases en anglès, per exemple, i el mateix nombre de traduccions d'aquestes frases al català i sobre aquesta base l'algoritme extreu patrons que aprèn. En conseqüència, per tal que l'algoritme faci bones traduccions, és fonamental que les dades amb què s'ha alimentat siguin molt abundants, prou diverses i representatives del conjunt de la societat, perquè, si no, amb tota probabilitat aquest sistema oferirà traduccions errònies o esbiaixades, per exemple, en clau de gènere (Costa-Jussà, 2020): que en les traduccions atribueixi un gènere estereotipat a paraules en anglès que no en tenen, com la traducció de *nurse* a infermera (i no infermer) i *lawyer* a advocat (i no advocada).

Els biaixos, doncs, són un dels principals riscos a l'hora d'integrar eines d'IA als processos i, quan es produeixen situacions com l'esmentada, la decisió automatitzada resulta més un judici de valor fet des del prejudici que una avaluació objectiva i imparcial, un fet que col·lideix amb el valor ètic i jurídic que diferents grups han de ser tractats de manera igualitària; per tant, posa en risc el principi de justícia.

Lògicament, com més impacte social ha de tenir la decisió algorítmica, més greus poden resultar les conseqüències del biaix. En apartats anteriors hem mencionat el cas de l'algoritme racista de la policia dels EUA o l'algoritme de selecció de personal d'Amazon que descartava perfils femenins per a posicions d'enginyeria. En tots els casos, aquests problemes es produeixen com a conseqüència dels biaixos de les dades amb què han estat alimentats: en el cas del sistema sexista de reclutament de personal, per exemple, passava perquè l'algoritme havia après de les dades amb què havia estat entrenat (dades de contractació de la mateixa empresa), la qual, per a aquesta mena de tasques, havia triat homes; per consegüent, la probabilitat que una candidata

femenina fos seleccionada per l'algoritme era pràcticament nul·la. D'una manera similar, el cas del sistema racista de la policia va passar perquè l'algoritme va aprendre de les dades el fet que es detenen més ciutadans negres. I com aquests, podem trobar cada cop més casos de delegació de decisions a la IA, com l'atorgament o no de crèdits financers, el càlcul del risc en les assegurances o el diagnòstic mèdic, entre d'altres, unes decisions que tenen una important transcendència social.

Aplicats a la informació, l'impacte que els algoritmes tenen en la societat es produeix en l'esfera pública de la qual depèn la salut de les nostres democràcies. Un dels potencials de la IA són els seus efectes a escala i la capacitat d'impacte de decisions algorítmiques errònies o de continguts directament falsos, que és, com hem vist, una de les principals preocupacions socials en l'àmbit de la comunicació.

“Quan t'enfrontes a una base de dades diversa de la qual vols obtenir una sèrie de respostes cal plantejar les preguntes més encertades per fer-li. Si les preguntes són esbiaixades, les respostes també ho seran, perquè és un automatisme, qui te les genera, per tant aquí entren en joc els valors del periodista que qüestiona la base de dades. L'honestedat és fonamental per evitar biaixos”.

Ximo Blasco, coordinador de Direcció d'Informatius de TV3 – CCMA

“Considerar l'explicabilitat -tenir evidència d'on surt qualsevol tipus d'informació- és fonamental per a tots els que ens dediquem a la gestió de la informació estadística i no estadística, no només els periodistes. El que passa és que hem passat de gestionar uns volums relativament petits d'informació a uns molt grans. El periodisme ha d'incorporar el domini d'aquestes tècniques de gestió de dades massives per minimitzar riscos”.

Albert Sabater

Un contingut esbiaixat, per tant, té el potencial afectar la qualitat de la informació que es difon amb les corresponents conseqüències en la formació dels públics i el seu dret a la informació. Tal com hem explicat en l'apartat d'ètica de la IA, com a conseqüència d'aquests riscos, la UE prendrà mesures específiques.

En un mitjà de comunicació, la casuística pot ser molt àmplia. A l'apartat sobre generació automatitzada de continguts hem descrit diferents exemples clars sobre aquests potencials

biaixos i errors com a conseqüència de les dades. Un cas molt diferent, però que també té l'origen a les dades d'entrada, es pot produir en processos editorials de selecció i distribució de continguts com els motors de recomanació o els generadors de sumaris, amb els quals es pot córrer el risc d'excloure informació rellevant des del punt de vista periodístic si aquestes només mostren els continguts que més consumeix el lector i s'han optimitzat únicament per aconseguir clics.

Seguint aquest exemple, si encarreguem a un algoritme la tasca de destacar les principals notícies del dia a la portada del web, a l'app o a un butlletí, part de les dades d'entrada seran l'univers de continguts entre el qual l'algoritme haurà de triar les més adients. Per fer la tria dels principals ítems es basarà en més dades, com quines han estat les peces més visitades per altres lectors, per exemple. Amb tota la informació amb què s'ha alimentat, l'algoritme decidirà quins continguts destaca en aquesta àrea de la interfície del web.

Què passaria si aquest algoritme es dissenya perquè només tingui en compte les dades de pàgines vistes com a criteri principal? Doncs que a la selecció del millor contingut del dia els usuaris hi trobaran, amb tota probabilitat, notícies morboses, successos, safareig... Si el que preteníem amb aquesta selecció era que l'usuari es fes una idea de la informació més important del dia, no ho estarem aconseguint perquè el nostre algoritme està esbiaixat.

“És fonamental supervisar les dades amb què s'alimenten els algoritmes: quines dades, com s'obtenen, qui les filtra, quin biaix poden tenir... en tots els processos: des de l'elaboració fins a la distribució, cal supervisar qui els ha generat, qui els controla, si es venen o no i finalment, com interactuen amb ells els usuaris”.

José Alberto García Avilés

Així doncs, com millor qualitat tinguin aquestes dades i millor supervisades estiguin, més acurat serà el resultat de la decisió algorítmica. Per tant, correspon als professionals responsables dels sistemes assegurar-se, en primer lloc, que les ordres que s'han donat a l'algoritme siguin les indicades per assolir la finalitat que es busca i, en segon lloc, que les dades amb què s'ha alimentat el sistema són adients, és a dir, que aquest pot obtenir la informació necessària per executar els procediments per als objectius pels quals s'ha habilitat. En tercer lloc, cal garantir la supervisió humana necessària del resultat del procés automatitzat i, per últim, com en qualsevol aplicació de la IA, és bàsica la transparència, és a dir, que expliquem abans o que puguem explicar després per què el nostre sistema ha pres una determinada decisió.

“Hi hauria d’haver un control sobre tot el procés: des de l’origen de les dades fins que arriben a la ciutadania. Com es recullen, com es tracten... Però això avui és inviable, perquè per als mitjans no és rendible”.

Karma Peiró

4.4.5.5 La qualitat passa per posar en valor les facultats humanes

Una de les principals preocupacions que manifesten els periodistes per l'ús de la IA als mitjans de comunicació és que l'automatització afecti la qualitat del producte periodístic. Tot i que el discurs dominant es resumeix en la idea que la incorporació de la IA es traduirà en temps per tal que els professionals puguin dur a terme tasques més creatives, la majoria dels enquestats tenen percepcions diferents. A les sessions i les entrevistes que vam dur a terme amb experts i mitjans, també va emergir aquest punt en diferents ocasions.

“Tot el que es pugui automatitzar s'automatitzarà, perquè comporta un estalvi de costos. Es diu que els robots faran les tasques més mecàniques i que els periodistes seguiran tenint el control, però hem d'anar amb compte perquè aquesta és una narrativa que ens venen les empreses d'IA. Va passar fa uns anys amb la convergència: si una persona podia fer la feina de tres, es contractava aquesta. Doncs si una màquina pot fer la feina de tres en moments de crisi de les empreses de comunicació... crec que la narrativa que ens estan venent és massa rosa”.

José Alberto García Avilés

La principal inquietud que evidencia el sector sobre aquest assumpte és la percepció que hi ha un risc de minva de la qualitat del producte periodístic com a conseqüència principalment d'un nou augment de la velocitat en els processos, la qual cosa dificulta aplicar el mètode verificador, a més dels potencials efectes exposats en apartats anteriors com els sistemes algorítmics esbiaixats o el perill que el periodisme renunciï a exercir la seva funció comunitària per excessos en la personalització. L'enquesta també mostra que el sector estableix una relació directa entre automatització i desinformació, fins i tot quan és el mateix mitjà el que fa servir la tecnologia.

“La IA aporta acceleració i immediatesa. La capacitat de ser més immediat que qualsevol font, per tant accelera tots els processos i disminueix la capacitat de filtrat, d'anàlisi, de comprovació i de verificació pròpia del periodista. En conseqüència, situa el professional en una posició molt delicada. Ha d'accelerar, ha de respondre primer que ningú, però ha de basar en màquines les seves eines de comprovació. Si ha d'aplicar el seu propi criteri no té temps de verificar”.

Joan Rosés

Determinar les limitacions de cadascuna de les aplicacions dels algoritmes per fer periodisme és fonamental per prendre decisions sobre com integrar l'automatització als fluxos de treball.

Tenint en compte que la computació consisteix a processar dades, les tasques en què s'obindrà una millor resposta de la IA seran, per consegüent, aquelles que siguin quantificables. Així, determinar el grau de delegació en els algoritmes haurà de passar per valorar amb rigor el potencial grau d'èxit de cada classe de tasca automatitzada.

Un dels principals reptes a l'hora de mantenir la qualitat, doncs, és entendre quines feines delegar en la IA o, dit d'una altra manera, què és allò que els humans fan millor que les màquines: “no es tracta només de donar sentit a la computació, sinó també de dotar de sentit el periodisme tenint en compte què és el que la computació pot fer i què no (...) L'articulació del periodisme computacional depèn en gran mesura de la juxtaposició entre l'instint humà i la velocitat i automatització mecàniques”, explica Taina Bucher (2017).

“En els aspectes on serà important el factor humà són el judici, el disseny dels algoritmes, com fem servir les bases de dades i llavors com aportem valor, com creem contingut amb més criteri, més ètic, més divers i més rellevant”.

Charlie Beckett

Sense aquest coneixement, l'automatització pot comprometre pràctiques associades al periodisme de qualitat com les comentades abans o d'altres com el component humà que afecta l'originalitat: l'estil atractiu o el grau d'investigació pròpia (Alsius, 1998). “Cal dissenyar fluxos de treball intel·ligents que aprofitin l'automatització sempre que sigui possible, però també la combinació amb l'esforç humà per garantir que la sortida compleixi amb les expectatives professionals de rigor i qualitat, malgrat això pugui limitar els beneficis de l'automatització que tenen a veure amb l'escalabilitat i la velocitat” (Diakopoulos et al., 2021; Trattner et al., 2021).

“Una de les qüestions que plantejaria sens dubte per progressar en periodisme ètic no és només l'aplicació dels sistemes d'IA, sinó la seva absència. Cal plantejar-se si veritablement volem automatitzar-ho tot. Llavors, s'haurà de decidir, a les redaccions i als despatxos, en quins processos no ha d'entrar cap màquina i tenir-ho molt clar”.

Albert Sabater

El valor del producte final anirà lligat al fet que la integració de la IA es porti a terme de manera estratègica, als objectius que el mitjà es proposi assolir i a la seva adequació als estàndards de qualitat del producte informatiu. Es tracta, en definitiva, de determinar com es reinvertirà el

temps que l'automatització estalvia tenint en compte, a més, que el factor humà continua sent necessari a diferents nivells i és fonamental per mantenir la qualitat.

En aquest sentit, convé considerar també que la IA genera noves tasques de caire més tècnic (Marconi, 2020), com l'elaboració de les estructures per als textos automatitzats, el domini de la gestió de bases de dades o el filtratge de possibles temes noticiables que proporcionen les eines de recopilació d'informació, entre d'altres.

“En un món ideal, hauríem de fer servir aquest temps que s'estalviarà en fer el que els humans fem millor, que és sent creatius, aplicant criteri editorial i comprensió humana de les històries i de les persones. Totes aquestes qualitats seran cada cop més importants, perquè si les màquines poden fer periodisme bàsic, qualsevol amb una màquina podrà fer aquest periodisme”.

Charlie Beckett

“La IA és un complement, així que per aprofitar-la no es tracta d'“enxufar-la” i esperar que estalviï un munt de temps i diners. De fet el que hem identificat és que també pot generar moltes noves tasques. Per exemple, una eina de descobriment de notícies que envia 20 possibles temes per valorar una vegada a la setmana pot arribar a generar una hora de treball cada cop. Potser se'n pot treure una bona història, genial, però aquí s'ha generat més feina per a éssers humans”.

Nicholas Diakopoulos

El primer estadi de digitalització que es va iniciar a les redaccions a finals dels anys noranta i s'ha consolidat als darrers temps ja va fer recaure en l'àrea de producció multitud de noves rutines associades a la publicació digital, les quals acaben repercutint en el temps disponible per a la generació de contingut i modifiquen la tipologia de feines que s'atribueixen al periodista: la ubicació dels components al web del mitjà o tractament de les imatges, l'optimització del contingut per a la seva distribució via xarxes o cercadors, amb totes les subtasques que pot comportar, per exemple, l'optimització del contingut per al posicionament en els motors de cerca (SEO): ús de les paraules clau, tags, hipervincles, complir amb paràmetres d'extensió dels textos, posar títols alternatius, etiquetar imatges...

L'automatització, llavors, implicarà més necessitat humana per dur a terme altres funcions tècniques que s'allunyen de les rutines periodístiques més tradicionals com les connexions amb les fonts, l'enfocament, la contextualització, l'estil, la profunditat en la investigació..., tots

aquests ingredients associats al periodisme de qualitat, allò que converteix la informació en el coneixement que habilita a prendre decisions.

“Entenent que les paraules clau són importants, els tags són importants, el títol SEO és important, els links, l’enriquiment, tenir una home ben cuidada i curada... arriba un moment que tot això distreu de la meva funció principal, que és parlar amb la gent i que m’expliqui coses que el públic vol saber. Si faig aquesta mena de tasques d’empaquetament 20 minuts al dia, és tolerable, si ho haig de fer 2 hores, és menys tolerable. Caldria buscar més automatització, per a depèn quines tasques”.

Àlex Gutiérrez - Cap de Mèdia - Diari Ara

En aquest punt resulta interessant la distinció que proposa Marconi (2020) entre periodisme automatitzat i periodisme augmentat. El primer, més associat a la generació automàtica de continguts simples que aporta més volum com la creació de textos, i el segon, que implica un ús de la tecnologia orientat a singularitzar el producte final: tasques complexes que requereixen molta intensitat computacional, com per exemple l’anàlisi algorítmica de grans bases de dades relacionades amb un projecte d’investigació.

A l’hora d’introduir la IA, doncs, el mitjà haurà de ser capaç de mantenir un bon equilibri estalvi-inversió que tingui en compte com l’eficiència que pot proporcionar l’automatització repercuteix en un producte ètic i, per tant, de qualitat i amb valor, components que, com dèiem al principi, poden resultar determinants també per a l’assoliment dels objectius comercials.

L’automatització de processos provocarà encara més canvis en l’organització dels equips de treball, i en l’àmbit de presa de decisió resultarà imprescindible disposar d’una visió estratègica molt clara a l’hora d’integrar nova tecnologia i fer reorganitzacions en moments de molta incertesa en el panorama comunicatiu, que segueix travessant per una crisi sense precedents, en la qual, si bé sembla que comencen a quallar certs models de negoci per subscripció (Cerezo, 2019; Northwestern Medill News Leaders Project 2019, 2019), també és cert que perviuen molts dubtes respecte a les possibilitats de supervivència de tot l’ecosistema mediàtic tradicional, almenys tal com l’havíem conegut durant el darrer segle.

“El que diran els mitjans és “dóna’m un model de negoci. O digues a tots els lectors que paguin per llegir-me, perquè jo necessito pagar sous i despeses a finals de mes”.

Karma Peiró

A les redaccions proliferaran els perfils tècnics. Els enginyers, dissenyadors dels algoritmes, agafen un clar protagonisme en els equips i, per tant, creix la seva participació en el resultat del producte periodístic i també la seva responsabilitat (Haapanen, 2021). Els periodistes, per la seva banda, seguiran modificant les seves rutines i hauran de treballar amb els desenvolupadors per tal d'adequar el producte als estàndards ètics. Alhora que es produeix aquesta redefinició, la formació resultarà cabdal, una capacitació de caràcter tecnològic per als periodistes i de caire ètic i editorial per als enginyers.

“Hem passat de gestionar uns volums relativament petits d’informació a uns molt grans. Per evitar perdre el control de la informació és necessari que els periodistes incorporin el domini d’aquestes tècniques de gestió de dades massives”.

Albert Sabater

“És important tenir en compte que aquesta tecnologia implica un alt risc de pèrdua de control del periodista. Això està passant a totes les professions. És el moment de plantejar un debat en la professió sobre les implicacions ètiques dels processos periodístics automatitzats. També hem d’especialitzar-nos per entendre com opera aquesta tecnologia, perquè conèixer els riscos passa per tenir una formació mínima”.

Karma Peiró

Una de les principals habilitats que els periodistes hauran d’adquirir és el pensament computacional, que consisteix a “formular problemes de manera que un sistema informàtic pugui ajudar a solucionar aquests problemes”, en paraules de Diakopoulos. A més, serà important conèixer conceptes fonamentals com les diferències entre automatització simple, IA, aprenentatge automàtic, aprenentatge profund i altres, que alhora contribuiran a establir una comunicació eficaç amb els desenvolupadors.

“Al final es tracta d’entendre les limitacions d’aquestes tecnologies i les limitacions de les versions quantitatives de la realitat que poden representar aquestes eines basades en IA. Cal saber el que és possible i el que no ho és. O saber què no és possible però pot ser-ho si es combina el factor humà amb el tecnològic d’una certa manera”.

Nicholas Diakopoulos

4.4.5.6 Poder de les plataformes i independència periodística

A més d'aquests grans reptes ètics, tant a la literatura relacionada, com a les enquestes i entrevistes portades a terme en el marc d'aquest estudi, hi ha algunes altres qüestions que apareixen amb certa recurrència. Una d'elles és la preocupació per una relació desigual entre les plataformes i el periodisme. Si ens ajustem al tema que ens ocupa no correspon aquí examinar com els mitjans algorítmics han influenciat la pràctica periodística. A l'apartat anterior ja hem apuntat algunes de les noves tasques que han adoptat les redaccions per optimitzar la difusió dels continguts a través dels nous *gatekeepers*. Totes aquestes pràctiques i les seves conseqüències en la qualitat del producte informatiu han estat ja força discutides en l'àmbit acadèmic, així com des dels mateixos mitjans (Marín García, 2019), que aborden amb prou dificultats en funció de la seva estratègia de negoci.

Com dèiem, el nostre objecte d'estudi s'allunya d'aquesta qüestió, però relacionat amb això, el sí que pertoca tractar com aquesta nova fase de digitalització que implica l'adopció de la IA pot fer créixer el poder de les plataformes i per consegüent la dependència dels mitjans cap a aquestes. D'entrada, tant Google com Meta aporten avui les majors sumes destinades a inversió en innovació en l'àmbit periodístic, en molts casos relacionat amb la capacitat tecnològica fent servir eines de les pròpies plataformes. Aquest finançament serveix per exercir un *soft power* que podria arribar a condicionar la independència dels mitjans.

El finançament de projectes periodístics, en termes generals, és percebut de manera positiva entre el sector, molt necessitat d'inversió i capacitat per l'escassa implicació estructural independent. Tanmateix també hi ha veus que consideren que el paper d'aquestes plataformes podria comprometre la independència d'aquells mitjans o projectes finançats (Fanta i Dachwitz, 2020; Mols, 2020). Precisament per enfortir la independència periodística, la UE va fer públic un nou pla d'acció dirigit als mitjans que en formen part per tal de protegir la seva "autonomia estratègica", que considera amenaçada degut, entre d'altres coses, a "les grans quotes de mercat que han adquirit les plataformes no comunitàries" (European Commission, 2020).

"Em preocupa que sovint no es vegi com pot afectar el finançament de Google als mitjans europeus i a la seva independència, quan és el primer que ens hauríem de preguntar".

Joan Rosés

“El que s'està plantejant és que Google o Facebook, que s'estan beneficiant del periodisme, acabin contribuint a la salut d'aquests mitjans. La qüestió és que aquestes plataformes mai arribin a ser una font de finançament tant important com perquè això faci que la dependència sigui excessiva. Per tant, quantitat que no sigui excessiva respecte del total d'ingressos que aquell mitjà pot obtenir”.

Ximo Blasco, coordinador de Direcció d'Informatius de TV3 – CCMA

“Per a les plataformes, la quantitat de diners que inverteixen en finançar el periodisme és molt baixa. Com de veritat ajuden als mitjans és proporcionant un accés increïble a les audiències i eines excel·lents com el cercador o bases de dades enormes que poden ser molt útils. Imagina fer periodisme sense els tipus de cercadors que han desenvolupat aquestes plataformes. Penso que les infraestructures que desenvolupen són més importants per al periodisme que el finançament directe”.

Charlie Beckett

A banda del finançament als projectes d'innovació periodística, les grans tecnològiques donen suport als mitjans de comunicació a través d'infraestructura sovint impossible de desenvolupar internament. Aquestes empreses també solen proporcionar força recursos formatius, dèiem, basats en eines pròpies per a la gestió algorítmica de la informació.

Una de les principals preocupacions dels mitjans té a veure amb l'ús d'aquestes infraestructures, serveis i eines oferts per empreses que proporcionen algoritmes i programari per dur a terme aplicacions d'IA que d'una altra manera serien inaccessibles com a conseqüència dels elevats costos que representa el seu desenvolupament. Només IBM, Microsoft, Google o Amazon disposen de recursos, personal expert i volum de bases de dades suficient (*La intel·ligència artificial a Catalunya: informe tecnològic*, 2019). Un dels problemes, diuen els mitjans, està en què per fer-les servir cal transferir dades pròpies sense garanties sobre com es gestionen posteriorment.

A més, algunes d'aquestes empreses poden tenir els seus servidors fora del territori de l'Espai Econòmic Europeu. Un extrem, les transferències internacionals de dades, regulat també pel RGPD (*Reglament general de protecció de dades*, 2016). Per evitar problemes, alguns mitjans han arribat a renunciar a millorar determinats serveis.

Segons un recent estudi de Felix Simon (2022), dependre de la infraestructura, serveis i eines de les grans tecnològiques comporta certs riscos, com per exemple la impossibilitat de canviar de proveïdor sense incórrer en costos elevats. A més, les plataformes tenen un control absolut sobre els serveis d'IA, fet que incapacita els mitjans de facto en la presa de decisions. Són només elles qui poden establir els límits del que és i, sobretot, del que no és possible. També existeix el risc que, un cop implementades amb tots els costos que pot implicar, aquestes estiguin subjectes a canvis o, en el pitjor dels casos, deixin de funcionar.

En relació directa amb les qüestions ètiques en l'ús de la IA, el fet de dependre de tecnologies d'altres fa encara més difícil que els mitjans de comunicació puguin entendre el funcionament de les eines i per tant això complicaria la possibilitat de retre comptes i d'assumir responsabilitats davant possibles errors o biaixos en els *outputs* dels algorismes.

El darrer dels grans riscos per a la independència del periodisme és el control que es pot arribar a cedir a les propietàries de les tecnologies i els serveis que cada cop s'integren a més àrees de l'organització. La cessió extra de poder a les plataformes es pot produir partir del coneixement que, a través de les tecnologies propietàries, van adquirint sobre l'organització interna dels mitjans o a partir de l'accés a dades sensibles, siguin de fonts, d'altres bases de dades o de la pròpia informació empresarial.

“Per a la majoria dels mitjans la IA encara queda massa lluny com per no haver-se de recolzar en les grans plataformes que presten el servei. Aquesta externalització, però, em genera alguns dubtes: de tipus editorial, de control de dades o de desapropiament de la pròpia intel·ligència generada. Perquè si la gràcia de la IA és que acumula saber... on es deposita? A dintre de la casa, o a fora?”.

Àlex Gutiérrez - Cap de Mèdia - Diari Ara

4.4.5.7 IA per enfortir els valors del producte periodístic

Un altre dels reptes que resulta interessant considerar des del punt de vista de l'ètica sorgeix en plantejar com la IA pot ajudar a dur a terme un periodisme més alineat amb els valors que regeixen la professió. Aquesta tecnologia també proporciona eines per conferir més protagonisme als grans principis en el producte informatiu.

Un clar exemple són les aplicacions d'IA que permeten detectar desinformació. *Full Fact*, al Regne Unit, ha desenvolupat diverses eines per donar suport a la verificació humana. Una d'elles té la capacitat de rastrejar la xarxa per identificar repeticions de declaracions falses. Vist que la desinformació sovint resideix en la repetició de falsedats, diuen, aquestes són com les males herbes, que s'han d'arrencar fins que deixen de créixer. Una aplicació de la IA per a verificar consisteix en un sistema que s'aplica a declaracions que es basen en estadístiques. La IA rastreja entre dades disponibles i n'avalua la veracitat. També fan servir una altra eina que transcriu en directe discursos i categoritza les diferents afirmacions que identifica en categories com quantitats, normes, correlació, personal (algú que fa o creu alguna cosa), prediccions entre d'altres (Smalley, 2022).

A més de la verificació automatitzada, també s'experimenta ja amb algoritmes que serveixen per identificar biaixos humans en els productes informatius, com l'*AJJO Project* (Beckett, 2020b) una aplicació que analitza tot el producte informatiu i avalua la presència de gènere en diferents dimensions com el nombre de cites atribuïdes a dones versus atribuïdes a homes; la llargada de les declaracions de dones front a les d'homes o la representació en imatges segons el gènere. El *Financial Times* en va desenvolupar una aplicació similar amb la finalitat determinar el grau de presència masculina front a la femenina en un producte informatiu i així detectar el nivell de presència per gènere tant a les imatges com en les cites textuais (Waterson, 2018).

“Està l'ètica de com generem la informació i l'ètica de com la processem. Són dos treballs complementaris. En un hem d'assegurar que la nostra acció és ètica i, en l'altre, que els tercers es comporten de forma ètica. El més normal, en els codis professionals, és ocupar-se de la primera part (què haig de fer per ser ètic). Però crec que en aquest context té sentit incloure aquesta altra acció més proactiva”.

David Casacuberta

CONCLUSIONS

Aquest procés de reflexió ens ha proporcionat respostes a les preguntes de recerca que formulàvem a l'inici d'aquest treball. En l'anàlisi teòrica que hem desenvolupat al segon capítol hem identificat que, a l'hora de conceptualitzar la intel·ligència artificial, existeix el risc d'assimilar el relat que impulsa el determinisme tecnològic. Aquest que considera la tecnologia com una finalitat en sí mateixa i l'únic motor de l'evolució i que alhora li atorga una autoritat perillosa, perquè la passivitat crítica humana que implica, pot ser feta servir de manera irresponsable o interessada. El risc per al periodisme d'assumir aquestes narratives és la disminució de la capacitat crítica per entendre i explicar que la tecnologia no és neutra, i per tant ometre que es pot prendre un paper actiu necessari per decidir com fer-la servir en benefici del bé comú. En adoptar el relat determinista també pot quedar ocult el fet que, guiada per valors de productivitat, aquesta tecnologia pot col·laborar a perpetuar relacions desiguals. Tot això és important per al periodisme com a mediador cultural i constructor de relats que promoguin un ús responsable de la tecnologia. També ho és per vigilar-ne els seus usos així com per i delegar-ne determinades tasques amb criteri. Si el periodisme adopta una mirada *dataísta* s'incapacitarà per vetllar pels usos erràtics, irresponsables o interessats que poden pretendre fer servir l'excés de credulitat davant la determinació maquinal. Els sistemes de decisió —amb o sense algoritmes— són susceptibles de comportar biaixos, i la tecnologia orientada a les finalitats humanes pot resultar una aliada a l'hora de millorar els processos. Tanmateix, caldrà trobar un equilibri. Persones humanes i algoritmes poden treballar en col·laboració aprofitant els grans beneficis de la tecnologia avançada en justa ponderació amb les facultats humanes, les més importants són la veritable consciència, la capacitat per comprendre i tractar amb els problemes d'altres persones amb empatia, el sentit comú i l'autoritat moral.

En el tercer capítol de la tesi hem analitzat com la IA, en el seu paper de medidora cultural, pot erosionar els valors democràtics. Tanmateix, de nou, el problema no és la tecnologia en sí, sinó el fet de concebre les seves funcionalitats per maximitzar objectius comercials. Aquest fet provoca que s'exprimeixin el potencial de la tecnologia i el disseny d'interfícies fins l'extrem que s'arriben a explotar vulnerabilitats humanes en favor de l'optimització de les mètriques de negoci. El fet de dissenyar els algoritmes mediadors per millorar l'*engagement* també revela la no neutralitat de la tecnologia en aquest àmbit cultural, perquè la gramàtica que busca les interaccions amb el contingut implica que certs temes o relats destaquin, al mateix temps que

n'oculta d'altres. Aquesta lògica amb finalitats comercials també incentiva determinats comportaments nocius, amb els conseqüents efectes que juguen en contra de l'evolució social o la idea de progrés que defensem en aquest treball (comportaments degradants com ara la tendència a la superficialitat, a l'exhibició de la intimitat o les bombolles d'informació en entorns en què avui hi socialitzen les criatures i altres persones intel·lectualment immadures).

Els continguts que més circulen solen ser més emocionals (perquè generen més interaccions) i el seu abast tot sovint acaba condicionat pels biaixos de les persones, que tendeixen a amplificar allò que ressona amb el seu parer. Tant els individus com, per extensió, les societats, són més vulnerables perquè els biaixos també són febleses: precisament en aquesta lògica emocional i d'explotació de biaixos humans tenen el seu origen les noves formes de desinformació així com el potencial de generar cambres d'eco o bombolles informatives que poden arribar a situar les persones en ecosistemes virtuals autoreferencials. Així, aquesta manca de diversitat podria contribuir a afeblir la cohesió social i, fins i tot, a atiar la polarització entre la ciutadania.

Les plataformes digitals que estan guiades per objectius comercials erosionen les llibertats també a través de la lògica extractivista de les dades personals que han inventat, creant perfils de les persones sense que aquestes siguin del tot conscients sobre fins a quin punt la seva informació personal i el seu comportament queden exposats i es posen a disposició de companyies, governs i altra mena d'organitzacions. Les plataformes posen a disposició potents eines de segmentació basades en aquests perfils que es poden exercir amb finalitats publicitàries que pot anar des de generar record de marca, passant per persuadir fins a altra mena d'accions que pretenen manipular aprofitant febleses que es poden inferir de les dades personals.

Aquest extractivisme ha promogut dinàmiques socials com el fet que cada cop més empreses i institucions s'enfoquin en l'obtenció—sovint gens transparent—de dades personals i per a tota mena de finalitats. Siguin legítimes, com les que pot tenir un ajuntament per dur a terme polítiques públiques, com perilloses, com les que poden tenir governs autoritaris a la recerca de més control.

La posició del periodisme davant l'escenari descrit passa per tres accions principals. D'entrada serà cabdal que les professionals es formin per assolir suficients coneixements sobre aquestes tecnologies i els seus riscos des d'un punt de vista ètic, i d'aquesta manera poder interpretar i transmetre la informació amb rigor.

En segon lloc, les redaccions hauran de desenvolupar les tècniques d'auditoria algorítmica com les exposades al quart capítol per tal de poder vigilar els usos de les tecnologies d'IA en la presa de decisions, especialment aquelles que s'apliquin en àmbits com la justícia, la salut, l'educació i d'altres que poden tenir efectes importants a la vida de les persones.

Finalment, per ser coherent, la professió periodística necessitarà ser capaç de fer servir aquesta mateixa tecnologia que haurà de comunicar i vigilar, tenint en compte les consideracions ètiques pertinents. Com hem vist amb detall a la recerca empírica, per poder emprar la IA amb criteri convindrà recórrer a la tradició eticista del periodisme ja que els codis deontològics que neixen dels grans principis de la professió periodística — veritat, justícia, llibertat i responsabilitat— resulten la manera més adient per concebre una ètica general de la comunicació. Alhora, també caldrà tenir en compte els principis de la ciberètica, que s'integren en els valors periodístics de la manera que hem detallat abans. Aquests principis tecnològics poden resumir-se en els que proposa la Comissió Europea: agència humana i supervisió, robustesa tècnica i seguretat, privadesa i govern de dades, transparència, diversitat, no discriminació i equitat, benestar social i mediambiental i retiment de comptes (High-Level Expert Group on AI, 2019).

Valdrà la pena considerar així mateix les principals preocupacions en relació a l'ús de la IA a les redaccions periodístiques que, com hem exposat, tenen a veure amb els riscos de pèrdua de qualitat com a conseqüència d'un excés d'automatització i/o de manca de supervisió dels sistemes automatitzats que, a més, disposen del potencial de generar errors i biaixos. L'automatització sense més criteri que el comercial en la distribució del contingut també comporta el risc de fomentar les cambres d'eco.

També s'ha observat que els usos que impliquen més riscos ètics són les aplicacions de generació de contingut automatitzat, les eines de distribució i personalització de la informació i la gestió de les dades personals dels usuaris per evitar problemes de violació de la privacitat. Així mateix, el periodisme haurà de mantenir-se atent a l'equilibri de relacions amb les plataformes, que són les principals finançadores de projectes periodístics i propietàries de programari i solucions específiques. Això podria implicar un excés de poder de les plataformes sobre el periodisme i per tant posar en risc la seva independència o afeblir-lo d'altres maneres. Aquests problemes són evitables fent una gestió adient que passi per valorar els beneficis de la tecnologia al temps que se'n fa una apropiació regida pels valors del periodisme i es fa servir en benefici de la funció social de la professió. Els grans valors que regeixen la pràctica de periodisme de qualitat que han de dirigir els usos de les tecnologies exponencials són: veracitat, responsabilitat, llibertat i

justícia. Amb l'arribada de la IA, els principis de privacitat i transparència, ja presents al codi deontològic, es reassignifiquen i adquireixen rellevància. La privacitat, perquè avui, a les redaccions, com en els altres sectors digitalitzats, es treballa amb moltes més dades personals que cal gestionar; la transparència, perquè es tracta d'un valor principal que està present tant en la tradició del periodisme, com en la ciberètica i resulta fonamental en ambdós casos per promoure la confiança de les seves comunitats. La transparència adquireix més aplicacions pràctiques amb l'adopció de la IA, per exemple a l'hora de comunicar com un algoritme ha donat lloc a un resultat determinat o fins a quin punt està present l'automatització en el producte o en quins continguts concrets i de quina manera. També serà important a l'hora de comunicar com el mitjà fa servir les dades personals dels seus usuaris.

Gràcies a la contribució dels principals mitjans de comunicació catalans, així com d'un comitè acadèmic interdisciplinari, hem pogut elaborar, amb un grup expert en ètica periodística membre del Consell de la Informació de Catalunya, una sèrie de recomanacions que haurien de resultar una eina adient en la qual les redaccions es puguin basar per avaluar i aplicar a les múltiples casuístiques que poden sorgir de l'ús de diferents eines i aplicacions amb IA.

Aquestes recomanacions han estat fetes des de la convicció que aquesta tecnologia pot convertir-se en una potent eina editorial sempre que els seus usos estiguin guiats pel compromís amb la qualitat informativa i la missió periodística de servei públic.

1. Qualitat de les dades i gestió responsable

Vetllar per l'origen i la diversitat de les dades i mantenir una vigilància constant sobre la seva representativitat

Els algoritmes funcionen amb dades. De la mateixa manera que verifiquem fonts, cal verificar les dades: origen, diversitat... amb especial atenció a la representativitat. Cal interrogar les bases de dades: és prou representativa del grup a qui al·ludeix? És prou completa per informar del tema del qual es vol informar? Un contingut esbiaixat té el potencial d'incidir en la qualitat de la informació que es difon i afectar el dret a la informació.

Hem de tractar a tots els grups amb respecte també a través de la gestió de les dades. L'eficiència que proporciona la IA es basa en gran part en la identificació de patrons entre aquestes dades. Cal estar molt atents perquè els estereotips se solen manifestar en la detecció tecnològica de patrons. Això comporta el perill de generar biaixos en la informació. En aquest sentit cal parar

especial atenció al tractament de les minories. Sexisme, racisme, etc. no són acceptables tampoc si emergeixen de manera automatitzada.

2. Supervisió dels processos

Assegurar la qualitat tècnica dels processos de tractament de les dades per tal de minimitzar els riscos i mitigar els errors

Cal monitoritzar tots els processos, des del disseny dels sistemes d'elaboració de la informació fins als que van adreçats a la seva difusió. Aquesta supervisió hauria d'incloure la realització de tests previs al llançament de noves línies de productes informatius. I convé reiterar-la un cop que es disposa del retorn ofert pels usuaris, conegut a través de les pautes de consum però també del *feedback* que se'ls demani expressament.

Ja des del disseny dels sistemes, cal fer-se preguntes de caràcter ètic: per a quins objectius optimitzem el sistema? només per aconseguir "clics"? o hi va implícita també la voluntat de fer bon periodisme? La voluntat d'aconseguir bons resultats comercials no hauria d'excloure, sinó que al contrari hauria d'acompanyar, el manteniment d'uns criteris deontològics.

3. Transparència i retiment de comptes

Fins allà on es pugui, donar a conèixer als usuaris l'existència dels algoritmes i els trets bàsics del seu funcionament

Encara que existeixi supervisió, els sistemes automatitzats poden cometre errors i reproduir biaixos. La transparència és la protecció més eficient contra aquests potencials errors. Algunes de les principals organitzacions del sector periodístic han optat per la transparència com una via per mitigar la pèrdua de credibilitat dels mitjans i potenciar la confiança de la ciutadania.

Els algoritmes posen en qüestió aquest objectiu com a conseqüència de l'opacitat amb què funcionen en les seves decisions automatitzades. La transparència és clau per construir confiança en el context actual en el qual prolifera la desinformació a l'esfera pública originada precisament per un ús de la IA sense valors per part de les plataformes, que optimitzen la tecnologia per assolir finalitats comercials. La transparència indica el respecte de l'empresa periodística envers els seus usuaris. Els permet jutjar el valor de la informació.

D'entrada i com a principi general, les dades introduïdes i la seva procedència han de ser transparents. Pel que fa al processos algorítmics, es evident que no tot es podrà explicar, per

raons industrials o simplement per la complexitat dels mecanismes. Cal determinar doncs de quina automatització s'ha informat a l'usuari. Per exemple, potser no cal informar que s'ha fet servir IA per fer una transcripció, però sí que cal comunicar-ho quan un text s'han generat de manera automàtica, o bé que una recomanació s'ha generat a partir d'un algoritme.

En tot cas els usuaris almenys haurien de tenir clar quin és el grau d'automatització present quan consumeixen un contingut, de manera que si hi troben inexactituds puguin saber a què atribuir-les i també com poden argumentar les seves reclamacions. Aquesta transparència és doncs un aspecte fonamental de l' "accountability" o retiment de comptes, un concepte cada cop més apreciat a la vida pública en general i al món periodístic en particular.

4. Gestió responsable de les dades i privacitat

Recollir només les dades personals necessàries, anonimitzar-les si no són rellevants i preservar-les d'un mal ús per part de tercers

La transparència de les dades reclamada al punt anterior òbviament ha de ser compatible amb la necessària preservació de la privacitat, especialment en tot allò que està regulat al respecte. Cal tenir present tant les exhortacions deontològiques relatives al dret a la privacitat com també la legalitat vigent, expressada sobretot al Reglament General de Protecció de Dades adoptat per la Unió Europea per preservar la protecció de les persones físiques en allò que respecta a les seves dades personals i a la circulació d'aquestes dades.

L'automatització fa trontollar la natura i l'abast dels valors de la privacitat. Com la transparència, la privacitat agafa rellevància. D'entrada, convé recollir només la informació estrictament necessària per a la finalitat plantejada. I després s'ha d'estar en condicions d'explicar com es recopila, com s'emmagatzema, qui té accés a aquella informació i de quina manera.

Molt més encara si es comparteix amb tercers. En aquest cas, és recomanable assegurar-se que aquests tercers també en facin l'ús necessari per a la finalitat amb què ha estat recollida.

Cal fer especial atenció a la gestió responsable de les bases de dades d'institucions. En primer lloc, plantejant-nos com les obtenim i si disposem del consentiment del propietari. Tot sovint les persones faciliten dades però potser no són conscients que poden arribar a les mans d'un mitjà de comunicació. En aquest sentit, també cal prendre totes les precaucions necessàries per protegir la intimitat de les persones, com anonimitzar, emmagatzemar de manera segura les bases de dades i limitar el seu ús i conservació.

5. Personalitzacions i recomanacions

Evitar que l'ús d'algoritmes per fer informació "a mida" contribueixi a socavar el pluralisme o perjudiqui persones vulnerables

Els algoritmes s'estan fent servir per proporcionar als usuaris dos tipus de serveis que poden ser molt útils però que comporten perills evidents d'un empobriment informatiu: d'una part la personalització de la informació en funció d'uns suposats interessos específics de cada usuari; i d'una altra part, les recomanacions sobre productes o serveis que s'emmotllin a les seves preferències. En un i altre cas, existeix el risc que aquestes propostes que es faciliten puguin redundar en una manca de perspectiva general, o una pèrdua de possibilitats que quedin eclipsades per les propostes fetes per l'algoritme. És allò que ha estat batejat com les "filter bubbles", una mena d'encapsulació mental de què seria objecte el ciutadà, amb la conseqüent minva de l'autonomia personal i, pitjor encara, del pluralisme.

És important doncs que, a l'hora de dissenyar serveis de personalització o de recomanació es procuri que l'automatisme no amagui o postergui la informació de major interès públic. Si els sistemes s'optimitzen per mostrar el més vist per altres o la mena de continguts que més consumeix un usuari, existeix el perill que ocultin una altra mena de continguts de l'actualitat o de l'agenda informativa.

És recomanable tenir això en compte des del mateix disseny dels sistemes. És possible incloure variables que fomentin la inclusió de continguts amb criteri editorial o el bandejament d'una altra mena de temes massa susceptibles de generar polarització. Les eines haurien de considerar la diversitat social i per tant no excloure grups minoritaris. Seria recomanable també que s'eviti promoure filtres bombolla que puguin contribuir a la radicalització o a actes de violència, en casos extrems.

Cal parar especial esment al fet que aquests sistemes tenen el potencial de manipular el comportament, entre d'altres coses, perquè retroalimenten el conegut com a "biaix de confirmació" (reafirmació en els propis punts de vista) o per altres factors que tenen a veure amb les tècniques de disseny persuasiu. Això pot afectar més a determinades persones vulnerables des del punt de vista psicològic, així com als menors d'edat.

6. Posar en valor el factor humà

No oblidar que és el professional qui té el tarannà ètic que no té la màquina

De moment, i mentre no es demostrï el contrari, les persones tenen unes capacitats superiors a les de les màquines, si més no en tot allò que afecta les decisions amb càrrega ètica. Alhora que les tasques s'automatitzen, cal promoure i potenciar les capacitats humanes dels equips. Si s'allibera el temps que l'automatització estalvia, pot semblar una bona idea invertir aquest temps a intentar que els periodistes facin allò que millor haurien de saber fer i que aporta més qualitat: interactuar amb les fonts, observar, investigar, posar en context, donar un enfocament humà als continguts, usar un estil atractiu, aplicar la creativitat...

La tecnologia, per ella mateixa, no té criteri ètic. Només el periodista pot imprimir al producte els principis del periodisme. Per això també és important que sigui capaç de supervisar la tecnologia, i per a això necessita capacitació.

7. Formació i promoció de la interdisciplinarietat en els equips

Assolir els graus de capacitació suficient per conjuminar el coneixement de la tècnica i l'aplicació dels principis ètics

Cal formació tècnica als periodistes per no perdre el control dels processos i ser conscients dels riscos que aquests comporten. Els professionals de la informació haurien d'estar en condicions d'auditar algoritmes propis i aliens i participar en decisions que tenen a veure amb l'automatització i com posar-la al servei de la missió del periodisme.

Per a això no cal que els periodistes aprenguin a programar. Però sí que és necessari que coneguin qüestions al seu abast com el llenguatge computacional i alguns conceptes bàsics. Han de poder entendre que es poden ajustar els paràmetres d'un algoritme i veure com canvien els resultats.

Els enginyers solen estar entrenats per a aconseguir *performance* (resultats en base a mètriques comercials). És desitjable que també rebin formació en els valors que regeixen el periodisme, que també es poden transformar en mètriques estratègiques de tipus més qualitatiu.

Per dotar la IA de valors, cal formació sobre els seus riscos ètics tant a enginyers com a periodistes. En aquest sentit és recomanable promoure la interdisciplinarietat en els equips i així buscar la complementarietat de les maneres de pensar. Per exemple, utilitzant projectes reals de R+D per fer formació.

Resulta habitual contractar sistemes desenvolupats per terceres empreses. La formació també es fa imprescindible per poder gestionar proveïdors d'eines externes i valorar les opcions disponibles des del punt de vista tecnològic aplicat al periodisme.

Quan els periodistes tenen capacitat d'auditar els algoritmes poden avaluar tant els propis com els aliens. A l'hora d'informar sobre tecnologia aliena, cal poder fer-ho amb independència. Això inclou els algoritmes de les plataformes, encara que aquestes hagin finançat formació o projectes d'innovació periodística del propi mitjà.

8. R+D i proactivitat

Promoure la recerca encaminada a explorar la convergència entre la eficàcia tècnica dels sistemes i els valors d'un periodisme ètic

És recomanable que les empreses promoguin o s'acullin a projectes R+D que contemplin com pot ajudar la IA a reforçar els principis del periodisme. És molt important investigar aplicacions com eines de verificació o d'altres per detectar els propis biaixos, per exemple.

En la mesura del possible, cal evitar serveis de tercers o el recurs al núvol si el mitjà desconeix què se'n fa de les dades que traspasarà. En tot cas, si s'ha de comptar amb eines externes, en general és preferible no esperar que arribin ofertes sinó que sempre serà millor actuar de manera proactiva, buscant acords per tal de participar en la creació d'aquestes eines, de manera que es puguin adaptar a la missió i els valors del propi mitjà i l'eficàcia de les quals pugui mesurar amb els seus indicadors. Per aconseguir això pot ser bo plantejar-se la col·laboració amb altres mitjans amb els mateixos problemes i també amb investigadors especialitzats o amb universitats.

BIBLIOGRAFIA

- Aburto, T. (2014, setembre). Cómo hacer trampas en redes sociales: «Astroturfing». *El Mundo*.
<https://www.elmundo.es/tecnologia/2014/09/05/5409990222601d101a8b4581.html>
- Agudo, U., i Matute, H. (2021). The influence of algorithms on political and dating decisions. *PLOS ONE*, 16(4), e0249454.
<https://doi.org/10.1371/journal.pone.0249454>
- Alsius, S. (1998). *Ètica i periodisme* (1. ed). Pòrtic.
- Alsius, S. (2011). *Cap a una gran base de dades per a l'estudi de l'ètica periodística*.
<https://doi.org/10.2436/20.3008.02.18>
- Alsius, S., Alcalá, F., Universitat Pompeu Fabra, i Grup de Recerca en Periodisme. (2010).
The ethical values of journalists: Field research among media professionals in Catalonia. Generalitat de Catalunya : Universitat Pompeu Fabra, GRP-Journalism Research Group.
- Amoedo, A., Vara-Miguel, A., Negredo, S., Moreno, E., i Kaufman, J. (2021). *A pesar de la débil confianza en las noticias, las marcas periodísticas y locales conservan credibilidad | Digital News Report España 2021*.
<https://www.digitalnewsreport.es/2021/a-pesar-de-la-debil-confianza-en-las-noticias-las-marcas-periodisticas-y-locales-conservan-credibilidad/>
- Amón, R. (2016, novembre 17). 'Posverdad', palabra del año | Internacional. *El País*.
https://elpais.com/internacional/2016/11/16/actualidad/1479316268_308549.html

- Angwin, J. L., Jeff; Mattu, Surya; Kirchner, Lauren. (2016, maig). Machine Bias. *ProPublica*. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- Arcos, E. (2022, maig 18). *La nueva campaña a favor de la privacidad de Apple te recuerda cómo venden tus datos*. Hipertextual. <https://hipertextual.com/2022/05/campana-privacidad-apple-datos>
- Arendt, H. (2012). *La condición humana* (N. R. Gil, Trad.; 7° impresión). Espasa Libros, S.L.U.
- Arnau, J. (2021, febrer 23). Misty, el robot que hace la vida más fácil a los mayores que viven solos en Barcelona. *El Español*. https://www.elespanol.com/invertia/disruptores-innovadores/autonomias/cataluna/20210223/misty-robot-facil-mayores-viven-solos-barcelona/560944535_0.html
- Asimov, I., 1920-1992. (1989). *Yo, robot / Isaac Asimov ; [traducción de Manuel Bosch Barrett]*. Edhasa,.
- Baeza-Yates, R., i Peiró, K. (2021, març 2). La IA pot crear un món més just? *CCCB LAB*. <https://lab.cccb.org/ca/la-ia-pot-crear-un-mon-mes-just/>
- Baeza-Yates, R., i Peiró, K. (2022, gener 18). Els algoritmes poden ser més justos que les persones? *CCCB LAB*. <https://lab.cccb.org/ca/els-algoritmes-poden-ser-mes-justos-que-les-persones/>
- Barceló, À. (Director). (2020, març 26). Lo cívico y lo civilizado. En *Hoy por hoy*. Cadena Ser. https://play.cadenaser.com/audio/1585219175_343937/
- Bauman, Z. (2010). *Mundo consumo* (A. Santos Mosquera, Trad.). Paidós.
- BCN Analytics. (2018, abril). *Data & Ethics*. [Vídeo]. <https://www.youtube.com/watch?v=VII8YWWD81U>
- Beckett, C. (2019). *New powers, new responsibilities. A global survey of journalism and*

artificial intelligence. The London School of Economics and Political Science.

- Beckett, C. (2020a). An algorithm for empowering public service news. *Polis*.
<https://blogs.lse.ac.uk/polis/2020/09/28/this-swedish-radio-algorithm-gets-reporters-out-in-society/>
- Beckett, C. (2020b). Crossing boundaries together to tackle journalism biases. *Polis LSE*.
<https://blogs.lse.ac.uk/polis/2020/12/08/crossing-boundaries-together-to-tackle-journalism-biases/>
- Bellio, N. (2021). In Catalonia, the RisCanvi algorithm helps decide whether inmates are paroled—AlgorithmWatch. *Algorithm Watch*.
<https://algorithmwatch.org/en/riscanvi/>
- Berger, J. (2014). *Contagioso: Cómo conseguir que tus productos e ideas tengan éxito* (J. Paredes, Trad.). Gestión 2000.
- Berger, P. L., i Luckmann, T. (1966). *The social construction of reality; a treatise in the sociology of knowledge* (1a ed.). Doubleday.
- Bigman, Y., Gray, K., Waytz, A., Arnestad, M., i Wilson, D. (2020). *Algorithmic Discrimination Causes Less Moral Outrage than Human Discrimination* [Preprint]. PsyArXiv. <https://doi.org/10.31234/osf.io/m3nrp>
- Boden, M. A. (2017). *Inteligencia artificial* (I. Pérez Parra, Trad.). Turner.
- Bogost, I. (2015, gener 15). The Cathedral of Computation. *The Atlantic*.
<https://www.theatlantic.com/technology/archive/2015/01/the-cathedral-of-computation/384300/>
- Bostrom, N. (2016). *Superinteligencia: Caminos, peligros, estrategias*. (M. A. Fernández, Trad.). Teell Editorial.
- Bradshaw, S., i Howard, P. N. (2019). *The Global Disinformation Order 2019 Global Inventory of Organised Social Media Manipulation*].
- Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., i Bavel, J. J. V. (2017). Emotion shapes

- the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences*. <https://doi.org/10.1073/PNAS.1618923114>
- Braojos Garrido, A., García, G., de Pablo, S., Paz, M. A., i Schulze, I. (1999). *Historia del periodismo universal* (J. L. Gómez Mompert i E. Marín Otto, Ed.). Editorial Síntesis
- Bridle, J. (2020). *La nueva edad oscura: La tecnología y el fin del futuro* (M. Pérez Sánchez, Trad.). Debate.
- Bruguera, J., i Fluvià i Figueras, A. (1996). *Diccionari etimològic* (1. ed). Enciclopèdia Catalana.
- Brynjolfsson, E., i McAfee, A. (2016). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies* (First published as a Norton paperback). W. W. Norton & Company.
- Bucher, T. (2017). 'Machines don't have instincts': Articulating the computational in journalism. *New Media and Society*, 19(6), 918-933.
<https://doi.org/10.1177/1461444815624182>
- Bustamante, J. (2010). La cuarta generación de derechos humanos en las redes digitales. *Telos*, 85. <https://telos.fundaciontelefonica.com/archivo/numero085/la-cuarta-generacion-de-derechos-humanos-en-las-redes-digitales/>
- Camus, A. (2021). *El hombre rebelde* (J. Escué, Trad.; Primera edición: mayo de 2021). Debolsillo.
- Carrie Wong, J. (2019, febrer). How Facebook and YouTube help spread anti-vaxxer propaganda. *The Guardian*.
<https://www.theguardian.com/media/2019/feb/01/facebook-youtube-anti-vaccination-misinformation-social-media>
- Carvajal, M., Mondéjar, D., Valero-Pastor, J. M., De-Lara, A., García-Avilés, J. A., i Arias-Robles, F. (2022). Las innovaciones periodísticas más destacadas en España (2010-2020): Características e impacto organizacional, industrial y social.

Profesional de la información, 31(3), Article 3.

<https://doi.org/10.3145/epi.2022.may.04>

Casacuberta, D. (1998). *¿Qué es una emoción?* [[Http://purl.org/dc/dcmitype/Text](http://purl.org/dc/dcmitype/Text), Universitat Autònoma de Barcelona].

<https://dialnet.unirioja.es/servlet/tesis?codigo=143362>

Casacuberta, D. (2003). *Creación colectiva: En Internet el creador es el público* (1. ed).

Gedisa Editorial.

Casacuberta, D. (2018, maig 9). Biaix en bucle: Alimentant la injustícia algorítmica. CCCB LAB. <https://lab.cccb.org/ca/biaix-en-bucle-alimentant-la-injusticia-algoritmica/>

Casacuberta, D. (2021). *La era de Casandra: Una apología del no saber*. Edicions UAB.

Castells, M. (2003). *La galaxia Internet* (1. ed. en bolsillo). Debolsillo.

Castells, M. (2012). *Redes de indignación y esperanza: Los movimientos sociales en la era de Internet*. Alianza.

Castillo, C. (2018). *La Oferta y Disponibilidad de Contenido Audiovisual en la Era de los Datos Masivos*.

Castillo, C. del. (2021, diciembre 28). España vigilará la Inteligencia Artificial como a los fármacos o los alimentos. *ElDiario.es*.

https://www.eldiario.es/tecnologia/espana-vigilara-inteligencia-artificial-farmacos-alimentos_1_8615818.html

Castillo, C. del. (2022, juny 27). España será el campo de pruebas de las nuevas normas de Inteligencia Artificial de la UE. *ElDiario.es*.

https://www.eldiario.es/tecnologia/espana-sera-campo-pruebas-nuevas-normas-inteligencia-artificial-ue_1_9120840.html

Castillo, M. (2018, abril 11). *Zuckerberg tells Congress Facebook is not a media company: «I consider us to be a technology company»*. CNBC.

<https://www.cnbc.com/2018/04/11/mark-zuckerberg-facebook-is-a->

technology-company-not-media-company.html

Catalonia AI. L'Estratègia d'Intel·ligència Artificial de Catalunya. (2020). Generalitat de Catalunya. Departament de Polítiques Digitals i Administració Pública.
<http://politiquesdigitals.gencat.cat/ca/tic/catalonia-ai>

Cavestany, M. (2022). *Publicació | LinkedIn*. LinkedIn.
https://www.linkedin.com/posts/martacavestany_esta-ma%C3%B1ana-de-camino-al-trabajo-enbcn-activity-6950440731459121152-k2KO/?utm_source=linkedin_share&utm_medium=ios_app

Cerezo, P. (2019). *Los Medios ante los modelos de suscripción*. UOC.

Chakrabarti, S., Stengel, L., i Solanki, S. (2018). *Duty, Identity, Credibility: Fake news and the ordinary citizen in India*.

Chang, K. (2021, juliol 22). Jeff Bezos viaja al espacio para impulsar los proyectos de su compañía de cohetes. *The New York Times*.
<https://www.nytimes.com/es/2021/07/22/espanol/jeff-bezos-espacio-amazon.html>

Civio. (2019). Que se nos regule mediante código fuente o algoritmos secretos es algo que jamás debe permitirse en un Estado social, democrático y de Derecho. *Civio*. <https://civio.es/novedades/2019/07/02/que-se-nos-regule-mediante-codigo-fuente-o-algoritmos-secretos-es-algo-que-jamas-debe-permitirse-en-un-estado-social-democratico-y-de-derecho/>

Civio. (2022a). El Consejo de Transparencia ya no se opone a nuestra petición de liberar el código fuente del bono social. *Civio*.
<https://civio.es/novedades/2022/06/08/el-consejo-de-transparencia-entiende-ahora-que-la-sentencia-que-denego-el-codigo-fuente-es-contraria-a-derecho/>

Civio. (2022b). La Justicia impide la apertura del código fuente de la aplicación que concede el bono social. *Civio*. <https://civio.es/novedades/2022/02/10/la-justicia-impide-la-apertura-del-codigo-fuente-de-la-aplicacion-que-concede-el->

bono-social/

Coenen, A. (2019, octubre 17). How The New York Times is Experimenting with Recommendation Algorithms. *New York Times Open*.
<https://open.nytimes.com/how-the-new-york-times-is-experimenting-with-recommendation-algorithms-562f78624d26>

Com funciona – OEIAC. (2022). <https://oeiac.cat/com-funciona/>

Cómo funcionan los algoritmos de búsqueda. (2020, juny 20). *Google*.
https://www.google.com/intl/es_es/search/howsearchworks/algorithms/

Constine, J. (2017). Facebook Now Has 2 Billion Monthly Users... and Responsibility. *Tech Crunch*. <https://techcrunch.com/2017/06/27/facebook-2-billion-users>

Coordinadora familiars residències 5+1. (2022). Coordinadora de residències 5+1.
<https://www.coordinadoraresidencias.com/la-coordinadora-en-los-medios>

Cortina, A. (2008). *La Escuela de Fráncfort: Crítica y utopía*. Síntesis.

Costa-Jussà, M. R. (2020). “Els algoritmes també poden servir per detectar desigualtats de gènere” / *Entrevistada per Patrícia Ventura Pocino*. [Les Dades obertes i la intel·ligència artificial, eines per a la igualtat de gènere].
http://governobert.gencat.cat/web/.content/01_Que_es/04_Publicacions/col·leccio_govern_obert/GovernObert_7/docs/Govern-obert-7_D22-9-acc.pdf

Crawford, A. (2017, març 7). Facebook failed to remove sexualised images of children. *BBC News*. <https://www.bbc.com/news/technology-39187929>

Cruanyes, T. (Editor). (2022, maig 4). El robot Pepper, un més en un residència d’avis. En *Telenotícies vespre*. TV3.
<https://www.ccma.cat/tv3/alcanta/telenoticies/el-robot-pepper-un-mes-en-un-residencia-davis/video/6157190/>

Dastin, J. (2018). Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters*. <https://www.reuters.com/article/us-amazon-com-jobs->

automation-insight-idUSKCN1MK08G

De Lara, A., García-Avilés, J.-A., i Arias-Robles, F. (2022). Implantación de la Inteligencia Artificial en los medios españoles: Análisis de las percepciones de los profesionales. *Textual & Visual Media*, 15, 1-17.
<https://doi.org/10.56418/txt.15.2022.001>

De Moneo, I. (2011). Egipto desaparece del mapa de Internet | Internacional. *El País*.
https://elpais.com/internacional/2011/01/28/actualidad/1296169207_850215.html

Derakhshan, H. (2015). *The Web We Have to Save. The rich, diverse, free web that I...*
<https://medium.com/matter/the-web-we-have-to-save-2eb1fe15a426>

Deuze, M., i Yeshua, D. (2001). Online Journalists Face New Ethical Dilemmas: Lessons From The Netherlands. *Journal of Mass Media Ethics*, 16(4), 273-292.
https://doi.org/10.1207/S15327728JMME1604_03

Diakopoulos, N. (2015). Algorithmic Accountability. *Digital Journalism*, 3(3), 398-415.
<https://doi.org/10.1080/21670811.2014.976411>

Diakopoulos, N. (2019). Automating the news: How algorithms are rewriting the media. En *Automating the News: How Algorithms Are Rewriting the Media* (p. 322). Harvard University Press.

Diakopoulos, N., i Friedler, S. (2016, novembre 17). How to Hold Algorithms Accountable. *MIT Technology Review*.
<https://www.technologyreview.com/2016/11/17/155957/how-to-hold-algorithms-accountable/>

Diakopoulos, N., Trielli, D., i Lee, G. (2021). *Towards Understanding and Supporting Journalistic Practices Using Semi-Automated News Discovery Tools*. 5, 30.
<https://doi.org/10.1145/3479550>

Diccionari de l'Institut d'Estudis Catalans. (2021). *Algorisme*. Recuperat el 10 d'octubre de 2021, de <https://dlc.iec.cat/Results?EntradaText=algoritme&OperEntrada=0>

- Diéguez, A. (2016). La singularidad tecnológica y el desafío posthumano. *Pasajes*, 50, 154-164. <https://www.jstor.org/stable/10.2307/pasajes.50.154>
- Díez, M. S. (2020, agost 22). *Estoy a favor de la tecnología, pero tiene que estar vinculada a un sistema político y económico distinto para lograr justicia*. ELDiario.es. https://www.eldiario.es/tecnologia/favor-tecnologia-vinculada-sistema-politico-economico-distinto-lograr-justicia_128_6137941.html
- Dörr, K. N., i Hollnbuchner, K. (2017). Ethical Challenges of Algorithmic Journalism. *Digital Journalism*, 5(4), 404-419. <https://doi.org/10.1080/21670811.2016.1167612>
- Draft text of the recommendation on the ethics of Artificial Intelligence*. (2021). UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000377897>
- Duch, L. (2001). *Antropología de la religión*. Editorial Herder.
- Duch, L. (2018). *Sortida del laberint: Una trajectòria intel·lectual: lliçons Ferrater Mora* (Primera edició). Fragmenta Editorial.
- Duch, L., i Chillón, A. (2012). *Antropología de la comunicación*. Editorial Herder.
- Duch, L i Chillón, A. (2010, agost 1). El desahucio de las humanidades. *La Vanguardia*, 25.
- Elon Musk [@elonmusk]. (2020, gener 17). @PRANSHUAGARWA13 @Erdayastronaut Yes [Tweet]. Twitter. <https://twitter.com/elonmusk/status/1217990910052458497>
- Erickson, A. (2017, gener 24). Three men raped a woman in Sweden on Facebook Live. *The Washington Post*. <https://www.washingtonpost.com/news/worldviews/wp/2017/01/24/three-men-raped-a-woman-in-sweden-on-facebook-live/>
- Esquirol, J. M. (2018). *La penúltima bondad: Ensayo sobre la vida humana* (Primera edición). Acantilado.

- Estrategia Nacional de Inteligencia Artificial*. (2020). Gobierno de España. Ministerio de Asuntos Económicos y Transformación Digital.
https://knowledge4policy.ec.europa.eu/ai-watch/spain-ai-strategy-report_en
- European Commission. (2018, agosto 22). *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679 (wp251rev.01)*.
<https://ec.europa.eu/newsroom/article29/items/612053>
- European Commission. (2020). *Digital Decade: Commission launches Action Plan to support recovery and transformation of the media and audiovisual sectors*.
<https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52020DC0784&from=EN>
- European Commission. (2021). *Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de Inteligencia Artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de la Unión*. European Commission.
- European Council. (2022). *Digital Services Act: Council and European Parliament provisional agreement for making the internet a safer space for European citizens*. European Council. <https://www.consilium.europa.eu/en/press/press-releases/2022/04/23/digital-services-act-council-and-european-parliament-reach-deal-on-a-safer-online-space/>
- Evgeny Morozov [@evgenymorozov]. (2022, gener 18). *I'm no publicity genius but it's obvious that, in announcing it ahead of time, people pushing Web3 have screwed it up already*. 1/4 [Tweet]. Twitter.
<https://twitter.com/evgenymorozov/status/1483461767162380298>
- Fanta, A., i Dachwitz, I. (2020). *Google, the media patron. How the digital giant ensnares journalism*. [Preprint]. SocArXiv. <https://doi.org/10.31235/osf.io/3qbp9>
- Feenberg, A. (2002). *Transforming technology: A critical theory revisited*. Oxford University Press.

- Ferràs, X. (2022, gener 8). Solo somos datos. *La Vanguardia*.
<https://www.lavanguardia.com/economia/20220109/7973452/datos-inteligencia-algoritmos-creacion.html>
- Finn, E. (2018). *La búsqueda del algoritmo: Imaginación en la era de la informática* (H. Castells Albareda, Trad.). Alpha Decay.
- Fishkin, R. (2019, agost 13). Less than Half of Google Searches Now Result in a Click. *SparkToro*. <https://sparktoro.com/blog/less-than-half-of-google-searches-now-result-in-a-click/>
- Ford, M. (2016). *El auge de los robots: La tecnología y la amenaza de un futuro sin empleo* (1ª ed. en España). Paidós.
- Gadamer, H. G. (2010). *Mito y razón*. (J. F. Zúñiga, Trad.). Paidós.
- Gadamer, H. G. (2018). *Estética y hermenéutica* (A. Ramos, Trad.; Cuarta edición). Tecnos.
- Garcés, M. (2022). *Un Món comú* (O. Valls, Trad.). Tigre de paper.
- García Aller, M. (2020). *Lo imprevisible: Todo lo que la tecnología quiere y no puede controlar*. Planeta.
- García Maldonado, A. (2021, gener 1). *Feliz (e imperfecto) 2021, p.* The Objective.
<https://theobjective.com/elsubjetivo/opinion/2021-01-01/feliz-e-imperfecto-2021/>
- García-Avilés, J. A. (2021). An Inquiry into the Ethics of Innovation in Digital Journalism. En M. Luengos i S. Herrera-Damas (Ed.), *News Media Innovation Reconsidered* (1a ed., p. 3-19). Wiley. <https://doi.org/10.1002/9781119706519.ch1>
- Gerber, N., Gerber, P., i Volkamer, M. (2018). Explaining the privacy paradox: A systematic review of literature investigating privacy attitude and behavior. *Computers & Security*, 77, 226-261. <https://doi.org/10.1016/j.cose.2018.04.002>
- Gibbs, S. (2016, desembre 22). Mark Zuckerberg appears to finally admit Facebook is a

media company. *The Guardian*.

<https://www.theguardian.com/technology/2016/dec/22/mark-zuckerberg-appears-to-finally-admit-facebook-is-a-media-company>

Gilbreath, A. (2017, agost). When Is an Internet Company Evil? *Longreads*, 29-29.

Gillespie, T. (2014). The Relevance of Algorithms. En T. Gillespie, P. J. Boczkowski, i K. A. Foot (Ed.), *Media Technologies* (p. 167-194). The MIT Press.

<https://doi.org/10.7551/mitpress/9780262525374.003.0009>

Gladwell, M. (2010). Why the revolution will not be tweeted. *The New Yorker*.

<https://www.newyorker.com/magazine/2010/10/04/small-change-malcolm-gladwell>

Gomis, L. (1987). *El medio media: La función política de la prensa*. Mitre.

Gonzalo, M. (2021, novembre 24). Qué es la Web3 y por qué todo el mundo cripto habla de ella. *Newtral*. <https://www.newtral.es/que-es-web3-cripto-web-3/20211124/>

GPT-3 [intel·ligència artificial]. (2020, setembre 8). A robot wrote this entire article. Are you scared yet, human? *The Guardian*.

<https://www.theguardian.com/commentisfree/2020/sep/08/robot-wrote-this-article-gpt-3>

Griffin, A. (2016, maig 19). If you've got LinkedIn, there's a good chance your password just got stolen. *The Independent*.

<https://www.independent.co.uk/tech/linkedin-hack-details-of-more-than-100-million-users-made-available-for-sale-on-the-internet-a7037876.html>

Guaglione, S. (2022, abril 29). How publishers are experimenting with more homepage personalization sections. *Digiday*. <https://digiday.com/media/how-publishers-are-experimenting-with-more-homepage-personalization-sections/>

Guardiola, I. (2019). *L'ull i la navalla: Un assaig sobre el món com a interfície*. Arcàdia.

- Guersenzvaig, A., i Sangüesa, R. (2022, maig 17). Los ‘chicos del coro’ de la inteligencia artificial. *Agencia SINC*. <https://www.agenciasinc.es/Opinion/Los-chicos-del-coro-de-la-inteligencia-artificial>
- Guess, A., Nyhan, B., Reifler, J., i Lyons, B. (2018). *Avoiding the echo chamber about echo chambers: Why selective exposure to like-minded political news is less prevalent than you think* (p. 25). Knight Foundation. https://kf-site-production.s3.amazonaws.com/media_elements/files/000/000/133/original/To_pos_KF_White-Paper_Nyhan_V1.pdf
- Haapanen, L. (2021). *Media councils and self-regulation in the emerging era of news automation* (p. 20).
- Habermas, J. (1981). *Historia y crítica de la opinión pública* (2da. edición). Gustavo Gili.
- Hagendorff, T. (2020). *The Ethics of AI Ethics: An Evaluation of Guidelines*. 30, 99-120. <https://doi.org/10.1007/s11023-020-09517-8>
- Han, B.-C. (2020a). *La desaparición de los rituales: Una topología del presente* (A. Ciria, Trad.; Primera edición). Herder.
- Han, B.-C. (2020b, maig 15). “El dataísmo es una forma pornográfica de conocimiento que anula el pensamiento” (C. Rendueles) [El País]. https://elpais.com/cultura/2020/05/15/babelia/1589532672_574169.html
- Han, B.-C. (2021). *No cosas: Quiebras del mundo de hoy* (J. Chamorro Mielke, Trad.).
- Hao, K. (2022). Artificial intelligence is creating a new colonial world order. *MIT Technology Review*. <https://www.technologyreview.com/2022/04/19/1049592/artificial-intelligence-colonialism/>
- Hao, K., i Hernández, A. P. (2022). How the AI industry profits from catastrophe. *MIT Technology Review*. <https://www.technologyreview.com/2022/04/20/1050392/ai-industry-appen-scale-data-labels/>

Hao, K., i Swart, H. (2022). South Africa's private surveillance machine is fueling a digital apartheid. *MIT Technology Review*.

<https://www.technologyreview.com/2022/04/19/1049996/south-africa-ai-surveillance-digital-apartheid/>

Hagood, F. (1995). The Media Lab at 10. *Wired*.

<https://www.wired.com/1995/11/media/>

Harari, Y. N. (2018). *21 lecciones para el siglo XXI* (J. Ros, Trad.). Debate.

Harris, T. (2017). *Cómo un grupo de compañías tecnológicas controlan billones de mentes cada día*. [Vídeo]. Ted Conferences.

https://www.ted.com/talks/tristan_harris_how_a_handful_of_tech_companies_control_billions_of_minds_every_day?language=es#t-776827

Harris, T. (2019). Our Brains Are No Match for Our Technology. *The New York Times*.

<https://www.nytimes.com/2019/12/05/opinion/digital-technology-brain.html#click=https://t.co/KPqbKtBV4P>

Hastings, R. (2017, abril 17). *Netflix's (NFLX) CEO Reed Hastings on Q1 2017 Results*.

Seeking Alpha. <https://seekingalpha.com/article/4062901-netflixs-nflx-ceo-reed-hastings-on-q1-2017-results-earnings-call-transcript>

Hays, K. (2022, agost 18). Facebook contractors learned they lost work with the company through a video call with anonymous representatives who said an «algorithm» chose random people to cut, workers say. *Business Insider*.

<https://www.businessinsider.com/facebook-contract-workers-accenture-austin-lost-jobs-2022-8>

Heinrichs, E. (2018). We need to talk! About artificial intelligence and ethics in

journalism. *Data Driven Investor*. <https://medium.datadriveninvestor.com/we-need-to-talk-about-artificial-intelligence-and-ethics-in-journalism-e2d52e5cd45f>

Helberger, N. (2019). On the Democratic Role of News Recommenders. *Digital*

Journalism, 7(8), 993-1012. <https://doi.org/10.1080/21670811.2019.1623700>

Helberger, N., i Diakopoulos, N. (2022). The European AI Act and How It Matters for Research into AI in Media and Journalism. *Digital Journalism*, 11.

Helberger, N., Karppinen, K., i D'Acunto, L. (2018). Exposure diversity as a design principle for recommender systems. *Information, Communication & Society*, 21(2), 191-207. <https://doi.org/10.1080/1369118X.2016.1271900>

High Level Group on Media Freedom and Pluralism. (2013). *A free and pluralistic media to sustain European democracy*. European Commission.

High-Level Expert Group on AI. (2019). Directrices éticas para una IA fiable. En *European Commission*. <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>

Hill, K. (2022, june 21). Microsoft Plans to Eliminate Face Analysis Tools in Push for 'Responsible A.I.'. *The New York Times*. <https://www.nytimes.com/2022/06/21/technology/microsoft-facial-recognition.html>

Hootsuite i We Are Social. (2022). *Digital Report 2022: El informe sobre las tendencias digitales, redes sociales y mobile*. <https://wearesocial.com/es/blog/2022/01/digital-report-2022-el-informe-sobre-las-tendencias-digitales-redes-sociales-y-mobile/>

How The New York Times prepared itself for a cookieless world. (2021, setembre 9). *WAN-IFRA*. <https://wan-ifra.org/2021/09/how-the-new-york-times-prepared-itself-for-a-cookieless-world/>

Huxley, A. (2005). *Un mundo feliz* (R. Hernández, Trad.). DeBolsillo.

IAB Spain. (2022, june 2). ADEX Benchmark 2021 study. *IAB Spain*. <https://iabspain.es/estudio/adex-benchmark-2021-study/>

Iglesias, C. (Amfitrió). (2022, febrer 14-a). *Behavioral Economics, con Ujué Agudo y Karlos G. Liberal* (Núm. 50). [Episodi de Podcast]. A Realworld. Apple podcasts. <https://podcasts.apple.com/es/podcast/realworld/id1329076851>

Iglesias, C. (Amfitrió). (2018, gener 5-b). *Customer Experience, en el mundo real* (Núm. 0001). [Episodi de Podcast]. A Realworld. Apple podcasts. <https://podcasts.apple.com/es/podcast/e001-customer-experience-en-el-mundo-real/id1329076851?i=1000431688801>

Iglesias, C. (Amfitrió). (2018, abril 9-c). *Customer Journey: Transformando Experiencia y Organización* (Núm. 009). [Episodi de Podcast]. A Realworld. Apple podcasts. <https://podcasts.apple.com/es/podcast/e009-customer-journey-transformando-experiencia-y-organizaci%C3%B3n/id1329076851?i=1000431688799>

Innerarity, D. (2018, maig). La democracia como interpretación. *El País*. https://elpais.com/elpais/2018/05/08/opinion/1525790580_635517.html

Innerarity, D. (2022, juny 4). Predecir el pasado. *La Vanguardia*. <https://www.lavanguardia.com/opinion/20220604/8315715/predecir-pasado.html>

Internet Live Stats. (2022). Internet Live Stats. <https://www.internetlivestats.com/>

Introducing Home and Feeds on Facebook. (2022, juliol 21). *Meta*. <https://about.fb.com/news/2022/07/home-and-feeds-on-facebook/>

Juntunen, L. (2022). *Digital challenges to ethical standards of journalism* (p. 42). Council for Mass Media in Finland. https://presscouncils.eu/userfiles/files/JSN_Digital%20Challenges_96ppi.pdf

Just, N, i Latzer, M. (2016). *Governance by Algorithms: Reality Construction by Algorithmic Selection on the Internet*. <https://doi.org/10.1177/0163443716643157>

Juste, M. (2021, febrer 10). La pandemia dispara el uso de las redes sociales, un 27%

más que hace un año. *Expansión*. <https://www.expansion.com/economia-digital/innovacion/2021/02/10/6022c89de5fdea59448b459b.html>

Kahneman, D. (2020). *Pensar rápido, pensar despacio* (J. Chamorro Mielke, Trad.). Debate.

Kahneman, D., Rosenfield, A. M., Gandhi, L., i Blaser, T. (2016, octubre 1). Noise: How to Overcome the High, Hidden Cost of Inconsistent Decision Making. *Harvard Business Review*. <https://hbr.org/2016/10/noise>

Kant, I. (2012). *Fundamentación de la metafísica de las costumbres* (R. Rodríguez Aramayo, Trad.; Segunda edición). Alianza editorial.

Kantor, J., i Sundaram, A. (2022, agost 14). Workplace Productivity: Are You Being Tracked? *The New York Times*. <https://www.nytimes.com/interactive/2022/08/14/business/worker-productivity-tracking.html>

Katz, E. (1996). And Deliver Us from Segmentation. *The Annals of the American Academy of Political and Social Science*, 546, 22-23.

Kelly, K. (2009). The New Socialism: Global Collectivist Society Is Coming Online. *Wired*. <https://www.wired.com/2009/05/nep-newsocialism/>

Kœnig, G. (2019). *La fin de l'individu: Voyage d'un philosophe au pays de l'intelligence artificielle*. Éditions de l'Observatoire.

Kovach, B., i Rosenstiel, T. (2014). *Los elementos del periodismo: Todo lo que los periodistas deben saber y los ciudadanos esperar* (A. Diéguez, Trad.). Aguilar.

Kovach, S. (2018, setembre 1). *Google CEO Sundar Pichai declines invite to testify in Senate*. <https://www.cnn.com/2018/09/01/google-ceo-sundar-pichai-declines-invite-to-testify-in-senate.html>

Krach, S., Paulus, F. M., Bodden, M., i Kircher, T. (2010). The rewarding nature of social interactions. *Frontiers in Behavioral Neuroscience*, 4, 1-3.

<https://doi.org/10.3389/fnbeh.2010.00022>

Krawford, K. (2021, juny 6). *Microsoft's Kate Crawford: 'AI is neither artificial nor intelligent'* (Z. Corbyn) [Entrevista].

<https://www.theguardian.com/technology/2021/jun/06/microsofts-kate-crawford-ai-is-neither-artificial-nor-intelligent>

Kropotkin, P. A. (2017). *Fijaos en la naturaleza: Ética : origen y evolución de la moral* (N. Tasin, Trad.). Pepitas de Calabaza.

Kurzweil, R. (2020). *La singularidad está cerca: Cuando los humanos transcendamos la biología* (C. García Hernández, Trad.; Primera edición 2012, reimpression 2020). Lola Books.

La intel·ligència artificial a Catalunya: Informe tecnològic. (2019). Generalitat de Catalunya. Unitat d'Estratègia i Intel·ligència Competitiva d'ACCIÓ i Secretaria de Polítiques Digitals.

Las desigualdades matan. (2022, maig 25). Oxfam International.

<https://www.oxfam.org/es/informes/las-desigualdades-matan>

Lassalle, J. M. (2019). *Ciberleviatán: El colapso de la democracia liberal frente a la revolución digital* (2a ed.). Arpa.

Latour, B. (1992). *Ciencia en acción: Cómo seguir a los científicos e ingenieros a través de la sociedad* (E. Aibar, R. Méndez, i E. Ponisio, Trad.; Primera Edició). Labor.

Lecher, C. (2018, març 21). A healthcare algorithm started cutting care, and no one knew why. *The Verge*.

<https://www.theverge.com/2018/3/21/17144260/healthcare-medicaid-algorithm-arkansas-cerebral-palsy>

Leonhard, G. (2018). *Tecnología versus Humanidad: El futuro choque entre hombre y*

- máquina*. (L. Castellanos i V. Ruiz, Trad.). The Futures Agency.
- Levy, S. (1984). *Hackers: Heroes of the computer revolution* (1st ed). Anchor Press/Doubleday.
- Lewis, P. (2017, octubre 6). «Our minds can be hijacked»: The tech insiders who fear a smartphone dystopia. *The Guardian*.
<https://www.theguardian.com/technology/2017/oct/05/smartphone-addiction-silicon-valley-dystopia>
- Lewis, S. C., i Westlund, O. (2015). Big Data and Journalism: Epistemology, expertise, economics, and ethics. *Digital Journalism*, 3(3), 447-466.
<https://doi.org/10.1080/21670811.2014.976418>
- Lippmann, W. (2003). *La opinión pública* (B. Guinea Zubimendi, Trad.). Langre.
- Llaneras, K. (2022, abril 14). ¿Crees que la inteligencia artificial no es creativa? Mira cómo escribe y cómo pinta. *El País*. <https://elpais.com/tecnologia/2022-04-14/crees-que-la-inteligencia-artificial-no-es-creativa-mira-como-escribe-y-como-pinta.html>
- Locke, J. (1999). *Ensayo sobre el entendimiento humano* (E. O’Gorman, Trad.; Segunda edición en español). Fondo de Cultura Económica.
- Lomas, N. (2017, març 21). After YouTube boycott, Google pulls ads from more types of offensive content. *TechCrunch*.
<https://social.techcrunch.com/2017/03/21/after-youtube-boycott-google-pulls-ads-from-more-types-of-offensive-content/>
- Lluís Duch. (2010, desembre 22). *Lluís Duch, maestro heterodoxo* (A. Chillón) [Entrevista]. [La Vanguardia].
<https://www.lavanguardia.com/cultura/20101222/54091764879/lluis-duch-maestro-heterodoxo.html>
- López, J. M. T., i Tejedor, S. (2019). Inteligencia artificial y periodismo [presentación del monográfico]. *Doxa Comunicación. Revista Interdisciplinar de Estudios de*

- Comunicación y Ciencias Sociales*, 163-168. <https://doi.org/10.31921/doxacom>
- Luscombe, R. (2022, juny 12). Google engineer put on leave after saying AI chatbot has become sentient. *The Guardian*.
<https://www.theguardian.com/technology/2022/jun/12/google-engineer-ai-bot-sentient-blake-lemoine>
- Madrigal, A. G. (2013, juny 10). *Skinner Marketing: We're the Rats, and Facebook Likes Are the Reward—The Atlantic*.
<https://www.theatlantic.com/technology/archive/2013/06/skinner-marketing-were-the-rats-and-facebook-likes-are-the-reward/276613/>
- Mántaras, R. (2021, setembre 13). El problema es el Dr. Frankenstein, no su monstruo. *La Vanguardia*.
<https://www.lavanguardia.com/ciencia/20210913/7696391/problema-dr-frankenstein-monstruo.html>
- Marconi, F. (2020). *Newsmakers: Artificial intelligence and the future of journalism*. Columbia University Press.
- Marcuse, H. (1993). *El Hombre unidimensional: Ensayo sobre la ideología de la sociedad industrial avanzada*. (Antonio Elorza, Trad.). Planeta-Agostini.
- Marín García, B. (2019). *La tiranía del clic*. Turner.
- Mars, A., i Giménez Cano, R. (2017). Las grandes tecnológicas admiten que toneladas de propaganda rusa invadieron sus plataformas durante las elecciones. *El País*.
https://elpais.com/internacional/2017/10/31/actualidad/1509406063_824896.html
- Martín, B. (2022, abril 21). *Qué Metaverso ni qué Metaversa* [Vídeo] SEAT:CODE.
<https://www.youtube.com/watch?v=D4lZCiZnROA>
- Martínez Ahrens, J. (2018). Cambridge Analytica: La compañía que burló la intimidad de 50 millones de estadounidenses. *El País*.

https://elpais.com/internacional/2018/03/20/estados_unidos/1521574139_109464.html

Marx, K. (2014). *El capital. Crítica de la economía política: el proceso de producción del capital tomo I, libro I tomo I, libro I* (I. Perrotini i R. Campa, Trad.). Fondo de Cultura Económica.

Masip, P., Suau, J., i Ruiz-Caballero, C. (2020). Incidental Exposure to Non-Like-Minded News through Social Media: Opposing Voices in Echo-Chambers' News Feeds. *Media and Communication*, 8(4), 53-62.

<https://doi.org/10.17645/mac.v8i4.3146>

Mayer, D. (2018, juny 4). *Facebook Shared Your Data With Phone Makers Like Apple* | *Fortune*. <https://fortune.com/2018/06/04/facebook-friends-data-apple-mobile-manufacturers/>

Northwestern Medill News Leaders Project 2019. (2019). Medill Local News Initiative.

<https://localnewsinitiative.northwestern.edu/research/news-leaders/revenue-models/>

Meta (2021, octubre 28). *The Metaverse and How We'll Build It Together*. [Vídeo].

<https://www.youtube.com/watch?v=Uvufun6xer8>

Meyer, R. (2018, gener 2). The Social-Media Star and the Suicide. *The Atlantic*.

<https://www.theatlantic.com/technology/archive/2018/01/a-social-media-stars-error/549479/>

Migallón, S. (2022, juny 28). Curso exprés sobre la consciencia para ingenieros de Google. *La Máquina de Von Neumann*.

<https://vonneumannmachine.wordpress.com/2022/06/28/curso-expres-sobre-la-consciencia-para-ingenieros-de-google/>

Mike Wallace (Director). (1958). *Aldous Huxley interviewed by Mike Wallace*.

<https://www.youtube.com/watch?v=alasBxZsb40>

- Mols, B. (2020, novembre 25). La inteligencia artificial en el periodismo: Con el poder vienen también las responsabilidades. *Global Investigative Journalism Network*. <https://gijn.org/2020/11/25/la-inteligencia-artificial-en-el-periodismo-con-el-poder-vienen-tambien-las-responsabilidades/>
- Mont'Alverne, C., Badrinathan, S., Arguedas, A. R., Toff, B., Fletcher, R., i Nielsen, R. K. (2022). *The Trust Gap: How and Why News on Digital Platforms Is Viewed More Sceptically Versus News in General* (p. 73). Reuters Institute Oxford.
- Moragas Spà, M. de. (1981). *Teorías de la comunicación / Miquel de Moragas Spa*. Gustavo Gili.
- Morozov, E. (2012). *El desengaño de internet: Los mitos de la libertad en la red* (E. G. Murillo, Trad.). Destino.
- Morozov, E. (2013, març). We are abandoning all the checks and balances. *The Guardian*. <https://www.theguardian.com/technology/2013/mar/09/evgeny-morozov-technology-solutionism-interview>
- Morozov, E. (2018). *Capitalismo big tech: ¿welfare o neufeudalismo digital?* (G. Maio, Trad.). Enclave de Libros.
- Mumford, L. (1998). *Técnica y civilización* (C. Aznar de Acevedo, Trad.; Tercera edició). Alianza.
- Mumford, L. (2010). *El Mito de la máquina* (A. Rigodón, Trad.; Tercera edició). Pepitas de Calabaza.
- Muñoz, J. M. (2011). *La llama que incendió Túnez*. https://elpais.com/diario/2011/01/23/domingo/1295758353_850215.html
- Narayanan, D. (2022, juliol 6). The Dangerous Populist Science of Yuval Noah Harari. *Current Affairs, March/April 2022*. <https://www.currentaffairs.org/2022/07/the-dangerous-populist-science-of->

yuval-noah-harari

Natale, S., i Guzman, A. L. (2022). Reclaiming the human in machine cultures:

Introduction. *Media, Culture & Society*, 44(4), 627-637.

<https://doi.org/10.1177/01634437221099614>

Nechushtai, E., i Lewis, S. C. (2019). What kind of news gatekeepers do we want

machines to be? Filter bubbles, fragmentation, and the normative dimensions of algorithmic recommendations. *Computers in Human Behavior*, 90, 298-307.

<https://doi.org/10.1016/j.chb.2018.07.043>

Negredo, S., Amoedo, A., Vara-Miguel, A., Moreno, E., i Kaufmann, J. (2020). *Los*

españoles conectados se informaron por igual en medios y redes sociales sobre coronavirus y covid-19 (Digital News Report España 2020). Reuters Institute

Oxford. <https://www.digitalnewsreport.es/2020/los-espanoles-conectados-se-informaron-por-igual-en-medios-y-redes-sociales-sobre-coronavirus-y-covid-19/>

Negroponte, N. (1997). *El mundo digital: Un futuro que ya ha llegado* (M. Abdala, Trad.).

Ediciones B.

Newberry, C. (2022, febrer 28). How the Facebook Algorithm Works in 2022. *Social Media Marketing & Management Dashboard*.

<https://blog.hootsuite.com/facebook-algorithm/>

Newman, N. (2021). *Journalism, Media and Technology Trends and Predictions 2021* (p. 40).

Newman, N. (2022). *Journalism, Media, and Technology Trends and Predictions 2022* (p. 48). Reuters Institute Oxford.

<https://reutersinstitute.politics.ox.ac.uk/journalism-media-and-technology-trends-and-predictions-2022>

- Newton, C. (2017). Instagram is pushing restaurants to be kitschy, colorful, and irresistible to photographers. *The Verge*.
<https://www.theverge.com/2017/7/20/16000552/instagram-restaurant-interior-design-photo-friendly-media-noche>
- Newton, C. (2020, desembre 31). Facebook's trust problem isn't about being understood. *The Verge*.
<https://www.theverge.com/interface/2020/1/31/21115104/facebook-mark-zuckerberg-liked-understood-trust>
- Notopoulos, K. (2018, febrer 15). How I Cracked Facebook's New Algorithm And Tortured My Friends. *Buzzfeed*.
<https://www.buzzfeednews.com/article/katienotopoulos/how-i-cracked-facebooks-new-algorithm-and-tortured-my>
- Núñez, C. A. V., i Beltrán, M. (2022, juliol 5). *¿Por qué nos preocupa solo Pegasus si estamos constantemente vigilados?* The Conversation.
<http://theconversation.com/por-que-nos-preocupa-solo-pegasus-si-estamos-constantemente-vigilados-184872>
- O'Reilly, T., i Battelle, J. (2009). *Web Squared Web 2.0 Five Years On*.
<https://learning.oreilly.com/library/view/-/9781449365714/?ar>
- Oxford Word of the Year 2016*. (s.d.). Oxford Languages. Recuperat 14 juliol 2022, de <https://languages.oup.com/word-of-the-year/2016/>
- Our Strategy*. (2022, març 24). The New York Times Company.
<https://www.nytc.com/press/our-strategy/>
- Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. Penguin Press.
- Reglament (UE) 2016/679 del Parlament Europeu i del Consell, de 27 d'abril de 2016, relatiu a la protecció de les persones físiques pel que fa al tractament de dades personals i a la lliure circulació d'aquestes dades i pel qual es deroga la Directiva*

95/46/CE (Reglament general de protecció de dades), (2016) (declaració del Parlament Europeu).

Parratt-Fernández, S., Mayoral-Sánchez, J., i Mera-Fernández, M. (2021). Aplicación de la inteligencia artificial al periodismo: Análisis de la producción académica. *El Profesional de La Información*, e300317.
<https://doi.org/10.3145/epi.2021.may.17>

Patino, B. (2020). *La civilización de la memoria de pez: Pequeño tratado sobre el mercado de la atención* (A. Martorell, Trad.).

Peirano, M. (2019). *El enemigo conoce el sistema: Manipulación de ideas, personas e influencias después de la economía de la atención*. Debate.

Peirano, M. (2022). *Contra el futuro: Resistencia ciudadana frente al feudalismo climático* (Primera edició). Debate.

Peiró, K. (2020). *Intel·ligència Artificial: Decisions Automatitzades a Catalunya* (p. 142). Autoritat Catalana de Protecció de Dades.

Pérez, M. H. (2019, novembre 5). Edward Snowden: “El modelo de negocio de Google, Amazon y Facebook es el abuso”. *El País*.
https://elpais.com/retina/2019/11/05/innovacion/1572936982_317364.html

Petit, M. (2018). Per una crítica de la raó algorítmica. Estat de la qüestió. *Quaderns del CAC*, XXI(44), 11.

Piketty, T. (2014). *El capital en el siglo XXI* (E. Cazneave-Tapie, Trad.; 1ª remp). Fondo de Cultura Económica.

Piketty, T. (2019). *Capital i ideologia* (J. Boixadós i I. Estany, Trad.; Primera edició). Edicions 62.

Porlezza, C. i Ferri, G. (2022). The Missing Piece: Ethics and the Ontological Boundaries of Automated Journalism. *ISOJ Journal*, 12(1), 71-98.

- Porter, M. E. (1998). *Competitive advantage: Creating and sustaining superior performance: with a new introduction* (1st Free Press ed). Free Press.
- Postman, N. (2012). *Divertirse hasta morir: El discurso público en la era del «show business»* (E. Odell, Trad.). Ediciones de la Tempestad.
- Postman, N. (2018). *Tecnópolis: La rendición de la cultura a la tecnología* (A. Almazán i S. Miras, Trad.). El Salmón.
- Poushter, J., Fagan, M., i Gubbala, S. (2022). *Climate Change Remains Top Global Threat Across 19-Country Survey*. Pew Research Center.
<https://www.pewresearch.org/global/2022/08/31/climate-change-remains-top-global-threat-across-19-country-survey/>
- Pearl, J. (2022, juny 25). “No veo impedimento para que los robots hagan lo mismo que los humanos, incluido sentir” (M. Rius) [Entrevista].
<https://www.lavanguardia.com/vida/20220625/8343723/inteligencia-artificial.html>
- Roca, G., Velasco, L., Paniagua, E., Gardó, H., Sabater, A., Donaldson, M., i Pérez, M. (2022, juliol). Dossier Les bretxes digitals. *Barcelona Metròpolis*, 123, 36.
- Morozov. (2014). *Un paseo por el lado oscuro de la red* (I. Sádaba) [Entrevista].
<http://www.circulobellasartes.com/revistaminerva/articulo.php?id=585>
- Rosés, J. (2022, març 29). Sustituciones invisibles. *Collateral Bits*.
<https://collateralbits.net/sustituciones-invisibles/>
- Ruiz-Caballero, C. (2016). *La digitalización del otro: Los retos de la democracia en la era del ciberespacio*. Milenio.
- Rusbridger, A. (2018, agost 18). Alan Rusbridger: “Los políticos que tienen éxito son los que apelan a las emociones. Trump es un fenómeno en ese aspecto” (J. Cruz)

[Entrevista].

https://elpais.com/elpais/2019/08/12/eps/1565610609_941165.html

Rushkoff, D. (2022, setembre 4). The super-rich 'preppers' planning to save themselves from the apocalypse. *The Observer*.

<https://www.theguardian.com/news/2022/sep/04/super-rich-prepper-bunkers-apocalypse-survival-richest-rushkoff>

Sabater, A., i De Manuel, A. (2021). *Intel·ligència artificial, ètica i societat: Una mirada i discussió a través de la literatura especialitzada i d'opinions expertes*. Observatori d'Ètica en Intel·ligència Artificial de Catalunya (OEIAC).

Salazar, I. A. (2020). Los robots y la Inteligencia Artificial. Nuevos retos del periodismo. *Doxa Comunicación. Revista Interdisciplinar de Estudios de Comunicación y Ciencias Sociales*, 295-315. <https://doi.org/10.31921/doxacom>

Sapolsky, R. (2012, febrer 15). *Dopamine Jackpot! Sapolsky on the Science of Pleasure [Video]*. California Academy of Sciences. <https://www.dailymotion.com/video/xh6ceu>

Savolainen, L., i Ruckenstein, M. (2022). Dimensions of autonomy in human–algorithm relations. *New Media & Society*, 14614448221100802. <https://doi.org/10.1177/14614448221100802>

Scavino, D. (2022). *Máquinas filosóficas: Problemas de cibernética y desempleo* (Primera edició). Editorial Anagrama.

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417-457.

Seetharaman, G. W., Jeff Horwitz and Deepa. (2021, setembre 14). Facebook Knows Instagram Is Toxic for Teen Girls, Company Documents Show. *Wall Street Journal*. <https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-11631620739>

- Segreti, G. (2016, agost 29). Facebook CEO says group will not become a media company. *Reuters*. <https://www.reuters.com/article/us-facebook-zuckerberg-idUSKCN1141WN>
- Serra, C. (2022, agost 8). Desaparecer. *El País*, 12.
- Sethumadhavan, A. (2021, agost 24). Is data privacy a thing of the past in a digital world? *World Economic Forum*. <https://www.weforum.org/agenda/2021/08/is-data-privacy-thing-of-the-past-in-digital-world/>
- Shaban. (2018, desembre 20). Amazon Alexa user receives 1,700 audio recordings of a stranger through 'human error'. *Washington Post*. <https://www.washingtonpost.com/technology/2018/12/20/amazon-alexa-user-receives-audio-recordings-stranger-through-human-error/>
- Shaw, R., Cook, C., Garcia, E., Gyulnazaryan, H., Melano, J., Parusinski, J., i Sabadan, A. (2021). *The next wave of disruption: Emerging market media use of artificial intelligence and machine learning. An overview of new technology implementation in emerging market media*. International Media Support (IMS).
- Silverman, C. (2016, novembre). This Analysis Shows How Viral Fake Election News Stories Outperformed Real News On Facebook. *Buzzfeed News*. <https://www.buzzfeednews.com/article/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook>
- Simon, F. M. (2022). Uneasy Bedfellows: AI in the News, Platform Companies and the Issue of Journalistic Autonomy. *Digital Journalism*, 0(0), 1-23. <https://doi.org/10.1080/21670811.2022.2063150>
- Simon, H. A i Newell, A. (1971). Human problem solving: The state of the theory in 1970. *American Psychologist*, 26(2), 145-159. <https://doi.org/10.1037/h0030806>
- Smalley, S. (2022, juliol 14). Is the future of fact-checking automated? *Poynter*. <https://www.poynter.org/fact-checking/2022/is-the-future-of-fact-checking-automated/>

- Soldevila, A., i Escudero, S. (2020, febrer). El Barça tiene contratada una empresa que llegó a lanzar mensajes en las redes contra algunos de sus jugadores. *Cadena Ser*.
https://cadenaser.com/ser/2020/02/17/deportes/1581937137_018966.html
- Soltani, J. V.-D., Jeremy Singer-Vine and Ashkan. (2012, desembre 24). Websites Vary Prices, Deals Based on Users' Information. *Wall Street Journal*.
<http://online.wsj.com/article/SB10001424127887323777204578189391813881534.html>
- Source Code Analysis of gstumbler* (p. 23). (2010). Stroz Friedberg.
- Smith, A. (1997). *La teoría de los sentimientos morales* (C. Rodríguez Braun, Trad.). Alianza.
- Stackpole, T. (2022, maig 10). What Is Web3? *Harvard Business Review*.
<https://hbr.org/2022/05/what-is-web3>
- Steels, L., i De Mantaras, R. L. (2018). The Barcelona declaration for the proper development and usage of artificial intelligence in Europe. *AI Communications*, 31(6), 485-494. <https://doi.org/10.3233/AIC-180607>
- Stelzner, M. (Amfitrió). (2017 novembre, 3). *The Facebook Algorithm Explained for Marketers* (Núm. 274). *A Social Media Examiner*. Apple podcasts
<https://podcasts.apple.com/es/podcast/the-facebook-algorithm-explained-for-marketers/id549899114?i=1000394369093>
- Striphas, T. (2015). Algorithmic culture. *European Journal of Cultural Studies*, 18(4-5), 395-412. <https://doi.org/10.1177/1367549415577392>
- Suárez, E. (2021, març 1). Cómo construir un modelo exitoso de suscripción. *Cuadernos de Periodistas*. <https://www.cuadernosdeperiodistas.com/como-construir-un-modelo-exitoso-de-suscripcion/>
- Sunstein, C. R. (1999). The Law of Group Polarization. *John M. Olin Law & Economics Working Paper*, 91.

<http://www.law.uchicago.edu/Publications/Working/index.html>

Tandoc, E., i Thomas, R. J. (2014). The Ethics of Web Analytics: Implications of using audience metrics in news construction. *Digital Journalism*, 3(2).

<https://doi.org/10.1080/21670811.2014.909122>

Tardáguila, C., Benevenuto, F., i Ortellado, P. (2018, novembre 17). Fake News Is Poisoning Brazilian Politics. WhatsApp Can Stop It. *The New York Times*.

<https://www.nytimes.com/2018/10/17/opinion/brazil-election-fake-news-whatsapp.html?module=inline>

Taylor, A. (2014). *The People's Platform: Taking Back Power and Culture in the Digital Age* (p. 288). Metropolitan Books.

The Meme Hustler. (2014, abril 8). The Baffler. <https://thebaffler.com/salvos/the-meme-hustler>

The OASI Register – The Observatory of Algorithms with Social Impact – OASI. (2021). Eticas Foundation. <https://eticasfoundation.org/oasi/register/>

Thumler (coord.), M., Chiusi, F., Fischer, S., Kayser-Bril, N., i Spielkamp, M. (2020). *Automating Society Report-2020*.

<https://automatingsociety.algorithmwatch.org>

Thurman, N. (2019). Personalization of News. *The International Encyclopedia of Journalism Studies*, 1-6. <https://doi.org/10.1002/9781118841570.iejs0052>

Tolan, S., Miron, M., Gómez, E., i Castillo, C. (2019). Why Machine Learning May Lead to Unfairness: Evidence from Risk Assessment for Juvenile Justice in Catalonia.

Proceedings of the Seventeenth International Conference on Artificial Intelligence and Law, 83-92. <https://doi.org/10.1145/3322640.3326705>

Torras, C. (2020). "Per eliminar els biaixos de gènere dels algoritmes cal que hi hagi més dones informàtiques" / Entrevistada per Patrícia Ventura Pocino. [Les Dades obertes i la intel·ligència artificial, eines per a la igualtat de gènere].

http://governobert.gencat.cat/web/.content/01_Que_es/04_Publicacions/colle

ccio_govern_obert/GovernObert_7/docs/Govern-obert-7_D22-9-acc.pdf

Torrubia, A. (Amfitrió). (2021 juny, 26.). *Filosofía Computacional – Santiago Sánchez Migallón* (s.n.). [Episodi de Podcast]. A Software 2.0. ivoox

https://www.ivoox.com/filosofia-computacional-santiago-sanchez-migallon-audios-mp3_rf_65732088_1.html

Trattner, C., Jannach, D., Motta, E., Costera Meijer, I., Diakopoulos, N., Elahi, M., Opdahl, A. L., Tessem, B., Borch, N., Fjeld, M., Øvrelid, L., De Smedt, K., i Moe, H. (2021). Responsible media technology and AI: Challenges and research directions. *AI and Ethics*. <https://doi.org/10.1007/s43681-021-00126-4>

Trielli, D., i Diakopoulos, N. (2020). *How Journalists Can Systematically Critique Algorithms*. 5. <http://www.nickdiakopoulos.com/wp-content/uploads/2020/02/How-journalists-can-systematically-critique-algorithms.pdf>

Tsotsis, A. (2010, setembre 10). Eric Schmidt: «We Know Where You Are, We Know What You Like». *TechCrunch*. <https://social.techcrunch.com/2010/09/07/eric-schmidt-ifa/>

Tsukayama, H. (2013, maig 15). Google I/O: The end of search as we know it? *Washington Post*. https://www.washingtonpost.com/business/technology/google-io-the-end-of-search-as-we-know-it/2013/05/15/d6ea4204-bd88-11e2-97d4-a479289a31f9_story.html

Tuchman, G. (1983). *La producción de la noticia: Estudio sobre la construcción de la realidad* (H. Borrat, Trad.). Gustavo Gili.

Tufekci. (2018, agost 14). How social media took us from Tahrir Square to Donald Trump. *MIT Technology Review*, 14. <https://www.technologyreview.com/s/611806/how-social-media-took-us-from-tahrir-square-to-donald-trump/>

- Tufekci, Z. (2013). "Not This One": Social Movements, the Attention Economy, and Microcelebrity Networked Activism. *American Behavioral Scientist*, 57(7), 848-870. <https://doi.org/10.1177/0002764213479369>
- Tufekci, Z. (2015). Algorithmic harms beyond Facebook and Google: Emergent challenges of computational agency L. *Colo. Tech. LJ*, 13(203), 20.
- Tufekci, Z. (2017). *Estamos construyendo una distopía solo para que la gente cliquee en los anuncios*. [Vídeo]. TED Conferences.

https://www.ted.com/talks/zeynep_tufekci_we_re_building_a_dystopia_just_to_make_people_click_on_ads?language=es
- Tufekci, Z., i Wilson, C. (2012). Social Media and the Decision to Participate in Political Protest: Observations From Tahrir Square. *Journal of Communication*, 62(2), 363-379. <https://doi.org/10.1111/j.1460-2466.2012.01629.x>
- Tuku, K. (2020, maig 17). *Así es el ambicioso plan de Jeff Bezos para establecer colonias humanas en el espacio y redefinir los límites de la arquitectura*. Business Insider España. <https://www.businessinsider.es/ambicioso-plan-jeff-bezos-colonizar-espacio-621391>
- Turing, A. (1950). Computing Machinery and Intelligence. *Mind*, 49, 433-460.
- Turkle, S. (1985). *The second self: Computers and the human spirit* (1. Touchstone ed). Simon & Schuster.
- Twitter. (2022). Twitter: Who we are. *Twitter*. <https://about.twitter.com/en/who-we-are/our-company>
- Vaidhyanathan, S. (2012). *The Googlization of everything: And why we should worry*. University of California press.
- Vara-Miguel, A., Amoedo-Casais, A., Moreno-Moreno, E., Negrodo-Bruna, S., i Kaufmann-Argueta, J. (2022). *Digital News Report España 2022. Reconectar con*

- las audiencias de noticias*. Servicio de Publicaciones de la Universidad de Navarra. <https://doi.org/10.15581/019.2022>
- Véliz, C. (2022, juny 25). Inteligencia artificial ¿para qué? *El País Semanal*.
<https://elpais.com/eps/2022-06-26/inteligencia-artificial-para-que.html>
- Ventura Pocino, P. (2021). *Algoritmes a les redaccions: Reptes i recomanacions per dotar la intel·ligència artificial dels valors ètics del periodisme* (p. 44). Consell de la Informació de Catalunya. https://fcic.periodistes.cat/wp-content/uploads/2022/02/algorismes_a_les_redaccions_CAT_.pdf
- Verbeek, P.-P. (2011). *Moralizing technology: Understanding and designing the morality of things*. The University of Chicago Press.
- Vidal Castell, D. (2020). Las Fake-news como Recurso de Legitimación de los Medios Convencionales. Análisis Discursivo de Artículos Publicados en los Principales Diarios de Barcelona (2017-2019). VV.AA. *Actas del VII Congreso Internacional de la AE-IC. Comunicación y Diversidad*, 2591-2615.
- Vidal Castell, D. Garde, C., i Ventura Pocino, P. (2022). Alteritmo y 'periodismo-otro' en la cultura del algoritmo. *Contribuciones desde la filosofía y el comparatismo literario ante la hegemonía de los lenguajes sintéticos*. [Manuscrit enviat per a publicació]. Departament de Mitjans, Comunicació i Cultura. Universitat Autònoma de Barcelona.
- Vidal i Castell, D. (2005). *El malson de Chandos: Aproximació a la crisi acadèmica i professional del periodisme des de la crisi postmoderna de la paraula*. Publicacions de la Universitat Jaume I.
- Vigdor, N. (2019, novembre 10). Apple Card Investigated After Gender Discrimination Complaints. *The New York Times*.
<https://www.nytimes.com/2019/11/10/business/Apple-credit-card-investigation.html>

- Villalvilla, E. L. (2022, agost 19). Comprar casa se complica: Cinco años más de ahorro para pagar la entrada de una vivienda. *El Mundo*.
<https://www.elmundo.es/economia/2022/08/19/62ffab0c21efa0ea4e8b45c5.html>
- Wakabayashi, D., i Maheshwari, S. (2019, febrer 20). Advertisers Boycott YouTube After Pedophiles Swarm Comments on Videos of Children. *The New York Times*.
<https://www.nytimes.com/2019/02/20/technology/youtube-pedophiles.html>
- Walsh, D. (2022). *Timnit Gebru: Ethical AI Requires Institutional and Structural Change*. Stanford HAI. <https://hai.stanford.edu/news/timnit-gebru-ethical-ai-requires-institutional-and-structural-change>
- Wardle, C., i Derakhshan, H. (2017). *Information Disorder: Toward an interdisciplinary framework for research and policy making*. Council of Europe.
<https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c>
- Warren, S. D., i Brandeis, L. D. (1890). Warren-brandeis.pdf. *Harvard Law Review*, 4(5), 193-220.
- Waterson, J. (2018, novembre 14). Financial Times tool warns if articles quote too many men. *The Guardian*.
<https://www.theguardian.com/media/2018/nov/14/financial-times-tool-warns-if-articles-quote-too-many-men>
- Weedon, J., Nuland, W., i Stamos, A. (2017). *Information Operations and Facebook* (p. 13). Facebook.
- Weil, S. (1996). *Echar raíces* (J. C. González Pont, Trad.). Trotta.
- Wiener, A. (2021). *Valle inquietante* (J. Calvo, Trad.). Libros del Asteroide.

Wiener, N. (1954). *The human use of human beings: Cybernetics and society*. Houghton Mifflin. <https://lccn.loc.gov/54003471>

Winner, L. (1980). *Do Artifacts Have Politics?* 121-136.

Wolfson, S. (2018, maig 24). Amazon's Alexa recorded private conversation and sent it to random contact. *The Guardian*.
<https://www.theguardian.com/technology/2018/may/24/amazon-alexa-recorded-conversation>

Wu, T. (2020). *Comerciantes de atención: La lucha épica por entrar en nuestra cabeza* (P. Zumalacárregui Martínez, Trad.). Capitán Swing.

Yan, H. (2017, gener 6). *Chicago torture video: 4 charged with hate crimes, kidnapping—* CNN. <https://edition.cnn.com/2017/01/05/us/chicago-facebook-live-beating/index.html>

Yun Chee, F. (2020). Facebook, Twitter, Google to report monthly on fake news fight, EU says. *Reuters*. <https://www.reuters.com/article/us-health-health-coronavirus-eu-disinfor/facebook-twitter-google-to-report-monthly-on-fake-news-fight-eu-says-idUSKBN23H1EH>

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power* (First edition). Public Affairs.

