

CONCLUSIONS. MECANISMES POLÍTICS, REPTES I RECOMANACIONS EN LA IA URBANA

Marta Galceran-Vercher

Investigadora sènior, Programa de Ciutats Globals, CIDOB

Alexandra Vidal D'oleo

Investigadora i gestora de projectes, Programa de Ciutats Globals, CIDOB

Es governs locals d'arreu del món van adoptant cada cop més els sistemes algorítmics per millorar la prestació dels serveis públics. No obstant això, l'evidència creixent assenyalava que aquests sistemes poden provocar danys no intencionats i ser implementats de manera poc transparent. En conseqüència, l'adopció de sistemes algorítmics sovint ha anat acompanyada del desenvolupament de principis rectors per a l'ús responsable de les tecnologies d'intel·ligència artificial (IA), principalment a escala nacional, supranacional o global. En són exemples notables els Principis d'IA de l'OCDE (2019), els Principis d'IA del G20 (2019), el grup de redacció del Conveni d'IA del Consell d'Europa (2022-2024), la declaració ministerial de l'Associació Global sobre Intel·ligència Artificial (2022), la declaració dels ministres del G7 (2023), la Declaració de Bletchley (2023), la Declaració Ministerial de Seül (2024), la Llei d'IA de la UE (2024) o l'Informe de l'ONU «Governing AI for Humanity» (2024). Tanmateix, aquests marcs normalment només proporcionen orientacions generals sobre el que constitueix un ús responsable de la IA i ofereixen poca orientació pràctica sobre com s'han d'aplicar aquests principis en contextos reals.

En resposta a aquests reptes, molts governs recorren a marcs normatius i eines polítiques per posar en pràctica aquests principis. Iniciatives com aquestes van sorgint amb molta rapidesa, tot i que varien considerablement quant a abast i enfocament. A més, gran part de l'anàlisi existent de les eines de polítiques del sector públic se sol centrar en perspectives d'abast nacional (per exemple, OECD, 2024), que sovint passen per alt el context i els reptes singulars als quals s'enfronten els governs locals.

Aquesta monografia CIDOB ha intentat superar aquesta limitació identificant els principals mecanismes polítics i marcs de governança que els governs locals aprofiten per garantir que adopten els sistemes algorítmics en consonància amb principis ètics bàsics, com ara la transparència i la rendició de comptes, l'equitat i la no-discriminació, la privacitat i la governança de dades, i la sostenibilitat. Aquesta anàlisi es complementa amb una sèrie d'estudis de casos que il·lustren com les principals ciutats implementen aquests mecanismes polítics a la pràctica, cosa que dona lloc a marcs locals integrals d'IA.

Aquesta monografia CIDOB identifica els principals mecanismes polítics i marcs de governança que els governs locals aprofiten per garantir que adopten sistemes algorítmics en consonància amb principis ètics fundamental.

Els criteris utilitzats per establir aquesta categorització es van centrar en la funció i els objectius principals dels mecanismes polítics.

En aquest capítol final, presentem una categorització dels mecanismes polítics identificats en aquesta publicació, juntament amb informació sobre quins mecanismes utilitzen més els governs locals i com concorden amb els principis ètics esmentats. També analitzem els reptes als quals s'enfronten els governs locals i oferim recomanacions per avançar cap a un ús més responsable dels sistemes d'IA en entorns urbans.

1. Categorització dels mecanismes polítics

Gràcies a una revisió exhaustiva dels mecanismes polítics presentats en els capítols anteriors d'aquesta monografia CIDOB, complementada amb una revisió bibliogràfica de les publicacions rellevants sobre mecanismes polítics per a governs locals i/o administracions públiques (com ara els informes de l'Ada Lovelace Institute, l'AI Now Institute i l'Open Government Partnership, 2021; Ben Dhaou *et al.*, 2024; Jordan *et al.*, 2024; United 4 Smart Sustainable Cities, 2024), hem identificat catorze categories diferents de polítiques que actualment implementen els governs locals arreu del món (vegeu la taula 1).¹

Els criteris utilitzats per establir aquesta categorització es van centrar en la funció i els objectius principals dels mecanismes polítics. Això inclou proporcionar orientació normativa per al desenvolupament i l'ús de sistemes d'IA, avaluar els riscos potencials dels algoritmes, garantir l'accés públic a la informació sobre els sistemes algorítmics i exigir rendició de comptes.

Altres possibles criteris s'haurien pogut centrar en si els mecanismes s'orienten cap a processos administratius interns (per exemple, directrius per al personal municipal o creació de comissions municipals d'IA) o accions externes, com la publicació de registres algorítmics públics o la imposició de prohibicions sobre aplicacions d'IA controvertides. A més, la nostra categorització no distingeix entre els mecanismes específics d'IA (per exemple, una estratègia d'IA) i els mecanismes indirectes que contribueixen a la governança ètica de la IA (per exemple, la importància de les dades en els sistemes d'IA).

1. Cal tenir en compte que la literatura sobre aquest tema és limitada i utilitza una sèrie de noms i termes per als diferents mecanismes polítics, però no hi ha un vocabulari comú per als seus components bàsics.

Taula 1. Categories de mecanismes polítics per a una governança responsable de la IA a les ciutats

CATEGORIA	DESCRIPCIÓ I FINALITAT	MECANISMES POLÍTICS
[1] Principis, estratègies i directrius	Documents polítics que ofereixen orientacions normatives no vinculants sobre els valors i els principis ètics a les administracions locals, i que presenten indicacions generals per desenvolupar i utilitzar la IA i alhora gestionar-ne els riscos associats. Tot i que el format varia, aquests documents solen identificar objectius polítics d'alt nivell i la seva rellevància per a l'ús dels sistemes algorítmics per part dels organismes públics. En alguns casos, també proporcionen orientacions pràctiques per implementar aquests principis a l'hora de dissenyar i desplegar aquests sistemes. En darrera instància, aquestes directrius estableixen estàndards normatius que permeten als organismes i a la ciutadania avaluar l'ús ètic dels sistemes algorítmics.	<ul style="list-style-type: none"> Plans d'acció, agendes, fulls de ruta, estatuts, manuals, estratègies d'IA ètica, etc. Definició de principis locals ètics per a la IA: declaració i/o aprovació. Directrius, guies i manuals sobre el desplegament de la IA ètica. Protocols interns per a l'ús de la IA. Marcos estratègics que no són específicament sobre IA que tenen un impacte en la governança de la IA (per exemple, marcs de governança de les dades o digitalització)
[2] Normativa i legislació local	Eines destinades a establir estàndards, lleis i regulacions que garanteixin el compliment i abordin els impactes socials.	<ul style="list-style-type: none"> Regulació i lleis locals (per exemple, normatives per garantir el dret a la justificació) Mecanismes de compliment legal: per garantir el compliment de marcs normatius regionals, nacionals o supranacionals Estàndards reguladors (per exemple, estàndards d'IA verda, estàndards de transparència) Adhesió als estàndards normatius internacionals Regulació de l'aplicació d'IA controvertida: prohibicions, moratòries, etc.
[3] Mecanismes de transparència i explicabilitat	Mecanismes per establir l'accés públic a la informació sobre sistemes i processos algorítmics. Tenen l'objectiu de proporcionar informació sobre sistemes algorítmics al públic en general (per exemple, les persones afectades, els mitjans de comunicació o la societat civil) perquè pugui conèixer l'ús d'aquests sistemes i exigir explicacions i justificacions relacionades amb aquest ús. Aquests mecanismes poden funcionar de manera independent o formar part de marcs més amplis per a la rendició de comptes algorítmica. Cal distingir els mecanismes de transparència pública dels drets d'audiència i d'explicació, que atorguen a les persones el dret a una explicació sobre les decisions algorítmiques concretes que s'han pres sobre elles.	<ul style="list-style-type: none"> Registres públics d'algorítmics Web municipal amb tota la informació pertinent sobre la IA Estàndards de transparència algorítmica Directori municipal d'eines d'IA adquirides per a ús intern Requisits per a la transparència del codi font Explicacions de lògiques algorítmiques
[4] Avaluació de l'impacte algorítmic	Instrument polític utilitzat pels organismes públics per avaluar els possibles riscos i danys dels sistemes algorítmics. Aquestes avaluacions tenen com a objectiu entendre, categoritzar i abordar els possibles efectes negatius dels algorítmics abans de ser desplegats o durant el procés de desplegament. Les avaluacions d'impacte algorítmic (AIA) es basen en marcs establerts d'altres camps, com ara les avaluacions d'impacte mediambiental i en matèria de drets humans i les avaluacions d'impacte de protecció de dades (AIPD).	<ul style="list-style-type: none"> Procediments d'avaluació i gestió del risc (inclouent-hi l'anàlisi de biaixos) Avaluació d'impacte en matèria de drets humans Mecanismes d'avaluació d'impacte mediambiental
[5] Auditories i inspecció reglamentària	Les auditories inclouen una sèrie de pràctiques adreçades a examinar com funciona un sistema algorítmic concret. Tenen l'objectiu d'entendre el funcionament del sistema i avaluar-ne el rendiment en funció d'uns estàndards normatius predefinits. Tot i que les auditories presenten algunes semblances amb les AIA, el context històric és diferent i normalment es duen a terme durant la implementació del sistema o després de la implementació. En canvi, les AIA normalment es duen a terme abans del desplegament o durant el desplegament. Les auditories les poden dur a terme entitats internes, externes o de tercers, en funció de l'abast i la naturalesa de l'avaluació. En una auditoria de tercers, una organització externa avalua el sistema basant-se únicament en els resultats del sistema. Una auditoria de segona part la du a terme un avaluador extern que té accés tant al fons (<i>backend</i>) del sistema com als seus resultats. Les auditories internes les duen a terme membres interns de l'organització.	<ul style="list-style-type: none"> Auditories de sistemes algorítmics Avaluació del procés Eines per a l'autoavaluació

[6] Supervisió humana, rendició de comptes, procediments d'audiència i apel·lació	Mecanismes per supervisar els sistemes d'IA i fer que siguin responsables. Concretament, es tracta de mecanismes que requereixen que les decisions preses amb l'ajuda de sistemes algorítmics segueixin procediments específics dissenyats per salvaguardar l'equitat i proporcionar vies perquè les persones cerquin reparacions en casos de resultats esbiaixats o erronis. Aquestes garanties procedimentals ofereixen a les persones o grups afectats la possibilitat de qüestionar o impugnar les decisions que els afecten.	<ul style="list-style-type: none"> • Elaboració d'informes i supervisió interna • Requisits de la intervenció humana (<i>human-in-the-loop</i>) • Procediments de resposta i objeccions per a la ciutadania • Deures de notificació de la resolució i audiència dels afectats • Obligacions d'aportar decisions motivades i explicacions d'una decisió • Mecanismes per garantir el dret de les parts afectades a presentar proves, apel·lar i impugnar decisions automatitzades
[7] Condicions d'adquisició pública	Les normes que regeixen l'adquisició de sistemes algorítmics per part dels governs i els organismes públics són fonamentals per garantir-ne l'ús responsable. Molts sistemes algorítmics utilitzats pels governs són subcontractats a proveïdors privats, ja sigui mitjançant la compra de productes o els contractes de servei. En conseqüència, els proveïdors tenen un paper important a l'hora de dissenyar i desplegar aquests sistemes. Les condicions dels contractes d'adquisició són fonamentals per configurar el desenvolupament i la implementació d'aquests sistemes. Es poden aplicar condicions específiques d'adquisició, com ara requisits de transparència i no-discriminació, per garantir que els sistemes adquirits compleixen els estàndards ètics i s'utilitzen de manera responsable.	<ul style="list-style-type: none"> • Clàusules d'adquisició pública • Directrius, marcs i protocols interns per a l'adquisició d'IA
[8] Òrgans assessors i supervisors externs/independents	Òrgans supervisors independents, que tenen la finalitat de supervisar i guiar l'ús de sistemes algorítmics per part dels organismes públics. Aquests mecanismes de supervisió independent tenen com a objectiu garantir la rendició de comptes mitjançant el seguiment de les actuacions dels organismes públics i la formulació de recomanacions, sancions o decisions sobre l'ús que fan dels sistemes algorítmics.	<ul style="list-style-type: none"> • Òrgans assessors: consells, comitès, juntes, xarxes, grups, etc.
[9] Aliances, comunitats de pràctiques i grups d'aprenentatge	Mecanismes adreçats a fomentar la cooperació i les associacions a escala local, nacional i internacional.	<ul style="list-style-type: none"> • Comunitats de pràctiques d'aprenentatge locals/nacionals/internacionals: xarxes de ciutats, grups de treball, etc. • Col·laboracions locals/nacionals/internacionals multiactors en l'àmbit de la IA: xarxes, plataformes, coalicions, etc. • Associació i col·laboracions publicoprivades
[10] Programes de capacitat	Mecanismes per millorar el coneixement i adquirir habilitats sobre la IA ètica. Aquestes iniciatives poden adreçar-se al personal municipal implicat, directament o indirectament, en el disseny, el desplegament o l'ús de sistemes algorítmics, així com al públic en general, per promoure una ciutadania informada i fomentar una comprensió més àmplia de l'ètica de la IA.	<ul style="list-style-type: none"> • Formació del personal municipal (enfocament sociotecnològic) • Creació d'un equip municipal d'IA • Òrgan municipal d'IA: òrgan o comissió transversal que coordina/supervisa l'ús municipal de la IA • Enfocament multidisciplinari: creació d'equips diversos
[11] Foment de la innovació, el coneixement i l'experimentació local	Mecanismes que ofereixen un espai per a l'experimentació, la innovació i les proves en entorns del món real.	<ul style="list-style-type: none"> • Promoció i col·laboració amb laboratoris, <i>hubs</i> i centres locals d'innovació d'IA • Observatori d'IA local • Entorns aïllats de proves reguladors locals d'IA (<i>regulatory sandboxes</i>) • Iniciatives per promoure els ecosistemes locals d'IA i donar-hi suport
[12] Participació de la comunitat	Eines per implicar els ciutadans, les comunitats i les parts interessades en els processos de presa de decisions en IA, fomentar converses i debats i garantir que les polítiques d'IA reflecteixin les preocupacions i les aportacions de la ciutadania.	<ul style="list-style-type: none"> • Participació popular: processos participatius, models de govern participatiu, sessions d'audiència pública, promoció d'espais de reflexió i debat, canals de comunicació amb la ciutadania, etc. • Formació dels ciutadans (alfabetització digital) • Comitès locals d'ètica d'IA
[13] Governança de dades		<ul style="list-style-type: none"> • Mesures de transparència de les dades com ara portals de dades obertes • Mecanismes d'intercanvi de dades • Drets sobre les dades • Protocols d'ús de les dades • Privacitat de les dades: lleis de privacitats de dades, marcs de dades sintètiques, etc. • Sistemes de governança de dades
[14] Altres mesures i polítiques		<ul style="list-style-type: none"> • Marcs i eines de prova • Incentius fiscals com ara desgravacions fiscals, subvencions, etc. • Programes de requalificació de treballadors • Marcs d'avaluació (per exemple, marcs d'avaluació de la IA mitjançant estrelles, certificacions verdes, etc.)

Font: Autores

2. Concordança dels mecanismes polítics amb principis ètics

L'anàlisi dels mecanismes polítics identificats al llarg d'aquesta publicació² palesa que es poden establir diverses tipologies tenint en compte el grau de concordança amb principis ètics concrets (vegeu la Taula 2), que s'examinen a continuació.

- a) **Mecanismes polítics que serveixen per defensar determinats principis ètics.** Per exemple, una avaluació d'impacte mediambiental compleix principalment propòsits de sostenibilitat, mentre que una mesura d'intervenció humana (*human-in-the-loop*) pot fomentar tant el principi de rendició de comptes com el d'equitat en garantir que algú supervisa el funcionament correcte d'un sistema d'IA i, en darre-ra instància, deriva la decisió final a un humà.
- b) **Mecanismes polítics que defensen tots els principis ètics:** degut al caràcter transversal que presenten o gràcies a la personalització, alguns mecanismes polítics poden fomentar tots els principis ètics en conjunt. Per exemple, mecanismes polítics com les estratègies d'IA es poden personalitzar per incloure tots els principis ètics. De la mateixa manera, un comitè de supervisió es pot encarregar de supervisar la protecció de la privacitat, la rendició de comptes o tot el paquet de principis ètics.

Malgrat el seu potencial, no tots són aplicats amb la mateixa freqüència per les ciutats que s'esforcen per establir marcs locals d'IA ètica. Alguns exemples destacats d'aquests mecanismes són:

- **Principis, estratègies i directrius:** un dels mecanismes més utilitzats per les ciutats d'arreu del món. Les ciutats implementen sistemàticament aquests mecanismes polítics des d'un punt de vista ètic per proporcionar-los una base i orientació. Aquests mecanismes són especialment utilitzats pels municipis per demostrar la seva voluntat de comprometre's amb un desplegament responsable de l'IA.
- **Clàusules d'adquisició públiques:** com que els municipis sovint no tenen els recursos necessaris per crear els seus propis sistemes d'IA interns, un altre mecanisme polític que s'aplica habitualment són les clàusules d'adquisició pública, que acaben esdevenint fonamentals i són una solució pràctica i popular. Permeten als municipis aprofitar el seu poder adquisitiu quan adquireixen sistemes d'IA i alhora promouen un desenvolupament ètic de la IA per part dels proveïdors del sector privat.
- **Mecanismes externs:** cada cop hi ha més ciutats que recorren a òrgans assessors i supervisors constituïts per experts externs i independents que assessoren el municipi sobre problemes ètics i supervisen l'ús que fan dels sistemes algorítmics. De la mateixa manera, molts municipis formen part d'**aliances, comunitats de pràctiques i grups d'aprenentatge** per abordar conjuntament els reptes i identificar maneres d'utilitzar la IA de forma segura.

L'anàlisi dels mecanismes polítics identificats al llarg d'aquesta publicació palesa que es poden establir diverses tipologies tenint en compte el grau de concordança amb principis ètics concrets.

2. Extrets únicament dels capítols (Part I) i dels casos pràctics (Part II) d'aquesta monografia.

- **Governança de dades:** tot i que la governança de les dades pot no percebre's d'entrada com un mecanisme polític directe per a la IA, assenta les bases per gestionar de manera correcta les dades dels ciutadans i salvaguardar-les, i és fonamental perquè els sistemes no siguin discriminatoris, la qual cosa la converteix en una peça clau per a un desplegament ètic de la IA. La governança de les dades és un pilar fonamental que permet la transparència de les dades, la protecció dels drets en matèria de dades, la privacitat de les dades i l'ús sostenible de les dades per als sistemes d'IA. En són exemples els protocols per a l'anonimització de dades personals o l'ús de dades sintètiques per entrenar els sistemes d'IA, per tal de resoldre problemes de privacitat i equitat.
- **Auditories:** tot i que els experts i els treballadors públics les consideren un dels mecanismes polítics fonamentals per garantir l'ètica, les auditories segueixen sent poc utilitzades. S'utilitzen poc en part degut a les limitacions imposades pels proveïdors privats d'IA i a la manca de capacitació tècnica interna.

c) Mecanismes polítics no associats a cap principi ètic específic.

Tot i que no estan vinculats a cap principi ètic específic, aquests mecanismes polítics es consideren clau per a un desplegament ètic de la IA. Estableixen els processos estructurats que cal seguir, en coordinen el desplegament o proporcionen l'especialització i els coneixements necessaris per prendre decisions de manera fonamentada. Alguns exemples destacats d'aquests mecanismes són:

- **Protocols interns per a l'ús de la IA:** la majoria de les ciutats desenvolupen protocols interns per guiar el seu ús de la IA, que proporcionen estructures pas a pas per facilitar-ne la implementació per part del municipi. Algunes ciutats, encara que una minoria, els complementen, a més, amb **protocols integrals per a l'adquisició pública d'IA**. Aquests protocols poden incloure avaluacions obligatòries d'impacte, anàlisi de biaixos i altres mecanismes polítics per garantir el respecte de principis ètics concrets.
- **Centres, hubs i laboratoris innovadors d'IA:** un nombre important de ciutats col·laboren amb centres, *hubs* o laboratoris innovadors d'IA, els fomenten o n'han establert per crear coneixement pràctic i desenvolupar solucions d'IA. La recerca que duen a terme aquestes institucions té una orientació ètica.
- **Programes de capacitació:** un dels mecanismes menys implementats és la creació d'**equips municipals destinats a la IA** amb l'especialització necessària per auditar o desenvolupar sistemes d'IA interns. Això es deu principalment a les limitacions tècniques i econòmiques d'una part dels municipis. En canvi, molts han creat òrgans municipals d'IA encarregats de coordinar l'ús de la IA entre els departaments. Aquests òrgans tenen un paper fonamental a l'hora de facilitar el seguiment transversal del desplegament de la IA a la ciutat, garantint un enfocament més organitzat i responsable de la governança de la IA.

Taula 2. Concordança dels mecanismes polítics amb principis ètics

MECANISMES POLÍTICS		PRINCIPIS ÈTICS			
		Rendició de comptes i transparència	Privacitat i governança de dades	Equitat i no-discriminació	Sostenibilitat
[1]	Estratègies d'IA	X	X	X	X
	Principis locals per a la IA	X	X	X	X
	Directrius, guies i manuals	X	X	X	X
	Protocol intern per a l'ús de la IA				
[2]	Regulacions i lleis locals	X	X	X	X
	Mecanismes de compliment legal	X	X	X	X
	Estàndards reguladors	X	X	X	X
	Estàndards normatius internacionals	X	X	X	X
	Regulació sobre l'aplicació d'IA controvertida		X	X	
[3]	Registre públic d'algoritmes	X	X	X	
	Web municipal amb tota la informació pertinent sobre la IA	X			
	Directori municipal d'eines d'IA adquirides per a ús intern	X			
[4]	Avaluació i gestió del risc	X	X	X	X
	Avaluacions d'impacte en matèria de drets humans	X	X	X	X
	Avaluació d'impacte ambiental				X
[5]	Auditories	X	X	X	X
	Eines per a l'autoavaluació	X	X	X	X
[6]	Elaboració d'informes i supervisió interna				
	Intervenció humana (human-in-the-loop)	X		X	
	Procediments de resposta i objeccions per a la ciutadania	X		X	
[7]	Clàusules d'adquisició pública	X	X	X	X
	Protocols interns per a l'adquisició de serveis d'IA				
[8]	Òrgans assessors i supervisors	X	X	X	X
[9]	Aliances, comunitats de pràctiques i grups d'aprenentatge	X	X	X	X
[10]	Formació del personal municipal	X	X	X	X
	Equip municipal d'IA				
	Òrgan municipal d'IA				
	Enfocament multidisciplinari			X	
[11]	Laboratoris, hubs i centres d'innovació d'IA				
	Observatori d'IA local				
	Entorns aïllats de proves reguladors (regulatory sandboxes)		X		
[12]	Participació popular	X		X	X
	Formació del públic (alfabetització digital)		X	X	X
	Comitès locals d'ètica d'IA	X	X	X	X
[13]	Governança de dades	X	X	X	X
[14]	Marcos i eines de prova	X	X	X	X
	Incentius fiscals (desgravacions fiscals, subvencions...)				X
	Programes de requalificació de treballadors				X
	Marcos d'avaluació	X	X	X	X

Font: les autores

Llegenda de la taula: groc (mecanismes polítics en consonància amb un principi ètic concret o diversos a l'hora); blau (mecanismes polítics transversal o personalitzables que serveixen per a tots els principis ètics); verd (mecanismes polítics no associats a cap principi ètic específic, però rellevants per a un desplegament responsable dels sistemes algorítmics en general). Gris clar (mecanismes polítics que no s'esmenten en els capítols); gris fosc (mecanismes polítics esmentats explícitament en els capítols; vegeu l'annex I).

Independentment del nivell d'especificitat, abast o eficàcia dels mecanismes polítics, les ciutats no els apliquen de manera uniforme, i s'observen diferències quant a la freqüència a causa de limitacions de recursos, limitacions tècniques o prioritats locals. La taula 3 compara les ciutats incloses en els estudis de cas i destaca la freqüència amb què s'apliquen determinats mecanismes polítics, quins són els més utilitzats i a quins es recorre poc sovint. Tot i que s'ha considerat rellevant comparar les ciutats analitzades a la part II d'aquesta publicació, cal reconèixer que les conclusions extretes dels estudis de cas seleccionats són limitades pel nombre reduït de les ciutats estudiades.

MECANISME POLÍTIC		CIUTATS					
		Barcelona	Amsterdam	Nova York	San José	Dubai	Singapur
[1]	Estratègies d'IA, agendes, plans d'acció, manuals, fulls de ruta, etc.	x	x	x	x	x	x
[1]	Principis locals per a la IA	x	x	x	x	x	x
[3]	Web municipal amb tota la informació pertinent sobre la IA	x	x	x	x	x	x
[13]	Portal de dades obertes (transparència de les dades)	x	x	x	x	x	x
[1]	Directrius, guies i manuals		x	x	x	x	x
[8]	Òrgans assessors i supervisors	x	x	x	x		x
[9]	Membre d'una comunitat de pràctiques	x	x	x	x		x
[1]	Protocols interns per a l'ús de la IA	x	x	x	x		
[4]	Avaluació i gestió del risc	x	x	x	x		
[7]	Clàusules d'adquisició pública	x	x	x	x		
[12]	Participació popular	x	x	x	x		
[3]	Registre públic d'algoritmes	x	x		x		
[4]	Avaluació d'impacte en matèria de drets humans	x	x		x		
[9]	Col·laboracions multiactors en l'àmbit de la IA		x		x		x
[10]	Formació del personal municipal (capacitació municipal)	x	x				x
[10]	Òrgan municipal d'IA	x		x	x		
[11]	Laboratoris, hubs i centres d'innovació d'IA		x			x	x
[13]	Altres polítiques de governança de dades (drets en matèria de dades, mecanismes d'intercanvi de dades, protocols de dades...)	x			x	x	
[2]	Estàndards de transparència algorítmica	x	x				
[6]	Elaboració d'informes i supervisió interna			x	x		
[6]	Procediments de resposta i objeccions per a la ciutadania		x		x		
[5]	Auditories	x	x				
[2]	Regulació sobre l'aplicació d'IA controvertida			x			
[3]	Director municipal d'eines d'IA adquirides per a ús intern			x			
[10]	Creació d'un equip municipal d'IA (capacitació municipal)		x				
[11]	Observatori d'IA local	x					
[11]	Entorns aïllats de proves reguladors (regulatory sandboxes)						x
[14]	Marc i eines de prova						x
[14]	Marc d'avaluació						x

Font: les autores

Nota: la llista de mecanismes polítics s'ha ordenat, en primer lloc, per freqüència descendent i, en segon lloc, per categorització.

3. Reptes i recomanacions

- a. **Hi ha poques referències sobre com aplicar de manera efectiva els principis ètics d'IA a escala local.** La majoria de les directrius, els estudis i els programes de capacitació existents no tenen en compte els reptes específics als quals s'enfronten les administracions urbanes, que sovint estan desconnectades de les estratègies i les polítiques nacionals. Aquesta discrepància s'agreuja encara més per la naturalesa heterogènia dels governs locals, que presenten importants diferències quant a mida, recursos i capacitació.

Recomanacions:

- Localitzar (és a dir, adaptar al context local) els principis ètics regionals, nacionals o globals i els mecanismes polítics de governança de la IA establint definicions locals d'èxit i identificant les prioritats locals.
- b. **Les administracions locals pateixen una important manca de talent i coneixements tècnics,** un problema que s'agreuja encara més per l'escassetat global de professionals d'IA, que dificulta atraure localment persones qualificades. En conseqüència, els governs municipals solen tenir un coneixement limitat de les possibles repercussions i implicacions dels sistemes algorítmics.

Recomanacions:

- Prioritzar els programes de capacitació com a part de les estratègies municipals i els marcs polítics per controlar i regular els sistemes algorítmics. Això hauria d'incloure l'assignació de recursos específics per a programes municipals de formació, la inversió en campanyes de conscienciació pública i la promoció d'iniciatives per adquirir coneixements i habilitats fonamentals al voltant de la IA ètica a l'Administració pública.
 - Per superar els reptes d'atraure talent local, els governs locals han d'invertir en estratègies que facilitin l'intercanvi i l'adaptació del coneixement dels agents locals. A més, forjar aliances i connexions sòlides amb xarxes d'intercanvi de coneixements pot ajudar a resoldre les deficiències quant a especialització.
 - Adoptar un enfocament holístic de la capacitació fomentant els debats públics i les iniciatives de sensibilització dins de les comunitats locals. Cal centrar aquests esforços a formar els ciutadans sobre les oportunitats i els riscos associats a l'ús de sistemes algorítmics.
- c. **Garantir la transparència i la rendició de comptes dels sistemes algorítmics** utilitzats pels governs locals planteja diversos reptes, com ara: la gestió de l'opinió pública i les possibles reaccions negatives respecte als sistemes d'IA externs, l'adaptació de la cultura organitzativa i de les pràctiques laborals per als sistemes d'IA interns, i el foment de la responsabilitat compartida a tota l'Administració pública (és a dir, la IA no hauria de ser considerada una responsabilitat exclusiva del departament de TIC).

Independentment del nivell d'especificitat, abast o eficàcia dels mecanismes polítics, les ciutats no els apliquen de manera uniforme, i s'observen diferències quant a la freqüència a causa de limitacions de recursos, limitacions tècniques o prioritats locals.

Recomanacions:

- Incorporar la transparència com a objectiu bàsic més enllà de les polítiques específiques d'IA. Això inclou fomentar una cultura de la transparència i rendició de comptes al llarg de tot el cicle de vida de la IA.
 - Fomentar l'ús de codi obert, que pot augmentar la confiança i permetre un major control i col·laboració.
 - Deixar clares les responsabilitats designant una persona de contacte específica o «líder del projecte d'IA» per a totes les iniciatives d'IA, de manera que es garanteixi la rendició de comptes i es faciliti la comunicació.
 - Crear múltiples canals per fer arribar comentaris i integreu avaluacions en les etapes clau del projecte per garantir la millora contínua i la capacitat de resposta.
 - Destinar part del pressupost a fases explicatives exhaustives, de manera que es garanteixi que les parts interessades, tant internes com externes, entenguin completament els sistemes d'IA i les seves implicacions.
- d .** Assegurar la **protecció de la privacitat** presenta reptes específics, com ara un panorama regulador complex en el qual els governs locals sovint tenen problemes per desenvolupar-se. Els problemes relacionats amb les dades estan estretament lligats als mecanismes polítics dissenyats per salvaguardar la protecció de dades. En particular, als entorns urbans hi ha disponibles poques dades d'alta qualitat, cosa que es pot atribuir a diversos factors: pràctiques inadequades de gestió de les dades, qüestions ètiques i riscos relacionats amb la recollida de dades a gran escala i un intercanvi de dades inadequat entre administracions per l'absència d'estàndards unificats i de marcs ben desenvolupats de governança de les dades.

Recomanacions:

- Generar conjunts de dades públiques d'alt valor millorant les pràctiques de recollida i gestió de dades per millorar la qualitat i la utilitat de les dades.
 - Promoure la interoperabilitat i la col·laboració intersectorial i entre organismes per facilitar un intercanvi fluid de dades.
 - Crear marcs segurs i transparents per compartir dades que garanteixin la protecció de la privacitat i alhora permetin la innovació.
 - Fomentar la innovació i l'experimentació en entorns controlats, com ara entorns aïllats de proves reguladors, per provar noves solucions basades en dades, de manera segura i responsable.
- e .** Els sistemes d'algoritmes poden reforçar les desigualtats urbanes existents i alhora crear noves formes de discriminació, d'aquí la importància de parar atenció a la **noció d'equitat i no-discrimi-**

nació quan l'Administració local desplega sistemes d'IA. Un repte concret en aquest sentit és que la discriminació automatitzada provocada per la IA és més abstracta, opaca, difícil de detectar (caixes negres) i a gran escala. Per tant, altera els recursos i els procediments jurídics tradicionals que solen utilitzar els governs locals per detectar-la, prevenir-la i corregir-la.

Recomanacions:

- Considerar les múltiples funcions de les administracions públiques –com a desenvolupadores, responsables del desplegament i reguladores– en dissenyar iniciatives per millorar l'equitat i la no-discriminació dels sistemes algorítmics.
 - Localitzar els marcs polítics existents per abordar la discriminació no intencionada en els sistemes algorítmics i garantir que s'adaptin als reptes específics dels entorns urbans.
 - Incorporar un enfocament holístic a la governança de la IA dins de les administracions locals, tenint en compte els impactes socioeconòmics al llarg de tot el cicle de vida de la IA, des del disseny fins al desplegament.
 - Garantir la diversitat entre els equips implicats en el disseny i el desplegament de sistemes algorítmics per reduir el risc de resultats esbiaixats i promoure solucions inclusives.
- f. El repte més important associat al principi de **sostenibilitat mediambiental de la IA** té a veure amb el fet que la IA per a la sostenibilitat sovint entra en conflicte amb la sostenibilitat de la pròpia IA. A més, hi ha pocs marcs per a la sostenibilitat de la IA amb un enfocament urbà.

Recomanacions:

- Avaluar l'impacte del cicle de vida complet dels sistemes d'IA per determinar si els seus beneficis superen els seus costos mediambientals. Minimitzar la petjada de carboni dels desplegaments de la IA a tota la ciutat prioritant els sistemes eficients energèticament, adoptant pràctiques de computació ecològica, utilitzant aprenentatge automàtic en dispositius petits (*tiny machine learning*) i alimentant els centres de dades amb energia renovable.
- Fomentar l'economia circular al voltant dels centres de dades reduint els residus electrònics. Promoveu l'aprovisionament responsable, la reutilització i el reciclatge del *hardware* de la IA. Fomentar la reutilització i el reciclatge de tecnologies, dades i infraestructures d'IA.
- Reutilitzar l'energia i els recursos utilitzats per la infraestructura d'IA i desplegar sistemes d'IA que s'integrin perfectament amb la infraestructura urbana existent, optimitzant tant el consum energètic com l'eficiència del sistema.

Finalment, és important reconèixer les limitacions de la recerca presentada en aquesta monografia CIDOB. L'estudi es va veure limitat tant pel període de temps relativament curt en què es va dur a terme com pels reptes inherents a la recollida d'informació en el context del GOUAI

(vegeu Galceran-Vercher i Vidal, 2024). En conseqüència, es reconeix que hi poden faltar alguns exemples clau de mecanismes polítics i marcs de governança algorítmics. A més, la majoria de les evidències que es presenten s'han extret de polítiques promogudes pel Nord Global, principalment a través d'intervencions liderades pels governs locals dels EUA i Europa. Aquest enquadrament geogràfic és una altra limitació que la futura recerca en el context del GOUAI intentarà abordar. Es reconeix que una anàlisi més sistemàtica de les polítiques i pràctiques de governança del Sud Global podria proporcionar nous coneixements i revelar diferents enfocaments polítics, prioritats i reptes de la implementació.

Referències bibliogràfiques

Ada Lovelace Institute, AI Now Institute and Open Government Partnership. «Algorithmic Accountability for the Public Sector», 2021

Ben Dhaou, S., et al. «Global Assessment of Responsible Artificial Intelligence in Cities: Research and recommendations to leverage AI for people-centred smart cities». Nairobi: Programa de las Naciones Unidas para los Asentamientos Humanos (ONU-Habitat), 2024

Jordan, C.; Glickman, J. i Panettieri, A. «AI in cities: Report and Toolkit». Washington: Liga Nacional de Ciudades i Google, 2024

OECD «OECD AI Policy Observatory: Catalogue of Tools & Metrics for Trustworthy AI», 2024

United 4 Smart Sustainable Cities. «Guiding principles for artificial intelligence in cities». Ginebra: Unión Internacional de Telecomunicaciones, 2024

ANNEX 1. Llista de mecanismes polítics esmentats en els capítols dels principis ètics (Part I)

Principis ètics	Mecanismes polítics
Rendició de comptes i transparència	<ul style="list-style-type: none"> • Avaluacions d'impacte [4] • Clàusules d'adquisició pública [7] • Auditories externes algorítmiques [5] • Registres d'algoritmes [3] • Normes de transparència [3] • Comitès multidisciplinaris de supervisió de la governança [8] • Processos participatius (en totes les etapes del cicle de vida de la IA) [12] • Disseny amb intervenció humana (<i>human-in-the-loop</i>) [6] • Formació i reciclatge laboral del funcionariat [10] • Contactes amb xarxes d'intercanvi de coneixement [9] • Col·laboració amb les parts interessades locals [9]
Privacitat i governança de dades	<ul style="list-style-type: none"> • Conformitat legal [2] • Sistemes de gestió de risc [4] • Sistemes de governança de dades [13] • Avaluacions d'impacte [4] • Auditories [5] • Repositoris d'algoritmes i registres de sistemes d'IA [3] • Sandboxes reguladors [11] • Estratègies urbanes d'IA [1] • Col·laboració multiactor [9]
Equitat i no-discriminació	<ul style="list-style-type: none"> • Estratègies d'IA [1] • Anàlisi de riscos i mecanismes de protecció [4] • Avaluacions d'impacte [4] • Normes locals per a la IA equitativa [2] • Normes d'adquisició pública per a una IA equitativa [7] • Legislació urbana per garantir el dret a la justificació [2] • Òrgans assessors multidisciplinaris [8] • Equips diversos i interdisciplinaris • Auditories [5] • Tècniques de mitigació en el cicle de vida de la IA [14] • Xarxes d'intercanvi de coneixement [9] • Formació del personal [10] • Formació de les comunitats ciutadania [12] • Principis [1] • Anàlisi de biaixos [4] • Protecció dels drets digitals [14]
Sostenibilitat	<ul style="list-style-type: none"> • Normes mediambientals, socials i de governança (per exemple, normes d'IA verda) [2] • Avaluacions d'impacte [4] • Supervisió i auditoria [5] • Consorci de ciutats per a la IA sostenible [9] • Incentius fiscals: desgravacions fiscals o subvencions [14] • Junes d'ètica de la IA local [12] • Lleis sobre la privacitat de les dades urbanes [13] • Participació pública [12] • Models de governança participativa [12] • Campanyes d'alfabetització digital [12] • Programes de reciclatge laboral [14] • Sistemes de classificació amb estrelles per a l'eficiència energètica de la IA [14] • Col·laboració internacional [9] • Estandardització internacional [2]

Font: les autores

Nota: els mecanismes polítics que es descriuen aquí mantenen la redacció original dels capítols dels principis ètics (Part I) i s'han classificat segons la categorització de les autores que es proporciona en el capítol final de la Monografia. Pot ser que els mecanismes no apareguin a l'apartat de «mecanismes polítics» dels capítols, però es troben al llarg del capítol en qüestió.

