



Co-funded by
the European Union

A result of the Erasmus+ project HEDY
KA220-HED-000029536 – Cooperation
partnership in the higher education.



BOOKLET

Un assaig sobre la vida a l'era de la IA



BAEHF

BOOKLET (LLIBRET) - Un assaig sobre la vida a l'era de la IA

Autors: Davide Careglio, Ana I. Alves Moreira, Cecilio Angulo Bahón, Federica Casaccio, Rozalina Dimova, Tihomir Dovramadjiev, Antonia Jakobi, Csaba Kollar, Ievgeniia Sukhovii, Gyula Szabó
2022

DOI: 10.5281/zenodo.7426885

<https://lifeintheaiera.eu/>

Finançat per la Unió Europea. Tot i això, els punts de vista i les opinions expressades pertanyen únicament als autors i no reflecteixen necessàriament les de la Unió Europea o l'Agència Executiva Europea d'Educació i Cultura (EACEA). Ni la Unió Europea ni l'EACEA no se'n poden fer responsables.

A result of the Erasmus+ project HEDY
KA220-HED-000029536 – Cooperation
partnership in the higher education.



**Co-funded by
the European Union**



El projecte HEDY

HEDY - La vida a l'era de la Intel·ligència Artificial (IA) és un projecte Erasmus+ de 2 anys que va començar el novembre de 2021. En el seu propi títol, ret homenatge a Hedy Lamarr, una actriu i inventora austríaca (1914-2000), cocreadora de la tecnologia de comunicacions sense fils, adoptada pel control de torpedes durant la Segona Guerra Mundial i usada actualment en xarxes mòbils, dispositius Bluetooth i Wi-Fi. El projecte HEDY destaca per ser una font d'informació gratuïta i accessible sobre les tecnologies digitals de la quarta revolució industrial (Indústria 4.0), en particular la IA, n'assenyala les possibles aplicacions futures, alhora que en destaca els possibles impactes socials. L'objectiu de HEDY és oferir una visió integral i compartida de com la IA està afectant les nostres vides i remodelant els nostres entorns socioeconòmics, culturals i humans mitjançant la promoció de la reflexió crítica, l'autoaprenentatge i el debat sobre aquests temes. El públic objectiu principal (però no exclusiu) d'aquest projecte és el que té relació amb l'educació superior.

Resum

El Booklet és un assaig que defineix la posició de HEDY sobre la vida a l'era de la IA i la justificació d'aquesta posició. En aquest treball, descrivim i presentem els resultats del nostre enfocament per construir la nostra argumentació sobre els desafiaments, les oportunitats i l'impacte esperat de la IA en quatre àrees: 1) negocis, 2) govern, 3) habilitats i competències, i 4) persones i estil de vida. Aquest enfocament és doble i consisteix a recopilar informació de dues fonts: i) Estat de l'art bibliogràfic, i ii) Interacció amb persones. La primera font permet recopilar l'estat actual del coneixement sobre l'impacte social de la IA. La segona font és un enfocament mixt que consta de qüestionaris i grups focals realitzats a cinc països europeus diferents amb experts i no experts en IA. Aquestes dues fonts brinden una contribució única al panorama de la IA en combinar investigacions d'avantguarda amb opinions de diferents actors socials i preguntes, inquietuds i idees debatudes.

Octubre 2022

Contingut

01	Introducció Quin és l'objectiu d'aquest Booklet i com es construeix?	5
02	Indústria 4.0/5.0 On som ara?	7
03	Breu història de la Intel·ligència Artificial És la IA un concepte nou?	10
04	Intel·ligència Artificial Quins són els impactes de la IA a la nostra societat?	12
05	Negocis La IA està millorant els ingressos o les qualitats laborals?	15
06	Governança la IA hauria d'estar lliure de regles?	18
07	Habilitats i competències La IA està afectant el sistema educatiu?	22
08	Persones i estil de vida La IA pot contribuir a millorar la vida de les persones?	26
09	Conclusions Què ens diu aquest estudi?	30
0A	Referències Vols saber les nostres fonts?	33
0B	Apèndix Vols saber-ne més?	39

1. Introducció

La tecnologia digital està envaint el nostre món, i s'utilitza en totes les dimensions de la vida, des de l'educació fins al treball, la salut o la governança. El desenvolupament de coneixements i habilitats és ara un procés que dura tota la vida i exigeix una alfabetització digital cada vegada més gran. No obstant això, per a alguns membres de la societat, com els 'nadius digitals', l'ús d'aquesta tecnologia és natural, en canvi no ho és necessàriament, per a les persones no digitalment natives (les anomenades immigrants digitals), . Com assegurarem que cada ciutadà desenvolupi les habilitats necessàries per mantenir-se inclòs en una societat cada cop més digital? I com aconseguir l'equitat en comptes d'amplificar les desigualtats? Assumint que la IA transformarà el mercat laboral, és rellevant imaginar el sistema educatiu en un món on el treball no és un factor central a la vida o on els treballs, com els coneixíem, ja no existeixen. Quin seria el paper de l'educació? Com podríem reorganitzar-la? Quins serien els seus objectius i quines necessitats atendria? I res millor que proposar l'ús de les tecnologies per conscienciar sobre la vida a l'era digital i desenvolupar habilitats per gaudir dels seus beneficis, però també afrontar els reptes que ofereix aquesta nova era.

Aquest és el leitmotiv de HEDY – La vida a l'era de la IA [1.1]. El projecte HEDY destaca per ser una font d'informació gratuïta i accessible sobre les tecnologies digitals de la quarta revolució industrial (Indústria 4.0), és a dir, la IA, en assenyalar-ne les possibles aplicacions futures, alhora que aclareix els possibles impactes. Per assolir aquesta fita es contemplen quatre resultats específics:



Un **Booklet** – Un assaig que defineix la posició de HEDY sobre la vida a l'era de la IA i la justificació d'aquesta posició. Al Llibret s'hi revisen aspectes com ara les característiques de la IA, els seus impactes positius, els riscos associats amb certs usos de la IA i s'hi identifiquen els desafiaments, les oportunitats i les repercussions esperades amb exemples paradigmàtics, amb l'ànim d'estimular la reflexió i el debat sobre temes de la societat del coneixement.



Un **Toolkit** –Una col·lecció d'eines audiovisuals que mostren com la IA podria abordar alguns dels problemes socials més desafiadors, així com els riscos que es mitigaran si la IA assoleix tot el potencial. El Toolkit està pensat per fer-se servir de manera flexible, i donar suport a activitats d'aprenentatge i debats. Inclourà pel·lícules, documentals de ficció i de fets reals, així com, ponències d'experts acuradament seleccionades, 5 dels quals seran produïdes pel mateix projecte.



Un **Massive Open Online Course (MOOC)** – Un curs per promoure un coneixement extens, la reflexió crítica i el debat sobre la IA, i el seu impacte clau a la societat. Aquest MOOC serà d'accés lliure i gratuït per als participants, amb fòrums interactius i estarà estructurat en 5 mòduls preparats pels experts. Els participants en el MOOC podran compartir idees i involucrar-se profundament en temes d'IA mitjançant una àmplia varietat d'activitats en línia.



Un **Guideline** – Una documentació concisa i fàcil de llegir que explicarà la naturalesa dels recursos d'aprenentatge preparats per aquest projecte HEDY, el millor ús dels actius produïts i un glossari amb termes clau sobre temes relacionats amb la IA. Aquesta guia ajudarà a crear una base sòlida per garantir l'aprofitament dels resultats de HEDY per part duna comunitat més àmplia durant i després del projecte.

En aquest document presentem el Booklet. Pretén ser una contribució al desafiament llançat per la Comissió Europea per construir un enfocament europeu sòlid de la IA, basat en l'estratègia del 2018 [1.2] i reforçat pel Llibre Blanc de la IA del 2020 [1.3]. També contribueix al marc ètic de la UNESCO [1.4], que proporciona, en la seva resolució, la base per fer que els sistemes d'IA funcionin pel bé de la humanitat, les societats, el medi ambient i els ecosistemes, i per prevenir possibles danys.

Com ja s'ha dit anteriorment, el seu objectiu és revisar les característiques de la IA, identificar desafiaments, oportunitats, riscos associats amb certs usos de la IA i repercussions esperades en quatre àmbits diferents de la nostra societat:

- **Negocis:** Amb una quantitat cada vegada més gran de dades disponibles i les preferències i la complexitat dels clients en constant evolució, les empreses ja no poden confiar en els mètodes comercials tradicionals per impulsar el seu creixement. Els canvis radicals que proporciona l'ús de la IA han obert un nou àmbit de possibilitats per fer-ho, a través d'informació processable obtinguda a partir de les dades dels clients.
- **Governança:** La paraula governança s'ha integrat al vocabulari comercial estàndard i inclou tots els processos de govern. És la manera com s'estructuren, mantenen i regulen les regles o accions i, sovint, com s'assigna la responsabilitat. El govern de la IA ha d'assegurar-se que s'estableixin límits dins de la tecnologia, és a dir, que l'ús de la IA es faci siguin criteris ètics.
- **Habilitats i competències:** És una certesa que la IA està començant a reemplaçar un nombre cada cop més gran de treballadors: el traspàs de tasques, fins ara fetes per mà d'obra humana, a les màquines està alterant les habilitats i competències necessàries en un mercat laboral global competitiu. En resposta als canvis inevitables en la força laboral, els sistemes educatius semblen necessitar un canvi per reflectir les habilitats i les competències que seran valuoses en una era dominada per la IA.
- **Persones i estil de vida:** La IA ha arribat per quedar-se a la nostra vida diària. La IA és una tecnologia veritablement transformadora amb efectes de gran abast, des de facilitar la nostra vida quotidiana amb recomanacions de cerca en línia i assistents de veu i inicis de sessió de reconeixement facial, fins a l'impuls d'enormes avenços en el tractament de la salut, la identificació de pandèmies i l'alleujament de la fam al món.

A més de descriure les aplicacions actuals i els impactes esperats de la IA, per a cada àmbit també identifiquem els temes més preocupants sobre la IA. Tot i això, el Llibret no pretén proporcionar solucions generals a aquestes preocupacions; per contra, la idea és oferir una manera atractiva d'estimular la reflexió i el debat sobre temes de la societat del coneixement, discutir els problemes ètics d'aquestes tecnologies digitals emergents i proporcionar exemples paradigmàtics.

Per assolir aquest objectiu i construir els nostres fonaments recopilem informació de dues fonts diferents: i) Revisió de l'estat de l'art, i ii) Interacció amb persones. La primera font ha consistit en recopilar l'estat actual del coneixement sobre les aplicacions i els impactes de la IA, mentre que la segona font ha consistit en recopilar informació mitjançant entrevistes a persones (tant expertes com no expertes en IA) a través de grups focals realitzades en cinc països europeus diferents. Aquestes dues fonts brinden una contribució única al panorama de la IA en combinar investigacions d'avantguarda amb opinions de primera mà i preguntes, inquietuds i idees debatudes entre les persones dels grups focals esmentats.

La resta del document està organitzat de la manera següent: L'apartat 2 descriu la quarta revolució industrial, és a dir, la digitalització i automatització de la fabricació mitjançant un canvi fonamental en la manera com es fabriquen els productes, i està profundament lligada a la gran acceptació de la IA en aquest àmbit. L'apartat 3 presenta una breu història de l'evolució de la IA. L'apartat 4 està dedicat a la definició d'IA i la identificació dels principals reptes i oportunitats en termes generals d'aquesta tecnologia. Els quatre apartats següents se centren, respectivament, en els quatre àmbits esmentats anteriorment: La IA en els negocis; en la governança; en les habilitats i competències; en les persones i estil de vida. Per acabar, l'apartat 9 resumeix les conclusions més importants del nostre treball i proporciona algunes recomanacions finals sobre com adaptar-se millor a l'era de la IA.

2. Indústria 4.0/5.0

És instructiu recordar que aquesta no és la primera vegada que la societat passa per una revolució industrial, sinó la quarta. Abans del segle XIX, un gran percentatge de la mà d'obra es concentrava al sector agrícola i l'ús de màquines mecàniques era molt limitat. Al llarg dels darrers dos segles, la humanitat ha viscut quatre revolucions industrials que han canviat radicalment la nostra vida i la nostra societat.



Existeixen tres raons per les quals la revolució actual (la indústria 4.0/5.0) no és només una continuació de la Revolució Digital, sinó el començament d'una de nova: l'impacte, la velocitat i l'abast. La taxa actual d'avenços no té precedents a la història. Els desenvolupaments tecnològics emergents en disciplines que inclouen IA, IoT, automòbils autònoms, robots, computació i comunicació quàntica, nanotecnologia, etc. fan que la indústria 4.0/5.0 evolucioni exponencialment en lloc de linealment com va passar amb les revolucions industrials precedents. A més, està causant canvis radicals en gairebé tots els negocis a tots els països, i l'amplitud i complexitat d'aquests desenvolupaments necessiten una revisió completa dels sistemes de producció, gestió i govern.

La Indústria 4.0/5.0, com les revolucions anteriors, té el potencial d'impulsar els nivells d'ingressos globals i millorar la qualitat de vida de les persones a tot el món [2.2]. Fins ara, els que poden pagar i accedir al món digital són els que més se n'han beneficiat. La tecnologia ha permès nous productes i serveis que milloren l'eficiència i el gaudi de la nostra vida diària. En el futur, s'espera que els avenços tecnològics donaran com a resultat una revolució del costat de l'oferta, amb beneficis a llarg termini en eficiència i producció. Els costos de transport i comunicació disminuiran, la logística i les cadenes de subministrament globals seran més eficients i els costos comercials disminuiran, obrint nous mercats i impulsant el desenvolupament econòmic.

Tot i això, no només s'esperen impactes positius. La revolució actual va acompanyada del perill de destruir llocs de treball. A mesura que l'automatització reemplaça la feina a tota l'economia, el desplaçament net d'empleats per màquines pot augmentar la bretxa entre els retorns del capital i del treball. Aquest, però, continua essent un tema obert, ja que també és possible que el desplaçament tecnològic dels empleats resulti en un augment net d'ocupacions segures i satisfactòries. Aquesta incertesa en el cas de la tecnologia basada en la IA s'analitza més en profunditat a l'apartat 5.

A més de l'economia, la preocupació social més important relacionada amb la indústria 4.0/5.0 és la desigualtat. Els proveïdors de capital intel·lectual i físic, inventors, accionistes i inversors, són els beneficiats més grans d'aquesta revolució. Això explica la creixent disparitat de riquesa entre qui depèn del capital i aquells que depenen del treball. Això ajuda a explicar per què tants empleats estan desil·lusionats i preocupats que els seus propis salaris, i els dels seus fills romanguin estancats. També explica per què les classes mitjanes del món estan insatisfetes i es consideren tractades injustament. Una economia en què el guanyador ho acapara tot amb accés restringit a la classe mitjana és una recepta per a l'estancament i el fracàs democràtic.

L'omnipresència de la tecnologia digital i la dinàmica de l'intercanvi d'informació representada per les xarxes socials també poden fomentar el descontentament. Per interactuar, estudiar i intercanviar coneixements, més del 30% de la població mundial utilitza avui dia les xarxes socials. L'ús d'aquestes eines en un món ideal permetrien la comprensió i la integració intercultural, però en el món real, poden fomentar i promoure expectatives equívokes del que constitueix l'èxit per a una persona o un grup, així com proporcionar plataformes per a la difusió de creences i ideologies radicals..

En aquest context, la IA és un dels impulsors clau de la quarta revolució industrial i ja ens envolta i afecta la nostra vida diària: vehicles autònoms, assistents virtuals, programari que tradueix, inverteix i suggereix preferències culturals... en són només alguns exemples. La capacitat computacional i d'emmagatzematge sense precedents, l'accés a quantitats massives de dades i la gran adopció de la IA en molts àmbits diferents estan creant una simbiosi entre els mons digital i biològic que estan canviant la manera com vivim i interactuem amb l'entorn.

En referència novament a les preguntes que plantegem en la introducció i en línia amb els potencials i riscos esmentats anteriorment, en els capítols següents intentem brindar algunes idees i arguments addicionals sobre la IA i els seus impactes. En primer lloc, presentem una breu descripció de l'evolució històrica de la IA, després identifiquem els desafiaments i les oportunitats de la IA en un context general, i finalment els detallem en quatre àrees diferents, que són, negocis, govern, educació i estil de vida.

En els capítols següents, s'il·lustra una breu història de la IA, després es presenta l'estat actual de la IA, identificant les principals aplicacions, desafiaments i oportunitats, i se n'analitzen l'impacte a la nostra societat en termes generals, així com en quatre àmbits diferents, és a dir, negocis, govern, habilitats i competències, i gent i estil de vida.

3. Breu història de la IA

La IA és una disciplina relativament nova, nascuda a mitjans del segle XX. Turing, sovint anomenat el “pare de la IA”, va publicar el 1936 una descripció matemàtica del que va anomenar una “màquina universal” [3.1] i va ser el primer a publicar un mètode (prova de Turing) a “Computing Machinery and Intelligence” [3.2] per determinar si una màquina és capaç de pensar.

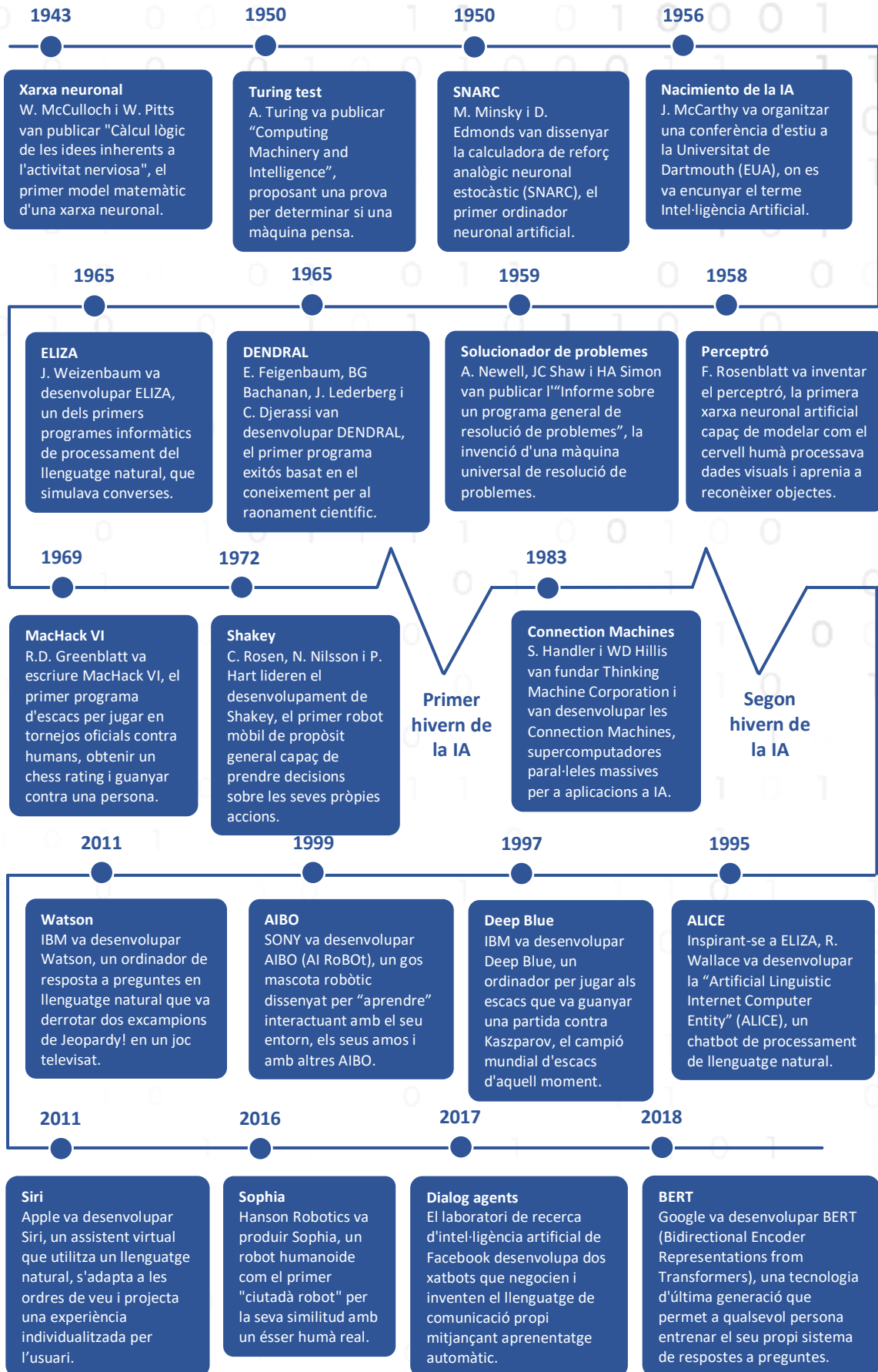
És difícil definir clarament la IA a causa de la diversitat d'aplicacions, problemàtiques i solucions en les que hi intervén, i la dificultat de distingir què és IA i allò que no ho és. IA és un terme encunyat per John McCarthy el 1955 i definit per ell mateix el 2007 [3.3] com “la ciència i l'enginyeria per fer màquines intel·ligents, especialment programes informàtics intel·ligents. Està relacionat amb el fet similar de fer servir ordinadors per comprendre la intel·ligència humana, però la IA no ha de limitar-se a mètodes que siguin biològicament observables”. L'Encyclopaedia Britannica [3.4] defineix la IA com “la capacitat d'un ordinador digital o un robot controlat per ordinador per fer tasques comunament associades amb éssers intel·ligents. El terme s'aplica sovint al projecte de desenvolupar sistemes dotats dels processos intel·lectuals característics dels humans, com la capacitat de raonar, descobrir significats, generalitzar o aprendre d'experiències passades”. El Cambridge Dictionary [3.5] defineix la IA en una frase més curta: “l'ús de programes informàtics que tenen algunes de les qualitats de la ment humana, com ara la capacitat de comprendre el llenguatge, reconèixer imatges i aprendre de l'experiència”. La definició més breu i senzilla és la que diu que “La IA no és intel·ligència biològica” [3.6].

En resum, no hi ha una definició universalment acceptada d'IA. IA és un terme general que es pot referir a un ampli camp de la ciència que abasta no només la informàtica. Si interpretem que la IA estudia com s'adquireix, processa, emmagatzema, utilitza... la informació en animals i màquines intel·ligents, òbviament se superposa amb diverses disciplines més antigues [3.7]: filosofia, matemàtiques i estadística, economia, neurociència, psicologia, biologia i ciències mèdiques, lingüística, informàtica, ciències tècniques, i ciències de la seguretat i la protecció.

La figura següent proporciona una descripció general i concisa de l'evolució històrica de la IA sense pretendre ser exhaustiva [3.8] - [3.10].

La primera fase d'expansió de la IA es va donar als anys 50/60. Durant aquells anys, es van produir grans avenços en molts aspectes que van establir les bases de la IA. Per exemple, la primera màquina de xarxa neuronal, SNARC, es creà el 1950. El primer assistent digital, ELIZA, es va desenvolupar el 1965. Tot i que la intenció del seu creador (Joseph Weizenbaum) era mostrar la superficialitat de la comunicació entre l'home i la màquina, va ser sorprès per la quantitat de persones que van atribuir sentiments humans al programa informàtic. El primer robot mòbil autònom d'ús general es va crear el 1972.

A aquesta primera fase li segueix, en la dècada dels 70, l'anomenat primer hivern de la IA, quan el desenvolupament es va alentir. Les previsions van resultar excessives, els enfocaments utilitzats fins aquell moment no van estar a l'alçada de les expectatives i la capacitat de les màquines va limitar les possibilitats. A la dècada dels 80, els sistemes experts van portar un breu auge, però gairebé de seguida va arribar el segon hivern de la IA. En els darrers anys, la IA ha iniciat una nova onada de progrés i entusiasme degut principalment a tres factors: 1) l'accés a ordinadors més potents; 2) la disponibilitat de grans bases de dades de fonts que inclouen comerç electrònic, empreses, xarxes socials, ciència i govern; 3) enfocaments i algorismes d'aprenentatge automàtic millorats.



4. Intel·ligència Artificial

La IA s'està estenent ràpidament i ja és present a totes les àrees de la vida. A continuació, sense pretendre ser exhaustius, es mostren exemples d'alguns dels usos de la IA.

- **Economia i finances:** Els sistemes basats en IA donen atenció al client, detecten anomalies i frau amb targetes de crèdit mitjançant el reconeixement de patrons i milloren la seguretat en diversos sectors comercials, inclosos el comerç al detall i les finances. Les organitzacions confien en la IA per rastrejar aquestes activitats mitjançant l'anàlisi del comportament de les transaccions. Les empreses empen mètodes científics de captura i anàlisi de dades per investigar patrons i preveure futurs moviments al mercat, i així millorar els seus beneficis al mercat borsari [4.1]. A l'apartat 5 s'analitzen més detalls sobre la IA als negocis.
- **Agricultura:** Les organitzacions estan utilitzant l'automatització i la robòtica per ajudar als agricultors en la gestió agrícola i facilitar altres tasques. Els sistemes d'IA troben maneres més eficients de protegir els cultius de les males herbes, combatre malalties animals i vegetals i monitoritzar el moviment, la temperatura i el consum d'aliments dels animals. Els sistemes d'IA permeten l'anàlisi d'imatges obtingudes via satèl·lit per a la coordinació de la producció a nivell regional i nacional, per exemple, identificar àrees propenses a la sequera [4.1].
- **Indústria de l'atenció sanitària:** Hi ha molts exemples de com la IA ha ajudat als pacients i s'espera que en el futur proper les aplicacions basades en IA podran millorar els resultats de salut i la qualitat de vida de les persones a tot el món. Les consultes personals o en línia, l'orientació sanitària personalitzada i els assistents virtuals redueixen les visites hospitalàries innecessàries. La IA utilitza la combinació de dades històriques i intel·ligència mèdica per al descobriment de nous fàrmacs, la detecció de malalties i la diagnosi [4.2].
- **Transport:** Amb el suport de la IA s'ha avançat enormement al sector dels vehicles autònoms (aviació en particular) [4.1]. Un dels exemples més clars són els cotxes autònoms i els drons no tripulats. Els automòbils actuals tenen funcions d'assistència al conductor basades en IA, com ara estacionament automàtic, controls avançats de la velocitat de creuer, interfícies de llenguatge natural i tecnologies d'assistència virtual.
- **Ciutats/edificis intel·ligents** [3.7][4.1]: El creixement esperat d'urbanització necessita gestionar la sostenibilitat ambiental, econòmica i social. Una ciutat intel·ligent/digital és un terme genèric pel que fa a una ciutat que utilitza tecnologies digitals per millorar l'eficiència operativa, compartir informació amb el públic, optimitzar l'ús dels recursos, brindar serveis d'alta qualitat i millorar el benestar de els ciutadans, incloent-hi gestió del trànsit, dels residus i dels serveis de seguretat (seguiment de zones d'alta criminalitat, sistema d'alerta primerenca d'incidents, etc.).
- **Educació:** La IA canviarà fonamentalment l'educació. La tecnologia revolucionarà la manera com els estudiants aprenen i transformarà la manera com els professors pensen sobre les pràctiques educatives [4.3]. Més detalls s'analitzen a l'apartat 7.
- **Robòtica:** Els robots sovint es relacionen amb la IA perquè és més fàcil per a l'ésser humà imaginar la IA en una estructura mecatrònica que també té un cos, especialment si el cos s'assembla a un humà/animal, per exemple, per a finalitats socials, terapèutics, o educatius. Generalment, els robots similars a màquines es troben a la indústria i la investigació, i es desenvolupen altres formes per a ús militar i mèdic.
- **Eines personals:** El reconeixement d'empremtes dactilars, l'assistent personal incorporat al telèfon mòbil, les aplicacions d'aprenentatge, l'escriptura automàtica de text parlat, els traductors automàtics d'idiomes i la mainadera virtual que parla amb els nens són només exemples de les nombroses funcions habilitades per a IA disponibles actualment per

simplificar la nostra vida diària

Els exemples anteriors demostren que la IA ja és present a totes les àrees de la nostra vida i està transformant radicalment el món. La humanitat s'enfronta a un desafiament existencial, davant del qual la consciència i una actitud activa pot generar un canvi positiu, en el qual la cooperació entre màquines i persones podria donar com a resultat un món ideal. En aquesta direcció, els participants dels grups focals han destacat la importància dels clàssics de la literatura de ciència ficció i robòtica per inspirar el desenvolupament de solucions del món real recolzades per la IA que ajudin a una evolució positiva de la humanitat.

“Crec que conec la diferència entre la ciència i la fantasia, però és la interacció entre aquests dos àmbits el que ha portat al desenvolupament de tots dos”

(expert/a)

La Recomanació [4.4] de l'OCDE deixa clar que el paper de la IA pot ser clau per donar forma positiva al futur, promoure el benestar de les persones i el benestar subjectiu, i contribuir al desenvolupament econòmic i aconseguir objectius sostenibles. Tot plegat va acompanyat de profunds canvis socials.

Sorgeix llavors la necessitat de que les persones es preparin per a aquests canvis, ja que si no en són capaces, és molt probable que es vegin perjudicats per la pròpia IA. Hi ha consens sobre que el debat ja no hauria de ser sobre si la IA és important o útil, sinó sobre com preparar la majoria de la societat per a aquest canvi i com més aviat millor.

“La societat no està preparada per als canvis ràpids provocats per la IA”

(expert/a)

A més, la IA també presenta desafiaments sense precedents, a mesura que aquesta nova tecnologia s'estén ràpidament al món, sorgeixen un munt de riscos ètics, morals i legals. A l'Apartat 6 es proporcionen detalls més específics sobre els riscos ètics.

En aquest sentit, Elon Musk, Stephen Hawking, Steve Wozniak (i molts altres) van signar una carta oberta a les Nacions Unides (ONU) [4.5] demanant imposar una prohibició sobre la militarització de la IA amb la capacitat d'apuntar i matar sense intervenció humana significativa. Tots els estats membres de l'Organització de les Nacions Unides per a l'Educació, la Ciència i la Cultura (UNESCO) van adoptar el 2021 [1.4] un acord històric que defineix els valors i els principis comuns necessaris per garantir el desenvolupament sostenible de la IA: un acord global sobre l'Ètica de la Intel·ligència Artificial [4.6].

És de destaca que en els nostres debats els participants condireren la possibilitat que la IA es torni tan autònoma que “sigui conscient d'ella mateixa” i evolucioni lliurement a formes que poden ser tant bones (millora de la seguretat) com dolentes (màquines que es revoltin contra els humans). També consideren que la millor solució és que, a mesura que la tecnologia vagi avançant, augmenti la nostra confiança en els dispositius basats en la IA, però que aquesta confiança només és possible si en qualsevol moment es pot tenir el control de la IA de manera segura.

“Allò que no entenem, generalment ho temem, i la nostra por sovint pren la forma de rebuig i odi. I així és amb la IA. Només hi ha un camí per a la humanitat: conèixer la intel·ligència artificial perquè pugui decidir conscientment si l'estima o l'odia”

(expert/a)

“Només crec en la intel·ligència artificial quan dóna el mateix resultat que la intel·ligència humana”.

(no expert/a)

La visió és que, en el futur –dins de 10 anys, el 2033–, els avenços poden ser tant bons com dolents, i que hi ha tants obstacles i barreres que és difícil fer una previsió realista basada en fets reals. Als participants dels grups focals els agradaria viure en un món més feliç, més segur i més còmode gràcies a la IA, però que això no comporti estar tan “acaronats” que l'experiència de la vida resulti “avorrida” en tots els aspectes.

“Estaria molt més tranquil si d'aquí 10 anys hi hagués una visió positiva del món i pogués dir que gràcies a la IA els meus pares estan segurs en una llar intel·ligent”.

(no expert/a)

Aleshores sembla urgent [4.7] el desenvolupament d'un nou paradigma en què la humanitat pugui definir la seva visió: la institucionalització de la IA. És important tenir en compte que la IA no ha de servir només els interessos d'un estrat reduït de la població (orientat a l'ús manipulador i amoral i als beneficis de la IA), sinó que ha de col·locar el bé públic per sobre dels interessos individuals. De fet, hi ha un nombre creixent de desenvolupaments -no només al camp militar- en els que no es respecten els aspectes ètics.

“Les recomanacions, regulacions i lleis sobre intel·ligència artificial seran bones si es compleixen. Si bé el càstig pot ser un element dissuasiu en alguns casos, malauradament es pot dir que, en general, cap càstig dissuadirà a aquells que desenvolupen IA amb males intencions i per als seus propis interessos”.

(expert/a)

La consciència individual dels canvis relacionats amb la IA es considera necessària. Això inclou fer de la ciutadania digital una part integral de la vida quotidiana, fomentant el diàleg social sobre la IA: les persones han de ser agents del canvi, en comptes de víctimes. Per tant, és essencial mantenir informació autèntica i de qualitat, i restaurar la confiança social mitjançant el desenvolupament d'un nou paradigma en què la humanitat pugui definir-ne la visió: la institucionalització de la IA i la realització d'una IA centrada en l'ésser humà [4.7].

“Abans era home contra home, després tecnologia contra tecnologia, i ara sembla que intel·ligència artificial contra intel·ligència artificial. Potser estariem millor si triem la cooperació en lloc de la lluita”.

(expert/a)

5. Negocis

Considerada un motor clau del creixement i desenvolupament econòmic futur, la IA s'ha convertit en la tecnologia principal per a un nombre cada vegada més gran de noves empreses, particularment a Europa [5.1]. La IA està conquerint totes les indústries a tot el món i motivant les empreses a competir per convertir-se en entitats centrades en la IA. L'entorn empresarial competitiu ha obligat els líders corporatius, empresaris, estratèges i investigadors a emprar la IA per desenvolupar noves solucions i generar noves fonts d'ingressos [5.2]. La Comissió Europea [5.3] indica que, el 2017, el 25% de les grans empreses de la UE i el 10% de les petites i mitjanes empreses van utilitzar anàlisi de Big Data. Només una de cada cinc petites i mitjanes empreses es va digitalitzar significativament, i un terç de la força laboral encara no té habilitats digitals fonamentals.

Mentrestant, els avantatges de les aplicacions de la IA són àmpliament reconeguts. Alguns exemples són:

Les empreses dels sectors de **comerç, agroalimentació i construcció** que han implementat IA reporten excel·lents resultats en el desenvolupament de productes o serveis, atraient nous clients i obrint nous mercats [5.3].

En **producció**, la IA ajuda a millorar la qualitat dels sistemes i, en conseqüència, la qualitat dels productes. La IA també permet la creació de productes altament personalitzats, que coincideixin amb els interessos dels clients.

El **màrqueting** és un dels camps més avançats pel que fa a la IA. Per exemple, mètodes basats en la IA poden ajudar a pronosticar si la despesa futura d'un nou client es reduirà o augmentarà després de la compra inicial.

A l'**administració d'empreses**, la IA es fa servir àmpliament en la gestió de recursos humans per millorar els processos de presa de decisions integrant sistemes tècnics, humans i organitzatius per aconseguir l'èxit estratègic d'una empresa.

La IA ha revolucionat les compres en línia. Les principals empreses de **comerç electrònic**, com Amazon, Alibaba, i eBay, fan ús de la IA per oferir recomanacions de productes que poden interessar els clients, cosa que es tradueix en un creixement significatiu dels ingressos [5.4].

La IA és un component crucial de les populars plataformes de xarxes socials, que s'utilitzen principalment amb finalitats comercials. LinkedIn fa servir la IA per proporcionar suggeriments de treball, suggerir nous contactes i diferents continguts d'interès personalitzats per a l'usuari [5.1].

La majoria dels experts entrevistats van coincidir en el fet que la IA agilitza les tasques, en particular, les que no requereixen processos complexos i, per tant, redueix els errors simples i estalvia temps

“Fem servir IA per tenir anàlisi de sentiments i tot això és avantatjós. El cost relacionat amb aquestes activitats pot ser un desavantatge, i trobar persones que puguin administrar aquests processos és problemàtic”.

(expert/a)

Si bé hi ha un interès creixent entre les empreses a invertir i incorporar la IA a les seves operacions, hi ha barreres importants a nivell organitzatiu que impedeixen que les empreses es

beneficiïn de tot el potencial de la IA. Algunes de les barreres principals són la manca d'una estratègia clara d'IA, la resistència cultural, la manca del talent necessari per a les solucions d'IA, la mida de l'empresa i les limitacions pressupostàries [5.2].

Els resultats mostren que a Europa la manca de finançament públic i de capital de risc sol assenyalar-se com l'obstacle principal per al desenvolupament de la IA, en particular, a les pimes i les empreses no tecnològiques [5.5]. En aquest sentit, Europa ha augmentat la seva inversió i aposta per la investigació a IA per augmentar el seu potencial de creixement tecnològic i posar-se al dia amb els països que lideren la carrera en IA. Segons la Comissió Europea [1.3], la IA pot impulsar productes i serveis innovadors en àrees on Europa destaca (maquinària, ciberseguretat, transport, agricultura, economia verda i circular, sanitat i sectors d'alt valor afegit com la moda i el turisme). El desenvolupament de tecnologies d'IA a Europa també implica una dependència menor de tecnologies estrangeres, la qual cosa és vital per a l'autonomia d'Europa, però també impulsar un desenvolupament de la IA amb valors europeus. Aquestes idees sobre com adoptar aquesta tecnologia en els negocis són essencials per guiar la legislació i garantir que la IA beneficiï tant els ocupadors com els empleats, fent que sigui fiable, fàcil d'usar i útil en el treball diari [5.6].

No obstant això, en aquest procés de transformació, el món dels negocis és força escèptic sobre aquest ràpid canvi tecnològic i els seus impactes [5.7]. Les preocupacions sobre la presa de decisions, la privadesa, l'ètica i la confiança també hi són presents i s'espera que augmentin en un futur proper. Tant els experts com els no experts van destacar als nostres grups focals la importància de l'habilitat de la IA per millorar les capacitats i perspectives en els negocis. Tot i això, també es va afirmar que la IA no ha de reemplaçar la capacitat humana i la presa de decisions, de manera que la interacció humana seguirà sent necessària.

“Els humans són els que millor poden detectar objectes en camps com la biomedicina, però ja podem fer servir la IA per a deteccions simples. I en camps mèdics més amplis, és important fer servir IA, però hem d'estar certs que és segura i no causarà grans errors. El coneixement expert no ha de ser reemplaçat per IA sinó millorat. Heu de donar recomanacions, però no heu de decidir. Preferiria que la IA millorés el que els humans poden fer. Estem molt lluny d'aquest punt on la IA pugui reemplaçar el coneixement humà en tots els casos. A més, pot ser perillós. La responsabilitat (dilema del tramvia) és un altre tema que cal tenir en compte a l'hora de prendre decisions a IA”.

(expert/a)

“La IA no es pot implementar completament ja que necessitem interacció i necessitem parlar amb persones reals”.

(no expert/a)

“La interacció humana és important, encara que es faci servir la IA. També hauríem de preservar una mica de llibertat en el procés de presa de decisions com a humans”

(no expert/a)

El tema més sensible en temes de privadesa és l'exposició d'informació personal. Per això, les empreses s'esforcen per establir objectius específics per guanyar-se la confiança. Per exemple, Google va declarar que no oferirà aplicacions basades en IA que causin, o puguin causar, un perjudici generalitzat i només procedirà si els beneficis superen els riscos, incloent-hi, però, límits de seguretat apropiats si hi ha un perill real de dany [5.8]. Si bé la majoria dels estats tenen lleis

de protecció de dades, la IA té el potencial de generar nous problemes de protecció de dades que no s'aborden a la legislació actual, cosa que planteja problemes ètics addicionals. La IA també podria utilitzar o generar tipus de dades personals que ara no es consideren, com a informació emocional, cosa que augmenta la magnitud del problema [5.9]. A la Secció 6 es presenten més detalls sobre els aspectes ètics..

Els estudis prediuen que la IA pot transformar el concepte de força laboral en extingir algunes feines. La investigació realitzada a [5.10] sobre l'impacte potencial de la IA a l'activitat econòmica mundial subratlla el gran potencial tant de desenvolupar nous llocs de treball com de reemplaçar alguns dels existents. S'estima que hi haurà més llocs disponibles per a enginyers, desenvolupadors de programari i professionals de les TIC a Europa en un futur proper. Segons l'informe sobre les tendències de capital humà de Deloitte [5.11], els llocs de treball en creació estaran més orientats als serveis, interpretatius i socials, i requeriran creativitat, empatia, comunicació i habilitats especials per resoldre problemes. L'informe d'Accenture [5.12] afirma que els llocs de treball impulsats per la IA seran formadors, explicadors i sustentadors. Aquests nous treballs inclouran la capacitació de sistemes d'IA, garantint que continuïn funcionant segons allò dissenyat i no aprenguin "coses incorrectes", tancant així la bretxa entre el negoci i la tecnologia. A la Secció 7 s'inclouen més detalls sobre el nou sistema laboral i educatiu relacionat amb la IA.

Els nostres entrevistats també van expressar la seva preocupació sobre l'optimització de costos que segueix a la implementació de certa IA i que dona com a resultat nivells més alts de desocupació per als humans.

“Les empreses poden beneficiar-se de la IA per renovar la seva imatge i fer una bona impressió. També optimitza els costos en recursos humans. És bo per a la perspectiva empresarial. Però quant a la RSE (responsabilitat social empresarial) i el treball humà, pot ser problemàtic, ja que la IA reemplaça algunes feines”.

(no expert/a)

L'opinió general és que tot depèn del sector.

“La IA pot obrir portes per a noves tasques, però en alguns camps simplement pot reemplaçar el treball humà i reduir la complexitat. Aleshores, realment depèn”.

(no expert/a)

Finalment, la IA serà una tendència de mercat i una oportunitat de negoci durant la propera dècada. Es preveu que aporti 15,7 bilions de dòlars al PIB mundial, incrementant-se un 14% per al 2030. Els analistes prediuen un augment de la producció de 6,6 bilions de dòlars, amb un augment del consum de 9,1 bilions de dòlars. Si Europa desenvolupa i difon la IA d'acord amb els seus actius actuals i la seva posició en relació amb la resta del món a l'àmbit digital, pot contribuir amb 2,7 bilions d'euros, és a dir, el 20%, a l'economia per al 2030, cosa que resultarà en un creixement compost anual de l'1,4 %. Aquest impacte seria gairebé el doble que el de les tecnologies anteriors de propòsit general que els països desenvolupats han adoptat en el passat [5.10].

6. Governança

Hores d'ara ja sabem que la IA és avui part de les nostres vides. Podem ser conscients de la seva presència i interactuar-hi, com, per exemple, quan demanem a Siri que ens busqui un restaurant d'acord amb les nostres preferències gastronòmiques. Però, en molts altres aspectes, no som plenament conscients de que la IA també hi és present. Les institucions financeres aprofiten la IA per identificar activitats potencialment fraudulentament als nostres comptes; les IA s'utilitzen per rastrejar i predir els impactes ambientals als camps agrícoles utilitzant dades d'escaneig per satèl·lit i monitoratge de la salut dels cultius i el sòl. Aquests són només alguns exemples i, segons diversos estudis, l'epidèmia de Covid-19 ha accelerat l'adopció de la IA a tots els sectors de l'economia [6.1].

D'altra banda, la IA no és tot diversió i jocs. Molts acadèmics assenyalen que la manera com es desenvolupen les eines d'IA ha de canviar a causa de les limitacions en la col·laboració i les suposicions de dades inexactes, com les expectatives poc raonables que prediuen sense prou solidesa l'ús de sistemes d'IA. Per exemple, la inacció davant dels prejudicis de la IA ha provocat moltes injustícies a grups socials i perfils racials i altres incidents preocupants. Els deepfakes i la capacitat de crear vídeos, imatges, text, discursos i altres formes de comunicació (social) falsos, però realistes han generat moltes preocupacions ètiques i legals darrerament sobre l'ús de la IA per manipular les percepcions humanes. En ciberseguretat, els atacants també tenen accés a eines d'IA, de manera que el joc del gat que empaita al ratolí continua. La videovigilància basada en IA per reconèixer les persones a través del seu rostre, parla, forma de caminar o moviment també ha plantejat algunes preocupacions sobre la privadesa. Amazon Alexa va suggerir recentment a una nena de 10 anys que toqués un endoll amb una moneda després que aquesta li demanés un repte [6.2].

En aquest escenari de pros i contres quan es tracta d'IA, la implantació d'una bona governança es considera fonamental. La governança fa referència a la formació, manteniment i regulació de regles o activitats, així com a l'assignació de responsabilitats [6.3].

Quan la IA i el terme governança es barregen, es poden trobar dues interpretacions diferents: i) L'ús de sistemes basats en IA a la governança, cosa que significa l'adopció de la IA en la prestació de serveis, la formulació de polítiques i la seva aplicació als ecosistemes del sector públic [6.4]; y ii) La governança de la IA, la qual cosa significa la promoció d'un marc institucional i legal adequat per al desenvolupament i l'ús de la IA [6.5]. Tot i això, no és possible mantenir una discussió sobre la IA en la governança sense considerar la governança de la IA, perquè funcionen com a vasos comunicants. Aquí entendrem la governança com el que es coneix com a "AI governance", una idea composta per tres components relacionats amb: a) la infraestructura, relativa a l'obtenció, emmagatzematge i processament de dades; b) l'aplicació relativa a la gestió de dades; i c) la utilització, relativa als processos de presa de decisions i avaluació basats en dades.

A la literatura es poden trobar moltes altres definicions d'"AI governance" [6.6] - [6.8]. Probablement, la definició més completa és la que es troba a [6.9], on s'afirma que "l'AI governance és un sistema de regles, pràctiques, processos i eines tecnològiques que es fan servir per garantir que l'ús de les tecnologies d'IA per part d'una organització s'alinea amb les estratègies, objectius i valors de l'organització; compleix els requisits legals; i obeeix als principis d'IA ètica seguits per l'organització". En poques paraules, l'AI governance ha de tancar la bretxa que hi ha entre la responsabilitat i l'ètica en l'avenç tecnològic [6.10], i assegurar que s'estableixin límits fiables dins de la tecnologia, perquè no causi mal i agravi encara més les desigualtats quan s'aplica incidentalment mentre opera.

“Es va finançar amb fons de paradisos fiscals un projecte de recerca per reconèixer una persona a partir de l'iris, per identificar les dones amb burca i saber si estaven o no amb el marit. A mi em va sorprendre molt [...]. Com fer-ho? Arribar fins aquí, sí? Arribar fins aquí, no? Quins límits?”

(no expert/a)

Hi ha un consens ampli sobre la necessitat de discutir aquests límits fiables en el desenvolupament de sistemes d'IA, ja que el seu ús pot tenir conseqüències negatives molt importants per a la vida de les persones, o reproduir models socials que es consideren moralment reprovables.

Tanmateix, els límits no són clars i és difícil establir o acordar un marc ètic, polític o regulatori que pugui definir el desenvolupament de formes d'IA que després puguin tenir un alt impacte en les decisions socials. Una de les dificultats que sorgeix en aquest sentit és la tensió entre les garanties per als ciutadans i, alhora, mantenir la competitivitat en recerca i innovació.

Els participants, tant experts com no experts, van destacar temes importants relacionats amb aquests dos àmbits específics:

- **Biaix de dades.** L'èmfasi està en la necessitat de garantir que les dades recopilades no estiguin esbiaixades per gènere, nivell socioeconòmic, ètnia, etc. La garantia de la diversitat de dades i la seva composició es refereix a l'ús d'IA en totes les etapes del procés: la recopilació de dades, la decisió en si o l'avaluació. L'ús de màquines per a la presa de decisions no està exempt de la ideologia subjacent a tota decisió. Aquestes ideologies poden representar interessos de diversos actors, ja siguin de caràcter polític, tècnic o econòmic. Aquest és un tema important a resoldre per garantir que les dades recopilades i el seu ús responguin als objectius per als que estan dissenyats.

“Els humans prenen moltes decisions sobre la base d'una ideologia (...) Una màquina també prendrà una decisió esbiaixada. Esbiaixat per qui? Per les dades, per l'enginyer que el va dissenyar o per l'empresa que hi ha al darrere, o per la ideologia de l'estat que el va finançar”

(expert/a)

- **Justícia.** Utilitzar la capacitat de manejar grans volums de dades i fer prediccions estadístiques es considera un valor important de la IA. Si les decisions afecten qüestions no substantives per a la vida de les persones, un possible error en les decisions de la IA es pot considerar un problema menor i, per tant, la IA es podria utilitzar per prendre decisions. No obstant això, en cas que les decisions afectin qüestions substantives de la vida de les persones, una decisió errònia podria tenir efectes terriblement injustos que podria condicionar la vida de la persona, per la qual cosa aquests casos les decisions no haurien de ser preses per sistemes d'IA.

“Al llarg dels anys hem construït un sistema judicial, que volem mantenir. Hi ha àrees on l'impacte [de les decisions preses per la IA] sobre les persones és molt important. La IA no hauria d'entrar en aquestes àrees”.

(no expert/a)

- **Privatització.** Certes empreses o corporacions estan acumulant molt coneixement algorítmic i dades sobre el comportament de la ciutadania, i no hi ha garanties que aquestes dades o aquest coneixement es faci respectant principis o valors ètics

consensuats. En aquest sentit, l'acumulació de dades i coneixements per l'ús d'IA per part d'entitats fora de l'àmbit de la supervisió governamental implica una privatització de la governança no hauria d'estar permès.

“Crec que s'haurien de crear institucions reguladores, de la mateixa manera que hi ha institucions que regulen els bancs i els auditen per veure què fan amb els diners. Haurien d'auditar aquestes empreses com Google, Netflix, etc., per veure quins són els seus algorismes i què estan fent realment amb la IA”.

(expert/a)

- **Automatització de decisions.** El debat sobre la limitació de l'automatització dels processos de presa de decisions no s'ha de circumscriure a la IA, de la mateixa manera que la IA no es pot considerar l'única responsable de l'automatització de les decisions. Però, la IA planteja problemes quan els que dissenyen un algorisme no poden explicar les seves decisions, així com quan els usuaris no coneixen els criteris que el dissenyador de la IA que ha desenvolupat l'algorisme. Podem caure en el que s'anomena la "síndrome de l'ordinador diu que sí" [6.11], on els empleats que han passat molt de temps bregant amb un sistema en què els errors són poc freqüents (com hauria de ser el cas dels sistemes d'IA) es tornen naturalment menys propensos a desafiar la correcció del sistema de manera progressiva.

“Hi ha algorismes usats en la IA que no estan degudament regulats, cosa que pot ser un problema. També sóc escèptic respecte que la IA no cometi errors humans. Però, i si l'algorisme d'IA es va entrenar de manera falsa o amb algun biaix (per exemple, racisme)? En aquest cas, la IA pot cometre errors. Aleshores, qui prendrà la darrera decisió, la IA o l'ésser humà? No sabem com la IA pren una decisió. Per què es va prendre qualsevol decisió, no ho sabem. Per tant, cal fer un esforç perquè els sistemes d'IA siguin transparents i així poder comprendre per què decideix a favor o en contra de les coses”.

(expert/a)

- **Llibertat.** L'amenaça a la llibertat que suposa l'ús de dispositius d'IA en els processos de presa de decisions es pot entendre des de dos nivells diferents. La primera dimensió es refereix a les estratègies que utilitzen la IA per aconseguir més publicitat o visualització, a partir d'algorismes que fan que els usuaris entrin en processos tipus bucle, que són utilitzats per empreses com ara Meta o Twitter. La segona dimensió està relacionada amb la capacitat de la IA per controlar les emocions i els sentiments. Tenint en compte el rastre digital que tots els ciutadans deixen en tots els seus moviments diaris, obtenir i utilitzar aquestes dades amb finalitats comercials o autoritàries pot resultar molt perillós.

“Tinc fills adolescents, que es creuen el que veuen: les notícies falses, el lleixiu que van prendre per curar-se del covid. Tinc una filla de 12 anys. Noto que la informació que reben és un perill brutal. La gent és impressionable i això és molt complicat. Quan comences a veure un contingut, quan som joves, busquem notícies que és el que esperes, som més influenciables. Si veus un vídeo que hi surt... Doncs dius 'Vull anar a Malibú', 'Vull una bossa de Prada'. L'algorisme et mou”.

(no expert/a)

Les mesures proactives de governança són cada cop més reconegudes com una característica diferenciadora per a les empreses que busquen establir una reputació de confiança.

Hi ha una sèrie de marcs internacionals sobre conceptes d'ètica i governança de la IA. La Unió Europea va emetre el Reglament General de Protecció de Dades (GDPR) que inclou un conjunt de regles que estableixen el dret del consumidor de no estar subjecte a una decisió automatitzada. Això no obstant, també va causar certa controvèrsia, ja que no atorga el dret a una explicació de la presa de decisions automatitzada [6.12]. En aquest sentit, és probable que la UE sigui la primera en promulgar una legislació reguladora de la IA [6.13]. La Llei de Responsabilitat Algorítmica [6.14] als EUA requereix que les principals empreses amb accés a grans quantitats de dades auditin els sistemes impulsats per IA per verificar l'equitat, la privadesa, la precisió i els riscos de seguretat. Una iniciativa notable és el Marc de Governança de la IA de Singapur. És el primer model desenvolupat a Àsia i la seva fortalesa rau en el fet que tradueix els principis en un marc pràctic i operatiu per a l'acció immediata i disminueix les barreres d'entrada a l'adopció d'IA. Aquest marc es basa en dos factors: i) les solucions d'IA han d'estar centrades en l'ésser humà i ii) les decisions preses o assistides per IA han de ser transparents, explicables i justes.

S'està treballant com arribar a un bon compromís entre les estratègies i els objectius de les empreses i els requisits legals ètics d'ús de la IA. Per exemple, la Universitat de Harvard [6.15] ha creat un mapa de visualització de 32 conjunts de principis d'IA. KPMG [6.16] proporciona quatre guies per ajudar les organitzacions a garantir el govern adequat dels algorismes d'IA. Google [6.5] destaca cinc àrees específiques en què l'orientació precisa i específica del context dels governs i la societat civil ajudaria a avançar en el desenvolupament legal i ètic de la IA.

La **responsabilitat** requereix una identificació clara de qui respon de les decisions i les accions en dissenyar, desenvolupar, operar i/o implementar sistemes d'IA. Han de ser persones o organitzacions les que siguin responsables en darrera instància dels actes dels sistemes d'IA, sense importar com de complexos siguin.

La **transparència** fa referència a la capacitat d'explicar per què un sistema d'IA es comporta de certa manera per augmentar la confiança de les persones en la precisió i l'adequació de les seves prediccions. En principi, com més sentiu els usuaris que entenen el comportament d'una IA, més inclinats estaran a usar-lo.

L'**equitat** ha de garantir que els sistemes d'IA siguin ètics, lliures de biaixos, lliures de prejudicis i que no s'utilitzin atributs protegits. Hi ha diferents punts de vista per definir l'equitat i es poden contradir entre ells. Això es pot mitigar si es decideix per endavant la tècnica exacta d'imparcialitat que es farà servir i es fa transparent.

La **seguretat** fa referència a prendre mesures contra l'abús inadvertit i intencional de la IA que pugui representar una amenaça per a les persones. Això, però, s'ha de fer de manera raonable, tenint en compte el dany potencial i la practicitat de les mesures preventives suggerides en termes tecnològics, legals, econòmics i culturals.

El **control humà** vol dir que les persones han d'estar en algun moment en el procés de presa de decisions d'un sistema automatitzat. En principi, independentment de com sigui de precís un sistema d'IA, la societat vol que un ésser humà prengui la decisió final.

La **universalitat** recomana la definició i l'aplicació d'estàndards tècnics, ètics i reguladors durant el desenvolupament, l'avaluació i el desplegament d'algorismes per tenir interoperabilitat, cooperació i un nivell específic de qualitat, seguretat i confiança.

7. Habilitats i competències

Tot i la seva joventut, la IA està influint extraordinàriament en el mercat laboral. En primer lloc, a causa de l'automatització, alguns treballs d'habilitats intermèdies estan desapareixent. En segon lloc, ara és més probable que les persones facin servir IA en la seva vida quotidiana i en el seu treball, ja que aproximadament el 50% de les organitzacions informen que usen IA [7.1]. Centrant-se en les feines que estan desapareixent a causa de la IA, els participants als grups focals van definir tres categories de feines que estan sent reemplaçades per la IA: 1) Feines repetitives/rutinàries on la IA, a més de substituir els humans, aconsegueix disminuir el risc d'errors i augmentar la productivitat; exemples d'aquesta mena de feines són els treballs de fàbrica i magatzem, conductors, empleats de peatge. 2) Feines administratives/d'oficina, que afecten secretàries/s, recrutadors, advocats, assessors fiscals i legals, assistents, traductors. 3) Feines físicament exigents, com ara la (des)càrrega de camions i (des)paletització de caixes.

“Per exemple, un treballador de magatzem que actualment utilitza un carretó elevador per moure palets aviat li tocarà gestionar una petita flota de robots de moviment autònom (AMR).

(expert/a)

A causa d'això, les institucions d'educació i formació s'han d'adaptar per proveir als estudiants de les habilitats i les competències que es necessiten en aquest món tan canviant. Això és particularment necessari i una manera de combatre la desconfiança i la por que puguin tenir les persones de que l'automatització i la digitalització reemplacin als humans [7.2] - [7.4]. De fet, la majoria dels europeus estan a favor de que els governs limitin la aplicació de l'automatització i la digitalització als llocs de treball per tal de protegir i mantenir les persones empleades [7.2], [7.4]. Tot i això, la naturalesa canviant de les modalitats de treball i el desplegament de noves tecnologies és inevitable, el 37 % dels participants de l'enquesta de Gartner 2019 CIO van afirmar que ja utilitzen IA i/o ho faran en un futur pròxim per intentar mantenir-se per davant de la seva competència, o almenys no quedar-se enrere [7.5]. A més, informes anteriors ja indicaven com les feines de les persones reemplaçables per la IA serien abandonades i substituïdes per la tecnologia, cosa que ha contribuït a corroborar la visió errònia i el temor de les persones respecte a la IA. Malgrat això, els informes més recents s'han concentrat en com la IA crea llocs de treball i/o permet que els treballadors puguin ser alliberats de treballs servils i/o perillosos per realitzar tasques més satisfactòries [7.6] - [7.8].

Es presta molta atenció al temor que els llocs de treball siguin reemplaçats per màquines i poca al fet que no són els llocs de treball en si mateixos els que es tornaran inherentment obsolets, sinó que la forma de treballar canviarà i que certes habilitats i competències guanyaran importància mentre que d'altres seran abandonades [7.9]. De fet, els beneficis d'usar sistemes d'IA depenen en gran mesura de les competències i habilitats dels qui els operin, de manera que la manca d'habilitats en l'ús de la IA és el problema número u de les empreses d'aquest sector [7.9]. Per tant, el desplegament de plans d'estudis educatius centrats en el coneixement de la IA i els seus desafiaments s'ha identificat com a crucial per abordar el canvi del sector laboral i de negocis [7.10].

A més, la IA pot ajudar a augmentar les habilitats i competències dels treballadors [7.8], [7.11]. Només cal mirar l'exemple de KONE, que va equipar els seus ascensors amb sensors IoT (Internet of Things) i va fer servir un sistema d'IA per analitzar les dades, cosa que va permetre als tècnics estar informats sobre possibles anomalies i realitzar un manteniment preventiu [7.11]. També, encara que no sigui estrictament una IA, l'entrenament amb simulació híbrida ha mostrat un gran

potencial a les professions d'infermeria i salut [7.12].

Segons un estudi recent [7.11], la IA pot funcionar com a motor d'innovació de les empreses per renovar competències. És a dir, destriar les habilitats i coneixements existents vigents de les obsoletes. Aquesta tria pot ser especialment útil perquè els gerents d'empreses identifiquin i prediguin l'impacte que pot tenir el desplegament de la IA a les competències de la empresa. Actualment, la majoria de les aplicacions d'IA fomenten les innovacions que milloren les competències, però en el futur es preveu un augment de les innovacions que, a causa de la millora de l'aprenentatge automàtic i la resolució de problemes amb l'ajut de la IA, eliminin competències obsoletes.

Una qüestió rellevant respecte a la IA consisteix identiicar aquelles habilitats i competències que són impossibles de replicar per robots. A [7.13], els autors afirmen que “la interacció social complexa i la creativitat són les coses més difícils d'automatitzar”. A causa d'això, és poc probable que els educadors siguin reemplaçats per la IA malgrat la seva creixent implantació a l'educació en part a causa de la pandèmia, que la va convertir en una necessitat [7.14].

“Crec que necessitem més gent tècnica, més coneixements sobre l'evolució de la societat [...]. I també per altra banda, a la gent que està més en el camp de la governança [...] que també entenguin aquest nou soci que té un abast a tot arreu... A nivell educatiu hem d'intentar fer un esforç per integrar aquesta IA a tota la base de coneixement existent”.

(expert/a)

El document de treball [7.14] destaca la idoneïtat de la IA per millorar la personalització i els resultats de l'aprenentatge. A més, facilita la possibilitat d'ajustar i personalitzar el progrés a la feina de cada persona, ja que permet desplegar plans d'aprenentatge, preferències i trajectòries individualitzades. Els experts van coincidir que la IA ja ha canviat l'entorn de l'educació superior, ja que ara tothom pot adquirir noves competències de forma individual.

“S'espera que la IA creï una educació flexible i personalitzada en analitzar i detectar els punts febles o diferents dels estudiants”.

(expert/a)

Això es pot fer, per exemple, mitjançant l'ajuda d'un MOOC, amb el qual no només es pot arribar a una gran quantitat d'estudiants, sinó que també inclou un enfocament d'aprenentatge flexible per a estudiants de tota mena. Pels professors, la tecnologia basada en la IA pot representar una eina per a l'avaluació de qualificacions, així com una ajuda per a la preparació de la lliçó i el seguiment dels grups de discussió. A causa dels components socioemocionals i creatius de l'ensenyament eficaç, els autors a [7.14] qualifiquen com a força improbable la substitució de docents per la IA. Això no obstant, subratllen la importància de que els docents s'adaptin a la nova era digital desenvolupant noves metodologies i millorant les seves habilitats.

“Podria ser útil fer algunes classes sobre IA als estudiants més joves perquè sàpiguen com bregar amb ella i coneixin els riscos/desafiaments. Ja la fem servir, però encara és complicat i no hi reflexionem prou. Potser pot ser útil tenir ensenyaments d'aquests per aprendre més sobre el tema”.

(no expert/a)

Una habilitat important que ha guanyat rellevància a l'era de la IA ha estat el pensament crític. En aquesta era en què les notícies falses es poden difondre fàcilment a través de les xarxes

socials, les habilitats de pensament crític i la validació de les fonts són primordials [7.15]. El pensament crític permet a les persones no sols arribar a conclusions basades en l'evidència actualment disponible, sinó també qüestionar-ne les creences quant a precisió i rellevància a causa de fonts més noves o diferents, i modificar-les en conseqüència [7.16]. A més, les habilitats socials continuen sent especialment rellevants per la seva dificultat per ser automatitzades. Altres habilitats pertinents són la capacitat d'adaptar-se al canvi, treball en equip, resolució de problemes, comunicació i ajuda als clients en la gestió de projectes i l'ús de TI [7.16]. Els acadèmics estan d'acord àmpliament que la integració d'aquests continguts és essencial per assegurar una transició beneficiosa cap a l'era de la IA [7.10][7.17].

En aquest sentit, un estudi recent a gran escala [7.17] va concloure que la necessitat d'habilitats físiques, cognitives bàsiques i manuals es reduirà a causa de la implantació de la IA. Els autors han identificat 56 DELTAS (elements diferents de talent, actituds i habilitats) dividits en 13 grups d'habilitats, que alhora es divideixen en quatre categories principals: Cognitiva, Interpersonal, Auto lideratge i Digital. Aquests inclouen el pensament crític esmentat anteriorment, el treball en equip, etc., però també d'altres com el coneixement i la ciutadania digital, l'ús i el desenvolupament de programari, i la comprensió dels sistemes digitals. Aquestes conclusions estan en línia amb altres investigacions recents, com [7.1], [7.9], [7.18], que destaquen la necessitat de dades, coneixement tecnològic i digital, així com competències en resolució de problemes, empatia, comunicació, innovació, pensament crític i treball en equip. Tot i això, els participants en els nostres grups focals també van assenyalar que aquestes habilitats no necessàriament es generalitzaran més:

“Hi ha una diferència entre allò que és important i allò que es promocionaria. I realment no sé quines competències... a part de les òbvies, OK, competències digitals i aquestes coses... que en realitat serien promogudes per un món basada a la IA en si mateix”.

(no expert/a)

A més, hi ha la percepció que el sistema educatiu és massa lent per canviar i ensenyar les habilitats i competències necessàries per a un món centrat en la IA.

“El sistema educatiu encara es basa en la lògica que molts coneixements han d'aprendre's de memòria mentre que ara són a l'abast de tothom en tot moment. Si bé hi ha hagut un petit canvi cap a una educació basada en competències, això no és suficient. Un enfocament més gran en l'educació tecnològica, així com en la resolució de problemes complexos, per exemple, a través del treball en projectes, hauria d'ocupar un lloc més important en l'estratègia educativa”.

(expert/a)

Si bé la implementació de la IA en el context educatiu ofereix una gran quantitat de beneficis, no s'han de passar per alt els problemes associats a aquest procés. D'acord amb [7.14], el desenvolupament d'una política pública integral de la implementació de la IA per fomentar el desenvolupament sostenible és fonamental. Tot i que la IA en si mateixa pot representar una oportunitat per a la inclusió, per exemple, a través de la possibilitat d'aprenentatge a distància, els desafiaments com l'accessibilitat a l'electricitat, el maquinari i Internet, els costos de les dades, les habilitats bàsiques de TIC, l'idioma i la idoneïtat cultural del contingut només són alguns aspectes que cal tenir en compte. Aquest és el cas dels països “menys desenvolupats”, que corren el risc de quedar-se encara més enrere si aquests desafiaments no es reconeixen i aborden de manera exhaustiva. De manera similar, la IA es basa en les dades que l'alimenten,

per tant, la qualitat i la inclusió han de ser una de les principals prioritats en desenvolupar tecnologies basada en la IA. Mitjançant la transparència i un codi ètic clar, es pot evitar que la IA perpetui les desigualtats [7.14].

Com a consideració final dels canvis actuals causats per la IA, també s'han de considerar els nous treballs creats per ella, ja que cal que les persones entrenin la IA (capacitin la IA perquè faci la feina prevista), expliquin els resultats de la IA (com la IA va arribar a una conclusió particular per al no expert) i mantenir l'ús responsable de la IA (garantir que els sistemes d'IA funcionin correctament, de manera segura i s'utilitzin de manera responsable) [7.8]. Aquests nous treballs necessiten noves habilitats i educació. Per això, alguns països, com Estònia amb el programa ProgeTigerProgramme llançat el 2012, estan començant a promoure la implementació de la programació i la robòtica en els plans d'estudis educatius per a estudiants de preescolar i primària, i també a nivell vocacional per preparar les persones per al mercat laboral del futur [7.1].

8. Gent i estil de vida

El fet que actualment la majoria de la població tingui un telèfon intel·ligent i/o un ordinador individual es pot considerar un requisit previ per al consum de certes aplicacions, programari i aplicacions que usin alguna forma d'IA [8.1], [8.2]. A la vida quotidiana, les persones i, en particular, els consumidors (la gent normal o els professionals) solen interactuar amb els principals tipus d'IA i aplicacions d'IA, de la següent manera [8.3]:

- **Tasques crucials de la IA.** La IA està relacionada amb el predomini de l'anàlisi intel·lectual, com ara la traducció automàtica, el reenviament i la recuperació automàtica d'informació, la comunicació verbal, la visió artificial i la recopilació de dades.
- **IA basada en eines específiques.** La diferència amb l'anterior és que aquí la IA està dissenyada per poder resoldre una classe més àmplia de problemes, com ara el càlcul evolutiu, el reconeixement de patrons, la programació heurística i l'enfocament multiagent.
- **IA segons el model de pensament desenvolupat.** La IA es caracteritza d'acord amb el model de pensament desenvolupat, com ara la cerca de solucions a l'espai en línia, la presentació del coneixement i l'aprenentatge automàtic.

Dirigit a les activitats diàries de la majoria de les persones, és a dir, els consumidors i, en part, també els desenvolupadors professionals, l'ús de la IA està associat amb certes aplicacions prioritàries [8.3]. Algunes de les implementacions populars d'IA (populars en volum de persones i en estil de vida) estan relacionades amb activitats com:

- **Visió per ordinador.** Aquesta tecnologia processa informació visual per extreure'n coneixement útil. Inclou moltes tasques [8.4]: descobriment de llocs, seguiment d'objectes, reconeixement de patrons, segmentació, estimació de la profunditat i de la distància, etc.
- **Identificació biomètrica.** Són molts i molt variats, com, per exemple: verificació de la identitat persona per la veu, reconeixement facial, identificació de l'iris, anàlisi de la composició química de la suor o de l'olor corporal, anàlisi de microvibracions dels dits i micro- moviments de les mans, anàlisi de freqüència cardíaca i la mida del cor, reconeixement d'empremtes dactilars, anàlisi de l'acció de l'usuari, forma de tecleig, escriptura a manual, anàlisi de la postura, identificació de llavis, prova d'ADN, etc. [8.5].
- **Processament del llenguatge natural, cerca i extracció d'informació de textos.** S'utilitzen per generar textos que són gairebé indistingibles dels humans en estil [8.6].
- **Reconeixement de veu.** És àmpliament utilitzat en centres de trucades, així com en aprenentatge d'idiomes estrangers.
- **Síntesi de veu.** Es pot utilitzar per canviar l'estil de la parla, generar diverses veus a partir d'un model, generar noves veus, modular l'entonació, adaptar-se a la veu del parlant i moltes altres.
- **Visió artificial.** És l'aplicació de la visió artificial a la indústria i la fabricació, per exemple, per comptar objectes en una cinta transportadora, llegir números de sèrie o buscar defectes a la superfície. Les màquines modernes ja reconeixen més del 90% dels objectes, és a dir, no només fixen la presència, sinó que també determinen exactament allò que veuen [8.7]. Un exemple és PowerAI Vision d'IBM [8.8].
- **Traducció automàtica.** Depenent dels idiomes a traduir, l'àrea temàtica i, de fet, com de similars siguin les dades utilitzades per entrenar els models de traducció automàtica, la

qualitat dels resultats dels diferents sistemes pot variar [8.9].

- **Generador de text.** Permet mesurar la qualitat dels models lingüístics, per exemple, mitjançant la probabilitat d'endevinar la paraula següent del context anterior (Perplexity Per Word). Per exemple, Google Brain permet tenir en compte de manera efectiva moltes correlacions remotes [8.9].
- **Sistemes de diàleg (xatbots).** Estan relacionats amb la interacció entre l'humà i un dispositiu (per exemple, automòbils, autobusos, camions, vaixells, etc.). Per finalitat, aquests sistemes es divideixen en tres grups: de finalitat general, adreçats i amb capacitat de diàleg [8.10].
- **Anàlisi d'emocions.** A través d'aquesta activitat, els usuaris poden determinar l'actitud o la reacció emocional del parlant. Exemples d'aplicacions populars són IBM Watson, Meaning Cloud i Salesforce Einstein.

Les perspectives de desenvolupament de la IA estan directament relacionades amb el desenvolupament de la tecnologia informàtica, les TIC, l'electrònica, l'automatització...[8.11] [8.12]. La seva aplicació serà cada cop més visible i formarà part substancial de les compres i el comerç en línia, la salut, el transport, la ciberseguretat i altres [8.13]. La IA deixarà de ser un servei per esdevenir una part important de la vida de les persones. Però també hi ha efectes negatius.

“La majoria de la societat s'ha centrat completament a la IA, s'ignoren altres tendències i fonaments importants en la vida de les persones”.

(expert/a)

El canvi en l'estil de vida actual i futur de les persones es torna cada cop més evident. En general, la IA s'associa a [8.13]:

- Exempció de persones d'activitats rutinàries, reemplaçament o reducció de la intensitat del treball intel·lectual en certes professions fins al reemplaçament total d'especialistes de certes professions amb dispositius intel·ligents;
- Construir un espai digital interactiu de tecnologia de la informació on les persones i les màquines intel·ligents col·laborin;
- Integració total de màquines intel·ligents com a robots en llocs complexos i perillosos per al treball, operacions de rescat i altres;
- Prendre decisions responsables en situacions i processos complexos;
- Augmentar l'eficiència del processament de la informació amb grans volums de dades;
- Millorar la qualitat de l'assistència en una àrea rutinària de la vida quotidiana;
- Millorar la qualitat de l'assistència professional.

Quan parlem d'IA relacionada amb les persones i l'estil de vida, també hem de tenir en compte l'actitud de les persones i les societats sobre aquest tema, cosa que inclou no només els aspectes purament tècnics i pràctics, sinó també psicològics i socials, així com a la zona de confort de l'individu i dels altres [8.14]. Algunes persones tendeixen a ser molt conservadores perquè no se senten còmodes amb la tecnologia que els treu les decisions de les mans.

“Ajuda una persona a la vida quotidiana, però la facilitació conduiria a l'habitació. Si les persones usen IA per a situacions més difícils i si un no té IA, la persona pot entrar en pànic”.

(no expert/a)

És clar que la IA és una tecnologia que evoluciona juntament amb la digitalització [8.15]. Les persones usen IA a la seva vida diària, però s'adonen que aquest procés de digitalització

intel·lectual ha de ser monitoritzat acuradament [8.16]. La IA pot tornar addictes a les persones i fer-les més mandroses, confiant en l'intel·lecte de la IA en lloc del seu intel·lecte natural. Això els facilitarà la vida, però també els privarà de la comunicació social i la proximitat entre les persones.

“El pensament humà limitat i la manca d'autodesenvolupament poden portar a la dependència humana de la presència d'intel·ligència artificial, cosa que condueix a una disminució en el desenvolupament de l'ésser humà”.

(expert/a)

Molts veuen l'impacte de la IA en les persones i el seu estil de vida en l'augment de les capacitats humanes, però alguns predeuen que la creixent dependència de les persones dels sistemes automatitzats socavarà la seva capacitat per pensar de forma independent, actuar i comunicar-se de manera efectiva amb els altres. Les persones poden perdre el sentit dels veritables valors de la vida. La IA atrau i dirigeix l'atenció de les persones en àrees on algú en parla malament, en lloc de dirigir la seva atenció a llocs que fomenten el desenvolupament personal. Fa que les persones siguin més dependents de la publicitat i de les influències externes. Això els separa de la vida en naturalesa.

La IA té un gran potencial per canviar la vida d'una persona i fer-la més productiva, eficient i fàcil. La vida continuarà canviant ràpidament i un ha de ser capaç d'adaptar-se a les noves condicions. Els avenços a IA afectaran el que realment significa ésser humà al segle XXI, però aquests canvis poden ser desiguals.

“Les persones més joves s'adaptaran més fàcilment i ràpidament als sistemes automatitzats, i a les persones grans els resultarà cada cop més difícil mantenir actualitzades les seves habilitats i coneixements per treballar-hi”.

(no expert/a)

Els canvis que portarà la IA tindran efectes positius i negatius a la vida diària de les persones. En resum, podem identificar els riscos i les oportunitats de la manera següent:

Positiu

Millores en l'eficiència del treball humà i l'augment del temps lliure, el benestar i la satisfacció de les persones.
Noves oportunitats i habilitats per millorar l'estil de vida i desenvolupar interessos i talents naturals.
Millores en les capacitats de seguiment i diagnòstic mèdic i proporcionar plans de tractament personalitzats.
Guany de temps i productivitat en el transport autònom.
Més seguretat en l'àmbit de la detecció de delictes.
Assistents virtuals per comprendre i realitzar millor tasques assignades a persones.
Sistemes automatitzats que canviaran l'esbarjo i la vida de les llars.
Millor connexió familiar reduint la tensió mental que produeixen les tasques de la llar.
Millor individualització de la formació basada en les necessitats i capacitats humanes.
Millores en les infraestructures (regulació del trànsit, millores en la cadena de subministrament, etc.).

Negatiu

Reorientació de plantilles, la qual cosa pot portar a la pèrdua de llocs de treball.

Implicacions econòmiques, legals, polítiques i reguladores que afectaran l'estil de vida no sempre en positiu.

Limitacions de l'ús de la IA per qüestions ètiques o legals.

Possibles violacions de la privacitat personal a causa de descontrols en la recopilació massiva de dades.

Invasió de l'espai personal i deteriorament de la privadesa digital.

Augment de la desigualtat socioeconòmica amb la desaparició de milions de llocs de treball de baixa/mitjana qualificació.

En definitiva, la IA planteja la qüestió del concepte que l'individu té d'ell mateix i de la seva llibertat. L'esperança és que la IA tingui un impacte més positiu que negatiu a les persones. La voluntat i la consciència dels éssers humans continuen essent l'origen de les intencions i els resultats. Les màquines es creen per oferir facilitat i eficiència en el camí de la intenció al resultat. La IA s'ha d'utilitzar per millorar les activitats i les experiències humanes, estalviar temps i augmentar el benestar de les persones. S'espera que en el futur la IA afecti encara més la vida quotidiana de les persones i una afectació més o menys positiva dependrà en gran mesura del control de les persones que dissenyen/manegen/gestionen els sistemes basats en la IA, així com de la seva convenient aplicació a la societat.

“Tan perillosa és l'estupidesa humana, que limita les funcions del cervell, com la inacció”.

(expert/a)

9. Conclusions

La IA va arribar per quedar-se a la nostra vida quotidiana. Aquesta és una conclusió òbvia i l'hem d'assumir. Aleshores, el tema és, com podem conviure amb aquest fet?

En aquest document, hem revisat moltes fonts bibliogràfiques i comentat el tema directament amb experts i no experts per comprendre millor la problemàtica, les preocupacions, les oportunitats que la IA pot implicar per/brindar a la humanitat i els seu impacte en la nostra societat. Arrel d'això, a continuació, proporcionem el conjunt de conclusions i pautes que n'hem extret.

En primer lloc, resumirem els nostres resultats en els quatre àmbits que s'han treballat:

Negocis: La IA es considera té un gran potencial pels negocis, especialment quan es tracta d'estalviar temps, facilitar decisions més ràpides, oferir solucions innovadores, augmentar la satisfacció del client, la productivitat, l'optimització dels costos i maximitzar les vendes, alhora que redueix la probabilitat d'error humà.

Els **principals riscos** en aquest àmbit estan relacionats amb trobar un equilibri entre els interessos econòmics de les empreses i la no vulneració dels drets dels ciutadans en matèria de privadesa i llibertat individual.

Governança: Hi ha un acord global sobre un conjunt bàsic de sis principis pel que fa al desenvolupament de la IA que són funcionalment independents dels algorismes, de la tecnologia i de l'àmbit, per assolir un bon compromís entre les estratègies i els objectius de l'empresa, els requisits legals i l'ètica, que són: responsabilitat, transparència, equitat, seguretat, control humà i universalitat.

Els **riscos principals** de la governança de la IA estan relacionats amb l'incompliment d'un o més d'aquests principis, que donen lloc a resultats incorrectes, biaixos, manca de llibertat, iniquitats, etc.

Habilitats i competències: La necessitat d'habilitats físiques, cognitives bàsiques i manuals es reduirà a causa de la presa de control per part de la IA, mentre que les competències digitals, el pensament crític, el treball en equip, l'empatia, les habilitats cognitives superiors, les habilitats tecnològiques, socials i emocionals tindran una major demanda. Els sistemes educatius necessitaran un canvi per reflectir aquests canvis i millorar les habilitats dels estudiants en aquest sentit.

El **risc més gran** és que els països "menys desenvolupats" corren el perill de quedar-se al marge de tot això i encara més endarrerits.

Persones i estil de vida: La IA hauria de contribuir a fer que la vida de les persones sigui més productiva, eficient, segura i fàcil. Podrem tenir assistència personalitzada de seguiment i diagnòstic mèdic, més temps lliure, possibilitat de desenvolupar els nostres interessos i talents naturals, una infraestructura millor i més ràpida, guany de temps i productivitat, una vida més segura i una individualització més gran.

El **risc més gran** en aquest sentit rau en el fet de que la IA pot estar esbiaixada i perpetuar o fins i tot augmentar la disparitat i iniquitat de gènere i raça.

En segon lloc, la nostra recerca en IA ens permet identificar que aquesta es concep de dues maneres diferents, restrictiva i disruptiva. Dues visions que apareixen a totes les nostres fonts, cosa que fa que el focus estigui en diferents temes i propostes. Tot i que aquestes dues visions es poden entendre com a contradictòries, en realitat són visions complementàries que permeten captar la complexitat d'opinions, inquietuds i propostes al voltant de l'ús dels sistemes d'IA.

Visió restrictiva: La IA com una tecnologia més i, per tant, s'ha de tractar com qualsevol altra tecnologia.

Visió disruptiva: La IA com una tecnologia diferent, que marca un abans i un després a la societat humana i la seva relació amb altres tecnologies.

Des d'un punt de vista restrictiu, la IA tendeix a veure's com una oportunitat: poder crear nous llocs de treball molt qualificats i remunerats, obrir nous mercats i oportunitats de negoci, fer la vida més fàcil i saludable, la burocràcia més ràpida, etc. La idea més comuna és que la IA és útil en els processos de presa de decisions. Els desafiaments que poden sorgir d'aquests processos es consideren limitats i possiblement superables simplement traçant una línia clara entre quan es pot fer servir la IA i quan no. En aquest sentit, es creu que la IA és particularment eficaç per a la gestió i l'anàlisi de dades, així com per a la generació d'informació per a la presa i l'avaluació de decisions, però no per a la presa de decisions automatitzada. En aquest sentit, es pensa que les decisions que tenen un impacte directe a les persones han de ser preses per persones.

En canvi, des d'una visió disruptiva de la IA, la impressió és que a les societats contemporànies qualsevol àmbit integra o integrarà la IA. Es pensa que, si bé no volem que la IA participi en molts aspectes de la nostra vida diària, és vital analitzar-ne els costos i beneficis, valorant el que passaria si els sistemes d'IA prenguessin decisions incorrectes. Això pot afectar diverses àrees del nostre entorn i és en aquestes àrees específiques on cal avaluar els riscos d'usar sistemes d'IA. La justícia, la privadesa de les persones (és a dir, la llibertat), els propis algorismes de disseny dels sistemes d'IA (incloent el biaix de les dades), la biomedicina, les finances són només alguns exemples de les àrees identificades.

Una preocupació compartida, majoritàriament associada a una visió restrictiva de la IA i que apareix tant explícitament com implícita a totes les fonts, té a veure amb la relació dels sistemes d'IA amb els imaginaris de la ciència ficció o amb la idea que la IA pot resoldre qualsevol tipus de problemes. S'han desenvolupat nombroses aplicacions en el camp de la IA i es poden fer servir en molts camps, però hi ha una bretxa significativa entre les funcionalitats i capacitats tècniques actuals i la narrativa del que la IA podria fer en el futur. Aquesta narrativa al voltant de la IA, que no es correspon amb els desenvolupaments actuals, considera dos tipus d'efectes negatius: i) la dificultat d'articular un debat públic provat sobre la rendició de comptes quan s'utilitzen formes d'IA en els processos de presa de decisions; ii) l'aparició d'una sèrie d'imaginariis catastròfics que generen reticències cap a la IA entre l'opinió pública i la ciutadania.

Val la pena esmentar també altres temes no tractats en aquest document. Un d'ells, probablement el més important, és l'impacte de la IA al nostre entorn ambiental i com pot ajudar al canvi climàtic. De fet, no hem aconseguit aprofundir prou en aquest tema als nostres grups focals perquè tampoc no va ser plantejat pels participants, cosa que pot indicar que el medi ambient encara no es percep com una preocupació prioritària. Tot i això, la IA és un important consumidor d'energia atesa la complexitat de l'aprenentatge automàtic i la inferència sobre grans volums de dades (Big Data), a més del fet que tot l'ecosistema de les TIC ja és un dels principals contribuents a les emissions de gasos d'efecte hivernacle (GEI) [9.1]. D'altra banda, la IA també s'ha presentat com a solució al canvi climàtic a causa de les seves capacitats multifuncionals, que inclouen el seguiment i la reducció d'emissions, permetent models econòmics creatius per ajudar el medi ambient i millorant la resiliència climàtica. Per exemple, un estudi encarregat per Microsoft [9.2] va concloure que l'ús d'IA per a aplicacions mediambientals té el potencial d'impulsar el PIB mundial entre un 3,1 i un 4,4% i, alhora, reduir les emissions globals de GEI entre un 1,5 i un 4,0 % pel 2030 al sector dels negocis, fins a 2,2% en energia i fins a 1,7% en transport.

Una opinió comuna és que les habilitats socials i intel·lectuals de les persones, com la creativitat, l'empatia, el treball en equip i la innovació, no poden ser substituïdes per la IA. A curt termini, la

visió és que hi haurà un resorgiment de les activitats artístiques. Això no obstant, aquesta visió sembla massa optimista: la IA ja és capaç de compondre simfonies, pintar quadres, escriure poemes, cançons i històries, i participar en jocs. Alguns països com Austràlia [9.3] ja van acceptar que un sistema basat en la IA es pot registrar com a inventor en una patent. En el futur, és probable que aquestes capacitats s'explorin encara més.

Una opinió comuna també és que la IA no ha de reemplaçar la capacitat humana i cal evitar que la llibertat humana per prendre decisions sigui influenciada per les eines impulsades per la IA. En particular, l'ètica és un problema recurrent plantejat a totes les nostres fonts. Fins i tot els experts que hem entrevistat consideren que no tenen prou coneixements per poder decidir sobre qüestions ètiques i socials, coneixements que s'haurien d'integrar de manera interdisciplinària.

Per evitar aquest tipus de narratives i els seus efectes negatius, es requereixen accions relacionades amb la informació i la participació ciutadana:

Informació: Assegurar que els mitjans de comunicació informin de manera ètica i honesta quan es parla de sistemes d'IA, que permeti una distinció clara entre visions futuristes especulatives, i desenvolupaments i possibilitats actuals. Cal desenvolupar una tasca pedagògica que permeti a la ciutadania conèixer com funciona la IA i quines aplicacions s'estan fent servir. Aquest coneixement pot augmentar la confiança de les persones en la precisió i la idoneïtat dels sistemes d'IA.

Participació: Involucrar la ciutadania en l'establiment de prioritats per al desenvolupament de la IA. Aquest es considera el valor afegit de l'estratègia europea de desenvolupament de la IA, davant d'altres estratègies que poden estar tecnològicament més avançades, com ara la de la Xina o la dels Estats Units. Es considera que l'estratègia europea pot incorporar com a valor afegit a la seva IA la integració dels ciutadans a establir àrees prioritàries on desenvolupar-la i/o aplicar-la.

En conclusió, la IA és una tecnologia que en el seu disseny i desenvolupament està tan allunyada de la vida quotidiana que els experts creuen que la ciutadania no està prou capacitada per prendre decisions sobre com fer servir la IA. Tot i això, al mateix temps, es considera necessari que els ciutadans prenguin decisions i decideixin el rumb de la IA. Per això, assenyalem la necessitat de formar els ciutadans en el funcionament, potencial i possibles efectes de la IA. Per tant, hem de proporcionar ensenyaments, cursos i capacitació a escoles i instituts d'educació superior per facilitar l'ús i l'acceptació de la IA per als joves i les futures generacions. De fet, aquest és el principal objectiu del projecte HEDY: Oferir un curs massiu obert en línia (MOOC) amb l'objectiu d'arribar a l'audiència d'educació superior i mostra la capacitat de la IA, l'oportunitat que té la nostra societat en aquest moment de canviar el nostre entorn per un de millor, però també els riscos a què ens enfrontem des de diferents punts de vista. Aquest material es complementarà amb el Rootkit: un conjunt de eines multimèdia de suport amb la capacitat de tenir un impacte més immediat i visual per a l'audiència.

A. Referències

- [1.1] HEDY project, *Life in the AI era*, KA220-HED 0C8D3623 - Cooperation partnerships in higher education, <https://lifeintheaiera.eu> , accessed on October 2022.
- [1.2] European Commission, *Artificial intelligence for Europe*, COM(2018) 237 final, Brussels, April 2018.
- [1.3] European Commission, *White paper on artificial intelligence - A European approach to excellence and trust*, COM(2020) 65 final, Brussels, February 2020.
- [1.4] UNESCO, *Recommendation on the ethics of artificial intelligence*, SHS/BIO/REC-AIETHICS, 2021, <https://unesdoc.unesco.org/ark:/48223/pf0000380455>, accessed on April 2022
- [2.1] Henning Kagermann, Wolf-Dieter Lukas, Wolfgang Wahlster, "Industrie 4.0: mit dem Internet der dinge auf dem weg zur 4. industriellen Revolution", VDI nachrichten, April, 2011, <https://web.archive.org/web/20130304101009/http://www.vdi-nachrichten.com/artikel/Industrie-4-0-Mit-dem-Internet-der-Dinge-auf-dem-Weg-zur-4-industriellen-Revolution/52570/1>, accessed on April 2022.
- [2.2] Thomas Philbeck, Nicholas Davis, "The fourth industrial revolution", *Journal of International Affairs*, vol. 72, no. 1, pp. 17–22, 2018.
- [3.1] Alan M. Turing, "On computable numbers, with an application to the entscheidungsproblem", in *Proceedings of the London Mathematical Society*, vol. s2-42, no, 1, pp. 230-265, November-December 1936.
- [3.2] Alan M. Turing, "Computing machinery and intelligence", *Mind*, vol. LIX, no. 236, pp. 433–460, October 1950.
- [3.3] John McCarthy, "What is artificial intelligence?", *Project JMC*, November 2007, <http://jmc.stanford.edu/articles/whatisai.html>, accessed on October 2022.
- [3.4] Encyclopedia Britannica, *Definition of artificial intelligence*, last update August 24, 2022, <https://www.britannica.com/technology/artificial-intelligence>, accessed on October 2022.
- [3.5] Cambridge Dictionary, *Definition of artificial intelligence*, Cambridge university press, <https://dictionary.cambridge.org/dictionary/english/artificial-intelligence>, accessed on October 2022.
- [3.6] Max, Tegmark, *Life 3.0: being human in the age of artificial intelligence*, Knopf, 1st edition, August 2017.
- [3.7] Csaba Kollár, "A mesterséges intelligencia és a kapcsolódó technológiák bemutatása a biztonságtudomány fókuszában", *Kiberbiztonság/Cybersecurity*, vol. 2, pp. 47-61, edited by Zoltán Rajnai, Doctoral School of Security Sciences, Budapest, Hungary, 2019.
- [3.8] Shelly Fan, *Will AI replace us: a primer for the 21st century*, The big idea series, Thames and Hudson, 1st edition, April 2019.
- [3.9] Council of Europe, *History of artificial intelligence*, December 2018, <https://www.coe.int/en/web/artificial-intelligence/history-of-ai>, accessed on October 2022.
- [3.10] Rebecca Reynoso, *A complete history of artificial intelligence*, G2 report, May 2021, <https://www.g2.com/articles/history-of-artificial-intelligence>, accessed on October 2022.

- [4.1] Csaba Kollár, “A mesterséges intelligencia kapcsolata a humán biztonsággal”, *Nemzetbiztonsági Szemle*, vol. 6, no. 1, 2018.
- [4.2] Peter Stone et al., *Artificial intelligence and life in 2030: one hundred year study on artificial intelligence*, Report of the 2015-2016 Study Panel, September 2016, <https://ai100.stanford.edu/2016-report>, accessed on October 2022.
- [4.3] Kathe Pelletier et al., *2021 EDUCAUSE horizon report - teaching and learning edition*, Educause publications, Horizon report, April 2021, <https://library.educause.edu/resources/2021/4/2021-educause-horizon-report-teaching-and-learning-edition>, accessed on October 2022.
- [4.4] OECD, *Recommendation of the council on artificial intelligence*, OECD/LEGAL/0449, adopted on 22/05/2019, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>, accessed on October 2022.
- [4.5] *An open letter to the United Nations convention on certain conventional weapons*, July 2015, <https://www.cse.unsw.edu.au/~tw/ciair/open.pdf>, accessed on October 2022.
- [4.6] United Nations, *193 countries adopt first-ever global agreement on the Ethics of Artificial Intelligence*, UN news, November 2021, <https://news.un.org/en/story/2021/11/1106612>, accessed on October 2022.
- [4.7] Omar Hatamleh, George Tilesh, *BetweenBrains: Taking back our AI future*, GTPublishDrive, May 2020.
- [5.1] Sandra Maria Correia Loureiro, João Guerreiro, Iis Tussyadiah, “Artificial intelligence in business: State of the art and future research agenda”, *Journal of Business Research*, vol. 129, pp. 911–926, May 2021.
- [5.2] James Eager et al., *Opportunities of artificial intelligence*, Study requested by the ITRE committee, European Parliament, June 2020, [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/652713/IPOL_STU\(2020\)652713_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/652713/IPOL_STU(2020)652713_EN.pdf), accessed on October 2022.
- [5.3] Laurent Probst et al., *EU businesses go digital: Opportunities, outcomes and uptake*, Digital transformation scoreboard 2018, European Commission, March 2019.
- [5.4] Neha Soni, Enakshi Sharma, Narotam Singh, Amita Kapoor, “Impact of artificial intelligence on businesses: from research, innovation, market deployment to future shifts in business models”, *arXiv.org*, May 2019, <http://arxiv.org/abs/1905.02092>, accessed on October 2022.
- [5.5] Awishkar Ghimire, Surendrabikram Thapa, Avinash Kumar Jha, Surabhi Adhikari, Ankit Kumar, “Accelerating business growth with big data and artificial intelligence”, in *Proceedings of the Fourth International Conference on IoT in Social, Mobile, Analytics and Cloud (I-SMAC 2020)*, Palladam, India, October 2020.
- [5.6] Mia Hoffmann Laura Nurski, *What is holding back artificial intelligence adoption in Europe?*, Bruegel policy contribution issue 24/21, November 2021, <https://www.bruegel.org/wp-content/uploads/2021/11/PC-24-261121.pdf>, accessed on October 2022
- [5.7] Fotis Kitsios, Maria Kamariotou, “Artificial intelligence and business strategy towards digital transformation: a research agenda”, *MDPI Sustainability*, vol. 13(4), 2025, February 2021.
- [5.8] Denise Carter, “How real is the impact of artificial intelligence? The business information survey 2018”, *Business Information Review*, vol. 35, no. 3, pp. 99-115, July 2018.

- [5.9] Bernd Carsten Stahl, *Ethical Issues of AI*, Artificial Intelligence for a Better Future, pp. 35-53, Springer, March 2021.
- [5.10] Jacques Bughin, Jeongmin Seong, James Manyika, Lari Hämäläinen, Eckart Windhagen, Eric Hazan, *Notes from the AI frontier tackling Europe's gap in digital and AI*, Discussion paper, Mckinsey Global Institute, February 2019.
- [5.11] Dimple Agarwal, Josh Bersin, Gaurav Lahiri, Jeff Schwartz, Erica Volini, *The rise of the social enterprise*, 2018 Deloitte global human capital trends, University Press, 2018, https://www2.deloitte.com/content/dam/insights/us/articles/HCTrends2018/2018-HCTrends_Rise-of-the-social-enterprise.pdf, accessed on October 2022.
- [5.12] İzzet Kılınc, Aslihan Ünal, "AI is the new black: effects of artificial intelligence on business world", *Journal of Contemporary Administrative Science*, vol. 2, no. 6, pp. 238-258, September 2019.
- [6.1] Melanie Malluk Batley, *AI adoption accelerated during the pandemic but many say it's moving too fast: KPMG survey*, Thriving in an AI World, KPMG study, March 2021, <https://info.kpmg.us/news-perspectives/technology-innovation/thriving-in-an-ai-world/ai-adoption-accelerated-during-pandemic.html>, accessed on October 2022.
- [6.2] BBC news, Alexa tells 10-year-old girl to touch live plug with penny, December 28, 2021, <https://www.bbc.com/news/technology-59810383>, Accessed on August 2022.
- [6.3] Marc Hufty, "Investigating policy processes: the governance analytical framework", *Research for Sustainable Development: Foundations, Experiences, and Perspectives*, pp. 403-424, edited by U. Wiesmann, et al., Geographica Bernensia, Bern, Switzerland, 2011.
- [6.4] Anneke Zuiderwijk, Yu-Che Chen, Fadi Salem, "Implications of the use of artificial intelligence in public governance: a systematic literature review and a research agenda", *Government Information Quarterly*, vol. 38, no. 3, July 2021.
- [6.5] *Perspectives on Issues in AI Governance*, Google report, <https://ai.google/static/documents/perspectives-on-issues-in-ai-governance.pdf>, accessed on October 2022.
- [6.6] James Butcher, Irakli Beridze, "What is the state of artificial intelligence governance globally?", *The RUSI Journal*, vol. 164, n. 5-6, pp. 88-96, November 2019.
- [6.7] Johannes Schneider, Rene Abraham, Christian Meske, Jan vom Brocke, "AI governance for businesses", *arXiv.org*, November 2020, <https://doi.org/10.48550/arXiv.2011.10672>, accessed on October 2022.
- [6.8] Alan F.T. Winfield, Marina Jirotko, "Ethical governance is essential to building trust in robotics and artificial intelligence systems", *Philosophical Transactions of the Royal Society A*, vol. 376, no. 2133, 20180085, November 2018.
- [6.9] Matti Mäntymäki, Matti Minkkinen, Teemu Birkstedt, Mika Viljanen, "Defining organizational AI governance", *AI and Ethics*, February 2022.
- [6.10] KOSA AI, *The importance of AI governance and 5 key principles for its guidance*, <https://kosa-ai.medium.com/the-importance-of-ai-governance-and-5-key-principles-for-its-guidance-219798c8f407>, accessed on October 2022.
- [6.11] Kevin Hoff, Masooda Bashir, "Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust", *Human Factors - The Journal of the Human Factors and Ergonomics Society*, vol. 57, no. 3, pp. 407-434, May 2015.
- [6.12] Sandra Wachter, Brent Mittelstadt, Luciano Floridi, "Why a right to explanation of automated decision-making does not exist in the general data protection regulation", *International Data Privacy Law*, vol. 7, no. 2, pp. 76-99, May 2017.

- [6.13] European Commission, *Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts*, COM(2021) 206 final, Brussels, April 2021.
- [6.14] US Congress, *Algorithmic Accountability Act of 2019*, H.R.2231, 116th Congress, April 2019, <https://www.congress.gov/bill/116th-congress/house-bill/2231>, accessed on August 2022.
- [6.15] Jessica Fjeld, Nele Achten, Hannah Hilligoss, Adam Nagy, Madhulika Srikumar, *Principled artificial intelligence: mapping consensus in ethical and rights-based approaches to principles for AI*, Berkman Klein Center Research Publication No. 2020-1, February 2020.
- [6.16] Martin Sokalski, *The shape of AI governance to come*, KPMG Insights, December 2020, <https://assets.kpmg/content/dam/kpmg/xx/pdf/2021/01/the-shape-of-ai-governance-to-come.pdf>, accessed on April 2022.
- [7.1] Kelly Shiohira, *Understanding the impact of artificial intelligence on skills development*, Education 2030, UNESCO-UNEVOC, March 2021.
- [7.2] Diego Rubio, Carlos Lastra, *Mapping European attitudes to technological change and its governance*, Center for the Governance of Change, European Tech Insights, 2019, <https://docs.ie.edu/cgc/European-Tech-Insights-2019.pdf>, accessed on October 2022.
- [7.3] Carlos Lastra, Diego Rubio, *Unveiling the technological future that citizens want and their concerns in a changing world*, Center for the Governance of Change, European Tech Insights, 2020, <https://docs.ie.edu/cgc/CGC-European-Tech-Insights-2020.pdf>, accessed on October 2022.
- [7.4] Oscar Jonsson, Carlos Luca de Tena, *Part I How the pandemic altered our relationship with technology*, Center for the Governance of Change, European Tech Insights, 2021, <https://docs.ie.edu/cgc/IE-CGC-European-Tech-Insights-2021.pdf>, accessed on October 2022.
- [7.5] Chris Howard, Andy Rowsell-Jones, *2019 CIO survey: CIOs have awoken to the importance of AI*, Gartner Research, January 2019.
- [7.6] Jenny Burke, Michael Coovert, Robin R. Murphy, Jennifer Riley, Erika Rogers, "Human-robot factors: robots in the workplace", in *Proceedings of the human factors and ergonomics society annual meeting*, vol. 50, no. 9, pp. 870-874, October 2006.
- [7.7] Albert Ellis, Steve Bates, *A changing perspective*, Harvey Nash/KPMG CIO Survey 2019, https://assets.kpmg/content/dam/kpmg/kz/pdf/2019/09/CIO-Survey_2019_ENG.pdf, accessed on October 2022.
- [7.8] H. James Wilson, Paul R. Daugherty, "Collaborative intelligence: Humans and AI are joining forces", *Harvard Business Review*, vol. 96, no. 4, pp. 115–123, July/August 2018.
- [7.9] Eduard Anton, Alina Behne, Frank Teuteberg, "The humans behind artificial intelligence – an operationalisation of AI competencies", in *Proceedings of the 28th European Conference on Information Systems (ECIS2020)*, June 2020.
- [7.10] David Chrisinger, "The solution lies in education: artificial intelligence & the skills gap", *On the Horizon*, vol. 27, no. 1, pp. 1–4, March 2019
- [7.11] Ulrich Paschen, Christine Pitt, Jan Kietzmann, "Artificial intelligence: building blocks and an innovation typology", *Business Horizons*, vol. 63, no. 2, pp. 147–155, November 2019.
- [7.12] Jean F. Byrd, Fabien Pampaloni, Linda Wilson, "Hybrid simulation", *Human simulation for nursing and health professions*, pp. 267–271, edited by L Wilson, L. Rockstraw, Springer, New York, USA, 2012.

- [7.13] Carl Benedikt Frey, *Technology trap: capital, labor, and power in the age of automation*, Princeton University Press, June 2019.
- [7.14] Francesc Pedró, Miguel Subosa, Axel Rivas, Paula Valverde, *Artificial intelligence in education: challenges and opportunities for sustainable development*, Education 2030, UNESCO, Paris, France, 2019.
- [7.15] Julian McDougall, "Media Literacy versus Fake News: Critical thinking, resilience and civic engagement", *Medijske studije*, vol. 10, no. 19, pp. 29–45, October 2019.
- [7.16] Lisa French, Mark Poole, *New competencies for media and communication in an AI era*, Humanistic futures of learning - perspectives from UNESCO chairs and UNITWIN networks, UNSCO, pp. 136–140, 2020.
- [7.17] Marco Dondi, Julia Klier, Frédéric Panier, Jörg Schubert, "McKinsey: these are the skills you will need for the future of work", *World Economic Forum*, June 2021, <https://www.weforum.org/agenda/2021/06/defining-the-skills-citizens-will-need-in-the-future-world-of-work/>, accessed on October 2022.
- [7.18] Giselle Rampersad, "Robot will take your job: Innovation for an era of artificial intelligence", *Journal of Business Research*, vol. 116, pp. 68–74, August 2020.
- [8.1] Wadzani A. Gadzama, Joseph Bitrus, Ngubdo Maigana A, "Global smartphone ownership, Internet usage and their impacts on humans", *Journal of Communications Networks*, vol. 1, no. 1, October 2017.
- [8.2] Noshir Kaka, Anu Madgavkar, Alok Kshirsagar, Rajat Gupta, James Manyika, Kushe Bahl, *Digital India: Technology to transform a connected nation*, McKinsey Global Institute, March 2019, <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/digital-india-technology-to-transform-a-connected-nation>, accessed on October 2022.
- [8.3] Moscow Industrial Development Agency, *Technologies of Artificial Intelligence*, 2019, <https://apr.moscow/content/data/6/11%20Технологии%20искусственного%20интеллекта.pdf>, accessed on October 2022.
- [8.4] Maxim Akimov, Herman Gref, Alexander Vedyakhin, *Analytical review of the global robotics market*, Sberbank Robotics Laboratory, Russia, 2019.
- [8.5] Tadviser, *Biometric identification technologies*, Russian analytical agency, November 2019, https://tadviser.com/index.php/Article:Biometric_identification_technologies, accessed on October 2022.
- [8.6] Bohdan Balov, *Convolutional neural networks from scratch*, Medium.com, February 2019, <https://medium.com/@balovbohdan/сверточные-нейронные-сети-с-нуля-4d5a1f0f87ec>, accessed on October 2022.
- [8.7] Center2M, *What is machine vision and how is it different from human vision? Now let's make it clear!*, Meduza.io, March 2019, <https://meduza.io/feature/2019/03/30/что-такое-машинное-зрение-и-чем-оно-отличается-от-человеческого-сейчас-об-этом-рассказывает>, accessed on October 2022.
- [8.8] IBM PowerAI Vision, <https://www.ibm.com/docs/en/mvi/1.1.0?topic=overview>, accessed on October 2022.
- [8.9] Igor Pivovarov (editor), *The Artificial Intelligence almanac*, Analytical Collection No. 2, Center of the National Technology Initiative, MIPT, Russia, September 2019.
- [8.10] Pierrick Milhorat, Stephan Schlögl, Gerard Chollet, Jérôme Boudy "Multi-step natural language understanding", in *Proceedings of 14th Annual Meeting of the Special Interest Group on Discourse and Dialogue* (SIGdial 2013), Metz, France, August 2013.

- [8.11] Boris Mirkin, *Artificial intelligence: history and current state*, Report at the second Nizhny Novgorod festival of science and art, November 2010, <https://www.hse.ru/data/2010/12/05/1209601907/MachineIntel.pdf>, accessed on October 2022.
- [8.12] Vamsi Vedantam, *Artificial intelligence in information and cyber security*, Tech Mahindra, January 2021, https://www.researchgate.net/publication/349350306_Artificial_Intelligence_in_Information_and_Cyber_Security, accessed on October 2022.
- [8.13] Chenzhuoer Li, Runjie Pan, Huiyu Xin, Zhiwen Deng, "Research on artificial intelligence customer service on consumer attitude and its impact during online shopping", in *Proceedings of 5th Annual International Conference on Information System and Artificial Intelligence (ISAI2020)*, Zhejiang, China, May 2020.
- [8.14] Vyacheslav Ovchinnikov, *Doroga v mir iskusstvennogo intellekta* [Road to the World of Artificial Intelligence], Institute of Economic Strategies, Moscow, Russia, 2017.
- [8.15] Maria José Sousa, Gabriel Osório de Barros, Nuno Tavares, "Artificial intelligence a driver for digital transformation", *Digital transformation and challenges to data security and privacy*, Chapter 14, pp. 234-251, edited by Pedro Fernandes Anunciação, Cláudio Roberto Magalhães Pessoa, George Leal Jamil, IGI Global, February 2021.
- [8.16] Philip Boucher, *Artificial intelligence: How does it work, why does it matter, and what can we do about it?*, Directorate-General for Parliamentary Research Services, European Parliament, June 2020, <https://data.europa.eu/doi/10.2861/44572>, accessed on October 2022.
- [9.1] Charlotte Freitag, Mike Berners-Lee, Kelly Widdicks, Bran Knowles, Gordon S. Blair, Adrian Friday, "The real climate and transformative impact of ICT: A critique of estimates, trends, and regulations", *Patterns review*, vol. 2, no. 9, September 2021.
- [9.2] Celine Herweijer, Benjamin Combes, Jonathan Gillham, *How AI can enable a sustainable future*, Microsoft and PwC report, April 2019, <https://www.pwc.co.uk/sustainability-climate-change/assets/pdf/how-ai-can-enable-a-sustainable-future.pdf>, accessed on October 2022.
- [9.3] Rebecca Currey, Jane Owen, "In the courts: Australian court finds AI systems can be inventors", *World Intellectual Property Organization magazine*, September 2021, https://www.wipo.int/wipo_magazine/en/2021/03/article_0006.html, accessed on October 2022.
- [A.1] Richard A. Krueger, Mary Anne Casey, *Focus groups: a practical guide for applied research*, Newbury Park, Sage Publications, August 2014.
- [A.2] Virginia Braun, Victoria Clarke, "Using thematic analysis in psychology", *Qualitative Research in Psychology*, vol. 3, no. 2, pp. 77-101, January 2006.

B. Material adicional

Metodologia

Recopilem informació de dues fonts diferents per construir els nostres fonaments: 1) Estat de l'art disponible a la literatura i 2) Grups focals.

La nostra primera font d'informació va provenir de la bibliografia disponible sobre IA, en general i sobre els quatre àmbits descrits anteriorment, en particular. L'objectiu era adquirir un coneixement profund del tema i comprendre'n el coneixement actual. Això ens va permetre: i) identificar teories, mètodes i opinions rellevants a l'estat de l'art existent i reportar-los en aquest Llibret; i ii) organitzar i guiar els participants dels grups focals amb aquestes bases ja apreses per obtenir la informació complementària que necessitàvem.

Per això busquem, llegim i avaluem més de 250 documents, entre treballs de recerca, articles de revistes, blocs d'experts, informes d'empreses, directrius d'agències, etc. per obtenir el coneixement adequat sobre els desafiaments i oportunitats de la IA, les aplicacions i els impactes esperats a curt i llarg termini. Aquests documents han estat seleccionats amb cura per sintetitzar l'estat de l'art, segons el seu impacte (p. e., nombre de cites, difusió en notícies i xarxes socials, etc.), autor/s, editorial i any de publicació.

La segona font va consistir a recopilar informació mitjançant entrevistes a persones a través dels grups focals amb experts i no experts a IA a 5 països europeus diferents. Un grup focal és un tipus de tècnica qualitativa de recopilació de dades, en què un grup de persones, guiades per un moderador, conversen i discuteixen al voltant d'un tema comú. Habitualment consisteix en un grup de 5 a 10 persones que no es coneixen entre si. Aquests individus es trien perquè tenen trets particulars que són rellevants pel tema del grup. En crear un entorn tolerant i obert que fomenti moltes perspectives i punts de vista, sense pressionar els membres per assolir un consens, el moderador utilitza el grup i la seva interacció per aprendre més sobre un tema en particular [A.1].

Durant el febrer del 2022, preparem i duem a terme dos grups de discussió diferents amb una metodologia i organització comuna a cadascun dels països dels socis del projecte: un grup focal només amb experts en IA i un altre grup de discussió només amb no experts en IA. El projecte defineix el terme expert com una persona amb títol universitari, treballant durant almenys 5 anys a l'àrea d'IA, societat digital, interacció humà robot o Indústria 4.0, i almenys 3 articles científics o professionals publicats. El nombre de participants per a cada grup es va fixar en un mínim de 5. A tots els grups hi va haver un moderador i un assistent: el primer dirigia i harmonitzava les discussions, mentre que el segon prenia notes. Cada grup de discussió va tenir una durada aproximada d'una hora i les sessions van ser enregistrades i després transcrites. En un cas no va ser possible organitzar el grup focal per problemes d'horaris relacionats amb la pandèmia, i vam substituir el grup de discussió per un qüestionari. El qüestionari es va crear amb preguntes similars a les utilitzades a la resta de grups de discussió. En resum, es van organitzar nou (9) grups focals i un qüestionari, i es van analitzar els resultats seguint les directrius disponibles a [A.2].

S'hi afegixen més detalls sobre totes les sessions:

- 1) Dos grups a Budapest (Hongria) sobre els desafiaments i oportunitats de la IA
 - 8 experts (7 homes, 1 dona) de 28 a 61 anys;
 - 5 no experts (2 homes, 3 dones) de 20 a 28 anys, estudiants universitaris matriculats a un master.

- 2) Dos grups a Münster (Alemanya) sobre IA als negocis
 - 7 experts (7 homes) de 26 a 50 anys;
 - 7 no experts (3 homes, 4 dones) de 22 a 30 anys, estudiants universitaris matriculats a un màster
- 3) Dos grups a Barcelona (Espanya) sobre IA a la governança
 - 9 experts (7 homes, 2 dones) de 35 a 70 anys;
 - 10 no experts (7 homes, 3 dones) de 22 a 70 anys, de la societat civil sense coneixements previs sobre IA.
- 4) Enfoc mixte a Lisboa (Portugal) sobre IA en habilitats i competències
 - 9 experts (6 homes, 3 dones) de 26 a 67 anys, a través d'un qüestionari en línia;
 - 5 no experts (3 homes, 2 dones) de 23 a 55 anys, estudiants de màster o recentment graduats a través d'un grup de discussió.
- 5) Dos grupos a Varna (Bulgària) sobre IA en persones i estil de vida
 - 6 experts (4 homes, 2 dones) de 29 a 59 anys;
 - 15 no experts (4 homes, 11 dones) de 20 a 24 anys, estudiants matriculats a una carrera universitària. en una carrera universitària.

Limitacions

És clar que aquesta metodologia té algunes limitacions. En primer lloc, és una anàlisi les conclusions de la qual permeten identificar les diferents interpretacions i arguments socialment disponibles sobre un tema, però, a diferència de l'anàlisi quantitativa, les seves conclusions no són representatives, només significatives. A més, hi ha la limitació de l'heterogeneïtat dels grups de discussió, ja que la majoria dels experts eren acadèmics i els no experts tenien títol universitari. Per tant, els resultats no poden representar els punts de vista de la població general sobre el tema. Tanmateix, val la pena esmentar que contrastem les opinions de les persones amb la literatura disponible i viceversa, per la qual cosa els nostres resultats tenen certa validesa i és probable que altres feines similars arribin a les mateixes conclusions.

Esborrany, accés a la versió més llarga

Aquest Llibret és el resultat final de gairebé un any de treball investigant l'impacte de la IA a la nostra societat. Per assolir aquest objectiu, vam produir durant aquest any molts altres treballs intermedis que finalment han estat harmonitzats i resumits en aquest document. Els lectors interessats poden voler obtenir accés a la primera versió més llarga del document, que està disponible al repositori de Zenodo (<https://www.zenodo.org>) amb DOI: 10.5281/zenodo.7243312.

Agraïments

Agraïm a totes les persones que ens han ajudat a preparar aquest Llibret i que generosament i obertament van participar en les nostres activitats.

Les icones al Capítol 1 han estat fetes per Chanut-is-Industries de www.flaticon.com.

La realització d'aquest Llibret ha estat finançat per la Unió Erasmus. Tot i això, els punts de vista i les opinions que s'hi expressen són responsabilitat únicament dels autors i no reflecteixen necessàriament les de la EACEA (Agència Executiva Europea d'Educació i Cultura) de la Unió Europea.

Autors (ordre alfabètic)



Davide Careglio received the M.Sc. double degree in telecom. engineering and electrical engineering from Technical University of Catalonia (UPC) (2000) and Politecnico di Torino (2001), respectively, and his PhD from UPC in 2005. He is currently a member of the Intelligent Data Science and Artificial Intelligence research centre (IDEAI) located in UPC. His research interests are focused on algorithms and protocols for computer networks with special interests in interoperability, control and management, planning and routing. He has been involved in several EU and industrial research projects.



Ana I. Alves Moreira has a Bachelor's degree in Psychology and a Master's degree in Psychology of Intercultural Relations from ISCTE-IUL for which she received an academic excellence award. She is a licensed psychologist by the Portuguese Order of Psychologists and is an International Affiliate Member of the American Psychological Association. She has been working as a Junior Project Manager in European Projects at AidLearn. and as a freelance research assistant, mostly related to data analysis/interpretation. Her research interests mostly lie in social cognition, culture and diversity, and social inclusion.



Cecilio Angulo Bahón received his Doctorate in Applied Mathematics from the UPC in 2001 where he is currently a Full Professor and founder of the IDEAI Research Center. He is also a member of the Ethics Committee in this university. His research interests include theoretical aspects of machine learning, social and cognitive robotics, reinforcement learning and human-robot interaction. He is the author of books on machine learning and robots, and has published nearly 300 articles in journals and conferences. He has led and participated in 47 competitive R&D projects, 21 of them funded by the European Commission.



Federica Casaccio holds a Bachelor Degree in Political Science and International Relations (Italy and Croatia) and a Master's Degree in Peace, Conflict and Development Studies (Spain), specialised in the analysis of peace education as a society's transformative tool. She has extensive expertise in research, advocacy and project management on the valorisation of education. Currently, she works as Strategic Development Officer at ACEEU, where she is involved in business development and in the management of several EU-funded projects within the cross-cutting themes of digitalization, entrepreneurship and innovation in HE.



Rozalina Dimova received the Diploma Eng. degree in electrical engineering from the Technical University Varna, the PhD degree from the Technical University of Sofia and has experience as a lecturer for more than 20 years in communication engineering and technologies. She is the head of Applied Health Technologies Center and former Rector of the Technical University of Varna. She has current research interests in ICT, AI and next generation networks. She is co-authored over 120 scientific publications and has participated either as project leader or member of scientific teams in 7 European and more than 50 national funded projects.



Tihomir Dovramadjiev has professional activity related to management, academic and scientific, educational, and other to the BAEHF and Technical University of Varna (TUV, Bulgaria). He received the Ph.D. in Ergonomics and industrial design (TUV/2012). He is an associate professor Dr. Eng. at TUV more than 10 years at Industrial Design Department. He is the author of the book "Ad-vanced technologies in Design", TUV, pp. 228, ISBN: 978-954-20-0771-5, 2017. He has participated in over 60 scientific publications (Incl. Springer & Elsevier). Based of ResearchGate data (2022), he has about 300k research reads.



Antonia Jakobi concluded her Bachelor's degree in Psychology in Germany, after which she proceeded to finalize her Master of Science degree in Psychology of Intercultural Relations in Portugal. Since then, she has been working as a Junior Project Manager in European Projects at AidLearn. Her work focuses on Erasmus+ funded European projects that deal with societal challenges and issues from an educational approach. Therefore, she has been involved in various projects on topics like gender equality, global citizenship and social inclusion. Her interests are psycho-societal issues, in particular social inclusion, diversity, and gender.



Csaba Kollar is a senior research fellow at the Óbuda University, leader of AI Workshop, lecturer and supervisor at the Doctoral School on Safety and Security Sciences, and at the National University of Public Service PhD School of Military Engineering. He is a communications engineering, certified communications specialist, head of electronic information security, consultant, coach, mediator. His research interests include the social aspects and economic impacts of the digital age, in particular the human dimension of information security and awareness, human-robot interaction, smart city, AI, social credit system, and domotics.



Ievgeniia Sukhovii is a philologist and international business engineer. She received her undergraduate degree in the field of English Philology from Namik Kemal University in Turkey, and her postgraduate degree in the field of International Business Engineering from the University of Montpellier in France. She completed two Erasmus+ exchange studies and two traineeships in Portugal, Poland and Germany. She worked as a language teacher, business and marketing developer and European project assistant. Her research interests include (but not limited to) entrepreneurship, education, business and marketing.



Gyula Szabó is a registered European ergonomist and assoc. professor at Óbuda University. He obtained an electrical engineering and teaching degree, a university doctorate at the Budapest University of Technology, and a PhD in ergonomics at the National University of Public Service. His primary research interest is the assessment of human factors at work, and he leads projects for their development. He participates in the scientific committee of conferences and proofreading journals. He is also an occupational safety specialist and engineer, specialist in investigating work accidents and ergonomics and human factors specialist.

HEDY

Life in the AI Era



ÓBUDAI EGYETEM
ÓBUDA UNIVERSITY



ACEEU
ACCREDITATION COUNCIL
FOR ENTREPRENEURIAL &
ENGAGED UNIVERSITIES



UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH



BAEHF