

## ARTICLE

# Intel·ligències artificials i el futur del treball: creixement desacoblat amb riscos existencials?

JOAN TORRENT-SELLENS

ESTUDIS D'ECONOMIA I EMPRESA, UNIVERSITAT OBERTA DE CATALUNYA (UOC)

### Resum

En aquest assaig prospectiu he organitzat l'evidència disponible sobre els tres tipus principals d'intel·ligències artificials (IAs) (predictiva, generativa i transformativa), i les he vinculat amb el pensament econòmic i estratègic a través dels valors que impulsen i dels efectes econòmics i laborals que generen. La idea de l'article és explicar la capacitat de transformació de cada tipus d'IA i analitzar com es vinculen amb el futur del treball i amb els riscos existencials o extrems d'origen laboral que poden generar. L'anàlisi ha revelat que en el context neoliberal o autocràtic predominant, els incentius per al desenvolupament de la IA predictiva i generativa decanten el seu ús amb criteris d'automatització i/o control. A més, estan en mans d'un nombre molt reduït d'empreses, s'articulen en base a models de negoci que cerquen rendes, substitueixen un nombre creixent d'habilitats cognitives sense promoure avenços prou significatius de productivitat, i tenen uns usos molt heterogenis i desiguals. La IA es va desalineant de les persones i del floriment social perquè tendeix a empobrir el treball. Tanmateix, l'arribada de la IA transformativa, una IA amb molts més agents, més potent i més hàbil per aprendre, millorar-se i replicar-se, adquirir recursos, produir idees i generar innovacions, podria impulsar un salt de nivell tant en termes positius com negatius. Una IA transformativa que augmentés exponencialment el treball cognitiu i que s'alineés amb el progrés humà podria generar una explosió en termes de producció de coneixement, idees, innovació i creixement econòmic. Però, aquesta mateixa IA transformativa pot ser utilitzada per a superar les persones i crear una economia de les màquines que assumeixi els nous valors de predicció, creació i transformació i desacobli el creixement. A través d'aquesta superioritat laboral i econòmica, la IAT podria colonitzar la humanitat, fer-la redundant en moltes dimensions de la vida i plantejar un oceà de riscos extrems i existencials per les persones. Seguint les recomanacions de la comunitat científica en IA i dels economistes de la tecnologia, aquest article és un crit d'alerta perquè comencem a organitzar i gestionar (no necessàriament prohibir) una economia política sobre l'adopció i els usos de la IA.

### Paraules clau

*Intel·ligència artificial (IA), IA predictiva (IAP), IA generativa (IAG), IA transformativa (IAT); Futur del treball; Riscos existencials; Creixement desacoblat; Economia política.*

---

# Introducció: economia política de les IAs

“Més resumidament, la IA podria augmentar el nivell de vida més que l'electricitat o Internet. Però també pot suposar riscos que superen els de les armes nuclears. A més, aquestes possibilitats, per probables o poc probables que siguin, estan correlacionades. És precisament l'estat del món en què la IA és prou potent com per generar augments profunds en els nivells de vida on sembla més probable que suposi també un risc existencial” (Charles I. Jones, *The A.I. dilemma: growth versus existential risk*, 2023, pàg. 1).

Durant els darrers anys una segona onada de transformació digital està modificant amb força l'economia, el treball i la societat. A l'epicentre d'aquesta transició digital (evolució des de la primera a la segona onada de la digitalització) hi ha les tecnologies i els desenvolupaments tècnics de la intel·ligència artificial (en endavant, IA). Tot i que hi ha moltes definicions d'IA, entre altres l'aproximació de la ciència computacional que ens diu que la IA és tot allò que les màquines fan i que encara no saben explicar bé com ho han fet, socialment es pot acceptar que la IA és “l'esforç per automatitzar tasques intel·lectuals mitjançant una màquina, que, governada per un algoritme, és capaç d'executar per si mateixa funcions que generalment requereixen la participació d'intel·ligència humana” (Torres, 2023, p. 28). De fet, aquesta idea seminal que ja va tenir John McCarthy, un dels pares de la IA, a meitats dels anys cinquanta del segle passat i que ens suggereix que la IA són màquines que actuen d'una manera que en diríem intel·ligent si aquest comportament el tingués una persona, ha anat passant l'escrutini del temps. Així que podem entendre la IA com a artefactes que usen tècniques molt diferents per a resoldre una gran quantitat de tasques, i que ho fan emulant capacitats humanes cognitives (López de Mántaras-Badia, 2023).

Per la seva banda, l'economia fa una aproximació a la IA, per descomptat, molt més prosaica i útil pels seus propòsits explicatius del comportament i la presa de decisions en l'esfera de la inversió, la producció, el consum i l'intercanvi. Econòmicament, podem entendre la IA com un procés d'innovació tecnològica. En particular, el procés de creació i desenvolupament de tecnologies digitalitzades amb aplicacions tècniques que simulen capacitats intel·ligents humanes. Així doncs, la IA es refereix a tot aquell estoc social de coneixement (tecnologia) que es fa servir per desenvolupar tota mena d'artefactes digitals (materials i immaterials) que, una vegada generats, s'apliquen a l'activitat econòmica (tècnica) i es basen en l'emulació (augment, substitució o desplaçament) de capacitats cognitives humanes (Torrent-Sellens, 2024a). Aquestes capacitats poden anar des de l'anàlisi de grans quantitats de dades (representació codificada de tot tipus d'esdeveniment numèric, textual, imatge, veu, etc), la detecció de patrons ocults o el suport expert en tota mena de decisions i comportaments. De fet, totes aquestes tècniques es basen en la capacitat de la IA per fer prediccions o IA predictiva (en endavant, IAP). La IAP aglutina tots aquells sistemes computacionals o algoritmes que utilitzen l'anàlisi avançada i tècniques basades en la lògica, inclosos l'aprenentatge automàtic i l'aprenentatge profund, per interpretar i predir esdeveniments, donar suport, automatitzar decisions i dur a terme algunes accions en contextos controlats (Girón-Sierra, 2023). Així doncs, el valor de predicció, que no sempre s'equival amb el valor d'explicació, és la força competitiva principal que les empreses extreuen de l'ús empresarial de la IAP amb importants conseqüències pels seus models de negoci i resultats, així com pels llocs de treball que s'hi associen (Agrawal et al., 2022a, 2022b).

Tanmateix, la recent irrupció de la IA generativa (en endavant, IAG), amb aplicacions com *ChatGPT*, *Copilot*, *Gemini* i altres, ha marcat un nou punt d'inflexió, clarament disruptiu, en la utilització econòmica i laboral de la IA. Aquest tipus d'IA, que també utilitza dades massives digitalitzades i algoritmes d'aprenentatge automàtic i profund, és capaç de generar tot tipus de documents digitalitzats, com textos, informes tècnics, treballs acadèmics, documents legals, resums, correus electrònics, traduccions, i fins i tot codi de programari o dissenys creatius, audiovisuals i artístics, que són emprats de manera creixent a tots els entorns laborals, i que fan encara molt més capaç a la IAP. De fet, la IAG va assolir la massa crítica de 100 milions d'usuaris,

considerada com el punt de partida per assolir la categoria de tecnologia disruptiva, en un temps rècord i inaudit per a cap altra tecnologia de la història (al voltant de tres mesos després del seu llançament) (Wang et al., 2024). De la mateixa manera, les primeres dades representatives ens confirmen uns usos individuals, educatius o laborals, situats a l'entorn del 20% del total poblacional als primers països analitzats, per bé que amb significatives diferències per edat, gènere i nivell educatiu (Fletcher i Nielsen, 2024). En aquest context, no és estrany que seguint l'estela de la IAP (Crafts, 2021; Goldfarb et al., 2023) però amb molta més velocitat d'implantació, la IAG estigui començant a rebre també la consideració de tecnologia d'utilitat general (*General Purpose Technology, GPT*, en terminologia anglosaxona) (Eloundou et al., 2024). Una GPT és una tecnologia d'ordre superior que, mitjançant la seva capacitat de connexió amb altres tecnologies, és capaç de generar processos de convergència tecnològica, innovacions derivades, complementarietats amb altres actius econòmics (especialment amb el capital intangible i l'humà) i, finalment, nous models de negoci, noves fonts de productivitat, noves palanques per al creixement econòmic i un nou cicle econòmic de llarga duració (Torrent-Sellens, 2022).

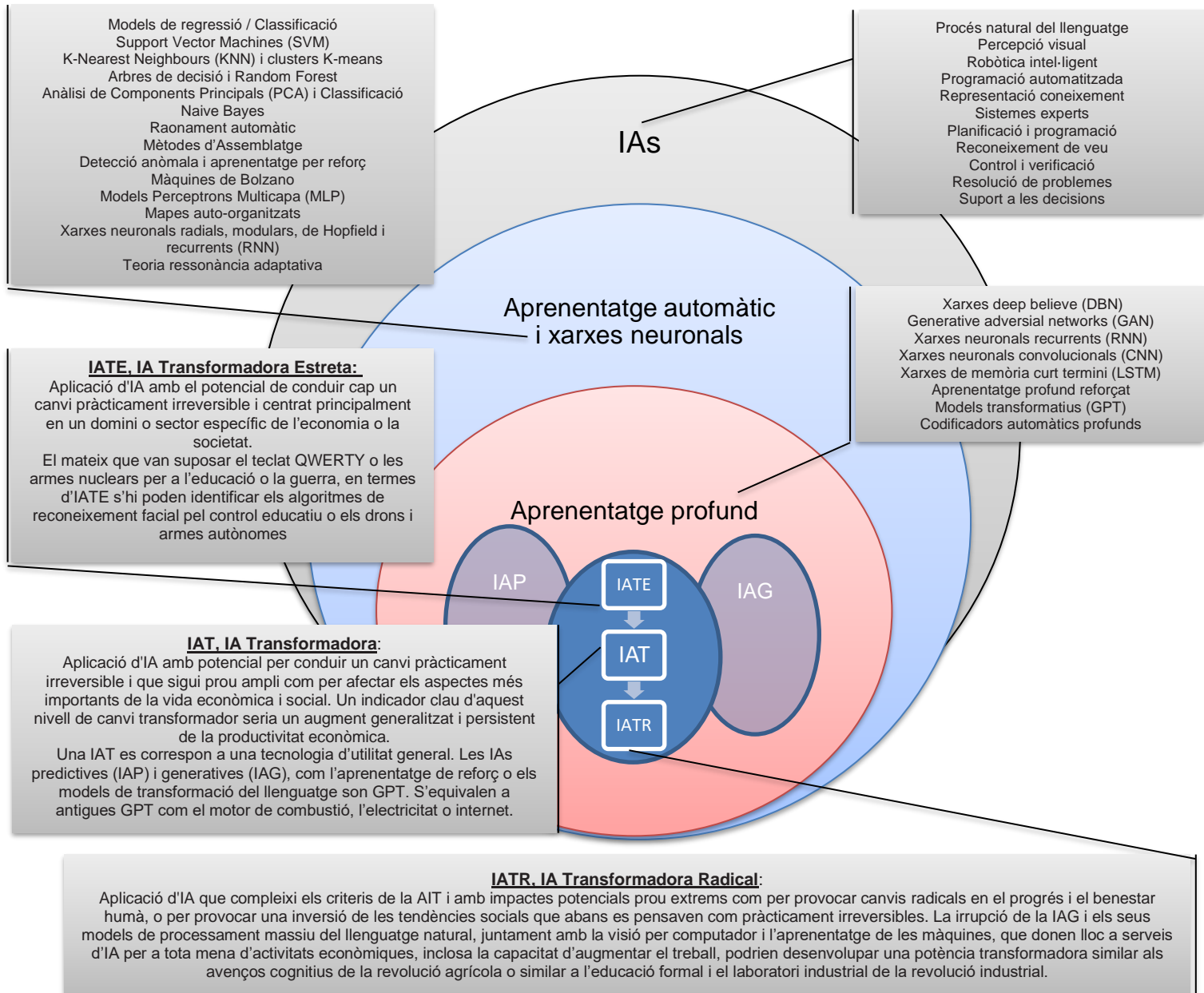
La recent consideració com a GPT de la IAG ens suggereix l'adveniment d'un nou tipus de valor que també adquireix una rellevància creixent per a la competitivitat i la sostenibilitat de les empreses i, en conseqüència, per l'ocupabilitat de les persones (Feuerriegel et al., 2024). El valor de creació, és a dir la utilització d'algoritmes de aprenentatge automàtic i profund per a crear de manera versemblant tota mena de documents digitalitzats de text, formulacions numèriques, àudio, imatges, obres artístiques o programari, entre altres, eleven la capacitat de la IA a un nivell superior de prestacions. De fet, la conjunció entre el valor de predicció i el valor de creació estan transformant de forma accelerada i profunda moltes tasques i gran part de professions de la indústria, els serveis professionals i a les empreses, alguns serveis personals, la recerca i l'educació, les activitats creatives i culturals, o l'assistència sanitària, entre moltes altres (Acemoglu et al., 2023).

Però tot sembla indicar que el progrés "artificial" no s'aturarà aquí. De fet, la seva pròpia naturalesa basada en l'aprenentatge continuat, i en la mesura que les capacitats de computació i de connectivitat responguin a unes necessitats d'anàlisi de grans dades cada vegada més massives i a major escala, farà que les futures generacions d'algoritmes portin la combinació entre la IAP i la IAG cap a una IA molt augmentada amb molts més agents, més potents i més hàbils per aprendre i replicar, millorar-se i adquirir recursos. Així que l'arribada d'una IA transformadora (en endavant, IAT) és, en els moments que els lectors llegeixen aquest article, molt més plausible del que era abans de l'arribada de la IAG i la seva connexió creixent amb les eines de la IAP. Podem conceptualitzar la IAT com una IA capaç d'actuar com un agent independent i autònom, que busca assolir els seus objectius, i que té un rendiment molt per sobre de l'assolit pel treball humà en una àmplia gamma de tasques, incloses moltes de les tasques essencials per a l'economia (Jones, 2023; Growiec, 2024).

La IAT és una IA molt capaç, extraordinàriament potent, però que no necessàriament té perquè arribar a la singularitat, la capacitat de les màquines per a superar en totes i cadascuna de les seves habilitats a les persones (Growiec, 2022). De fet, la IAT es pot definir a través dels nivells de profunditat de les transformacions que genera. La investigació de l'àmbit ha distingit tres nivells d'IAT en funció de sí el seu potencial de transformació és estret, mitjà o radical (Gruetzemacher i Whittlestone, 2022). La transformació estreta es vincula amb canvis de molt calat, però circumscrits en un domini o sector específic de l'economia. Els algoritmes de reconeixement facial emprats als sistemes educatius o els drons i armes parcialment autònomes en seguretat i defensa, serien exemples d'IAT estreta (en endavant, IATE). A l'alta costat de l'evolució hi trobem les IAT radicals (en endavant, IATR). Radicals en el sentit que son capaces de generar impactes tant potents que modifiquin molt substancialment el progrés social i el benestar de les persones. La combinació entre diferents tecnologies d'IAP i d'IAG, per exemple, el processament del llenguatge natural, la visió per ordinador i l'aprenentatge robòtic aplicats conjuntament en molts sectors d'activitat, sí que podrien conduir a un escenari de transformacions econòmiques i socials de magnitud semblant o superior a les

assolides durant les revolucions industrials. Al mig es situa la IAT. És una IA que es consolida com a tecnologia d'utilitat general i que genera augments molt substancials dels indicadors de resultat econòmic, com la productivitat o el treball, però sense alterar significativament a curt termini la cultura, les institucions o l'estructura social. A la figura 1 es representen les diferents tipologies d'IA, així com les habilitats emulades, les tècniques utilitzades i la seva capacitat transformadora.

**Figura 1. Intel·ligències artificials, habilitats i capacitats de transformació econòmica**



Font: Elaboració pròpia.

En aquest article em proposo reflexionar a propòsit de l'emergència d'un conjunt convergent d'IATs que, basades en els avenços exponencials de la potència predictiva i generativa, fossin capaces d'alterar molt significativament, tal i com preveuen molts experts de IA (Grace et al., 2024; Roser, 2024) i, després de moltes reticències inicials (Jorgenson, 2021), cada vegada més economistes del canvi tecnològic digital, la vida laboral i econòmica de les persones (Korinek i Juelfs, 2022; Bonney et al., 2024). Com tota ruptura tecnològica, l'emergència de la IAT podria treballar a favor de les persones i, gràcies a la seva potència transformadora,

generar un nou cicle de creixement econòmic i prosperitat material per a les properes dècades o, per al contrari, constituir-se com un dels riscos existencials més rellevants que ha hagut d'afrontar la humanitat en la seva història. De fet, tot sembla indicar que ambdues coses estan directament relacionades i que, suposant que encertem, una IAT que fos capaç de treballar pel progrés social compartit no està exempta de grans riscos econòmics i socials que haurem d'aprendre a gestionar (Dwivedi et al., 2023).

En el context laboral, la investigació de l'àmbit comença a donar crèdit a la possibilitat que la IA acabi augmentant molt considerablement les seves capacitats per a emular el treball cognitiu humà, així com l'aplicació intencionada d'aquest coneixement en forma d'innovació. Una vegada les màquines siguin capaces de produir i aplicar idees a l'activitat econòmica per sí soles, els límits al creixement establerts per la quantitat i la qualitat dels professionals de l'R+D+i poden afeblir-se moltíssim (Korinek, 2023). Aquests nous processos d'automatització i de creació de coneixement (auto-rèplica i auto-millora del capital) impulsarien una gran acceleració i el renaixement del creixement econòmic (Trammell i Korinek, 2023). Però en la mesura que moltes de les habilitats del treball humà s'anessin depreciant, aquest creixement econòmic s'aniria desacoblant de les persones, sobretot si aquestes no fossin capaces d'assumir nous nínxols competencials i canviant de tasques i d'ocupacions. És precisament en aquest sentit que aquests avenços espectaculars de la IAT no estan exempts de riscos. Tot el contrari, hi estan directament vinculats. Si es confirmen els pronòstics i emergeix una IAT capaç d'impulsar una revolució càmbrica a l'esfera econòmica i laboral (Pratt, 2015), els riscos d'usos malintencionats i erronis, o fins i tot, aproximacions ben intencionades però poc centrades en les persones i el treball humà augmentat, o en les repercussions socials de la IAT, podrien generar riscos molt importants per la humanitat (Wang et al., 2024). Com que estem tractant amb una tecnologia transformadora molt poderosa, els riscos que s'utilitzi inadequadament o amb males intencions, amb conseqüències desastroses per les persones i l'entorn socio-ambiental, també són molt elevats. Així doncs, és molt important anar seguint quina és la velocitat d'implantació i de transformació de la IAT, avaluar constantment els seus efectes i externalitats, fer prospectiva de futur, i estar preparats tant per la promoció i la gestió dels seus beneficis, com per a la lluita i la governança dels seus riscos existencials (Bengio et al., 2024).

Com tot exercici de prospectiva, incert per naturalesa, en aquest article em basaré en l'evidència empírica que hem acumulat a propòsit de les principals conseqüències laborals dels usos de la IAP i, en menor mesura, de la IAG que és molt més recent. En aquest context, i per a fer una prospecció útil, és especialment important recuperar moltes de les aproximacions i conceptes de l'economia política i l'economia evolutiva. Sense una aproximació plural, que tingui en compte tots els fonaments emocionals, institucionals, ètics, socials, històrics, culturals i polítics del comportament i la presa econòmica de decisions amb presència de tecnologia, i que sigui capaç de superar l'aproximació neoclàssica, únicament focalitzada en decisions racionals, aïllades, egoistes i calculadores, és impossible fer una diagnosi acurada de les possibles conseqüències de la IAT. En articles anteriors ja he defensat la necessitat de construir una nova economia política de la digitalització i quines haurien de ser les seves peces principals. A mode de resum:

“La nova economia política de la digitalització representa una nova proposta científica perquè l'economia recuperi la seva finalitat social. S'anomena economia política perquè expandeix i supera els postulats de l'economia neoclàssica: 1) estenent-se més enllà dels preceptes individuals, racionals i restrictius de *l'Homo economicus*; 2) assumint que la presa de decisions econòmiques amb presència de tecnologia també considera factors i contextos grupals, dinàmics, emocionals, ètics, culturals, institucionals, històrics i polítics; 3) creuant-se amb tot tipus de ciències socials, humanes, jurídiques i experimentals; 4) combinant mètodes científics positius amb finalitats i objectius normatius; i 5) essent especialment sensible a les ideologies, figures i narracions dels seus diferents components” (Torrent-Sellens, 2024b, p. 2).

De fet, i seguint aquesta aproximació, els dos principals propòsits i relats de la IA tenen a veure amb les principals funcions que la tecnologia tradicionalment ha assumit en la dinàmica econòmica: l'automatització, que substitueix o desplaça el treball, i l'augment del treball. Molts economistes afegirien el control a aquesta dupla d'objectius (Muldoon et al, 2024). Per motius d'espai en aquesta comunicació no l'abordaré. Centrant-

nos en aquestes dues funcions, en aquest article revisaré l'evidència que tenim sobre els efectes de les IAs (IAP i IAG) en el món del treball. Una vegada revisada aquesta evidència, i tenint en compte les principals característiques que han seguit el processos d'adopció tecnològica digital durant els darrers anys, analitzaré sí d'aquests resultats se'n poden extreure conclusions vàlides per a la IAT i el futur del treball. Aquesta anàlisi es circumscriurà en algunes de les principals dimensions del treball, no totes, com la quantitat, qualitat i retribució del treball.

## 1. Transformació digital i treball

“En lloc de fixar-nos tant en la intel·ligència artificial, hauríem de preguntar-nos per la seva utilitat per als éssers humans, cosa que és, precisament, la definició d'utilitat de les màquines.” (Daron Acemoglu i Simon Johnson, *Power and progress*, 2023, pàg. 5).

Durant les últimes quatre dècades de paradigma digital, les tecnologies de la primera onada digital o de la informació i la comunicació (TIC), com els ordinadors, internet o els telèfons intel·ligents, han estat capaces d'impulsar la productivitat del treball i convertir-se en un motor del creixement econòmic (Vu et al., 2020), sobretot gràcies als avenços en la computació, l'anàlisi i gestió de la informació, i les capacitats de comunicació. El valor de càlcul, informació i comunicació impulsats per les TIC han estat capaços tant d'augmentar com de desplaçar i substituir el treball de les persones. Com totes les tecnologies anteriors, el vincle entre les tecnologies digitals i el treball s'estableix a partir d'un procés primigeni de desplaçament d'habilitats. El canvi tecnològic digital que esbiaixa habilitats modifica les capacitats i competències que han de desenvolupar les persones per treballar amb solvència utilitzant les noves tècniques digitals. Seguint amb l'exemple clàssic de la investigació de l'àmbit: les competències demandades als treballadors per al seu desenvolupament professional utilitzant llapis, màquines d'escriure, ordinadors personals, internet o la IA són clarament diferents. A partir d'aquest procés inicial de biaix, se sol produir que el canvi tècnic va de la mà d'unes noves o majors necessitats de formació i capacitació dels treballadors, així com d'unes empreses o organitzacions més innovadores, flexibles i obertes al canvi (Torrent-Sellens, 2008).

Hem après que, en les seves primeres passes, el biaix digital ha fomentat l'ocupabilitat dels treballadors més preparats per al nou paradigma, en general els més ben formats (digital i no digitalment), i ha debilitat les oportunitats per començar o continuar treballant, així com les condicions de treball, de les persones menys preparades (Card i Di Nardo, 2002). De fet, tenim molta evidència que ens posa de relleu que, precisament a partir de la irrupció de l'era digital, els mercats de treball es veuen immersos en processos creixents de polarització, cosa que ha reduït la participació dels salaris, en detriment del capital, en la renda nacional (Acemoglu i Restrepo, 2022). Aquesta polarització està directament vinculada amb l'orientació del biaix d'habilitats, clarament enfocat cap a la substitució d'habilitats i tasques rutinàries (cognitives i no cognitives) i cap a l'avenç del treball cognitiu i no rutinari (Autor et al., 2003). En aquest sentit, la creació de nous llocs de treball s'ha anat desplaçant des de les ocupacions de producció i d'oficina amb remuneració mitjana, típiques de l'era madura de la societat industrial, cap a les ocupacions professionals ben remunerades i, també cap als serveis de baixos salaris, tot confirmant la polarització del treball a l'era digital (Autor et al., 2024). I encara més, les tecnologies digitals solen exercir efectes laborals diferents en funció del seu propòsit. De fet, es constata un efecte substitució entre els usos laborals de la tecnologia que tenen com a propòsit l'automatització o el control, i que debiliten tant la quantitat com la qualitat del treball, i els usos laborals de la tecnologia que tenen com a objectiu l'augment del treball fent possibles tant avenços de la quantitat com de la qualitat del treball.

Un gran exemple d'això són els diferents usos tècnics d'una mateixa tecnologia digital, la robòtica industrial. La robotització exerciria efectes positius sobre la productivitat i efectes mixtos, sense caure en la pèrdua o precarització massiva, sobre l'ocupació. Tot i substituir ocupació industrial de baixa qualificació, la robòtica

industrial generaria efectes desbordament i externalitats cap a la qualitat de l'ocupació i els salaris dels empleats industrials més resilients i amb més habilitats per a les complementaritats amb la digitalització, i cap a la quantitat i qualitat d'ocupació als serveis (Kariel, 2021; Klenert et al., 2023). Com a contrapartida, la polarització laboral perjudicaria als empleats, empreses, sectors i territoris amb pitjors treballs, productivitat i retribucions associades una menor capacitat dinàmica per interactuar amb la robòtica. Tots aquests resultats ens indiquen que l'automatització robòtica intel·ligent impulsaria la productivitat de l'empresa, però pel que fa als seus efectes sobre l'ocupació, tindriem efectes mixtos i molt dependents del propòsit, l'entorn econòmic i el model de generació de valor existent a les empreses (Ballestar et al., 2021a). En investigacions desenvolupades per al cas de la indústria espanyola i en el període 2000-2016, ja vàrem constatar que l'adopció d'un robot industrial generava una pèrdua mitjana de 12 llocs de treball a l'empresa industrial d'Espanya. En canvi, quan l'adopció de robòtica industrial es combinava amb la formació universitària dels treballadors, el resultat era totalment diferent. Aquest efecte combinat acaba per generar un efecte net positiu de 6 treballadors industrials més (Camiña et al, 2020; Ballestar et al., 2021b).

Acabem de constatar que aquest desplaçament d'habilitats vinculat amb la transformació digital a les empreses és clarament heterogeni (Koch et al. 2021). En general però es pot afirmar que impulsa la productivitat i el rendiment de les empreses més competitives, en particular les que són capaces d'establir millors relacions de complementaritat entre la transformació digital, el capital humà i el replantejament estratègic i organitzatiu, i tendeix a destruir creativament les empreses menys preparades per la transformació digital. De fet, la investigació de l'àmbit sol parlar també del canvi tècnic digital que esbiaixa organitzacions, en el sentit que per a ser capaces d'introduir amb èxit la tecnologia digital i les noves habilitats de les persones, les empreses i organitzacions han d'adquirir també tot un nou conjunt de capacitats dinàmiques (Caroli i Van Reenen, 2001; Piva et al., 2005). De fet, l'existència d'importants relacions de complementaritat entre l'ús de la tecnologia, la formació i l'aprenentatge dels treballadors, el canvi organitzatiu i les innovacions en els models de negoci, són una part primordial de l'explicació dels resultats de la interacció entre tecnologia i treball. I, d'això en tenim múltiples exemples històrics. Perquè els beneficis de l'electrificació es generalitzessin van haver de passar gairebé una generació de treballadors, directius i models de negoci. De la mateixa manera, molts canvis tecnològics de la primera onada digital, com el comerç electrònic, no van generar impactes positius sobre el treball fins que persones i organitzacions es van adaptar i superar les corbes d'aprenentatge del nou context tecnològic (Brynjolfsson et al., 2021).

Les capacitats dinàmiques són capacitats d'ordre superior que ajuden a crear, reconfigurar i aprofitar recursos i capacitats de les organitzacions. Per tant, és possible definir les capacitats dinàmiques com: "les capacitats evolutives de les empreses o organitzacions per integrar, construir i reconfigurar competències internes i externes amb l'objectiu d'abordar entorns que canvien ràpidament" (Teece et al., 1997). En conseqüència, les capacitats dinàmiques reflecteixen la preparació d'una empresa o d'una organització per aconseguir formes noves i innovadores d'avantatge competitiu ateses les dependències de la trajectòria i les posicions del mercat. Segons les capacitats dinàmiques, les empreses o les organitzacions aconseguen avantatges competitius a través de l'aprenentatge i el suport de l'experimentació, la recombinació de recursos, el creixement de nous productes i la transformació del sistema existent de producció. Les capacitats dinàmiques inclouen habilitats, processos, procediments, estructures organitzatives, regles de decisió i disciplines que donen suport a la detecció, explotació i reconfiguració de capacitats (Teece, 2014). No hi ha dubte que, en el context de la digitalització i les seves causes i conseqüències, les capacitats dinàmiques són de gran utilitat. De fet, la investigació de l'àmbit ja fa temps que ha vingut assenyalant tot un conjunt de capacitats dinàmiques digitals que serien molt adequades per què organitzativament les empreses entomessin amb èxit els reptes de la transformació digital (Warner i Wäger, 2019; Ellström et al., 2022).

En resum, hi ha un conjunt multidimensional de factors que expliquen la relació final –positiva o negativa– entre una tecnologia i l'ocupació. La tecnologia per se no és mai la causa única de cap resultat laboral. Són

els usos de la tecnologia, les tècniques que aquesta adopta, i la seva imbricació amb factors personals, organitzatius, empresarials, econòmics, socials, laborals, culturals i, per descomptat, històrics, polítics i institucionals els que acaben determinant aquesta relació. En general, i almenys fins a la segona onada de la digitalització, la imbricació entre tecnologia –també la digital de primera onada– i el conjunt de dimensions i factors vinculats amb la quantitat d'ocupació havia funcionat en el llarg termini. És a dir, que l'ocupació perduda en el biaix inicial d'habilitats havia estat més que compensada per un conjunt d'efectes d'eficiència, salaris, consum i empenedoria i innovació que havien impulsat la creació neta d'ocupació a llarg termini (Balsmeier i Woerter, 2019). Tanmateix, aquestes efectes nets sobre la quantitat de treball no s'haurien traslladat en igual mesura a la qualitat de l'ocupació ni als salaris mitjans, tot generant notables polaritzacions i desigualtat d'arrel laboral. A més a més, la situació podria estar empitjorant. Pel que fa la quantitat de treball, tot sembla indicar que aquests mecanismes compensatoris no estarien funcionant tan bé quan ens referim a les tecnologies de la segona onada digital, i en especial quan ens centrem en algunes de les tecnologies de la IA (Virgilio et al., 2024). Serà aquesta vegada diferent?

## 2. IAs, automatització i augment del treball

“Jo diria que el nivell de vida als països progressistes d'aquí a cent anys serà entre quatre i vuit vegades superior a l'actual. No hi ha res de sorprenent en això, fins i tot a la llum del nostre coneixement actual. No seria absurd contemplar la possibilitat d'un progrés més gran. Arribo a la conclusió que, suposant que no hi hagi guerres importants ni un augment accelerat de la població, el problema econòmic es pot resoldre, o al menys estar en vies de solució, en un termini de 100 anys” (John Maynard Keynes, *Las posibilidades económicas de nuestros nietos 1930*, a *Ensayos de Persuasión*, 1982 (1931), pàg. 325).

Acabem de constatar que les complementarietats tecnològiques, formatives i organitzatives estan a la base de l'explicació de la transformació de l'ocupació i les relacions laborals en període digital. En concret, disposem de proves suficients a propòsit de que les complementarietats entre el canvi tecnològic digital i el canvi organitzatiu, quan s'orienten cap a la mateixa direcció, acaben reforçant les transformacions de les habilitats i les competències requerides en la força de treball. Aquest seria el fenomen principal a través del qual la primera onada de la digitalització va exercir efectes directes i indirectes positius sobre la quantitat de feina. Però l'impuls i la creació de nous llocs de treball més qualificats i, en principi, no sempre, de més qualitat i millor retribuïts, és només una part de l'engranatge laboral d'una economia. Aquesta creació d'ocupació també coincideix amb la destrucció o el deteriorament de llocs de treball, que abans eren bons i que amb la digitalització han deixat de ser-ho (Ballestar et al., 2020). Per tant, l'anàlisi econòmica s'ha d'ocupar tant de les ocupacions guanyades com de les perdudes o deteriorades, i molt especialment dels factors, institucions i polítiques que cal impulsar perquè les onades digitals, especialment la transició digital liderada per la IA, pugui oferir oportunitats laborals en condicions d'igualtat i sense discriminacions per a tothom. Per descomptat, això també implica preocupar-se per la desocupació o la subocupació tecnològica, la qualitat del treball híbrid entre persones i màquines, o la polarització salarial, les seves causes, externalitats, complementarietats, conseqüències i solucions (Virgilio et al., 2024).

### 2.1. IAs, tasques, ocupacions i treballs

En aquest context, la irrupció laboral de la IA ha generat una gran controvèrsia a propòsit de sí serà capaç de crear i/o millorar els llocs de treball o, per al contrari, estem davant d'una tecnologia que, a través dels típics processos d'automatització i control, desplaçarà, substituirà i/o empitjorarà els llocs de treball. Aquesta controvèrsia té la seva lògica per que la IA està impulsant tots aquests efectes, positius i negatius, a la vegada (Acemoglu i Johnson, 2023). A través del biaix inicial d'habilitats i la substitució o la complementarietat de tasques d'intel·ligència humana, la IA pot generar tant processos de substitució i desplaçament del treball



humà (automatització robòtica) com processos de complementarietat d'intel·ligències (algoritmes de predicció o de generació). De fet, el més habitual és que ho faci tot alhora. La IA pot augmentar tasques i fluxos de treball (fulls de càlcul i comptables), els pot reduir (centres logístics), els pot reconstituir, eliminant i creant noves tasques alhora (radiòlegs) o els pot substituir (conductors d'autobús).

De fet, és possible resumir el panorama laboral vinculat amb la IA afirmant que es caracteritza pels típics efectes positius de la innovació tecnològica sobre el treball, generats a través dels canvis d'habilitats, l'avenç de la productivitat, la creació de nous llocs de treball i la millora de salaris. Aquesta és la situació de molts dels treballadors del coneixement, que s'identifiquen a través d'elevats nivells educatius, la intensitat tecnològica dels seus sectors d'activitat i els usos elevats de la digitalització en les seves ocupacions, generalment segures i protegides (Torrent i Ficapal, 2009). I encara més, la satisfacció i la salut laboral dels treballadors del coneixement és clarament superior i va molt més enllà dels típics factors financers i materials de la resta de treballadors que no utilitzen el coneixement de manera tant intensiva als seus llocs de treball (Viñas-Bardolet et al., 2020). Però, la IA també ha arribat amb força a les ocupacions amb menys intensitat de coneixement, consolidant també un escenari complementari d'efectes negatius. La generalització de processos digitals d'automatització i control del treball a través de la IA, també té com a resultat la destrucció o el desplaçament de les ocupacions menys qualificades o l'empitjorament de les condicions de treball. Aquesta darrera també es dona en algunes ocupacions del coneixement. En conseqüència, la satisfacció i la salut de molts treballadors empitjora (Acemoglu i Restrepo, 2022; Bessen et al., 2023).

Per entendre tot aquest desencaix sol ser d'utilitat pensar en els llocs de treball com una successió de tasques que realitzen els treballadors. I, per dur-les a terme, poden interactuar o no amb la IA (Acemoglu i Restrepo, 2019). I és precisament en aquest punt, on apareix una dimensió econòmica importantíssima que separa els camins de la IAP i la IAG en el context organitzatiu i empresarial. Mentre que la decisió d'adoptar IAP és una decisió corporativa, la decisió d'utilitzar IAG, fins al dia d'avui, encara és una decisió individual. Seguint aquest fil argumental, normalment la IAP es sol implantar amb propòsits corporatius i amb criteris d'automatització i/o control, reemplaçant inicialment l'ocupació rutinària, cognitiva i no cognitiva, cosa que redueix o desplaça els llocs de treball i els salaris dels treballadors dels nivells ocupacionals baixos i mitjans, fent créixer la polarització (Frey i Osborne, 2017; Autor et al., 2020). De fet, la relació entre la IAP i la creació de treball cada dia que passa segueix més una forma d'U, tot i que per motius clarament diferenciats. En el cas dels trams més baixos de l'estructura de qualificacions i d'ocupacions es crea ocupació en funció de la relació entre el cost salarial i el cost/efectivitat de la IAP, en aquest cas molt més favorable per als salaris d'aquests col·lectius que estan evolucionant clarament a la baixa en termes reals (Cnossen, 2024; Georgieff, 2024). En canvi, en els trams més elevats de l'estructura ocupacional, es crea ocupació perquè els treballadors disposen de les competències necessàries, moltes d'elles socials però també competències STEM, perquè l'adopció organitzativa de la tecnologia sigui efectiva encara que la relació entre cost salarial i el cost/efectivitat de la IAP sigui favorable a la inversió en IAP (Aleekseva et al., 2021; Fontanelli et al., 2024). Aquesta dinàmica on simultàniament les tasques i ocupacions més rutinàries s'automatitzen, cosa que empeny als treballadors cap a la realització de tasques i d'ocupacions més cognitives i difícilment emulables per la IA, s'ha anomenant la cursa contra la màquina (Acemoglu i Restrepo, 2018).

En canvi, la recent irrupció de la IAG podria contrapesar en part les tendències cap a l'automatització de la IAP. Això és així per diversos motius. En primer lloc, perquè la decisió d'adoptar la IAG se situa més en el terreny individual, de manera que és fa difícil pensar que les persones utilitzaran la IAG per realitzar tasques que les substitueixi enlloc de complementar-les. De fet, la primera investigació de l'àmbit, de naturalesa experimental i contrastada per algunes ocupacions molt concretes, apunta clarament en aquesta línia. S'han observat clares vinculacions dels usos dels grans models del llenguatge (LLMs) amb avenços significatius de la qualitat, la rapidesa i la productivitat de les tasques, especialment per als treballadors novells i amb menys habilitats. A més, també s'han identificat importants avenços en la difusió corporativa del coneixement, des

dels treballadors més experimentats cap els que tenen menys experiència, o perfils d'adopció d'IAG diferenciats en funció d'uns usos en forma de centaure (auto-delegació d'unes tasques cap a la persona i d'altres tasques cap a la IAG) o de *cyborg* (integració de la IA a totes les tasques realitzades per les persones) (Brynjolfsson et al., 2023; Dell'Acqua et al., 2023). En segon lloc, les primeres investigacions més generals que han analitzat els riscos per la quantitat i la qualitat del treball de l'exposició a la IAG obtenen resultats heterogenis (Gmyrek et al., 2023). Per bé que les ocupacions vinculades amb el treball d'oficina estarien altament o moderadament exposades a la IAG (amb un 24% i un 58% de tasques, respectivament), la resta d'ocupacions tindrien un nivell d'exposició molt més baix, situat entre un 1% i un 4% de tasques en el cas de l'exposició alta i al voltant del 25% de tasques en el cas de l'exposició mitjana. I en tercer lloc, i per al nivell empresarial més global, la investigació ha confirmat l'evidència preliminar en el sentit que la IAG s'associa amb la rapidesa en el rendiment, el cost unitari i l'eficàcia de les tasques. Tanmateix, sembla que, tot i la seva lliure disposició, la distribució dels guanys de la IAG és molt asimètrica i beneficia a poques empreses, especialment les que disposen dels millors recursos i capacitats en termes d'accés a les dades, talent i capital humà, i els mecanismes de governança adequats per a gestionar la IAG (Bughin, 2024).

A diferència d'altres tecnologies digitals orientades cap a l'automatització de tasques, com els fulls de càlcul, els robots pre-programats o els algorismes d'IA personalitzada a les plataformes de *streaming*, a mesura que la IA va millorant el seu rendiment de la mà de les dades massives i la potència de càlcul, la seva capacitat es multiplica exponencialment (Growiec, 2024). De fet, la creixent convergència entre IAP i IAG augura un salt de nivell d'aquesta tecnologia amb potencials de transformació molt elevats, tant que determinarien l'emergència de la IAT. La combinació entre la IAG (i els seus algorismes massius i profunds de processament del llenguatge o de qualsevol altra dada digitalitzada,) i la IAP (i els seus algorismes, supervisats o no supervisats, que son capaços d'analitzar quantitats ingents de dades i detectar-ne patrons i comportaments previstos), està començant a generar profundes transformacions en les habilitats, l'estructura, les contraprestacions i l'organització del treball. Les dades que disposem sobre els usos i efectes de la IAT encara presenten uns resultats molt incipients i incerts, assenyalant repercussions més intenses per al nivell de tasques que per al nivell d'ocupacions. Als EUA al voltant del 27% de les empreses que utilitzen els models d'IA més avançada informen que substitueixen les tasques dels seus treballadors, però, de moment, només un 5% d'aquests experimenten canvis d'ocupació a causa de l'ús de la IAT. Tanmateix, es preveu que aquestes participacions augmentin fins a gairebé el 35% de les empreses i el 12% dels treballadors desplaçats, respectivament, durant els propers anys (Bonney et al., 2024). Ja ningú dubta que la IAT és capaç d'automatitzar moltes tasques cognitives sofisticades i no rutinàries realitzades fins avui només per a professionals qualificats (Korinek i Juelfs, 2022; Korinek, 2023). Tampoc hi ha massa dubte que en un futur no massa llunyà, algunes de les tasques de nova creació ja no les faran persones, sinó que seran assumides íntegrament per una IAT, capaç d'automatitzar parts integrals de la producció (Bengio et al., 2024). Això suposaria l'emergència d'un nou valor, el valor de transformació, que seria capaç de generar avantatges competitius per les organitzacions a través de l'assumpció per part de la IA de noves tasques, decisions estratègiques i activitats del negoci en alguns contextos determinants (Growiec, 2022). Però, com es relacionarà aquest nou valor amb el creixement econòmic i, de retruc, amb la dinàmica del treball?

## 2.2. IAs, treball cognitiu augmentat i creixement econòmic (desacoblat?)

El creixement econòmic és el resultat combinat de l'acumulació de factors productius i la innovació tecnològica. Tradicionalment, el creixement econòmic ha estat dirigit principalment per l'avenç del treball cognitiu de les persones que s'ha vist augmentat per les diferents onades de tecnologia (Torrent-Sellens, 2004). Tanmateix, tot i ser el principal puntal del creixement, el treball cognitiu augmentat per la tecnologia també s'ha constituït en un factor d'escassetat, cosa que ha limitat l'avenç de la riquesa material. De fet, i contràriament al que s'esperava als inicis del nou paradigma digital, el creixement exponencial dels fluxos d'informació, comunicació i coneixement no ha estat capaç de revertir l'escassetat d'idees, que han restringint

l'avenç del creixement econòmic (Bloom et al., 2020). De fet, el desacoblament entre les capacitats de les persones i de la tecnologia per a generar, processar i comunicar informació i coneixement només ha començat a ser possible amb la irrupció de les primeres onades d'IATs (Nordhaus, 2021). Precisament, aquesta capacitat que té la IAT per a desacoblar la producció d'idees i de coneixement de les persones és el principal fonament de la revitalització i l'acceleració del creixement econòmic que esperen molts dels economistes de la IA per al mitjà termini (Brynjolfsson et al., 2019; Trammell i Korinek, 2023). De fet, la IAT i la definitiva superació de les capacitats limitades de les primeres tecnologies digitals (automatització parcial i necessitats de supervisió humana per a prendre decisions i fer accions) podria solventar la creixent divergència que hi ha entre l'avenç del PIB mundial (que es duplica cada 23 anys) i la capacitat acumulada de comunicació, emmagatzematge i processament de tota mena de dades digitals (que es duplica cada 2 anys). Així que, l'emergència de la IAT promet un salt de nivell per a la producció d'idees, augmentant la recerca, l'escala de la investigació i la facilitació de la innovació, i planteja escenaris de substitució del treball humà en molts conjunts de tasques, tot eliminant molts dels límits de la cognició humana. Si la IAT és capaç de produir idees i desenvolupar innovacions, el creixement econòmic podria explotar i fer un salt de nivell. Tot i amb això, les capacitats cognitives revolucionàries de la IAT topen amb un conjunt de límits físics i econòmics que no fan molt plausible una explosió imminent del creixement econòmic en el curt termini (Almeida et al., 2024).

Això no significa en cap cas que l'avenç exponencial de les capacitats cognitives de la IAT i les seves aplicacions no siguin capaces de pressionar amb força els factors de creixement. De fet, només cal que millorin la seva capacitat d'agència (és a dir d'actuació en forma d'agent econòmic), de sentit comú i de comprensió del món real per a començar a descabdellar processos d'IAT radicals (Bubeck et al., 2023). Però, la IAT ja funciona com una tecnologia d'utilitat general, i amb múltiples disruptcions en marxa que van més enllà de la millora sistemàtica i que plantegen escenaris de creuament de punts de no retorn pel que fa a la superació dels límits humans d'acumulació de coneixement i noves idees (Kosinski, 2024). A títol d'exemple, algunes d'aquestes capacitats que, sorprenentment, han adquirit els algorismes d'IAT:

“Concretament, han après de manera inesperada a fer aritmètica, recuperar una paraula de lletres codificades, construir mapes conceptuals fonamentats, resoldre tasques de comprensió de múltiples llenguatges que cobreixen temes com la història o el dret, fer raonaments en diversos passos, seguir instruccions, codificar i executar programes informàtics” (Growiec, 2024, p. 1695).

En la mateixa línia sobre els avenços disruptius previstos per a la IAT en un futur proper, i en una recent investigació prospectiva realitzada a gairebé 2.800 investigadors en IA:

“Les previsions agregades donen al menys un 50% de possibilitats que els sistemes de IA avançada aconseguixin diverses fites per al 2028, inclosa la construcció autònoma d'un lloc de processament de pagaments des de zero, la creació d'una cançó indistingible d'una nova cançó elaborada per un músic humà, i la descàrrega i ajust autònom d'un model de llenguatge massiu. Si la ciència continua sense interrupcions, la probabilitat que les màquines sense ajuda superin als humans en totes les tasques s'estima en un 10% per al 2027 i en un 50% per al 2047. Aquesta darrera estimació és 13 anys inferior a l'assolida en una enquesta similar feta només fa un any. Tanmateix, s'ha pronosticat que la probabilitat que totes les ocupacions humanes esdevinguin totalment automatitzades assolirà un 10% el 2037, i el 50% com a molt tard el 2116 (comparat amb el 2264 de l'enquesta de 2022). (Grace et al., 2024, p. 1).

També hi ha però elements del context econòmic, social i polític que podrien impulsar o contrarestar aquests factors tecnològics que apunten cap una gran revitalització de la dinàmica econòmica. Un primer element restrictiu rellevant són els extraordinaris incentius que estableix la lògica dels efectes de xarxa i plataforma en els models de negoci privats de la IA que, a dia d'avui, controlen l'evolució de la tecnologia. Les típiques dinàmiques digitals del *guanyador s'ho emporta tot* en un context on el desenvolupament tècnic de la IAT només és possible per part d'algunes, molt poques, empreses *superstar*, podria conduir l'economia global cap una mena d'imperi corporatiu (a l'estil de la *Tyrell Corporation* a la pel·lícula *Blade Runner*, Torrent-Sellens et al., 2022b) on la primera entitat que desplegui la IAT assoleixi una avantatge competitiva de tal magnitud que

li permeti controlar la dinàmica de l'economia global. I, a través del poder econòmic, projectar totes les seves preferències sobre el futur de la humanitat. Tenim molta evidència disponible que ens certifica que els augments excessius del poder de mercat restringeixen el creixement econòmic, la mobilitat social i el benestar compartit (Autor et al., 2020; Eeckhout, 2022). En segon lloc, hi ha moltes incerteses sobre com s'establiria la relació entre creixement econòmic i benestar social compartit, en un context de desacoblament entre les fonts del creixement (fonamentalment tecnològiques) i la necessitat social de distribuir els beneficis d'aquest creixement a les persones. L'experiència ens assenyalava que sempre que la narrativa econòmica del moment no ha pogut justificar l'existència de grans desigualtats, el descontent social s'ha desbordat en forma de conflictes civils (molts violents) de tota mena (Piketty, 2019). Com a contrapunt a aquests aspectes negatius, l'emergència d'IATs es podria vincular amb un nou floriment de la humanitat en la mesura que contribuís decisivament al desenvolupament científic i tecnològic en termes de salut planetària, benestar humà i medi ambient (Bostrom, 2014). Així doncs, al costat de tot el potencial de la IAT per a impulsar una onada de creixement econòmic desacoblat sense precedents a través de la producció artificial d'idees i d'innovació, hi ha tot un conjunt d'elements vinculats amb la producció, organització i coordinació econòmica i social de la mateixa IAT que podrien afeblir o, fins i tot, eliminar aquest gran potencial. La investigació de l'àmbit sol parlar de desalineament entre IAs, persones i societat, cosa que impulsaria l'emergència de riscos existencials o extrems per a la humanitat (Turchin i Denkenberger, 2020; Bengio et al., 2024).

### 3. Treball amb IAs, alineament i riscos existencials o extrems

“Al·legar que els treballadors desplaçats per les màquines trobaran feina, no és molt més sensat que considerar que els cavalls desplaçats per vehicles puguin ser ocupats a la indústria automotriu” (Wassily Leontief, 1973).

Una vegada revisats els principals efectes de les diferents IAs sobre la quantitat, qualitat i retribució del treball, cosa que ens ha permès identificar per a cadascuna d'elles clares tendències diferenciades, a continuació abordarem com aquestes tendències es vinculen amb els principals reptes socials que generen i amb els riscos existencials o extrems per a les persones. Amb aquesta aportació he pretès organitzar l'evidència disponible sobre els diferents tipus de IAs, vincular-les amb el pensament econòmic i estratègic a través dels valors que impulsen i els efectes econòmics que generen, explicar la seva capacitat de transformació i analitzar com aquesta es vincula amb el futur del treball i amb els riscos existencials o extrems d'origen laboral per les persones. La clau de volta de l'anàlisi ha estat que en el context neoliberal o autocràtic predominant on s'insereixen la IAP i la IAG, els incentius per al seu desenvolupament es decanten clarament cap a la seva utilització en processos d'automatització o control. D'aquesta manera, el desenvolupament tecnològic es desalinea de les persones i del floriment social perquè tendeix a empobrir el treball. L'arribada de la IAT podria impulsar un salt de nivell tant pel que fa a les seves implicacions positives per les persones com pel que fa a les negatives. A l'apartat anterior he revisat les extraordinàries implicacions que una IAT que augmenta el treball cognitiu i s'alinea amb el progrés humà podria generar en termes de producció de coneixement, idees, innovació i creixement econòmic. Tanmateix, la mateixa IAT pot ser utilitzada per assolir un creixement econòmic desacoblat que substitueix massivament el treball cognitiu, generant escenaris de desocupació, desqualificació o pèrdua de sentit del treball. Sense dubte, aquest context de desalineament entre els objectius de les persones i les màquines, suposa un risc existencial o extrem per les persones.

Tot i que hi ha un conjunt d'esdeveniments molt negatius que poden suposar riscos greus per al futur de la humanitat, com els vinculats amb l'emergència climàtica, la investigació prospectiva defineix els riscos existencials com a aquells esdeveniments provocats per l'acció humana i que poden culminar si es porten a l'extrem en la seva extinció (Ord, 2020). Segona aquesta literatura, i en contraposició amb el baix interès que té entre economistes, polítics i la societat en general, les estimacions disponibles assenyalen que els riscos que la humanitat no sobrevisqui al segle XXI són elevats, amb una probabilitat del 16,7%. La llista de riscos

existencials o extrems que tenen més probabilitats de materialitzar-se en el futur està encapçalada per l'emergència d'una IAT no alineada, seguida ja a més distància per la guerra nuclear i les pandèmies dissenyades. De fet, la investigació ha observat riscos existencials i extrems diferenciats per a cadascuna de les tipologies i evolucions de la IA, cosa que dóna sentit a un anàlisi diferenciat entre IAP, IAG i IAT (Turchin i Denkenberger, 2020). De la mateixa manera, la investigació de l'àmbit ha estat capaç d'identificar quatre categories de riscos existencials principals vinculats amb els usos de la IA (Hendrycks et al., 2023): 1) ús maliciós, quan la utilització intencionada de la IA és per causar danys (per exemple el bio-terrorisme o l'estat de vigilància i manipulació); 2) carrera d'IA, quan la pressió competitiva dels mercats obliga als agents (empreses i persones) a desplegar IA insegures o a cedir el seu control als algoritmes que aprenen (com el cas de la substitució massiva de treball per l'automatització o el transhumanisme); 3) riscos organitzatius, quan els factors humans i els sistemes complexos de IA, poden augmentar la possibilitat d'accidents catastròfics (per exemple, cultures de seguretat dèbils o sistemes d'IA manipulats); i 4) IA canalla, quan la intenció és desenvolupar i controlar agents molts més intel·ligents que els humans (per exemple, quan la IA reproduïx objectius defectuosos o la cerca de poder). Tenint en compte aquestes aproximacions, a continuació revisaré els riscos existencials amb vincle laboral per a cadascuna de les tres IAs analitzades.

El vincle entre IAP, automatització i augment del treball ens determina dos objectius clars. La IAP generalment fa créixer la productivitat mitjana (producte per tasca), però no té per què fer créixer la productivitat marginal (augment de productivitat per tasca). Perquè la IAP generi més i millor ocupació cal que apareguin noves tasques que justifiquin avenços combinats de productivitat mitjana i marginal, de manera que es justifiquin noves ocupacions i contractacions. I perquè la IAP vagi més enllà dels típics processos d'automatització i costos, i sigui capaç d'augmentar i millorar el treball de manera massiva, li calen nous enfocaments basats en la utilitat de les màquines i la complementarietat d'intel·ligències (Acemoglu i Johnson, 2023). És a dir, redirigir els usos de la IAP cap la utilitat de les persones, i redirigir les habilitats de les persones cap als usos de la IAP. Aquestes és el principal repte social que la IAP té plantejat a l'actualitat si vol evitar el risc extrem de no generar prou destrucció creativa que impulsi la dinàmica econòmica (Aghion et al., 2019).

A banda d'aquestes qüestions també és important tenir en comte els efectes de la IAP sobre la polarització i la desigualtat entre persones i empreses. Com a resultat de les seves economies d'escala i d'àmbit, i del conjunt d'efectes de xarxa, plataforma, biaix i polarització que sustenten la IAP, les empreses poden trobar moments i enfocaments de predicció i construir models de negoci digitals automatitzats (eficients), barats (amb costos marginals de reproducció tendents a zero) i molt escalables (amb rendiments creixents). Aquesta eficiència i escalabilitat a partir de l'analítica i la gestió de grans dades és molt més plausible en les grans empreses que en les petites, i és molt més probable que desplaci als treballadors amb qualificacions baixes i mitjanes (Torrent-Sellens, 2024b). De fet, aquesta polarització i desigualtat és el principal risc existencial que s'associa actualment amb la IAP. És un risc controlat, sobretot si la política pública hi intervé protegint i formant el treball, però també real. Segons les darreres estimacions per la Unió Europea, les innovacions relacionades amb l'automatització augmenten significativament la probabilitat d'atur de llarga durada, mentre que l'adopció de tecnologies digitals d'automatització, entre elles la IAP, augmenta significativament les possibilitats de caure en la inactivitat (Castellani i Lamperti, 2024). A la taula 1 es reproduïxen les principals característiques de les IAs analitzades, així com els principals reptes i riscos existencials i extrems que se les hi associen.

Pel que fa al vincle entre la IAG i el treball, l'evidència relativa a la quantitat de treball ens suggereix que aquest vincle podria ser molt favorable, però també, i en comparació amb la IAP, amb un augment molt notable dels riscos laborals i socials extrems que s'hi vinculen. La IAG i el valor de creació que se l'hi associa es podria constituir en una eina molt vàlida per augmentar el treball, mentre que a l'hora seria capaç d'automatitzar tasques cognitives rutinàries que alliberessin espai per el desplaçament cap a tasques creatives i de més valor. De fet, la IAG podria ser un bon terreny de col·laboració entre persones i màquines, si les primeres poden decidir amb quines tasques col·laboren i amb quines tasques la IAG és molt més efectiva. Tot i amb

això, la investigació sobre els efectes laborals de la IAG és molt emergent. Ens falten encara molts estudis per a confirmar o no aquesta tendència inicial, en especial a mesura que els grans proveïdors de IAG estandarditzin models de IAG més potents de pagament, i traslladin novament la decisió d'adopció de les persones cap a les corporacions, augmentant de nou la desigualtat d'accés i la distribució desigual de beneficis. A més, els efectes de la IAG sobre la qualitat, les retribucions i la salut del treball romanen molt incerts. Pràcticament no disposem d'evidència sobre com són els nous llocs de treball creats amb IAG, ni quins són o seran els seus efectes potencials sobre el canvi d'habilitats, la intensitat laboral i l'autonomia quan aquesta tecnologia es vagi integrant plenament als llocs de treball (Margaryan, 2023).

De moment, el que sí sabem és que l'ocupació que més està utilitzant la IAG, el treball d'oficina, és la que té més risc de substitució i la que presenta més problemes de qualitat i salut laboral, entre altres coses perquè, sota la patina de la generalització de la IA, moltes persones continuen realitzant tasques rutinàries, repetitives o desagradables i en condicions precàries, per exemple en treballs cognitius com la vigilància o control de continguts (Muldoon et al., 2024). Però també sabem que als EUA un 80% dels llocs de treball podran veure exposats al menys un 10% de les seves tasques a la IAG, guanyant eficiència sense perdre qualitat. Seguint aquests projeccions, un 46% dels llocs de treball dels EUA podrien veure afectades el 50% de les seves tasques durant els propers anys (Eloundou et al., 2024). En la mateixa línia se situen les primeres aproximacions realitzades per als països de l'OCDE, amb una exposició mitjana (20% de tasques que es poden fer un 50% més ràpid) a la IAG que podria arribar al 20% dels treballadors (OECD, 2024). De fet, aquesta exposició podria arribar a ser alta (50% de tasques que es poden fer un 50% més ràpid) per un forquilla regional situada entre el 16% i el 70% dels treballadors en cas de continuar la convergència transformadora entre IAP i IAG. En aquest context, una novetat important és que s'accentuen les diferències regionals i dels mercats de treball local perquè la IAG incideix especialment en el treball cognitiu no rutinari i en sectors intensius en coneixement:

“A tota l'OCDE, prop d'una quarta part dels treballadors estan exposats a la IA generativa, cosa que significa que un 20% (o més) de les tasques laborals es podrien fer al menys un 50% més ràpid amb l'ajuda de la IA generativa. L'exposició a la IA seguirà creixent a mesura que es desenvolupi o s'integri nou software amb tecnologies d'IA generativa, i la proporció de treballadors que podrien estar altament exposats (el 50% de les seves tasques podrien fer-se al menys un 50% més ràpid amb la IA generativa) possiblement oscil·larà entre el 16% i més del 70% a les regions de l'OCDE. A diferència de les tecnologies de l'automatització anteriors, la IA generativa es destacarà en la realització de tasques cognitives, no rutinàries, cosa que canvia l'exposició del mercat laboral regional, ja que les regions que concentren sectors com l'educació, les TIC o les finances són les més exposades a la IA generativa (OCDE, 2024, pàg. 12)

De fet, i atesa l'acceleració del canvi laboral induït per la IAG, el repte de la capacitació dels treballadors i els canvis del treball cognitiu és el principal repte social que s'hi vincula. Durant els propers anys, les necessitats de formació, ensinistrament i recapacitació en noves competències i habilitats cognitives que els treballadors necessitaran per interactuar amb la IAG seran enormes. En cas de no aconseguir-ho, els riscos extrems d'una desqualificació massiva de les persones, amb una acceleració de la desigualtat i polarització social, fa créixer la versemblança de conflictes i riscos existencials de tota manera. I anant més enllà del treball, però considerant aspectes socials que s'hi vinculen, la investigació de l'àmbit també ha identificat riscos existencials importants en les desigualtats d'accés a la IAG en l'educació o la salut, a més de ser utilitzada de manera creixent per a la desinformació o la lluita contra drets individuals, civils i polítics (Capraro et al., 2024). En aquest context, la recuperació del valor públic de la IAG, amb el co-disseny i la creació de nous algorismes generatius de codi obert i amb propietat pública de les dades, que complementin les ofertes privades, esdevindrà rellevant per a la construcció d'un progrés social compartit que redueix el risc de col·lisió social amb origen laboral (Margaryan, 2023).

Finalment, la possible irrupció d'una IAT que impulsi un creixement econòmic desacoblat i un progrés social desalineat de les persones augmenta ostensiblement la versemblança de nous riscos extrems i existencials

per a la humanitat. Segons la investigació de l'àmbit, no hi ha gaires dubtes entre els investigadors i desenvolupadors de IA que tècnicament l'emergència d'una IAT multi-agent, molt capaç i amb habilitats per a emular la majoria de competències humanes necessàries per la dinàmica econòmica, especialment la producció d'idees i la generació d'innovacions, és possible. De fet, amb la irrupció dels grans models del llenguatge i la IAG, les estimacions de l'arribada de la IAT han baixat uns deu anys de mitjana, i a l'actualitat se situen a la dècada del 2030 i principis del 2040 (Grace et al., 2024; Roser, 2024). La clau de volta de l'emergència tècnica de la IAT està molt vinculada amb el que s'anomena la hipòtesi d'escala. És a dir, que l'augment de la capacitat (dades, capes, neurones i paràmetres) dels algoritmes actuals d'aprenentatge profund, en especial les xarxes neuronals convolucionals o les xarxes generatives inverses, els faci capaços d'avançar en l'aprenentatge significatiu i de generar noves propietats transformatives (Bubeck et al., 2023). A més a més, encara que l'aprenentatge profund necessiti replantejaments de disseny, no hi ha lleis físiques (i encara menys des de l'aparició de la computació quàntica) que impedeixin l'optimització constant de les xarxes neuronals profundes perquè segueixen preveient i creant tota mena d'artefactes de manera molt més ràpida i eficient que el cervell humà.

L'altre element rellevant per l'emergència d'una IAT amb riscos existencials és el desalineament amb el benestar de les persones i el floriment social. De fet, si no som capaços de resoldre el problema de la desalineació, l'emergència de riscos extrems i existencials és pot donar per segura (Bostrom, 2014; Grawiec, 2024). A l'actualitat tenim moltes proves de la disfunció humana i del creixent nombre de riscos extrems vinculats amb la IAP i la IAG, encara que en termes socials i polítics se'ls hi presti poca atenció (Turchin i Denkenberger, 2020; Gruetzemacher i Whittlestone, 2022). Es produeixen bàsicament amb criteris tècnics i d'optimització per part de col·lectius de professionals de les ciències de la computació i les dades. Aquesta aproximació basada en criteris de sol·lucionisme tecnològic genera un ampli ventall de biaixos i una baixa presència de l'ètica social. S'apliquen i es desenvolupen amb criteris d'automatització i control, segueixen models de negoci que creen riquesa però que també busquen rendes, estan en mans d'un grapat de grans empreses *superstars* tecnològiques, substitueixen un nombre creixent d'habilitats humanes, i tenen uns usos molt heterogenis i dispars entre persones, empreses, col·lectius i territoris (Jones, 2023; Torrent-Sellens, 2024b). Si a aquestes disfuncions de base hi afegim la possibilitat tècnica de l'emergència d'una IAT i el seu valor de transformació associat, que siguin capaços de superar la humanitat "només" en la producció d'idees i la generació d'innovacions, desalineant i desacoblant el creixement econòmic de les persones, els riscos per a la humanitat serien enormes. L'emergència d'una economia de les màquines basada en la seva capacitat superior a les persones per a predir, crear i transformar la riquesa, a banda de que és un risc existencial per sí mateix, tindria uns efectes laborals i socials mai vistos. Per començar, la necessitat de redefinir què és el treball i com s'assignen persones i màquines en un context de depreciació del treball humà, quins són els seus incentius i resultats, i com s'articula socialment (Korinek i Juelfs, 2022). I per acabar, la necessitat de definir la vida individual i social en un context transhumanista de fusió persona-màquina, o de gestionar l'emergència de contextos post-antropocè amb la pèrdua de privilegis per part de l'espècie humana enfront una espècie de màquines superiors (Belk et al., 2020). No cal ni plantejar escenaris de singularitat o de consciència artificial. A través de l'economia, les màquines podrien colonitzar la humanitat i fer-la redundant o supèrflua en moltes, sinó en totes, les altres dimensions de la vida, en particular en la dimensió laboral. I, com que, les societats s'articulen a partir del treball, això suposaria un terrabastall sense precedents.

És obvi que és urgent començar a organitzar i gestionar aquests riscos, doncs han deixat de ser ciència ficció tal i com mostra l'alerta creixent de la comunitat científica (Bengio et al., 2024) inclosos molts dels economistes de la tecnologia (Jones, 2023; Eloundou et al. 2024). De fet, ja comencem a disposar d'uns primers models que suposen capacitat d'aprenentatge social sobre els riscos existencials, i intenten compaginar la protecció dels punts forts de la IAT, en especial la seva capacitat per impulsar salts de nivell de la productivitat i el creixement econòmic, i per a frenar els seus usos socialment inadequats. Les calibracions generades ens confirmen que l'elecció socialment òptima de la IAT és gradual i convexa, de manera que si l'adopció de la

IAT genera danys socials proporcionals als guanys de productivitat, el creixement s'ha d'acompassar al ritme d'adopció. Però, aquest equilibri és inestable quan les empreses no internalitzen les externalitats dels costos socials que generen i quan la regulació independent no és suficient per a garantir la optimalitat (Acemoglu i Lensman, 2024). De fet, el que ens suggereix aquest model i d'altres investigacions abordades amb una aproximació més social i laboral (Gmyrek et al., 2023; Saetra, 2023) és que la governança dels riscos econòmics i laborals extrems de la IAT depèn molt de com s'abordi la seva adopció i necessita fer-se ja (Grace et al., 2024; Roser, 2024). I fer-se amb una aproximació integral, d'una economia política total que tracti totes i cadascuna de les dimensions tècniques, de comportament, econòmiques, històriques, culturals, socials, institucionals i polítiques, i tots els nivells (micro, meso i macro) de la vida digital de les persones. Qui avisa, no és traïdor. Continuarà...

**Taula 1.** IAs i treball: característiques, reptes socials i riscos existencials

Tipus d'IA	Característiques	Reptes socials	Riscos existencials
<b>IAP: IA Predictiva</b>	Adopció per decisió corporativa  Valor de predicció  Complementarietats persones, organització i IA  Usos d'automatització i control  Augments de la productivitat mitjana i del treball qualificat	Augment sostingut de la productivitat marginal  Capacitat per a crear noves tasques i ocupacions  Redirecció de la IAP cap a les necessitats de les persones i les empreses  Redirecció de les persones i les empreses cap al treball híbrid amb la IAP	Substitució massiva de tasques rutinàries  Avenç de la polarització laboral i la desigualtat salarial  Polarització empresarial: poder de mercat i polític de les grans superstars
<b>IAG: IA Generativa</b>	Adopció per decisió individual  Valor de creació  Complementarietats persona-IA (centaures vs cyborgs)  Usos d'augment del treball  Augments de rapidesa, qualitat, i productivitat tasques	Reconduir la distribució molt desigual dels beneficis de la IAG  Entomar la gran recapacitació del treball cognitiu  Augmentar les competències STEM i socials de les persones per interactuar amb tota la cadena d'IAG  Construcció d'IAGs de valor públic	Empobriment de la qualitat del treball  Polarització i desigualtat entre persones i empreses  Extensió de la desigualtat d'accés a altres àmbits socials (educació i salut)  Polarització política
<b>IAT: IA Transformativa</b>	Adopció per decisió individual, corporativa i social  Valor de transformació  Complementarietats persona-organització-societat-IA  Usos de transformació social	IAT ètica, inclusiva i responsable al nivell personal i social  Economia política de la producció i l'adopció de la IAT  Grans avantatges amb riscos extrems  Gestió del creixement desacoblat	Redefinició del treball i del seu paper econòmic i social  Gestió de la pèrdua de valor econòmic i laboral de les persones  Control del poder social i polític de la IAT  Preparar-se per la vida transhumana i el post-antropocè

Font: Elaboració pròpia.

## Conclusió

La contribució d'aquest article ha estat organitzar i aportar noves perspectives econòmiques sobre les tipologies d'IA, el futur del treball i els riscos existencials i extrems que s'hi vinculen. He conclòs que a mesura que augmenta de capacitats i va aprenent, la IA es va apoderant per a emular un nombre cada vegada més creixent d'habilitats humanes, i accelerant el seu potencial de substitució del treball cognitiu no rutinari de les persones. Tècnicament no hi ha dubte que l'emergència d'una IA transformadora, que sigui capaç de produir



idees, generar innovacions i desacoblar el creixement econòmic de les persones, serà un fet a les properes dècades. Per si mateixa aquesta IA transformadora suposa un risc existencial perquè, a través de la seva superioritat per activar els valors de predicció, creació i transformació, s'apoderaria de l'avantatge econòmic per sobre les persones, fent-les redundants en altres dimensions de la vida social, com la laboral. Aquesta capacitat acabaria per desalinejar totalment la progressió de la IA de les persones, ja en l'actualitat molt desenfocada resultat dels seus usos molt desiguals, en mans de poques corporacions i amb propòsits bàsics d'automatització i control. Però, si fóssim capaços d'alinejar i fer treballar la IA transformativa a favor de les persones, podríem acabar impulsant una nova era de creixement i floriment social, gràcies a la seva capacitat per a impulsar salts de nivell de la productivitat, el creixement econòmic, el benestar social i la protecció del medi ambient.

Per a perseguir aquest futur tant esperançador i brillant, i atesa l'acceleració de les tecnologies intel·ligents, és molt important el que comencem a fer ja des d'ara mateix. A través del dilema de Collingridge, segons el qual és molt més fàcil regular la tecnologia en la seva infància, precisament quan és més difícil saber quins seran els seus efectes (Collingridge, 1980; Engström i Strimling, 2020), i de la dependència de la trajectòria, segons la qual una vegada les empreses han invertit i après sobre una matriu tecnològica és molt difícil que canviïn de tecnologia (Arthur, 1994; Aghion et al., 2019), sabem que molt probablement l'alineació de la IA amb les persones i les societats humanes està a les nostres mans si actuem amb determinació i celeritat. Cal doncs articular polítiques públiques que prioritzin la investigació sobre els riscos existencials i la seguretat de la IA. Sortosament, no partim des de zero. Tenim l'experiència històrica de les dues revolucions industrials i de gairebé quatre dècades de paradigma digital amb orientació neoliberal que ens ha ensenyat els punts forts i febles del capitalisme digital (Acemoglu and Johnson, 2024). Hem après de la rellevància de la productivitat marginal, la destrucció creativa i les complementaritats entre tecnologia, persones, organitzacions i institucions (Torrent-Sellens 2024a). És veritat que repte actual no té parangó perquè mai la humanitat havia hagut de gestionar una tecnologia amb un potencial cognitiu tant gran, ni amb riscos existencials tan rellevants. Però si com capaços de plantejar una economia política de la IA transformativa que contempli tots els elements individuals, grupals, racionals, emocionals, ètics, culturals, institucionals, jurídics, històrics i polítics que expliquen la presa de decisions i l'acció econòmica amb presència de tecnologia a tots els nivells, i a més ho fem combinant ciència positiva amb finalitats (equitat, biaix, privadesa, seguretat) i objectius normatius clars (per exemple en termes de llibertat, democràcia, sostenibilitat, benestar i justícia), les probabilitats d'èxit per alinear tecnologia i societat augmenten ostensiblement (Brynjolfsson i Mitchell, 2024). La història no es repeteix, però rima. Leontieff deia que si els cavalls haguessin estat persones i haguessin pogut votar, igual la tracció equina hagués tingut un final diferent. El futur no està escrit.

## Bibliografia

- ACEMOGLU, D., LENSMAN, T. (2024). <<Regulating transformative technologies>>, *American Economic Review: Insights* (vol. 6, núm. 3, pàg. 359-376).
- ACEMOGLU, D., JOHNSON, S. (2024). <<Learning from Ricardo and Thompson: Machinery and labor in the early Industrial Revolution and in the age of artificial intelligence>>, *Annual Review of Economics* (vol. 16, núm. 1, pàg. 597-621).
- ACEMOGLU, D., AUTOR, D.A., JOHNSON, S. (2023). *Can we have pro-worker AI? Choosing a path of machines in service of minds*. Cambridge, MA: MIT Shaping the Future of Work Initiative. Disponible a: <https://shapingwork.mit.edu/wp-content/uploads/2023/09/Pro-Worker-AI-Policy-Memo.pdf>.
- ACEMOGLU, D.; JOHNSON, S. (2023). *Poder y progreso. Nuestra lucha milenaria por la tecnología y la prosperidad*. Barcelona: Deusto / Planeta.
- ACEMOGLU, D.; RESTREPO, P. (2022). <<Tasks, automation, and the rise in US wage inequality>>, *Econometrica* (vol. 90, núm. 5, pàg. 1973-2016).
- ACEMOGLU, D.; RESTREPO, P. (2019), <<Automation and new tasks: how technology displaces and reinstates labor>>, *Journal of Economic Perspectives* (vol. 33, núm. 2, pàg. 3-30).

- AGHION, P., BERGEAUD, A., BOPPART, T., KLENOW, P.J., LI, H. (2019). <<Missing growth from creative destruction>>, *American Economic Review* (vol. 109, núm. 8, pàg. 2795-2822).
- AGRAWAL, A., GANS, J., GOLDFARB, A. (2022a). *Prediction machines. The simple economics of artificial intelligence (updated and expanded)*. Cambridge, MA: Harvard Business Press.
- AGRAWAL, A., GANS, J., GOLDFARB, A. (2022b). <<From prediction to transformation>>, *Harvard Business Review* (vol. 100, num. 11/12, pàg. 100-109).
- ALEKSEEVA, L.; AZAR, J.; GINÉ, M.; SAMILA, S.; TASKA, B. (2021). <<The demand for AI skills in the labor market>>, *Labour Economics* (vol. 71, article 102002).
- ALMEIDA, D., NAUDÉ, W., SEQUEIRA, T.N. (2024). <<Artificial intelligence and the discovery of new ideas: Is an economic growth explosion imminent?>> *IZA Discussion Papers* (núm. 16766).
- ARTHUR, W.B. (1994). *Increasing returns and path dependence in the economy*. Ann Arbor: University of Michigan Press.
- AUTOR, D.H., CHIN, C., SALOMONS, A., SEEGMILLER, B. (2024). <<New frontiers: The origins and content of new work, 1940–2018>>, *Quarterly Journal of Economics* (vol. 139, núm. 3, pàg. 1399-1465).
- AUTOR, D.; DORN, D.; KATZ, L.F.; PATTERSON, C.; VAN REENEN, J. (2020). <<The fall of the labor share and the rise of superstar firms>>, *Quarterly Journal of Economics* (vol. 135, núm. 2, pàg. 645-709).
- AUTOR, D.H., LEVY, F., MURNANE, R.J. (2003). <<The skill content of recent technological change: An empirical exploration>>, *Quarterly Journal of Economics* (vol. 118, núm. 4, pàg. 1279-1333).
- BALLESTAR, M.T., DÍAZ-CHAO, Á., SAINZ, J., TORRENT-SELLENS, J. (2021). <<Impact of robotics on manufacturing: A longitudinal machine learning perspective>>, *Technological Forecasting and Social Change* (vol. 162, article 120348).
- BALLESTAR, M.T., CAMIÑA, E., DÍAZ-CHAO, Á., TORRENT-SELLENS, J. (2021b). <<Productivity and employment effects of digital complementarities>>, *Journal of Innovation & Knowledge* (vol. 6, núm. 3, pàg. 177-190).
- BALLESTAR, M.T., DÍAZ-CHAO, Á., SAINZ, J., TORRENT-SELLENS, J. (2020). <<Knowledge, robots and productivity in SMEs: Explaining the second digital wave>>, *Journal of Business Research* (vol. 108, pàg. 119-131).
- BALSMEIER, B., WOERTER, M. (2019). <<Is this time different? How digitalization influences job creation and destruction>>, *Research Policy* (vol. 48, núm. 8, article 103765).
- BELK, R.W., HUMAYUM, M., GOPALDAS, A. (2020). <<Artificial life>>, *Journal of Macromarketing* (vol. 40, núm. 2, pàg. 221-236).
- BENGIO, Y., HINTON, G., YAO, A., SONG, D., ABBEEL, P., DARRELL, T., ... MINDERMANN, S. (2024). <<Managing extreme AI risks amid rapid progress>>, *Science* (vol. 384, núm. 6698, pàg. 842-845).
- BESSEN, J.; GOOS, M.; SALOMONS, A.; VAN DEN BERGE, W. (2023). <<What happens to workers at firms that automate?>>, *Review of Economics and Statistics*. Doi: [10.1162/rest\\_a\\_01284](https://doi.org/10.1162/rest_a_01284).
- BLOOM, N., JONES, C.I., VAN REENEN, J., WEBB, M. (2020). <<Are ideas getting harder to find?>>, *American Economic Review* (vol. 110, núm. 4, pàg. 1104-1144).
- BONNEY, K., BREAUX, C., BUFFINGTON, C., DINLERSOZ, E., FOSTER, L., GOLDSCHLAG, N., HALTIWANGER, J., KROFF, Z., SAVAGE, K. (2024). <<The impact of AI on the workforce: Tasks versus jobs?>>, *Economics Letters* (vol. 244, art. 111971).
- BOSTROM, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford: Oxford University Press.
- BRYNJOLFSSON, E., MITCHEL, M. (2024). *Intelligence and the future of work*. Washington: The National Academies Press. Doi: [10.17226/27644](https://doi.org/10.17226/27644).
- BRYNJOLFSSON, E.; ROCK, D.; SYVERSON, Ch. (2021). <<The productivity J-curve: how intangibles complement general purpose technologies>>, *American Economic Journal: Macroeconomics* (vol. 13, núm. 1, pàg. 333-372).
- BRYNJOLFSSON, E.; ROCK, D.; SYVERSON, Ch. (2019). <<Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics>>. A: AGRAWAL, A., GANS, J., GOLDFARB, A. (Eds.). *The economics of artificial intelligence: An agenda*. Chicago: University of Chicago Press (pàg. 23-57).
- BUBECK, S., CHANDRASEKARAN, V., ELKAN, R., GEHRKE, J., HORVITZ, E., KAMAR, E., ... ZHANG, Y. (2023). <<Sparks of artificial general intelligence: Early experiments with gpt-4>>, *ArXiv Preprint ArXiv* (núm. 2303.12712).
- BUGHIN, J. (2024). <<What drives the corporate payoffs of using generative artificial intelligence?>>, *Structural Change and Economic Dynamics* (vol. 71, pàg. 658-668).
- CAMIÑA, E., DÍAZ-CHAO, A., TORRENT-SELLENS, J. (2020). <<Automation technologies: Long-term effects for Spanish industrial firms>>, *Technological Forecasting and Social Change* (vol. 151, núm. 119828).
- CAPRARO, V., LENTSCH, A., ACEMOGLU, D., AKGUN, S., AKHMEDOVA, A., BILANCHINI, E., .. VIALEC, R. (2024). <<The impact of generative artificial intelligence on socioeconomic inequalities and policy making>> *PNAS Nexus* (vol. 3, núm. 6, pàg. 191-210).

- CARD, D., DiNARDO, J.E. (2002). <<Skill-biased technological change and rising wage inequality: Some problems and puzzles>>, *Journal of Labor Economics* (vol. 20, núm. 4, pàg. 733-783).
- CAROLI, E., VAN REENEN, J. (2001). <<Skill-biased organizational change? Evidence from a panel of British and French establishments>>, *Quarterly Journal of Economics* (vol. 116, núm. 4, pàg. 1449-1492).
- CASTELLANI, D., LAMPERTI, F. (2024). *Aggregate megatrends and the risk of labour market exclusion across Europe*. Brussels: European Union and European Research Executive Agency. Disponible a la web: <https://projectwelar.eu/wp-content/uploads/2024/11/D4.1-Aggregate-Megatrends-and-the-Risk-of-Labour-Market-Exclusion-Across-Europe.pdf>.
- CNOSSSEN, F. (2024). <<Tasks, wages and new technologies>>, *Industrial Relations*. Doi: [10.1111/irel.12380](https://doi.org/10.1111/irel.12380)
- COLLINGRIDGE, D. (1980). *The social control of technology*. Londres: Frances Pinter.
- CRAFTS, N. (2021). <<Artificial intelligence as a general-purpose-technology: An historical perspective>>, *Oxford Review of Economic Policy* (vol. 37, núm. 3, pàg. 521-536).
- DELL'ACQUA, F., McFOWLAND III, E., MOLLICK, E.R., LIFSHITZ-ASSAF, H., KELLOGG, K., RAJENDRAN, S., KRAYER, L., CANDELON, F., LAKHANI, K.R. (2023). <<Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality>>, *Harvard Business School Technology & Operations Management Unit Working Paper* (núm. 24-013).
- DWIVEDI, Y.K., KSHETRI, N., HUGHES, L., SLADE, E.L., JEYARAJ, A., KAR, A.K., ... WRIGHT, R. (2023). <<"So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy>>, *International Journal of Information Management* (vol. 71, article 102642).
- ECKHOUT, J. (2022). *La paradoja del beneficio. Cómo las empresas exitosas amenazan la economía*. Barcelona: Taurus / Penguin Random House.
- ELLSTRÖM, D., HOLSTRÖM, J., BERG, E. (2022), <<Dynamic capabilities for digital transformation>>, *Journal of Strategy and Management* (vol. 15, pàg. 272-286).
- ELOUNDOU, T., MANNING, S., MISHKIN, P., ROCK, (2024). <<GPTs are GPTs: Labour market impact potential of LLMs>>, *Science* (vol. 384, núm. 6702, pàg. 1306-1308).
- ENGSTRÖM, E., STRIMLING, P. (2020). <<Deep learning diffusion by infusion into preexisting technologies. Implications for users and society at large>>, *Technology in Society* (vol. 63, article 101396).
- FEUERRIEGEL, S., HARTMANN, J., JANIESCH, C., ZSCHECH, P. (2024). <<Generative AI>>, *Business & Information Systems Engineering* (vol. 66, núm. 1, pàg. 111-126).
- FLETCHER, R., NIELSEN, R.K. (2024). What does the public in six countries think of generative AI in news? Oxford: Reuters Institute for the Study of Journalism. Doi: [10.60625/risj-4zb8-cq87](https://doi.org/10.60625/risj-4zb8-cq87).
- FONTANELLI, L., CALVINO, F., CRISCUOLO, C., NESTA, L., VERDOLINI, E. (2024). <<The role of human capital for AI adoption: evidence from French firms>>, *SSRN* (núm. 5010602).
- FREY, C.B., OSBORNE, M.A. (2017). <<The future of employment: How susceptible are jobs to computerisation?>>, *Technological Forecasting and Social Change* (vol. 114, pàg. 254-280).
- GEORGIEFF, A. (2024). <<Artificial intelligence and wage Inequality>>, *OECD Artificial Intelligence Papers* (núm. 13).
- GIRÓN-SIERRA, J.M. (2023). *Introducción a la inteligencia artificial. La tecnología que nos cambiará para siempre*. Còrdova: Sekoitia / Almuzara.
- GMYREK, P., BERG, J., BESCOND, D. (2023). <<Generative AI and jobs: A global analysis of potential effects on job quantity and quality>>, *International Labor Organization Working Paper* (núm. 96). Doi: [10.54394/FHEM8239](https://doi.org/10.54394/FHEM8239).
- GOLDFARB, A., TASKA, B., TEODORIDIS, F. (2023). <<Could machine learning be a general purpose technology? A comparison of emerging technologies using data from online job postings>>, *Research Policy* (vol. 52, núm. 1, article 104653).
- GRACE, K., STEWART, H., SANDKÜLER, J.F., THOMAS, S., WEINSTEIN-RAUN, B., BRAUNER, J. (2024). <<Thousands of AI authors on the future of AI>>, *ArXiv Preprint ArXiv* (núm. 2401.02843).
- GROWIEC, J. (2024). <<Existential risk from a transformative AI: An economic perspective>>, *Technological and Economic Development of Economy* (vol. 30, núm. 6, pàg. 1682-1708).
- GROWIEC, J. (2022). <<Automation, partial and full>>, *Macroeconomic Dynamics* (vol. 26, núm. 7, pàg. 1731-1755).
- GRUETZEMACHER, R., WHITTLESTONE, J. (2022). <<The transformative potential of artificial intelligence>>, *Futures* (núm. 132, art. 102884).
- HENDRYCKS, D., MAZEIKA, M., WOODSIDE, T. (2023). <<An overview of catastrophic AI risks>>, *ArXiv Preprint ArXiv* (núm. 2306.12001).
- JONES, C.I. (2023). <<The A.I. dilemma: Growth versus existential risk>>, *National Bureau of Economic Research Working Paper* (núm. 31837).

- KARIEL, J. (2021). <<Job creators or job killers? Heterogeneous effects of industrial robots on UK employment>>, *Labour* (vol. 35, núm. 1, pàg. 52-78).
- KLENERT, D., FERNÁNDEZ-MACÍAS, E., ANTÓN, J.I. (2023). <<Do robots really destroy jobs? Evidence from Europe>>, *Economic and Industrial Democracy* (vol. 44, núm. 1, pàg. 280-316).
- KOCH, M., MANUYLOG, I., SMOLKA, M. (2021). <<Robots and firms>>. *Economic Journal* (vol. 131, pàg. 2553-2584).
- KORINEK, A. (2023). <<Language models and cognitive automation for economic research>>, *National Bureau of Economic Research Working Paper* (núm. w30957).
- KORINEK, A., JUELFIS, M. (2022). <<Preparing for the (non-existent?) future of work>>. A: Bullock, J.B. (Ed.). *The Oxford Handbook of AI governance*. Oxford: Oxford University Press, pp. 746-776.
- KOSINSKI, M. (2024). <<Evaluating large language models in theory of mind tasks>>, *Proceedings of the National Academy of Sciences* (vol. 121, núm. 45, article e2405460121).
- LARSON, E. J. (2023). *El mito de la inteligencia artificial. Por qué las máquinas no pueden pensar como nosotros lo hacemos*. Barcelona: Shackleton. Books.
- LÓPEZ DE MÁNTARAS I BADIA, R. (2023). *100 coses que cal saber sobre intel·ligència artificial*. Valls: Cossetània edicions.
- MARGARYAN, A. (2023). <<Artificial intelligence and skills in the workplace: An integrative research agenda>>, *Big Data & Society* (vol. 10, núm. 2, article 20539517231206804).
- MAZZUCATO, M. (2019). *El valor de las cosas. Quién produce y quien gana en la economía global*. Barcelona: Taurus / Penguin Random House.
- MULDOON, J., GRAHAM, M., CANT, C. (2024). *Feeding the machine. The hidden human labour powering AI*. Edinburgh: Canongate Books.
- NORDHAUS, W.D. (2021). <<Are we approaching an economic singularity? Information technology and the future of economic growth>>, *American Economic Journal: Macroeconomics* (vol. 13, núm. 1, pàg. 299-332).
- OECD (2024). *Job creation and local economic development 2024: The geography of generative AI*. París: OECD Publishing. Doi: [10.1787/83325127-en](https://doi.org/10.1787/83325127-en).
- ORD, T. (2020). *The precipice: Existential risk and the future of humanity*. París: Hachette.
- PIKETTY, T. (2019). *Capital e ideología*. Barcelona: Deusto / Planeta.
- PIVA, M., SANTARELLI, E., VIVARELLI, M. (2005). <<The skill bias effect of technological and organisational change: Evidence and policy implications>>, *Research Policy* (vol. 34, núm. 2, pàg. 141-157).
- PRATT, G.A. (2015). <<Is a Cambrian explosion coming for robotics?>>, *Journal of Economic Perspectives* (vol. 29, núm. 3, pàg. 51-60).
- ROSER, M. (2024). <<AI timelines: What do experts in artificial intelligence expect for the future?>>, *Our World in Data*. Oxford: Oxford Martin School. Disponible a: <https://ourworldindata.org/ai-timelines>.
- SAETRA, H.S. (2023). <<Generative AI: Here to stay, but for good?>>, *Technology in Society* (vol. 75, article 102372).
- SULEYMAN, M., BHASKAR, M. (2023). *La ola que viene. Tecnología, poder y el gran dilema del siglo XXI*. Barcelona: Penguin Random House.
- TEECE, D.J. (2014). <<The foundations of enterprise performance: dynamic and ordinary capabilities in an (economic) theory of firms. *Academy of Management Perspectives* (vol. 28, núm. 4, pàg. 328-352).
- TEECE, D.J., PISANO, G., SHUEN, A. (1997). <<Dynamic capabilities and strategic management>>, *Strategic Management Journal* (vol. 18, núm. 7, pàg. 509-533).
- TORRALBA, F. (2022). *L'ètica algorítmica*. Barcelona: Edicions 62.
- TORRENT-SELLENS, J. (2024a). <<Homo digitalis: narrative for a new political economy of digital transformation and transition>>, *New Political Economy* (vol. 29, núm. 1, pàg. 125-143).
- TORRENT-SELLENS, J. (2024b). <<Digital transition, data-and-tasks crowd-based economy, and the shared social progress: unveiling a new political economy from a European perspective>>, *Technology in Society* (vol. 79, article 102739).
- TORRENT-SELLENS, J. (2024c). <<Los límites de la gestión desadherida del conocimiento. El caso de la UOC>>, a TORRENT-SELLENS, J. (2024). *Capital, trabajo y valor en la sociedad digital de mercado. Ensayos para una economía más sabia y sostenible en el siglo XXI*. València: Tirant lo Blanch, pàg. 317-376.
- TORRENT-SELLENS, J., DÍAZ-CHAO, A., MIRÓ-PÉREZ, A.P., SAINZ, J. (2022b). <<Towards the Tyrell corporation? Digitisation, firm-size and productivity divergence in Spain>>, *Journal of Innovation & Knowledge* (vol. 7, núm. 2, article 100185).

- TORRENT-SELLENS, J. (2022a). <<Digitalitzats i digitalitzades del món, a les xarxes! De la quarta i abundant revolució industrial a la ruptura digital i l'escassetat superstar>>, *Oikonomics. Revista d'Economia, Empresa i Societat* (núm. 17, març, pàg. 1-12).
- TORRENT-SELLENS, J., FICAPAL-CUSÍ, P. (2009). *TIC, conocimiento, redes y trabajo*. Barcelona: Edicions de la Universitat Oberta de Catalunya.
- TORRENT-SELLENS, J. (2008). <<Canvi tecnològic digital esbiaixador d'habilitats (e-SBTC), ocupació i salaris: Un estat de la qüestió>>, *UOC Papers, Revista sobre la Societat del Coneixement* (núm. 6, març, pàg. 1-14).
- TORRENT-SELLENS, J. (2004). *Innovació tecnològica, creixement econòmic i economia del coneixement*. Barcelona: Consell de Treball, Econòmic i Social de Catalunya.
- TORRES, J. (2023). *La intel·ligència artificial explicada als humans*. Barcelona: Plataforma editorial.
- TRAMMELL, P., KORINEK, A. (2023). <<Economic growth under transformative AI>>, *National Bureau of Economic Research Working Paper* (núm. w31815).
- TURCHIN, A., DENKENBERGER, D. (2020). <<Classification of global catastrophic risks connected with artificial intelligence>>, *AI & Society* (vol. 35, núm. 1, pàg. 147-163).
- VIÑAS-BARDOLET, C.; TORRENT-SELLENS, J.; GUILLEN-ROYO, M. (2020). <<Knowledge workers and job satisfaction: evidence from Europe>> *Journal of the Knowledge Economy* (vol. 11, pàg. 256-280).
- VIRGILIO, G.P.M., SAAVEDRA-HOYOS, F., BAO-RATZEMBERG, C.B. (2024). <<The impact of artificial intelligence on unemployment: A review>>. *International Journal of Social Economics*. Doi: [10.1108/IJSE-05-2023-0338](https://doi.org/10.1108/IJSE-05-2023-0338).
- VU, K., HANAFIZADEH, P., BOLIN, E. (2020). <<ICT as a driver of economic growth: A survey of the literature and directions for future research>>, *Telecommunications Policy* (vol. 44, núm. 2, article 101922).
- WANG, S., COOPER, N., EBY, M. (2024). <<From human-centered to social-centered artificial intelligence: Assessing ChatGPT's impact through disruptive events>>, *Big Data and Society*. Doi: [10.1177/20539517241290220](https://doi.org/10.1177/20539517241290220).
- WARNER, K.S.R., WÄGER, M. (2019). <<Building dynamic capabilities for digital transformation: An ongoing process of strategic renewal>>, *Long Range Planning* (vol. 52, pàg. 326-349).