



INTEL·LIGÈNCIA ARTIFICIAL GENERATIVA. OPORTUNITATS, RISCOS I REPTES EN POLÍTIQUES PÚBLIQUES

Informe EPTA 2023

Aquest informe ha estat elaborat per iniciativa del CAPCIT, que ha exercit la presidència de l'EPTA el 2023. Les aportacions les ha sintetitzat l'equip editorial: Bjørn Bedsted i Nicklas Bang Bådum (DBT, Dinamarca), Reinhard Grünwald i Steffen Albrecht (TAB, Alemanya), Tore Tennøe (NBT, Noruega), i Ferran Domínguez i Clara Marsan (CAPCIT, Catalunya).

Publicat per:

European Parliamentary Technology Assessment

(EPTA) eptanetwork.org

Barcelona, octubre del 2023

Disponible en línia a:

Parlament de Catalunya. Consell Assessor del Parlament sobre Ciència i Tecnologia (CAPCIT)

www.parlament.cat/capcit

INTEL·LIGÈNCIA ARTIFICIAL GENERATIVA. OPORTUNITATS, RISCOS I REPTES EN POLÍTIQUES PÚBLIQUES

Informe EPTA 2023

Taula de continguts

Prefaci	7
Resum executiu	9
Per què és rellevant parlar d'intel·ligència artificial (IA) generativa?	9
Com aborda l'avaluació tecnològica la IA generativa?	10
Què pensen els parlamentaris sobre la IA generativa?	11
IA generativa i democràcia	13
IA generativa i salut	14
IA generativa i educació	16
IA generativa i treball	17
Perspectives	19
Àustria	20
La IA a Àustria	20
Rellevància i debat social i polític	22
El paper de l'AT en els debats	22
Bibliografia	23
Catalunya	25
Consell Assessor del Parlament sobre Ciència i Tecnologia (CAPCIT)	25
Intel·ligència artificial generativa per a l'educació: el cas de Catalunya	25
Dinamarca	29
Intel·ligència artificial generativa i democràcia: diputats danesos i una perspectiva de país	29
Problema(es) en joc	29
Reaccions de la societat sobre IA generativa i democràcia	30
El paper de l'AT en el debat danès sobre IA generativa i democràcia	31
Parlament Europeu	33
Intel·ligència artificial generativa. Oportunitats, riscos i reptes polítics	33
Descripció del sistema/problema en qüestió	33
Rellevància i debat social i polític	35
El paper de l'AT en els debats	37
França	39
Informe EPTA 2023 sobre intel·ligència artificial generativa	39
Contribució de l'Oficina Parlamentària d'Avaluació Científica i Tecnològica (OPECST) de França	39
Alemanya	41
ChatGPT es troba amb reaccions diverses	41
Adopció i rellevància econòmica de la IA (generativa) a Alemanya	41
Un debat viu davant de moltes incògnites	42
Què hi ha en joc en la política i en els diferents àmbits polítics?	44
Referències	45
Grècia	47
Comissió parlamentària grega sobre avaluació tecnològica	47
Intel·ligència artificial i ocupació	47

Japó	51
Contribució de la Biblioteca de la Dieta Nacional a l'informe de l'EPTA del 2023	51
Intel·ligència artificial generativa. Oportunitats, riscos i reptes polítics	51
 Lituània	 56
IA generativa i treball	56
 Països Baixos	 62
Intel·ligència artificial generativa: oportunitats, riscos i reptes polítics	62
Descripció del sistema/problema en qüestió	62
Rellevància i debat social i polític	63
Paper de l'AT en els debats	65
Referències	66
 Noruega	 68
Consell Noruec de Tecnologia (Teknologirådet)	68
Desenvolupaments clau sobre IA generativa i democràcia a Noruega	68
 Portugal	 74
Intel·ligència generativa artificial i treball	74
Bibliografia	81
 Suècia	 83
Antecedents de la conferència de l'EPTA sobre IA generativa i democràcia	83
 Suïssa	 86
IA generativa i democràcia a Suïssa	86
 Regne Unit	 90
Informe EPTA 2023 - Oficina Parlamentària de Ciència i Tecnologia	90
La IA generativa i el mercat laboral del Regne Unit i un escrutini parlamentari més ampli de la IA	90

Prefaci

Distingits lectors,

La intel·ligència artificial (IA) generativa és un subconjunt o branca especialitzada de la IA que s'ocupa específicament de la generació de continguts o dades que sovint són creatius, semblants als humans o originals. Els sistemes d'IA generativa estan dissenyats per a crear continguts nous, com ara text, imatges, música o vídeo, que poden imitar la creativitat humana o generar continguts contextualment adequats. Com a conseqüència d'això, les tecnologies d'IA generativa han suscitat una gran atenció els darrers anys a causa de la seva capacitat per a crear continguts realistes i sovint convincents en diversos àmbits creatius.

Les tecnologies d'IA continuen avançant i afectant diverses esferes de la societat, com ara l'economia, la privadesa, la seguretat o la seva dimensió ètica. En aquest context, els parlaments estan cridats a jugar un paper fonamental en la navegació pel complex panorama de la IA. Els òrgans elegits democràticament han de prendre la iniciativa en la regulació d'aquesta qüestió tan candent. En fer-ho actuen com la veu de la ciutadania i se'ls confia l'elaboració i l'aprovació de lleis que puguin abordar amb eficàcia els riscos relacionats amb la IA.

A més, els parlaments ofereixen una plataforma per a debats informats, inclusivament i transparents, que permeten als experts, a les parts interessades i als ciutadans contribuir a l'elaboració de polítiques sobre IA. La seva participació garanteix que la normativa sigui exhaustiva, amb visió de futur i reflex dels valors de la societat, i garanteix l'equilibri entre el foment de la innovació i la protecció davant possibles danys.

La IA té el potencial de remodelar les nostres societats en el seu conjunt. Per aquest motiu, la participació activa dels parlaments és fonamental per a salvaguardar l'interès públic i garantir que les tecnologies d'IA es desenvolupin i s'utilitzin de manera responsable.

Teniu a les vostres mans els esforços col·laboratius de l'EPTA (European Parliamentary Technology Assessment), la xarxa parlamentària europea d'avaluació tecnològica, presentats en la conferència anual del 2023 «Intel·ligència artificial generativa: oportunitats, riscos i reptes en polítiques públiques (L'hora de l'avaluació tecnològica als parlaments)», que ha tingut lloc a la seu del Parlament de Catalunya, a Barcelona, el 9 d'octubre de 2023.

Aquest esforç col·lectiu reuneix l'experiència, les alternatives i les perspectives de diversos parlaments en un esquema de governança multinivell.

Aquest informe subratlla la importància de la col·laboració per a abordar problemes globals complexos i, al mateix temps, serveix com a full de ruta perquè els responsables polítics i el públic en general naveguin per la frontera de la IA d'una manera responsable. Junts estem configurant el futur de la IA defensant els principis ètics de la IA, promovent la transparència i guiant el desenvolupament de tecnologies d'IA que beneficiïn la societat en el seu conjunt.

Us desitjo una instructiva i agradable lectura!

Anna Erra i Solà

Presidenta del Parlament de Catalunya i presidenta del CAPCIT
(Consell Assessor del Parlament sobre Ciència i Tecnologia)

Barcelona, octubre del 2023

Resum executiu

Per què és rellevant parlar d'intel·ligència artificial (IA) generativa?

En l'última dècada, els mètodes d'aprenentatge profund (un subcamp de l'aprenentatge automàtic, al seu torn un subcamp de la IA) han experimentat millores notables en la precisió i la generalització, i han donat lloc a avenços en l'assistència sanitària (per exemple, en la diagnòsi del càncer) o en la ciència (per exemple, en la predicció meteorològica o el modelatge de proteïnes).

Alguns dels desenvolupaments més reeixits que utilitzen una metodologia d'aprenentatge profund són els models de llenguatge extens (LLM). Aquests models estan orientats a generar seqüències de paraules que siguin plausibles, però no necessàriament veraces. El propòsit dels desenvolupadors és aconseguir bones aproximacions estadístiques del que els humans podrien escriure a partir de la gran quantitat de textos escrits per humans utilitzats per a entrenar el model.

Diverses empreses han creat aplicacions a partir dels LLM, com ara ChatGPT o Bard. Aquestes aplicacions tenen finalitats diverses, com ara cercar informació, resumir textos o crear contingut a partir d'un diàleg amb un usuari. Una part del públic ha percebut aquests sistemes com un exemple d'intel·ligència general i això ha provocat que la comunitat científica advertís d'aquesta percepció errònia.

Hi ha una sèrie de qüestions que cal posar sobre la taula per a crear un debat sobre l'ús correcte d'aquest grup d'aplicacions.

Sense fiabilitat. Com s'ha esmentat anteriorment, el resultat que generen aquests sistemes no pretén ser veraç sinó versemblant. Aquesta és una propietat arriscada, atès que els ciutadans desinformatats poden prendre decisions basades en informació errònia produïda per aquests sistemes. S'han d'introduir les garanties adequades per a protegir-los. Als LLM hi manca la connexió amb les fonts d'informació que tenen els motors de cerca actuals.

Comportament de la caixa negra. Aquests models són el resultat d'un tediós procés d'optimització que calcula milers de milions de paràmetres que després s'utilitzen per a generar els resultats. La capacitat d'aquests sistemes per a explicar per què generen alguna cosa és extraordinàriament limitada, atès que el que generen és el resultat de milions d'operacions aritmètiques sobre aquests paràmetres.

Rigidesa. Els sistemes s'entrenen amb un conjunt de dades per a fixar els paràmetres, i després no hi ha cap entrenament addicional i no s'hi introdueixen més canvis. Si limitem l'entrenament a documents anteriors a una data determinada, la informació dels documents generats després d'aquesta data no influirà en el que generin.

Consum d'energia elevat. El cost de l'entrenament dels sistemes comporta milers de megawatts. Així, atès que els sistemes s'han de tornar a entrenar sovint per la manca d'adaptabilitat, el consum s'ha de repetir cada vegada.

Sistemes de propietat. La majoria d'aquestes aplicacions pertanyen a un nombre reduït d'empreses que es poden permetre l'elevat cost del desenvolupament. Aquest negoci legítim s'ha de controlar per a protegir els ciutadans de possibles biaixos i mals usos. Per exemple, hi ha opacitat pel que fa a les dades utilitzades per a entrenar els sistemes i, per tant, no és clar si aquests sistemes respecten regulacions com el Reglament general de protecció de dades.

Drets d'autor. Atès que els textos es generen a partir dels documents que s'havien usat per a l'entrenament i que no hi ha cap manera viable de rastrejar-los, sembla que, en aquest cas, la noció dels drets d'autor està en joc. S'ha de debatre sobre com s'han de preservar els drets dels creadors humans.

En resum, tot i que els nous desenvolupaments basats en mètodes d'aprenentatge profund estan produint avenços significatius en diverses àrees de la ciència, el seu ús en el context dels LLM s'ha d'estudiar de prop per a evitar la desinformació i les amenaces de ciberseguretat.

Com aborda l'avaluació tecnològica la IA generativa?

Vam fer a ChatGPT la pregunta «Què és l'avaluació tecnològica?», i aquests són el principi i el final de la seva resposta: «L'avaluació tecnològica, sovint anomenada *avaluació tecnològica (AT)* o *avaluació de la tecnologia*, és un procés sistemàtic i interdisciplinari que s'utilitza per a avaluar i comprendre els impactes, les conseqüències i les implicacions potencials de les noves tecnologies. Comporta l'anàlisi detallada dels desenvolupaments tecnològics, les innovacions o les solucions tecnològiques proposades, per a avaluar-ne els aspectes socials, econòmics, ambientals, ètics i polítics. L'objectiu principal de l'avaluació tecnològica (AT) és informar els responsables de la presa de decisions, responsables polítics i ciutadans sobre els riscos i els beneficis associats amb l'adopció i la implementació de tecnologies específiques (...) Té un paper crucial en la configuració del discurs públic, el desenvolupament de polítiques i les decisions reglamentàries, amb relació a les noves tecnologies i les emergents».

Atès que es tracta d'una definició general que és prou detallada (cal tenir en compte l'enumeració oberta de l'abast, l'objecte, el públic i l'impacte de l'AT), és difícil de criticar.

En altres informes de l'EPTA, hem afirmat que l'AT explora la relació entre tecnologia i societat, i hem indicat que té per objectiu saber com els desenvolupaments tecnològics actuals afecten el món en què vivim i que pretén contribuir a la formació de l'opinió pública i política. Així, el deure dels membres de la xarxa EPTA és informar i assessorar els parlaments sobre les interrelacions entre CTI (ciència, tecnologia i innovació) i política i societat.

En l'informe de l'EPTA del 2018 s'hi diu: «La intel·ligència artificial (IA) s'està convertint en una de les tecnologies més discutides que existeixen». L'informe de l'EPTA d'aquest any es basa en avaluacions anteriors, aprofundint en les implicacions socials en constant evolució d'aquesta tecnologia transformadora.

A partir de les idees i les lliçons apreses dels informes dels membres de l'EPTA, aquesta introducció fa una anàlisi matisada dels impactes multifacètics de la IA. Navega pel complex panorama de l'adopció de la IA, tenint en compte les seves influències en les estructures econòmiques, els mercats laborals, la sanitat, l'ètica, la democràcia i la governança. Els informes dels membres de l'EPTA destaquen la capacitat de la IA per a agreujar les desigualtats i els dilemes ètics.

L'informe austríac assenyalava la necessitat de l'AT de reaccionar ràpidament als nous productes d'IA, com ara els bots de conversa d'IA, i afirma que «l'AT ha de ser encara més ràpida que abans».

En aquest sentit, l'informe suís diu que «atesa la gran diferència entre la velocitat inusual d'aquestes tecnologies i el temps necessari per a adoptar possibles mesures contra els seus riscos, sembla crucial que l'AT pugui reaccionar d'una manera responsable i proactiva». L'informe també diu: «L'AT ha de contribuir d'una manera decisiva a la legitimació democràtica de l'ús i la regulació de la IA generativa. Per mitjà de proporcionar informació independent que tingui en compte les múltiples facetes de l'impacte de la IA generativa en la societat, l'AT pot ajudar els qui prenen decisions (tant ciutadans com actors polítics) a formar-se una opinió i a prendre decisions informades sobre aquests desenvolupaments tecnològics».

Com escriu el Consell de Tecnologia de Dinamarca (DBT), «l'AT ha de jugar un paper a l'hora d'equipar els diputats per a entendre millor la tecnologia» i recorda les capacitats de previsió de l'AT per a «fer prediccions informades sobre com la IA generativa pot afectar les nostres societats».

L'informe alemany escrit per l'Oficina d'Avaluació Tecnològica del Bundestag Alemany (TAB) subratlla el fet que «s'han de posar a prova els procediments d'avaluació tecnològica establerts». L'informe alemany també assenyalava el potencial de la IA generativa en l'àmbit de l'AT i el procés de presa de decisions polítiques.

L'informe portuguès inclou una sèrie de preguntes per a l'AT, com ara si «els desenvolupadors d'IA haurien de publicar o restringir els seus models?» o «fins a quin punt les plataformes i els desenvolupadors d'IA poden crear associacions significatives que puguin ajudar a detectar i eliminar el contingut que no sigui autèntic?».

L'informe de l'Institut Rathenau (Països Baixos) serveix com a crida a l'acció a l'AT, amb les paraules següents: «L'AT pot ampliar el debat sobre la IA generativa i el treball». L'informe explica que «l'AT mostra que l'impacte potencial de la IA generativa en el mercat laboral no és inequívoc, sinó variat i complex, amb possibles resultats tant positius com negatius».

L'informe grec, centrat també en les implicacions laborals, posa èmfasi sobre el fet que «les conseqüències de les noves tecnologies no es determinen per endavant. S'han d'examinar i controlar».

L'informe japonès mostra les preocupacions i pors creixents pel que fa a la proliferació sense control de tecnologies d'IA i demana «debats i mesures nous» que «equilibrin els riscos, al mateix temps que s'aprofita el potencial de la IA generativa». L'informe japonès també demana que s'actui per a «educar la població en l'ús de la IA generativa i desenvolupar recursos humans per a crear-los».

L'informe suec exposa que «la necessitat d'un coneixement rellevant de la IA s'ha de satisfer mitjançant l'educació i la formació, l'educació contínua i la recerca».

L'informe del Consell Noruec de Tecnologia (NBT) comparteix la seva experiència pel que fa a la participació de les parts interessades en el camp de la IA generativa en «un esforç continu per a implicar la ciutadania, les parts interessades i els afectats, per a ajudar a oferir un assessorament per a polítiques futures als parlamentaris en el proper trienni». Sovint, l'AT comporta la participació d'un grup divers de parts interessades, inclosos experts, responsables polítics, representants del sector, comunitats afectades i públic en general. Això garanteix que es tingui en compte una àmplia gamma de perspectives en el procés d'avaluació.

L'informe del Panel per al Futur de la Ciència i la Tecnologia (STOA) del Parlament Europeu se centra a demostrar que els experts en AT han posat l'èmfasi en l'àmbit de la legislació de la IA. Quin és el risc dels sistemes d'IA d'ús general? Quin tipus de normes s'han d'establir per a la IA? L'informe posa èmfasi en la necessitat d'un desenvolupament responsable de la IA i marcs reguladors reflexius per a garantir la distribució equitativa dels beneficis i mitigar les conseqüències no desitjades.

L'informe del Regne Unit explica el debat del Govern i el Parlament del Regne Unit sobre la manera de regular la IA, amb posicions que defensen la regulació de la IA per mitjà d'un codi de conducta en lloc de basar-se en normes jurídiques com un enfocament pragmàtic i flexible per a abordar el panorama de la ràpida evolució de la IA. Ara bé, també hi ha posicions que defensen que «es garanteixi el dret de consulta i notificació als treballadors si l'aplicació de la tecnologia en el lloc de treball comporta la vigilància d'un treballador o un canvi significatiu de la seva feina».

En el mateix sentit, l'informe lituà afirma que «és molt important considerar diversos mecanismes de regulació i governança en el camp de la IA generativa» i «preparar la societat per a l'ús de la intel·ligència artificial».

L'informe català (CAPCIT) segueix la mateixa línia i diu que «s'ha de posar en marxa una regulació per a evitar els possibles perills dels sistemes d'IA generativa», perquè, si ens fixem en tecnologia d'IA generativa com els LLM, trobem que, com subratlla aquest informe: «Són sistemes sense incorporació que no estan preparats per a resoldre els reptes dels sistemes robòtics. Pel que fa a les capacitats intel·lectuals, aquests sistemes mostren capacitats de raonament molt superficials; desconeixen el seu món, no tenen sentit comú i no han experimentat res».

Què pensen els parlamentaris sobre la IA generativa?

Servir el legislatiu per mitjà de l'AT exigeix un bon coneixement del context del legislador. És tan important saber com fer arribar la informació científica i tecnològica als parlamentaris com saber quins són els coneixements del legislador i el context jurídic i polític en què es troba, perquè tota l'AT s'ha de contextualitzar d'acord amb les necessitats legislatives si vol ser útil i contribuir als debats jurídics i a la redacció jurídica de les activitats parlamentàries.

Amb la IA hi ha una necessitat més gran de saber la posició dels parlamentaris, atès que és una tecnologia que evoluciona a una velocitat tan ràpida que la societat, i, per tant, també els parlaments, queden fàcilment enrere pel que fa al coneixement de la tecnologia, els usos i els reptes. La ràpida evolució d'aquesta tecnologia dificulta tenir-ne un coneixement profund actualitzat i, en conseqüència, un coneixement sobre les polítiques i normatives que haurien de ser a l'ordre del dia.

Què saben els parlamentaris sobre la IA generativa? Per a respondre aquesta pregunta, hem donat l'oportunitat a tots els parlaments de l'EPTA de respondre sis preguntes, cadascuna relacionada amb els àmbits que es tracten en aquest informe i en la conferència de l'EPTA d'enguany.¹

Tenint en compte les respostes a la pregunta sobre el concepte d'IA generativa, es veu clarament que la majoria de parlamentaris l'associen amb la generació de contingut (text, àudio i vídeo) per una màquina sense supervisió humana. La majoria dels parlamentaris van subratllar que la IA, i, en particular, la IA generativa, és una eina i, com qualsevol eina, pot oferir oportunitats però també comporta assumir riscos. De tota manera, aquesta resposta tan equilibrada a la primera pregunta tendeix a no ser-ho tant tan bon punt els hem preguntat sobre la relació entre la IA generativa i la democràcia, i també amb àrees polítiques clau com la salut, l'educació i el mercat de treball. En aquest cas, les respostes mostren que hi ha més preocupació sobre els riscos que comporta l'ús de la IA generativa que no pas idees sobre els beneficis que pot aportar a les societats. Fem una ullada breu a aquesta qüestió.

Quan es tracta d'IA generativa i democràcia, la majoria de les respostes connecten aquests dos conceptes i, generalment, ho fan posant èmfasi en els perills que preveuen. Un concepte molt comentat és la desinformació, juntament amb el fet que la IA generativa té una gran capacitat per a potenciar els biaixos socials; en moltes de les respostes, s'esmenta la necessitat d'algun tipus de control de qualitat de les dades que alimenten la IA generativa. A més, sembla que alguns parlaments ja són actius en aquest àmbit, i ho demostren les respostes dels seus diputats en l'àmbit dels LLM. En aquest sentit, els parlamentaris han manifestat la necessitat d'avaluar si hi ha d'haver un LLM en la seva pròpia llengua per tal d'alimentar la IA generativa amb els valors i dades socials d'aquest país determinat (per exemple, Noruega). Moltes respostes relacionaven els LLM amb la manca de control de qualitat i la «mentalitat» de les dades, cosa que pot conduir a fomentar valors que xoquen amb els de les democràcies. No obstant això, molts parlamentaris han subratllat que també esperen que tingui un impacte positiu i, fins i tot, poder-lo utilitzar en la seva tasca parlamentària.

Pel que fa a la IA generativa i la salut, els parlamentaris han relacionat la IA generativa (i, en general, la IA) amb diagnòstics i tractaments més bons. No obstant això, en les respostes dels diputats el més comú ha estat la por que les dades personals es puguin veure compromeses, i també la necessitat que les decisions finals quedin sempre en mans humanes. D'altres també han subratllat el fet que els biaixos en les dades sanitàries poden contribuir a augmentar aquests biaixos en el sistema sanitari i en la recerca sanitària.

Si ens fixem ara en la IA generativa i l'educació, probablement es tracta d'una de les àrees en què la majoria de parlamentaris han mostrat una comprensió clara del repte real que, a causa de la IA generativa, ja afronta el sistema educatiu. És a dir, tot i reconèixer que és una eina que pot ajudar en les tasques educatives, s'ha subratllat la necessitat de formar professors i dissenyar un programa educatiu que tingui en compte aquesta tecnologia i, d'aquesta manera, que sigui capaç d'ensenyar i avaluar sense que la IA generativa impedeixi que aquests procediments s'assoleixin.

Finalment, si tenim en compte les respostes a la relació de la IA generativa i el mercat de treball, els parlamentaris han tendit a equilibrar més els avantatges i els riscos de la IA generativa. Tot i que subratllen que aquesta tecnologia substituirà tasques més que no pas llocs de treball, també subratllen que es crearan nous llocs de treball, i assenyalen que el que és crucial en aquest àmbit és que una societat tingui la capacitat d'oferir les competències per a aquests llocs de treball nous i, especialment, garantir que ningú es queda enrere per no tenir coneixements tecnològics.

-
1. Les preguntes fetes als parlamentaris de l'EPTA van ser:
 - i. Què és el primer que us ve al cap quan sentiu «intel·ligència artificial generativa»? Si no n'esteu segur, què us suggereix la intel·ligència artificial?
 - ii. Creieu que hi ha una relació entre IA generativa i democràcia? Podríeu explicar breument què en penseu?
 - iii. Si haguéssiu de pensar en àrees polítiques on probablement s'haurà de tractar la IA generativa, quines serien?
 - iv. Us imagineu com pot influir la IA generativa en el desenvolupament/estructura del mercat laboral / sector de l'ocupació/treball?
 - v. Sabeu si el Parlament ha tractat qüestions relacionades amb la IA generativa o si té previst de fer-ho?
 - vi. Quina mena de coneixements necessiteu per a prendre decisions sobre la IA generativa? Com us pot ajudar la vostra institució d'AT?

IA generativa i democràcia

Les eines d'IA generativa poden reduir les barreres de la participació democràtica, fer que els serveis públics siguin més eficients i inclusivament i oferir a la població d'accedir més fàcilment al coneixement i millorar les habilitats. Les noves eines de generació de textos i les eines dissenyades d'una manera universal per a la cerca, la traducció i la simplificació de textos complexos podrien millorar l'accés dels ciutadans a la informació i capacitar més persones per a expressar les seves opinions sobre qüestions socials. Els diputats danesos subratllen que la IA crea noves possibilitats de participació democràtica mitjançant, per exemple, enquestes públiques per a conèixer les opinions dels ciutadans.

Desinformació, parcialitat i confiança en les institucions democràtiques

La IA generativa pot augmentar dràsticament el ritme i el volum de la desinformació i la informació falsa, els hipertrucatges i les notícies falses en línia. L'esfera pública digital podria inundar-se de contingut automatitzat, singularitzat i manipulador, cosa que dificultaria cada vegada més que la població identifiqués si la informació és vertadera o falsa, o generada per humans o per la IA.

Les noves eines es poden utilitzar per confondre els ciutadans, agreujar la tensió social i augmentar la polarització política, i, d'aquesta manera, socavar la integritat dels processos democràtics i la confiança en les institucions. D'acord amb un estudi sobre hipertrucatges entre diputats suïssos, molts representants consideren que els hipertrucatges i el contingut d'IA són un risc elevat i concret per a la democràcia suïssa, amb l'exemple d'un partit polític que va utilitzar la IA per a generar un cartell de campanya en les eleccions federals actuals.

A més, la recerca mostra que els LLM predominants tenen biaixos tant en les dades com en l'entrenament, i, per tant, poden perpetuar diversos estereotips socials. Els LLM també poden al·lucinar i generar contingut fals que sembla cert. Així, els resultats dels sistemes d'IA poden influir i manipular les opinions d'una manera inadvertida o intencionada, cosa que suposa una amenaça per a l'autonomia individual i els processos democràtics.

El biaix també pot conduir a decisions errònies i, fins i tot, discriminació, si els utilitzen, per exemple, institucions del sector públic que desconeixen o no poden detectar els errors produïts pels sistemes. Això pot debilitar la confiança dels ciutadans en les institucions democràtiques.

Control democràtic de la IA

Molts països europeus debaten actualment sobre com es pot reforçar la governança democràtica de la IA sense que això sigui un obstacle per a la innovació. Actualment, les grans empreses tecnològiques són les responsables dels grans avenços en la IA. Aquestes empreses tenen coneixements, recursos i experiència en sistemes d'IA amb el potencial d'impactar i transformar les societats. No obstant això, els models posats a disposició del públic són opacs. Per a reforçar els mecanismes de supervisió i control públics, serà crucial que les democràcies estableixin un marc normatiu sòlid que regeixi tant el desenvolupament com el desplegament de la IA.

Actualment, la UE està negociant la Llei de la IA, un marc legislatiu basat en el risc que estableix estàndards de transparència, fiabilitat i retiment de comptes. Els sistemes d'IA que poden afectar significativament la democràcia, l'estat de dret o els drets i les llibertats individuals es prohibiran o es consideraran d'alt risc. Els sistemes d'alt risc estaran subjectes a obligacions estrictes, com ara sistemes adequats d'avaluació i mitigació de riscos, conjunts de dades d'alta qualitat, documentació i registre detallats i mecanismes adequats de transparència i supervisió humana.

Dins l'OCDE i el Consell d'Europa també hi ha debats sobre normativa i, a escala nacional, actualment s'estudien complements normatius als marcs internacionals. Les directrius ètiques sobre l'ús responsable de la IA, a més dels marcs legals, també poden ajudar a abordar reptes comuns com ara la desinformació, la privadesa i la protecció de dades, els biaixos, la propietat intel·lectual i la desigualtat lingüística. Per exemple, el Centre Nacional d'Ètica de Dinamarca actualment elabora recomanacions ètiques sobre l'ús de la IA generativa i l'Agència de Digitalització de Noruega ha elaborat directrius per a l'ús responsable de la IA en el sector públic.

Es pot construir una infraestructura d'IA a Europa?

El ChatGPT s'ha caracteritzat per ser multilingüe, però monocultural. En general, els ciutadans i les autoritats tenen una visió limitada del funcionament dels LLM. Quan les eines d'IA generativa s'incorporen als serveis digitals quotidians, és crucial conèixer les funcions dels models per a salvaguardar principis fonamentals com ara la privadesa, la transparència i la fiabilitat. Una infraestructura d'IA accessible i segura també podria donar més suport als idiomes poc representats, representar valors culturals i permetre que l'Administració pública fos conscient del potencial de la IA generativa.

Països com Noruega, Suècia i el Regne Unit actualment exploren opcions per a construir LLM nacionals. El desenvolupament d'aquests models podria ser una manera d'enfortir la supervisió democràtica del disseny i el desplegament de la IA generativa i reduir la dependència dels actors externs. No obstant això, hi ha moltes preguntes sense resoldre, per exemple, qui hauria de construir, operar i accedir a aquest model. A més, construir un model des de zero requereix molts recursos i comporta costos elevats. Diverses contribucions a aquest informe de l'EPTA destaquen que l'accés a les dades i una potència informàtica suficient també són essencials perquè els països i les regions es beneficiïn de l'adopció de la IA generativa.

IA generativa i salut

Els sectors sanitaris de molts països europeus estan sota pressió a causa de l'envelliment de la població, la manca generalitzada de personal qualificat i l'augment dels costos dels tractaments nous i més especialitzats.

Així, per a l'atenció sanitària, l'augment de les tecnologies d'IA generativa no podria haver arribat en un moment millor, i els polítics i les empreses tecnològiques ja elogien les solucions d'IA perquè els consideren un ingredient clau per a afrontar aquests reptes. En el Procés d'Hiroshima del G7 sobre IA generativa, sis dels set enquestats consideren que la millora de l'atenció sanitària és una de les oportunitats clau de la IA generativa.²

Tal com assenyala un estudi de l'NBT, la IA és esperançadora per a l'àmbit de l'assistència sanitària.³ Els assistents sanitaris digitals poden ajudar el personal sanitari a fer diagnòstics, seleccionar el tractament, controlar el pacient i advertir les complicacions. Mitjançant l'aprenentatge automàtic, poden analitzar literatura mèdica, interpretar imatges i altres dades i explorar milers de registres de pacients. Això pot esdevenir la clau per a aconseguir tractaments que s'adaptin més a cada individu i que depenguin menys de l'experiència del metge.

Més concretament, l'anomenada *IA mèdica generalista*⁴ pot redactar automàticament informes de radiologia que tinguin en compte la història del pacient, ajudar els equips quirúrgics en les consultes orals i visualitzacions anotades, donar suport en la decisió clínica al costat del pacient amb explicacions i recomanacions detallades per a l'atenció futura i redactar notes del pacient i informes d'alta.

A més, els sistemes informàtics amb IA poden parlar l'idioma dels pacients i respondre d'una manera ràpida i precisa quan s'hi posen en contacte. Aquestes primeres línies digitals poden oferir una assistència sanitària més bona a tot el país o la regió i permetre que el personal sanitari passi menys temps al telèfon i en pugui dedicar més al tractament. Els rellotges i les polseres amb sensors poden registrar-ho tot, des de la freqüència cardíaca fins al to de veu. La IA interpreta les dades i proporciona informació contínua sobre la salut física i mental dels usuaris. Això fa que sigui més fàcil controlar els pacients a casa, detectar la malaltia abans i començar el tractament més ràpidament.

2. https://read.oecd-ilibrary.org/science-and-technology/g7-hiroshima-process-on-generative-artificial-intelligence-ai_bf3c0c60-en?utm_source=pocket_saves#page13

3. Teknologirådet 2022: Artificial intelligence in the clinic – Six trends for the health service of the future [Inteligència artificial a la clínica. Sis tendències per al sistema de salut del futur]. <https://media.wpd.digital/teknologiradet/uploads/2023/01/Artificial-Intelligence-in-the-Clinic.pdf>

4. Moor, M., Banerjee, O., Abad, Z. S. H., et al. Foundation models for generalist medical artificial intelligence [Models de fundació per a la IA mèdica generalista]. En: *Nature* 616, 259-265 (2023). <https://doi.org/10.1038/s41586-023-05881-4>

A més d'augmentar l'eficiència i reduir les tasques manuals, la IA generativa ja ha demostrat la seva aplicabilitat en la recerca sanitària, tant per a millorar la nostra comprensió del nostre cervell i el nostre cos, i també per a desenvolupar nous tractaments, metodologies i eines diagnòstiques i predictives. La filial d'Alphabet DeepMind va anunciar el 2022 que la seva eina AlphaFold havia cartografiat totes les proteïnes conegudes, cosa que és un punt d'inflexió per a entendre la nostra biologia, amb una immensitat d'usos potencials en termes de diagnòstic i tractaments. El descobriment de fàrmacs es pot accelerar amb la IA generant nous candidats a fàrmacs i predient quin disseny de proteïnes serà més eficaç.

La societat s'enfronta a reptes jurídics, ètics i tècnics importants perquè la IA sigui un èxit en l'assistència sanitària. Per tenir èxit amb la IA generativa en l'atenció sanitària, les dades de salut han de ser accessibles més fàcilment, d'una manera que sigui compatible amb una protecció de dades adequada. Això serà un repte quan les mesures clíniques es combinin amb els registres dels pacients i les dades genòmiques, i també amb les dades de comportament dels sensors. A més, amb els models extensos es corre un risc més gran d'exposar dades sensibles dels pacients a partir dels conjunts de dades per a l'entrenament.⁵

Les dades sanitàries històriques tenen biaixos, com ara una sobrerrepresentació dels homes occidentals. L'aprenentatge automàtic es basa generalment en dades sanitàries històriques, cosa que els pot augmentar segons l'escala del model. Quan s'introdueixen solucions d'IA per a l'atenció sanitària, s'ha de garantir, mitjançant la validació i l'auditoria contínua, que no contribueixen a la discriminació, per exemple, per raó de gènere o ètnia.

En aquest sentit, hi ha d'haver un debat sobre fins a quin punt un metge ha d'entendre el procés automatitzat. Es tracta tant del que un metge ha de saber per a actuar d'una manera responsable i significativa davant els resultats del procés, com de garantir la transparència i la confiança envers els pacients mitjançant la capacitat d'entendre i explicar els antecedents del resultat d'un procés automatitzat o parcialment automatitzat.

Els sistemes d'IA no sempre poden explicar per què han arribat a un resultat d'una manera que els humans puguin entendre, com ara que algú corri un risc alt de patir una malaltia. Per tant, és fonamental desenvolupar nous mètodes i estàndards sobre com els models d'IA comuniquen els seus consells o idees. A més, com millors siguin els assistents digitals, més difícil serà saber quan el sistema hauria de ser anul·lat pels humans. Per tant, calen rutines per a mantenir la competència dels metges i infermeres. És important que el personal tingui els recursos necessaris per a utilitzar els sistemes correctament i una comprensió clara del que fa el sistema. Un element addicional que cal tenir en compte és el garantiment que la implementació de la IA no comporti un augment tan gran de la velocitat dels processos que els metges no tinguin l'oportunitat real de fer un judici humà.

Les noves capacitats de diagnòstic i tractament han fet que la pressió pressupostària sobre els hospitals i l'atenció s'hagi incrementat d'una manera important. L'augment de la capacitat diagnòstica i predictiva que s'espera que potencii la IA generativa significa que es diagnosticaran més malalties, la qual cosa augmentarà aquesta pressió.

Al mateix temps, la medicina de precisió i especialitzada és proporcionalment més cara que la medicina general perquè l'escala de producció és més petita i sovint és més cara. D'aquesta manera, això se suma a la necessitat ja existent de prioritzar diagnòstics i tractaments.

Per tant, en conclusió, encara queden per aclarir algunes qüestions fonamentals en l'ús de la IA generativa en l'atenció sanitària. Algunes d'aquestes qüestions no són exclusives d'aquest sector, però en l'atenció sanitària sovint hi ha en joc la vida o la mort, la qual cosa fa més difícil trobar l'equilibri per a una implementació responsable i exitosa.

5. Moor, M., Banerjee, O., Abad, Z. S. H., *et al.* Foundation models for generalist medical artificial intelligence [Models de fundació per a la IA mèdica generalista]. En: *Nature* 616, 259–265 (2023).

IA generativa i educació

És difícil sobreestimar l'impacte de la IA generativa en l'educació. Fa més de quaranta anys que es debat sobre l'ús de la IA en l'educació. Amb l'arribada de ChatGPT (i altres generadors de text potents), les diverses visions pel que fa a les futures tecnologies d'aprenentatge –dels estudiants que tenen un company d'aprenentatge personal a la IA que ajuda a evitar les avaluacions formals– es van convertir en opcions realistes i fàcilment disponibles per primera vegada. Tot i això, malgrat la seva capacitat per a simular d'una manera convincent l'ús humà del llenguatge natural, els sistemes d'IA generativa encara mostren una capacitat de raonament molt superficial i els manca l'experiència del món real. L'ús en contextos educatius pot tenir efectes profunds, directes i indirectes, i requereix un gran escrutini. La Llei d'IA que ha proposat la UE, per exemple, classifica l'ús de la IA en l'educació com d'alt risc. La qüestió de com ha d'afrontar el sistema educatiu aquest desenvolupament tecnològic adquireix, doncs, una nova urgència.

En el debat, i en les respostes dels diputats a les preguntes de l'EPTA per a l'informe d'enguany, hi sorgeixen dos temes principals: el paper de l'educació en la preparació de les generacions futures per a un món on els humans i les màquines intel·ligents interactuïn (aprendre sobre la IA generativa), i el paper dels sistemes basats en IA en la millora de l'educació (aprendre amb IA).

La necessitat d'aprendre (i educar) sobre la IA és reconeguda per la majoria de diputats que van respondre a la pregunta sobre IA i educació. Això és aplicable als estudiants a les escoles, però també a la formació contínua de professionals i a la formació dels professors, les «habilitats d'ús dels quals han d'augmentar significativament», com diu un diputat alemany. També hi ha debat sobre l'edat en què ha de començar aquesta educació i sobre les competències que s'han d'ensenyar –competències per a crear IA, però també per a tractar-la, també d'una manera crítica. «S'ha de garantir que les futures generacions entenguin els riscos de la IA des de ben petits, però que al mateix temps en puguin aprofitar els beneficis», diu un diputat austríac. Un diputat alemany subratlla que els estudiants encara han d'aprendre a resoldre problemes pel seu compte. I s'apunta que les escoles també tenen un paper important per a «difondre democràticament» les competències a molta gent, com diu un diputat danès.

Les contribucions dels països a aquest informe de membres individuals de l'EPTA detallen una sèrie d'iniciatives polítiques per a desenvolupar programes educatius sobre IA. Per exemple, a Suècia, s'espera que les universitats ofereixin programes i cursos relacionats amb les noves necessitats del mercat laboral. Al Regne Unit, el Comitè de Ciència i Tecnologia de la Cambra dels Lords ha identificat diverses accions que el Govern hauria de dur a terme per a abordar el dèficit de competències CTEM existent, inclosa l'oferta de cursos per sota del nivell de grau i la millora dels programes d'aprenentatge. La contribució de Dinamarca veu la necessitat d'implicar els ciutadans i les parts interessades en el desenvolupament de principis per a aprendre sobre la IA generativa i utilitzar-la en el sistema educatiu. I Lituània està adoptant un enfocament altament inclusiu, que implica no només la comunitat experta i les empreses, sinó també els filòsofs en el debat sobre la coexistència dels humans i la IA. Finalment, però no menys important, una majoria de dos terços dels ciutadans europeus demana més educació i formació per a desenvolupar les habilitats digitals, segons els resultats d'enquestes citats pel Parlament Europeu (STOA).

Pel que fa a l'aprenentatge amb IA, la majoria de diputats hi veuen oportunitats i n'hi ha molt pocs que consideren una prohibició absoluta. La majoria de governs semblen estar d'acord amb la visió positiva; les autoritats educatives d'Alemanya, per exemple, van respondre a ChatGPT permetent-ne l'ús com a ajuda en els exàmens, i molts educadors estan oberts a utilitzar la IA. La IA generativa té el potencial de millorar l'aprenentatge, permetre una participació més gran i reduir les desigualtats, per mitjà d'adaptar el procés d'aprenentatge a les necessitats individuals, inclosos els aprenents amb necessitats especials (per exemple, amb dislèxia). «Hi ha possibilitats enormes amb aquest recurs», diu un diputat danès. Ara bé, per tal d'aprofitar aquestes oportunitats, els diputats diuen que cal que hi hagi determinades regulacions. S'han de protegir les dades personals d'alumnes i professors, i també els drets especials dels menors. Els sistemes d'IA necessiten controls de qualitat i certificacions «per a ajudar els consells escolars, els directors i els professors a decidir si utilitzen o no un producte» (diputat alemany). Els educadors han de ser conscients que les eines basades en IA poden «esdevenir molt addictives» (comentari dels diputats noruecs) i les institucions públiques haurien d'evitar dependre de proveïdors concrets. En canvi, un diputat austríac suggereix promocionar productes de codi obert. Es demana un marc jurídic per a no «deixar-ho en mans del mercat»,

com diu un diputat holandès. A més, la IA generativa, amb la seva gran quantitat de dades d'entrenament de fonts que sovint no es revelen, planteja problemes pel que fa als drets d'autor i a la compensació justa dels creadors, tal com destaca la contribució del Japó a aquest informe.

A més d'aprendre sobre la IA i aprendre amb IA, es pot identificar un tercer problema: la IA generativa pot canviar l'educació de maneres molt fonamentals («la IA transforma l'educació»). Una qüestió és el paper dels exàmens i les qualificacions educatives formals: «Molts estudiants han començat a utilitzar ChatGPT per a completar les tasques, i possiblement copien en situacions d'avaluació», tal com van saber els diputats noruecs. De fet, hi ha el cas d'un estudiant alemany que van enxampar copiant a la selectivitat amb l'ajuda de ChatGPT. Com a conseqüència d'això, «l'avaluació del que s'ha après probablement s'ha de fer d'una manera diferent en el futur», segons un diputat austríac. «Després de tot, volem conèixer les habilitats dels aprenents, no de la IA» (diputat alemany). Un altre assumpte és la revisió dels currículums i dels mètodes d'ensenyament de totes les assignatures, no només de l'educació en matèria de mitjans de comunicació. «La IA també tindrà un impacte en els continguts i les competències que hem d'ensenyar als estudiants» (diputat alemany). «Que els estudiants passin molt de temps editant la gramàtica potser no és la millor manera d'emprar el temps», diu un diputat danès. A Catalunya, el Departament d'Educació ja ha començat a estudiar la introducció de la IA en els currículums de primària i batxillerat. I el Ministeri d'Educació, Cultura, Esports, Ciència i Tecnologia del Japó va publicar directrius provisionals sobre l'ús de la IA generativa en l'educació (primària i secundària) el juliol del 2023 i va instar les universitats a elaborar les seves pròpies directrius.

Tot i que el caràcter disruptiu de la IA generativa en l'educació és àmpliament reconegut, cal veure quin paper jugarà en relació o en conjunció amb altres tipus de sistemes d'IA més tradicionals.

IA generativa i treball

Juntament amb les imatges de Terminator i SkyNet, la perspectiva de l'atur massiu induït per la IA és probablement una de les representacions més comunes de la IA en l'imaginari públic i un dels impactes potencials que més sovint ha arribat a les redaccions dels mitjans de comunicació. Per aquest motiu, no és sorprenent veure que l'impacte de la IA generativa en l'ocupació en general és un tema que té un lloc destacat en les converses amb polítics que es presenten en els informes dels països següents.

L'impacte de la IA generativa en l'ocupació és lluny de ser cert i és més complex que els escenaris binaris de l'atur massiu o la creació de noves indústries. Actualment, hi ha algunes professions que pateixen cada vegada més la manca de treballadors qualificats i hi ha més sectors que esperen experimentar aquesta escassetat en un futur no gaire llunyà. Aquí les esperances i les prediccions són que la IA generativa, juntament amb, per exemple, la tecnologia robòtica, puguin alleujar aquesta escassetat. Quan es parla dels beneficis de la IA, sovint s'esmenta el seu potencial en radiologia. Es tracta d'un cas en què, potencialment, la IA generativa pot alleujar l'escassetat de mà d'obra, com al Regne Unit, on el Reial Col·legi de Radiòlegs ha estimat que tres de cada quatre hospitals no tenen prou radiòlegs.

A l'altra cara hi ha el fet que la por a l'atur massiu no és del tot injustificada. En diversos sectors l'automatització de llocs de treball amb IA generativa és un escenari molt probable i alguns llocs de treball que avui hi existeixen desapareixeran. El que encara no queda clar és si dins aquests mateixos sectors i empreses s'hi crearan llocs de treball nous que els treballadors substituïts per IA puguin desenvolupar, si aquests treballadors tindran les competències i els requisits previs per a recapacitar-se i passar a altres sectors o si seran substituïts definitivament en el mercat laboral. Per a alguns, les dues primeres possibilitats no són factibles.

Una variable addicional de l'equació sobre l'impacte de la IA en l'ocupació és la probabilitat que apareguin, com un subproducte de la IA generativa, nous productes i serveis, el desenvolupament i utilització dels quals fomenti nous tipus de llocs de treball en empreses existents i en indústries noves.

L'impacte de la IA generativa en l'ocupació va molt més enllà de la qüestió de l'atur. La implementació de la IA generativa en els llocs de treball comportarà un canvi en la naturalesa de les tasques. Més enllà de l'automatització total, com s'ha descrit anteriorment, es poden imaginar almenys dos escenaris. En

un escenari, la tasca en si s'automatitzarà i la nova feina consistirà en administrar el procés automatitzat, per exemple, des de fer soldadura real fins a l'administració digital de la màquina de soldadura. En un altre escenari, el sistema d'IA generativa ajudarà a dur a terme la feina com, per exemple, a l'hora de prendre decisions amb el suport de la IA en l'àmbit de la radiologia, en què la IA pot mostrar al radiòleg àrees d'interès o preocupació. En ambdós casos, hi haurà la necessitat fonamental de recapacitació i formació contínua dels treballadors, per tal de fer aquestes feines d'una manera responsable i eficient i amb una alta qualitat.

La IA generativa és diferent de les tecnologies d'automatització anteriors, atès que és probable que afecti de la mateixa manera els treballadors menys qualificats i els més qualificats, com es veu en l'exemple dels radiòlegs anterior. A continuació, es mostren només alguns exemples de com la IA generativa ja afecta diversos tipus de feina. En l'àmbit del periodisme ja hi ha informes generalitzats d'organitzacions de mitjans de comunicació que utilitzen la IA generativa per a elaborar notícies. Part de la vaga dels guionistes de Hollywood se centra en la preocupació pel fet que la IA generativa es pugui utilitzar per a fer gran part de la feina en el desenvolupament de guions. En la comptabilitat, és evident que la IA hi té un gran potencial, i ja s'ha utilitzat durant força temps en el sector financer. Gran part del treball de recerca legislativa que fan els advocats es pot dur a terme més ràpidament i amb menys cost mitjançant la IA generativa. El disseny gràfic i les il·lustracions es poden fer cada cop d'una manera més automàtica.

En general, l'esperança és que la IA generativa i els sistemes que la utilitzen siguin capaços de fer-se càrrec de feines avorrides, brutes i perilloses; que alliberi els treballadors per a fer feines més significatives i gratificants, i, en molts sectors, que també permeti als treballadors centrar-se en la seva tasca principal en lloc de tasques secundàries com la documentació o les tasques de seguiment. Però també hi ha dubtes. Un dubte és el fet que la integració de la IA generativa doni lloc a una acceleració dels processos de treball. Una conseqüència d'això és que cada treballador, especialment els més qualificats, hagin de dur a terme i completar les tasques a un ritme més ràpid, cosa que pot comportar diversos problemes segons el sector. La qualitat del treball corre el risc de ressentir-se, simplement perquè no hi hauria prou temps per a completar-lo. També s'ha de tenir en compte el risc addicional que comporta la implementació pel fet que si es posa tant d'èmfasi en la velocitat no serà possible fer una supervisió humana adequada dels processos. També hi ha un risc per al benestar mental associat a aquest escenari futur. L'OMS ja ha qualificat l'estrès d'epidèmia de salut del segle XXI, i una acceleració més gran de les tasques podria augmentar aquesta tendència. Per tant, cal un debat fonamental sobre com volem assignar els recursos que es poden alliberar mitjançant la implementació de la IA generativa. Això connecta amb el debat social més ampli sobre l'equilibri entre la vida laboral i la vida familiar, una setmana laboral de quatre dies i l'equitat econòmica a la societat, en vista de la concentració creixent de la riquesa i l'augment de la bretxa salarial entre els empleats i els alts nivells de direcció.

Amb relació a aquesta disparitat hi ha l'ús de la IA com a eina de gestió, que ja s'ha observat. L'ús de la IA per a la gestió no és en si mateix problemàtic ni preocupant. La IA generativa podria ajudar els directius a tenir una visió general més bona de les tasques actuals i futures dels treballadors i a fer un retorn i una assignació de tasques més adequats i personalitzats. Si s'utilitza correctament, podria ser un auxiliar de la gestió, que ajudi els directius a millorar les condicions de treball i el benestar entre els subordinats. Però l'economia de bolos ja ha demostrat com la IA generativa també es pot utilitzar per a degradar les condicions laborals. La IA s'ha utilitzat per a dissenyar horaris gairebé impossibles de complir, per a crear mecanismes de qualificació opacs que castiguen els retards en els lliuraments independentment de la causa i per a fomentar estructures dissenyades per algoritmes amb la finalitat de garantir que les bonificacions només les puguin aconseguir pocs empleats. A més d'aquestes condicions laborals deshumanitzadores, la gestió basada en les dades també pot portar a pràctiques qüestionables de vigilància dels empleats.

L'últim aspecte important fa referència a les condicions laborals que hi ha darrere el desenvolupament de la IA generativa. La producció d'IA generativa es basa en grans quantitats de dades anotades. Quan aquestes dades no estan ben documentades en les bases de dades, hi ha d'haver un treball humà important per a fer aquesta anotació. Les empreses d'IA més grans han optat per subcontractar aquesta tasca a empreses del sud global que exploten els treballadors, gairebé sense condicions laborals, i utilitzen un

enfocament de gestió algorítmica similar a les plataformes descrites anteriorment.⁶ Per tant, es tracta de fer que les empreses d'IA garanteixin unes condicions laborals justes per a les persones que anoten i entrenen els models.

Tant el debat anterior com els informes dels països següents destaquen que la IA generativa encara es troba en una fase d'incertesa pel que fa als impactes que tindrà en l'ocupació. Els potencials positius són immensos, però aprofitar-los alhora que es frenen els possibles efectes negatius requereix una consideració política acurada, especialment sobre com s'han de compensar els impactes negatius en l'ocupació alhora que s'estimula la innovació i l'aparició de nous serveis, i com s'han d'assignar els dividends d'una eficiència més gran.

Perspectives

La IA generativa té el potencial per a influir en molts àmbits de la nostra societat i la nostra economia. El desenvolupament tecnològic es produeix a un ritme cada cop més ràpid. La qüestió de si la IA generativa es pot utilitzar d'una manera responsable i sostenible, i com, continua oberta. L'AT parlamentària afronta el repte constant d'estar al dia per a oferir als diputats orientacions i coneixements per a la presa de decisions oportuns i rellevants, i contribuir a un debat social fructífer sobre aquest tema.

6. <https://fair.work/en/fw/publications/fairwork-cloudwork-ratings-2023-work-in-the-planetary-labour-market/>

Aportació de l'Institut d'Avaluació Tecnològica a l'informe EPTA 2023 sobre intel·ligència artificial generativa: oportunitats, riscos i reptes polítics
Intel·ligència artificial generativa i democràcia

Àustria

Autor: Walter Peissl

Estimulat per la propera generació de sistemes d'intel·ligència artificial generativa, el desenvolupament de la intel·ligència artificial (en endavant, IA) i la seva percepció social ha entrat en una nova etapa, també a Àustria. La IA ja s'està utilitzant de diverses maneres, per exemple, per a detectar estructures en grans quantitats de dades (reconeixement de patrons), per extreure'n conclusions i així preparar la presa de decisions. Amb els sistemes d'IA generativa, la IA s'endinsa en noves aigües. Aquesta nova generació de bots de conversa fomenta les interaccions humà-màquina basades en el llenguatge natural i facilita el processament de sol·licituds (també anomenades *avisos*). Això és possible gràcies al model de llenguatge extens (en endavant, LLM). En resposta a peticions concretes, la IA facilita la generació de nous continguts, però basant-se en el processament de les dades existents, ja sigui un text, una imatge, un vídeo o un àudio. A mesura que tot tipus de contingut es fa més fàcil de generar, això comporta reptes addicionals per a la societat, especialment per a la democràcia, és a dir, com més contingut hi ha, també hi ha més desinformació. ChatGPT, un bot de conversa d'IA de l'empresa OpenAI, és el producte més conegut d'aquest tipus d'IA. Altres empreses també han entrat en aquest mercat.⁷

La IA a Àustria

Com en molts altres països europeus, el llançament de ChatGPT ha tingut molt de ressò en els mitjans de comunicació, en la població en general i en l'àmbit polític a Àustria. Els sistemes d'IA generativa representen un nou desenvolupament amb efectes socials que poden ser disruptius. Tanmateix, aquesta enorme atenció pública i les altes expectatives també enfosqueixen la visió de les avaluacions realistes de les possibilitats i els límits d'aquests sistemes. A Àustria, l'alt nivell d'atenció pública també ha donat lloc a un debat públic sobre el finançament de la investigació en IA a través dels mitjans de comunicació, i se n'han posat de manifest els possibles déficits.⁸

Els principals actors del panorama d'IA austríac són el Ministeri Federal de Finances (BMF), amb el secretari d'estat de Digitalització; el Ministeri Federal de Treball i Economia (BMAW), i el Ministeri Federal d'Afers Socials, Salut, Atenció i Protecció del Consumidor (BMSGPK). Tots els partits polítics del Parlament contribueixen activament al debat. IA-Àustria⁹ és un grup de reflexió i un lobby per a l'aplicació de la IA a Àustria que aplega actors provinents de la ciència, els negocis, l'educació i la societat. Els principals actors del món acadèmic són diverses universitats i altres institucions de recerca com ara l'Institut de Ciència i Tecnologia d'Àustria (ISTA), l'Institut Austríac de Tecnologia (AIT) i l'Acadèmia de Ciències d'Àustria (OeAW). I s'està fundant una nova universitat especialitzada a Linz, l'Institut de Ciències Digitals d'Àustria (IDSA).¹⁰

A Àustria el sector de recerca i desenvolupament no està a l'avantguarda mundial pel que fa a la IA, però té un cert potencial. Per exemple, a més d'alguns grups de recerca universitaris, hi ha algunes empreses emergents d'IA. No obstant això, els òrgans de seguiment informen que la intel·ligència artificial es manté en el primer lloc entre tendències d'innovació i tecnologia en el panorama de les empreses emergents austría-

7. <https://www.pcguides.com/apps/chat-gpt-alternative/>

8. <https://science.orf.at/stories/3218860/>

9. <https://aiaustria.com/>

10. <https://www.bmbwf.gv.at/Themen/HS-Uni/Aktuelles/idsa.html>

ques.¹¹ Un estudi recent detecta reptes considerables.¹² A més, es pot destacar un estudi fonamental sobre el potencial de les aplicacions d'IA en el sector dels mitjans, concretament AI.AT.Media, que va ser dut a terme pel Laboratori de Mitjans de l'Agència Austríaca de Premsa el 2021 (Krawarik *et al.*, 2021). Al principi de 2023, l'Agència Austríaca de Premsa (APA) va presentar la seva futura estratègia d'IA «APA Trusted AI for periodism and communication» [IA de confiança per al periodisme i la comunicació] (APA, 2023).

Com en molts països, a Àustria la IA es considera una tecnologia de propòsit general amb la corresponent importància per a la ubicació i la competitivitat empresarial del país (Huang/Peissl, 2023). La base de la política d'IA austríaca és l'Estratègia nacional d'IA de 2021 - Missió d'intel·ligència artificial per a Àustria 2030 (prevista per al 2030). Se centra en la consecució dels tres objectius següents:

1. Té com a objectiu un desplegament ampli de la IA orientat al bé comú, fet de manera responsable sobre la base dels drets fonamentals i humans, els valors fonamentals europeus i el futur marc legal europeu.
2. Àustria s'hauria de situar com un lloc de recerca i innovació per a la IA en àrees clau i àmbits de fortalesa.
3. I mitjançant el desenvolupament i l'ús de la IA, s'hauria de garantir la competitivitat de la tecnologia austríaca i la ubicació empresarial (BMK, 2021).

La Missió d'intel·ligència artificial per a Àustria té com a objectiu examinar l'ús segur de les aplicacions d'IA en els processos administratius públics. Per a promoure un sector públic modern i eficient, l'aprenentatge automàtic basat en dades i la IA s'utilitzaran per a automatitzar (parcialment) els processos de treball i la presa de decisions a l'Administració (BMK, 2021).

La IA en els debats del Parlament austríac

La qüestió de la IA s'ha plantejat diverses vegades al Parlament austríac, incloses preguntes sobre la IA i la pornografia infantil, l'ús de la IA en contextos educatius, la IA i la seguretat interior, els impactes en el màrqueting i el comportament dels consumidors i molts més problemes.¹³ Més recentment, va tenir lloc un fòrum obert sobre l'impacte de la intel·ligència artificial en la societat i la democràcia específicament centrat en ChatGPT. Per casualitat, l'endemà de l'acte, la Comissió Parlamentària de Recerca, Innovació i Digitalització també va tenir una compareixença d'experts sobre la qüestió.

El qüestionari de l'EPTA es va enviar als membres del Consell Assessor Parlamentari per a la Prospecció i l'Avaluació Tecnològica. L'enviament es va superposar amb l'inici de les vacances d'estiu del Parlament, de manera que, de les possibles cinc, només dues respostes van arribar a temps a l'Institut d'Avaluació Tecnològica (en endavant, ITA), provinents dels grups parlamentaris Verds i Socialdemòcrates. Els dos diputats coincideixen a predir que la IA té potencial per a perjudicar la democràcia. També hi veuen un potencial per a promoure la democràcia. En particular, assenyalen les notícies falses i els hipertrucatges i el perill que comporten per a la democràcia. En conseqüència, la comunicació política, el mercat laboral, la política educativa i l'ús militar de la IA es consideren àrees especialment sensibles. Un diputat cita la protecció de dades pel que fa a les dades d'educació i salut, per a garantir sistemes i dades de formació d'alta qualitat, evitar la dependència de les corporacions de tecnologia de la informació i promoure solucions de codi obert. Els possibles guanys en qualitat i eficiència s'haurien de distribuir d'una manera socialment equitativa i, sobretot en el cas de les aplicacions en l'àmbit mèdic, la responsabilitat final hauria de recaure en el metge. Al Parlament austríac ja hi ha hagut alguns debats sobre la IA (vegeu més amunt). Pel que fa als coneixements necessaris, ambdós grups de diputats van indicar que volien estar informats sobre els aspectes tècnics fonamentals, les implicacions i les possibles àrees d'aplicació. També es van reconèixer

11. <https://austrianstartupmonitor.at/wp-content/uploads/2023/04/ASM-2022.pdf>

12. <https://www.ipa.fraunhofer.de/de/Publikationen/studien/ki-austria.html>

13. https://www.parlament.gv.at/suche?VTS_01searchType=all&VTS_01searchScope=all&VTS_01category=Parlamentskorrespondenz&VTS_01gp_liste=XXVII&VTS_01search=Intel·ligència artificial

els valors subjacents en el disseny d'aquests sistemes i se li va donar un paper important a l'avaluació tecnològica (en endavant, AT).

Finalment, però no menys important, els informes de seguiment lliurats al Parlament semestralment cobrien molts aspectes de la IA, inclosa la IA generativa,¹⁴ els hipertrucatges, la IA en salut, els riscos de la IA i altres. Aquestes línies temàtiques van despertar un interès especial entre els diputats. Per exemple, l'hipertrucatge es va tractar en una moció parlamentària.¹⁵

Rellevància i debat social i polític

ChatGPT va conscienciar molt la societat austríaca. Moltes notícies i comentaris en els mitjans van obrir el camí a debats polítics i públics al Parlament. Van conduir a debats sobre casos d'ús específics en ONG, institucions de formació d'adults i altres entorns. No obstant això, en una enquesta del març de 2023,¹⁶ al voltant del 18% dels enquestats va dir que ja feia servir ChatGPT, i la xifra dels menors de vint-i-vuit anys (generació zeta) se situava al voltant del 34%. En general, els austríacs enquestats utilitzen bots de conversa per a traduir i compondre texts en llengües estrangeres (75%), com a motor de cerca (73%) o per a compondre texts inventats amb finalitats d'entreteniment (61%). Un gran escepticisme és evident pel que fa a les pors de frau (78%), les notícies falses (77%), la manca de privadesa (67%) i l'impacte ambiental (61%). El 81% demana regulacions més estrictes de ChatGPT i IA per a garantir la privadesa, la neutralitat i la correcció.¹⁷

El govern federal (anterior) va establir el Consell Austríac de Robòtica i IA i va presentar el llibre blanc «How to Shape Austria's Future with Robotics and AI Positively» [Com donar forma positiva al futur d'Àustria amb la robòtica i la IA] el 2018.¹⁸ A més, s'ha encarregat un estudi sobre la IA a Àustria.¹⁹ El 2021, l'anterior govern federal va publicar finalment la «Artificial Intelligence Mission Austria 2030» [Missió d'intel·ligència artificial per a Àustria 2030] (BMK, 2021).

De moment no hi ha cap legislació especial sobre la IA a Àustria. Igual que altres països de la UE, Àustria participa en el desenvolupament de la llei de la IA en l'àmbit europeu. A més, hi ha una proposta de la Secretaria d'Estat de Digitalització de crear una autoritat d'IA austríaca amb caràcter de servei per a preparar els ciutadans i les empreses. Per exemple, recopilaria totes les institucions que hagin obtingut la certificació d'IA. També s'està plantejant un segell de qualitat austríac per a la IA. L'obligació de transparència també es considera cabdal quan s'aborda la IA.²⁰ Des de l'àmbit acadèmic, hi va haver un llibre blanc de 2021 de la Universitat Tècnica de Viena i la Universitat Johannes Kepler de Linz en aquest sentit, que suposadament impulsava el primer certificat comercial per a sistemes d'IA, que es va debatre àmpliament a tota la UE.²¹

El paper de l'AT en els debats

Fa anys l'ITA que s'ocupa de la IA en general i de qüestions específiques (sobre treball, institucions de benestar, educació, etc.). Concretament sobre ChatGPT, l'ITA va publicar una breu informació el maig de 2023 per al Parlament austríac com a part de l'activitat de seguiment semestral (vegeu més amunt). Col·

14. https://www.parlament.gv.at/document/fachinfos/zukunftsthemen/143_generative-ki.PDF

15. https://www.parliament.government.at/PAKT/VHG/XXVII/E/E_00104/

16. <https://science.apa.at/power-search/14071413255035448015>

17. https://www.ots.at/presseaussendung/OTS_20230328_OTS0064/pwc-chatgpt-ki-studie-fast-zwei-drittel-der-oesterreicherinnen-sind-fuer-verbot-an-schulen-jeder-dritte-fuerchtet-um-arbeitsplatz

18. <https://www.bmk.gv.at/themen/innovation/publikationen/forschungspolitik/ki/whitepaper.html>

19. <https://www.bmk.gv.at/themen/innovation/publikationen/ikt/ai/zahlen.html>

20. https://www.parlament.gv.at/aktuelles/pk/jahr_2023/pk0740

21. https://www.jku.at/fileadmin/gruppen/219/LIT_AI_Lab/News_Seite/White_Paper_-_Trusted_Artificial_Intelligence_-_Towards_Certification_of_Machine_Learning_Applications_web_s.pdf i <https://www.elise-ai.eu/events/first-commercial-certificate-for-ai-systems-now-on-the-market>

legues de l'ITA van fer aportacions a la compareixença d'experts del TAB per a preparar el primer informe parlamentari sobre IA generativa per al Bundestag alemany.

Entre els estudis de l'ITA sobre la IA en els últims cinc anys hi ha:

- «Critical AI Literacy» [Alfabetització bàsica en IA] (2023-2024)²² (Strauß, 2021)
- «Automating Welfare» [Automatització del benestar] (2022-2025)²³
- Intel·ligència artificial: «Towards more comprehensibility and transparency» [Cap a més comprensibilitat i transparència] (2021-2022) (Udrea *et al.*, 2022)²⁴
- «Technology Assessment in Austrian AI Start-ups» [Avaluació tecnològica a les empreses emergents d'IA austríaques] (2020-2021) (Tanja Sinozic *et al.*, 2022)²⁵
- «The AMS [Labour Market Service Authority] algorithm» [L'algoritme de l'Autoritat del Servei del Mercat de Treball] (2019-2020) (Allhutter *et al.*, 2020)²⁶
- «Artificial intelligence and labour» [Intel·ligència artificial i treball] (2018-2019) (Cas/Krieger-Lamina, 2020)²⁷

A més dels informes anteriors i nombroses invitacions a grups d'experts (per exemple, a la UNECSO), conferències públiques, jornades, etc., l'ITA va ser convidat a comparèixer a una sessió sobre IA generativa de la Comissió Parlamentària d'Investigació, Innovació i Digitalització (amb experts externs i personal de l'ITA). Els detalls funcionals bàsics dels LLM, el panorama europeu i les reflexions ètiques sobre la IA, en general, van ser la base dels debats. Els temes de transparència, desenvolupament de competències i investigació van ser fonamentals per a moltes preguntes de diputats. A més, sota l'epígraf «Aptitud en la IA», els diputats també van debatre l'acceptació i la competència dels usuaris amb els experts.²⁸ L'ITA també va ser convidat pel Tribunal de Comptes d'Àustria per a impartir una ponència sobre els usos de la IA a l'hora de fiscalitzar i la fiscalització de la IA en el sector públic.

Les lliçons apreses de l'AT

El fet que la política i la ciència tenen lògiques i línies temporals diferents s'ha tornat a evidenciar en aquesta qüestió. L'AT sempre intenta estar a punt amb la informació en el moment adequat, però de vegades l'AT ha de ser encara més ràpid que abans. El llançament públic de ChatGPT i la seva capacitat disruptiva va produir una gran demanda política per a saber-ne més. Així, sempre que l'AT ja hagi fet algun treball anterior en un aspecte similar o més general, és més fàcil reaccionar ràpidament a un tema nou en particular, com ara els LLM. Tractar molts altres camps de la tecnologia també pot ajudar a fer les preguntes adequades en un tema específic. En definitiva, l'experiència metodològica i procedimental ajuda tant com el coneixement en una àrea concreta de recerca.

22. <https://www.oeaw.ac.at/en/now/projects/current-projects/cail-critical-ai-literacy>

23. <https://www.oeaw.ac.at/en/now/projects/auto-welf>

24. <https://www.oeaw.ac.at/en/ita/projects/artificial-intelligence>

25. <https://www.oeaw.ac.at/en/ita/projects/artificial-intelligence-technology-assessment-in-austrian-ai-start-ups>

26. <https://www.oeaw.ac.at/en/now/projects/ams-algorithm>

27. <https://www.oeaw.ac.at/en/ita/projects/the-social-effects-of-artificial-intelligence>

28. https://www.parlament.gv.at/aktuelles/pk/jahr_2023/pk0740

Bibliografia

- Allhutter, D., Mager, A., Cech, F., Fischer, F. i Grill, G., 2020, *Der AMS-Algorithmus – Eine soziotechnische Analyse des Arbeitsmarktchancen-Assistenz-Systems (AMAS)* [L'algoritme AMS: Una anàlisi socio-tècnica del Sistema d'Assistència i d'Oportunitats del Mercat de Treball (AMAS)], novembre del 2020, Viena: ITA.
- APA, 2023, *Leitlinie zum Umgang mit Künstlicher Intelligenz* [Directrius per a tractar la intel·ligència artificial] APA - Agència Austríaca de Premsa.
- BMK, 2021, *Strategie der Bundesregierung für Künstliche Intelligenz – Artificial Intelligence Mission Austria 2030* [Estratègia del Govern federal per a la intel·ligència artificial: Missió d'intel·ligència artificial per a Àustria 2030] (AIM AT 2030), <https://www.bmk.gv.at/themen/innovation/publikationen/ikt/ai/strategie-bundesregierung.html>.
- Cas, J. i Krieger-Lamina, J., 2020, «KI und Arbeitswelt» [La IA i el món laboral], a: Christian M.; Mader C.; Cas J.; Abou-Chadi T.; Bernstein A.; BraunBinder N.; Dell'Aglio D.; Fabián L.; George D.; Gohdes, A. (Ed.): *Wenn Algorithmen für uns entscheiden: Chancen und Risiken der künstlichen Intelligenz* [Quan els algoritmes decideixen per nosaltres: oportunitats i riscos de la intel·ligència artificial], Zurich: vdf, 144-164.
- Huang, L. i Peissl, W., 2023, «Artificial Intelligence – A New Knowledge and Decision-Making Paradigm?» [Intel·ligència artificial: Un nou paradigma de coneixement i presa de decisions?], a: Hennen, L., Ladikas, M., Hahn, J., Peissl, W., Est, R. v. i Lindner, R. (Eds): *Technology Assessment in a Globalized World – Facing the Challenges of Transnational Technology Governance* [Avaluació de la tecnologia en un món globalitzat: enfrontar els reptes de la governança tecnològica transnacional]: Springer, 175-201, <https://link.springer.com/content/pdf/10.1007/978-3-031-10617-0.pdf>.
- Krawarik, V., Schell, K., Ertelthaler, V., Thallinger, G. i Bailer, W. (BMK), 2021, *AI.AT.Media – AI and the Austrian Media Sector: Mapping the Landscape, Setting a Course* [AI.AT.Media. IA i el sector dels mitjans de comunicació a Àustria: Radiografiar el paisatge, fixar el rumb].
- Strauß, S., 2021, «Don't let me be misunderstood – Critical AI literacy for the constructive use of AI technology» [No deixeu que se'm malinterpreti: Alfabetització bàsica en la IA per a l'ús constructiu de la tecnologia d'IA], *TATuP - Zeitschrift Für Technikfolgenabschätzung In Theorie Und Praxis* 30, 44-49.
- Tanja Sinozic, Steffen Bettin i Udrea), T., 2022, *Cultures of responsibility for AI start-ups* [Cultures de responsabilitat per a les empreses emergents d'IA], encarregat per: ITA, Viena.
- Udrea, T., Fuchs, D. i Peissl, W., 2022, *Künstliche Intelligenz – Verstehbarkeit und Transparenz* [Intel·ligència artificial: Comprensió i transparència], Viena: ITA.

Catalunya

Consell Assessor del Parlament sobre Ciència i Tecnologia (CAPCIT)

Intel·ligència artificial generativa per a l'educació: el cas de Catalunya

De què es tracta?

El llançament de ChatGPT ha permès que milions de persones experimentin la intel·ligència artificial (en endavant, IA) per primera vegada. La IA és un somni ancestral de la humanitat: delegar les nostres tasques desagradables en màquines perquè puguem dedicar el nostre temps a gaudir de la vida. ChatGPT és només un dels molts exemples d'IA generativa. Abans de l'onada actual de sistemes semblants a ChatGPT i durant més de quaranta anys, el camp de la IA ha intentat emular la intel·ligència seguint un enfocament simbòlic. El nostre raonament i pensament sembla basat en símbols, com a paraules que representen conceptes i relacions entre si. El programa de recerca de la IA simbòlica consistia a donar instruccions als ordinadors, escrites en llenguatges que combinaven aquests símbols, perquè els ordinadors poguessin decidir sobre les accions. No obstant això, la IA simbòlica ha demostrat que funciona molt malament quan s'enfronta a tasques per a les quals cal la percepció humana: veure, escoltar o sentir. Els humans fem aquestes tasques molt bé, però són difícils d'expressar com a manipulacions de símbols.

Una manera alternativa explorada en IA durant força temps és modelar computacionalment les estructures que veiem al cervell, les anomenades *xarxes neuronals*. Una neurona busca patrons en els senyals de les seves connexions entrants i, quan es detecta un patró, s'envia un senyal a les connexions sortints. Un comportament individual força senzill que, quan s'escala a milers de milions de neurones interconnectades, produeix el complex comportament combinat d'un cervell humà. Aquestes neurones s'estructuren en capes de senyals d'entrada i de sortida. Aquí és on apareix el concepte d'*IA generativa*. Quan es donen determinades entrades a la primera capa de neurones, aquests sistemes propaguen els senyals per a «generar» un senyal de sortida a la capa final.

Des dels anys 40, hem produït diversos models informàtics de neurones i hem desenvolupat algorismes per a determinar els paràmetres que hi ha en les connexions entre aquestes neurones. Els paràmetres determinen quantitativament les relacions entre els senyals entrants i sortints d'una sola neurona. L'accés a les dades i la potència de càlcul ha permès als investigadors de desenvolupar xarxes amb milers de milions de paràmetres. ChatGPT té 1.500 milions de paràmetres i GPT4 1,7 bilions de paràmetres. Nombres increïbles.

Com funciona? En termes senzills, en el context de la generació de text, aquestes grans xarxes aprenen les paraules més probables que continuen una seqüència determinada de paraules i utilitzen trucs algorítmics per a seleccionar-ne una. Els sistemes iteren el procés afegint la paraula acabada de generar a la seqüència d'entrada anterior per a generar una altra paraula, etcètera. Com que el text sobre el qual s'entrenen les xarxes és generat per humans (web, llibres, diaris), el text generat «sembla» generat per humans, i de la mateixa manera per a imatges, combinacions d'imatges i text o vídeos.

La IA generativa ha tingut un impacte social important en l'educació des del principi. Com que el seu rendiment era tan semblant al contingut generat per humans, els educadors tenien por de no poder distingir el contingut creat per humans del contingut generat per una IA. Alguns països van optar per prohibir aquests sistemes a l'aula, per exemple, alguns estats dels EUA; semblava que els sistemes intentaven detectar contingut generat per IA, per exemple, DetectGPT per a text. Les coses han canviat en els últims mesos i els educadors veuen aquests sistemes i la IA en general com una oportunitat en l'educació.

Quin és l'estat de la qüestió

L'any 2019, la Generalitat de Catalunya va definir l'estratègia catalana d'IA, anomenada [CATALONIA.AI](#), al voltant de diversos eixos. Entre aquests eixos, hi ha el suport a l'ecosistema, l'impuls a la recerca i la innovació, l'atracció de talent i la promoció d'un enfocament ètic de la IA. Aquesta estratègia identifica quatre mecanismes: CIDAI, per a promoure la innovació; AIRA, per a promoure la recerca i l'atracció de talent; OEIAC, per a supervisar un enfocament ètic de la IA, i DCA-AI, per a integrar les comunitats.

Catalunya té un ecosistema vibrant, amb quinze centres de recerca que treballen en IA i grups de recerca a totes les universitats. Des del 1994, Catalunya té la seva pròpia associació independent d'IA, ACIA, que és membre de l'Associació Europea d'IA, EurAI.

Segons l'ACCIÓ, actualment Catalunya té un sector TIC fort amb més de 22.000 empreses que representen el 12% del PNB, amb un creixement molt ràpid. Tot i que ens falten dades més recents, ja el 2018 el sector de la IA representava més 1.000 milions d'euros i donava feina a més de 8.000 persones. Amb la seva infraestructura de recerca, industrial i informàtica, Catalunya és un centre atractiu per a les empreses d'IA.

Des que van aparèixer els primers treballs sobre l'aprenentatge profund –la base tecnològica de la IA generativa–, diversos grups de recerca catalans hi han fet aportacions importants. L'investigador català amb més impacte en IA generativa és, sens dubte, Oriol Vinyals, format a la Universitat Politècnica de Catalunya (en endavant, UPC), que actualment treballa a Google. El seu treball ha influït en la recerca posterior a la UPC i a molts altres centres de recerca catalans. Els investigadors catalans han estat fent recerca fonamental i aplicant models generatius per a crear continguts com ara material artístic, música o imatges.

Un projecte molt significatiu de producció de models de llenguatge extens (en endavant, LLM) per al català és el projecte Aina, coordinat pel Barcelona Super Computing Center (BSC). El projecte Aina té un finançament raonable i els seus models s'obriran a tothom, incloses les empreses interessades a desenvolupar aplicacions.

Pel que fa a l'educació, el sector educatiu, inclòs el Departament d'Educació, està reaccionant positivament a l'onada d'IA i d'IA generativa i veu oportunitats en l'aplicació de la IA en l'educació. Un senyal molt positiu és que el mes de maig passat el Departament d'Educació va obrir per primera vegada una convocatòria de recerca per a nivells no universitaris i inclou la intel·ligència artificial com un dels seus objectius de recerca.

Grups d'interès a Catalunya

Són nombrosos els actors del panorama de la IA generativa a Catalunya. En primer lloc, hi ha dos departaments governamentals principals amb responsabilitats. El Departament de Recerca i Universitats, per mitjà del director general de Recerca, i el Departament d'Empresa i Treball, per mitjà de la Secretaria de Polítiques Digitals, són els encarregats d'aplicar l'estratègia catalana en IA.

En l'àmbit acadèmic, diversos departaments universitaris estan investigant la IA i l'apliquen en diverses àrees del coneixement. En l'estratègia catalana sobre IA, com s'ha dit més amunt, els actors industrials s'organitzen al voltant del CIDAI i els acadèmics al voltant d'AIRA. ACIA, l'associació civil sobre IA, ha aplegat investigadors d'aquest camp des de 1994.

Des del març de 2023, la ciutat de Barcelona és un node de la Xarxa Europea ELLIS. ELLIS Barcelona està cridada a ser la principal interlocutora en IA generativa. El node integra investigadors dels principals centres de recerca i universitats amb un fort reconeixement internacional en IA, especialment en aprenentatge automàtic.

Moltes empreses han establert departaments d'investigació sobre IA a Catalunya o han mostrat un fort interès a utilitzar la IA en el seu negoci, com ara Amazon, Huawei, Microsoft, NTT i SAP.

En l'àmbit educatiu, cal destacar que, el març de 2023, el Departament d'Educació i el Departament de Recerca i Universitats van signar un conveni per a facilitar les activitats de recerca a l'ensenyament obligatori. La IA és entre les prioritats establertes en aquest acord.

Importància per a Catalunya

Com ja s'ha dit, el sector TIC és molt dinàmic i estratègic per a Catalunya. Hi ha un teixit ric d'importants desenvolupadors de programari i de petites empreses emergents amb una forta capacitat per a generar solucions noves i innovadores. Barcelona es troba entre les cinc ciutats preferides d'Europa per a establir una empresa. La IA està creant i continuarà creant una quantitat important de llocs de treball qualificats. L'onada d'IA generativa, sens dubte, alimentarà la creació de noves empreses i desenvolupaments. L'estratègia catalana d'IA ha identificat clarament la importància de la IA.

L'aplicació de la IA a l'educació també és molt important per a Catalunya. La UNESCO, ja l'any 2019, en un document signat per més de cent països, va advertir sobre la responsabilitat dels governs d'introduir la IA en els currículums. A mesura que el ritme de desenvolupament s'accelera, calen noves reflexions sobre com adaptar els continguts als conceptes d'IA generativa. Catalunya té una llarga tradició de mètodes educatius innovadors, i la introducció de la IA en els currículums tant de primària com de batxillerat ha cridat l'atenció dels educadors i ja és objecte d'estudi per alguns grups de treball del Departament d'Educació.

Debat en curs

Una gran part de la població percep els sistemes d'IA generativa com l'inici de la IA, de manera que es compleix la promesa de la IA. Aquesta percepció genera un cert debat dins de la comunitat d'IA.

Com en ocasions anteriors en la història de la IA, un èxit particular amb una tecnologia específica fa que la gent cregui que «això és tot», hem aconseguit «la veritable IA». Si ho analitzem amb una mica de detall, podem veure que no és així (encara). Les capacitats d'intel·ligència humana que la IA vol automatitzar es poden dividir en físiques i intel·lectuals. Una veritable IA hauria de fer el mateix que els humans en ambdós àmbits. Vegem els models de llenguatge extens (també coneguts com LLM, per exemple, ChatGPT) o els sistemes de generació d'imatges (com MidJourney). Són sistemes sense corporació que no estan preparats per a resoldre els reptes dels sistemes robòtics. Pel que fa a les capacitats intel·lectuals, aquests sistemes mostren capacitats de raonament molt superficials; desconeixen el seu món, no tenen sentit comú i no han experimentat res. Manipulen el llenguatge natural de maneres impressionants, però no són el camí integral cap a la veritable IA. En poden ser un component, però no la solució completa.

Tenint en compte les expectatives que havia generat ChatGPT en la societat, molts investigadors i personatges públics rellevants van signar una carta emesa el març de 2023 en què demanaven una moratòria sobre l'alliberament públic dels sistemes d'IA generativa fins que hi hagués una regulació adequada. Aquesta carta va generar una onada d'interès en els mitjans de comunicació i reaccions per part de les grans empreses, que van argüir que s'autoregularien millor. Aquestes reaccions s'han vist com una maniobra per a evitar la regulació als EUA.

Un altre aspecte significatiu del debat públic ha estat el posicionament dels tres «pares» de l'aprenentatge profund que van rebre el premi Turing (una mena de premi Nobel d'informàtica), Bengio, Hinton i LeCun. L'aprenentatge profund és la tecnologia darrere de tots els sistemes d'IA generativa. Bengio va signar la carta esmentada anteriorment, Hinton va renunciar a Google i després va advertir sobre els perills de desinformació que podrien comportar els LLM. LeCun proposa tornar a propostes més antigues en IA basades en arquitectures cognitives per a superar les limitacions de la xarxa neuronal.

El debat a Catalunya es manté en la mateixa línia, i els investigadors creuen que s'ha de posar en marxa una regulació per a evitar els possibles perills dels sistemes d'IA generativa.

IA al Parlament

El Parlament de Catalunya encara no ha legislat directament en aquest àmbit. Dit això, al llarg dels darrers cinc anys la IA ha aparegut cada cop més en els debats parlamentaris, especialment amb relació a les anomenades *mocions* i *resolucions* que s'utilitzen per a afavorir una determinada acció del Govern. Des del 2019, el Parlament ha demanat a la Generalitat que estableixi una estratègia sobre IA seguint el pla de co-

ordinació d'IA de la Comissió Europea (Moció 66/XII).²⁹ Posteriorment, durant la pandèmia, el Parlament va demanar al Govern que utilitzés la tecnologia d'IA per a fer vigilància epidemiològica (Resolució 940/XII).³⁰ En termes diferents, també es va plantejar la IA com una oportunitat per a promoure l'ús de la llengua catalana i per a garantir que els ciutadans poguessin interactuar amb els bots de conversa en la seva llengua materna (Resolució 911/XII).³¹ Finalment, es poden destacar dues iniciatives parlamentàries més tenint en compte l'enfocament que tenen de la IA. D'una banda, el Parlament ha destacat la necessitat de treballar per un ús ètic de les dades, i també per garantir el respecte dels drets humans dins dels sistemes d'IA (Resolució 482/XIV). D'altra banda, el Parlament ha assenyalat la relació entre IA i el sector turístic (Moció 247/XIV), i aquesta és la iniciativa parlamentària més recent, que data del juliol de 2023.³²

El paper de l'avaluació tecnològica en els debats

El coneixement actual que mostren els responsables polítics a Catalunya sobre la IA generativa demostra que el terme encara no forma part del vocabulari parlamentari consolidat i que sovint es confon amb altres termes com IA, aprenentatge automàtic o desinformació. Hi ha la idea compartida que el Parlament ha d'abordar la qüestió i que el CAPCIT ha de tenir un paper central a l'hora d'informar els polítics. Sovint se subratlla el costat negatiu de la IA, especialment pel que fa als perills per a la democràcia. Les oportunitats que representa són evidents per a la majoria dels responsables polítics, i també la necessitat que hi hagi una regulació. Les oportunitats identificades són nombroses i inclouen la millora de la salut, l'educació i l'Administració pública. En el cas concret de l'educació, hi ha un consens que el seu ús és inevitable, però no hi ha consens sobre com s'ha d'introduir en el sistema educatiu. En l'aplicació de la IA a la salut es planegen la protecció de la privadesa de les dades i algunes preocupacions ètiques. L'impacte en el mercat de treball es veu principalment com oportunitats de millora i increment de l'eficiència. També s'ha assenyalat la possibilitat d'abordar reptes com ara el canvi climàtic.

Així, es percep que la IA en general i la IA generativa en particular són tecnologies que cal explicar als legisladors i a la societat catalana per l'impacte que ja tenen i que s'espera que tinguin en un futur proper. Per tant, el CAPCIT està cridat a tenir un paper cabdal, amb l'aportació de contribucions científiques per a guiar la presa de decisions polítiques en aquest camp.

29. [Publicació de l'aprovació en el Ple](#)

30. [Publicació de l'adopció en el Ple](#)

31. [Publicació de l'adopció en el Ple](#)

32. [Publicació de l'adopció en el Ple](#)

Dinamarca

Intel·ligència artificial generativa i democràcia: diputats danesos i una perspectiva de país

Aquest article introdueix i aborda diverses perspectives sobre intel·ligència artificial (en endavant, IA) generativa i democràcia. Dinamarca és interessant per a estudiar aquests temes, ja que el país és elogiat pel seu alt nivell de digitalització, a més d'haver ocupat un lloc destacat en nombrosos índexs de democràcia. Amb una tecnologia clau com ara la IA generativa, es produiran canvis significatius i, per tant, l'impacte en la societat i la democràcia és molt rellevant. Aquest article començarà amb una panoràmica general dels temes que connecten la IA generativa i la democràcia, tal com van sorgir en sis entrevistes amb diputats danesos. Posteriorment, els desenvoluparem fent referència a les reaccions d'altres parts interessades de la societat. Finalitzarem aquest article identificant el paper que pot tenir l'avaluació tecnològica (en endavant, AT) per a preparar els polítics perquè puguin abordar la IA generativa.

Problema(es) en joc

La posició digital i democràtica en què es troba Dinamarca col·loca el país en una situació única cada vegada que sorgeixen tecnologies habilitadores clau. L'aparició de la IA generativa no n'és una excepció. Per tant, és natural estudiar la manera com els polítics s'imaginen que serà l'impacte de la IA generativa en la democràcia a Dinamarca. Amb la nostra relació inherent amb la digitalització, no és estrany que els diputats danesos en general tinguin una actitud positiva i oberta envers la IA generativa. Tot i que no es va parlar gaire directament sobre la relació entre la IA generativa i la democràcia, es van tractar temes de geopolítica, llibertat d'expressió i educació pública com a preocupacions pel que fa a l'impacte negatiu que podria tenir la IA generativa en la societat.

Actituds positives sobre la IA generativa i la democràcia

Hi ha visions sobre la manera en què la IA generativa pot influir positivament en la democràcia. Es fa referència als factors d'accessibilitat i eficiència en àmbits en què l'aspiració és que la IA generativa pugui funcionar com a assistent personal. Començant pels mateixos polítics, poden utilitzar i ja utilitzen la IA generativa en l'acompliment de les seves tasques quotidianes. La utilitat positiva de la IA generativa no solament es pot trobar en el treball dels diputats, sinó també en la societat. La IA generativa pot empoderar les persones amb discapacitat, per exemple. La IA generativa també permet a les persones amb dislèxia d'expressar-se per escrit amb molta més confiança. A una escala social més àmplia, els diputats danesos veuen que la IA generativa pot obrir noves possibilitats per a copsar les opinions dels ciutadans sobre diversos temes. Els exemples esmentats durant les entrevistes són que la IA generativa pot permetre una nova forma d'enquestes públiques per a copsar l'opinió dels ciutadans, i que les eines d'IA seran més barates i accessibles, cosa que permetrà que més població participi en la conversa democràtica. A més, es va apuntar que les eines d'IA generativa podrien ajudar els ciutadans a adquirir més coneixements, i, per tant, capacitats, que poden empoderar la persona i potenciar-ne la participació en la democràcia.

Preocupació per l'impacte de la IA generativa en la democràcia

Dos temes de preocupació recurrents en les converses amb els diputats van ser la geopolítica i la llibertat d'expressió. Tot i que no era una preocupació directa, un altre tema persistent va ser la necessitat de fer canvis en el sistema d'educació primària perquè la societat estigui actualitzada amb el desenvolupament tecnològic.

La geopolítica ja no és només una qüestió de poder militar o de fronteres geogràfiques. Les tecnologies també tenen un paper cabdal en les tensions i els debats geopolítics. L'impacte de la IA generativa en l'àmbit de la geopolítica és motiu de preocupació. Els danesos utilitzen cada cop més les plataformes de xarxes socials per a trobar informació i notícies. Tal com són avui dia, aquestes plataformes ja han estat susceptibles a la propagació de desinformació, però amb la IA generativa encara és més fàcil generar desinformació plausible a gran escala i distribuir-ne per mitjà d'aquestes plataformes. La preocupació entre els diputats entrevistats era que això permetia als trols sancionats per règims estrangers influir en el debat polític danès, però també influir en la visió i la comprensió del món dels ciutadans danesos. Per exemple, amb les tecnologies de l'hipertrucatge, es fa cada cop més difícil distingir entre contingut fals i legítim, i aquest és un problema cabdal per al debat polític democràtic nacional. L'altra cara d'això és que les xarxes socials s'estan convertint en un indicador cada cop més pobre del sentiment popular general, i els polítics corren el risc de ser enganyats si utilitzen contingut de les xarxes socials per a navegar.

La difusió de la desinformació en les plataformes de xarxes socials planteja un dilema: d'una banda, les persones tenen dret a expressar les seves creences, encara que estiguin «equivocades», però, d'altra banda, cal protegir les persones de ser enganyades. El debat és com regular la IA generativa per a evitar la propagació de més desinformació, sense vulnerar el dret a la llibertat d'expressió, que a Dinamarca és un valor molt apreciat i una pedra angular de la nostra democràcia.

L'última preocupació està relacionada amb l'educació primària i la necessitat de garantir que la tecnologia i la comprensió de la tecnologia esdevinguin una part integrada de l'educació democràtica que reben tots els infants a Dinamarca. La preocupació és que l'expansió d'aquests coneixements fonamentals és lenta i va endarrerida en relació amb els desenvolupaments reals de la societat. Això no està relacionat amb la IA generativa específicament, sinó que és una preocupació general pel desenvolupament tecnològic en general.

Reaccions de la societat sobre IA generativa i democràcia

No només els diputats danesos participen en la qüestió de la IA generativa. Altres participants també han contribuït de diverses maneres al debat sobre la IA generativa i els tres temes de geopolítica, llibertat d'expressió i educació primària. En aquest apartat destacarem alguns dels participants que han estat actius en el debat sobre la IA generativa. El focus principal se centra en les aportacions als tres temes ja tractats anteriorment.

Les reaccions de la societat sobre la IA generativa són relativament noves, ja que aquesta tecnologia fa molt poc que s'ha fet accessible al públic amb ChatGPT. Diversos actors han investigat recentment o tot just estan començant a investigar la IA generativa i la implicació que té en la societat.

La qüestió de l'explotació de les dades dels usuaris per part de les grans empreses tecnològiques i com protegir els drets dels usuaris va ser debatuda el juny de 2023 per un grup d'experts establert pel Parlament danès, que va publicar tretze recomanacions per a regular i controlar les qüestions relacionades amb el model de negoci de les grans empreses tecnològiques. Volen evitar l'explotació dels infants per part de les grans empreses tecnològiques i evitar que influeixin en el comportament de la gent en línia, que té un vessant democràtic. El següent tema en el qual se centra el grup d'experts és com l'ús de la IA, i de la IA generativa específicament, per les grans empreses tecnològiques afecta la societat i com es pot gestionar.

La qüestió de la desinformació ha cridat l'atenció en diversos contextos. El Centre Nacional d'Ètica danès ha aconsellat sobre la regulació de la desinformació en línia i aviat començarà una investigació més àmplia sobre la IA generativa i la seva implicació en la societat, amb l'objectiu de publicar les seves recomanacions ètiques sobre l'ús de la IA generativa l'estiu de 2024. En un estudi sobre la vida digital dels infants aborda la democràcia, la desinformació i la responsabilitat digital, i conclou que la desinformació és present en vuit de cada deu infants, tot i que els infants no estan preocupats per la desinformació. L'estudi és del Projecte Algoritmes, Dades i Democràcia (Projecte ADD), un actor que ha reconegut la necessitat de fer una atenció continuada, per exemple per mitjà de publicacions i esdeveniments, sobre com la IA afecta la societat.

Estudia àmpliament la manera en què la nostra democràcia es veu influïda pel desenvolupament tecnològic. Està organitzant un esdeveniment perquè els actors socials debatin qui és el responsable de limitar la propagació de la desinformació en la nostra societat en un món d'IA generativa, que tindrà lloc el primer trimestre del 2024 en col·laboració amb el Consell de Tecnologia de Dinamarca.

El Consell de Ministres Nòrdic ha participat en el debat sobre l'afectació de la IA generativa en la democràcia. Com que les democràcies dels països nòrdics comparteixen una característica clau de tenir un alt nivell de confiança per part de la ciutadania, el Consell ha expressat una gran preocupació per un possible mal ús de la IA generativa per a manipular o fins i tot socavar les converses democràtiques als països nòrdics. Això va molt en línia amb la preocupació que van expressar els parlamentaris danesos.

Entre altres actors que han entrat en el debat, i especialment en la qüestió de la IA generativa relacionada amb l'educació primària, hi ha els mestres a Dinamarca. El seu enfocament és una mica diferent de la manca d'integració de la tecnologia en el currículum, ja que la seva preocupació és més aviat com integrar, o no integrar, les eines d'IA generativa en la docència. El Ministeri d'Educació s'ha negat a fer directrius clares que els professors hagin de seguir, però ha anunciat que aviat es faran recomanacions per a l'ús d'eines d'IA generativa.

Dins de l'àmbit educatiu, s'ha debatut molt com regular la IA generativa per a utilitzar-la amb finalitats educatives i utilitzar-la en els exàmens. La majoria dels debats sobre el seu ús s'han fet amb relació a l'educació primària, secundària i secundària superior, mentre que l'ús per als exàmens també s'ha debatut amb relació a les universitats. La necessitat de comprendre la tecnologia ha estat un debat constant durant diversos anys a Dinamarca, i ha ressorgit amb força amb l'aparició de la IA generativa. El que ha tingut poca o cap conscienciació és la necessitat de millorar les competències generals per a totes les persones, independentment de la seva posició en la societat.

Tot i que ja no és un actor actiu, val la pena esmentar la comissió independent TechDK de Djøf, amb membres d'empreses, mitjans de comunicació, universitats, ONG i altres associacions. Criticava molt els polítics per la seva falta d'actuació en el camp de les grans empreses tecnològiques i les tecnologies emergents. Va impulsar directament una col·laboració entre les grans empreses tecnològiques, els mitjans de comunicació i les universitats amb l'objectiu de combatre els perfils falsos en línia i la difusió d'informació falsa. A més, va fomentar l'educació dels estudiants de primària i secundària pel que fa a la identificació de la propaganda en línia i les normes de llibertat d'expressió i censura.

Els actors són actius en diversos graus amb relació a aspectes de la IA generativa i, sobretot, sobre la IA en general, però hi ha molt debat en fòrums tancats i exclusius on el públic poques vegades és convidat. A part de l'estudi de població del Projecte ADD i d'alguns estudis sobre el comportament en línia dels infants, a Dinamarca no s'ha estudiat gaire la percepció dels ciutadans amb relació a la IA generativa. El Consell de Tecnologia de Dinamarca ha fet en dues ocasions en els darrers anys una campanya de sensibilització sobre el tema de la IA en dos camps específics. Primerament, sobre la qüestió de la IA en general i la IA per a la investigació mèdica i sanitària durant el Projecte cervell humà 2019-2020. Més endavant, amb el desenvolupament d'un bot de conversa basat en IA per a identificar la desinformació en el projecte Titan el 2023. Aquests exercicis de participació pública s'han dut a terme per a obtenir informació sobre les opinions dels ciutadans danesos (i dels ciutadans de països europeus seleccionats) sobre el desenvolupament i l'ús de la IA en diversos contextos que tenen una usabilitat més àmplia que per als projectes de recerca respectius. Els coneixements resultants d'aquests esforços donen una idea sobre la percepció de la societat pel que fa a una tecnologia emergent com la IA generativa.

El paper de l'AT en el debat danès sobre IA generativa i democràcia

Hi ha una assumpció general que Dinamarca té el potencial de convertir-se en un líder de la transformació digital que aportarà la IA generativa. No obstant això, els parlamentaris són conscients que amb el ràpid desenvolupament tecnològic, Dinamarca ja està dos passos enrere en l'adaptació de les legislacions a la nova realitat que comporta la IA generativa. Tot i que els diputats sembla que estan d'acord a fer un seguiment de

la propera legislació de la UE sobre IA generativa (la llei de la IA), veuen que Dinamarca sola o en col·laboració amb els veïns nòrdics pot dur a terme actuacions sobre la pròpia regulació que la UE (encara) no cobreix.

És evident que els polítics no poden ser experts en tots els camps en què estan implicats i, per tant, hi ha una tasca ingent perquè la informació externa s'introdueixi al sistema polític perquè estiguin informats sobre els matisos de qualsevol tema, i la IA generativa no n'és una excepció. Ni la IA ni específicament la IA generativa són conceptes nous, però per als diputats entrevistats és un camp en què encara no estan familiaritzats del tot. I, per aquest motiu, l'AT té un paper a l'hora de preparar els diputats per a entendre millor la tecnologia. Era clar que els diputats, tot i haver provat i utilitzat serveis d'IA generativa, no distingien entre IA generativa i IA general. Cal ajudar a desenvolupar el vocabulari tècnic i la imaginació dels diputats, perquè puguin entendre adequadament en quin aspecte la IA generativa és un canvi radical. El que l'AT té com a finalitat única en aquest sentit és proporcionar anàlisis holístiques de la tecnologia, tant amb relació a aplicacions específiques com en general per a la societat. A partir de les entrevistes dels diputats, va quedar clar que les capacitats de previsió de l'AT, per a fer prediccions informades sobre les possibles afectacions de la IA generativa en les nostres societats en una àmplia gamma de paràmetres a mitjà i llarg termini, són necessàries per a oferir als parlamentaris una comprensió de la tecnologia que els prepari adequadament per a establir una guia sobre com volem utilitzar la tecnologia i quins possibles resultats volem limitar. Part d'això també consisteix a presentar oportunitats concretes que aporta la IA generativa, bons exemples de casos d'altres països i coneixements sobre quines competències necessiten les autoritats en el futur. Hi ha una necessitat d'informació tàcita, fàcilment accessible i en un llenguatge planer que els polítics puguin fer servir com a base per a participar en converses sobre IA generativa.

Les tecnologies emergents requereixen debats democràtics amplis i oberts que impliquen múltiples actors per a descobrir els matisos dels seus possibles avantatges i riscos, i aquest és un altre paper clau que podria tenir l'AT. En una societat sistemàticament digitalitzada com la danesa, és probable que la IA generativa tingui un impacte immediat i ampli. Per aquest motiu, la participació dels ciutadans i les parts interessades amb relació a la IA generativa és molt rellevant per a Dinamarca. Això amb relació als principis generals d'ús, construcció de perspectiva, establiment de l'agenda de recerca oberta i participació en temes relacionats amb sectors específics; per exemple, principis per a utilitzar i aprendre sobre la IA generativa en el sistema educatiu. A Dinamarca hi ha grans esperances que la tecnologia d'IA generi molts beneficis per a la societat, però adonar-se d'aquestes esperances depèn de portar els ciutadans i les parts interessades a un debat sobre què es pot fer i com. De la mateixa manera, hi ha una gran oportunitat d'incloure la societat en la definició de visions positives de la tecnologia, i entendre quin és el grau d'acceptació de la tecnologia.

Com es pot veure en aquest document, hi ha diverses esperances i preocupacions sobre com pot la IA generativa afectar la societat danesa i la nostra democràcia. Des de preocupacions sobre la influència exterior a causa de les tensions geopolítiques, fins a les esperances que la IA generativa pugui potenciar i augmentar la inclusió dels ciutadans vulnerables. Les possibilitats són moltes, però totes convergeixen en un punt: per mitjà de la participació en converses democràtiques amb múltiples actors, els diputats danesos poden estar més preparats per a prendre decisions informades que impedeixin que les nostres preocupacions es converteixin en realitat i obrir el camí perquè les nostres esperances esdevinguin el futur.

Parlament Europeu

Intel·ligència artificial generativa. Oportunitats, riscos i reptes polítics

Contribució del Panel per al Futur de la Ciència i la Tecnologia del Parlament Europeu (STOA)

Descripció del sistema/problema en qüestió

De què es tracta?

La Intel·ligència Artificial Generativa (IA) és una branca de la IA que se centra en el desenvolupament d'algoritmes capaços de generar resultats originals i creatius, com ara textos, imatges, música, etc. Un dels exemples més destacats d'IA generativa són els models de llenguatge extens (Large Language Models, LLM), que han estat al capdavant dels avenços recents en la investigació sobre IA. Els models de llenguatge extens, com ara ChatGPT d'OpenAI, es creen utilitzant tècniques d'aprenentatge profund i s'entrenen en conjunts de dades massives que contenen grans quantitats de text de fonts diverses. Aquests models estan dissenyats per a generar textos similars als humans mitjançant la predicció de la paraula següent o de la seqüència de paraules més probable atès un context específic. Els LLM han demostrat capacitats impressionants en diverses aplicacions, com ara la comprensió del llenguatge natural, la resposta a preguntes, la generació de continguts i la traducció.

Què en pensen els diputats?

Les respostes següents s'han elaborat a partir dels comentaris d'alguns dels nostres diputats:

1. Què és el primer que us ve al cap quan sentiu «intel·ligència artificial generativa»?

- Models de fundació
- IA per a usos generals
- Models de llenguatge extens
- GPT
- Més enllà de la IA discriminatòria
- Qüestions serioses sobre la naturalesa de la realitat i el seu impacte en les nostres societats
- La por de quedar atrapat en un món fictici

2. Creieu que hi ha una relació entre IA generativa i democràcia? Podríeu explicar breument què en penseu?

- La intel·ligència artificial generativa es podria utilitzar per a crear una quantitat massiva de notícies falses, publicacions a les xarxes socials i altres formes de contingut que difonguin informació errònia i sembrin discòrdia. Això podria socavar la confiança del públic en les institucions democràtiques i dificultar la presa de decisions informades.
- La IA generativa es podria utilitzar per a enviar missatges personalitzats dissenyats per a influir en les opinions de les persones. Es podria utilitzar també per a influir en les eleccions, promoure polítiques específiques o, fins i tot, reprimir la participació electoral.
- D'altra banda, la intel·ligència artificial generativa es podria utilitzar per a crear materials educatius, millorar el compromís cívic i facilitar la comprensió de problemes polítics complexos.
- La IA generativa es podria utilitzar per a analitzar dades governamentals i identificar patrons i tendències que, d'altra manera, serien difícils de veure. Podria ajudar a millorar la transparència i la rendició de comptes del Govern.

- Per descomptat! Amb una IA generativa que domini bé el llenguatge, del qual depenen en última instància les nostres converses polítiques i les nostres democràcies, és possible que no sempre es pugui determinar si estem parlant amb un humà o amb una IA.
- Com diu Yuval Harari, la IA pot fomentar una intimitat falsa amb els humans per a intentar convèncer-los d'idees, opinions o candidats concrets. Tot plegat és un greu problema quan es tracta del nostre discurs públic i, per extensió, de les nostres democràcies.

3. Si haguéssiu de pensar en àrees polítiques on probablement s'haurà de tractar la IA generativa, quines serien?

- Privadesa/hipertrucatges
- Ocupació/potencial per a automatitzar moltes feines
- Educació/aprenentatge personalitzat, garantir que tots els estudiants tinguin accés a eines d'IA generativa
- Sanitat: descobriment de fàrmacs i medicines personalitzats, garantiment de la seguretat i l'ètica
- Propietat intel·lectual / protegir els creadors i fomentar alhora la innovació
- Ensenyament
- Indústries de la creació
- Drets de propietat intel·lectual
- Discriminació
- Desinformació

4. Us imagineu com pot influir la IA generativa en el desenvolupament/estructura del mercat laboral / sector de l'ocupació/treball?

- Molts treballs d'oficina requeriran menys mà d'obra.
- Les tecnologies actuals d'IA podrien automatitzar almenys el 20% de les hores de treball.
- Els treballs manuals seran molt més lucratius.
- Sens dubte, els models lingüístics i la generació d'imatges, àudio o vídeo tenen el potencial de sacsejar el mercat laboral, amb conseqüències tant positives com negatives. Tot i que les feines menys qualificades poden estar encara més amenaçades per l'automatització, les feines altament qualificades també poden estar en risc.

5. Quina mena de coneixements necessiteu per a prendre decisions sobre la IA generativa? Com us poden ajudar les institucions d'AT?

- Necessitem formació pràctica en l'ús (i disseny) d'eines GenAI per al personal perquè puguem experimentar els efectes de GenAI de primera mà i no només mitjançant informes.
- Almenys coneixements bàsics sobre el funcionament i la tecnologia que hi ha darrere de la IA generativa, per exemple, com s'entrenen els models i produeixen contingut, i coneixements sòlids sobre com s'utilitzen en àmbits concrets de la societat i del treball, inclosa informació sobre el seu impacte potencial.
- Sens dubte, les institucions d'AT hi poden ajudar amb els seus coneixements i experiència a l'hora de donar suport als diputats en l'elaboració de les polítiques. Això pot incloure tallers, sessions informatives, però també demostracions pràctiques d'IA generativa per a confrontar els diputats amb la tecnologia.

Quins són els principals actors?

Les tecnologies d'intel·ligència artificial (IA) de caràcter general, com ara ChatGPT, estan transformant ràpidament la manera com es construeixen i es despleguen els sistemes d'IA. Tot i que s'espera que aquestes tecnologies aportin grans beneficis en els propers anys, estimulant la innovació en molts sectors com la indústria i l'educació, la seva naturalesa disruptiva planteja preguntes entre el públic general, les ONG i els responsables polítics sobre la privadesa i els drets de propietat intel·lectual, la responsabilitat i la rendició de comptes, i les preocupacions sobre el seu potencial per difondre informació falsa i desinforma-

ció. Els legisladors de la UE han d'aconseguir el delicat equilibri entre fomentar el desplegament d'aquestes tecnologies i assegurar alhora les garanties adequades.

Per què és important per a la Unió Europea?

Per a maximitzar els beneficis i mitigar els riscos relacionats amb l'adopció d'IA generativa, la Unió Europea hauria de fer un balanç de les inversions i de les actuacions polítiques actuals i previstes, com ara l'estratègia europea d'IA, el mercat únic digital i els programes Horizon Europe. Aquests esforços es poden aprofitar per a donar suport a la investigació, el desenvolupament i la innovació en IA generativa, alhora que s'aborden les implicacions ètiques, legals i socials d'aquestes tecnologies. Els principals àmbits d'interès haurien de ser els següents:

- Proporcionar una capacitat informàtica suficient per a la R+D europea: invertir en una infraestructura informàtica d'alt rendiment serà essencial per a avançar en el desenvolupament de models de llenguatge extens (LLM) i garantir la competitivitat de la recerca europea en IA. (Per exemple, admetre llengües poc representades: fomentar el desenvolupament de LLM i tecnologies d'IA que parin atenció al panorama lingüístic divers d'Europa, i ajudin a mitigar la desigualtat lingüística i promoure la preservació del patrimoni cultural.)
- Directrius ètiques i marcs legals: establir directrius ètiques i un marc legal europeu comú, robust i preparat per al futur que abordi qüestions relacionades amb la desinformació, la privadesa i la protecció de dades, el biaix, la propietat intel·lectual i la desigualtat lingüística.
- Incentius per al desenvolupament responsable de la IA: promoure el desenvolupament de sistemes d'IA que prioritzin l'equitat, la transparència i la responsabilitat, alhora que minimitzin els impactes negatius sobre la societat.

Per a garantir l'autonomia estratègica i la competitivitat global de la Unió Europea en IA generativa, és crucial invertir en recerca i desenvolupament, fomentar les col·laboracions entre el món acadèmic i la indústria i promoure el creixement de les empreses emergents europees d'IA. El desenvolupament de LLM i tecnologies basats a la UE pot ajudar a reduir la dependència dels actors externs i fomentar un ecosistema competitiu.

Rellevància i debat social i polític

Hi ha un debat en curs sobre l'impacte de la IA generativa en les nostres societats i democràcies?

Les característiques clau identificades en els models d'IA de caràcter general (la seva extensió, opacitat i potencial per a desenvolupar capacitats inesperades més enllà de les previstes pels seus productors) plantegen una multitud d'interrogants. Hi ha estudis que documenten que els models de llenguatge extens (LLM), com ara ChatGPT, presenten riscos ètics i socials. Poden discriminar injustament i perpetuar estereotips i biaixos socials, utilitzar llenguatge tòxic (per exemple, incitar a l'odi o a la violència), presentar un risc per a la informació personal i sensible, proporcionar informació falsa o enganyosa, augmentar l'eficàcia de les campanyes de desinformació i provocar una sèrie de danys en la interacció entre humans i ordinadors (com ara portar els usuaris a sobreestimar les capacitats de la IA i utilitzar-la de manera no segura). Malgrat els intents dels enginyers per mitigar aquests riscos, els LLM, com ara GPT-4, continuen plantejant desafiaments per a la seguretat i els drets fonamentals dels usuaris (per exemple, produir textos subtilment falsos o mostrar una habilitat més gran per a oferir consells il·lícits) i poden generar continguts nocius i delictius.

Atès que els models d'IA de caràcter general s'entrenen mitjançant el raspat, l'anàlisi i el processament de les dades disponibles públicament a la Internet, els experts en privadesa adverteixen que sorgeixen problemes de privadesa amb relació al plagi, la transparència, el consentiment i les bases legals per al tractament de les dades. Aquests models són un repte per als sistemes educatius i per als recursos comuns

com ara els repositoris públics. A més, l'aparició de LLM [planteja](#) moltes qüestions, fins i tot pel que fa a la infracció dels drets de propietat intel·lectual i la distribució de materials amb drets d'autor sense permís. Alguns experts [adverteixen](#) que la creativitat generada per IA podria alterar significativament les indústries creatives (en àrees com el disseny gràfic o la composició musical, per exemple). Demanen incentius per a reforçar la innovació i la comercialització de la creativitat generada per IA, d'una banda, i mesures per a protegir el valor de la creativitat humana, de l'altra. També s'ha [plantejat](#) la qüestió amb relació al règim de responsabilitat que s'ha d'aplicar quan els sistemes d'IA d'ús general causen danys. També s'espera que aquests models tinguin un impacte significatiu en el mercat laboral, fins i tot pel que fa a [tasques laborals](#).

En aquest context, els experts [argumenten](#) que hi ha una gran necessitat de controlar la difusió d'eines d'IA de caràcter general, atès el seu impacte en la societat i l'economia. També [reclamen](#) la supervisió i el seguiment dels LLM mitjançant mecanismes d'avaluació i prova, [subratllen](#) el perill de permetre que aquestes eines quedin en mans d'unes poques empreses i governs, i [destaquen](#) la necessitat d'avaluar les complexes dependències entre les empreses en desenvolupament i les empreses que despleguen eines d'IA d'ús general. Els experts en IA també [reclamen](#) una pausa de sis mesos, almenys, en l'entrenament de sistemes d'IA més potents que GPT-4. (Reunió informativa de l'EPRS)

Hi ha resultats d'enquestes entre la població sobre l'ús o opinions sobre la IA generativa?

Un eurobaròmetre no específic sobre IA generativa.

Al juny del 2023 es va publicar una enquesta més general sobre la «[Dècada digital](#)». Segons aquesta enquesta, la majoria dels europeus creu que una connectivitat avançada (76%) i una ciberseguretat més forta (77%) milloraran significativament l'ús diari de les tecnologies digitals. A més, dos terços dels ciutadans europeus (67%) demanen més educació i formació per a desenvolupar les seves habilitats digitals. Així mateix, dos terços dels ciutadans (66%) consideren que les tecnologies digitals tindran un paper important en la lluita contra el canvi climàtic.

Finalment, més del 80% dels europeus creu que els estats membres haurien de col·laborar més per avançar en l'accés a les tecnologies digitals, i també per estimular la innovació i perquè les empreses siguin competitives a nivell mundial. Quan es tracta d'una transformació digital basada en valors, només la meitat dels europeus considera que els drets i els principis digitals estan ben protegits a Europa. Més d'un terç dels ciutadans (36%) creu que cal fer més i, en diversos aspectes, menys de la meitat considera que l'aplicació dels drets i principis digitals al seu país és satisfactòria. Els ciutadans estan especialment preocupats quan es tracta de garantir entorns i continguts digitals segurs per a infants i joves o d'aconseguir el control de les pròpies dades o del llegat digital.

El 2020 la Direcció General de Comunicacions, Xarxes, Continguts i Tecnologia de la Comissió Europea (Connect) va publicar una [Enquesta a les empreses europees sobre l'ús de tecnologies basades en la intel·ligència artificial](#). Segons l'enquesta, el coneixement de la IA és alta a tota la UE (78%). Quatre de cada deu empreses (42%) han adoptat almenys una tecnologia d'IA, el 25% n'ha adoptat almenys dues. Mentre que el 18% té plans d'adoptar la IA en els propers dos anys, el 40% no l'ha adoptat ni té previst fer-ho. L'adopció a nivell de cada tecnologia encara és relativament baixa: des del 3% per a l'anàlisi de sentiments fins al 13% per a la detecció d'anomalies i l'optimització de processos/equips. L'estratègia d'aprovisionament més comuna és externa, ja que el 59% de les empreses de la UE que utilitzen IA adquireixen programari o sistemes preparats per a utilitzar. Tres barreres internes clau en l'adopció de la IA són les dificultats per a contractar personal nou amb les habilitats adequades (57%), el cost d'adopció (52%) i el cost d'adaptar els processos operatius (49%). Reduir la incertesa pot ser beneficiós, ja que les empreses consideren que la responsabilitat per danys potencials (33%), l'estandardització de dades (33%) i els obstacles reguladors (29%) són els principals reptes externs per a l'adopció de la IA.

Hi ha alguna legislació vigent sobre aquesta qüestió?

La [regulació de la intel·ligència artificial](#) ha d'estar estretament lligada amb una regulació adequada de les dades i de la seva governança, així com amb el concepte de [responsabilitat](#). Després d'aprovar recentment una legislació com la [DSA](#), la [DMA](#), la [Llei de dades](#) i la [Llei de governança de les dades](#), la Unió Europea està enllestint el treball sobre la [Llei d'intel·ligència artificial](#), que establirà els mecanismes gene-

rals per a regular l'aplicació i el desenvolupament de la IA. Avaluarà com s'ha d'aplicar aquesta regulació a la IA generativa i si calen instruments reguladors addicionals. El Parlament Europeu supervisarà l'aplicació del programa de recerca i innovació de la UE Horizon, que facilita i finança la investigació sobre intel·ligència artificial a Europa. Supervisarà el Programa de polítiques per al 2030 [Camí cap a la dècada digital](#), amb l'objectiu que el 75% de les empreses de la UE utilitzin Cloud / AI / Big Data el 2030. La Unió Europea establirà polítiques europees en àrees fortament afectades per la IA generativa: economia creativa, educació, àrees socials i culturals i tecnologies sanitàries.

Hi ha actualment propostes polítiques o legislatives sobre aquests temes?

El projecte de Llei d'IA de la UE és el primer intent d'aprovar una regulació horitzontal per a la IA. El marc legal proposat se centra en la utilització específica dels sistemes d'IA i els riscos associats. La Comissió Europea proposa establir una definició tecnològicament neutra dels sistemes d'IA en la legislació de la UE i establir una classificació per als sistemes d'IA amb diferents requisits i obligacions adaptats a un «enfoque basat en el risc». Els sistemes d'IA que presentin riscos «inacceptables» estarien prohibits. S'autoritzaria una àmplia gamma de sistemes d'IA «d'alt risc», però subjectes a un conjunt de requisits i obligacions per a accedir al mercat de la UE. Els sistemes d'IA que només presenten un «risc limitat» estarien subjectes a obligacions de transparència molt lleus. El Consell va acordar la posició general dels estats membres de la UE al desembre del 2021. El Parlament Europeu va crear una comissió especial ([AIDA](#)) sobre aquesta qüestió, que va presentar un [informe final](#) sobre la intel·ligència artificial en l'era digital a l'abril del 2022. El Parlament va votar la seva posició al juny del 2023.

Els legisladors de la UE estan negociant en aquests moments per finalitzar la nova legislació, amb esmenes substancials en la proposta de la Comissió, que inclouen la revisió de la definició de sistemes d'IA, l'ampliació de la llista de sistemes d'IA prohibits i la imposició d'obligacions a la IA d'ús general i als models d'IA generativa com ara ChatGPT.

El paper de l'AT en els debats

El vostre institut s'ha ocupat d'aquest tema?

L'STOA ha publicat més de 30 [publicacions sobre intel·ligència artificial](#) des del 2016. A més dels estudis estàndard, l'STOA ha publicat un conjunt d'altres publicacions, com 'What if?' [I si?] i resums de publicacions que ofereixen una visió general de la conscienciació sobre les tendències tecnocientífiques actuals i rellevants, així com publicacions de blocs, articles, entrevistes, vídeos i pòdcasts (vegeu [Repositori IA](#)).

L'AT ha tingut un impacte en els debats en curs?

L'STOA ha fet servir cada cop més les xarxes socials i altres canals per a comunicar les seves activitats. Això inclou el compte X (anteriorment Twitter) de STOA (@EP_ScienceTech) i les publicacions al bloc EPRS, que informen sobre notícies, projectes i esdeveniments de manera accessible. L'STOA també continua produint pòdcasts periòdics, sovint basats en les seves publicacions i vídeos breus de sensibilització. Els esdeveniments de STOA es transmeten en directe per Internet i sovint s'acompanyen de publicacions en directe de @EP_ScienceTech, que permeten la interacció simultània amb les parts interessades, els experts i la ciutadania. Durant les seves reunions a les setmanes de plens d'Estrasburg, els membres del panel STOA solen escoltar les [presentacions](#) sobre els resultats dels projectes de STOA en curs i proporcionar comentaris i orientació addicional per a completar els informes. Com que l'STOA va ser una de les primeres entitats del Parlament Europeu a treballar en IA, les seves conclusions han informat i alimentat l'activitat parlamentària dels eurodiputats. Això va ser especialment rellevant per als debats que tenien lloc en diverses comissions, com ara la comissió especial sobre IA (AIDA) i les comissions legislatives que lideren el treball sobre la Llei d'IA (la Comissió de Llibertats Civils, la Comissió del Mercat Interior i la Comissió d'Afers Jurídics).

Quines són les lliçons apreses de l'AT?

Els legisladors de la UE estan immersos actualment en llargues negociacions per a definir un marc normatiu de la UE per a la IA que sotmetria els sistemes d'IA d'«alt risc» a un conjunt de requisits i obligacions en la UE. L'abast exacte d'una proposta de lleï d'intel·ligència artificial (Llei de la UE sobre IA) és motiu de controvèrsia. Tot i que la proposta original de la Comissió Europea no contenia cap disposició específica sobre tecnologies d'IA d'ús general, el Consell de la UE ha proposat que haurien de ser considerades. Mentrestant, els científics han advertit que qualsevol plantejament que classifiqui els sistemes d'IA com d'alt risc o no en funció de la seva finalitat crearia una llacuna per als sistemes d'ús general, ja que la futura llei d'IA regularia els usos específics d'una aplicació d'IA però no els models de base subjacents.

En aquest context, alguns dictàmens científics, com el de l'Institut del Futur de la Vida, han demanat que la IA de caràcter general s'inclouï en l'àmbit d'aplicació de la llei d'IA. Alguns acadèmics favorables a aquest enfocament han suggerit modificar la proposta en aquest sentit. Helberger i Diakopoulos proposen considerar la creació d'una categoria de risc separada per als sistemes d'IA de caràcter general. Estarien subjectes a obligacions i requisits legals que s'ajustessin a les seves característiques, i a un sistema de seguiment del risc sistèmic similar al que estableix la Llei de serveis digitals (DSA). Hacker, Engel i Mauer argumenten que la llei d'IA s'hauria de centrar en aplicacions específiques d'alt risc d'IA de caràcter general i incloure obligacions en matèria de transparència, gestió de riscos i no-discriminació. Les normes de moderació de continguts de la DSA (per exemple, els mecanismes de notificació i actuació, i els indicadors de confiança) s'haurien d'ampliar per a cobrir aquest tipus d'IA d'ús general.

Küspert, Moës i Dunlop reclamen que la regulació d'IA d'ús general estigui preparada per al futur i que, entre altres coses, abordi la complexitat de la cadena de valor, tingui en compte les estratègies de codi obert i adapti el compliment i l'aplicació de les polítiques als diferents models de negoci. Per a Engler i Renda, la llei hauria de desincentivar l'accés de les API per a l'ús d'IA de caràcter general en sistemes d'IA d'alt risc, introduir compromisos poc estrictes per als proveïdors de sistemes d'IA de caràcter general (com ara un codi de conducta voluntari) i aclarir les responsabilitats dels actors al llarg de les cadenes de valor.

Tot i que l'STOA pot esbossar diverses opcions des d'un punt de vista tècnic (per exemple, sobre l'abast de la regulació de la IA, la definició d'IA i la classificació dels riscos), és evident que els diputats també poden tenir en compte altres perspectives o elements en el seu treball legislatiu. No obstant això, l'àmplia activitat de l'STOA, que inclou l'avaluació tecnològica i els estudis de prospectiva científica, juntament amb la comunicació científica i tecnològica, constitueixen un actiu essencial per la seva contribució decisiva a sensibilitzar sobre l'impacte dels nous avenços i promoure l'elaboració de polítiques basades en l'evidència. L'STOA també comparteix els seus coneixements i experiència amb col·legues dels parlaments nacionals per mitjà d'una sessió informativa regular en línia, «Fonts de coneixement sobre IA», que s'ofereix als parlaments nacionals.

França

Informe EPTA 2023 sobre intel·ligència artificial generativa Contribució de l'Oficina Parlamentària d'Avaluació Científica i Tecnològica (OPECST) de França

L'Oficina Parlamentària d'Avaluació Científica i Tecnològica (OPECST) de França va començar a tractar la intel·ligència artificial (IA) a partir del 2016, i, el 14 de març de 2017, va aprovar un informe **titulat «Cap a una intel·ligència artificial controlada, útil i desmitificada»**.

El 2018, França va decidir afrontar aquest repte amb la implementació d'un **pla nacional** i la creació d'una **coordinació intergovernamental** i un **comitè d'ètica nacional específic**.

Entre el 2017 i principi del 2020, els **avenços en xarxes neuronals profundes basades en transformadors** van fer possibles **nous sistemes d'IA generativa**, que utilitzen indicacions de llenguatge natural. En conseqüència, es van començar a desenvolupar models de llenguatge extens entrenats amb grans quantitats de dades no etiquetades. Models com ara els **bots de conversa** (ChatGPT 3 i 4, llançat per OpenAI, Microsoft Bing Chat, Google Bard i PaLM, DeepMind Gato, Meta LLaMA, etc.) o els **sistemes d'art de text a imatge** (Stable Diffusion, Midjourney, DALL-E, etc.) s'han fet molt coneguts entre la població. Una empresa emergent francesa d'IA generativa, Mistral AI, aquest any ha recaptat 105 milions de dòlars durant el mes posterior a la seva creació, la qual cosa és un rècord. A diferència de ChatGPT o Bard, el seu model anomenat Mistral 7B és de codi obert i està destinat a desenvolupadors, que podran utilitzar-lo, millorar-lo i comercialitzar-lo com vulguin, gràcies a una llicència molt permissiva. Però és un llenguatge «petit» basat en 7.000 milions de paràmetres, molt lluny dels centenars de milers de milions de paràmetres de GPT4, d'OpenAI, o PaLM, de Google.

Com a conseqüència d'aquest nou context de sistemes generatius d'IA, la Mesa de l'Assemblea Nacional i la Mesa del Senat han demanat recentment a l'OPECST que **elabori un nou informe** sobre qüestions relacionades amb la IA generativa. Aquesta petició política al més alt nivell parlamentari mostra fins a quin punt la tasca actual de l'EPTA és crucial.

Aquest tema s'ha d'**analitzar i desmitificar a fons**, perquè l'auge de les tecnologies d'IA representa **un canvi important, que transforma profundament les nostres societats i economies**. La IA està subjecta a les limitacions de l'acceptabilitat social, a causa de les **visions alarmistes**. Els perills potencials de la IA han reforçat les pors i angoixes davant el desplegament de tecnologies d'IA.

Els cicles d'esperança i decepció que marquen la història de la IA suggereixen que hem de ser prudents i tenir expectatives realistes d'aquestes tecnologies: de períodes d'opinions polaritzades, **ansietats excessives**, i també **esperances excessives**, ja se n'han vist abans. És imprescindible fer un recorregut per la història. El concepte d'IA fa referència a **les diverses tecnologies que van sorgir a la segona meitat del segle XX**, basades en l'ús **d'algoritmes**. Aquestes tecnologies, amb combinacions en evolució constant, ja **s'implanten en un ampli ventall de sectors** i han donat lloc a oportunitats sense precedents per a revolucionar el nostre entorn vital i millorar les nostres vides, sobretot pel que fa a l'atenció a les persones.

Els desenvolupaments en aquest camp poden ser ràpids i les aplicacions sectorials actuals o futures tenen un abast considerable en diversos àmbits, com ara **educació, ciència, medi ambient, energia, transport, aeronàutica, agricultura, comerç, finances, defensa, seguretat, comunicacions, lleure, salut, dependència, discapacitat** i molts altres. **El debat públic no es pot dur a terme d'una manera racional** sense una bona comprensió de les tecnologies, els mètodes científics i els principis de la IA.

El progrés de la IA planteja preguntes de les quals la societat ha de ser conscient: quines són les **noves oportunitats** i els nous **riscos**? Quina és la posició de França i d'Europa en la cursa mundial que ha començat? Quin lloc han d'ocupar la recerca pública i la recerca privada? Quins **principis ètics, jurídics i polítics** haurien de guiar aquestes tecnologies? La regulació s'hauria de fer a escala nacional, comunitària o internacional?

De fet, amb aquestes tecnologies ens enfrontem a **reptes ètics, jurídics, econòmics, socials i científics**. Alguns dels reptes són els següents:

- El paper dominant de la **recerca privada**, liderada per empreses nord-americanes i potencialment xineses, que accelera la transició cap a una economia globalitzada dominada per **plataformes**, com ara Google, Facebook o Amazon, etc., i que pot **perjudicar d'una manera crítica la governança democràtica**.
- **Problemes amb els algoritmes**, amb riscos relacionats amb resultats de la IA esbiaixats (conseqüència de les dades i els algoritmes esbiaixats), amb el fenomen de les caixes negres d'algoritmes i amb les bombolles d'informació o els filtres bombolla, que també poden **perjudicar la governança democràtica**.
- Finalment, però no menys important, **les transformacions del mercat laboral**: la IA substituirà molts treballadors i també crearà llocs de treball nous, i sovint canviarà el contingut de la nostra feina i la nostra manera de treballar, cosa que podria tenir efectes en termes de desigualtat social.

La IA també es veu com una oportunitat per a les nostres societats i economies; no és **ni una recerca estèril ni un pla per a substituir l'home per la màquina**. Avancem cap a una **intel·ligència humana augmentada** en lloc de **una intel·ligència artificial que competeixi amb els humans**. Els avenços en IA són, en primer lloc, **beneficiosos**. Tanmateix, no es pot negar que **també comporten riscos**. Aquests riscos es poden **anticipar, identificar i mitigar, i cal fer-ho**.

L'adveniment de la **superintel·ligència** no forma part d'aquests riscos a curt i mitjà termini. Encara és incert si aquesta amenaça es farà realitat a llarg termini. Pel que fa a la seva imminència a curt o mitjà termini, profetitzada per diverses figures mediàtiques, és només una **fantasia**.

Per tant, cal **anar més enllà de les aparences** i veure la **realitat científica** que hi ha darrere de les esperances i angoixes que s'han expressat sobre el desenvolupament de la IA.

Aquest treball de desmitificació ha de ser **col·lectiu, interdisciplinari i internacional**. Per a evitar el desencís futur, cal garantir un **control continu d'aquestes tecnologies i els seus usos**.

És per això que l'OPECST, en el seu informe de 2017, va defensar una IA **controlada, útil i desmitificada**.

Controlada, perquè aquestes tecnologies hauran de ser **tan segures, transparents i justes com sigui possible**.

Útil, perquè si bé han de **respectar els valors humanístics**, en última instància, han de **beneficiar els ciutadans**.

Desmitificada, perquè les dificultats de l'acceptabilitat social de la IA són, en gran part, de resultes de **visions alarmistes infundades i incomprensió**.

En lloc d'informar d'un **enfrontament hipotètic entre homes i màquines**, que és més aviat ciència-ficció distòpica, el proper treball de l'OPECST hauria d'aclarir les condicions i directrius per a **la complementarietat futura entre l'home i la màquina**.

Alemanya

Informe EPTA 2023

Intel·ligència artificial generativa - Punt de vista d'Alemanya / TAB

ChatGPT es troba amb reaccions diverses

«Només era qüestió de temps» es va dir als mitjans de comunicació quan, al maig del 2023, un estudiant d'una escola secundària d'Hamburg (Gymnasium) va ser enxampat fent trampes a l'examen final amb l'ajuda de ChatGPT.³³ En realitat, aquest ús de la IA generativa no era inesperat: els experts havien advertit que els estudiants farien que ChatGPT els escrigués les redaccions (Weßels, 2022), els periodistes havien demostrat que es podien obtenir bons resultats amb la darrera versió en diverses assignatures,³⁴ algunes escoles dels EUA ja tenien prohibicions explícites,³⁵ i l'ús d'aquestes ajudes no autoritzades als exàmens era un dels escenaris previstos en l'informe de la TAB sobre ChatGPT per al Bundestag alemany (Albrecht, 2023.76f.). La reacció de l'autoritat escolar, però, va ser força sorprenent: no va endurir les regles ni les sancions contra l'ús de ChatGPT als exàmens, sinó que en va permetre l'ús, sempre que es documentés.³⁶ Es va demanar als professors que adaptessin els exàmens en conseqüència i comprovessin directament els coneixements dels estudiants si desconfiaven de la seva honestetat a l'hora de citar les fonts.

Aquesta reacció és bastant simptomàtica de la postura alemanya sobre la nova generació de sistemes d'IA generativa. D'una banda, la seva introducció fou rebuda amb reserves: ChatGPT, per exemple, va ser qualificat de defectuós (en les seves respostes), de perillós per a la seguretat pública i l'opinió pública, i d'haver de ser tractat amb precaució a causa dels efectes en gran part desconeguts del sistema en diversos àmbits de la societat i també pel destí desconegut de les dades introduïdes en el sistema. D'altra banda, ChatGPT va captivar fins i tot els experts que havien seguit de prop el desenvolupament de la IA generativa. La qualitat dels textos generats en resposta a simples sol·licituds breus va significar un gran salt endavant respecte als sistemes d'IA anteriors. El sistema va demostrar ser capaç de simular de manera convincent tant la comprensió com la producció d'enunciats, i exercia així el paper d'interlocutor. Aviat es va fer evident que la tecnologia subjacent a ChatGPT es podria utilitzar no només per a resoldre diverses tasques basades en text, sinó també per a la generació i la manipulació d'imatges. En poc temps es va llançar tota una família de nous i potents sistemes d'IA, anomenats IA generativa o models de fundació. Simples prohibicions o recomanacions de no utilitzar els sistemes no serien una opció viable. Més aviat es va produir un intens debat en l'opinió pública alemanya sobre com aprofitar al màxim el potencial de la IA generativa en diversos camps d'aplicació sense caure en els seus riscos i inconvenients. I es va debatre sobre quin tipus de guies calia establir per a orientar el desenvolupament tecnològic i les aplicacions del sistema en una direcció beneficiosa per al conjunt de la societat.

Adopció i rellevància econòmica de la IA (generativa) a Alemanya

Tot i que la introducció de ChatGPT va ser seguida arreu del món pel creixement més ràpid de qualsevol aplicació de consum fins al moment³⁷ i el debat públic posterior a Alemanya va créixer enormement (com en altres països), la seva adopció real a Alemanya va ser menys dinàmica. A finals de gener del 2023, se-

33. <https://www.ndr.de/nachrichten/info/Hamburg-Mit-Chat-GPT-im-Abi-gemogelt,ndrinfo46278.html>

34. <https://www.br.de/nachrichten/netzwelt/chatgpt-ki-besteht-bayerisches-abitur-mit-bravour,TfB3QBw>

35. <https://ny.chalkbeat.org/2023/1/3/23537987/nyc-schools-ban-chatgpt-writing-artificial-intelligence>

36. <https://www.abendblatt.de/hamburg/kommunales/article238612213/Nun-doch-Schueler-duerfen-Chat-GPT-bei-Abitur-Pruefung-benutzen.html>

37. <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>

gons una enquesta (CAIS, 2023), aproximadament una quarta part de la població adulta havia sentit parlar de ChatGPT i l'11% va dir que ja havia utilitzat el sistema, amb un 3% que n'indicava un ús habitual. A finals d'abril del 2023, el 83% havia sentit parlar de ChatGPT i el 23% l'havia utilitzat (almenys una vegada), la majoria amb finalitats privades (Bühler, 2023). Pel que fa a l'ús empresarial, una enquesta del juny del 2023 va mostrar que el 13% de totes les empreses alemanyes tenien algun tipus de tecnologia d'IA en ús (no necessàriament IA generativa), amb una major adopció observada a la indústria (17% en ús, 13% pensen fer-la servir, Schaller, *et al.* 2023, pàg. 5).

Al voltant d'un terç de la població d'Alemanya considera positivament l'ús de la intel·ligència artificial en general, amb un 45% que és escèptic i un 10% que s'hi nega (març del 2023, Continental 2023).³⁸ La majoria dels enquestats temen que la IA pugui costar llocs de treball o que es pugui tornar incontrolable.³⁹ La introducció de ChatGPT no ha canviat gaire aquest panorama general: per al 53% dels enquestats res no ha canviat; el 18% ara en té opinions més positives, enfront del 14% les opinions dels quals s'han tornat més negatives. Mentre que la majoria de la gent assumeix que la IA canviarà la seva vida l'any 2030, les opinions estan dividides pel que fa a la direcció del canvi: un terç espera que la vida canviï a millor, un 23% a pitjor i un 30% no ho sap, cosa que posa de manifest una gran quantitat d'incertesa sobre el desenvolupament futur de la IA i el seu ús.⁴⁰

Aquesta postura general de la societat xoca en certa manera amb el fet que Alemanya tingui un panorama de recerca en IA viu i consolidat des de fa temps. Tot i que gairebé no hi ha grans empreses tecnològiques com als EUA (a excepció de SAP), el nombre de publicacions d'institucions de recerca alemanyes sobre IA coincideix amb el dels EUA i Israel i és superior al d'altres països, almenys si es mesura en proporció a la població.⁴¹ Hi ha diversos programes de finançament de la investigació i una estratègia nacional d'IA,⁴² en què es van gastar més de 1.000 milions d'euros en aquest marc entre el 2019 i el final del 2022 (Bertschek, 2023). Però quan es tracta de traduir l'excel·lència de la investigació en històries d'èxit econòmic, Alemanya està clarament endarrerida respecte de les principals nacions. Aquesta afirmació s'aplica al nombre de patents, i encara més pel que fa a les solucions d'IA llestes per a la comercialització, segons el president de Bitkom, una associació que representa les empreses alemanyes d'economia digital.⁴³ No obstant això, amb l'empresa Aleph Alpha, un important actor europeu en el camp dels sistemes d'IA generativa està registrat a Alemanya, comercialitzant activament la seva tecnologia d'IA com a «sobirana» i «europea».⁴⁴

Un debat viu davant de moltes incògnites

Els investigadors d'IA i les empreses d'IA, representades pels seus executius i també per l'associació alemanya d'IA (una xarxa de 400 empreses d'IA), es trobaven entre els actors més destacats en el debat sobre la IA generativa que es va iniciar a finals del 2022. Van advertir que la llei europea d'IA prevista podria regular en excés la IA generativa i, per tant, inhibir el creixement d'una economia europea d'IA. També van destacar les qualitats dels sistemes d'IA desenvolupats en el marc del RGPD i altres marcs europeus, especialment pel que fa a l'aplicació de la IA al sector industrial, centrant-se en la confidencialitat i la res-

38. Es troben resultats similars en una enquesta de l'abril del 2023, amb un 30% que aprova el desenvolupament de tecnologies basades en IA, un 37% que el desaprova i un 30% neutral (Fox *et al.*, 2023). En una enquesta de finals d'abril del 2023, només el 18% dels enquestats va estar d'acord que la IA faria del món un lloc millor, però el 49% veu possibilitats de fer-la servir en la vida privada (Bühler, 2023).

39. A l'enquesta de Fox *et al.* (2023), només un terç dels enquestats temen la pèrdua de llocs de treball, els principals temors són la IA com a amenaça per a la humanitat, la manipulació del debat públic i la vigilància generalitzada.

40. Els resultats de Fox *et al.* (2023) són similars, amb un 14% dels enquestats que espera un canvi positiu «en els propers deu anys», un 39% un canvi negatiu, un 38% una barreja d'ambdós i un 8% que no ho sap.

41. <https://www.plattform-lernende-systeme.de/ki-monitoring.html>

42. https://www.bmbf.de/bmbf/shareddocs/downloads/files/201201_fortschreibung_ki-strategie.pdf?__blob=publicationFile&v=2

43. <https://www.bitkom.org/Presse/Presseinformation/KI-Deutschland-Spitze-Forschung-Nachholbedarf-Wirtschaft>

44. <https://www.aleph-alpha.com/>

ponsabilitat en el tractament de dades sensibles (Aleph Alpha, 2023). Altres grups d'interès també van expressar les seves opinions. Artistes i altres creadors van assenyalar la necessitat de protegir els seus drets, davant el temor que els sistemes d'IA es poguessin utilitzar per a automatitzar el seu treball i privar-los de la seva base econòmica (utilitzant el seu treball com a dades de formació), i provocar també un augment de la desinformació i manipulació (Iniciative Urheberrecht, 2023; DJV, 2023). Els editors temien que els bots de conversa que informaven de notícies de manera autònoma sense pagar drets de llicència poguessin posar encara més en risc el seu model de negoci, a més dels reptes que planteja el mercat digital (Tagesspiegel Background, 2023). D'altra banda, molts educadors es van mostrar oberts als nous avenços i van demanar que s'experimentés amb la nova tecnologia per a determinar les condicions per a un ús productiu i segur en l'àmbit educatiu (per exemple, Salden/Leschke, 2023, Mohr *et al.*, 2023). Alguns esperaven que la IA generativa pogués estimular encara més el debat sobre noves formes d'avaluar el progrés dels estudiants, un debat que ja fa temps que es manté a Alemanya. Tanmateix, els educadors també van advertir que els resultats de la IA generativa no són fiables i que les dades dels usuaris necessiten protecció.

El Parlament alemany va participar activament en aquest debat, intentant reunir les diferents perspectives en un diàleg constructiu. Ja l'any 2018 es va crear una Comissió d'Estudi sobre Intel·ligència Artificial. El seu informe final, publicat l'any 2020, no pretenia ser un full de ruta per a desenvolupaments futurs, sinó més aviat com un memoràndum d'entesa i un pas intermedi en les consideracions polítiques sobre la IA. Els darrers anys es van encarregar a la TAB diversos projectes sobre l'impacte de la IA en, per exemple, l'exèrcit (Grünwald/Kehl, 2020), l'educació,⁴⁵ l'Administració pública (Evers-Wölk *et al.*, 2022) i la producció de continguts multimèdia que semblin fidels a la realitat (hipertrucatges),⁴⁶ per citar-ne només alguns. Amb la Comissió d'Afers Digitals, des del 2014 hi ha una comissió permanent per a tractar les qüestions relacionades amb la digitalització i la infraestructura digital. Van ser aquesta comissió i la Comissió d'Educació, Recerca i Avaluació Tecnològica els que a la primavera de 2023 van celebrar audiències d'experts sobre ChatGPT i IA generativa per mantenir els parlamentaris informats sobre els últims desenvolupaments i explorar possibles necessitats d'acció reguladora. La TAB va rebre l'encàrrec d'estudiar els fonaments tecnològics de ChatGPT i d'identificar les possibles implicacions de les seves aplicacions en diverses àrees de la societat (Albrecht, 2023). Aquestes activitats dirigides per la comissió i els debats ocasionals sobre IA generativa en el Ple del Bundestag són només l'inici de noves activitats, tal com van respondre els diputats en el qüestionari per a aquest informe. En paraules d'un diputat, «l'evolució en l'àmbit de la IA generativa seguirà sent objecte d'un seguiment periòdic i *ad hoc*».

Malgrat aquestes activitats i el consens sobre la gran rellevància social del desenvolupament de la IA generativa, encara no se n'han explorat els impactes potencials. Atesa l'orientació de molts sistemes d'IA generativa sobre el llenguatge i la comunicació (i, per tant, els fonaments de la vida social), probablement afectaran molts àmbits de la societat. Entre els potencials identificats en l'informe de la TAB (Albrecht, 2023) hi ha la generació automàtica de textos o respostes en camps com l'atenció al client, el periodisme, els assumptes legals i fins i tot la literatura i la investigació. Però aquest ús podria tenir el preu d'un augment (potencialment massiu) de textos de qualitat qüestionable, el desplaçament de llocs de treball creatius, la pèrdua de les competències corresponents i, potencialment, fins i tot la pèrdua de confiança en l'autenticitat de les publicacions. Els sistemes d'IA es podrien construir per a resoldre de manera autònoma cada cop més tasques, convertir-se en una interfície còmoda per a l'ús de sistemes informàtics en general, però també confondre's amb persones amb un caràcter i intencions individuals. Com a nota positiva, els sistemes multimodals i multilingües com ara ChatGPT podrien facilitar la comunicació entre llengües i millorar la inclusió de persones amb un deteriorament cognitiu o d'aprenentatge. Tanmateix, el biaix derivat de les dades d'entrenament o el caràcter de caixa negra de l'aprenentatge automàtic també podrien donar lloc a decisions errònies o a discriminació. A més, s'espera que la intel·ligència artificial generativa pugui empitjorar la seguretat informàtica pel fet de potenciar els actors maliciosos. L'enorme quantitat de recursos (energia

45. https://www.tab-beim-bundestag.de/english/projects_application-potentials-and-challenges-of-artificial-intelligence-in-education.php

46. https://www.tab-beim-bundestag.de/english/digital-society-and-economy_deepfakes-legal-and-societal-challenges-as-well-as-innovation-potentials.php

elèctrica, aigua de refrigeració, mà d'obra humana) necessaris per al manteniment de sistemes tan grans va contra els esforços de sostenibilitat.

Què hi ha en joc en la política i en els diferents àmbits polítics?

Les respostes dels diputats alemanys entrevistats per a aquest informe de l'EPTA confirmen l'opinió que hi ha molt en joc quan es tracta d'IA generativa. Esperen «enormes possibilitats», però també veuen riscos en moltes aplicacions potencials dels sistemes d'IA. La IA generativa podria aportar un «canvi revolucionari» en moltes ocupacions, podria «reforçar processos i opinions arrelats» o conduir a «decisiones nocives» en medicina. Tractar amb IA generativa podria «esdevenir una de les habilitats clau per al món laboral del segle XXI». Curiosament, tot i assumir implicacions de gran abast, els diputats no se sumen a l'enrenou generat en el debat alemany al voltant de la IA generativa. Més aviat, consideren amb cautela les oportunitats i els riscos que es deriven de la innovació. Alguns apunten a la llarga història de la IA i del debat polític corresponent, altres a les nombroses qüestions obertes i necessitats de coneixement. «L'impacte en àrees polítiques específiques és difícil de preveure», afirma un diputat pel que fa a l'impacte de la IA generativa.

Pel que fa a les implicacions per al sistema polític i la democràcia en general, tots els diputats veuen possibles efectes negatius d'un ús més estès de la IA generativa. El president federal alemany ja havia destacat alguns dels reptes que planteja per a la democràcia l'auge de la IA generativa, és a dir, que distorsiona la nostra relació amb els fets, que podria soscavar la presa de decisions legitimada democràticament i fomentar la manipulació i les pràctiques malicioses (Oficina del President Federal, 2023). Les respostes dels diputats apunten en una direcció similar, i emfatitzen que la IA generativa podria facilitar la producció massiva de desinformació i de continguts (textuals, visuals i àudio) adreçats a danyar les estructures democràtiques, donant lloc a «més converses ocioses» que siguin «menys fiables». Es pot utilitzar per a manipular els ciutadans, ja que és difícil «determinar si [el contingut produït per IA generativa] reflecteix la realitat o és simplement artificial». Si les organitzacions de mitjans utilitzen sistemes d'IA en lloc de periodistes per a produir o distribuir notícies, la pluralitat d'opinió es podria veure afectada encara més del que ja ho està. Els possibles biaixos dels sistemes d'IA podrien distorsionar el treball dels parlaments. I com que els models s'entrenen estadísticament amb coneixements del passat, confiar en la seva producció podria portar a l'abandonament de les posicions minoritàries o dissidents en els debats polítics. En canvi, «la democràcia s'ha de guiar pel compromís d'objectius contraposats». Tanmateix, la majoria de parlamentaris també veuen un paper potencial per a la IA generativa en la política i en el treball parlamentari, és a dir, podria ajudar a preparar discursos, mocions, preguntes o declaracions, podria facilitar la comunicació amb els ciutadans i a les xarxes socials i podria donar suport a la investigació sobre temes polítics i la traducció de textos.

Les **implicacions en el mercat laboral** es consideren molt contingents i incertes, amb alguns diputats que esperen tant guanys com pèrdues pel que fa al nombre de llocs de treball, altres esperen «canvis revolucionaris» pel que fa a la qualitat del treball. Cinc dels sis diputats entrevistats esperen que hi haurà canvis fonamentals, i tres d'ells consideren important que s'ofereixin opcions de qualificació i formació perquè la mà d'obra en el futur sigui capaç de cobrir les necessitats futures.

En **àrees polítiques com l'educació i la salut**, hi ha algunes diferències interessants en les expectatives dels diputats sobre l'impacte de la IA generativa. Semblen més prudents amb les aplicacions a la medicina, mentre que en l'educació ja s'aplica la IA generativa (com mostra el cas esmentat a la introducció) i el focus se centra més en configurar els processos i les condicions per a un ús responsable de la tecnologia.

En **l'àmbit de l'educació**, hi ha un consens entre els diputats que l'ús d'IA generativa presenta tant oportunitats com riscos. Per als estudiants, podria permetre experiències d'aprenentatge més personalitzades i millorar l'alfabetització digital. També podria ajudar a reduir la desigualtat i millorar l'accés a l'educació dels col·lectius desfavorits. Tanmateix, també hi ha el risc que els estudiants simplement ometin l'aprenentatge o el pensament individual i, en canvi, confiïn en els resultats de la IA. A mesura que els textos produïts per la IA generativa s'assemblen cada cop més als textos humans, és més difícil per als educadors avaluar les competències pròpies dels estudiants. Per a fer front a aquesta nova situació, els professors necessiten una formació específica sobre com utilitzar la IA: com ensenyar amb l'ajuda de la IA, però també com en-

senyar les habilitats necessàries per a utilitzar la IA. També han d'adaptar els seus exàmens per a avaluar aquestes noves competències. Els consells escolars i els directors necessiten suport per a triar les eines adequades (per exemple, sistemes de certificació per a productes d'IA). Les dades personals i els drets dels estudiants, especialment en el cas dels menors, s'han de protegir si s'aplica la IA en contextos educatius.

El paper de la **IA generativa en la sanitat**, en opinió dels diputats, és el de donar suport als metges, sempre que la responsabilitat i el control siguin sempre d'un humà. «Sempre hi ha d'haver la possibilitat d'invalidar una decisió presa per la IA», afirma un diputat. Per tal d'aprofitar els beneficis potencials de l'ús de la IA en l'assistència sanitària en forma de millors diagnòstics i tractaments, cal que hi hagi una regulació sòlida que garanteixi que el tractament de dades personals sensibles sigui acurat i legítim, que les dades d'entrenament per a aplicacions de salut d'IA siguin de màxima qualitat i lliures de prejudicis (gènere, edat, raça, etc.), i que l'ús de la IA sigui transparent i se'n pugui fer la traçabilitat. «Ha de quedar clar al llarg de tot el procés quines decisions/recomanacions són de la IA i quines són d'un humà». S'han d'abordar les qüestions ètiques i minimitzar, mitjançant la regulació, el risc que les companyies d'assegurances o els empresaris facin un mal ús dels coneixements generats per la IA. Igual que en l'educació, l'ús de la IA generativa «s'ha d'integrar en la formació clínica dels professionals mèdics per a preparar-ne un ús més ampli i donar-hi forma d'una manera responsable».

En general, els diputats se centren en qüestions de «regulació reflexiva» de la IA generativa, és a dir, una comunicació transparent de l'ús de la IA a les plataformes de mitjans socials, evitar la discriminació i assegurar la traçabilitat dels resultats, i també en les mesures de protecció ètiques, legals i democràtiques adequades. Com que l'impacte de la IA és difícil de predir, es considera que cal més investigació en aquest àmbit. Aquí és on els diputats veuen el paper de l'avaluació tecnològica: hauria d'informar els parlamentaris sobre el funcionament dels algorismes i les dades utilitzades, i hauria de supervisar els avenços tècnics per a ajudar-los a mantenir-se al dia. Hauria de proporcionar coneixement sobre els potencials i els riscos de la IA generativa i, en general, sobre l'impacte real i potencial en els diferents àmbits d'aplicació. Podria proporcionar aquest coneixement en forma d'escenaris, i aquest coneixement podria servir «com a punts de partida importants per als debats socials». L'avaluació tecnològica podria basar-se en les experiències d'altres països per a conèixer els impactes i l'eficàcia de les mesures reguladores. I es considera «imprescindible» a l'hora de plantejar-se «quines normatives són possibles/necessàries perquè puguem aprofitar les oportunitats i reduir els riscos que es derivaran de l'ús de la IA generativa».

Per a l'avaluació tecnològica (parlamentària), hi ha molt en joc per a satisfer aquestes expectatives de gran abast. Com que l'evolució és molt ràpida, cal posar a prova els procediments establerts d'avaluació tecnològica. Les avaluacions s'han de fer gairebé en temps real a mesura que avança el desenvolupament tecnològic. D'altra banda, el potencial de la IA generativa per a millorar l'eficàcia de diverses tasques de comunicació podria potenciar i millorar el treball de les institucions parlamentàries d'AT, així com el procés polític en conjunt. Tanmateix, si és possible utilitzar la IA generativa d'una manera responsable i sostenible, i com fer-ho, és una qüestió oberta.

Referències

- Albrecht, S. (2023): ChatGPT und andere Computermodelle zur Sprachverarbeitung – Grundlagen, Anwendungspotenziale und mögliche Auswirkungen [ChatGPT i altres models informàtics per al processament del llenguatge: conceptes bàsics, potencial d'aplicació i possibles efectes]. Berlín: Oficina d'Avaluació Tecnològica del Bundestag Alemany (TAB)
- Aleph Alpha (2023): Declaració escrita de Jonas Andrulis, fundador i soci d'Aleph Alpha GmbH. Document de la Comissió 20(23)157 del 24 de maig de 2023, Berlín: Bundestag alemany, Comissió d'Afers Digitals, <https://www.bundestag.de/resource/blob/949934/ad322fe569ad8c40df1688665f9e7c22/Stellungnahme-Andrulis-data.pdf>
- Bertschek, I. (2023): «Jetzt bloß nicht den Anschluss verlieren! – Status quo, Potenziale und Herausforderungen von Künstlicher Intelligenz» [No perdís la connexió ara! *Statu quo*, potencial i reptes de la intel·ligència artificial]. En: Wirtschaftsdienst 103(8), p. 518-520 (DOI: 10.2478/wd-2023-0149)

- Bühler, J. (2023): ChatGPT & Co.: Sicherheit von generativer Künstlicher Intelligenz [ChatGPT & Co.: Seguretat de la intel·ligència artificial generativa]. Associació TÜV, roda de premsa, 11 de maig de 2023. https://www.tuev-verband.de/fileadmin/user_upload/Content_local/Studien_local/TUEV-Verband_PK_Praesentation_ChatGPT_11_05_2023_final.pdf
- Oficina del President Federal (2023): Federal President Frank-Walter Steinmeier at a dinner for members of the Order Pour le mérite for Sciences and the Arts on 11 June 2023 at Schloss Bellevue [El president federal Frank-Walter Steinmeier en un sopar per a membres de l'Orde Pour le Mérite per a les ciències i les arts l'11 de juny de 2023 al Schloss Bellevue]. https://www.bundespraesident.de/Shared-Docs/Downloads/DE/Reden/2023/06/230611-AE-Orden-Pour-le-merite-Englisch.pdf?__blob=publicationFile
- Centre d'Estudis Avançats d'Internet (CAIS) (2023): ChatGPT, how many people already know you? [ChatGPT, quanta gent ja et coneix?]. Nota de premsa, 3 de febrer de 2023. <https://www.cais-research.de/en/news/chatgpt-how-many-people-already-know-you/>
- Continental (2023): Umfrage Künstliche Intelligenz [Enquesta d'intel·ligència artificial]. https://cdn.continental.com/fileadmin/_imported/sites/corporate_international/german/hubpages/10_20pressestudien_und_publicationen/umfrage_kuenstliche_intelligenz_2023/20230301-continental-umfrage-ki.pdf
- Associació de Periodistes Alemanys (DJV) (2023): Positionspapier bezüglich des Einsatzes Künstlicher Intelligenz im Journalismus [Document de posició sobre l'ús de la intel·ligència artificial en el periodisme]. https://www.djv.de/fileadmin/user_upload/INFOS/Themen/Medienpolitik/DJV-Positionspapier_KI_2023-04.pdf
- Evers-Wölk, M.; Kluge, J.; Steiger, S. (2022): Künstliche Intelligenz und Distributed-Ledger-Technologie in der öffentlichen Verwaltung [Intel·ligència artificial i tecnologia de comptabilitat distribuïda a l'Administració pública]. Berlín: Oficina d'Avaluació Tecnològica del Bundestag Alemany (TAB)
- Fox, P.; Privitera, D.; Reuel, A. (2023): So denken die Deutschen über KI [Així és com pensen els alemanys sobre la IA]. Informe KIRA. Berlín: Centre de Riscos i Impactes de la IA
- Grünwald, R.; Kehl, C. (2020): Autonome Waffensysteme – Endbericht zum TA-Projekt [Sistemes d'armes autònoms. Informe final sobre el projecte d'AT]. Berlín: Oficina d'Avaluació Tecnològica del Bundestag Alemany (TAB)
- Initiative Urheberrecht [Iniciativa sobre drets d'autor] (2023): Urheber:innen und Künstler:innen fordern Maßnahmen zum Schutz vor generativer KI in der Europäischen KI-Verordnung [Autors i artistes demanen mesures de protecció davant la IA generativa en el Reglament europeu sobre IA]. 19 d'abril de 2023, https://urheber.info/media/pages/diskurs/ruf-nach-schutz-vor-generativer-ki/03e4ed0ae5-1681902659/finale-fassung_de_urheber-und-kunstler-fordern-schutz-vor-gki_final_19.4.2023_12-50.pdf
- Mohr, G.; Reinmann, G.; Blüthmann, N.; Lübcke, E.; Kreinsen, M. (2023): Übersicht zu ChatGPT im Kontext der Hochschullehre [Visió general de ChatGPT en el context de la docència universitària]. Hamburg: Centre d'Ensenyament i Aprenentatge Universitari d'Hamburg
- Salden, P.; Leschke, J. (2023, eds): Didaktische und rechtliche Perspektiven auf KI-gestütztes Schreiben in der Hochschulbildung [Perspectives didàctiques i jurídiques sobre l'escriptura assistida per IA en l'educació superior]. Bochum: Centre de Didàctica de la Ciència a la Universitat del Ruhr Bochum, DOI: 10.13154/294-9734
- Schaller, D.; Wohlrabe, K.; Wolf, A. (2023): «KI, Cloud Computing und Blockchain – wo steht die deutsche Wirtschaft?» [IA, computació al núvol i cadena de blocs: on és l'economia alemanya?] A: IFO Schnelldienst 76(8), 16 d'agost de 2023, pàg. 3-9
- Tagesspiegel Background (2023): Chatbot-Suche: Alarm in den Verlagen [Cerca de bots de conversa: alarma a les editorials]. Correu electrònic - Butlletí Tagesspiegel Background Digitalisierung & KI, 13.3.2023
- Weßels, D. (2022): ChatGPT – ein Meilenstein der KI-Entwicklung [ChatGPT: una fita en el desenvolupament de la IA]. <https://www.forschung-und-lehre.de/lehre/chatgpt-ein-meilenstein-der-ki-entwicklung-5271>

Grècia

Comissió parlamentària grega sobre avaluació tecnològica Intel·ligència artificial i ocupació

Sens dubte, el desenvolupament de les tecnologies de la informació és una evolució d'una importància cabdal en l'era moderna. Així doncs, parlem d'algoritmes, intel·ligència artificial (en endavant, IA) i quarta revolució industrial. L'erupció de les tecnologies de la informació està provocant turbulències derivades dels canvis en la manera d'organitzar el treball.

I. La IA i l'empresa

Els algoritmes es coneixen des de fa molt de temps. Actualment, s'utilitzen en diverses empreses modernes com a part del desenvolupament de programari de sistemes informàtics que tenen com a objectiu imitar el pensament humà.

1. El desenvolupament de la IA

La IA s'està convertint en una realitat. L'ús d'algoritmes i de la IA s'està desenvolupant en tots els sectors de la vida contemporània. En el sector laboral és a l'arrel de qüestions cabdals.

Pel que fa a l'ocupació i la retallada de llocs de treball, la qüestió és especialment important. Tanmateix, les respostes són divergents. Alguns consideren que és esperable un augment de l'atur. Altres argumenten que la pèrdua de llocs de treball es compensarà amb la creació de nous llocs de treball relacionats amb la nova tecnologia. Tanmateix, limitarem aquest estudi a l'examen de qüestions relacionades amb la IA al servei de les funcions de gestió.

2. La IA al servei de les funcions de gestió

a. La contractació de persones

De resultes de la IA es poden avaluar gairebé a l'instant una sèrie de dades dels currículums dels treballadors o d'altres fonts per a determinar a quins treballadors se'ls contractarà directament o a quins treballadors se'ls farà una entrevista. En una empresa moderna sovint és un algoritme el que, directament o indirectament, decideix qui es contractarà. Tot seguit, també és un algoritme el que proposa la retribució del treballador en funció de les seves capacitats. Portant-ho al límit, podem considerar que el departament de recursos humans és substituït per l'algoritme durant el període de contractació. D'altra banda, sovint l'empresa intenta, mitjançant la recerca en xarxes socials professionals, identificar professionals competents als quals podria presentar una oferta de feina. En aquest cas, és l'algoritme el que examinarà els perfils d'un nombre molt gran de treballadors, susceptibles de satisfer millor les seves necessitats.

Tot i així, la IA no es limita a l'inici de la vida professional. Pot acompanyar el treballador durant tota la seva trajectòria a l'empresa.

b. L'execució del contracte de treball

Sabem que diverses empreses tenen instal·lats des de fa temps sistemes de seguiment dels treballadors i de valoració de la seva activitat. Les dades registrades fan referència a la productivitat, els moviments o les valoracions tant dels seus superiors jeràrquics com dels clients de l'empresa.

En una empresa moderna es pretén que aquestes dades tinguin un paper molt més important. En lloc del control personal, s'estableix un sistema de recollida i anàlisi de dades personals dels treballadors. Les dades analitzades per a fer el seguiment de l'execució del contracte de treball poden provenir tant de l'interior com de l'exterior de l'empresa.

En una empresa tecnològicament avançada no només tot està programat, sinó que tot està meticulosament controlat. A més, les empreses tecnològicament avançades, especialment les plataformes digitals, adopten sistemes d'avaluació derivats dels clients per mitjà d'algoritmes per a afavorir els treballadors o, per contra, per a desfavorir-los.

Posteriorment l'anàlisi d'aquestes dades pot ajudar a fixar el salari del treballador amb l'avaluació de les dades d'oferta i demanda del mercat de treball. D'aquesta manera pot decidir sobre el pagament de les bonificacions de productivitat. També pot decidir sobre la promoció del treballador a partir de diverses dades que descriuen la seva activitat a l'empresa. Finalment, l'algoritme, que acompanya el treballador al llarg de la seva carrera, es pot utilitzar en el context d'un acomiadament.

c. L'acomiadament

L'algoritme es pot utilitzar per motius disciplinaris, sobretot tenint en compte les avaluacions dels clients. Aquest procediment s'utilitza especialment per als treballadors externs. De fet, com ja hem observat, els usuaris de la plataforma estan constantment cridats a avaluar els treballadors.

D'aquesta manera, l'algoritme pot decidir sobre un acomiadament. El problema es fa encara més important quan es té en compte que aquestes valoracions dels clients sovint són estandarditzades, impersonals, prenen una aparença d'objectivitat i sovint el treballador no té ni la possibilitat de respondre-hi. La necessitat de transparència tant pel que fa a les decisions de gestió com als paràmetres utilitzats pel sistema d'avaluació és evident.

II. Protecció dels treballadors

Els algoritmes que permeten delegar tasques, prèviament acomplertes per persones, a sistemes automàtics cada cop més «autònoms» plantegen qüestions legals i ètiques. La delegació de la presa de decisions als algoritmes no implica que aquestes decisions puguin defugir la responsabilitat humana.

1. Els riscos derivats de la IA

De fet, els algorismes no sempre són objectius i imparcials, fins i tot encara que sovint l'«objectivitat» dels sistemes automàtics estigui programada. Els usuaris tendeixen a prendre el resultat generat per una màquina com una veritat indiscutible, en contraposició al judici humà, sempre percebut com a fallible. Tanmateix, els algoritmes inclouen valors introduïts per persones.

a. Les relacions de poder a l'empresa

La IA, en la interacció amb el flux de dades massives, la intel·ligència ambiental, la computació ubíqua i el núvol accentuen el gran canvi actual pel que fa al tractament de dades personals: mai abans no s'havia recollit aquest volum de dades de tantes persones, emmagatzemades en tants llocs, analitzades i processades. Així, l'entorn de la informació es modifica profundament.

Els principis que recull la legislació en matèria de tractament de dades personals poden ser tensionats pel desenvolupament de la IA, gran consumidora d'aquestes dades.

En aquest context, la qüestió de la discriminació adquireix una importància especial.

b. El risc de discriminació

En una empresa que utilitza IA, hi ha un risc important de discriminació contra determinades categories de treballadors. El risc encara augmenta més si es té en compte que aquestes discriminacions no són evidents. Un procediment de presa de decisions basat en algoritmes pot estar basat en biaixos socials. D'altra banda, determinats prejudicis es poden reproduir mitjançant avaluacions dels clients, que posteriorment seran processades per un algoritme que condueixi a la discriminació. Les avaluacions dels clients de la plataforma poden comportar, sense necessàriament estar vinculades a mala fe, discriminacions inconscients o no reconegudes que poden comportar posteriorment l'extinció del contracte de treball. En general, és probable que l'expansió dels algoritmes i de la IA creï o reforci la discriminació, cosa que pot generar preocupacions. Per tant, és fàcil entendre que la gestió d'aquestes qüestions no es pot deixar a la discreció

de l'algoritme i dels seus responsables. Per tant, és inevitable cercar formes tradicionals i «modernes» de protegir els treballadors.

2. Les vies de protecció

Les formes de protecció dels treballadors davant l'ús indegut de la IA a l'empresa poden derivar tant de normes de caràcter general com de normes més específiques de les relacions laborals.

a. Les disposicions generals

Les normes generals relatives a la protecció de dades personals són susceptibles d'aplicar-se a l'ús de l'algoritme. De fet, l'algoritme i la IA no poden funcionar sense dades personals. Aleshores, el que sovint és cabdal no és tant l'algoritme com a tal, sinó les dades que processa. Garantir el dret a l'autodeterminació informativa és, doncs, la primera tasca.

Això significa, en primer lloc, garantir que les persones estiguin sempre degudament informades. Són molt útils els principis reconeguts per la legislació sobre tractament de dades personals, com ara els de legalitat i transparència. L'ús de les dades personals necessàries per al funcionament dels algoritmes només es pot regir per la llei relativa a la protecció de la privadesa, encara que no s'orienti directament al tractament algorítmic. El tractament de les dades ha de ser adequat i rellevant amb relació a les finalitats per a les quals s'hagin recollit les dades. Les dades també han de ser exactes i completes. També s'ha de garantir el dret de rectificació. Finalment, els sistemes de tractament de dades han de ser transparents.

L'article 22 del Reglament general de protecció de dades també estableix que cap decisió, i per tant cap decisió de l'empresari, es pot basar exclusivament en el tractament automatitzat de dades personals. És cert que els algoritmes i la IA tenen l'objectiu comú d'executar automàticament tasques o operacions humanes, és a dir, delegar-les en sistemes automàtics. Tanmateix, cal garantir que la recomanació que proporcionen els algoritmes només es faci servir com a ajuda per a la presa de decisions i l'acció humana.

Així doncs, una atenció més exhaustiva als procediments interns de l'empresa adquireix una importància especial pel que fa a la protecció dels drets dels treballadors i, subsegüentment, de les seves dades personals.

b. Les normes i els procediments propis de l'empresa

A l'hora de protegir els drets dels treballadors davant els riscos derivats de l'ús de la IA, s'han de tenir en compte les particularitats de l'empresa per a donar respostes satisfactòries.

Els sistemes de gestió de personal que utilitzen algoritmes han de ser, en primer lloc, transparents i comprensibles. Tots els paràmetres de funcionament de l'algoritme i la persona que n'ha fet la tria han de ser clars i comprensibles. L'article 9 de la Llei grega 4961/2022 estableix que tota empresa, si utilitza un sistema d'IA que afecti qualsevol procés de presa de decisions sobre treballadors o candidats a treballadors que tingui un impacte en les condicions de treball, la selecció, la contractació o l'avaluació, ha de proporcionar, abans del primer ús, informació suficient i clara a cada treballador que inclogui, com a mínim, els paràmetres en què es basa la decisió, i ha de garantir el compliment del principi d'igualtat de tracte i no-discriminació a la feina.

D'altra banda, les institucions existents, siguin de l'àmbit nacional i sectorial o a escala empresarial, poden contribuir posteriorment al compliment de la normativa de tractament de dades personals. Aquí és on el paper dels sindicats pot ser molt útil. L'empresa ha d'estar preparada per a explicar als representants dels treballadors els principis que informen les seves decisions, encara que es basin en l'algoritme. Els treballadors, per mitjà dels seus representants, han de saber quines decisions, referents a la seva feina, es basen en les seves dades personals.

En general, l'empresari i tots els usuaris dels sistemes algorísmics són responsables del seu funcionament i dels resultats de l'avaluació emesos. No podran invocar l'objectivitat de l'algoritme per a evitar el control i la informació. Com que darrere de l'algoritme sempre hi ha una persona que el desenvolupa d'una manera determinada, a partir d'uns criteris determinats i a partir de dades concretes facilitades per l'empresari, aquest darrer ha d'explicar com funciona el seu algoritme.

En qualsevol cas, la decisió de l'empresari sobre els seus treballadors o col·laboradors no es pot basar exclusivament en el tractament automatitzat de dades personals, per la qual cosa s'ha d'implicar els òrgans interns de l'empresa.

Conclusió

Avui dia la IA és una realitat en funcionament dins de l'empresa. No ens podem oposar a aquest desenvolupament. Tanmateix, les conseqüències de les noves tecnologies no es determinen per endavant. S'han d'examinar i controlar.

Japó

Contribució de la Biblioteca de la Dieta Nacional a l'informe de l'EPTA del 2023

Oficina de Recerca i Referència Legislativa (ORRL), Biblioteca de la Dieta Nacional (BDN), Shinya Nakamura i Hiroko Azuma

Intel·ligència artificial generativa. Oportunitats, riscos i reptes polítics

Descripció del sistema/problema en joc

Amb la ràpida difusió de la IA generativa, hi ha el debat de si és acceptable utilitzar-la en l'educació japonesa. S'ha observat que utilitzar la IA generativa en l'educació té el risc que es redueixi la creativitat dels estudiants, s'infringeixin els drets d'autor i es filtri informació. D'altra banda, si s'utilitza adequadament i tenint en compte aquests riscos, la IA generativa podria ser una eina útil per a donar suport a l'aprenentatge i contribuir a millorar l'eficiència de la feina dels professors.

Aquest article presenta algunes iniciatives del Govern i les universitats amb l'objectiu d'establir regles per a l'ús de la IA generativa en l'educació al Japó. Com a qüestió relacionada, parlarem breument de la IA generativa i els drets d'autor al Japó.

Rellevància i debat socials i polítics

(1) Situació i debat en l'educació primària i secundària

Des de l'abril del 2023, el Ministeri d'Educació, Cultura, Esports, Ciència i Tecnologia (MEXT) estudia aquestes qüestions a fi d'establir directrius per a l'ús de la IA generativa en l'àmbit educatiu, a partir d'entrevistes amb professors en actiu, entre d'altres, i també dels debats d'un grup d'experts format per professors universitaris, alcaldes, professors i superintendents d'educació.

El 26 de maig, el Consell de l'Estratègia de la IA, un grup d'experts que analitza la política en matèria d'IA del Govern, va fer un resum provisional de qüestions relacionades amb la IA, que incloïa la gestió de la IA generativa en l'àmbit educatiu. El Consell va expressar la seva preocupació pel fet que si els estudiants utilitzen la IA generativa per a fer els deures serà més difícil fer-ne una avaluació adequada i pel fet que si la utilitzen per a les redaccions i els treballs en disminuirà la creativitat. D'altra banda, va destacar que l'ús de la IA generativa podria ser avantatjós, atès que permetria ajustar els mètodes d'ensenyament al nivell de comprensió de cada estudiant i comprovar en detall l'eficàcia de l'aprenentatge per mitjà d'exàmens senzills, amb la qual cosa milloraria l'eficàcia educativa i es reduiria la càrrega de feina dels professors.⁴⁷

El 4 de juliol, com a resultat dels estudis esmentats, el MEXT va publicar algunes directrius per a escoles de primària i secundària sobre l'ús de la IA generativa en l'educació (taula 1). Aquestes directrius es recopilen provisionalment com a materials de referència perquè el personal de les escoles avaluï l'adequació de l'ús de la IA generativa i no pretenen imposar una prohibició o un mandat uniforme sobre el seu ús.⁴⁸

47. Consell de l'Estratègia de la IA, «Tentative Discussion Paper on AI» [Document d'anàlisi provisional sobre la IA], 26 de maig de 2023, p. 12. Lloc web de l'Oficina del Govern (en japonès) <<https://www8.cao.go.jp/cstp/ai/ai-senryaku/2kai/ronten.pdf>>

48. Tentative Guidelines for the Use of Generative AI at the Primary and Secondary Education Level [Directrius provisionals per a l'ús de la IA generativa en l'educació primària i secundària], 4 de juliol de 2023. Lloc web del Ministeri d'Educació, Cultura, Esports, Ciència i Tecnologia (MEXT) (en japonès) <https://www.mext.go.jp/content/20230704-mxt_shuukyo02-000003278_003.pdf>

Taula 1: continguts principals de les directrius presentades pel MEXT per a centres de primària i secundària sobre l'ús de la IA generativa

Concepte bàsic	
• En l'actualitat, és adequat començar amb un ús limitat i de prova en algunes escoles mentre se'n verifica l'efectivitat de l'ús en determinades situacions (ús amb la finalitat de determinar si la IA és eficaç en les activitats educatives i l'avaluació de l'aprenentatge, subjecte al compliment de les restriccions d'edat i a altres condicions d'ús). • Totes les escoles han de millorar l'educació per a desenvolupar la capacitat d'utilitzar la informació, incloent-hi l'hàbit de verificar l'autenticitat de la informació, i per a millorar les qualitats i habilitats que es necessiten en l'era de la IA. • S'han de fer esforços per a promoure la formació dels professors i l'ús adequat de la IA a l'escola, a fi de millorar l'alfabetització en IA i els estils de treball dels professors.	
Alguns exemples d'ús inadequat	Alguns exemples d'ús adequat
• Permetre que els estudiants utilitzin la IA lliurement en una etapa en què no han desenvolupat completament la seva capacitat d'utilitzar la informació. • Presentar treballs generats amb IA a concursos, redaccions, etc., com si fossin productes propis dels estudiants. • Facilitar l'ús de la IA, des del principi, en situacions com la creació de poemes i haikus (poesia breu) i l'expressió i valoració de la música, l'art, etc. • Facilitar l'ús de la IA als estudiants abans de fer servir els llibres de text o altres materials didàctics de qualitat garantida. • Utilitzar els resultats de la IA només quan els professors avaluen els aprenentatges dels estudiants. Etc.	• Utilitzar les respostes que genera la IA, inclosos els errors, com a materials didàctics per a fer que els estudiants en coneguin la naturalesa i les limitacions. • Utilitzar la IA per a trobar els punts de vista que falten i aprofundir en el debat quan es resumeixen pensaments i idees en un grup. • Fer que els estudiants utilitzin la IA per a l'expressió oral en anglès i que els estudiants estrangers la utilitzin per a l'aprenentatge del japonès. • Fer que els alumnes utilitzin la IA en el procés de revisió de les frases que creen, amb l'objectiu d'aprendre a utilitzar-la. • Utilitzar la IA per a la programació avançada. Etc.

Alguns experts han manifestat que, pel que fa a l'ús de la IA generativa en l'educació primària i secundària, cal garantir que els professors tinguin oportunitats per a adquirir coneixements sobre la IA generativa i compartir exemples concrets del seu ús en cada etapa del desenvolupament dels estudiants.

(2) Situació i debat en l'educació superior

Algunes enquestes fetes a estudiants universitaris suggereixen que la IA generativa s'utilitza per a preparar treballs i altres entregues, adquirir habilitats de programació, preparar cartes de presentació a l'hora de buscar feina⁴⁹ i simular entrevistes de feina.

Des del març del 2023, moltes universitats han enviat comunicats als estudiants sobre l'ús de la IA generativa. Tot i que el contingut d'aquests comunicats varia, generalment es poden dividir en dues categories: els que demanen que se'n faci un ús prudent i els que fomenten que se'n faci un ús adequat i es tinguin en compte els riscos.

Com a exemple de la primera categoria, la Universitat Sophia (una gran universitat privada de Tòquio) ha emès una directriu per als seus estudiants en què indica que no poden utilitzar les frases, els resultats de

49. Al Japó, és molt comú que, cada any, les empreses contractin grups d'estudiants que estan a punt de graduar-se. Mentre encara són a la universitat, els estudiants envien un currículum amb una carta de presentació a l'empresa en què volen treballar, hi fan una entrevista i reben una oferta de feina.

càlculs, etc., creats amb IA generativa en els seus treballs, tesis, etc., sense el permís del professorat, atès que no els han elaborat ells mateixos.⁵⁰

La Universitat de Nagasaki (una universitat nacional situada a la regió de Kyushu) també ha distribuït directrius que prohibeixen als estudiants apropiar-se del contingut obtingut de la IA generativa per als seus treballs, tesis, etc., i que recullen algunes preocupacions sobre els problemes que l'ús de la IA generativa pot causar en l'aprenentatge dels estudiants i l'avaluació del seu rendiment, i també sobre riscos com la filtració d'informació personal i la infracció dels drets d'autor.⁵¹

Un exemple d'això és la Universitat de Tòquio, que, després d'advertir sobre la credibilitat de la informació i el risc de filtració d'informació confidencial i personal, afirma que la simple prohibició de l'ús de la IA generativa no solucionarà el problema i que els seus estudiants i el professorat haurien de trobar, d'una manera proactiva, mètodes per a un ús adequat, noves tecnologies, nous sistemes jurídics i nous sistemes socials i econòmics.⁵²

La Universitat de Waseda també adverteix que l'enviament de treballs i altres documents creats amb IA generativa es tractaria com si s'hagués copiat, però afirma que cal familiaritzar-se amb els usos adequats de la IA generativa.⁵³

A més de les directrius per als estudiants, algunes universitats han elaborat directrius per al professorat. Per exemple, les directrius de la Universitat de Tohoku per al professorat i el personal, basades en la premissa que no és realista eliminar completament l'ús de la IA generativa per part dels estudiants, indiquen les mesures que els professors poden prendre, tant a classe com en les entregues. Les directrius donen exemples de mesures, com ara que els estudiants facin els treballs a l'aula i els presentin oralment abans de qualificar-los.⁵⁴

El juliol del 2023, a partir de les actuacions independents de moltes universitats, el MEXT va enviar un avís en què instava les universitats a prendre la iniciativa per a elaborar unes directrius per a gestionar la IA generativa d'acord amb les seves realitats educatives i a fer-les arribar als seus estudiants i professors.⁵⁵

Un expert en tecnologia educativa posa de manifest que en la formació universitària del futur els professors hauran d'esforçar-se més per a dissenyar les tasques. No només hauran de centrar-se en els resultats dels treballs i exàmens, sinó també en el procés d'aprenentatge, amb un enfocament que canviï de simplement «el que s'ha après» a «com s'ha après».

(3) Debats relacionats amb els drets d'autor

Tal com apunten les directrius esmentades anteriorment, l'ús de la IA generativa amb relació a la Llei de drets d'autor és controvertit. A continuació, es presenta una breu explicació de les disposicions de la Llei de drets d'autor del Japó i l'impacte que la Llei té en les activitats creatives.

50. «Regarding ChatGPT and other AI chatbots (Generative AI)» [Sobre ChatGPT i altres bots de conversa d'IA (IA generativa)], 27 de març de 2023. Lloc web de la Universitat Sophia <<https://piloti.sophia.ac.jp/assets/uploads/2023/03/27162218/83431423966a7acb8457a729683dd1d8.pdf>>

51. «The Use of ChatGPT and Other Generative AI» [L'ús de ChatGPT i altres IA generatives], 25 d'abril del 2023. Web de la Universitat de Nagasaki (en japonès) <<https://www.nagasaki-u.ac.jp/ja/news/news3948.html>>

52. «About Generative AI (ChatGPT, BingAI, Bard, Midjourney, Stable Diffusion, etc.)» [Sobre la IA generativa (ChatGPT, BingAI, Bard, Midjourney, Stable Diffusion, etc.)], 3 d'abril de 2023. Lloc web d'Utelecon de la Universitat de Tòquio (en japonès) <<https://utelecon.adm.u-tokyo.ac.jp/docs/20230403-generative-ai>>

53. «About the Use of Generative Artificial Intelligence (ChatGPT, etc.)» [Sobre l'ús de la intel·ligència artificial generativa (ChatGPT, etc.)], 21 d'abril de 2023. Lloc web de la Universitat de Waseda <<https://www.waseda.jp/top/en/news/77786>>

54. «Notes on the use of ChatGPT and other generative AI (for teachers)» [Notes sobre l'ús de ChatGPT i altres IA generatives (per a professors)], 31 de març de 2023. Lloc web de la guia de classes en línia de la Universitat de Tohoku (en japonès) <<https://olg.cds.tohoku.ac.jp/forstaff/ai-tools>>

55. «Handling of the Teaching and Learning Aspects of Generative AI in Universities and Colleges of Technology (Notification)» [Gestió dels aspectes d'ensenyament i aprenentatge de la IA generativa a les universitats i escoles tecnològiques (notificació)], 13 de juliol de 2023. Lloc web de MEXT (en japonès) <https://www.mext.go.jp/content/20230714-mxt_senmon01-000030762_1.pdf>

(a) Explicació de l'Agència d'Afers Culturals

En resposta a la ràpida propagació de la IA generativa, l'Agència d'Afers Culturals (una oficina externa del MEXT) va organitzar un seminari obert al públic en general el juny del 2023. En el seminari, s'hi va explicar la relació entre la IA generativa i els drets d'autor pel que fa a la fase de desenvolupament i la fase d'ús de la IA (taula 2).⁵⁶

Taula 2: relació entre la IA generativa i els drets d'autor, d'acord amb el resum de l'Agència d'Afers Culturals

	Fases de desenvolupament i aprenentatge de la IA	Fases de generació de contingut i fases d'utilització de contingut generat amb IA
Acció	- Recollir i duplicar obres amb drets d'autor com a dades d'entrenament i crear conjunts de dades d'entrenament. - Utilitzar conjunts de dades per a entrenar i desenvolupar la IA.	- Generar imatges amb IA. - Carregar i publicar les imatges generades. - Vendre reproduccions de les imatges generades (per exemple, col·leccions d'il·lustracions).
Interpretació de la Llei de drets d'autor	- En principi, l'explotació sense la finalitat de gaudir dels pensaments o sentiments expressats en una obra, com ara l'anàlisi de dades per al desenvolupament de la IA, es pot fer sense l'autorització del propietari dels drets d'autor (article 30-4 de la Llei de drets d'autor). - La limitació de drets d'autor anterior no s'aplica a «casos que perjudiquin d'una manera desproporcionada els interessos del propietari dels drets d'autor», com en la reproducció d'una base de dades comercial amb finalitats d'aprenentatge de la IA.	- La infracció dels drets d'autor en els casos en què s'utilitza la IA per a generar imatges, etc., s'ha de jutjar de la mateixa manera que en els casos normals, com quan una persona dibuixa una imatge sense utilitzar la IA. - La càrrega o la venda d'obres generades amb IA que, sense l'autorització del propietari dels drets d'autor, es consideri que són similars a obres existents, o que en depenen, constitueix una infracció dels drets d'autor i està subjecta a reclamacions per danys, a ordres judicials i a sancions penals.

(b) Antecedents de l'esmena de la Llei de drets d'autor

El Japó, històricament, ha fomentat una legislació adequada per al desenvolupament de la tecnologia de la informació. D'acord amb la modificació del 2018 de la Llei de drets d'autor, en principi, ja no cal obtenir l'autorització del propietari dels drets d'autor si s'utilitzen obres amb drets d'autor en les fases d'aprenentatge de la IA.

En el moment de la modificació d'aquesta llei, el Consell d'Afers Culturals del MEXT va debatre sobre les tecnologies com la IA generativa, però es diu que el risc d'infracció dels drets d'autor no es va comunicar totalment als titulars dels drets. Alguns propietaris de drets van argumentar que era injust que fins i tot els serveis comercials estiguessin subjectes a la limitació dels drets d'autor, però això no es va incloure a l'esmena.

(c) Debat

Hi ha disposicions similars a la UE i al Regne Unit, i també al Japó, per a permetre l'anàlisi de dades per al desenvolupament de la IA sense l'autorització del propietari dels drets d'autor. Tot i que la directiva sobre drets d'autor de la UE exigeix que la finalitat de l'anàlisi de dades sigui la recerca acadèmica i la Llei de drets d'autor del Regne Unit exigeix que les finalitats no siguin comercials, la Llei de drets d'autor japonesa permet l'anàlisi de dades per a una àmplia gamma de finalitats, incloses les comercials. A més, la Llei

56. Divisió de Drets d'Autor, Agència d'Afers Culturals, «AI and Copyright» [IA i drets d'autor], juny del 2023. Lloc web de l'Agència d'Afers Culturals (en japonès) <https://www.bunka.go.jp/seisaku/chosakuken/pdf/93903601_01.pdf>

de drets d'autor japonesa no permet als propietaris de drets d'autor rebutjar la recollida i l'anàlisi d'obres amb drets d'autor (opció d'exclusió). Per aquests motius, alguns creadors japonesos han expressat la seva preocupació.

En el Programa estratègic de propietat intel·lectual 2023, del 9 de juny de 2023, el Govern va revisar el rumb anterior de la seva política, que s'havia centrat en la promoció de la IA. El Govern no ha indicat clarament quins casos entrarien en la categoria de «casos desproporcionadament perjudicials per als interessos dels propietaris dels drets d'autor», però aclarirà la qüestió segons els casos concrets, tenint en compte la protecció dels drets dels creadors.⁵⁷

3. Paper de l'AT en els debats

L'Oficina de Recerca i Referència Legislativa (ORRL) de la Biblioteca Nacional de la Dieta del Japó (BDN) va publicar un informe de recerca titulat «Perspectives on Artificial Intelligence / Robotics and Work/Employment» [Perspectives sobre intel·ligència artificial / robòtica i treball/ocupació],⁵⁸ amb la intenció de donar suport a les deliberacions de la Dieta. Aquest informe de recerca es va encarregar a la Universitat de Tòquio, i un equip de recerca format per vint-i-tres experts d'especialitats i afiliacions variades va investigar el nivell d'adopció de la IA i la robòtica, l'estat de la recerca i els problemes que han sorgit, amb l'objectiu de contribuir al debat sobre l'ocupació i el treball en una futura societat de la IA i la robòtica. Aquest informe de recerca es va fer arribar directament als membres de la Dieta, i també està publicat al seu lloc web.

En el moment de la investigació, sovint es deia que el Japó era optimista quant a la introducció de la IA i els robots al lloc de treball. El Japó esperava que la IA i els robots resolguessin el problema de l'escassetat de mà d'obra causada per la disminució de la natalitat i l'envelliment de la població. En canvi, altres països estaven més preocupats per la possibilitat que les màquines assumissin les feines que actualment fan els humans.

Gairebé al mateix temps, a fi de donar unes directrius ètiques a la societat i al Govern pel que fa a l'ús de la IA, al març del 2019 el Govern japonès va publicar «Social Principles of Human-Centric AI» [Principis socials de la IA centrada en l'ésser humà].⁵⁹ El president del consell d'experts que es va encarregar d'aquesta publicació recorda que en aquell moment la IA no es trobava en una fase en què pogués amenaçar la creativitat humana o suposar una amenaça per a la societat en general, tot i que podia oferir un rendiment superior en un nombre limitat d'aplicacions.

La situació ha canviat dràsticament des de l'any passat amb el llançament de ChatGPT i altres tipus d'IA generativa. Hi ha inquietud pel fet que la IA generativa pugui amenaçar els llocs de treball en oficines, inclòs el professorat universitari, que requereixen coneixements i intel·ligència integrats, i per l'amenaça cada cop més gran que les notícies falses generades amb la IA generativa suposen per a la democràcia. Calen noves mesures i debats per a equilibrar els riscos, al mateix temps que s'aprofita el potencial de la IA generativa. Com a base per a aquests debats, també és important educar la població en l'ús de la IA generativa i desenvolupar recursos humans per a crear-los. L'ORRL també farà la recerca i proporcionarà la informació que siguin necessàries per a donar suport a les deliberacions de la Dieta a fi d'abordar les situacions noves que provoquin els avenços en la tecnologia de la IA generativa.

57. Seu de l'Estratègia de Propietat Intel·lectual, Programa estratègic de propietat intel·lectual 2022, 9 de juny de 2023 (en japonès) <https://www.kantei.go.jp/jp/singi/titeki2/kettei/chizaikaku_kouteihyo2023.pdf>

58. «Perspectives on Artificial Intelligence / Robotics and Work/Employment» [Perspectives sobre intel·ligència artificial / robòtica i treball/ocupació], materials de recerca, 2017-5, Tòquio: Oficina de Recerca i Referència Legislativa, Biblioteca Nacional de la Dieta, 2018. (en japonès) <<https://dl.ndl.go.jp/pid/11065181>> La versió en anglès traduïda per l'autor està disponible a <<http://hdl.handle.net/10367/11072>>.

59. «Social Principles of Human-Centric AI» [Principis socials de la IA centrada en l'ésser humà], del Consell de Promoció de l'Estratègia d'Innovació Integrada <<https://www.cas.go.jp/jp/seisaku/jinkouchinou/pdf/humancentricai.pdf>>

Lituània

IA generativa i treball

Una de les qüestions àmpliament debatudes entre els investigadors és si la IA generativa comportarà un atur massiu perquè substituirà els llocs de treball de la majoria de la gent i quins seran a més els seus efectes en termes de desigualtat social. Es pot preveure una gran substitució de treballadors qualificats per algorismes d'IA? Podem esperar una creació equivalent de nous llocs de treball? La IA generativa canviarà molt probablement la naturalesa del nostre treball i la manera com treballem. De quines eines disposem per a abordar aquesta qüestió?

Descripció del sistema/problema en qüestió

De què es tracta?

La IA generativa és un subtipus d'IA que és responsable de crear contingut nou i original. El terme «generatiu» es refereix al fet que aquestes eines poden identificar patrons en conjunts de dades massives i generar contingut nou, una habilitat que sovint s'ha considerat exclusivament humana.⁶⁰

La IA generativa és una tecnologia amb un ràpid desenvolupament que utilitza models entrenats per a generar continguts originals en una varietat de formes, des del text escrit fins al vídeo, la música i, fins i tot, codi de programari. La IA generativa té un gran potencial per als negocis, especialment en el màrqueting, les vendes i els serveis. Models avançats com ara el ChatGPT, el DALL-E i el Midjourney impulsen les possibilitats de la IA generativa. Això és el que fan:

El ChatGPT permet als bots de conversa i als assistents virtuals proporcionar textos semblants als humans a fi de personalitzar experiències i automatitzar converses.

El DALL-E genera imatges realistes a partir de descripcions textuais, cosa que és útil per a indústries com el comerç electrònic, el marxandatge, la publicitat i el disseny.

El Midjourney facilita la generació i manipulació de dades per a augmentar-ne el nombre i crear continguts.⁶¹

Utilitzant models de llenguatge i conjunts de dades molt fornits, la tecnologia pot crear instantàniament contingut únic que pràcticament no es pot distingir del treball humà i que en molts casos és més precís i atractiu. A la pràctica, la IA està representada principalment per l'aprenentatge automàtic, i concretament per les xarxes neuronals. Simulen la lògica utilitzant estadístiques i matemàtiques basades en quantitats massives de dades. El recent augment de l'interès per la IA es deu al gran salt en el desenvolupament de la IA generativa i les seves capacitats, i sobretot a la seva accessibilitat per part d'un públic ampli que ja ha esdevingut part integrant de la vida pública.

Com amb qualsevol nova tecnologia, amb la IA generativa també hi ha hagut molta especulació, incertesa i preocupació perquè una eina així pugui agreujar la desigualtat social o crear noves formes d'estratificació social eliminant feines que normalment fan els humans. De fet, juntament amb els indubtables beneficis de les tecnologies d'IA per al progrés social i econòmic de la humanitat, també és necessari avaluar els riscos i els desavantatges de la IA generativa i com aquesta tecnologia pot afectar la vida de les persones i el conjunt de la societat en el futur.

60. Kweilin Ellingrud; Saurabh Sanghvi; G the urn Singh Dandona, Anu Madgavkar, Michael Chui, Olivia White, Paige Hasebe. IA GENERATIVA I EL FUTUR DEL TREBALL A AMÈRICA, 26 de juliol de 2023 | Informe <https://www.mckinsey.com/mgi/our-research/generative-ai-and-the-future-of-work-in-america?stcr=D2B0E2A3B-2674FEDB1394DB1C8BBE552&cid=other-eml-alt-mip-mck&hlid=90da8f0dc9cd4d5f822b2888cbc44006&hctky=12636281&hdpid=8f9b6ed9-12c0-4de5-b2cd-85de357d05f4>

61. Set coses que cal saber sobre la IA generativa: una guia pràctica per a líders empresarials, <https://www.avaus.com/blog/7-things-you-need-to-know-about-guia-practica-generativa-ai-a-per-dirigents-empresaris/>

Quin és l'estat actual del vostre país (desenvolupament, implantació, ús...)?

A mesura que els avenços tecnològics canvien el panorama del treball modern, la IA generativa s'ha convertit en una eina transformadora amb un potencial inigualable. No hi ha dubte que en un futur proper els camps de la IA i de la IA generativa ocuparan una part cada cop més gran de l'economia lituana, per la qual cosa és important tenir en compte que cap professió no és immune a l'automatització o la transformació dels llocs de treball, ja que aquestes tecnologies aniran millorant constantment. Per tant, cal suposar que la majoria de professions es mantindran, però els requisits de competències per als empleats canviaran. Disposar de les eines que facin realitat el potencial de la IA generativa és una circumstància que aviat començarà a passar gradualment de la categoria de «serà un avantatge» a la categoria de «competència requerida». De la mateixa manera que moltes feines avui dia requereixen unes bones competències informàtiques i una familiaritat amb almenys mètodes de treball bàsics en paquets ofimàtics generals, els empleats en el futur esperaran dels seus empleats que siguin competents en la utilització de la funcionalitat de la IA generativa. Així, l'ús tant de la IA com dels seus subtipus en tota activitat humana esdevé gradualment la norma i una necessitat en la lluita competitiva.

Les principals orientacions de la recerca lituana en IA inclouen l'aprenentatge automàtic, l'enginyeria ontològica, la modelització semàntica del coneixement i els sistemes de bases de dades, la modelització conceptual del dret, la mineria de dades i el descobriment de coneixement, les xarxes neuronals artificials, la identificació de dades, la lògica matemàtica aplicada, etc.⁶² Per a desenvolupar el mercat de la IA a Lituània i enfortir l'ecosistema de recerca i innovació del país a escala nacional i internacional, la Universitat de Tecnologia de Kaunas i els seus socis estan creant un centre d'excel·lència mixt en IA per al desenvolupament sostenible. S'espera que aquest centre augmenti les oportunitats de digitalitzar els sectors de la producció, la salut, el transport i l'energia de Lituània i que garanteixi el desenvolupament de la recerca, el desenvolupament experimental i la innovació mitjançant l'aplicació de solucions basades en tecnologies d'IA.⁶³

Avui dia, les indústries líders en la implantació de tecnologies d'IA a Lituània són el sector financer, les tecnologies de la informació i la comunicació i la biotecnologia. Les solucions d'IA també s'implanten amb èxit als sectors de l'energia, el transport i la salut. Moltes indústries, com ara la dels serveis en el sector de la construcció, continuen sent conservadores pel que fa a la digitalització. Les taxes d'implantació de solucions d'IA són encara baixes en aquestes àrees, que tenen, però, oportunitats força importants per al desenvolupament tecnològic.

Aquestes són algunes zones de Lituània on aquestes tecnologies podrien crear nous llocs de treball o substituir els existents. En les condicions actuals, tenint en compte el ràpid desenvolupament de la IA generativa, cal reforçar la implantació de solucions d'IA aplicades en el sector real a Lituània i accelerar la transformació digital dels sectors econòmics prioritaris. No obstant això, el principal factor a tenir en compte és que l'impacte d'aquestes tecnologies en el mercat laboral lituà dependrà de molts aspectes, inclosos el progrés tecnològic, la regulació i l'interès i l'acceptació públics.

Què en pensen els vostres diputats?

El desenvolupament de la IA i els seus subtipus té el potencial de canviar el model econòmic, tecnològic i polític de la societat. Per tant, és important que el legislador prevegi com aquestes tecnologies afectaran els ciutadans i com fer que aquests canvis siguin més fàcils i menys arriscats i que garanteixi les responsabilitats corresponents.

Els parlamentaris creuen que avui hi ha una amenaça real relacionada amb la transformació del mercat laboral i l'atur digital, perquè, per exemple, la transició tecnològica fa desaparèixer algunes professions i hi ha un risc de biaix i imprecisió en l'ocupació.

62. Dr. Tomas Lavišius; Giedrius Kanapka. Líder de la recerca: Dr. Alvidas Lukošaitis. ACCENTS I DESENVOLUPAMENT DE L'ENTORN LEGAL DE LA IA. Tendències, riscos, recomanacions, decisions a la UE, alguns estats membres, altres països i organitzacions internacionals. 09.01.2023, Unitat de Recerca i Departament d'Informació i Comunicació del Parlament de la República de Lituània (Seimas). Aquest estudi el va encarregar la Comissió per al Futur del Seimas.

63. A Lituània s'està creant un centre d'excel·lència en IA. https://www.15min.lt/verslas/naujiene/mokslas-it/lietuvoje-kuriamas-dirbtinio-intelekt-ekscelesencijos-centras-1290-2005470?utm_medium=copied.

Es creu que la introducció de la IA generativa té un impacte diferent en els homes i les dones, ja que la feina de les dones és més del doble de susceptible a l'automatització i la naturalesa del treball en si es qualifica com una activitat d'oficina. Per tant, malgrat el potencial generador de la IA, encara hi ha moltes llacunes en aquestes tecnologies que cal abordar, perquè aquests sistemes estan entrenats per a utilitzar grans conjunts de dades que poden contenir informació esbiaixada o incorrecta. Atès el progrés de la IA en la creació artística, sempre hi haurà preguntes sobre l'existència de drets de propietat intel·lectual, ja siguin drets d'autor, patents, etc.

Tot i que la IA fa presagiar eficiència, la supervisió humana continua sent fonamental per a reduir el biaix. Hi hauria d'haver un filtre d'ajust primari quan s'utilitzen els sistemes d'IA en qualsevol nivell. Aquí cal trobar un equilibri raonable entre les activitats de la IA generativa i les activitats humanes, fent realitat els objectius de la IA a fi de simplificar i optimitzar els tràmits socials, administratius o empresarials.

Tot i que la IA pot substituir totalment o parcialment els humans en diferents llocs de treball, també ofereix grans oportunitats per a crear-ne de nous. Tanmateix, cal fer una crida als empresaris perquè quan es preparin per implantar la IA pensin abans en els empleats. Vols que la gent sàpiga per endavant el que ha de venir. A més, els empresaris haurien de facilitar també totes les condicions per a formar els seus empleats. Així mateix, si s'ha de fer res s'ha de fer d'una manera transparent i clara, en coordinació amb els treballadors i els seus representants.

Com la IA generativa pot ajudar en el treball del Seimas (el coautor d'aquesta resposta va ser el ChatGPT. El ChatGPT no va escriure la resposta sencera, però es va encarregar de gestionar el blog creatiu, generar frases individuals i paràgrafs del text i considerar diversos casos d'ús d'IA generativa. Escriure aquesta resposta amb ChatGPT és un signe benvingut d'interacció entre l'home i l'ordinador que es pot convertir en la nova norma.)

La IA generativa pot ajudar en el treball del Seimas de diverses maneres, millorant els processos, l'eficiència i la presa de decisions. A continuació s'especifiquen algunes maneres amb què la IA es pot aplicar al Parlament de Lituània:

Anàlisi i previsió de dades. La IA es pot utilitzar per a analitzar grans conjunts de dades a fi d'aïllar tendències i identificar qüestions crítiques. Això ajudaria els diputats del Seimas a entendre millor l'actualitat i a prendre decisions amb coneixement de causa.

Suport a la decisió. Els sistemes d'IA poden proporcionar informació i recomanacions als diputats tenint en compte diverses alternatives polítiques, econòmiques i socials. Això els podria ajudar a prendre les millors decisions.

Automatització de sistemes. La IA es pot utilitzar per a automatitzar tràmits administratius del Seimas, com ara la gestió documental, la traducció, l'organització de reunions i altres feines habituals, estalviant d'aquesta manera temps i recursos humans.

Traducció i comunicació. Les tecnologies de traducció basades en la IA poden facilitar la comunicació amb països estrangers, permetent als diputats de comunicar-se amb els d'altres països i seguir els afers internacionals.

Atenció al ciutadà. Els llocs web i les aplicacions poden utilitzar robots del llenguatge o bots de conversa basats en sistemes d'IA per a oferir als ciutadans informació sobre les activitats del Seimas, els tràmits parlamentaris i la legislació.

Seguretat. La IA es pot aplicar per a supervisar la seguretat del sistema d'informació del Seimas i detectar-hi possibles llacunes.

Anàlisi i informes. La IA pot generar informació detallada i informes sobre les activitats del Seimas i sobre les decisions parlamentàries i el seu impacte en la societat, facilitant l'avaluació del procés de presa de decisions.

Visualitzacions i sistemes d'informació. La IA pot crear visualitzacions digitals interactives que permetin als diputats comprendre més fàcilment la informació complexa i trobar la necessària.

Anàlisi del llenguatge legislatiu. La IA es pot utilitzar per a analitzar la legislació i el seu llenguatge a fi de garantir que sigui clar i comprensible per als ciutadans.

Traçabilitat de decisions. L'ús de la tecnologia de cadena de blocs, que també pot ser compatible amb la IA, pot garantir la traçabilitat i la transparència de les decisions.

És important assenyalar que la IA generativa en el treball parlamentari s'ha d'aplicar d'una manera responsable, tenint en compte els aspectes ètics, de privacitat i de seguretat. També és important garantir que els diputats i el personal estiguin degudament formats i entenguin com s'han d'utilitzar aquestes tecnologies d'una manera eficaç, productiva i segura.

Quins són els principals actors?

La IA és potser l'únic gran camp de la tecnologia que avui ja cobreix tota la gamma d'activitats econòmiques, des de l'agricultura fins a l'Administració pública. Per tant, per a garantir que la IA estigui alineada amb els valors i els objectius humans cal entendre les necessitats i les expectatives dels diversos actors afectats o implicats en la IA i els seus projectes prototipus.

Una de les dificultats a l'hora d'identificar els actors de la IA generativa és que, atesa la magnitud de les conseqüències socials que comporta i les àrees econòmiques que implica, la majoria, si no totes, de les persones (societat i usuaris), de les organitzacions (científiques i educatives, empresarials, no governamentals, etc.) i de les institucions polítiques (parlaments i governs nacionals i internacionals) ho són.

Tots aquests actors creen un ecosistema per al desenvolupament i l'aplicació de la IA generativa en què tots col·laboren per desenvolupar i utilitzar aquesta tecnologia, tenint en compte, respectivament, els aspectes científics, empresarials, ètics i legals.

Per què és important per al vostre país?

El més important és que la IA generativa té el potencial de canviar radicalment la manera com la nostra societat i cadascun dels seus membres treballen, aprenen, es comuniquen i creen. Tanmateix, juntament amb els seus grans beneficis, també planteja reptes ètics, de seguretat i legals. Per tant, és important seguir de prop el desenvolupament i l'ús de la IA per a garantir que s'aplica d'una manera responsable i ètica.

Rellevància i debat socials i polítics

La IA generativa ha esdevingut un tema important en el debat públic i polític a causa de la seva influència en diversos àmbits de la vida pública i de l'Estat. Amb cada nou avenç tecnològic sorgeix la por que les màquines substitueixin els humans. Històricament, quan les màquines s'han utilitzat per a substituir treballadors a fi d'augmentar la productivitat i reduir costos s'han creat tot un seguit de nous llocs de treball que abans no existien. Per tant, fins a quin punt les noves tecnologies alteraran seriosament el mercat laboral dependrà, en molts casos, del desenvolupament de la tecnologia, la regulació governamental i les prioritats de la societat. Per a resoldre aquest repte, hem d'implantar aquestes importants mesures:

Formació i educació. Les inversions en educació i formació són necessàries per a permetre que les persones s'adaptin a les noves condicions de treball i a les tecnologies. La formació ha de ser contínua i s'ha de centrar en el desenvolupament de les competències necessàries per al futur mercat laboral.

Regulació del mercat de treball. Un factor important pel que fa a l'impacte de la IA en el mercat laboral serà, sens dubte, la política reguladora de l'Estat. És evident que qualsevol amenaça important que conduïxi a un augment significatiu i a llarg termini de les tensions socials relacionades amb l'atur, amb els efectes a llarg termini sobre l'ocupació i amb el capital humà ha de ser gestionada per l'Estat. Tanmateix, hem de reconèixer també que al llarg dels nou mesos des que la IA generativa va aparèixer a la societat no s'han produït canvis dràstics en el mercat laboral, tot i que els creadors de la IA generativa envaeixen ara també la creativitat «humana».

El ràpid creixement de la IA generativa també demostra l'acceptació i la confiança del públic en general en les tecnologies d'IA. El fet que tants usuaris hagin adoptat tan ràpidament el bot de conversa habilitat per IA demostra que la integració de la IA en la vida quotidiana és cada cop millor i més fàcil. Aquesta tendència és un indicador positiu per al futur de les aplicacions d'IA generativa, ja que mostra la voluntat d'explorar i utilitzar solucions d'IA en diversos camps.

Aquestes preguntes i aquests temes mostren que la IA generativa té un impacte significatiu en el mercat laboral i requereix un enfocament equilibrat que inclogui la justícia social, la protecció dels treballadors i el desenvolupament econòmic. Els debats sobre aquests temes es produeixen cada cop amb més freqüèn-

cia a la societat, entre els líders polítics i els representants empresarials a fi de trobar un equilibri entre els beneficis que pot aportar la IA generativa i els reptes que planteja per al mercat laboral.

Hi ha un debat en curs sobre l'impacte de la IA generativa en les nostres societats i democràcies?

Predir exactament com la IA generativa afectarà la democràcia és molt difícil, ja que les seves aplicacions potencials són en gran part desconegudes i aparentment il·limitades. Aquesta tecnologia no és només una aplicació més; més aviat, és una tecnologia fonamental, l'aparició de la qual s'assembla més a la del mateix internet. Per tant, pot afectar la democràcia tant directament, com ara canviant els mecanismes electorals i de governança, com indirectament, amenaçant de canviar els fonaments dels ecosistemes de la informació.

Els beneficis i els riscos d'utilitzar la IA generativa en el context de la democràcia i desenvolupar polítiques i legislació que garanteixin la responsabilitat i la seguretat d'aquesta tecnologia per a la societat.

En un procés democràtic és important garantir que la gent tingui l'oportunitat de participar en la presa de decisions relacionades amb la IA generativa. Això pot incloure debats públics, enquestes, participació i altres mitjans que permetin als ciutadans expressar les seves opinions. Un excés de regulació pot obstaculitzar el desenvolupament de la tecnologia i destruir-ne el potencial sense que s'hagi aprofitat. Un dels factors més importants que cal abordar és el desenvolupament de nous paràmetres millorats, classificadors estandarditzats i objectius per a avaluar el rendiment dels models d'IA generativa i la qualitat dels conjunts de dades.

Hi ha resultats d'enquestes entre la població sobre l'ús o les opinions amb relació a la IA generativa?

No n'hi ha cap.

Hi ha cap legislació vigent?

Tampoc no hi ha cap regulació específica en aquesta direcció a Lituània. Per tant, els problemes de l'ús tant simple com generatiu de la IA encara es troben a la «zona grisa». Per descomptat, hi ha iniciatives individuals d'investigadors i polítics, per la qual cosa s'espera que un determinat enfocament estatal de la IA i els seus subtipus es pugui desenvolupar en un futur proper, ja que cada cop més sorgeixen importants i controvertides qüestions legals pel que fa al desenvolupament, l'aplicació o l'ús d'aquest producte.

A més, cal centrar-se en l'aplicació de la llei de la IA de la UE, que aviat s'aprovarà. La llei de la IA estableix tot un seguit d'excepcions destinades als estats i que s'han d'utilitzar amb prudència (per exemple, el mode *sandbox*, excepcions per a proveïdors i usuaris d'IA a petita escala, article 55, etc.).

Actualment hi ha propostes polítiques o legislatives sobre aquests temes?

Com passa amb la majoria dels fenòmens de l'economia digital, la naturalesa jurídica de la IA i els seus subtipus no és òbvia. Per tant, el desenvolupament de la IA dicta la necessitat de trobar noves solucions per a la regulació legal de la tecnologia. Entre els temes rellevants de regulació normativa que sorgeixen de l'ús de la tecnologia d'IA, es distingeixen els següents:

- Privacitat de dades
- Seguretat i responsabilitat
- Tecnologia de dades massives
- Propietat intel·lectual
- Ètica, etc.

Què es pot dir sobre els *inputs* basats en la ciència i l'evidència (organismes implicats, enfocaments integrals o especialitzats, etc.) i els *inputs* democràtics (participació pública, transparència, etc.) que guien la presa de decisions polítiques sobre aquest tema?

Avui dia hi ha tot un debat sobre problemes que han anat molt més enllà dels límits de la tecnologia mateix i estan relacionats amb el desenvolupament de tota la comunitat humana. Les tecnologies en el camp

de la IA emergeixen més ràpidament que no el temps que la societat té per a entendre-les. Per tant, els *inputs* basats en la ciència i l'evidència i els *inputs* democràtics sobre el tema de la IA generativa són dos aspectes importants que haurien de servir de guia per als responsables polítics. Aquests dos principis diferents però complementaris ajudarien a garantir que les tecnologies d'IA generativa s'utilitzin d'una manera ètica, responsable i en benefici de la societat.

El paper de la IA en els debats

La vostra institució ha abordat aquest tema?

Aquest tema es va tractar en la darrera sessió de primavera del Seimas. Vam iniciar converses amb experts sobre els possibles riscos de l'ús d'eines d'IA generativa, la seva certificació, etc. Vam parlar de si calen mesures reguladores nacionals, que podrien garantir que els sistemes d'IA utilitzats compleixen la llei i l'ètica i són segurs i en general fiables.

Una de les preguntes importants era a qui caldria responsabilitzar si la IA generativa comportés resultats amb conseqüències jurídiques significatives. En la reunió de la Comissió, la conclusió unànime va ser que els responsables polítics, els científics i els líders de la indústria haurien de treballar plegats per a desenvolupar polítiques i mesures per a garantir l'ús responsable de la IA generativa. Tot el que sabem amb certesa ara mateix és que la IA no va enlloc i que amb el pas del temps la gent hi haurà de confiar cada cop més. En segon lloc, podem confiar en la dita de moda que la IA no substituirà els humans però que els humans amb IA substituiran els humans sense IA. Per tant, els qui no només puguin utilitzar aquestes eines, sinó també desenvolupar-les, seran els qui n'obtinguin més beneficis.

La IA ha tingut un impacte en els debats en curs?

Amb els grans canvis tecnològics, és important entendre que realment parlem de canvis tan grans i fonamentals, en les tecnologies d'IA, que canvien no només una indústria concreta, sinó que de fet canvien l'organització o l'existència de molts sectors en general. S'assembla a la introducció de l'electricitat en l'economia, i en aquest sentit té molt en comú amb els grans canvis tecnològics que s'han produït en el passat, units pel seu abast i el fet que acaben englobant gairebé totes les activitats humanes.

És necessari que tothom, inclosos la comunitat d'experts, l'Estat, el sistema educatiu i les autoritats reguladores, prengui mesures per a preparar la societat per a l'ús de la IA. La comunitat d'experts i les empreses han de reflexionar sobre l'esquema aplicable a la implantació de la IA en la vida quotidiana. Els filòsofs han de transmetre als ciutadans el seu raonament sobre la coexistència entre els humans i la IA. El sistema educatiu s'ha de reformar perquè ensenyi als nens els coneixements i les competències que calen no només per al present, sinó també per al futur proper i llunyà, i perquè ofereixi a les persones oportunitats d'aprenentatge al llarg de tota la vida. Seria deure de l'Estat vetllar per la seguretat i la protecció de la població davant els riscos i perills relacionats amb la IA.

Quines són les lliçons apreses de la IA?

Avui dia és molt important considerar diversos mecanismes de regulació i governança en el camp de la IA generativa, ja que la revolució tecnològica ha avançat cap a una d'intel·lectual. Fins ara qualsevol nova tecnologia ha tingut per objecte alliberar l'home del treball físic monòton, mentre que les xarxes neuronals li permeten alliberar-se del treball mental tediós.

Països Baixos

Intel·ligència artificial generativa: oportunitats, riscos i reptes polítics

Capítol de l'informe sobre els Països Baixos: IA generativa i treball

Descripció del sistema/problema en qüestió

De què es tracta?

La intel·ligència artificial generativa ha estat en el punt de mira des de la introducció d'aplicacions com ara ChatGPT, Bard GitHub, Copilot i DALL-E.

La intel·ligència artificial generativa fa referència a sistemes que poden generar contingut automàticament, com ara text o imatges, per exemple, una carta de sol·licitud, un missatge a les xarxes socials o continguts de marca en línia. Així, la IA generativa pot efectuar diverses tasques no rutinàries,⁶⁴ mentre que les tecnologies anteriors es limitaven sobretot a automatitzar tasques rutinàries.⁶⁵ Aquesta qualitat versàtil és alhora una font de grans expectatives però també de profundes preocupacions per al futur mercat laboral. No obstant això, de moment els estudis que informen d'efectes empírics encara són, i comprensiblement, limitats.

En els debats anteriors sobre robotització i automatització, es podien distingir tres categories d'impacte potencial en el treball.⁶⁶ A grans trets, els experts esperen aquest mateix tipus de repercussions per a la IA generativa, però encara se'n desconeix l'ordre de magnitud.

- **Quantitat de treball:** les noves tecnologies poden provocar la desaparició de tasques i llocs de treball, la qual cosa pot donar lloc a un atur (massiu) o a una polarització laboral en què determinats grups es vegin especialment afectats.⁶⁷ D'altra banda, a mesura que sorgeixen noves tasques o necessitats, les noves tecnologies fins ara també han creat noves tasques i llocs de treball.
- **Qualitat del treball:** les noves tecnologies poden perjudicar els llocs de treball, i donar lloc a feines estressants, deshumanitzadores i angoixants. D'altra banda, la nova tecnologia també es pot utilitzar per automatitzar tasques repetitives, avorrides o perilloses, i deixar als treballadors temps per a altres tasques.
- **Competències i habilitats necessàries:** els dos impactes anteriors demostren que les noves tecnologies sovint comporten un paper diferent per a les persones en el procés laboral, per al qual han d'adquirir competències i habilitats específiques. La qüestió és si finalment tothom podrà mantenir-se al dia en aquesta transició.

Quin és l'estat actual del vostre país (desenvolupament, desplegament, ús)?

Els Països Baixos tenen una posició sòlida en determinades disciplines científiques d'IA,⁶⁸ però no tenen empreses que desenvolupin grans models de llenguatge o sistemes d'IA generatius, com molts altres països d'Europa.

64. Brynjolfsson, Li en Raymond (2023)

65. Autor, David H. 2015. Why Are There Still So Many Jobs? The History and Future of Workplace Automation [Per què encara hi ha tantes feines? Història i futur de l'automatització del lloc de treball]. *Journal of Economic Perspectives*, 29 (3): 3-30.

66. En general, cal reorganitzar els processos de treball i les organitzacions abans que es puguin incorporar les funcionalitats de les noves tecnologies (Institut Rathenau 2015)

67. OCDE (2019) Job polarisation and the middle class. New evidence on the changing relationship between skill levels and household income levels from 18 OECD countries [Polarització laboral i classe mitjana. Noves proves sobre la relació canviant entre els nivells de competència i els nivells d'ingressos de les llars de 18 països de l'OCDE]. <https://doi.org/10.1787/4bf722db-en>

68. Institut Rathenau (2021)

No hi ha enquestes oficials que indiquin l'acceptació dels usuaris. A les xarxes socials, blocs, etc., els usuaris holandesos informen amb entusiasme de les maneres en què incorporen l'ús de la IA generativa en el seu treball diari. ChatGPT té 100 milions d'usuaris a tot el món, però no se sap quants usuaris són d'Europa o dels Països Baixos.

Què en pensen els vostres diputats?

Pel que fa a la IA generativa, consideren que hi ha tant oportunitats com riscos. Difereixen en la relació entre la IA generativa i la democràcia. Hi ha qui diu que aquesta tecnologia redueix el llinar de difusió d'informació falsa. Un altre subratlla que, segons la seva opinió, la tecnologia en si no és cap perill, però sí que ho és fer-ne un mal ús.

Creuen que la IA generativa tindrà un impacte en totes les àrees polítiques. Pel que fa a l'impacte en el treball, esperen un possible millor encaix entre l'oferta i la demanda, la pèrdua de llocs de treball «antics» i la demanda de nous llocs de treball.

Quins són els principals actors?

Empresaris i treballadors (tant assalariats com autònoms/empresaris), sovint representats per organitzacions comercials i sindicats, institucions educatives per a la (re)formació dels treballadors, el Ministeri d'Afers Socials i Ocupació, el Ministeri d'Educació, Cultura i Ciència i el Ministeri d'Economia tenen tradicionalment un paper important en els debats sobre els riscos i les oportunitats de la tecnologia en el mercat laboral.

El Consell Social i Econòmic Holandès (SER) és un assessor clau del Govern i s'espera que el Govern li demani assessorament sobre l'impacte de la IA generativa en l'economia i el mercat laboral dels Països Baixos, com a part de l'estratègia del Govern actualment en desenvolupament. Es preveu que l'estratègia del Govern es publiqui a finals del 2023.

Per què és important per al vostre país?

Innegablement, els avenços tecnològics han afectat i seguiran afectant la nostra feina. A més, actualment hi ha una gran escassetat de mà d'obra en el mercat laboral holandès. En el sector públic, el problema de l'escassetat de personal és agut i s'espera que continuï sent un problema en els propers anys. El Consell Social i Econòmic Holandès va recomanar organitzar el treball d'una manera més «intelligent» i tècnicament innovadora per a estalviar mà d'obra.⁶⁹ Per aquest motiu, les organitzacions busquen solucions innovadores per a resoldre l'escassetat de personal, incloses les possibilitats de la IA generativa.

Rellevància i debat social i polític

Hi ha un debat permanent sobre l'impacte d'aquestes noves tecnologies en les nostres societats i democràcies?

Sí, alguns parlamentaris han posat la IA generativa a l'agenda política. En concret, van expressar la seva preocupació per la privadesa i els drets d'autor amb relació a la IA i van demanar al Govern que formulés una estratègia per a la IA generativa.⁷⁰

69. Valuable work: public service under pressure [Treball valuós: servei públic sota pressió] SER (2023), [Waardevol werk: publieke dienstverlening onder druk - Oplossingsrichtingen voor de arbeidsmarktkrapte \(ser.nl\)](#)

70. Moció esmenada dels diputats Dekker-Abdulaziz i Rajkowski sobre una visió global dels nous productes d'IA (corresponent a 26643-998) ([4 d'abril de 2023](#)), Resposta a les preguntes de la diputada Leijten sobre les preocupacions de centenars d'experts en tecnologia sobre l'evolució de la intel·ligència artificial ([Cambra de Representants, 27 de juny de 2023](#)), Moció dels diputats Dekker-Abdulaziz i Rajkowski sobre l'exigència de marques d'aigua en els tex-

En la premsa holandesa hi ha un debat actiu entre els acadèmics i les parts interessades sobre les oportunitats i els riscos potencials de la IA generativa. Nombrosos expolítics i líders d'opinió han signat una petició, iniciada, entre altres, per dos exdiputats centrats en la digitalització, que conté una crida urgent a l'acció. Aquests debats i peticions inclouen l'impacte potencial en el mercat laboral, però aborden una varietat més àmplia de qüestions socials i ètiques.

Hi ha alguna legislació vigent sobre aquesta qüestió?

Actualment no hi ha cap llei específica relacionada amb la IA generativa i el mercat laboral. A Europa, la Llei d'IA s'està negociant amb el Parlament Europeu, el Consell Europeu i la Comissió Europea, per a veure com es pot incloure la IA generativa en aquesta nova llei. A més, els marcs legals existents sobre protecció de dades, no-discriminació, responsabilitat, drets d'autor i propietat intel·lectual cobreixen, o poden cobrir, qüestions relacionades amb la IA generativa. Per exemple, fins a quin punt es permet als desenvolupadors d'IA generativa (específicament models de llenguatge extens) utilitzar les obres creatives de molts escriptors, artistes, etc.?

A més, les lleis específiques relacionades amb els empresaris i els empleats, com ara la Llei del comitè d'empresa holandès i les «bones pràctiques laborals» del Codi civil holandès, poden cobrir algunes qüestions relacionades amb la introducció i l'ús de sistemes d'IA generativa en el lloc de treball.

Hi ha actualment propostes polítiques o legislatives sobre aquests temes?

El Govern holandès està treballant en una estratègia sobre la IA generativa, a petició de la Cambra de Representants.⁷¹ «Estableix, entre altres coses, quines accions potencials pot prendre el Govern holandès per a garantir que aquesta tecnologia s'incorpori a la societat d'una manera responsable i fa una anàlisi més profunda de les qüestions de drets d'autor relacionades amb la IA generativa.»⁷²

La Llei europea d'IA, actualment en negociació, també serà aplicable als Països Baixos, un cop entri en vigor.

Què es pot dir sobre les aportacions basades en la ciència/evidència?

El llançament públic de ChatGPT i altres sistemes d'IA generativa és encara molt recent, de manera que hi ha pocs estudis empírics sobre l'impacte en l'ocupació i els mercats laborals. S'han elaborat alguns estudis de casos inicials, però no específicament per al mercat laboral holandès.

La IA generativa és capaç de crear contingut completament nou, com ara text, imatge o so. Per exemple, es pot fer servir per a elaborar notícies, articles, publicitat, resums, receptes, codi informàtic o fins i tot peces musicals senceres. Aquestes aplicacions farien molt més fàcil, si no redundant, la feina de periodistes, guionistes, venedors, programadors i músics. El mateix passa amb altres professions.

La IA generativa es pot utilitzar per donar suport al treball humà. Un estudi recent dels EUA mostra que els treballadors l'utilitzen per a generar idees, així com per a finalitats comunicatives, com ara escriure esborranys de missatges i correus electrònics, informes, resums o per a extreure errors d'un text o per a oferir suggeriments de millora.⁷³ Un altre estudi va demostrar que la IA generativa ja s'havia fet servir per a oferir suggeriments de trucades i possibles respostes en temps real als operadors d'atenció al client.⁷⁴ Això no

tos i imatges generats per IA en què s'hagin utilitzat fonts protegides per drets d'autor ([Cambra de Representants, 6 de juliol de 2023](#))

71. Moció esmenada dels diputats Dekker-Abdulaziz i Rajkowski sobre una visió integral dels nous productes d'IA (corresponent a 26643-998) ([4 d'abril de 2023](#))

72. Resposta a les preguntes del diputat Van Raan sobre la restricció de ChatGPT per motius de privadesa ([Cambra de Representants, 21 d'agost de 2023](#)). Vegeu també: Carta de compilació «Algoritmes de regulació» ([Cambra de Representants, 7 de juliol de 2023](#))

73. Cardon *et al.* (2023)

74. Brynjolffson *et al.* (2023)

només va provocar un augment de l'eficiència, sinó que també va fer que les converses es desenvolupessin més fluidament i els empleats experimentessin menys tensió social i emocional.

Actualment, les expectatives al voltant de la IA generativa es veuen matisades pel fet que el contingut generat sovint encara és susceptible de contenir errors i «al·lucinacions».⁷⁵ La tecnologia produeix informació que sembla coherent però que en realitat és inexacta i, a més, sovint incorrecta. Per tant, sovint encara es requereix la verificació del resultat i la possible edició per part d'un humà.

Dels estudis disponibles sorgeixen les oportunitats següents:⁷⁶

- Guanyos en eficiència i productivitat;
- Reforçar l'autonomia i augmentar la productivitat dels treballadors poc qualificats;
- Impacte positiu en la qualitat del treball;
- Solució a la rigidesa del mercat laboral.

Riscos possibles:⁷⁷

- Pèrdua de llocs de treball i/o ocupació;
- Desajust entre l'oferta i la demanda de competències i habilitats;
- Grups vulnerables que cauen pel camí;
- Impacte negatiu en la qualitat del treball.

Paper de l'AT en els debats

El vostre institut s'ha ocupat d'aquest tema?

L'Institut Rathenau contribueix en l'estratègia sobre la IA generativa que el Govern holandès està desenvolupant actualment. Ho fem explorant la naturalesa del desenvolupament de la tecnologia, identificant aplicacions prometedores i riscos amb relació als valors públics i formulant opcions polítiques sobre aquesta base.

Prèviament, l'Institut Rathenau va publicar sobre l'impacte de la tecnologia en el treball i el mercat laboral:

- *Own rhythm of algorithm? Algorithmic management beyond the platform economy* [Té un ritme propi l'algoritme? Gestió algorítmica més enllà de l'economia de la plataforma] Institut Rathenau i TNO, tardor del 2023
- *Valued at work. Limits to digital monitoring at the workplace using data, algorithms and AI* [Valorats a la feina. Límits al monitoratge digital en el lloc de treball mitjançant dades, algoritmes i IA] ([Institut Rathenau, 2020](#))
- *Robotisering en automatisering op de werkvloer. Bedrijfskeuzes bij technologische innovaties* [Robotització i automatització en el lloc de treball. Opcions empresarials pel que fa a les innovacions tecnològiques] ([Institut Rathenau, 2018](#))
- *Werken aan de robotsamenleving. Visies en inzichten uit de wetenschap over de relatie technologie en werkelijkheid* [Vers una societat robotitzada. Visions i intuïcions de la ciència sobre la relació entre tecnologia i realitat] ([Institut Rathenau, 2015](#)) (sol·licitat pel Parlament holandès)

L'AT ha tingut un impacte en els debats en curs?

En el debat públic, gran part del focus actual se centra en la possible pèrdua de llocs de treball. No obstant això, encara no hi ha proves empíriques. En el passat, hi ha hagut grups vulnerables els llocs de treball

75. Heikkilä (2023) Why You Shouldn't Trust AI Search Engines [Per què no hauríeu de confiar en els motors de cerca d'IA]. MIT Technology Review; Peng *et al.* (2023) Check your facts and try again: improving Large Language Models with external knowledge and automated feedback [Comproveu les vostres dades i torneu-ho a provar: millora dels models de llenguatge extens amb coneixements externs i comentaris automatitzats].

76. o.a. Brynjolfsson *et al.* (2023), Noy en Zhang (2023) Cardon *et al.* (2023) en Alshurafat (2023)

77. Gmyrek *et al.* 2023; Hatzius *et al.* 2023

dels quals eren més susceptibles a la pèrdua d'ocupació i que lluiten per adquirir noves habilitats per als nous tipus de llocs de treball que crea la tecnologia. Tot i que les noves tecnologies han estat beneficioses per al conjunt de la societat, aquests grups van requerir una atenció addicional per part dels responsables polítics pel que fa a la seguretat social.

Els nostres estudis demostren que les noves tecnologies també poden tenir un impacte significatiu en la qualitat del treball i en les competències que s'exigeixen. Amb els nostres estudis anteriors, vam poder demostrar que els partits polítics més orientats a l'esquerra consideren que la robotització i l'automatització també creen oportunitats per al mercat laboral i l'economia, i fan que els partits orientats a la dreta estiguin més atents als grups vulnerables del mercat laboral. També vam cridar l'atenció sobre altres qüestions socials i ètiques relacionades amb la robotització i l'automatització, més enllà dels impactes en els llocs de treball.

Quines són les lliçons apreses de l'AT?

La IA generativa és una tecnologia de ràpid desenvolupament i la seva direcció i funcions són difícils de predir. Per tant, és important que es faci una investigació sobre com canvia els llocs de treball i el mercat laboral.

L'AT pot ampliar el debat sobre la IA generativa i el treball. L'AT mostra que l'impacte potencial de la IA generativa en el mercat laboral no és inequívoc, sinó variat i complex, amb possibles resultats tant positius com negatius. A més de centrar-se en la qualitat del treball, és important invertir en ensenyar noves habilitats als empleats o fins i tot oferir-los una oportunitat per a canviar el camp en què treballen.

Referències

Selecció d'articles

- Alshurafat, H. (2023) The usefulness and challenges of chatbots for accounting professionals: Application on ChatGPT [La utilitat i els reptes dels chatbots per als professionals de la comptabilitat: Aplicació a ChatGPT]. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4345921
- Brynjolfsson, E., Li, D. i Raymond, L. R. (2023) Generative AI at work [IA generativa en el treball] (núm. w31161). Oficina Nacional d'Investigació Econòmica.
- Cardon, P. W., Getchell, K., Carradini, S., Fleischmann, C. i Stapp, J. (2023) Generative AI in the Workplace: Employee Perspectives of ChatGPT Benefits and Organizational Policies [IA generativa al lloc de treball: Perspectives dels empleats sobre els beneficis i les polítiques organitzatives de ChatGPT].
- Chohan, Usman W. (2023) Generative AI, ChatGPT, and the Future of Jobs [IA generativa, ChatGPT i el futur dels llocs de treball]. Document acadèmic SSRN. Rochester, NY, 29 de març de 2023. <https://doi.org/10.2139/ssrn.4411068>.
- Davenport, T.H. i Mittal, N. (2022) How generative AI is changing creative work [Com la IA generativa està canviant el treball creatiu]. Harvard Business Review. <https://hbr.org/2022/11/how-generative-ai-is-changing-creative-work>
- Eloundou, Tyna, Sam Manning, Pamela Mishkin i Daniel Rock (2023) GPTs Are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models [Els GPT són GPT: Una primera mirada a l'impacte potencial en el mercat laboral dels models de llenguatge extens]. arXiv, 23 de març. 2023. <https://doi.org/10.48550/arXiv.2303.10130>.
- Gmyrek, P., Berg, J., Bescond, D. (2023) Generative AI and Jobs: A global analysis of potential effects on job quantity and quality, ILO Working Paper 96 [IA generativa i llocs de treball: Una anàlisi global dels efectes potencials sobre la quantitat i la qualitat de l'ocupació, document de treball 96 de l'OIT] (Ginebra, OIT). <https://doi.org/10.54394/FHEM8239>
- Hatzius, J., J. Briggs, D. Kodnani, F. Pierdomenico (2023) The potentially large effects of artificial intelligence on economic growth [Els efectes potencialment importants de la intel·ligència artificial sobre el creixement econòmic]. Goldman Sachs.

Noy, Shakked i Zhang, Whitney (2023) Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence [Evidència experimental sobre els efectes de productivitat de la intel·ligència artificial generativa] (1 de març de 2023). Disponible a SSRN: <https://ssrn.com/abstract=4375283> or <http://dx.doi.org/10.2139/ssrn.4375283>

Zarifhonarvar, Ali (2023) Economics of ChatGPT: A Labor Market View on the Occupational Impact of Artificial Intelligence, ZBW [Economia de ChatGPT: Una visió del mercat laboral sobre l'impacte laboral de la intel·ligència artificial, ZBW] - Centre d'Informació Econòmica de Leibniz, Kiel, Hamburg

Altres publicacions

<https://www.nytimes.com/2023/05/20/business/dealbook/the-optimists-guide-to-artificial-intelligence-and-work.html>

<https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-AI-the-next-productivity-frontier#introduction>

<https://www.axelera.ai/how-will-generative-ai-revolutionize-our-work/>

<https://www.forbes.com/sites/joemckendrick/2023/07/05/generative-ai-adds-new-dimension-to-productivity-that-were-just-beginning-to-understand/?sh=2a087e9964cb>

<https://www.businessinsider.com/chatgpt-jobs-at-risk-replacement-artificial-intelligence-ai-labor-trends-2023-02?international=true&r=US&IR=T>

<https://www.brookings.edu/articles/will-generative-ai-kill-jobs/>

Noruega

Consell Noruec de Tecnologia (Teknologirådet)

Desenvolupaments clau sobre IA generativa i democràcia a Noruega

De què es tracta?

La democràcia noruega és d'alta qualitat i es caracteritza per alts nivells de confiança i transparència de la societat. Tanmateix, l'aparició de la intel·ligència artificial generativa (IA) té el potencial de desafiar aspectes importants de la democràcia:

- Accés a informació fiable i veraç
- Confiança en les institucions públiques, com els mitjans de comunicació
- Autonomia individual
- Governança democràtica de la tecnologia d'IA
- Serveis públics de confiança

Per tant, serà important tant aprofitar les oportunitats que ofereixen els avenços tecnològics com establir mecanismes per a prevenir-ne el mal ús i generar resiliència.

IA generativa: Què hi ha en joc per a Noruega com a democràcia?

L'ús generalitzat d'eines d'IA generativa pot afectar i desafiar moltes funcions bàsiques de les societats democràtiques:

Accés a informació fiable

En les societats democràtiques, la capacitat d'accedir a la informació és crucial perquè les persones puguin prendre decisions amb coneixement de causa i expressar-se. Internet s'ha convertit en una infraestructura fonamental per a les democràcies modernes i un requisit previ per a la llibertat d'expressió i informació. Actualment, la quantitat de continguts en línia generats per IA augmenta ràpidament. Les eines d'IA generativa poden crear imatges fotorealistes, clonar o crear àudio de veu sintètica i escriure text de manera convincent amb aparença humana. Els continguts dels mitjans de comunicació es poden produir en massa d'una manera més fàcil, més barata i a una escala molt més gran que abans. La tecnologia és accessible a tothom, inclosos els actors amb males intencions. Les noves eines d'IA es poden utilitzar per a difondre desinformació, notícies falses i hipertrucatge. La qualitat i la credibilitat del contingut generat per IA també augmenten, cosa que dificulta esbrinar si un article de notícies, un usuari en línia o una imatge ha estat generat per IA o no.

Distingir la veritat de la mentida

Distingir la veritat de les mentides pot ser cada cop més difícil en un ecosistema d'informació digital inundat de continguts generats per IA. Les eines d'IA generativa es poden fer servir per a produir grans quantitats de desinformació, notícies falses i hipertrucatge en línia, que poden amenaçar l'accés de les persones a informació veraç, confondre i influir en el públic, augmentar la polarització política i fins i tot afectar les eleccions democràtiques. El contingut enganyós també es pot difondre sense voler. Com que els models de llenguatge extens no estan entrenats per a dir la veritat, poden «al·lucinar» i generar continguts falsos que semblen reals.

Confiança en els mitjans de comunicació

Els ciutadans noruecs tenen un alt nivell de confiança en les institucions públiques, especialment en comparació amb altres països.⁷⁸ No obstant això, la institució pública amb menys confiança a Noruega, se-

78. Noruega es troba entre els països de l'OCDE amb els nivells més alts de confiança pública: el 77% de la població declara que confia en el Govern, mentre que la mitjana de l'OCDE és del 47%.

gons l'Enquesta de confiança de l'OCDE del 2022, són els mitjans de comunicació.⁷⁹ Entre els factors que van contribuir a la manca de confiança hi havia les xarxes socials, la preocupació per les notícies falses i els estils de comunicació política. La IA generativa pot inundar l'entorn informatiu de continguts sintètics i erosionar encara més la confiança. Garantir que la societat pugui accedir a una varietat de mitjans de comunicació de confiança continuarà sent una mesura important per a contrarestar la propagació d'informació falsa i la desinformació en línia, a més de fomentar la resiliència entre la població mitjançant el pensament crític i la crítica de les fonts.

Autonomia individual

ChatGPT d'OpenAI i My AI de Snapchat representen un canvi cap a l'intercanvi d'informació en l'esfera privada, mitjançant interaccions entre humans i bots de conversa. El públic no té cap manera d'accedir a la informació sobre les interaccions, ni quina informació es comparteix, ni l'impacte significatiu d'aquestes interaccions en les opinions i percepcions de la gent. Les investigacions mostren que els joves tendeixen a confiar més en la informació proporcionada pels bots de conversa que en els humans.⁸⁰

Un altre repte important és el biaix implícit en els models de llenguatge extens. En teoria, ChatGPT hauria de ser políticament neutre. Tanmateix, segons investigacions recents, el bot de conversa respon d'acord amb els valors polítics liberals.⁸¹ Com a resultat del biaix dels sistemes d'IA, l'opinió de les persones i del públic es pot veure influenciada, o fins i tot manipulada, mitjançant les interaccions amb les noves eines d'IA. Un experiment amb 1.500 participants va demostrar que les actituds dels participants estaven sotmeses a una influència oculta (anomenada *persuasió latent*) mitjançant interaccions amb bots de conversa preprogramats.⁸²

Concentració de poder i governança democràtica

El desenvolupament i el desplegament d'eines d'IA generativa està liderat avui per empreses tecnològiques privades nord-americanes, com OpenAI, Google i Microsoft, per esmentar-ne algunes. Aquestes empreses disposen d'informació sobre una tecnologia amb potencial per a influir en la societat. No obstant això, el públic, els investigadors acadèmics i les empreses més petites només tenen accés a la informació que les empreses d'IA decideixen revelar. Això pot socavar una competència lleial, distorsionar la capacitat del públic per a innovar amb la tecnologia i debilitar els mecanismes públics de supervisió i control. A Noruega es debat actualment la necessitat d'un model de llenguatge extens nacional (vegeu la discussió a continuació).

Serveis públics fiables i de gran qualitat

Les eines d'IA generativa poden oferir a la societat noves oportunitats per a millorar la participació democràtica i la governança. Per exemple, els governs i les institucions del sector públic poden utilitzar les noves eines per a garantir serveis més inclusius, eficients i accessibles. Un exemple és l'ús de la tecnologia de la parla sintètica per a garantir una millor informació i serveis públics per a les persones amb discapacitat visual.

Segons l'Agència Noruega per a la Gestió Pública i Financera, la confiança pública està interrelacionada amb la capacitat de les autoritats per a oferir serveis bàsics al públic i adaptar-se i respondre a les necessitats dels ciutadans. De cara al futur, això indica que la ciutadania hauria d'aprofitar els avenços tecnològics per a millorar els serveis públics. No obstant això, la pressa per fer ús de noves eines d'IA podria comprometre altres interessos i valors i, potencialment, augmentar el risc d'ús arbitrari del poder, la discriminació i

79. Enquesta de confiança de l'OCDE (2022), <https://www.oecd.org/publications/drivers-of-trust-in-public-institutions-in-norway-81b01318-en.htm>

80. Brandtzæg, Skjuve, Dysthe & Følstad (2021) When the Social Becomes Non-Human: Young People's Perception of Social Support in Chatbots, CHI Conference Paper [Quan el fet social deixa de ser humà. Percepció dels joves del suport social als bots de conversa, document de la conferència CHI] <https://dl.acm.org/doi/10.1145/3411764.3445318>

81. <https://www.uea.ac.uk/news/-/article/fresh-evidence-of-chatgpts-political-bias-revealed-by-comprehensive-new-study>

82. Jakesch, Buschek i Naaman (2022) Interacting with Opinionated Language Models Changes User's Views [Interactuar amb models lingüístics canvia l'opinió de l'usuari], https://mauricejakesch.com/assets/pdf/aimc_influence.pdf

el tracte injust. Les institucions públiques que utilitzen noves eines d'IA tenen un paper fonamental a l'hora de garantir que la tecnologia sigui digna de confiança, fiable i transparent.

Avenços principals: augment de l'ús d'eines d'IA generativa en diferents sectors a Noruega

Els mitjans digitals estan aprofitant la IA generativa i en regulen l'ús

A mesura que l'ús d'eines d'IA generativa es generalitza, els mitjans de comunicació noruecs s'hi han hagut d'adaptar ràpidament. Molts mitjans de comunicació han començat a experimentar amb la IA generativa. Per exemple, una nova característica comuna en els articles de notícies és un resum de l'article en qüestió, generat per IA. Quan l'utilitza el diari noruec Aftenposten, diu: «Aquest resum ha estat creat mitjançant la intel·ligència artificial i la qualitat ha estat assegurada pels periodistes d'Aftenposten». Una altra eina que fan servir alguns mitjans de comunicació és la veu sintètica o clonada, que permet escoltar la notícia en lloc de llegir-la.⁸³

A més d'experimentar amb les noves eines d'IA, els mitjans de comunicació també han desenvolupat directrius internes per a regular-ne l'ús. Per exemple, el canal de radiodifusió noruec TV2 té directrius que afirmen que tots els continguts generats per IA s'han de contrastar amb altres fonts creïbles i s'han de revisar manualment abans de publicar-los. Per a mantenir la integritat de la fotografia en el periodisme, es limita l'ús d'imatges i vídeos generats amb IA. Si s'utilitzen aquestes imatges o vídeos, s'han d'etiquetar explícitament com a generats per IA i incloure detalls de les eines utilitzades. Altres mitjans de comunicació noruecs també adopten directrius similars.

El sector públic està explorant noves possibilitats d'IA, amb precaució

L'Agència de Digitalització de Noruega i el Consorci de Recerca en Intel·ligència Artificial de Noruega han elaborat una guia sobre l'ús de la IA en el sector públic.⁸⁴ Tot i que la majoria de projectes que impliquen IA generativa encara es troben en una fase exploratòria, demostren com les institucions aborden els nous avenços tecnològics. Per exemple, l'Administració Tributària noruega està experimentant amb la transcripció automàtica de trucades del públic, combinada amb l'anàlisi de dades en temps real, com a mesura per a donar suport als assessors de l'agència a l'hora d'oferir una orientació eficaç als seus clients. La Direcció General de Salut noruega està explorant l'ús d'IA generativa per a oferir una millor informació a infants i joves amb necessitats complexes. El servei postal noruec està investigant si els models de llenguatge extens i ChatGPT poden millorar els seus serveis d'atenció al client.

A més, l'Agència Noruega de Digitalització ha desenvolupat directrius per a l'ús responsable de la IA en el sector públic.⁸⁵ A part d'oferir orientació sobre el desenvolupament i l'ús de la IA i la gestió del risc, les directrius inclouen consells explícits sobre l'ús d'eines d'IA generativa. Per exemple, si s'utilitzen eines generatives per a generar imatges, les institucions haurien de tenir en compte les qüestions legislatives no resoltes amb relació als drets d'autor. També es recomana a les institucions que siguin conscients que els models d'IA generativa no estan entrenats per a dir la veritat ni s'alineen necessàriament amb els valors socials. Pel que fa a les eines de generació de text, les directrius descriuen usos específics que poden ser beneficiosos per a les institucions públiques. Es recomana precaució quan les institucions públiques utilitzen eines d'IA generativa per a interactuar amb el públic. En aquests casos, el públic sempre ha de rebre informació del fet que està interactuant amb una eina d'IA.

83. Vegeu els experiments d'Aftenposten per a crear una veu sintètica per a les notícies: <https://futurereport.schibsted.com/an-ai-voice-makes-news-accessible-to-everyone/>

84. Panorama de l'ús de la IA en el sector públic de Noruega: <https://data.norge.no/kunstig-intelligens>

85. Orientació sobre l'ús de la IA generativa en el sector públic: <https://www.digdir.no/kunstig-intelligens/bruk-av-generativ-kunstig-intelligens-i-offentlig-sektor/>

Ús entusiasta però amb reticències de les eines d'IA generativa entre els parlamentaris

El nostre qüestionari per als diputats sobre IA generativa indica que l'ús d'eines d'IA s'està generalitzant. Tot i que la majoria de parlamentaris han experimentat amb les eines per a generar tant text com imatges, sembla que s'utilitza principalment per a l'entreteniment. Alguns van esmentar que era més complicat utilitzar la IA generativa per a escriure discursos que escriure'ls ells mateixos, ja que coneixen millor el seu públic. Un representant va reconèixer que la IA podia generar un discurs comparable al seu, però va optar per no utilitzar-lo perquè li semblava que era «incorrecte» o «poc autèntic».

Molts representants van destacar l'impacte positiu dels mitjans de comunicació que fan servir la tecnologia per a resumir les notícies. També es va destacar que les noves eines d'IA es podrien utilitzar per a comunicar polítiques i decisions polítiques al públic. Tanmateix, hi ha una aprensió compartida sobre un possible ús indegut, especialment pel que fa a la informació falsa, la desinformació i l'hipertrucatge. A més, els representants van expressar la seva preocupació pel biaix implícit en els models de llenguatge extens, que podrien amplificar trets negatius de la societat, com ara l'homofòbia, el sexisme i els racismes.

Rellevància i debat social i polític

Creació d'un model de llenguatge extens per al noruec

Un debat polític clau és si Noruega hauria de desenvolupar un model de fundació de llenguatge extens per al noruec. La IA generativa basada en models de llenguatge extens s'integrarà en la major part de serveis digitals quotidians, tant públics com privats. Per tant, qui pugui dissenyar, distribuir, controlar i gestionar aquests models impactarà i influirà en la qualitat, la fiabilitat, les tendències ètiques i culturals d'aquests serveis. Són qüestions de política democràtica, ja que es refereixen al poder, la responsabilitat i l'equitat d'accés als serveis i institucions públiques.

Els models de llenguatge extens predominants, com ara GPT-4, s'entrenen principalment amb fonts en anglès. Tenen inclinacions culturals nord-americanes, ja que són desenvolupats i propietat d'empreses tecnològiques americanes. La informació que es comparteix sobre com es dissenyen, s'entrenen i s'ajusten aquests models és limitada. També ho és la informació sobre el processament de dades i les funcions del model des d'un punt de vista tècnic.

Un model de fundació de llenguatge extens per al noruec podria:

- Augmentar la supervisió democràtica del disseny i el desplegament de la IA generativa.
- Garantir la diversitat lingüística, amb un millor rendiment en llengua noruega, inclosos els dialectes i les llengües minoritàries.
- Reflectir i mantenir el context cultural i els valors noruecs en l'era digital.
- Protegir la privadesa, la transparència i la fiabilitat, assegurant el coneixement de la recollida i el processament de dades, d'acord amb la legislació i les normes nacionals.
- Contrarestar la monopolització del mercat i reduir la dependència de Noruega de la infraestructura d'IA desenvolupada i desplegada per empreses tecnològiques estrangeres.
- Millorar la innovació dels sectors públic i privat i el desenvolupament empresarial a Noruega proporcionant a les empreses un accés segur, previsible i holístic a un model noruec.

Construir un model noruec des de zero requereix recursos i costos. Una alternativa podria ser cooperar amb altres països nòrdics o europeus. Una altra qüestió es refereix a quants projectes de models de llenguatge extens amb suport del Govern hi hauria d'haver: un o diversos? Qui hauria de gestionar i fer funcionar aquests models i qui hauria de tenir-hi un accés prioritari?

Avui, diverses universitats noruegues han iniciat projectes per a construir els seus propis models. No obstant això, no hi ha cap garantia que aquests models assoleixin el nivell de qualitat desitjat. A més, per a entrenar aquests models es necessiten conjunts de dades en noruec més nombrosos i adaptats, i també competències especialitzades en IA. Finalment, hi ha la qüestió d'accedir a una potència informàtica suficient i als superordinadors.

Es necessita una regulació nacional de la IA?

Actualment, la UE negocia el projecte de llei d'intel·ligència artificial (AI Act). És probable que les negociacions finalitzin el 2023. Si és així, la llei entrarà en vigor el 2026. És probable que la llei d'IA es converteixi en llei noruega i una autoritat de supervisió l'haurà de fer complir a nivell nacional. Aquesta autoritat pot exigir que es retirin del mercat productes específics d'IA i tindrà accés a informació sobre els productes d'IA, com ara el codi font i les dades d'entrenament.

La necessitat d'una legislació nacional addicional per a regular la IA a Noruega és part d'un debat en curs, tant en el Parlament com en el Govern. Tot i que la llei d'IA implicarà determinades prohibicions i diversos requisits de productes nous, les autoritats nacionals encara podran decidir en quins sectors es pot i s'ha d'utilitzar la tecnologia, amb quina finalitat i com, especialment en àrees com ara l'aplicació de la llei, l'educació i l'Administració pública.

Una agència nacional per a contrarestar la desinformació i augmentar la resiliència?

Tot i que les institucions de mitjans fiables tenen un paper important en la lluita contra la informació falsa, no hi ha cap autoritat clarament definida per a abordar la propagació de les informacions falses i la desinformació a Noruega. En canvi, Suècia ha establert una Agència de Defensa Psicològica per a contrarestar la desinformació i reforçar la resiliència de la població.⁸⁶ Com que les eines d'IA generativa poden augmentar la quantitat de continguts sintètics i falsos en línia, podria caldre adoptar mesures addicionals per a garantir un sistema d'informació sòlid i sostenible. La Comissió de Preparació Total va proposar recentment la creació d'una agència nacional amb la responsabilitat general d'avaluar i contrarestar les amenaces de les operacions d'influència, centrada a reforçar la resiliència de la població davant la desinformació.⁸⁷ El llibre blanc de la Comissió es troba actualment en tràmit de consulta pública.

Una moratòria sobre l'ús de la IA en el sector públic?

Actualment hi ha dues propostes relacionades amb la IA generativa que es discutiran al Parlament aquesta tardor. El Partit d'Esquerra Socialista suggereix una moratòria sobre l'ús de la IA en el sector públic fins que no s'estableixin regulacions i directrius.⁸⁸ A la seva proposta també s'inclou un llibre blanc sobre la IA i com afectarà la societat. El Partit Liberal, en canvi, es mostra crític amb una moratòria.⁸⁹ En lloc d'això, proposen el desenvolupament d'un pla governamental per a impulsar la investigació en IA, que inclogui l'ampliació del nombre d'estudiants rellevants en camps de recerca com ara el dret, l'ètica i les matèries STEM. Al setembre, abans que aquesta proposta es discutís al Parlament, el primer ministre de Noruega va anunciar 1.000 milions de corones noruegues per a finançar la investigació i la innovació en IA.⁹⁰

El paper de l'avaluació de la tecnologia

El Consell Noruec de Tecnologia (NBT) fa molts anys que treballa en qüestions relacionades amb la IA. Actualment, la junta impulsa diversos projectes sobre IA generativa, a més de conscienciar sobre la qüestió el Parlament i els responsables polítics.

Agenda al Parlament i per als responsables polítics

L'NBT ha acollit diverses reunions del grup de treball parlamentari sobre tecnologia, sobre temes com l'avenç dels models de llenguatge extens i el seu impacte a Noruega, la regulació digital, la ciberseguretat, l'impacte de la IA en l'esfera pública, i també la llei d'IA. A més dels parlamentaris i els seus assessors,

86. Agència Sueca de Defensa Psicològica: <https://www.mpf.se/en/>

87. El llibre blanc de la Comissió de Preparació Total: <https://www.regjeringen.no/contentassets/4b9ba57bebae44d-2bebf845ff6cd5f5/no/pdfs/nou202320230017000dddpdfs.pdf>

88. La proposta del Partit d'Esquerra Socialista: <https://www.stortinget.no/no/Saker-og-publikasjoner/Publikasjoner/Representantforslag/2022-2023/dok8-202223-232s/>

89. La proposta del Partit d'Esquerra: <https://www.stortinget.no/no/Saker-og-publikasjoner/Publikasjoner/Representantforslag/2022-2023/dok8-202223-273s/>

90. <https://www.nrk.no/norge/regjeringen-med-milliardsatsning-pa-kunstig-intelligens-1.16546093>

també han assistit a les reunions parts interessades rellevants i experts. Per a la preparació de les reunions, s'han publicat diversos informes d'orientació política sobre aquests temes per als diputats.

L'interès per les qüestions relacionades amb la IA generativa ha estat massiu. Al llarg del 2023, l'NBT ha fet més de cinquanta presentacions sobre IA generativa, per a parlamentaris, ministeris, Administració de justícia, sector públic i privat, i també grups d'interès rellevants.

Un grup d'experts de base àmplia

L'NBT ha establert un grup d'experts que inclou investigadors en IA, dret, ètica i ciències socials, desenvolupadors d'IA i una de les institucions del sector públic més grans de Noruega. El grup identificarà i debatrà els problemes més urgents relacionats amb la IA generativa, en particular pel que fa a com afectarà la tecnologia en el lloc de treball, l'Administració pública, la seguretat, l'educació i les professions creatives. També formen part del projecte consultes més àmplies amb les parts interessades rellevants, com ara l'Administració pública i experts en temes relacionats amb la IA i el canvi climàtic. A la tardor es publicarà un informe sobre l'impacte de la IA generativa amb opcions polítiques per al Parlament.

Anàlisi de la IA generativa des del punt de vista dels drets humans

En cooperació amb la Institució Nacional Noruega de Drets Humans, l'NBT treballa actualment en una avaluació i anàlisi dels efectes de la IA generativa sobre la democràcia i la llibertat d'expressió, utilitzant els drets humans com a punt d'entrada per a l'avaluació tecnològica. L'anàlisi es publicarà en un informe i es presentarà en un acte amb participació parlamentària.

Cap al futur: un laboratori de democràcia

La IA generativa forma part cada cop més de la vida quotidiana de les persones. La nova tecnologia introdueix consideracions i dilemes socials relacionats amb els valors, l'ètica i la responsabilitat. Per aquest motiu, l'NBT planeja un esforç continu per a implicar la ciutadania, les parts interessades i els afectats per a ajudar a formular futurs consells polítics per als parlamentaris en el proper trienni. El Laboratori de Democràcia de la IA farà servir diversos mètodes de participació.

Portugal

Intel·ligència generativa artificial i treball

António Moniz,⁹¹ Marta Candeias,⁹² Nuno Boavida⁹³
(Observatori d'Avaluació Tecnològica, OAT/CICS.NOVA)

Descripció del sistema/problema en qüestió

De què es tracta?

Les aplicacions d'intel·ligència artificial (IA) generativa fa poc que estan disponibles, de manera que l'avaluació d'impacte encara és molt limitada, però es poden percebre algunes implicacions i riscos, que s'han anat documentant. Per a avaluar el potencial d'eines d'IA generativa com ara ChatGPT d'OpenAI i Bard de Google hem fet un petit exercici de prova. Hem demanat a ambdues eines d'IA generativa que cerquin aplicacions en les quals puguin tenir impacte i hem obtingut les següents:

Per a ChatGPT: atenció al client, educació, salut, creació de continguts, assistència personal i serveis financers. «Aquests són només alguns exemples de les possibles aplicacions de ChatGPT. A mesura que la tecnologia continuï evolucionant i millorant, és probable que veiem usos encara més innovadors i creatius d'aquest potent model de llenguatge» (maig de 2023).

Per a Bard: educació, atenció al client, creació de continguts, investigació, assistència personal, investigació mèdica i auxili en cas de catàstrofe. «Aquestes són només algunes de les moltes possibles aplicacions futures de Bard. A mesura que Bard es continuï desenvolupant, és probable que encara s'utilitzi de més maneres per a millorar les nostres vides» (9 de setembre de 2023).

Comparant ambdós resultats, vam observar que els camps d'aplicació són molt similars, la qual cosa significa que se'n poden anticipar els possibles impactes. Què en podem concloure?

En l'àmbit de l'**atenció al client**, es pot esperar que tinguí un impacte significatiu en molts sectors. Encara hi ha molt marge per als errors, però probablement afectarà tota la humanitat. Segons ChatGPT, es pot utilitzar en comerç electrònic, telecomunicacions i sanitat. Aquesta eina pot entendre consultes en llenguatge natural, proporcionar respostes rellevants als clients i millorar l'eficiència i la qualitat del servei al client. Bard pot respondre preguntes, solucionar problemes i resoldre qüestions de manera oportuna i eficient. Bard també es pot utilitzar per a personalitzar l'experiència del client, de manera que els clients se sentin valorats i apreciats.

Pel que fa a l'**educació**, ChatGPT es pot utilitzar com a tutor o mentor virtual per a ajudar els estudiants amb els deures proporcionant comentaris personalitzats i responent preguntes relacionades amb diverses matèries. També es pot utilitzar com a eina d'aprenentatge d'idiomes, ja que permet als estudiants de practicar les seves habilitats lingüístiques per mitjà de la conversa. Bard es pot utilitzar per a crear experiències d'aprenentatge interactives per a estudiants de totes les edats. Pot generar contingut personalitzat que s'adapti a les necessitats i els interessos individuals de cada alumne. Bard també es pot utilitzar per a proporcionar observacions i orientació als estudiants i ajudar-los a aprendre de manera més eficaç. Tanmateix, hi ha una forta oposició per part dels professors. Hi ha hagut debats en diverses universitats sobre aquest tema. Hi ha hagut un interès important entre els periodistes. També pot fer augmentar les preocupacions que els professors han plantejat sobre el possible mal ús d'aquestes eines.

Pel que fa a la **salut**, ChatGPT es pot utilitzar per a oferir assessorament mèdic i suport als pacients, especialment en situacions en què els pot ser difícil accedir als serveis mèdics presencials. També es pot

91. abm@fct.unl.pt

92. marta.candeias@tecnico.ulisboa.pt

93. nuno.boavida@fcsh.unl.pt

utilitzar per a oferir suport i teràpia de salut mental. Bard es pot utilitzar per a investigar condicions i tractaments mèdics. També es pot utilitzar per a desenvolupar nous fàrmacs i teràpies. Això el converteix en una eina poderosa per a metges, investigadors i pacients que busquen maneres més adients de millorar la salut i el benestar. També pot reduir la burocràcia entre els treballadors sanitaris.

Pel que fa a la **creació de contingut**, ChatGPT es pot utilitzar per a generar contingut per a diverses indústries, com ara el periodisme, el màrqueting i l'entreteniment. Pot ajudar a automatitzar tasques com ara la redacció d'articles, les publicacions en les xarxes socials i els guions de vídeo. Bard es pot utilitzar per a generar contingut creatiu, com ara articles, publicacions de bloc, guions i poemes. També es pot utilitzar per a traduir llengües, escriure diversos tipus de contingut creatiu i respondre preguntes de manera informativa. Això el converteix en una eina valuosa per a escriptors, periodistes i altres professionals creatius. Però també pot destruir els modes convencionals de producció creativa (en música, rodatges, escriptura).

Pel que fa a la **investigació**, només Bard va esmentar que pot accedir i processar informació des de diverses fonts, inclosos llibres, articles i llocs web. Això el converteix en una eina potent per a estudiants, investigadors i qualsevol altra persona que necessiti trobar informació de manera ràpida i senzilla. L'aplicació d'IA generativa i de models de llenguatge extens (en endavant, LLM) pot conduir en última instància a resultats positius. És difícil imaginar fins a quin punt afectarà les tasques i fins i tot el projecte de recerca.

- La cerca de documentació per a la investigació científica es pot millorar amb l'ajuda d'IA generativa.
- Es pot donar suport a l'execució de models o simulacions en un entorn de laboratori.
- La gestió i l'administració de la recerca (financera, programació, tràmits, tasques administratives, planificació) també serà un camp d'aplicació actual.

Els **assistents personals** encara estan en fase de desenvolupament per part dels desenvolupadors d'aplicacions d'IA generativa. El ChatGPT es pot utilitzar com a assistent personal per a ajudar les persones a gestionar tasques, horaris i cites. També es pot integrar en altres dispositius intel·ligents, com ara altaveus intel·ligents i sistemes domòtics, per a oferir una experiència d'usuari més satisfactòria. Bard es pot utilitzar com a assistent personal per a ajudar amb tasques com ara programar cites, organitzar viatges i gestionar les finances. També es pot utilitzar per a fer recordatoris, configurar alarmes i controlar dispositius domèstics intel·ligents. Això el converteix en una eina valuosa per a persones ocupades que volen estalviar temps i simplificar les seves vides. Pot tenir una acceptabilitat més alta a Portugal.

L'auxili en cas de catàstrofe és una aplicació hipotètica que seria molt útil, però cal veure què es pot fer realment, i també la resistència que hi tindran els treballadors de primers auxilis (bombers, policies, paramèdics, etc.). Precisament Bard va apuntar que es pot utilitzar per a proporcionar ajuda en cas de catàstrofes naturals com ara huracans, inundacions i terratrèmols. Es pot utilitzar per a coordinar els esforços de socors, proporcionar informació als supervivents i ajudar a reconstruir comunitats.

Pel que fa a **serveis financers**, només ChatGPT va apuntar que es pot utilitzar per a oferir assessorament financer i suport a persones i empreses. Pot ajudar amb tasques com ara pressupostos, assessorament d'inversió i planificació fiscal. Els desenvolupaments d'IA són més antics que la generació ChatGPT. Els seus impactes en el personal dels serveis financers van ser impressionants (desdibuixats per l'epidèmia de la Covid-19) i s'havien previst des de les entrevistes del 2018. Segons un director de tecnologia entrevistat el 2018, «l'aplicació de molts algorismes es va suspendre per a mantenir la pau social fins que la gent va arribar a l'edat de jubilació». A més, s'espera que hi hagi una transformació important en el sector, a causa de l'acceleració del ritme d'implantació de més programari i al fàcil accés als professionals de les tecnologies de la informació i la comunicació (en endavant, TIC), que encara comportarà més reducció dels treballadors del sector.

Segons l'Organització per a la Cooperació i el Desenvolupament Econòmic (en endavant, OECD), «tot i que l'impacte de l'última onada d'IA generativa encara no és del tot clar, les primeres estimacions de l'exposició ocupacional a la IA que tenen en compte les capacitats de models de llenguatge extens com ChatGPT arriben a conclusions similars a les d'estimacions anteriors d'exposició a la IA: són principalment les ocupacions ben remunerades que requereixen estudis o formació superiors a la mitjana les que estan més exposades a la IA» (OECD, 2023: 96).

Quin és l'estat de la qüestió al vostre país (desenvolupament, desplegament, ús)?

Segons l'últim informe de l'índex d'economia i societat digitals (DESI, 2022), les competències digitals a Portugal estan al voltant dels nivells mitjans de la UE o per sobre. Però les habilitats relacionades amb la creació de continguts digitals bàsics encara són molt baixes a Portugal (60%, molt per sota de la mitjana de la UE, gairebé el 70%). Al mateix temps, hi ha alts nivells d'inversió en el desenvolupament d'habilitats de competències digitals a Portugal (vegeu seguidament les decisions polítiques recents). Finalment, això significa que les eines d'IA generativa poden proporcionar suport per a desenvolupar la capacitat de «crear i editar contingut digital, millorar i integrar informació i contingut en un cos de coneixement existent, alhora que s'entén com s'han d'aplicar els drets d'autor i les llicències i se sap com donar-los instruccions comprensibles per a un sistema informàtic» (DESI, 2022: 8). Algunes aplicacions ja s'estan executant en algunes empreses: bots de conversa, com el de Logo (assegurances) i altres serveis bancaris. També hi ha alguns debats científics significatius en universitats, empreses i alguns grups d'interès rellevants (APDSI, APIA, GEE). Hi ha hagut un interès general en aquests debats, principalment entre els periodistes, si es té en compte el nombre de programes de televisió i articles de premsa publicats des que la IA generativa ha arribat a tots els ciutadans.

Què en pensen els vostres diputats?

De moment no hem tingut respostes. Tot i així, la narrativa sobre les tecnologies actualment ja inclou la IA, però de manera limitada. Els diputats no tenen coneixement ni accés a un bon informe sobre l'estat de la qüestió pel que fa a la IA generativa, més enllà de les declaracions de regulació en l'àmbit de la Unió Europea (UE).

Qui són els principals grups d'interès?

Partits polítics, sindicats i patronals.

APDSI – Associació per al Foment de la Societat de la Informació

APIA - Associació Portuguesa d'Intel·ligència Artificial

FCT – Fundació Ciència i Tecnologia

FCCN – Fundació Nacional per a la Informàtica Nacional

GEE – Estudis Econòmics del Ministeri d'Economia

OCS/CNCS – Observatori de la Ciberseguretat

Per què és important això per al vostre país?

El juny de 2019, el Govern portuguès va presentar l'estratègia nacional d'intel·ligència artificial ([AI Portugal 2030](#)) per a exposar els reptes i les oportunitats del creixent ecosistema d'IA a Portugal. Aquesta estratègia presenta el pla per a fomentar l'ús de la IA en el sector públic i privat durant els propers anys. El pla concentra les accions en la inclusió, l'educació, la qualificació, l'especialització i la recerca, ja que les persones són el principal motor d'un desplegament reeixit de la IA (AI watch).⁹⁴

- L'estratègia nacional d'IA de Portugal no revela xifres financeres ni estimacions, per a aplicar-la.
- Els nivells més baixos d'escolarització poden implicar oportunitats competitives per a les aplicacions d'IA generativa en la majoria de casos en què sigui possible la substitució humana.
- Menys força dels sindicats en els procediments de negociació en sectors clau.
- Els sindicats no tenen una posició rellevant sobre IA generativa.
- Els empresaris tenen informació conservadora sobre el potencial d'aquestes aplicacions.
- No hi ha estudis anticipats sobre l'aplicació de la IA generativa en el mercat laboral portuguès i l'impacte consegüent.
- A l'Administració pública s'han aplicat moltes solucions d'IA i més que se'n preveuen (per exemple, en el sistema de justícia).

94. https://ai-watch.ec.europa.eu/countries/portugal/portugal-ai-strategy-report_en

Rellevància i debat social i polític

Resultats de les enquestes

Hi ha un debat en curs sobre l'impacte de la IA generativa en les nostres societats i democràcies? Hi ha resultats d'enquestes fetes entre la població sobre l'ús o opinions pel que fa a la IA generativa?

Hi ha alguns debats en el món acadèmic i les associacions sectorials sobre la IA. Hi ha molts debats entre els creadors d'opinió sobre la IA generativa. Els interessos dels periodistes sobre les conseqüències tecnològiques per a la seva professió han donat lloc a un debat important sobre la IA general a Portugal. Hi ha enquestes públiques sobre temes generals, com ara les següents:

- INE (2020a) Inquérito à Utilização de Tecnologias da Informação e da Comunicação pelas Famílias - 2020 [Enquesta sobre l'ús de les tecnologies de la informació i la comunicació per les famílies - 2020]. Institut Nacional d'Estadística. Disponible a: https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaques&DESTAQUESdest_boui=415621509&DESTAQUESmodo=2&xlang=pt

Aquesta enquesta, que es va fer el 2020, tracta sobre la societat de la informació i el coneixement a Portugal. En concret, es tracta d'una enquesta sobre l'ús de les tecnologies de la informació i la comunicació en les famílies. Es va concloure que els usuaris d'Internet i del comerç electrònic havien augmentat significativament. El percentatge d'usuaris per a finalitats educatives s'havia més que doblat.

- INE (2020b) Inquérito à Utilização de Tecnologias da Informação e da Comunicação nas Empresas - 2020 [Enquesta sobre l'ús de les tecnologies de la informació i la comunicació a les empreses - 2020]. Institut Nacional d'Estadística. Disponible a: https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaques&DESTAQUESdest_boui=415622957&DESTAQUESmodo=2&xlang=pt

L'any 2020, al voltant del 97% de les empreses amb deu treballadors o més i el 42,8% dels treballadors utilitzaven un ordinador amb connexió a Internet amb finalitats professionals. El mateix any, un 61,1% de les empreses van declarar que tenien un lloc web propi o del grup econòmic al qual pertanyien. Així mateix, un 13,0% de les empreses amb deu treballadors o més utilitzaven dispositius o sistemes interconnectats que es podien supervisar o controlar remotament per mitjà d'Internet (l'Internet de les coses) i un 9,1% utilitzaven robots industrials o de serveis. El 2019, un 4,5% de les empreses amb deu treballadors o més van utilitzar la impressió 3D, amb impressores 3D de l'empresa o per mitjà de serveis d'impressió prestats per altres empreses.

Hi ha cap legislació vigent?

Els textos legislatius següents són exemples de normativa relacionada amb la IA i aprovada en els darrers anys a Portugal:

Establiment del programa «Qualifica Indústria», adreçat a les microempreses, petites i mitjanes empreses (PIME) dels sectors industrials, amb l'objectiu de donar suport als processos de qualificació i requalificació dels treballadors per a evitar l'atur futur [Ordenança 282/2023, del 14 de setembre (Ministeri de Treball)].⁹⁵

Definició dels models oficials i exclusius de la targeta ciutadana, els elements de seguretat física que la componen, els requisits tècnics i de seguretat que s'han d'observar en la captura de la imatge facial i les empremtes dactilars del titular de la sol·licitud, i també les mesures concretes per a la inclusió dels ciutadans amb necessitats especials en la societat de la informació que s'han de complir en la prestació de serveis d'acompanyament ciutadà (Ordenança 286/2017,⁹⁶ del 3 d'abril, per la qual es modifica el Codi de treball i la legislació relacionada dins del marc de l'agenda del treball digne).⁹⁷

Definició de l'ús de càmeres portàtils per a ús individual dels agents de policia (Decret Llei 2/2023, del 2 de gener).⁹⁸

95. <https://diariodarepublica.pt/dr/detalhe/portaria/282-2023-221642378>

96. <https://diariodarepublica.pt/dr/legislacao-consolidada/portaria/2017-205638899>

97. <https://diariodarepublica.pt/dr/detalhe/declaracao-retificacao/13-2023-213650804>

98. <https://diariodarepublica.pt/dr/detalhe/decreto-lei/2-2023-205557199>

Estratègia de desenvolupament de la base tecnològica i industrial de la defensa 2023-2033 (Resolució 52/2023, del 5 de juny, del Consell de Ministres).⁹⁹

Estratègia nacional de ciberdefensa (Resolució 106/2022, del 2 de novembre, del Consell de Ministres).¹⁰⁰

Reglament del sistema d'incentius «Empreses 4.0» (Ordenança 135-A/2022).¹⁰¹

Llei de comunicacions electròniques, que transposa les directives 98/84/CE, 2002/77/CE i (UE) 2018/1972, Llei 16/2022.¹⁰²

Actualment hi ha propostes polítiques o legislatives sobre aquests temes?

Recentment, hi ha un interès creixent en la IA generativa i general de les parts interessades relacionades amb la jurisdicció. S'han aprovat i establert algunes noves propostes legislatives. La majoria segueixen el marc de la legislació de la UE.

Tal com s'esmenta en l'últim número de la publicació *Employment Outlook* [Perspectives d'ocupació], de l'OEDE, «la IA té la capacitat d'automatitzar completament les decisions relacionades amb l'ocupació, incloent-hi quines persones que busquen feina veuen una vacant, la selecció de candidats en funció del seu currículum, l'assignació de tasques a la feina i les bonificacions, decisions de promoció o formació. Tot i que aquesta capacitat allibera temps perquè els directius se centren més en els aspectes interpersonals de la seva feina, planteja la qüestió de si les decisions que tenen un impacte significatiu en les oportunitats i el benestar de les persones a la feina s'han de prendre sense cap implicació humana, o almenys la possibilitat que un humà hi intervingui» (OEDE, 2023: 191). En general, continua el mateix document, «hi ha nous esforços legislatius que impedirien l'adopció d'eines de presa de decisions totalment automatitzades en entorns d'alt risc com el lloc de treball perquè exigeixen una supervisió humana (és a dir, un «humà en el cercle»)». Per exemple, la proposta de 2021 per a la llei d'IA de la Unió Europea que regula els sistemes d'IA es va posar a disposició per als vint-i-set estats membres de la UE per a abordar els riscos per a la seguretat, la salut i els drets fonamentals, incloses disposicions específiques per a l'ús de determinades aplicacions d'IA d'alt risc en el lloc de treball. Aquest també va ser el cas de Portugal.

Recentment, s'han incorporat uns quants canvis en el principal instrument de legislació laboral (Codi del treball). No obstant això, no hi ha gaires canvis introduïts pel que fa a les aplicacions d'intel·ligència artificial. Malgrat tot, els canvis que s'estan aplicant a partir dels recents canvis de l'abril de 2023 (Llei 13/2023) són molt importants i influeixen en l'actuació de les organitzacions que representen els treballadors.

El primer fa referència a les normes legals que regulen els contractes de treball. Tot i que el nou apartat 3.o esmenta l'exemple de les plataformes digitals, és important saber que l'ús d'algoritmes, d'intel·ligència artificial o de qualsevol altra eina relacionada no poden permetre, per dir-ho d'una manera simple, l'acomadament. Si això es pogués utilitzar en plataformes digitals en què s'hagués establert un contracte de treball, actualment el Codi de treball no ho permetria. Però finalment podem considerar tots els altres sectors. Així, per exemple, quan una empresa o organització utilitza els seus propis serveis o els de consultors de gestió de personal que utilitzen «algoritmes, intel·ligència artificial o qualsevol altra eina relacionada» per a canviar els «estàndards legals que regulen els contractes de treball», els treballadors poden estar protegits pel Codi de treball.

El segon, relatiu a l'article 24, fa referència al dret a la igualtat en l'accés a l'ocupació i al treball. La protecció també entra en joc aquí quan es demostra que la presa de decisions basades en algoritmes o altres sistemes d'intel·ligència artificial mina les disposicions legals relatives a l'exercici d'una activitat professional per part d'una persona estrangera o apàtrida, o les disposicions relatives a la conciliació de l'activitat professional amb la vida familiar. En general, podem dir que actualment hi ha una prevenció contra la presa de decisions de gestió dels recursos humans basades en algoritmes o altres sistemes d'intel·ligència artificial.

Els altres dos casos estan relacionats amb les organitzacions de representants dels treballadors. En el primer (article 424), els comitès de treballadors ara tenen dret a informar-se sobre els «paràmetres, els cri-

99. <https://diariodarepublica.pt/dr/detalhe/resolucao-conselho-Ministros/52-2023-213918210>

100. <https://diariodarepublica.pt/dr/detalhe/resolucao-conselho-Ministros/106-2022-202899924>

101. <https://diariodarepublica.pt/dr/legislacao-consolidada/portaria/2022-207756412>

102. <https://diariodarepublica.pt/dr/legislacao-consolidada/lei/2022-187527517>

teris, les normes i les instruccions amb relació als algorismes o altres sistemes d'intel·ligència artificial que afecten la presa de decisions sobre l'accés i el manteniment del lloc de treball, i també les condicions de treball, inclosa la creació de perfils i el control de l'activitat professional». Aquesta estipulació sembla molt interessant, ja que restringeix clarament la possibilitat d'aplicar «paràmetres, criteris, normes i instruccions» sobre «l'accés i el manteniment del lloc de treball, i també les condicions de treball, inclosa la creació de perfils i el control de l'activitat professional» que es basen en procediments automatitzats i programables. Tanmateix, aquesta restricció només es pot produir quan els comitès d'empresa són capaços de reconèixer aquests tràmits o la possibilitat que es facin servir.

Això implicarà que debatran aquests temes internament. Per exemple, ja hi ha una certa atenció en la relació entre la digitalització del treball i l'ocupació, però la implantació d'aquesta relació no sempre és evident. En el segon cas (article 466), són els sindicats els que hauran d'interioritzar aquest debat.

Aportacions basades en ciència/evidència i aportacions democràtiques

L'Observatori de la Ciberseguretat, que pertany al Centre Nacional de Ciberseguretat, té un enfocament multidisciplinari del fenomen de la seguretat informàtica, i integra en el seu camp d'observació diverses àrees de coneixement. Aquestes àrees són societat, economia, polítiques públiques, ètica i dret, riscos i conflictes, i també innovació i tecnologies de futur. A continuació, posen a disposició un conjunt de documents públics, d'entitats nacionals i internacionals, que serveixen de font per a estudiar cadascun d'aquests aspectes de la ciberseguretat a Portugal. Per exemple, han fet una enquesta, en col·laboració amb la Direcció General d'Educació, dirigida al professorat no universitari i centrada en les conductes relacionades amb la ciberseguretat que ha experimentat la professió docent durant el segon semestre del curs 2019/2020, durant el qual la formació a distància va substituir la docència presencial, arran de la pandèmia de la Covid-19.¹⁰³

El debat del marc

Els investigadors d'OpenAI van col·laborar amb el Centre de Seguretat i Tecnologia Emergent de la Universitat de Georgetown i l'Observatori d'Internet de Stanford per a investigar com es podria fer un ús indegut dels models de llenguatge extens amb finalitats de desinformació. La col·laboració va incloure un taller de l'octubre de 2021 que va aplegar trenta investigadors de desinformació, experts en aprenentatge automàtic i analistes de polítiques, i va culminar amb un informe de coautoria basat en més d'un any de recerca. Aquest informe descriu les amenaces que comporten els models de llenguatge per a l'entorn de la informació si s'utilitzen per a augmentar les campanyes de desinformació i introdueix un marc per a analitzar-ne les possibles atenuacions del risc.¹⁰⁴ En aquest article, els autors afirmen que «moltes de les propietats dels grans models generatius no s'entenen del tot. De la mateixa manera, encara falta claredat tant pel que fa a l'estructura com als impactes de moltes operacions d'influència, que es duen a terme en secret. La claredat sobre el grau d'amenaça que representen les operacions d'influència continua essent elusiva» (Goldstein *et al.*, 2023: 65).

Com també van revelar les millores recents del text generat per màquina, podem presenciar un augment dels abusos dels models de generació de llenguatge natural (NLG), com ara la pesca per correu electrònic, la desinformació, les ressenyes de productes fraudulents, la deshonestat acadèmica i el correu brossa tòxic. Les activitats de l'Observatori Nacional de Ciberseguretat han treballat en això a Portugal.¹⁰⁵ El seu informe sobre tecnologies emergents (Aguiar *et al.*, 2023) esmenta l'apartat 2 de l'article 9 de la Carta portuguesa dels drets humans a l'era digital, que intenta abordar la qüestió de les decisions automatitzades. «Es poden destacar alguns problemes pràctics, com ara compartir entorns de l'internet de les coses impulsats per IA en què diverses persones utilitzen l'entorn sense necessitat d'identificar-se. No obstant

103. <https://www.cnscs.gov.pt/en/observatory/#inqueritos>

104. <https://openai.com/research/forecasting-maluse>

105. Vegeu més informació a <https://www.cnscs.gov.pt/en/observatory/#relatorios>

això, una fiscalització de les decisions preses és fonamental per a poder revisar quina va ser la base d'una decisió i com es va produir, i els marcs legals en joc sembla que es demostrin insuficients per a les múltiples situacions que es preveuen per al futur» (Aguar *et al.*, 2023: 105).

A causa d'aquests problemes que aporten els models de generació de llenguatge natural, «abordar el risc d'abús és cabdal per a maximitzar els possibles avantatges de la tecnologia de generació de llenguatge natural, alhora que se'n minimitzen els danys, un principi clau per a una IA fiable» (Crothers, Japkowicz i Viktor, 2023: 1). Aquests autors canadencs també prediuen com n'és, d'important i decisiu, aquest tema. Subratllen que «predir la manera en què és probable que es desenvolupin els abusos i entendre com han de ser les millors defenses és essencial per a permetre a la humanitat d'obtenir els beneficis positius d'aquesta tecnologia, alhora que se'n minimitzen els possibles perjudicis. Hem de caminar amb cautela per l'era de les paraules de silici» (Crothers, Japkowicz i Viktor, 2023: 2).

Aquest debat és molt rellevant tenint en compte l'impacte de risc d'aquesta tecnologia i els problemes relacionats amb el seu ús. L'informe sobre l'àmbit social de l'Observatori Nacional de Ciberseguretat ha esmentat aquest tema. Aquest informe, a més, a banda d'analitzar les dades d'actituds, comportaments, formació i conscienciació respecte a la ciberseguretat, conté com a part destacada alguns indicadors correlacionats amb possibles impactes de l'Estratègia nacional de seguretat ciberespacial, un nou capítol sobre els usos de les tecnologies digitals en general, l'aprofundiment de l'enquesta sobre accions de sensibilització i un conjunt de recomanacions.

Paper de l'avaluació tecnològica en els debats

El vostre institut s'ha ocupat d'aquest tema?

L'Observatori d'Avaluació Tecnològica del centre de recerca CICS.NOVA, que és membre de l'EPTA, dirigeix el projecte InteliArt sobre les implicacions de la IA en el treball, l'ocupació i les relacions laborals, i es va iniciar l'any 2021. L'equip de recerca va publicar un article en una revista Q1 (Moniz, *et al.*, 2022) i ja ha guanyat dos premis (Candeias *et al.*, 2022).

L'OAT/CICS.NOVA també executa altres projectes en què s'aborda el tema. Alguns projectes estan més relacionats amb el treball de la plataforma digital (Moniz *et al.*, 2023; Boavida *et al.*, 2023) o amb l'aplicació de la IA en entorns industrials (Candeias, 2023; Moniz, 2018; Moniz i Krings, 2016).

L'avaluació tecnològica ha tingut un impacte en els debats en curs?

La investigació desenvolupada a OAT/CICS.NOVA s'ha difós i comunicat en molts fòrums científics, i també en moltes xarxes socials.

Quines són les lliçons apreses de l'avaluació tecnològica?

La pàgina web d'OpenAI¹⁰⁶ diu que «només perquè una atenuació del risc pugui reduir l'amenaça de les operacions d'influència habilitades per IA no vol dir que s'hagi de posar en marxa. Algunes atenuacions comporten els seus propis riscos negatius». També es plantegen qüestions importants en la pàgina web esmentada: És factible l'atenuació del risc des d'una perspectiva política, jurídica i institucional? Requereix una coordinació costosa, s'incentiven els actors clau per a implementar-la i es pot aplicar d'acord amb la llei, la regulació i els estàndards de la indústria? Aquestes són preguntes rellevants per a una anàlisi d'avaluació tecnològica.

Les qüestions finals de l'estudi de Goldstein *et al.* (2023) són reveladores. Presenten les qüestions següents: «Els desenvolupadors d'IA haurien d'alliberar o de restringir els seus models? Els investigadors d'Internet haurien de publicar tàctiques observades de propagandistes o mantenir-les en secret? Fins a quin punt les plataformes i els desenvolupadors d'IA poden conformar associacions significatives que puguin ajudar a detectar i eliminar contingut no autèntic? En el nivell més ampli, la implicació reflexiva amb totes aquestes qüestions, tant de persones de les indústries rellevants com d'observadors neutrals de tercers, és

106. <https://openai.com/research/forecasting-maluse>

una necessitat cabdal» (Goldstein *et al.*, 2023: 66). I aquestes són qüestions que l'avaluació tecnològica (AT) pot abordar una vegada es mostri clarament on són els principals factors de risc i els elements que requereixen un procés de decisió que impliqui diferents actors socials i parts interessades.

Bibliografia

- Acemoglu, D.; Lensman, T. (2023). *Regulating Transformative Technologies* [Regular les tecnologies transformadores], 6 de juliol, Massachusetts Institute of Technology (MIT), 2023 <https://toddlenman.com/files/research/tech-reg-8.pdf>
- Aguar, R.L. *et al.* (2023). *Tecnologias Emergentes* [Tecnologies emergents], Lisboa, CNCS, 115 p.
- Boavida, N.; Roque, I.; Moniz, A. B. (2023). «Collective Voice and Organizing in Digital Labour Platforms» [Veu col·lectiva i organització en les plataformes digitals laborals a Portugal]. A: *Journal of Labor and Society*: 1-25.
- Candeias, M. (2023). *Impact assessment of AI-enabled automation on the workplace and employment. The case of Portugal* [Avaluació d'impacte de l'automatització habilitada per la IA en el lloc de treball i l'ocupació. El cas de Portugal]. Caparica, tesi doctoral, UNL, <http://hdl.handle.net/10362/156952>
- Candeias, M.; Boavida, N.; Moniz, A. B. (2022). «Automation trends in Portugal: implications in productivity and employment» [Tendències de l'automatització a Portugal: implicacions en la productivitat i l'ocupació]. A: *GEE Paper*, núm. 165 https://www.gee.gov.pt/RePEc/WorkingPapers/GEE_PA-PERS_165.pdf
- Candeias, M.; Moniz, A. B.; Boavida, N. (2022). «Digital Transformation in the Automotive Sector in Portugal: Data Analysis of Industrial R&D Projects» [Transformació digital en el sector de l'automoció a Portugal: Anàlisi de dades de projectes industrials d'R+D]. A: *Inclusive Futures for Europe: Addressing the Digitalisation Challenges*, Actes de la conferència científica BEYOND 4.0 Sofia 2021. Sofia: Prof. Editorial Marin Drinov de BAS. [https://press.bas.bg/en/eBooks-105/show-106\(55\)](https://press.bas.bg/en/eBooks-105/show-106(55)).
- CGTP (2021). *Livro Verde Sobre o Futuro Do Trabalho - Apreciação Da CGTP-IN* [Llibre verd sobre el futur del treball - Avaluació del CGTP-IN]. Lisboa: CGTP. <http://www.cgtp.pt/cgtp-in/areas-de-accao/accao-reivindicativa/16100-o-futuro-do-trabalho-tem-de-ser-de-progresso-e-justica-social>.
- Crothers, E.; Japkowicz, N.; Viktor, H. (2023). *Machine Generated Text: A Comprehensive Survey of Threat Models and Detection Methods* [Text generat per una màquina: Una enquesta exhaustiva de models d'amenaça i mètodes de detecció] [arXiv:2210.07321v4](https://arxiv.org/abs/2210.07321v4)
- DESI (2022). *Digital Economy and Society Index 2022, Human Capital* [Índex d'economia i societat digitals 2022, capital humà]
- Duarte, J. B., *et al.* (2019). *O Futuro Do Trabalho Em Portugal: O Imperativo Da Requalificação*. Nova SBE, *Confederação Empresarial de Portugal* [El futur del treball a Portugal: L'imperatiu de requalificació]. Lisboa: NOVA SBE & CIP. https://www.cip.org.pt/wp-content/uploads/2019/10/Relatório-FoW_NSBE-CIP.pdf.
- Goldstein, J.A. *et al.* (2023). *Generative Language Models and Automated Influence Operations: Emerging Threats and Potential Mitigations* [Models de llenguatge generatiu i operacions d'influència automatitzades: Amenaces emergents i possibles atenuacions del risc], arXiv: 2301.04246v1
- Krings, B. J.; Moniz, A. B.; Frey, P. (2021). «Technology as enabler of the automation of work? Current societal challenges for a future perspective of work» [La tecnologia com a facilitadora de l'automatització del treball? Reptes socials actuals per a una perspectiva futura del treball]. A: *Revista Brasileira de Sociologia*, 9: 206-229.
- Mateus, A. *et al.* (2017). *Avanço Da Economia Digital* [Avenç de l'Economia Digital]. Lisboa: EY - Augusto Mateus & Associados
- Ministeri de Treball, Solidaritat i Seguretat Social de Portugal (2021). *Livro Verde Sobre o Futuro Do Trabalho* [Llibre verd sobre el futur del treball 2021]. Lisboa: Ministeri de Treball, Solidaritat i Seguretat Social
- Moniz, A. B.; Candeias, M.; Boavida, N. (2022). «Changes in productivity and labour relations: artificial intelligence in the automotive sector in Portugal» [Canvis en productivitat i relacions laborals: intel·ligència en el sector de l'automoció a Portugal]

- cia artificial en el sector de l'automoció a Portugal]. A: *Int. J. Automotive Technology and Management*, 22: 222-244, DOI: [10.1504/IJATM.2022.124366](https://doi.org/10.1504/IJATM.2022.124366)
- Moniz, A.B. (2018). *Robótica e trabalho. O futuro hoje* [Robòtica i treball. El futur avui]. Lisboa: Glaciar/FLAD http://www.glaciar.com.pt/detalhe_edicao.php?cd_edicao=65
- Moniz, A. B. (2023). «Inteligência artificial no Código de Trabalho» [La intel·ligència artificial en el Codi del treball] <https://www.esquerda.net/opinio/inteligencia-artificial-no-codigo-de-trabalho/87454>
- Moniz, A. B.; Krings, B. J. (2016). «Robots working with humans or humans working with robots? Searching for social dimensions in new human-robot interaction in industry.» [Robots que treballen amb humans o humans que treballen amb robots? Recerca de dimensions socials en la nova interacció humà-robot en la indústria]. A: *Societies*, 6 (3) <https://doi.org/10.3390/soc6030023>
- Moniz, A. B.; Boavida, N.; Makó, C.; Krings, B. J.; P. Sanz de Miguel (ed.) (2023). *Digital Labour Platforms. Representing Workers in Europe* [Plataformes laborals digitals. Representació dels treballadors a Europa]. Gaia: Humus/CICS.NOVA
- OCDE (2023). *OECD Employment Outlook 2023: Artificial Intelligence and the Labor Market* [Perspectives d'ocupació de l'OEDE 2023: Intel·ligència artificial i mercat de treball]. París: OECD Publishing <https://doi.org/10.1787/08785bba-ca>

Suècia

31-08-2023

Antecedents de la conferència de l'EPTA sobre IA generativa i democràcia

Introducció

A continuació es mostra una breu descripció de diverses decisions i iniciatives a Suècia abans de la conferència de l'EPTA del 2023 sobre IA generativa i democràcia, compilada per la Secretaria d'Avaluació i Recerca (ERS) del Parlament suec.

La presentació es basa principalment en declaracions del Parlament (Riksdag) i del Govern suecs en el camp de la IA. Fins ara no ha estat possible implicar cap diputat de Suècia per participar a la reunió del Consell de l'EPTA, motiu pel qual no ha estat possible recopilar les seves experiències i opinions. El memòr·ndum se centra en qüestions amb rellevància social i política.

En general, es pot dir que, com en la majoria de la resta de països, el debat públic sobre la IA i la seva importància amb relació a àrees polítiques diverses, com ara la democràcia, ha estat ampli i viu. La discussió sobre els reptes i les oportunitats que ofereixen la IA i la IA generativa és de gran actualitat entre els investigadors, els mitjans de comunicació i les persones amb responsabilitat política. Les diverses maneres de regular la IA ha estat una qüestió àmpliament debatuda, però de moment no s'ha pres cap decisió en aquest sentit. El debat gira primordialment al voltant de la IA i, en un nombre menor de casos més concrets, al voltant de la IA generativa. A continuació es descriuen algunes de les directrius i declaracions adoptades pel Riksdag i el Govern en el camp de la IA. Al Parlament encara no ha dut a terme cap AT formal en el camp de la IA generativa, però l'ERS ha elaborat documents objectius sobre la IA per a diverses comissions, per exemple recentment en el marc del Dia de la Recerca del Riksdag, el juny del 2023, quan la Comissió del Mercat de Treball i la Comissió de Transports i Comunicacions van organitzar seminaris sobre les conseqüències de la IA en els seus respectius àmbits de treball.

Decisions al Riksdag sobre la IA: Declaració sobre el llibre blanc de la Comissió Europea sobre la IA

El 2020, la Comissió d'Educació va emetre una declaració¹⁰⁷ sobre el llibre blanc de la Comissió Europea sobre la IA, COM (2020) 65. La Comissió va acollir amb satisfacció el llibre blanc i un debat sobre l'enfocament futur d'Europa de la IA. La Comissió va veure la necessitat de mobilitzar recursos i revisar quines mesures són necessàries perquè la UE pugui reforçar la seva capacitat en el camp de la IA. La Comissió va subratllar que és de gran importància que la recerca en IA es prioritzi a escala europea. Això s'aplica al suport tant a la recerca bàsica com a la recerca aplicada, com també als esforços dirigits a promoure l'excel·lència. La Comissió va considerar important garantir que es tingui en compte tota la cadena des de la recerca fins al mercat, per exemple facilitant la cooperació entre entorns empresarials i acadèmics. Alhora, va considerar de gran importància que els esforços no estiguin inequívocament dirigits a comercialitzar aplicacions d'IA. Els resultats de la recerca en IA també tenen un gran potencial per ser aplicats en el sector públic, i s'ha de promoure el suport a la recerca amb beneficis per a la societat. A més, la Comissió es va mostrar convençuda que les dades accessibles són d'una importància cabdal i que és important assegurar el treball per a dades obertes i d'alta qualitat a la UE.

La Comissió va assenyalar que la IA, com tota tecnologia, existeix per a servir les persones, i no al revés. Alhora, la IA comporta certs riscos especials, pot tenir efectes no desitjats i es pot utilitzar d'una manera

107. Declaració 2019/20:UbU20

que pot ofendre persones i grups. També pot haver-hi un excés de confiança en la tecnologia, cosa que pot comportar conseqüències imprevistes. Des d'aquesta perspectiva, un marc normatiu deficient no només significa que la UE corre el risc de quedar-se endarrerida perquè molt poca gent utilitzi solucions d'IA, sinó que també comporta un risc potencial per als drets fonamentals. Per tant, és d'una importància fonamental que un futur marc normatiu tingui en compte els aspectes ètics de la tecnologia de la IA i que la legislació, tant a escala sueca com europea, sigui adequada per a garantir que els riscos de la IA es minimitzen. En aquest context, la Comissió va assenyalar la importància d'integrar una perspectiva d'igualtat i drets en el treball continuat.

La Comissió també ha volgut destacar el tema dels riscos dels monopolis tecnològics i de dades per tal d'evitar que les dades que es poden utilitzar per al bé públic es limitin a uns quants actors. Això s'aplica també a les qüestions relacionades amb la propietat de les dades, que ha de ser clara perquè la informació recollida no s'utilitzi d'una manera indeguda. Per tant, la Comissió va acollir amb satisfacció un debat continuat sobre la transparència en els sistemes d'IA.

Política governamental sobre IA

El 2018, el Govern suec va presentar una orientació nacional per a la IA. Un objectiu general és convertir-se en un líder mundial a l'hora d'aprofitar les oportunitats que ofereix la transformació digital. El Govern ha identificat la necessitat de desenvolupar un enfocament nacional de la IA a Suècia. L'objectiu del pla és definir una orientació general per al treball relacionat amb la IA a Suècia i establir les bases per a les futures prioritats. L'objectiu del Govern és fer de Suècia un líder en l'aprofitament de les oportunitats que pot oferir l'ús de la IA, amb la finalitat d'enfortir el benestar i la competitivitat del país.

Suècia ha de desenvolupar la seva oferta a llarg termini de coneixement i experiència en el camp de la IA a fi d'aprofitar-ne els beneficis. La necessitat d'un coneixement rellevant de la IA s'ha de satisfer mitjançant l'educació i la formació, l'educació contínua i la recerca. També calen innovacions i iniciatives pel que fa als àmbits d'ús per a promoure projectes d'aplicació primerenca. Suècia ha de garantir l'accés a dades i infraestructures, com ara la potència de càlcul, a més dels marcs nacionals, europeus i internacionals adequats. Per tant, els agents públics haurien de donar suport activament a les aplicacions d'IA fent disponibles les dades rellevants i creant una infraestructura digital nacional, tenint en compte les qüestions de seguretat i integritat.

A la llum de la transformació social que la IA comporta, és important treballar per una política d'IA coherent i estratègica que tingui per objectiu crear un clima segur i favorable per a la digitalització i aprofitar les oportunitats que ofereix la IA. El desenvolupament i l'ús de la IA s'han de guiar per normes i principis ètics orientats a aprofitar-ne els beneficis, minimitzant alhora els riscos tant per a la societat com per a les persones. Aquesta no és només una qüestió per a investigadors i enginyers, sinó per a un ampli ventall de professions. Per tant, uns marcs adequats de principis, normes, estàndards i regles són un requisit previ important si Suècia vol adonar-se dels beneficis que la IA té per a la societat. Aquests marcs han de trobar l'equilibri entre les necessitats fonamentals de privacitat, d'ètica, de confiança i de protecció social i l'accés a les dades necessàries per a treure profit del potencial de la IA. Així mateix, són importants els marcs normatius a escala europea i internacional, com ara les normes de transferència de dades entre fronteres.

Diverses iniciatives del Govern per a promoure la capacitat de l'Administració pública per a utilitzar la IA

El Govern ha encarregat¹⁰⁸ a diversos organismes públics (de fiscalitat, mercat laboral, empreses o govern digital) que promoguin la capacitat de l'Administració pública per a utilitzar la IA. El Govern creu que és important que l'Administració pública desenvolupi més eines comunes per a utilitzar la IA. Durant l'en-

108. Decisió del Govern del 17 de juny de 2021, ref. 2021/01825

càrrec,¹⁰⁹ s'ha posat de manifest que dins l'Administració pública hi ha una gran demanda de suport integral i concret per a desenvolupar solucions d'IA i entorns tècnics segurs i elaborar guies en IA. També hi ha una gran necessitat per part de les agències públiques d'una governança i una coordinació clares a fi de preparar l'Administració pública per al proper reglament de la UE.¹¹⁰ A Suècia s'han fet grans inversions en recerca sobre IA en el sector privat, però manquen iniciatives semblants en el sector públic. Cal continuar treballant per a posar a disposició de tots els actors de l'Administració pública una infraestructura d'IA que permeti desenvolupar serveis impulsats amb aquesta tecnologia. Ha de ser possible per a tothom, d'una manera igualitària, a tot el país i en diferents sectors experimentar i desenvolupar solucions d'IA amb dades obertes o de propietat. Una infraestructura com aquesta comporta que més actors de l'Administració pública es puguin adonar del potencial de la IA amb un baix cost i en un entorn segur i provat.

Anteriorment, el Govern va encarregar¹¹¹ a diverses universitats importants que duguessin a terme una iniciativa de desenvolupament de competències en el camp de la IA, i recentment les autoritats d'educació superior han tingut l'encàrrec d'analitzar l'impacte de la IA en l'oferta de programes i cursos d'educació superior amb relació a la necessitats del mercat laboral. El Govern considera important que l'oferta de programes i cursos d'educació superior estigui en consonància amb les noves tecnologies a fi de satisfer les necessitats del mercat laboral.

109. Informe final: Encàrrec per a promoure la capacitat de l'Administració pública per a utilitzar la IA. Agència de Govern Digital, gener del 2023

110. Proposta de Reglament sobre normes harmonitzades sobre IA i pel qual es modifiquen determinats actes legislatius de la Unió (COM/2021/206 final)

111. Sol·licitud d'una iniciativa de desenvolupament de competències en el camp de la IA, ref. U2018/02719/UH

Suïssa

IA generativa i democràcia a Suïssa

Situació actual i problemes en joc

Com en altres països de l'EPTA, la IA generativa ha rebut una atenció creixent a Suïssa des del llançament de ChatGPT, on ràpidament va guanyar molts usuaris. En l'àmbit polític, els primers debats han estat intensos, però, fins a la data d'aquest informe, no s'han adoptat mesures importants. Al Parlament, diversos diputats han impulsat iniciatives (vegeu més avall). Un d'ells va fer que ChatGPT escrigués una interpellació al Consell Federal, com a experiment.¹¹² Atès que les eleccions federals suïsses seran a l'octubre del 2023, els partits polítics van iniciar una debat no oficial sobre un possible acord amb relació a un ús just de la IA durant la campanya.¹¹³ El juliol del 2023, un partit polític va dissenyar un cartell de campanya amb intel·ligència artificial, que va ser polèmic per les possibles conseqüències que el contingut creat amb IA podria tenir en aquest context (per exemple, si es tractava d'un tipus de notícia falsa o si, en realitat, no és res de nou).¹¹⁴

Suïssa té dos mecanismes de democràcia directa fortament arrelats, que són les iniciatives populars i els referèndums (facultatius o obligatoris).¹¹⁵ Es pot preveure que la IA generativa també tindrà efectes sobre la democràcia per mitjà d'aquests canals, que involucren en el procés de presa de decisions els ciutadans individuals i grups polítics, i també els partits i representants. Un estudi de TA-SWISS sobre hipertrucatges i contingut creat amb IA, que es publicarà properament, ha identificat diversos riscos per a la democràcia suïssa. Concretament, es podrien utilitzar com a mitjans de manipulació, desinformació i *astroturfing* (la simulació de moviments de base per mitjà de bots), o com a instruments per a desacreditar o assetjar els oponents. A més, podrien tenir finalitats desestabilitzadores, com ara agreujar les tensions socials o crear una desconfiança general en les fonts d'informació més comunes.¹¹⁶ Pel que fa a les oportunitats futures per a la democràcia, de moment són menys tangibles, però podrien incloure un accés més ampli a la informació política, gràcies a noves eines de cerca i resum, o la simplificació de textos complexos. La generació de textos també podria resultar útil per a persones o grups amb poca experiència política, poques habilitats d'escriptura o pocs mitjans financers, la qual cosa podria reduir certes desigualtats de recursos i eliminar barreres a la participació. No obstant això, aquests beneficis són polèmics per si mateixos, sobretot pel que fa a la qualitat dels resultats de la IA, i també el paper de les màquines i les empreses privades d'IA en la vida política.¹¹⁷

112. RTS, 17 de març de 2023 «Une intervention parlementaire rédigée par l'intelligence artificielle» [Una intervenció parlamentària escrita per intel·ligència artificial], a <https://www.rts.ch/info/suisse/13817951-le-parlement-a-definitivement-adopte-la-reforme-du-2e-pilier.html>. Vegeu la interpellació 23.3147 (<https://www.parlament.ch/en/ratsbetrieb/suche-curia-vista/geschaefte?AffairId=20233147>).

113. RTS, 22 de maig de 2023 (<https://www.rts.ch/info/suisse/14042381-une-charte-de-bonne-conduite-sur-lintelligence-artificielle-lidee-fait-son-chemin-dans-les-partis.html>).

114. *Le Temps*, 7 de juliol de 2023 (<https://www.letemps.ch/suisse/l-annonce-du-plr-et-la-question-de-l-intelligence-artificielle-en-politique>).

115. Per a una breu descripció d'aquests instruments, vegeu les entrades del lèxic de termes parlamentaris *iniciatives populars* i *referèndums*: <https://www.parlament.ch/en/über-das-parlament/parlamentswörterbuch>.

116. Vegeu Murat Karaboga, Nula Frei, Frank Ebbers, Greta Runge, Michael Friedewald, Daniel Vogler, Adrian Rauchfleisch, Manuel Puppis i Patrick Raemy, «Deepfakes und manipulierte Realitäten» [Hipertrucatges i realitats manipulades], TA-SWISS (publicació el 2024).

117. Per a un experiment suís, vegeu *Politbot*, un bot de conversa entrenat amb textos de partits polítics que respon a les consultes sobre les seves posicions amb l'ajuda de ChatGPT, que té per objectiu fer la política «més accessible» per a les eleccions del 2023 (<https://politbot.ch/>).

Una visió de les opinions dels diputats suïssos

TA-SWISS va encarregar una enquesta als diputats sobre hipertrucatges, que es va fer al febrer del 2023.¹¹⁸ Actualment, els autors de l'enquesta encara n'estan avaluant els resultats. Alguns resultats inicials suggereixen que els vint-i-tres enquestats consideren que els hipertrucatges i el contingut creat amb IA són un risc elevat i concret per a la democràcia i les institucions suïsses. Tot i que alguns consideren que aquests riscos són part del debat polític, altres no consideren que ho siguin (o no ho consideren en una mesura significativa). A més, els enquestats no han indicat que els percebin com una oportunitat per a la política.

Per a elaborar aquest informe, també vam demanar a quatre membres de la Comissió de Ciència, Educació i Recerca que emplenessin el qüestionari de l'EPTA sobre la IA generativa. Dos d'ells van respondre-hi. Les primeres coses que els venen al cap en llegir el terme són OpenAI, ChatGPT, creació de contingut i augment de productivitat. Tots dos veuen un impacte evident en la **democràcia**. També esmenten diversos riscos, com ara notícies falses més sofisticades en les campanyes, la manipulació de l'opinió, els atacs a la reputació i la pèrdua de la confiança en els càrrecs electes i les institucions. Tanmateix, segons un d'ells, si s'utilitza adequadament, la IA generativa també pot tenir un impacte positiu en la presa de decisions i generar nous coneixements. Pel que fa a la necessitat de regular la IA generativa en els sectors **educatiu i sanitari**, un diputat recomana un enfocament que tingui en compte la classificació dels riscos: l'autoregulació seria suficient per a aplicacions de baix risc, mentre que els usos amb més impacte, com les aplicacions en l'àmbit de la vigilància o les assegurances, els haurien d'examinar les comunitats interessades, i les aplicacions d'impacte alt, com el diagnòstic mèdic o els vehicles autònoms, també s'haurien de regular (amb coordinació internacional). Per a l'altre diputat, la IA generativa requereix que desenvolupem nous mètodes d'aprenentatge per a fomentar el pensament independent i repensar el mandat educatiu de l'Estat, i, en el sector sanitari, per a ajustar les mesures de protecció de la intimitat. Pel que fa als efectes de la IA generativa en l'**ocupació**, un diputat veu les oportunitats per a alliberar temps per a tasques més exigents i satisfactòries, i també per a compensar l'escassetat d'etiquetes. Tanmateix, la quantitat de llocs de treball afectats actualment estaria subestimada, especialment en el cas dels llocs de treball altament qualificats. En el context de la contractació de personal, un diputat veu que hi ha potencial per a trobar-ne de més adequat i que l'equivalència de titulacions jugaria un paper cada cop més important. Ara bé, les trobades cara a cara continuarien essent indispensables per a garantir l'autenticitat. L'altre diputat desconfia dels biaixos en el procés i considera necessari un marc responsable, fins i tot a escala internacional. Pel que fa als **debats al Parlament suís**, tots dos parlamentaris fan referència a iniciatives que s'han impulsat (vegeu més avall). Un d'ells considera que la IA, en general, és una part important de l'estratègia digital de Suïssa.¹¹⁹ Segons el parer de l'altre, una llei de la IA general per a Suïssa seria controvertida, mentre que l'enfocament «neutre respecte a la tecnologia» de Suïssa encara seria un mitjà adequat per a garantir les oportunitats i reaccionar d'una manera responsable als nous desenvolupaments. No obstant això, Suïssa quedaria enrere pel que fa a la infraestructura de dades, encara que el Parlament adoptes mesures en el futur.

Debats polítics i socials

Els debats sobre la IA generativa a Suïssa van començar fa uns quants anys, inicialment a un nivell més aviat teòric (per exemple, sobre si la IA serà capaç d'escriure una novel·la per si sola), i van augmentar significativament amb la difusió recent d'aplicacions concretes per al públic en general. El 2022, es van citar àmpliament generadors d'imatges com Midjourney, LensaAI o DALL-E, juntament amb qüestions relacionades amb la transformació de la comunicació, l'art i la creativitat.¹²⁰ Des del llançament de ChatGPT, la IA ge-

118. Vegeu Murat Karaboga, Nula Frei, Frank Ebberts, Greta Runge, Michael Friedewald, Daniel Vogler, Adrian Rauchfleisch, Manuel Puppis i Patrick Raemy, «Deepfakes und manipulierte Realitäten» [Hipertrucatges i realitats manipulades], TA-SWISS (publicació el 2024).

119. Per a l'Estratègia Suïssa Digital oficial, vegeu <https://digital.swiss/en/>.

120. Aquestes preguntes les abordarà l'estudi sobre cultura i digitalització de TA-SWISS (publicació el 2024, <https://www.ta-swiss.ch/en/culture-and-digitalisation>).

nerativa s'ha convertit en un dels temes més destacats dels mitjans de comunicació, les xarxes socials, els debats públics i la comunicació corporativa. Així, l'atenció inicial a ChatGPT es va estendre gradualment a la IA generativa, i fins i tot a la IA en general.

Aquesta dinàmica es reflecteix a nivell polític. Al Parlament, l'any 2023, s'han impulsat diverses iniciatives,¹²¹ moltes de les quals plantegen preguntes al Consell Federal, especialment pel que fa a les seves avaluacions dels riscos i les oportunitats de la IA generativa, i també les mesures planificades per a contenir els riscos, però, alhora, donar suport a la recerca i la innovació. També s'han demanat més estudis i més recerca, especialment pel que fa a la necessitat de regulació; les conseqüències de la IA generativa per a l'ocupació, l'educació i la ciberseguretat; els possibles danys a les persones; la posició de Suïssa respecte a la llei de la IA i de la comissió sobre IA del Consell d'Europa, i els ajustos a l'estratègia digital del país. La regulació és un tema central, amb la típica divisió entre les demandes d'una regulació més estricta, d'una banda, i la preocupació que això pugui obstaculitzar els beneficis de la innovació, de l'altra.¹²²

Pel que fa a la regulació de la IA, Suïssa oficialment ha seguit una estratègia sectorial i neutra quant a la tecnologia, en lloc d'establir disposicions globals sobre la IA. Mentre preparem aquest informe (a l'agost del 2023), el Consell Federal està a l'espera que finalitzi l'elaboració de la llei de la IA de la UE per a avaluar-ne l'impacte a Suïssa, i també la necessitat de fer ajustaments legislatius. L'abril del 2023, el Consell Federal va encarregar a l'Oficina Federal de Comunicacions l'elaboració d'un projecte de llei sobre la regulació de les plataformes de comunicació.¹²³ Les mesures que el Consell Federal havia adoptat anteriorment inclouen les directrius sobre intel·ligència artificial per a la Confederació i¹²⁴ la creació de la Xarxa de Competències per a la Intel·ligència Artificial (CNAI) el 2022,¹²⁵ i també la creació de diversos grups de treball a tota l'Administració federal i entre les parts interessades. En l'escena internacional, Suïssa contribueix activament a les iniciatives de la IA de l'OCDE, el Consell d'Europa, la UNESCO i la Unió Internacional de Telecomunicacions.¹²⁶

Atès que el debat sobre la IA generativa es troba en la fase preliminar, sembla massa aviat per a avaluar la implicació de la ciència i el públic en el procés de la presa de decisions. Molt sovint, els mitjans de comunicació consulten científics de diverses disciplines. Al Parlament, la Comissió de Ciència, Educació i Investigació del Consell Nacional va organitzar una audiència científica sobre la IA generativa a l'agost del 2023, a la qual va convidar TA-SWISS. Seria difícil poder fer un seguiment sistemàtic dels altres intercanvis entre la ciència i la política sobre aquesta qüestió. Pel que fa a l'àmbit públic, moltes entitats de la societat civil, del sector informàtic i de l'economia han estat molt actives a l'hora d'adoptar una posició sobre aquesta qüestió (oportunitats i riscos, mesures per adoptar). Atesa la manca d'enquestes exhaustives, per ara és difícil avaluar l'actitud dels ciutadans envers aquesta qüestió. Actualment s'està duent a terme una enquesta sobre l'actitud de la població envers els hipertrucatges i el contingut creat amb IA, per encàrrec de TA-SWISS.¹²⁷

121. Vegeu, concretament, els elements 22.1074, 23.3147, 23.3201, 23.3516, 23.3583, 23.3644, 23.7270, 23.3806, 23.3806, 23.3806, 24.923, 82.32, 82.3861, 23.3930, 23.438 i 23.7049 en la base de dades pública del Parlament suís <https://www.parlament.ch/en>.

122. Vegeu, per exemple, el debat a *Inside IT* entre els diputats F. Grüter, «Zu viel Regulierung bei KI bremst die Innovation» [Massa regulació en IA frena la innovació] (<https://www.inside-it.ch/parldigi-direkt-zu-viel-regulati-on-bei-ki-bremst-die-innovation-20230419>) i M. L. Marti, «KI-Regulierung muss Innovation nicht behindern» [La regulació de la IA no ha d'obstaculitzar la innovació] (<https://www.inside-it.ch/parldigi-direkt-ki-regulation-mus-s-innovation-nicht-behindern-20230510>).

123. Vegeu la nota de premsa de l'OFCOM del 5 d'abril de 2023 (<https://www.bakom.admin.ch/bakom/en/homepage/ofcom/ofcom-s-information/press-releases/nsb.msg-id-94116.html>).

124. Disponible a <https://www.sbf.admin.ch/sbf/en/home/eri-policy/eri-21-24/cross-cutting-themes/digitalisation-eri/artificial-intelligence.html>.

125. <https://cnaai.swiss/en/>.

126. Vegeu la declaració del Consell Federal, del 26 d'abril de 2023, sobre la iniciativa 23.3201 (<https://www.parlament.ch/en/ratsbetrieb/suche-curia-vista/geschaefte?AffairId=20233201>).

127. Vegeu Murat Karaboga, Nula Frei, Frank Ebberts, Greta Runge, Michael Friedewald, Daniel Vogler, Adrian Rauchfleisch, Manuel Puppis i Patrick Raemy, «Deepfakes und manipulierte Realitäten» [Hipertrucatges i realitats manipulades], TA-SWISS (publicació el 2024).

El paper de l'AT en els debats

TA-SWISS ha posat en marxa diversos projectes relacionats amb la IA generativa. Com s'ha esmentat anteriorment, el 2024 es publicarà un estudi complet sobre hipertrucatges i contingut creat amb IA, amb recomanacions per als polítics, les parts interessades i el públic sobre la base d'una visió general dels riscos i les oportunitats, a més d'una anàlisi tècnica i jurídica i una enquesta sobre les percepcions dels ciutadans.¹²⁸ Paral·lelament, a l'abril del 2023, TA-SWISS va publicar un article breu sobre ChatGPT i models de llenguatge extens, que presenta el funcionament i les principals qüestions que plantegen aquestes aplicacions.¹²⁹ El 2024 es podria iniciar un estudi més extens dels models lingüístics. Els estudis previs d'interès per a la relació entre democràcia i IA inclouen un estudi general sobre la IA (2020),¹³⁰ i també estudis sobre els efectes de la digitalització en la democràcia (2021)¹³¹ i el reconeixement facial i de la parla (2022).¹³² A més, l'impacte de la IA generativa en l'art i la cultura també s'està debatent al nostre projecte en curs sobre cultura i digitalització.¹³³

Pel que fa al paper de TA-SWISS com a assessor dels actors polítics, els dos diputats que van respondre al qüestionari de l'EPTA van expressar la necessitat de més informació, concretament, sobre les qüestions següents: l'evolució de les pràctiques a Suïssa (especialment en empreses, institucions i administracions), els aspectes jurídics i tècnics i els resultats de la recerca en curs (per exemple, sobre mitjans de control, traçabilitat de dades o transparència dels algorismes).

D'una manera més general, des del llançament de ChatGPT TA-SWISS sovint rep sol·licituds per a fer contribucions sobre la IA generativa en esdeveniments públics i acadèmics, als mitjans de comunicació i al Parlament (a la Comissió de Ciència, Educació i Recerca del Consell Nacional). Això posa de manifest el valor afegit de la perspectiva de l'AT sobre aquesta qüestió. Atesa la gran diferència entre la velocitat inusual d'aquestes tecnologies i el temps necessari per a adoptar possibles mesures contra els seus riscos, sembla crucial que l'AT pugui reaccionar d'una manera responsable i proactiva. Tenint en compte les sol·licituds que hem rebut, també sembla imprescindible un enfocament interdisciplinari, atès que les preguntes es refereixen al funcionament tècnic d'aquestes tecnologies, al marc normatiu, als seus efectes en la societat i l'economia i a qüestions ètiques. En últim lloc, però no per això menys important, l'AT ha de contribuir d'una manera decisiva a la legitimació democràtica de l'ús i la regulació de la IA generativa. Per mitjà de proporcionar informació independent que tinguin en compte les múltiples facetes de l'impacte de la IA generativa en la societat, l'AT pot ajudar els qui prenen decisions (tant ciutadans com actors polítics) a formar-se una opinió i a prendre decisions informades sobre aquests desenvolupaments tecnològics.

128. Vegeu Murat Karaboga, Nula Frei, Frank Ebbers, Greta Runge, Michael Friedewald, Daniel Vogler, Adrian Rauchfleisch, Manuel Puppis i Patrick Raemy, «Deepfakes und manipulierte Realitäten» [Hipertrucatges i realitats manipulades], TA-SWISS (publicació el 2024).

129. Versió francesa: <https://www.ta-swiss.ch/fr/chatgpt>; versió alemanya: <https://www.ta-swiss.ch/chatgpt>.

130. Markus Christen, Clemens Mader, Johann Cäs, Tarik Abou-Chadi, Abraham Bernstein, Nadja Braun Binder, Daniele Dell'Aglio, Luca Fábian, Damian George, Anita Gohdes, Lorenz Hilty, Markus Kneer, Jaro Krieger-Lamina, Hauke Licht, Anne Scherer, Claudia Som, Pascal Sutter i Florent Thouvenin, «Wenn Algorithmen für uns entscheiden: Chancen und Risiken der künstlichen Intelligenz» [Quan els algorismes decideixen per nosaltres: oportunitats i riscos de la intel·ligència artificial], TA-SWISS (2020, <https://www.ta-swiss.ch/en/artificial-intelligence>).

131. Urs Bieri, Edward Weber, Nadja Braun Binder, Sébastien Salerno, Tobias Keller i Manuela Kälin, «Digitalisierung der Schweizer Demokratie – Technologische Revolution trifft auf traditionelles Meinungsbildungssystem» [La digitalització de la democràcia suïssa. La revolució tecnològica es troba amb el sistema tradicional de formació de l'opinió]; Nora Räss, Ira Differding i Jasmin Odermatt, «Jugend, politische Partizipation und Digitalisierung. Eine Analyse der digitalen politischen Partizipation junger Menschen in der Schweiz» [Joventut, participació política i digitalització. Anàlisi de la participació política digital dels joves a Suïssa]; Anna Boos, Ramona Sprenger, Jeannie Schneider, Basil Rogger, René Odermatt, David Simon, «Szenarien zu Demokratie und Digitalisierung. Ein partizipatives Zukunftsexperiment für die Schweiz» [Escenaris sobre democràcia i digitalització. Un futur experiment participatiu per a Suïssa], TA-SWISS (2021, <https://www.ta-swiss.ch/digitale-demokratie>).

132. Murat Karaboga, Nula Frei, Frank Ebbers, Sophia Rovelli, Michael Friedewald i Greta Runge, «Automatisierte Erkennung von Stimme, Sprache und Gesicht. Technische, rechtliche und gesellschaftliche Herausforderungen» [Reconeixement automàtic de la veu, l'idioma i la cara. Reptes tècnics, legals i socials], TA-SWISS (2022, <https://www.ta-swiss.ch/en/speech-speaker-facial-recognition>).

133. Stefano Kunz, Jens Meissner i Ramona Sprenger, «Kunst & Digitalisierung in Musik, Theater und visuellem Design» [Art i digitalització en la música, el teatre i el disseny visual], TA-SWISS (publicació el 2024, títol provisional, <https://www.ta-swiss.ch/en/culture-and-digitalisation>).

Regne Unit

Informe EPTA 2023 - Oficina Parlamentària de Ciència i Tecnologia

Oliver Bennett MBE, cap de l'Oficina Parlamentària de Ciència i Tecnologia

Devyani Gajjar, assessor en ciències físiques i l'àmbit digital

La IA generativa i el mercat laboral del Regne Unit i un escrutini parlamentari més ampli de la IA

Quin és l'estat actual del vostre país (desenvolupament, desplegament, ús)?

Demanda de coneixements especialitzats sobre tractament de dades

Per a tota la mà d'obra del Regne Unit hi ha una demanda creixent de coneixements especialitzats sobre tractament de dades, inclosos els coneixements sobre intel·ligència artificial, que tenen el potencial d'aportar beneficis econòmics i socials.

L'evidència suggereix que la disponibilitat de persones amb coneixements especialitzats sobre tractament de dades al Regne Unit és insuficient per a satisfer la demanda. D'acord amb un estudi del 2021, al Regne Unit era poc probable que l'oferta de científics de dades de les universitats superés els 10.000 l'any, mentre que potencialment hi havia com a mínim 178.000 posicions d'especialistes en dades sense ocupar.

La recerca mostra que determinats grups (com ara les dones, les persones d'origen ètnic minoritari i les persones amb discapacitat) tenen poca representació en la mà d'obra especialitzada en el tractament de dades. La manca de diversitat en la mà d'obra potencialment pot fer augmentar les desigualtats i els prejudicis existents.

Els esforços per a reduir el dèficit de coneixements es poden veure obstaculitzats per la definició incoherent dels coneixements sobre tractament de dades, la cultura organitzativa, la disponibilitat de professors especialitzats de primària i secundària i les barreres per a la mobilitat de les persones entre sectors.

Les iniciatives per a augmentar el nombre de persones amb coneixements sobre tractament de dades inclouen els cursos pont, els centres de formació per a estudiants de doctorat, les plataformes digitals de millorament de competències, els aprenentatges i els visats per a atraure talent internacional.

Desplegament i ús de la IA

L'Artificial intelligence sector study 2022 [Estudi del sector de la intel·ligència artificial 2022], publicat per l'Oficina d'Intel·ligència Artificial i el Ministeri de Ciència, Innovació i Tecnologia va posar de manifest que hi havia 3.170 empreses d'IA treballant al Regne Unit, que van tenir uns ingressos de 10.600 milions de lliures esterlines relacionats amb la intel·ligència artificial. D'aquestes empreses, el 60% són empreses dedicades a la IA i el 40% duen a terme activitats relacionades amb la IA com a part d'una oferta de productes o serveis més àmplia. L'estudi calcula que el sector ocupa 50.040 equivalents a jornada completa en llocs de treball relacionats amb la IA, el 53% dels quals són en empreses dedicades a la IA.

Desenvolupament de models de fundació (subjacents als sistemes d'IA generativa recents)

Els avenços recents en sistemes d'IA generativa que es poden utilitzar amb finalitats generals (com ara ChatGPT) es poden etiquetar com a models de fundació. Aquests models de fundació es desenvolupen utilitzant grans quantitats de dades i recursos informàtics. Actualment, el Regne Unit no té capacitats per a desenvolupar models de fundació sobiranament. L'abril del 2023, el primer ministre va anunciar 100 milions de lliures esterlines de finançament per a un grup de treball del model de fundació per a «garantir les capacitats sobiranes i l'adopció àmplia de models de fundació segurs i fiables».

Impacte potencial de la IA generativa al mercat laboral del Regne Unit

Un informe del juny del 2023 de KPMG sobre la IA generativa i el mercat laboral del Regne Unit va identificar que el 2,5% de totes les tasques les podria dur a terme la IA, i que això podria tenir un impacte en el 40% dels llocs de treball del Regne Unit en la propera dècada. L'informe prediu que la meitat de l'impacte de la substitució de la IA generativa es podria compensar amb la creació de noves tasques en els llocs de treball afectats. L'informe fa una previsió de les tres principals aplicacions de la IA generativa al mercat laboral: les tasques de classificació i resum (com la cerca en els documents), la creació de contingut tècnic (la codificació) i els treballs subjectius (com el dibuix, la presentació i el contingut de màrqueting).

Què en pensen els vostres diputats?

Hi ha hagut un interès important al Parlament del Regne Unit, amb diversos debats parlamentaris i informes de comissions. El debat ha inclòs l'anàlisi de què s'ha de fer per a garantir que el Regne Unit es pugui beneficiar de la nova tecnologia d'IA i, alhora, gestionar les possibles conseqüències negatives per als treballadors.

Una investigació del 2022 de la Comissió de Ciència i Tecnologia de la Cambra dels Lords va posar de manifest que hi havia un desajust entre l'escala del dèficit de competències en ciència, tecnologia, enginyeria i matemàtiques (CTEM) del Regne Unit i les solucions plantejades pel Govern. La Comissió va concloure que:

«Hi ha un desajust entre l'escala del dèficit de competències del Regne Unit i les solucions proposades pel Govern, sobretot tenint en compte l'ambició del Regne Unit per ser una superpotència científica i tecnològica. Les polítiques del Govern són inadequades i fragmentàries. Eliminar el dèficit requereix alinear les prioritats d'alt nivell d'una sèrie de ministeris. La coordinació és crucial perquè la disponibilitat de mà d'obra qualificada no impedeixi el creixement de les indústries CTEM».

La Comissió va identificar diverses actuacions que el Govern ha de dur a terme, agrupades en quatre àmbits:

- Política d'immigració per al talent CTEM. Per exemple, acords de visats per a atraure i retenir talent internacional.
- Quantificar i abordar el dèficit de competències nacional. Per exemple, oferir cursos per sota del nivell de grau per a permetre que els treballadors es reciclin o desenvolupin noves competències CTEM, publicar una avaluació del dèficit de competències que especifiqui la manera en què les polítiques i iniciatives poden contribuir a resoldre el dèficit, o millorar els programes d'aprenentatge.
- Contractar i retenir professors i educadors de ciència. Per exemple, oferir un sou més alt als professors especialistes, millorar la retenció dels professors o donar suport a professionals qualificats perquè esdevinguin professors.
- La precarietat i l'atractiu de les carreres acadèmiques en l'àmbit CTEM. Per exemple, l'assessorament professional per a doctors i investigadors postdoctorals, els programes de beques per a permetre que els doctors i postdoctorats passin temps a la indústria, el millorament de les condicions de treball i, en particular, la creació de més beques de recerca d'una durada més llarga.

El 2023, [un informe de la Comissió d'Empresa, Energia i Estratègia Industrial](#) va destacar l'impacte que la IA podria tenir en la productivitat al Regne Unit. Es va referir a una recerca de Deloitte que exposava que «fins al 2035, la IA podria augmentar la productivitat del mercat laboral del Regne Unit en un 25%» i que «quatre de cada cinc organitzacions del Regne Unit van dir que l'ús d'eines d'IA havia fet que els seus empleats fossin més productius, havia millorat la presa de decisions i havia fet els seus processos més eficients». També va assenyalar que la IA i les tecnologies que hi estan relacionades poden tenir un impacte positiu a l'hora d'ajudar a accedir al mercat laboral les persones que, si no, tindrien dificultats per a trobar feina i mantenir-la, com ara les persones amb discapacitat.

La Comissió va fer diverses recomanacions, que incloïen les següents:

- «La introducció d'un requisit perquè les empreses duguin a terme avaluacions d'impacte per a entendre l'abast i les conseqüències de l'ús de les noves tecnologies en el lloc de treball» i «que es garanteixi el dret de consulta i notificació als treballadors si l'aplicació de la tecnologia al lloc de treball comporta la vigilància d'un treballador o un canvi significatiu de la seva feina».
- La introducció d'una «nova estratègia de pobresa de dades i competències digitals que, entre altres coses, estableixi com s'ha de donar suport als treballadors en el desenvolupament de les competències digitals».
- Treballar per a abordar «la manca d'inversió en formació i recapacitació al lloc de treball».
- La necessitat que els reguladors contractin personal addicional amb expertesa tècnica.
- Que el Govern «faci consultes sobre un codi de pràctiques sobre l'ús de la tecnologia de vigilància al lloc de treball que es pugui aplicar».
- Que el Govern creï un grup de treball «per a avaluar les implicacions de la tecnologia en el lloc de treball, per a valorar si l'aplicació de les lleis en matèria laboral és efectiva i per a fer recomanacions sobre si cal legislació addicional».

Hi ha legislació vigent?

Actualment, la IA està regulada per mitjà dels marcs jurídics existents.

El Govern va publicar la seva posició sobre la regulació de la IA el març del 2023: [A pro-innovation approach to AI regulation \[Un enfocament a favor de la innovació en la regulació de la IA\]](#). L'informe afirmava que el seu enfocament amb relació a la IA era:

- A favor de la innovació: permetre en lloc de sufocar la innovació responsable.
- Proporcionat: evitar càrregues innecessàries o desproporcionades per a les empreses i els reguladors.
- Fiable: abordar els riscos reals i fomentar la confiança pública en la IA per a promoure'n l'adopció.
- Adaptable: permetre adaptar-nos d'una manera ràpida i eficaç per a seguir el ritme de les oportunitats i riscos emergents a mesura que evolucionen les tecnologies d'IA.
- Clar: facilitar als actors del cicle de vida de la IA, incloses les empreses que la utilitzen, saber quines són les regles, a qui s'apliquen, qui les fa complir i com les han de complir.
- Col·laboratiu: fomentar que el Govern, els reguladors i la indústria treballin junts per a facilitar la innovació en IA, generar confiança i garantir que la veu del públic s'escolti i es tingui en compte.

El Govern va assenyalar que, si s'identifiqués un buit en el sistema normatiu, la legislació vigent s'adaptaria. No obstant això, va quedar clar que «en un principi, no pretenem introduir legislació nova» perquè «si ens precipitéssim a legislar massa aviat, ens arriscaríem a imposar càrregues indegudes a les empreses».

El primer ministre ha anunciat que vol fer del Regne Unit «la llar geogràfica de la regulació global en matèria de la seguretat de la IA» i té previst d'organitzar una cimera mundial sobre seguretat de la IA al novembre.

Paper de l'AT en el debat parlamentari

L'Oficina Parlamentària de Ciència i Tecnologia (POST) ha donat un suport considerable a l'examen de la IA que ha dut a terme el Parlament del Regne Unit.

La POST ha elaborat publicacions encarregades per altres membres del personal parlamentari per a donar suport a les consultes i debats parlamentaris, com ara:

- [Data Science Skills in the UK Workforce](#) [Competències en ciència de dades en la mà d'obra del Regne Unit]
- [Digital technology in freight](#) [Tecnologia digital en el transport de mercaderies]
- [Automation in military operations](#) [Automatització en les operacions militars]
- [How is artificial intelligence governed in Australia, France, Italy and Singapore?](#) [Com es controla la intel·ligència artificial a Austràlia, França, Itàlia i Singapur?]

La POST també ha donat suport *ad hoc* a les comissions i biblioteques en l'examen de la IA que han dut a terme, com ara la revisió de notes i informes, la proposta de temes de recerca, l'assessorament tècnic i la proposta de compareixents.

El treball de la POST s'ha citat en informes parlamentaris rellevants recents, que inclouen:

- Comissió de Cultura, Mitjans de Comunicació i Esports de la Cambra dels Comuns, [Connected tech: smart or sinister? \(HC 157\)](#) [Tecnologia connectada: intel·ligent o sinistra? (HC 157)], 7 d'agost de 2023
- Comissió de Cultura, Mitjans de Comunicació i Esports de la Cambra dels Comuns, [Connected tech: AI and creative technology \(HC 1643\)](#) [Tecnologia connectada: IA i tecnologia creativa (HC 1643)], 30 d'agost del 2023
- Resum de recerca de la Biblioteca de la Cambra dels Comuns, [Artificial intelligence and employment law](#) [Intel·ligència artificial i dret laboral], 11 d'agost del 2023
- [Potential impact of artificial intelligence on the labour market - House of Commons Library](#) [Impacte potencial de la intel·ligència artificial en el mercat laboral - Biblioteca de la Cambra dels Comuns]
- [Debate on Artificial Intelligence – House of Commons Library](#) [Debat sobre la intel·ligència artificial – Biblioteca de la Cambra dels Comuns]

Actualment, la POST prepara més treballs sobre IA, que inclouen:

- [Artificial Intelligence \(AI\) Technologies](#) [Tecnologies d'intel·ligència artificial (IA)]
- [Artificial Intelligence in Weapon Systems](#) [Intel·ligència artificial en sistemes d'armes]

