



Taula d'entitats
del Tercer Sector Social
de Catalunya

m4social

INFORME

**RADAR D'ALGORITMES
D'INTEL·LIGÈNCIA ARTIFICIAL
I PROCESSOS DE DECISIÓ
AUTOMATITZADA PER A
L'ACCÉS ALS DRETS SOCIALS
A CATALUNYA**

GENER DE 2024

CRÈDITS

Estudi impulsat per:

Taula d'entitats del Tercer
Sector Social de Catalunya

Coordinació:

Roger Civit
Laura Castro
Xènia Carné
Gerard Sentís

Autoria de l'estudi [KSNET]:

Romén Adán González
Mireia Álvarez
Elena Costas

Peer reviewers [KSNET]:

Beatriz Belmonte Acero
Manuel González Noriega

Disseny: Toni Sánchez Poy

Llicència: CC BY-NC-SA 4.0



Barcelona, Febrer 2024

ÍNDEX

Resum	4
-------------	---

Document 1. La intel·ligència artificial en les administracions públiques

1. Introducció	7
2. Aclariment de conceptes	10
2.1 Conceptes clau	10
2.2 Un exemple de l'aplicació	12
3. L'origen de la intel·ligència artificial i els últims avenços	16
4. Governança i estratègies nacionals en l'àmbit de la intel·ligència artificial	21
4.1 El rol dels Estats en la governança de la IA	23
4.2 El rol dels Estats en l'aplicació de la IA en el Sector Públic	24
4.3 L'Estratègia Nacional d'Espanya	25
5. L'enfocament centrat en la persona i la llei de la intel·ligència artificial de la Unió Europea	27
5.1 L'enfocament centrat en la persona	27
5.2 La Llei de la IA de la Unió Europea	30
5.3 Les accions d'Espanya	34
<i>El Reial decret de sandbox d'Espanya</i>	<i>34</i>
<i>L'Estatut de l'Agència Espanyola de Supervisió de la Intel·ligència Artificial (AESIA)</i>	<i>35</i>
6. La intel·ligència artificial a les administracions públiques	36
6.1 Què entenem per IA en el context de les AAPP?	36
6.2 Aplicacions de la IA a les AAPP	37
6.3 Casos internacionals	43
6.4 Aplicacions en l'àmbit social	48
7. Els beneficis i riscos de l'aplicació de la intel·ligència artificial a les administracions públiques en l'àmbit social	51
8. Conclusió	56
Referències	57

Document 2. Radar d'algoritmes i processos de decisió automatitzada per a l'accés als drets socials de la ciutadania

1. Introducció	63
2. Metodologia	65
3. El disseny del radar d'algoritmes	67
3.1. Camps del Radar d'algoritmes	67
4. Resultats	74
Fitxes informatives	78
1. Intel·ligència col·lectiva - Sistema d'identificació de demandes, Wproblemes i respostes (DPR automàtic)	78
2. Processador de casos IRIS	82
3. Sistema de suport a la classificació d'incidències - MARIO	83
4. Sistemes conversacionals, infoCanal Alzheimer i Fons 0-16	84
5. Transcripció d'informes d'emergència social (CUESB)	86
6. RisCanvi	87
7. Motor de Regles de Prestacions Socials	90
8. Recomanador de Prestacions	91
9. Identificació necessitats en la Intervenció Social	92
10. Xatbots conversacionals de l'Administració Oberta de Catalunya (AOC)	94
11. Servei de videoidentificació per obtenir l'idCAT Mòbil	98
12. Servei d'automatització dels informes de pobresa energètica	102
5. Valoracions finals	108

RESUM

En l'última dècada, la intel·ligència artificial (IA) s'ha desenvolupat de forma exponencial. Una evolució que ha estat principalment impulsada pels progressos en el camp de l'aprenentatge profund (*deep learning*). La IA generativa, tant de text com d'imatge o àudio, es troba avui en un nivell de desenvolupament difícilment imaginable pocs anys enrere. D'altra banda, els sistemes que utilitzen IA estan cada vegada més integrats en la vida quotidiana. Des d'assistents personals intel·ligents, sistemes de recomanació en plataformes, bots, models que ajuden en el diagnòstic i tractament de malalties, o en la traducció i transcripció de textos.

A més, **les possibles aplicacions de la IA posseeixen un potencial transformador, no sols en el sector privat, sinó també en el sector públic i, en particular, en l'àmbit social.** La capacitat per a simplificar processos o proveir millors serveis públics pot facilitar i agilitzar tant el treball del personal de l'administració, com la relació dels ciutadans amb aquesta. No obstant això, l'ús i aplicació de sistemes d'IA, a més de beneficis, comporta igualment una sèrie de riscos.

En aquest context, **des de la Taula d'entitats del Tercer Sector Social de Catalunya s'ha impulsat el projecte Radar d'algoritmes d'intel·ligència artificial i processos de decisió automatitzada per a l'accés a drets socials de la ciutadania de Catalunya.** Aquest projecte compta amb un doble objectiu: d'una banda, ajudar a la societat en general, i a les entitats del tercer sector en particular, a entendre el context actual per a poder fixar una posició compartida. En segon lloc, mapar tots aquells sistemes d'intel·ligència artificial o de decisió automatitzada que estan sent utilitzats actualment per part de les Administracions Públiques catalanes en l'àmbit social.

Per aquesta raó, **el present treball es compon de dos documents.** El primer d'ells, ***La Intel·ligència Artificial en les Administracions Públiques***, lliurat al setembre de 2023, és un informe de caràcter divulgatiu sobre els orígens, evolució, governança, beneficis, riscos i aplicació de la intel·ligència artificial per part de l'Administració Pública.

El segon document, ***Radar d'algoritmes i processos de decisió automatitzada per l'accés als drets socials de la ciutadania***, finalitzat el desembre de 2023, reuneix un total de 12 sistemes identificats de 4 administracions públiques catalanes diferents. Per a cadascun d'aquests algoritmes s'ha elaborat una fitxa informativa que recull informació rellevant i variada. Aquest Radar és el primer repositori de sistemes d'intel·ligència artificial i processos automatitzats en l'àmbit social utilitzats per Administracions Públiques de Catalunya que s'ha creat a dia d'avui.

En un context de profundes transformacions és fonamental poder identificar allò que és rellevant, els avenços i els riscos que comporten, i com i què s'espera aconseguir d'allò què s'està desenvolupant. Amb aquest treball es busca contribuir a la divulgació sobre Intel·ligència Artificial així com a la discussió pública, principalment donant una major difusió pública a aquells sistemes d'intel·ligència artificial que estan sent utilitzats per part de l'administració.

DOCUMENT 1

LA INTEL·LIGÈNCIA ARTIFICIAL EN LES ADMINISTRACIONS PÚBLIQUES

SETEMBRE DE 2023

1. INTRODUCCIÓ

‘Intel·ligència Artificial’ (IA) s’ha convertit en un terme tan popular com (potser) poc útil. Un concepte paraigua que sovint costa esbrinar a què està fent referència quan s'utilitza. Si bé fa anys que convivim amb els avenços científics en el camp de la IA, amb el seu ús en el sector privat i la seva integració en moltes de les aplicacions i webs que utilitzem en el dia a dia, fins a l'actualitat els mitjans de comunicació i l'opinió pública encara no hi havien posat especial atenció. Probablement, tot ha canviat amb l'aparició del ChatGPT, una interfície amigable que permet a qualsevol persona amb accés a Internet interactuar directament i conscientment amb un sistema d'Intel·ligència Artificial. Un sistema que, en pocs segons, produeix respostes similars a les que podria proporcionar una persona.

Això ha generat molts dubtes en el debat públic. La majoria de persones perdran les seves feines? Acabarà la IA amb el pensament crític? O inclús amb la humanitat? Tanmateix, és important generar espais de debat asossegats sobre la Intel·ligència Artificial. **De què parlem exactament quan parlem d'IA? Quins són els seus orígens i quin ha estat el seu desenvolupament? Quin ús se li ha donat en el sector privat en l'última dècada? Com es podria beneficiar l'Administració pública i quines aplicacions innovadores s'estan realitzant en altres països? Quins beneficis pot aportar la IA a la societat? Quins riscos existeixen i com hem d'afrontar-los i regular-los?, etc.**

El projecte AI Impacts va realitzar el 2022 la segona edició de l'*Expert Survey on Progress in AI*¹, on varen entrevistar a 738 investigadors experts en *Machine Learning* sobre els impactes de la Intel·ligència Artificial en el futur. Tot i que no deixa de ser un judici valoratiu, la meitat dels experts en Intel·ligència Artificial creu que la probabilitat que l'efecte a llarg termini de la IA avançada en la humanitat sigui 'extremadament dolenta (per exemple, l'extinció humana)' és solament del 5% o menys. Encara més, un de cada quatre experts la redueix a un 0%.

Si bé la Intel·ligència Artificial té un potencial de canvi molt destacat en el sector privat, el sector públic no n'està exempt. Diverses Administracions Públiques ja compten amb aquesta tecnologia per a ajudar-les en diferents tasques. De fet, cal assenyalar que, actualment, la principal tecnologia emprada pel sector públic són els sistemes de decisió automatitzada (o ADMS, per les seves sigles en anglès), que poden (o no) utilitzar IA.

¹ Es pot consultar [aquí](#).

Els sistemes de decisió automatitzada (ADMS) són aquells que tenen capacitat per a prendre decisions a través de mitjans tecnològics sense la intervenció humana. Poden estar basats en qualsevol mena de dades, com aquelles facilitades pels individus a través de qüestionaris, dades observades com les de geolocalització, o dades inferides a partir d'unes certes característiques². Com s'explica a l'informe *Automating Society* d'AlgorithmWatch³, els ADMS poden (o no) utilitzar tècniques d'intel·ligència artificial, és a dir, poden recórrer a procediments simples d'anàlisis basats en regles i automatitzacions, o implementar tècniques sofisticades com el processament de llenguatge natural, l'analítica predictiva o la visió per computadora. En qualsevol cas, els autors advoquen per adoptar un enfocament holístic, d'aquí l'ús de la paraula *sistemes* en comptes de *technologies*. Al fer referència a un sistema, no es pren en consideració únicament l'algoritme o tècnica de forma aïllada, sinó tot el procés de presa de decisió, el diagnòstic que explica per què es va adoptar aquesta solució, l'algoritme en si, les dades, el codi, la forma en què va ser desenvolupat (per part d'una companyia pública o privada) i finalment implementat. Des d'un enfocament holístic es pot comprendre el procés integral des del moment que s'identifica una problemàtica o necessitat, fins a implementar finalment una solució, passant per tots els passos intermedis necessaris.

Així doncs, en aquest informe s'utilitza el concepte d'IA, com un concepte ampli que inclou tant els ADMS com la Intel·ligència Artificial, donat que actualment els ADMS es troben amb major freqüència dins del Sector Públic, i presenten oportunitats i riscos semblants als de la IA.

Al situar la IA com un recurs del Sector Públic a l'hora de, per exemple, assignar ajuts socials entre la ciutadania, és imprescindible que les entitats del tercer sector coneguin com s'estan utilitzant aquests sistemes en l'àmbit dels drets socials. Així doncs, **la finalitat d'aquest informe és ajudar a la societat en general, i a les entitats del Tercer Sector en particular, a entendre el context actual per poder fixar una posició compartida**. Si les entitats volen ser escoltades, és fonamental que comparteixin i transmetin una visió similar, tant en termes de diagnòstic com de propostes.

Sota aquest punt de partida, l'informe s'estructura de la següent manera: en el segon apartat es defineixen breument una sèrie de conceptes que apareixen al llarg de l'informe. Al tercer capítol s'exposa l'origen i desenvolupament de la IA fins a l'actualitat. A continuació, es presenta un breu

² Aquesta definició s'ha extret del document *Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679* elaborat per el Data Protection Working Party, un òrgan assessor de la Unió Europea. Pot consultar-se [aquí](#).

³ Disponible [aquí](#).

exemple per aclarir el concepte d'entrenament d'una IA. Als apartats 4 i 5 es posa el focus sobre la governança de la IA, el paper dels governs i la Llei d'IA de la Unió Europea. El capítol 6 presenta aplicacions de sistemes que utilitzen IA per part d'Administracions Públiques (AAPP), tant internacionals com a Catalunya i Espanya, en diversos àmbits. Finalment, el capítol 7 fa un repàs sobre els principals beneficis i, sobretot, riscos de l'aplicació de la IA per part de les AAPP en l'àmbit social.

2. ACLARIMENT DE CONCEPTES

La Intel·ligència Artificial (IA) és un camp ampli i complex, amb molts subapartats i tècniques associades. Aquest concepte paraigua aglutina una àmplia diversitat de tècniques: aprenentatge automàtic (*Machine Learning*), processament de llenguatge natural, el reconeixement de veu, els sistemes experts, la robòtica o la visió per computadora. Abans de començar, i per evitar confusions, a continuació es descriuen alguns dels conceptes utilitzats al llarg d'aquest informe.

2.1 Conceptes clau

- ▮ **Algoritme:** En termes generals, un algoritme és un conjunt de regles o instruccions que, pas a pas, resolen un problema o aconseguen un objectiu. En el context de la IA, els algoritmes sovint es refereixen als mètodes utilitzats per construir i entrenar models d'aprenentatge automàtic. Hi ha una gran quantitat d'algoritmes diferents i, a l'hora de crear un model d'IA, el científic de dades escull el més adequat per al problema que es pot resoldre.
- ▮ **Dades:** Sense dades no hi ha Intel·ligència Artificial. Les dades són la font d'informació. Bona part de la creació d'un model es centra en la recollida i neteja de les dades. Les dades poden ser imatges, text o una matriu de dades (com un full d'Excel). Independentment del seu format original, la informació sempre es converteix en números perquè l'algoritme escollit pugui aprendre.
- ▮ **Entrenament:** És la fase de creació del model. Després d'haver recollit i processat una quantitat de dades determinada, aquestes són passades a un algoritme concret perquè 'apregui' a reconèixer els patrons. Per exemple, si es disposa d'informació sobre la ubicació, les característiques i el preu d'habitatges, es podria entrenar un model perquè apregui a predir el preu d'un habitatge nou a partir de la seva ubicació i característiques.
- ▮ **Intel·ligència Artificial (IA):** No hi ha consens a la comunitat científica sobre com definir el terme IA. No obstant això, la Comissió Europea ha proposat una definició operacional que permet discernir responsabilitats legals, i que està basada en una concepció de l'OCDE. L'article 3(1) de la proposta de Llei d'Intel·ligència Artificial estableix que un 'sistema d'Intel·ligència Artificial' implica:

‘Un software que és desenvolupat amb tècniques i mètodes específics i que pot, per a un conjunt donat d’objectius definits per un humà, generar sortides com a contingut, prediccions, recomanacions o decisions que influeixen l’ambient amb el qual interactua.’

El llistat de tècniques s’inclou en l’Annex 1 de la proposta de la Llei d’Intel·ligència Artificial i inclou:

a) Enfocaments d’aprenentatge automàtic, inclòs l’aprenentatge supervisat, no supervisat i de reforç, utilitzant una àmplia varietat de mètodes, inclòs l’aprenentatge profund;

b) Enfocaments basats en la lògica i el coneixement, inclosa la representació del coneixement, la programació (lògica) inductiva, les bases de coneixement, els motors d’inferència i deducció, el raonament (simbòlic) i els sistemes experts;

c) Enfocaments estadístics, estimació bayesiana, mètodes de cerca i optimització.

► **Machine Learning (Aprentatge Automàtic):** És un subcamp de la IA que es centra en el disseny de sistemes que poden aprendre de les dades. En altres paraules, en comptes de programar directament totes les accions que realitzarà el model (com succeïa en els anys 60 i 70), els sistemes d’aprenentatge automàtic s’entrenen amb un conjunt de dades, aprenent patrons que després s’utilitzen per fer prediccions o classificacions. Es consideren algoritmes de *Machine Learning* els algoritmes de regressió lineal, regressió logística, *k-nearest neighbours*, arbres de classificació o el *random forest*, entre molts altres.

► **Deep Learning (Aprentatge Profund):** És un subcamp de l’aprenentatge automàtic que es centra en les xarxes neuronals artificials amb moltes capes (o ‘profundes’). Les xarxes neuronals són un altre tipus d’algorisme dins del camp de la IA. Aquests models són capaços d’aprendre patrons complexos i subtils en les dades. S’utilitzen en una àmplia varietat d’aplicacions, incloent-hi el reconeixement de veu, la visió per computadora i el processament del llenguatge natural.

► **Algoritmes de classificació i predicció:** Els termes ‘classificació’ i ‘predicció’ s’utilitzen en el camp de l’aprenentatge automàtic per descriure dos tipus diferents de problemes que aquests algoritmes poden abordar.

► Els **algoritmes de classificació** s’utilitzen per predir/identificar la categoria o la classe d’un objecte o un esdeveniment. Un exemple de classificació podria ser la detecció de correu brossa, on es classifiquen els correus electrònics en ‘correu brossa’ o ‘no correu brossa’. O,

per exemple, un model que decideix si a una persona li correspon o no, una ajuda social. Hi ha diferents tipus d'algoritmes de classificació com la regressió logística, els arbres de decisió o el veí més proper (sovint abreujat k-NN).

- ▮ Els **algoritmes de predicció** se centren a predir un valor continu. Per exemple, es poden utilitzar per predir el preu d'un habitatge donades les seves característiques, o quant pujarà la inflació el pròxim mes. En aquest conjunt trobem algoritmes com la regressió lineal o els arbres de regressió.

▮ **IA Generativa o Models fundacionals:** Fa referència als sistemes d'IA que poden crear contingut nou i original. Això pot abastar des de la generació de text i la creació d'imatges fins a la música i altres tipus de contingut. El model GPT-4 de OpenAI, que està darrere de ChatGPT, és un exemple d'IA generativa en el camp del text.

▮ **IA simbòlica:** Fa referència als desenvolupaments de la primera onada (anys 60 i 70). La IA simbòlica es refereix als enfocaments per desenvolupar màquines intel·ligents mitjançant la codificació dels coneixements i l'experiència dels experts en conjunts de regles que poden ser executades per la màquina.

2.2 Un exemple de l'aplicació

Un cop definits els conceptes que s'utilitzaran al llarg d'aquest informe, i per acabar d'entendre com funciona aquesta tecnologia, a continuació posem un exemple senzill i hipotètic que ens serveixi per tenir una intuïció sobre la majoria de sistemes que utilitzen algoritmes de *Machine Learning*, responnent a la següent pregunta: Com s'entrenen aquests models d'aprenentatge automàtic o *Machine Learning*?

Suposem que una entitat del Tercer Sector porta a terme un programa de classes de reforç de matemàtiques i llengua per a alumnes en situacions socioeconòmiques complexes. Col·laboren amb més de 100 instituts als quals envien, en un primer moment, un únic professor de reforç. Quan es considera necessari, l'entitat envia un nou professor a un institut per poder millorar la qualitat i l'impacte del projecte.

No obstant això, el procés per decidir si enviar un professor extra de reforç és costós i comporta un temps considerable. El professor del projecte ha de complimentar un informe, experts de l'entitat han d'acudir a l'institut i avaluar el cas, etc. Amb l'experiència, l'entitat ha detectat que el procés és tan costós que no estan sent capaços de donar respostes a temps. Ar-

ribats a aquest punt, decideixen recolzar-se d'un sistema d'Intel·ligència Artificial.

Per sort, l'equip de dades de l'entitat ha anat recopilant informació dels diferents projectes, anotant el municipi on es troba l'institut, la renda per càpita del barri, la nota mitjana dels alumnes o la taxa d'abandonament escolar. A més, han inclòs una columna al seu Excel que indica si s'ha enviat o no un professor de reforç extra al projecte.

Projecte	Municipi de l'institut	Renta per càpita del barri de l'institut	Nota mitjana dels alumnes	Taxa de abandonament escolar de l'institut	Professor de reforç extra
1	A	25	7	7	No
2	B	10	5,6	43	Sí
3	B	8	8	33	Sí
4	C	17	7,2	20	No
...
50	D	16	7,5	15	No

A partir d'aquesta informació, l'equip de dades decideix entrenar un model de *Machine Learning* que predigui si és necessari o no assignar un professor extra de reforç a un projecte. Els analistes de dades no decidiran quin criteri seguir, sinó que deixaran que sigui el model qui 'aprengui' de les dades. És a dir, el model identificarà sota quines circumstàncies passades l'entitat va decidir enviar un nou professor de reforç. Aquest procés d'aprenentatge és el que es denomina com l'entrenament del model, de manera que el model aprendrà de les següents dades:

Variables explicatives				Outcome
Municipi de l'institut	Renta per càpita del barri de l'institut	Nota mitjana dels alumnes	Taxa de abandonament escolar de l'institut	Professor de reforç extra
A	25	7	7	No
B	10	5,6	43	Sí
B	8	8	33	Sí
C	17	7,2	20	No
...
D	16	7,5	15	No

La idea de l'entrenament rau en el fet que l'algoritme aprengui a predir el resultat (*outcome*) més probable a partir de les variables explicatives amb què compta. Per exemple, els projectes d'uns determinats municipis, amb una taxa d'abandonament per sota de cert valor i una nota mitjana concreta tendeixen a tenir assignat un professor de reforç.

Com saber si l'algoritme dona un bon resultat? L'equip de dades comparà els resultats reals (si el projecte té o no assignat un professor de reforç) amb els estimats per l'algoritme per a cada projecte. Si el model amb prou feines encerta un 30% dels casos, el model serà descartat ja que no aporta cap utilitat a l'entitat. L'equip de dades provarà amb diferents models i ajustos fins a donar amb aquell que aconsegeixi un major grau d'encert.

Un cop entrenat el model, l'equip de dades l'aplica sobre els restants 50 projectes pels quals no saben si enviar o no un professor de reforç. El model ha estimat que 5 instituts necessitarien un professor de reforç. Ara bé, l'entitat, en lloc d'haver analitzat els 50 projectes cas per cas per valorar en quin seria necessari comptar amb un professor extra, se centrarà en els 5 identificats pel model com a prioritaris.

Projecte	Municipi de l'institut	Variables explicatives			Outcome
		Renta per càpita del barri de l'institut	Nota mitjana dels alumnes	Taxa de abandonament escolar de l'institut	Professor de reforç extra
51	B	22	7,4	12	No
52	A	8	5	31	Sí
...	D
99	D	12	5,2	6	Sí
100	C	25	6,5	25	No

És important destacar que segons quin algoritme de *Machine Learning* s'hagi utilitzat, l'equip de dades sabrà explicar per què el model estima que es necessita un professor de reforç o no. Els models més precisos, és a dir, aquells que més encerten, solen ser més opacs, 'caixes negres', de manera que l'equip de dades tindrà dificultats per explicar a l'entitat el perquè del resultat. D'altra banda, els algoritmes més senzills són més explicables, però poden cometre molts més errors.

En conclusió, quin ha estat el procés seguit en aquest exemple hipotètic? En primer lloc, l'entitat va detectar una carència: alguns projectes podrien demandar un professor de reforç, però no compten amb recursos suficients per analitzar individualment cada cas.

En segon lloc, a partir de la informació d'altres projectes, per als quals van decidir en el seu dia si enviar-hi un professor de reforç o no, creen un model de *Machine Learning* capaç d'aprendre quines característiques són més rellevants a l'hora de prendre la decisió i, per tant, capaç de predir per si sol si, en altres casos on es compti amb les mateixes variables, s'hauria d'enviar un professor extra de reforç o no. Cap equip d'experts de l'entitat ha decidit quines característiques són les rellevants per prendre la decisió. Ha estat l'algoritme qui ha detectat els patrons per si mateix, aprenent què ha tingut en compte l'entitat en el passat a l'hora de prendre aquesta decisió.

Això permet poder estimar en pocs segons i per a tots els projectes si necessiten o no una persona de reforç, sense haver de dur a terme una avaluació individual i molt costosa de cada centre. L'entitat guanya en eficiència i efectivitat, podent posar el focus sobre aquells projectes que més ho necessiten.

3. L'ORIGEN DE LA INTEL·LIGÈNCIA ARTIFICIAL I ELS ÚLTIMS AVENÇOS

La història del desenvolupament de la Intel·ligència Artificial és una història d'anades i vingudes, de moments àlgids i moments d'estancament. Des de models més bàsics fins als sofisticats sistemes d'aprenentatge profund. Un avenç no lineal que ha 'explorat' en l'últim lustre. Ara bé, com hem arribat fins aquí?

Els orígens de la IA es remunten a la meitat del segle XX quan, **a la dècada de 1950**, Alan Turing va proposar la idea de les 'màquines de pensar'. El 1956 es va celebrar la Conferència de Dartmouth, on es va encunyar el terme 'Intel·ligència Artificial'. John McCarthy, Marvin Minsky, Nathaniel Rochester i Claude Shannon, organitzadors de la conferència, són considerats alguns dels pares fundadors de la IA.

Un dels primers models d'aprenentatge automàtic o *Machine Learning* va ser el perceptró. Un algoritme d'aprenentatge supervisat que servia per resoldre tasques senzilles de classificació. Tot i que inicialment va generar molt entusiasme, aviat es van fer evidents les seves limitacions a l'hora de resoldre problemes més complexos.

Als anys 60 i 70, la IA va experimentar el que es coneix com la seva primera 'època daurada', amb avenços en àrees com el processament del llenguatge natural i els sistemes experts. Durant aquesta època, els investigadors creien que estaven a les portes de crear màquines que podrien entendre i respondre al llenguatge natural, resoldre problemes complexos i aprendre de la seva experiència. No obstant això, aquests sistemes no 'aprenien' de la mateixa manera que ho fan les modernes tècniques de *Machine Learning*, com les xarxes neuronals. Es basaven en regles explícitament codificades per humans. No aprenien de les dades ni milloraven el seu rendiment amb l'experiència.

Aquestes limitacions van comportar, de nou, que l'entusiasme inicial comencés a esvaïr-se a **mitjans de la dècada dels 70**, donant lloc a 'l'hivern de la IA'. Com a resultat, es va reduir el finançament per a la investigació i el camp es va convertir en un tema menys popular en l'àmbit acadèmic.

La dècada dels 80 va marcar una fita important en el desenvolupament de les xarxes neuronals i, per extensió, de la Intel·ligència Artificial. Fins a aquest moment, l'aplicació dels algoritmes de xarxes neuronals s'havia vist limitat per les dificultats per entrenar-les de manera efectiva, en par-

ticular per a les xarxes de diverses capes o 'profundes'. No s'havia trobat una solució eficient a l'hora d'entrenar aquests models.

La situació va canviar quan es va produir un avenç crucial en la forma d'entrenar aquestes xarxes: l'algoritme de propagació cap enrere, o *backpropagation*, publicat per Rumelhart, Hinton i Williams el 1986. Aquest algoritme va obrir la porta a l'entrenament eficaç de xarxes neuronals més profundes i complexes, cosa que finalment va portar al desenvolupament de les tècniques d'aprenentatge profund que veiem avui.

No obstant això, tot i que l'algoritme de *backpropagation* va ser un avenç important, les limitacions en la capacitat de computació de l'època i la falta de grans conjunts de dades va impedir la generalització de les xarxes neuronals. En altres paraules, als anys 80 ja existien algoritmes d'IA amb un gran potencial, però escassejaven les bones dades i ordinadors prou potents.

Aquesta limitació va continuar durant **la dècada dels 90**, tot i que això no va significar que no es produïssin avenços importants. El *Machine Learning*, i especialment l'aprenentatge supervisat, va començar a guanyar més rellevància durant aquests anys. Es van desenvolupar i refinar una sèrie d'algoritmes, com les Màquines de Vectors de Suport (SVM), que van proporcionar un rendiment sòlid en una gran varietat de tasques.

Ara bé, no va ser fins a **la primera dècada dels anys 2000**, amb l'arribada de la computació d'alt rendiment i el *Big Data*, quan es van crear les condicions adequades perquè les xarxes neuronals, i per tant l'aprenentatge profund, comencessin a mostrar el seu veritable potencial. La popularització d'Internet i la proliferació de dispositius digitals van generar una gran quantitat de dades, proporcionant el combustible necessari per als algoritmes de *Machine Learning* i *Deep Learning*.

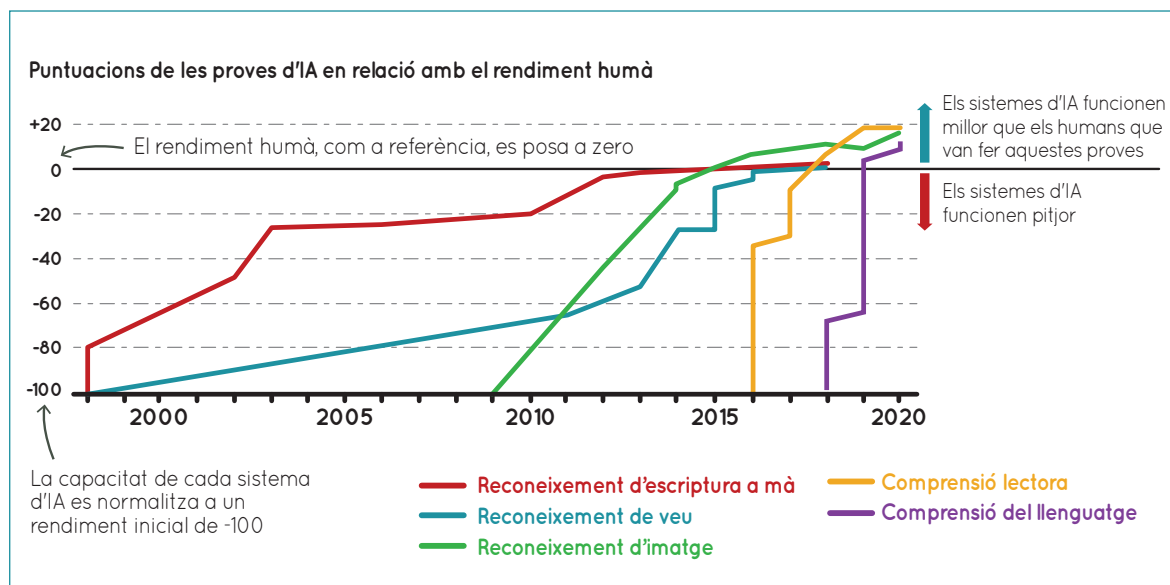
El 2006, Geoffrey Hinton va publicar un article que detallava com entrenar xarxes neuronals profundes, la qual cosa va marcar l'inici de l'era moderna de l'aprenentatge profund. Des de llavors, els avenços en la IA han estat ràpids i continus.

La dècada del 2010 va ser un període de gran avenç en el camp de la Intel·ligència Artificial. Les xarxes neuronals convolucionals (CNN), introduïdes per Yann LeCun, han transformat el camp del reconeixement d'imatges convertint-se en l'eina estàndard per a tasques de visió per computadora com el reconeixement d'objectes en imatges. En el camp del processament del llenguatge natural (NLP), les xarxes neuronals recurrents (RNN) van demostrar ser molt efectives i van permetre avenços significatius. A més, el 2017 es va produir un avenç que marcaria un abans i un després:

Google Research va publicar l'arquitectura de xarxa neuronal Transformer. Els Transformers s'han utilitzat per construir els models GPT i LaMDA, que estan darrere de les interfícies de ChatGPT d'OpenAI i BARD de Google. Models capaços de produir respostes i solucions al mateix nivell o superior a un humà.

El següent gràfic probablement sigui el millor resum possible de la ràpida evolució de la IA en els últims anys. Publicat per *Our World in Data* a l'article 'The brief history of artificial intelligence: The world has changed fast - what might be next?', mostra la puntuació màxima aconseguida per un model d'IA en diferents tasques al llarg del temps. Una puntuació igual a 0 equival a la puntuació obtinguda per un humà, de manera que un cop un model d'IA aconseguixi un resultat superior a 0, implica que ha superat les capacitats d'una persona.

Gràfic 1. Les capacitats de reconeixement d'imatges i llenguatge dels sistemes d'IA han millorat ràpidament.



Font: Extret de Roser (2022b)

L'any 2020, diferents models d'IA eren capaços de puntuar igual o millor que un humà en tasques com el reconeixement de veu, d'escriptura o d'imatges, la comprensió lectora i la comprensió del llenguatge. En aquests últims àmbits, l'avenç és sorprenentment ràpid, superant als humans en pocs anys.

Fa temps que la IA s'utilitza per a una miriada d'activitats. Fixa el preu del bitllet d'avió, monitoritza el nostre comportament als aeroports i assisteix

al pilot de l'avió durant el vol. Hi ha models d'IA que determinen la concessió d'un préstec, d'una ajuda o d'un lloc de treball. També s'utilitzen per transcriure i traduir text. Els assistents virtuals com Alexa estan a milers de llars. Les sèries, pel·lícules i vídeos que ens recomana Netflix o Youtube estan basats en IAs anomenades sistemes de recomanació. **La IA no és ni molt menys una tecnologia de futur, és una tecnologia de present que ja ha impactat en les nostres vides i que fa temps que ho fa.**

Segons l'informe de l'Observatori Nacional de Tecnologia i Societat (ONT-SI) sobre l'ús de la Intel·ligència Artificial de les empreses espanyoles, el 11,8% de les empreses amb més de 10 empleats feien ús d'aquesta tecnologia l'any 2022 (un 50% més que l'any anterior), i un 4,6% en el cas de les microempreses (un creixement aproximat d'un 30%). Espanya, el 2021, se situava en la mitjana de la UE27, tot i que lluny del nivell d'adopció dels països líders com Dinamarca (24%), Portugal (17%) o Finlàndia (16%).

Entre les empreses, destaca l'adopció de la IA per automatitzar fluxos de treball o ajudar en la presa de decisions (46,2% de les empreses) i per identificar objectes o persones en funció d'imatges (39,7%). Més d'un 30% utilitzen tècniques de processament del llenguatge natural, més d'un 20% utilitzen tecnologies per generar llenguatge escrit o parlat, i al voltant d'un 30% apliquen tècniques de *Machine Learning* (o aprenentatge automàtic). A més, aquests percentatges no deixaran de créixer. En l'estratègia Espanya Digital 2026 del Govern es fixa la meta que per a 2025, el 25% de les empreses espanyoles facin ús d'Intel·ligència Artificial i *big data*.

No obstant això, **no va ser fins a l'aparició de la IA generativa, i en concret de la famosa aplicació ChatGPT, que la IA va traspasar l'opinió pública.** ChatGPT va assolir 1.000.000 d'usuaris en només 5 dies. Assolir aquesta xifra va dur dos mesos i mig a Instagram, 5 mesos a Spotify, 10 mesos a Facebook i tres anys i mig a Twitter. Un èxit històric.

La capacitat del model GPT, i la seva fàcil interacció amb l'usuari a través de la interfície ChatGPT, ha sorprès el món sencer. Un model d'IA capaç de resoldre tasques per a les quals no va ser entrenat. Realitza traduccions raonablement bones, sintetitza informació, resol dubtes, genera codi en gairebé qualsevol llenguatge de programació, i un llarg etcètera. A més, la generació de text no és l'únic de què és capaç la Intel·ligència Artificial. La IA generativa inclou tant la generació de text, com d'imatges, sons o vídeos. En el camp de la generació d'imatges destaca DALL-E d'OpenAI, capaç de generar i/o editar imatges a partir de text, o Stable Diffusion, que a més és de codi obert.

Avui dia, tots aquests sistemes presenten certes limitacions. Les imatges generades poden contenir errors. ChatGPT pot al·lucinar informació⁴ o errar en càlculs matemàtics. No obstant això, no convé perdre la perspectiva temporal: aquesta és la primera generació d'aquests models, creats en només uns pocs anys. És possible que, malgrat tractar-se de limitacions complexes, puguin quedar resoltes amb el temps. Probablement aleshores serà quan podrem avaluar el veritable potencial dels sistemes d'IA.

Així doncs, al llarg de les dècades la IA ha passat de ser un somni a una realitat que toca gairebé cada aspecte de les nostres vides. El reconeixement d'imatges, la transcripció d'àudios, la traducció de textos, realitzar tasques de predicció o la generació de text o vídeos són àmbits que han quedat dominats per models d'Intel·ligència Artificial.

Com no pot ser d'una altra manera, i com veurem en capítols posteriors, la IA també es pot aplicar en el context de l'Administració Pública, oferint oportunitats per millorar l'eficiència, la precisió i l'abast dels serveis públics. Però també generant riscos i desafiaments. **Entendre el context present és fonamental per planificar les accions de futur en l'àmbit del tercer sector.**

4 En el context de models de llenguatge, 'al·lucinar informació' fa referència al fet que el model pot generar o proporcionar detalls que no estan recolzats per dades reals o informació veritable que hagi après. Això pot ser causat per la pròpia naturalesa de l'entrenament del model, atès que no està enfocat a 'memoritzar' dades ni fets, sinó que el model veu vastes quantitats de text i aprèn a predir la paraula següent en una seqüència.

4. GOVERNANÇA I ESTRATÈGIES NACIONALS EN L'ÀMBIT DE LA INTEL·LIGÈNCIA ARTIFICIAL

Com es prenen les decisions? Estan alineades amb valors ètics? Qui les pren? Com s'implementen i qui les supervisa? En definitiva, per què és important parlar de la governança de la IA? D'una banda, la Intel·ligència Artificial té el potencial per transformar les nostres societats. **Aplicar una governança adequada en el moment correcte pot contribuir a aconseguir que la transformació sigui positiva, prevenint i minimitzant els riscos i potenciant els beneficis.** D'altra banda, tot sistema d'IA que sigui utilitzat per una empresa o institució pública hauria de poder estar sotmès a un marc universal que aprovés una adequada utilització.

Per aquests motius, després d'haver indagat en el desenvolupament i evolució de la IA, és indispensable conèixer com està sent la interacció entre els governs, empreses, societat civil i ciutadania i, per tant, com s'està configurant l'ambient regulador i discursiu al voltant de la IA.

La governança de la Intel·ligència Artificial s'executa en diferents nivells i àmbits: pot desenvolupar-se a escala internacional o nacional, pot provenir i estar dirigida exclusivament al sector privat, al sector públic o a tots dos. Engloba aspectes com la definició d'estàndards ètics, la promulgació de lleis i regulacions, i la implementació de sistemes de supervisió i control per prevenir i mitigar possibles danys o abusos.

En aquest sentit, existeixen diferents tipus d'iniciatives de governança global en l'àrea de la IA. Algunes d'aquestes iniciatives parteixen de les empreses privades. Per exemple, a Microsoft han impulsat l'Oficina d'IA Responsable (ORA) i el Comitè d'AI i Ètica en Enginyeria i Investigació (Aether) per garantir internament que els sistemes d'IA es desenvolupin de forma responsable. O la Partnership on AI fundada el 2016 per Amazon, Facebook, Google, DeepMind, Microsoft, IBM i Apple (que va entrar el 2017) que agrupa més de 100 socis de l'àmbit acadèmic, societat civil i indústria, en una organització sense ànim de lucre que advoca per la recerca en bones pràctiques i l'ús responsable de la IA. No obstant això, en els darrers temps s'estan produint canvis. Companyies com Microsoft, Twitter o Twitch estan prescindint dels treballadors que estudiaven el costat negatiu de la IA.⁵

⁵ Informació extreta de l'article del The Washington Post, *As AI booms, tech firms are laying off their ethicists*, disponible a: <https://www.washingtonpost.com/technology/2023/03/30/tech-companies-cut-ai-ethics/>.

També poden provenir d'organismes internacionals com les Recomanacions en IA de l'OCDE del 2019, o les Recomanacions sobre l'ètica de la Intel·ligència Artificial de la UNESCO del 2021. En tots dos casos no deixen de ser normatives no vinculants que fonamentalment al·ludeixen a principis ètics i morals.

No obstant això, pel que fa al poder, estem en un sector amb una concentració considerablement elevada. Un reduït nombre d'empreses controlen els recursos que faciliten la creació dels sistemes d'IA, incloent-hi els recursos físics, epistèmics i informàtics. A més, la forma en què es distribueix la IA, normalment en format de plataformes, les converteix en uns actors importants. Aquesta circumstància proporciona un paper clau a les principals firmes, que poden tractar d'impulsar les expectatives, els marcs de debat i els discursos sobre el present i el futur de la IA, la visió, els interessos i les propostes.

El principal risc d'aquesta situació és que s'implementin tecnologies que, per manca de visions crítiques, produeixin més riscos que beneficis. **Per aquest motiu és fonamental que altres actors, com el Tercer Sector, s'organitzin, coordinin i fixin una postura conjunta.** Una postura a favor dels avenços i millores que pugui portar la tecnologia, però previsorament davant dels seus riscos i que ofereixi garanties en la seva aplicació.

Fins ara, dos gegants tecnològics han pres clarament la davantera. Microsoft s'ha aliat amb OpenAI. Aquest laboratori és el precursor de l'eina més famosa, ChatGPT, però també d'altres com DALL-E 2 (capaç de generar imatges a partir d'unes poques paraules) i Whisper (transcripció d'àudio en anglès a un nivell humà). Després d'invertir 10.000 M€ en OpenAI, Microsoft està integrant la IA en tots els seus productes: Windows, Word, etc. Per la seva banda, Google ha respost aliant-se amb DeepMind, una empresa especialitzada en IA amb seu a Londres. Google ha publicat l'eina Bard, competidor directe de ChatGPT. Dos gegants tecnològics i dos laboratoris de recerca d'IA que (per ara) estan liderant la competició.

Tot i això, en els últims mesos Meta ha irromput amb força. Ha publicat models de generació de llenguatge, generació d'imatge, generació de text a veu i viceversa i generació de música. La gran diferència de Meta respecte als seus competidors és que ha apostat per ser *open source* (de codi obert), de manera que qualsevol programador del món pot accedir al model i al codi, i pot contribuir a millorar-lo i desenvolupar-lo. Això augmenta considerablement la capacitat per crear models més eficients, més robustos i més segurs.

4.1 El rol dels Estats en la governança de la IA

Ara bé, quin rol estan prenent els Estats? Un dels canvis fonamentals que s'està produint és el paper que estan assumint els països. Els Estats de la Unió Europea estan liderant aquest nou rol, tant pel seu impuls conjunt a la regulació de la IA com per les publicacions de les seves estratègies nacionals en aquest àmbit.

En un estudi recent⁶, els països eren classificats en funció de l'enfocament i accions per les quals es decantaven en les seves estratègies nacionals:

- › Els **països desenvolupadors**: prioritzen un rol de l'Estat com a facilitador i iniciador dels projectes d'IA a través de la innovació i l'assignació directa de recursos. Aquest enfocament seria el predominant entre els països del bloc postsoviètic i d'Àsia de l'est, on l'Estat assumiria el rol principal en el desenvolupament.
- › Els **països que assumeixen un rol de promoció**: atorguen el lideratge al sector privat. Països com els Estats Units, el Regne Unit o Irlanda prenen un enfocament descentralitzat i deleguen el rol clau al sector privat, no implicant l'Estat i apostant per l'autoregulació de la indústria. Tot i amb enfocaments diferents, aquests països prioritzarien la innovació per sobre de la protecció.
- › Els **països que aposten pel control**: els països de la Unió Europea situen el centre d'atenció sobre el control, reflectint la forta posició de la unió pel que fa a la regulació. Els Estats volen garantir la protecció de la societat dels riscos de la IA mitjançant rigorosos marcs reguladors, prioritzant la protecció per sobre de la innovació.

El rol de l'Estat és rellevant perquè obre o tanca finestres d'oportunitat per a actors com les organitzacions del tercer sector. En un context en què el Govern delegui el lideratge en el sector privat, serà més difícil per a les entitats que la seva visió i perspectiva sigui influent. No obstant això, aquest no és el context europeu, la qual cosa **obre potencialment la porta a les entitats per fer arribar tant les seves preocupacions com fer contribucions significatives.** No només en termes legislatius, sinó també de discurs.

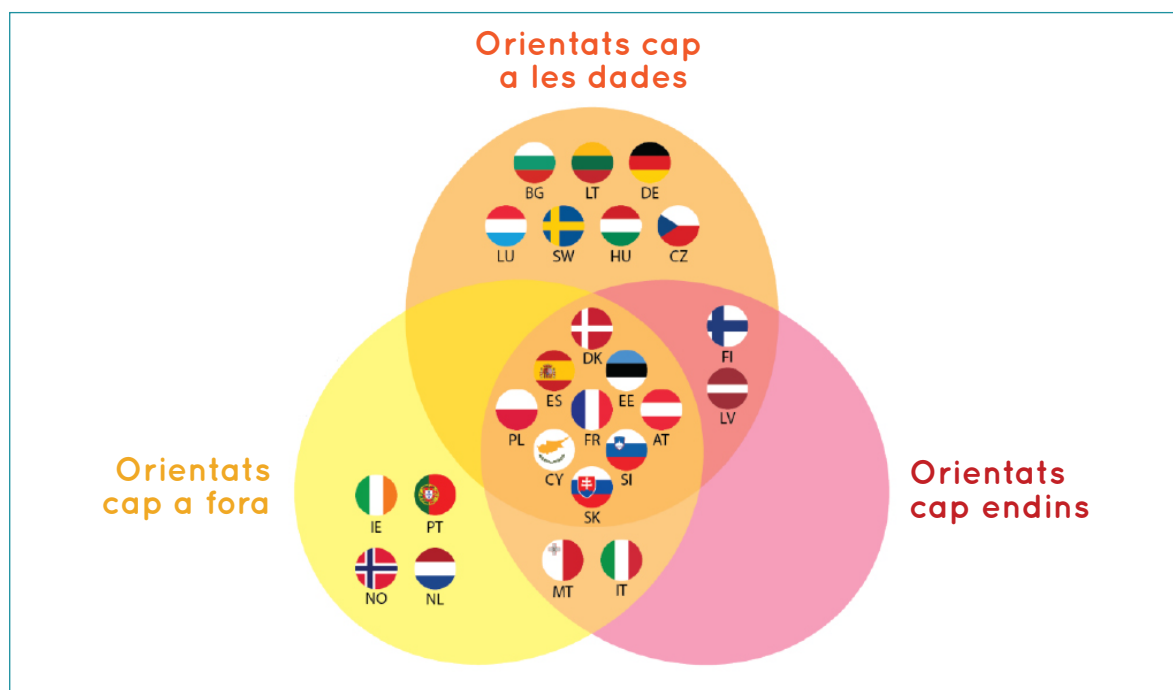
6 Gleb Papyshv & Masaru Yarime (2023) The state's role in governing artificial intelligence: development, control, and promotion through national strategies. *Policy Design and Practice*, 6:1, 79-102, DOI: 10.1080/25741292.2022.2162252

4.2 El rol dels Estats en l'aplicació de la IA en el Sector Públic

L'informe *European Landscape on the Use of Artificial Intelligence by the Public Sector*, elaborat pel servei de supervisió de desenvolupament i impacte de la IA de la Comissió Europea, l'Ai Watch, analitza 24 estratègies nacionals dels països de la UE (més Noruega) i les classifica en tres tipus ideals en funció del grau d'implicació, accions i iniciatives del sector públic. Aquest tipus d'exercicis són molt útils atès que permeten posicionar els governs de forma relativa a les accions de la resta. Saber quin paper està adoptant el nostre país és necessari per comprendre quin paper pot jugar el Tercer Sector en la governança de la IA.

La primera tipologia d'Estats són aquells **orientats cap a fora (externally-oriented)**. Aquests governs posen el focus en la cooperació públicoprivada per desenvolupar la IA en el sector públic. Assumeixen que l'estat no té suficient capacitat i competències i que els sistemes són massa complexos per a enfrontar-se a la IA de forma autònoma. Per aquest motiu, posen un major èmfasi en crear un ecosistema de start-ups i empreses que desenvolupin IA per al sector públic, promovent la col·laboració públicoprivada. En aquest grup es troben països com Irlanda, Portugal, Noruega i els Països Baixos.

Gràfic 2. Exercici d'agrupació de les estratègies nacionals.



Font: Extret de Tangi et al (2022)

D'altra banda, els **països orientats cap a les dades (*data-oriented*)** prioritzen facilitar l'accés i millorar la qualitat de les dades, eliminant les barreres i millorant les infraestructures per al desenvolupament de la IA en el sector públic. Tornar disponibles més bases de dades, fomentar la compartició de dades entre institucions públiques o millorar les pràctiques de recollida i govern de les dades serien les seves accions principals. Entre els països que encaixen en aquest tipus ideal hi ha Alemanya o Suècia.

Els Estats amb una estratègia **orientada cap endins (*internally-oriented*)** són aquells que se centren a millorar la capacitat interna de l'Estat com a instrument per estimular el desenvolupament de la IA. Crear noves unitats o institucions públiques especialitzades en IA, millorar el capital humà del funcionariat mitjançant formació i augmentar el coneixement intern sobre la IA mitjançant esdeveniments serien les seves accions predilectes. No obstant això, cap país aposta exclusivament per aquestes mesures com a solució única. Per exemple, Finlàndia i Letònia ho combinen amb una estratègia orientada cap a les dades.

Finalment, fins a 9 dels Estats analitzats, entre els quals es troba Espanya, integren els 3 enfocaments en les seves estratègies nacionals. Per tant, **existeix una clara aposta en la majoria dels països per treballar conjuntament amb el sector privat en l'impuls de la IA en el sector públic, la qual cosa és raonable, ja que és on està la major part de capital humà i tecnològic en l'actualitat.**

4.3 L'Estratègia Nacional d'Espanya

En el cas particular d'Espanya, el Govern Estatal va publicar la seva Estratègia Nacional d'Intel·ligència Artificial al desembre del 2020, amb l'objectiu de desenvolupar un marc polític que defineixi les diverses accions que les administracions governamentals emprendran per facilitar el desenvolupament i desplegament de la IA en l'economia i la societat.

A l'estratègia nacional destaquen objectius vinculats a la formació i la investigació, com és l'impuls al desenvolupament del capital humà en IA mitjançant formació, atracció de talent i ocupació qualificada, juntament amb el desenvolupament d'una excel·lència científica per promoure Espanya com a país líder.

També es fixa com a objectiu impulsar el desplegament i ús de tecnologies d'IA tant en el sector públic com en el privat, i situar Espanya com a líder en el desenvolupament d'eines, tecnologies i aplicacions per a la projecció i ús de l'idioma espanyol en IA.

Al mateix temps que es promou el desenvolupament del capital humà i tecnològic, també es fixen objectius ètics, com garantir un marc ètic que delineï els drets individuals i col·lectius i construeixi un entorn de confiança, així com garantir la inclusió en l'economia impulsada per la IA –inclusió de gènere, digital o territorial.

Pel que fa a mesures directament relacionades i amb un impacte en el sector públic, l'Estratègia Nacional d'Espanya presenta iniciatives polítiques en totes les àrees analitzades per AI Watch:

- (1) Augmentar el coneixement dels empleats públics sobre la IA;
- (2) Millorar la capacitat interna, formant a les funcions en IA i creant nous departaments especialitzats.
- (3) Mesures per millorar la qualitat i l'accés a les dades.
- (4) Pel que fa a l'ètica i al marc legal, es vol desenvolupar tant un marc ètic com reformar les lleis per a l'intercanvi de dades.
- (5) Pel que fa al finançament i processos burocràtics, hi ha un compromís per finançar projectes d'IA, estimular les *start-ups* del *GovTech* i revisar els processos burocràtics per agilitzar-los.
- (6) En l'àrea del *learning by doing*, Espanya vol regular i liderar la implementació dels *sandboxes*.

Per tant, independentment de la traducció definitiva en iniciatives polítiques i marcs normatius, i sempre que no canviïn les prioritats o els partits de govern, **existeix a priori una clara aposta per a que Espanya sigui líder en el desenvolupament i aplicació de la Intel·ligència Artificial, tant en el sector privat com públic.**

5. L'ENFOCAMENT CENTRAT EN LA PERSONA I LA LLEI DE LA INTEL·LIGÈNCIA ARTIFICIAL DE LA UNIÓ EUROPEA

En el capítol 3 s'ha exposat el desenvolupament històric i evolució de la Intel·ligència Artificial, des de mitjan segle XX fins avui. Després d'indagar en aquesta perspectiva, el focus del capítol 4 s'ha posat en discernir com els governs, les empreses i la societat civil estan interactuant entre si, assentant les bases de la governança de la IA del futur, i, particularment, quina aplicació de la IA a les Administracions Públiques estan seguint els Estats.

El desenvolupament de qualsevol sector econòmic està fortament determinat pel marc regulador en el qual opera. En l'àmbit de la Intel·ligència Artificial la llei que més debat està generant pel seu potencial per impulsar (o soscarar) el desenvolupament de la IA als països europeus és la Llei Europea d'Intel·ligència Artificial.

Aquesta Llei parteix d'una visió molt concreta: l'enfocament centrat en la persona. Una visió que advoca per un progrés 'útil' per a la persona. És a dir, entén la IA com un mitjà per millorar el benestar, i no com un fi en si mateix. En aquest apartat s'exposa aquest enfocament per, un cop comprès el punt de partida, explicar les claus de l'esborrany de Llei que ja ha estat aprovat.

5.1 L'enfocament centrat en la persona

El 2018, la Comissió Europea va crear el Grup d'Experts d'Alt Nivell en Intel·ligència Artificial amb l'objectiu de desenvolupar dos informes: (1) una guia ètica sobre la IA (AI Ethics Guidelines) i (2) recomanacions de polítiques públiques i inversió.⁷

El grup va publicar un primer esborrany de la guia ètica el desembre de 2018, que va ser sotmès a consulta amb els stakeholders i els estats membres. L'abril de 2019 es va publicar el document definitiu: Directrius Ètiques per a una IA Fiable (Ethics Guidelines For Trustworthy AI). Aquest document, que fixa també la posició de la Comissió Europea, transmet un mis-

⁷ Els informes del grup d'experts es poden consultar en: <https://digital-strategy.ec.europa.eu/es/policies/expert-group-ai>

satge clar: la IA no és un fi en si mateixa, sinó una eina que ha de servir a les persones amb l'objectiu últim de millorar el benestar. En altres paraules, **el desenvolupament i la implementació de la IA s'han de fer posant les persones al centre. Per aconseguir-ho, la confiança en la IA és un prerequisit.**

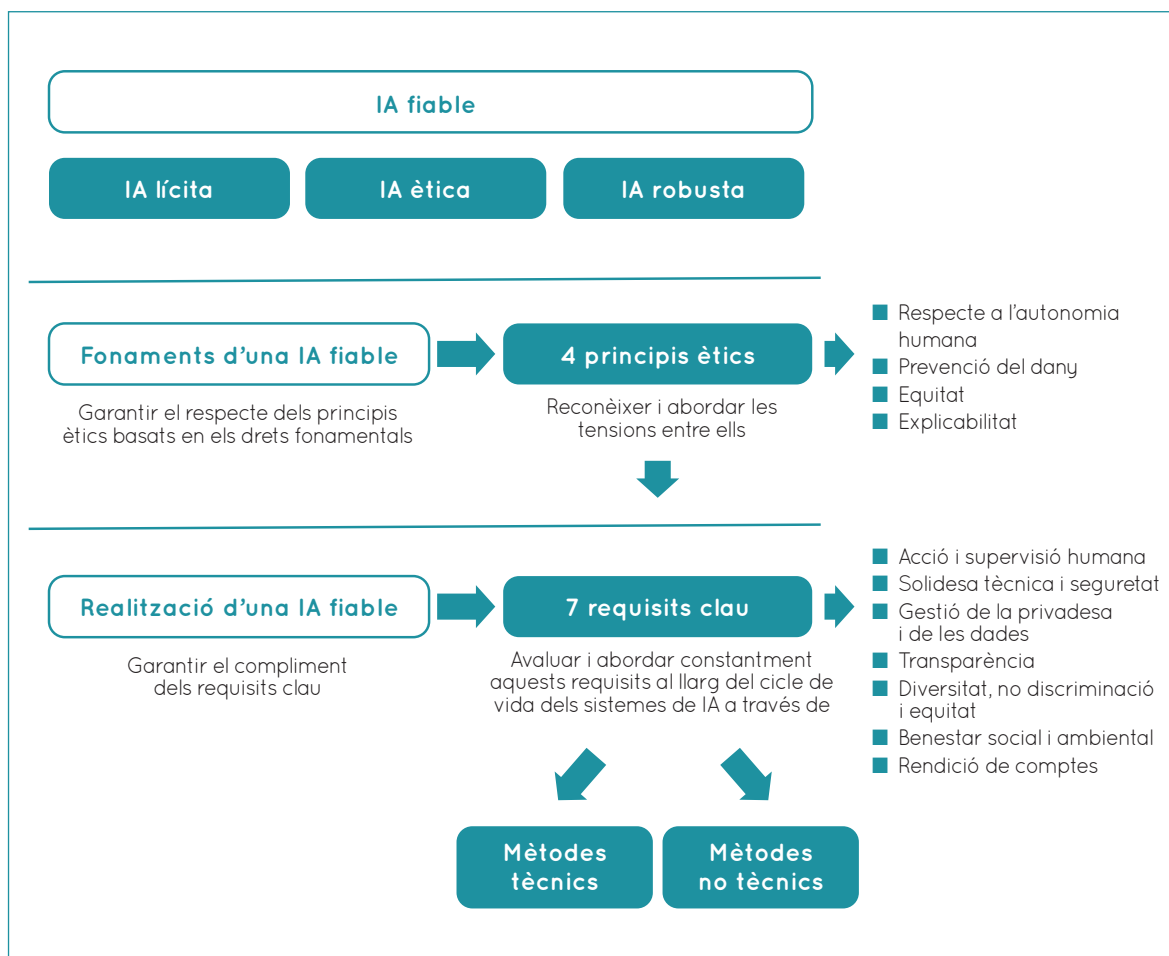
La IA té el potencial de transformar significativament la societat, millorant el benestar individual i social. Però, perquè això succeeixi, els sistemes d'IA han d'estar centrats en la persona i han de basar-se en un compromís amb el seu ús al servei de la humanitat i el bé comú. El moment present és clau perquè ens trobem davant d'una important finestra d'oportunitat per donar forma al futur (i present) desenvolupament de la Intel·ligència Artificial, el que implica tant tractar de maximitzar els seus beneficis com prevenir i minimitzar els seus riscos.

Ara bé, com es pot garantir el desenvolupament d'una IA centrada en les persones? La proposta del grup de treball comença per establir quins drets fonamentals i principis ètics hauria de respectar qualsevol sistema que utilitzi Intel·ligència Artificial. Entre ells hi hauria el respecte a la llibertat individual, la democràcia, la igualtat o l'Estat de Dret. A partir de tots els drets i principis identificats, el grup de treball resol que una IA és fiable si compleix tres components essencials:

-) Lícita** (*lawful*): compleix totes les lleis i regulacions aplicables.
-) Ètica** (*ethical*): s'adhereix a principis i valors ètics.
-) Robusta** (*robust*): és robusta tant des d'una perspectiva tècnica com social.

Una IA lícita, ètica i robusta serà una IA fiable, que respecti els drets fonamentals i els principis ètics esmentats. Aquests 3 components abstractes s'aterren en 7 requisits que tot sistema d'IA hauria de complir, i es resumeixen en el següent gràfic i descriuen a continuació:

Gràfic 3. Marc per a una IA fiable



Font: Extret i adaptat de HLEGAL (2019)

- (1) **Acció i supervisió humanes:** el benestar de l'usuari ha de ser central per a la funcionalitat del sistema i la supervisió humana ha d'estar assegurada.
- (2) **Solidesa tècnica i seguretat:** una IA fiable requereix algoritmes segurs, fiables i robustos que facin front als errors i les inconsistències que puguin sorgir, així com contra possibles atacs.
- (3) **Privacitat i governança de dades:** el respecte de la privacitat, la qualitat i la integritat de les dades ha d'estar garantit.
- (4) **Transparència:** inclou la traçabilitat, l'explicabilitat i la comunicació.
- (5) **Diversitat, no discriminació i equitat:** inclou la prevenció de prejudicis injustos, l'accessibilitat i el disseny universal, i la participació de les parts interessades.

- (6) Benestar social i ambiental: els impactes han de ser considerats en termes socials i no només individuals, per la qual cosa s'han de tenir en compte tant els seus impactes socials com mediambientals.
- (7) Rendició de comptes: han d'existir mecanismes que assegurin la rendició de comptes dels resultats, abans i després de la seva implementació. Les auditories (externes i internes) són fonamentals en aquest aspecte. També s'han d'identificar els potencials impactes negatius, així com garantir les compensacions i reparacions.

Per tant, la visió de la Comissió Europea es pot resumir en la següent idea: **atesa la capacitat transformadora dels sistemes d'Intel·ligència Artificial, ens trobem davant d'una finestra d'oportunitat per definir com ha de desenvolupar-se i implementar-se, de manera que ens assegurem que els sistemes d'IA futurs situaran sempre les persones en el centre, orientant els avenços cap a la millora del benestar individual i social, i prevenint des d'ara mateix els seus potencials usos i dissenys amb conseqüències negatives.** Des d'aquesta visió s'està desenvolupant la Llei d'Intel·ligència Artificial de la Unió Europea.

5.2 La Llei de la IA de la Unió Europea

El juny de 2023 es va aprovar al Parlament Europeu l'esborrany de la Llei d'Intel·ligència Artificial de la Unió Europea (EU IA Act), una Llei que havia estat preparant-se durant dos anys. Ara han començat les negociacions entre els Estats, durant les quals Espanya podria exercir un paper clau sota la presidència del Consell de la Unió Europea, i s'espera que a finals d'any es publiqui el text definitiu.

Aquest text legislatiu no és rellevant únicament per l'impacte directe que podria tenir sobre el nostre país, sinó perquè és un cas de referència a escala mundial. La regulació que està desenvolupant la Unió Europea (i també la Xina) servirà a la resta de països per prendre nota tant dels encerts com dels errors.

La regulació de la IA ha estat una qüestió política central a la Unió en els últims anys i aquest esborrany és el resultat d'aquest esforç. Després de la regulació 'tova' no vinculant de 2019 a través de les Directrius Ètiques per a una IA Fiable, la Comissió va passar al 2021 cap a un enfocament legislatiu fent una crida a l'adopció d'un nou marc regulador. Partint d'un enfocament que situa les persones en el centre, és a dir, de garantir el desenvolupament d'una IA fiable, es busca crear una regulació que com-

pleixi dos objectius paral·lels: (1) promoure el desenvolupament de la IA i (2) gestionar els seus riscos associats.

En altres paraules, aquesta Llei proposa regular els usos de la IA i no la IA en si mateixa, buscant aconseguir que la regulació no quedi obsoleta amb els avenços de la tecnologia. Ara bé, abans d'entrar en aquesta qüestió, com es defineix la Intel·ligència Artificial? En el punt 1) Article 3 es defineix com 'sistema d'Intel·ligència Artificial':

El software que es desenvolupa emprant una o diverses de les tècniques i estratègies que figuren a l'annex I i que pot, per a un conjunt determinat d'objectius definits per éssers humans, generar informació de sortida com a continguts, prediccions, recomanacions o decisions que influeixin en els entorns amb els quals interactua.

Aquests sistemes d'IA especificats a l'Annex I són:

- a) Les estratègies d'aprenentatge supervisat, inclosos l'aprenentatge supervisat, el no supervisat i el realitzat per reforçament.
- b) Les estratègies basades en la lògica i el coneixement, especialment la representació del coneixement, la programació (lògica) inductiva, les bases de coneixement, els motors d'inferències i deducció, els sistemes experts i de raonament (simbòlic).
- c) Les estratègies estadístiques, estimació bayesiana, mètodes de cerca i optimització.

Per tant, els productes comercials d'empreses privades i aplicacions d'IA en l'Administració Pública que facin ús d'algunes d'aquestes estratègies quedaran regulades (en principi) per aquesta llei.

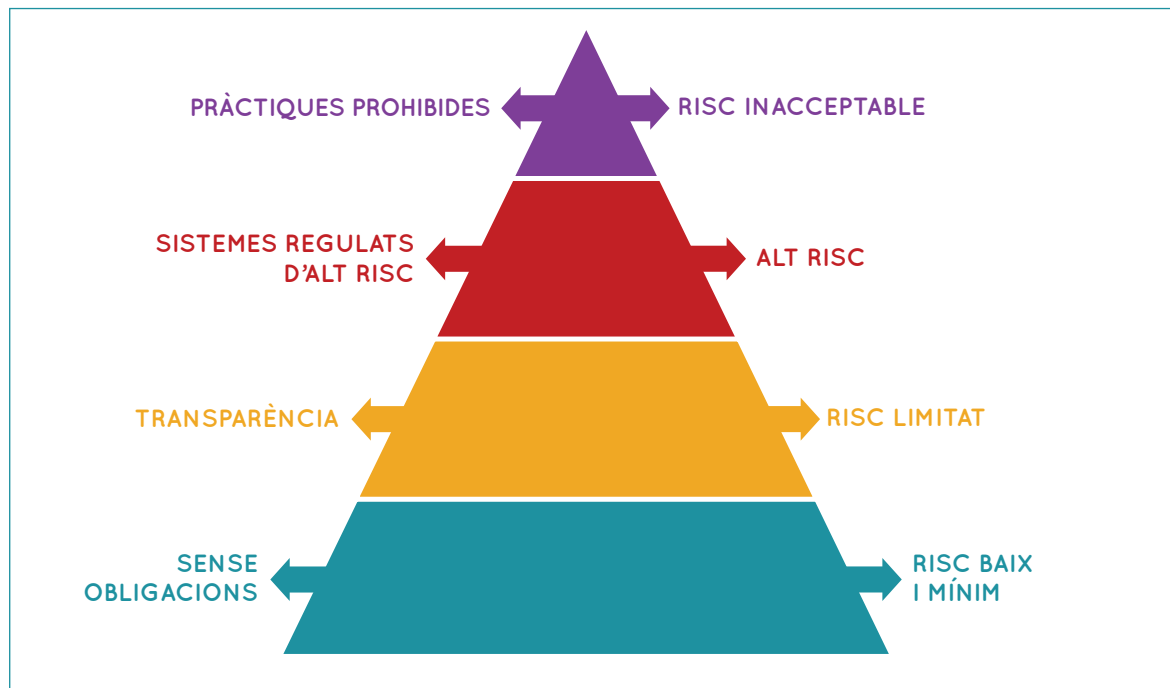
Els objectius específics del reglament són:

- 1) Garantir que els sistemes d'IA introduïts i utilitzats en el mercat de la UE siguin segurs i respectin la legislació vigent en matèria de drets fonamentals.
- 2) Garantir la seguretat jurídica per facilitar la inversió i innovació en IA.
- 3) Millorar la governança i l'aplicació efectiva de la legislació vigent en matèria de drets fonamentals i els requisits de seguretat aplicables als sistemes d'IA.
- 4) Facilitar el desenvolupament d'un mercat únic per fer un ús legal, segur i fiable de les aplicacions d'IA i evitar la fragmentació del mercat.

Com s'ha comentat, aquesta legislació es basa en una classificació dels riscos que la Intel·ligència Artificial podria suposar per a la seguretat, la salut o els drets fonamentals d'una persona en cada cas, i determina obli-

gacions per a proveïdors i usuaris en funció dels riscos que suposin els usos de la IA.

Gràfic 4.



Font: Llei d'Intel·ligència Artificial de l'UE

En primer lloc, els **sistemes d'IA de risc baix o mínim** no estaran subjectes a cap obligació, i podran ser desenvolupats i utilitzats en la UE sense més requisits legals. No obstant això, la Llei proposa la creació de codis de conducta.

En segon lloc, estarien els **sistemes d'IA de risc limitat**. Els responsables estarien sotmesos a unes obligacions generals de transparència. Per exemple, hauran de presentar informació per comunicar l'ús d'un sistema d'IA quan aquest interactuï amb humans, o si s'utilitza per generar o manipular imatges, àudios o vídeos ha de ser obligatori informar que s'ha generat per mitjans automatitzats.

En tercer lloc, es troben **els sistemes d'IA d'alt risc**. En aquests casos l'esborrany cita explícitament que 'els requisits relatius a les dades d'alta qualitat, la documentació i la traçabilitat, la transparència, la vigilància humana, la precisió i la solidesa són estrictament necessaris per reduir els riscos de la IA per als drets fonamentals i la seguretat'.

Els sistemes considerats d'alt risc s'especifiquen en el punt 1 de l'article 6 del Títol III i a l'Annex III. Sense enumerar tots els casos possibles, alguns exemples són:

-) Sistemes utilitzats com a components de seguretat d'un producte o com un producte que cau sota la legislació harmonitzada de salut i seguretat de la Unió (joguines, aviació, cotxes, etc.).
-) La identificació biomètrica i categorització de les persones en temps real o de forma retroactiva.
-) Sistemes utilitzats per a la gestió i operació d'infraestructura crítica (trànsit de carreteres, subministrament d'aigua, gas, calefacció i electricitat).
-) Sistemes utilitzats per avaluar els estudiants i determinar l'accés a institucions educatives.
-) Sistemes utilitzats per seleccionar o descartar treballadors o avaluar candidats.
-) Sistemes utilitzats per avaluar el risc d'incidència criminal d'una persona.
-) Sistemes utilitzats per les autoritats públiques per verificar l'autenticitat de documents de viatge.

Tots aquests sistemes s'estructuren en vuit àrees que podrien ser actualitzades si fos necessari: la identificació biomètrica i la categorització de persones; la gestió i operació d'infraestructura crítica; l'educació i formació; l'ocupació i gestió del treball; l'accés a serveis privats essencials i serveis públics; compliment de la Llei; migració, asil i fronteres; i administració de justícia.

Aquests sistemes estaran sotmesos a tota una sèrie de noves regles. Els proveïdors d'aquests sistemes hauran de registrar-los en una base de dades de la UE gestionada per la Comissió abans de comercialitzar-los per a ser sotmesos a una avaluació de conformitat *ex ante*. A més, hauran de complir una sèrie de requisits i obligacions, en particular, en matèria de gestió de riscos, proves, solidesa tècnica, formació i transparència, supervisió humana i ciberseguretat (articles 8 a 15).

Finalment, estan els **sistemes d'IA que suposen un risc inacceptable**. Aquests sistemes estan directament prohibits, ja que el seu ús es considera inacceptable per ser contrari als valors de la Unió. Queden prohibits els següents usos:

- 1) L'ús de tècniques subliminals que transcendeixin la consciència d'una persona per alterar el seu comportament o provocar perjudicis físics o psicològics.
- 2) La IA que aprofiti alguna de les vulnerabilitats d'un grup específic de persones a causa de la seva edat o discapacitat física o mental per alterar el seu comportament.

- 3) La utilització de sistemes d'IA per part de les autoritats públiques amb la finalitat d'avaluar o classificar la fiabilitat de persones físiques.
- 4) L'ús de sistemes d'identificació biomètrica remota 'en temps real' en espais d'accés públic (excepte en casos concrets com la recerca de víctimes d'un delictes o prevenció d'una amenaça)

D'altra banda, com a mesura de suport a la innovació, a l'article 53 es recull l'establiment d'un entorn controlat de proves d'IA (*sandbox*) que serà establert per un o més estats membres o pel Supervisor europeu de protecció de dades. Aquest entorn facilitarà el desenvolupament, prova i validació dels sistemes d'IA abans de ser introduïts en el mercat. En aquesta qüestió, Espanya vol posicionar-se com a país de referència, tal com es veurà al pròxim apartat.

Finalment, la Llei proposa la creació d'un Comitè Europeu d'Intel·ligència Artificial que actuaria com a consell assessor de la Comissió amb l'objectiu de contribuir a la cooperació efectiva entre les autoritats nacionals i la Comissió, coordinar i contribuir a les anàlisis sobre qüestions del mercat intern, i assistir a les autoritats nacionals i a la Comissió sobre l'aplicació d'aquesta regulació. Aquest Comitè estaria compost per representants de les autoritats nacionals. Aquestes autoritats hauran d'estar establertes o designades per cadascun dels Estats.

5.3 Les accions d'Espanya

Fins ara, a més dels documents polítics com l'Estratègia Nacional, Espanya ha aprovat dos decrets d'especial interès per visualitzar els passos que està donant el país en aquest àmbit. L'estiu del 2023 es van aprovar dos decrets rellevants: d'una banda, aquell que regula l'aplicació del *sandbox* a Espanya i, de l'altra, els estatuts de l'Agència supervisora.

El Reial decret de *sandbox* d'Espanya

El juny de 2022 el Govern d'Espanya i la Comissió Europea van presentar el projecte pilot per posar en marxa el primer *sandbox* regulador de la UE sobre IA, que serà finançat mitjançant els Fons de Recuperació i Resiliència amb 4,3 milions d'euros fins al 2025.

Un any després, el juny de 2023, el Govern va publicar l'esborrany del Reial decret regulador del *sandbox* d'Intel·ligència Artificial. L'objectiu és proporcionar a les empreses, especialment a les petites i mitjanes empreses i les *startups*, un entorn controlat de proves que garanteixi el desenvolupament d'una tecnologia responsable i centrada en la persona, mitigant els riscos potencials per a la salut, la seguretat i els drets fonamentals.

La participació en aquest entorn controlat de proves se circumscriu a sistemes d'IA que siguin classificats d'alt risc, a sistemes d'IA de propòsit general, models fundacionals i sistemes d'IA generativa. D'altra banda, s'exclou la participació de sistemes amb finalitat científica, de defensa o de seguretat nacional.

L'objectiu final del *sandbox* és **que les empreses que estiguin desenvolupant aquest tipus d'aplicacions tinguin un entorn on provar-les i avaluar-les, de forma que s'asseguri que s'estan complint tots els requisits que exigirà el futur Reglament europeu de la IA de cara a poder comercialitzar el seu producte.**

L'Estatut de l'Agència Espanyola de Supervisió de la Intel·ligència Artificial (AESIA)

L'agost del 2023 s'aprova el Reial decret que dona el vistiplau a l'estatut de l'Agència Espanyola de Supervisió d'Intel·ligència Artificial (AESIA), adscrita al Ministeri d'Afers Econòmics. Amb la seva creació, Espanya s'ha convertit en el primer país europeu a comptar amb una entitat d'aquestes característiques. Aquesta agència serà l'encarregada de supervisar l'aplicació de la normativa comunitària en matèria d'Intel·ligència Artificial.

6. LA INTEL·LIGÈNCIA ARTIFICIAL A LES ADMINISTRACIONS PÚBLIQUES

6.1 Què entenem per IA en el context de les AAPP?

Els apartats previs, on s'ha explicat què és la Intel·ligència Artificial, analitzat la seva evolució, la governança i l'esborrany de la Llei d'Intel·ligència Artificial de la Unió Europea, han servit per entendre de què estem parlant quan ens referim a Intel·ligència Artificial, situar el debat i contextualitzar el moment polític i social en què ens trobem actualment.

Arribats a aquest punt, el següent pas és posar el focus sobre un aspecte d'especial interès per al Tercer Sector: com s'està aplicant la Intel·ligència Artificial en les Administracions Públiques? En altres paraules, en quins processos i serveis públics s'està incloent l'IA amb l'objectiu de millorar-ne l'efectivitat? En aquest apartat s'explicaran, a través de casos reals, com s'està utilitzant la IA des del sector públic i, en particular, en l'àmbit dels drets socials.

Com hem comentat, la Intel·ligència Artificial també ha arribat a les Administracions Públiques. No obstant això, és important diferenciar el procés de digitalització que està experimentant el sector públic de la pròpia introducció d'Intel·ligència Artificial en aquest àmbit. La transformació digital de les Administracions Públiques comprèn un conjunt de mecanismes tecnològics que poden, o no, incloure Intel·ligència Artificial.

En aquest sentit, cal mencionar que la principal tecnologia emprada pel sector públic són els sistemes de decisió automatitzada (o ADMS, per les seves sigles en anglès), definits com 'uns processos pensats per optimitzar la presa d'una decisió mitjançant l'ús d'algoritmes i grans quantitats de dades'⁸. Aquests sistemes poden utilitzar IA (concepte ja definit anteriorment), encara que hi ha d'altres que no.

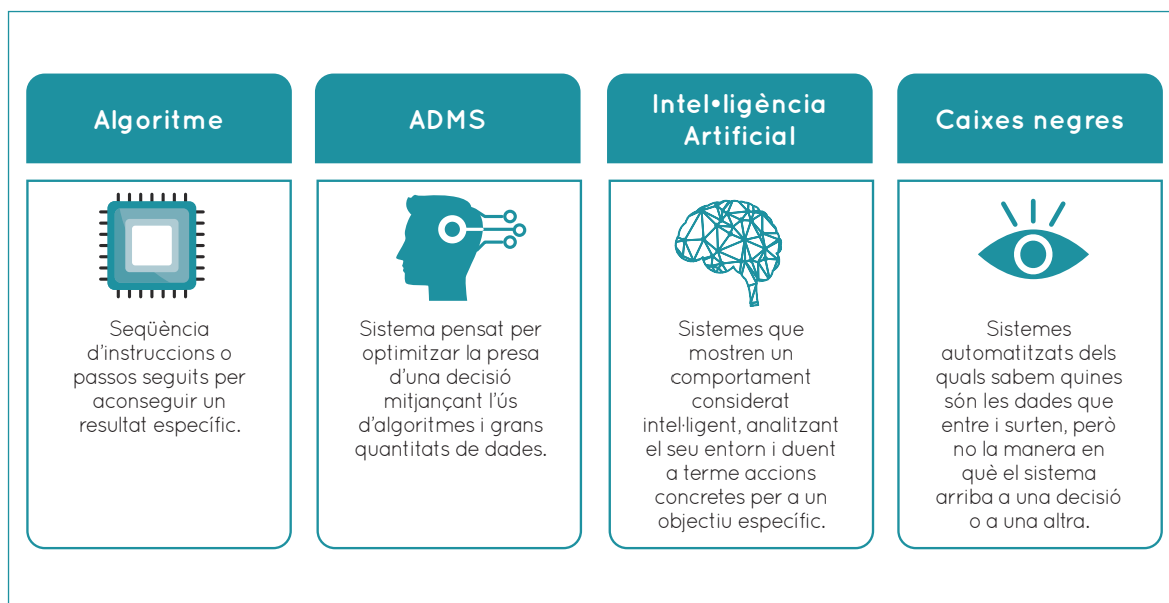
Finalment, un altre concepte vinculat als ADMS és el de caixa negra, que es refereix a quan els procediments i les variables que usen els algoritmes dels ADMS no són fàcils d'observar i, per tant, són poc transparents per les persones.

En el present informe, doncs, i tal com s'ha avançat a la introducció, es fa servir el concepte d'IA com un concepte ampli que inclou tant els ADMS i la Intel·ligència Artificial, donat que actualment els ADMS es troben amb

⁸ AlgorithmWatch, 2019, citat per Jiménez Arandia, P., 2023.

major freqüència dins del Sector Públic, i presenten oportunitats i riscos semblants a les proposades per la IA.

Gràfic 5.



Font: Elaboració pròpia a partir de Jiménez Arandia, P. (2023).

6.2 Aplicacions de la IA a les AAPP

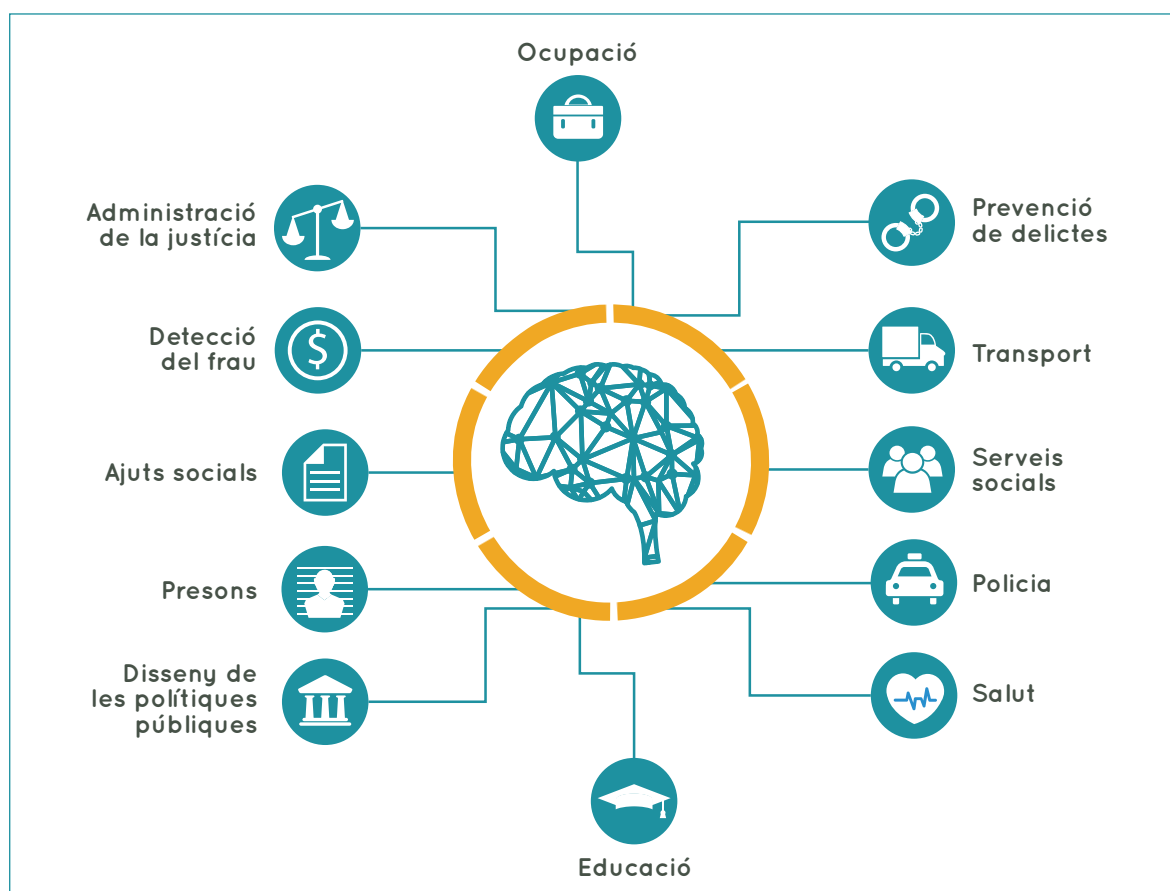
L'ús dels ADMS i IA dins les Administracions Públiques ha crescut durant els últims anys. Com s'ha mencionat anteriorment, el govern es vincula amb la IA a través de dues línies d'acció: per una banda, treballa per establir un marc legislatiu que guii el seu desenvolupament; i per l'altra, treballa per definir en quines tasques la seva implementació pot ser-ne útil en termes d'eficiència i efectivitat. En aquest apartat, doncs, s'exposarà en quines funcions dins de l'Administració Pública s'ha començat a treballar amb IA.

Les experiències internacionals sobre les primeres aplicacions d'IA dins del Sector Públic mostren com aquesta pot tenir un gran impacte en la millora de les operacions internes, de la presa de decisions, dels serveis públics així com de la confiança en el govern. Aquesta evidència, per tant, assenyala el gran potencial que té la IA dins el marc de les Administracions Públiques.

La Intel·ligència Artificial permet, sempre que es mantinguin uns principis ètics i propis d'un Estat democràtic, que els processos actuals realitzats per l'Administració Pública siguin més eficients i precisos. Així com es desenvoluparà amb més detall en la Secció 7, els nombrosos beneficis de la IA en el sector públic poden ajudar a dissenyar millors polítiques públiques, millorar la comunicació i el compromís amb els ciutadans, i millorar la qualitat

i velocitat dels serveis públics. Per aquest motiu, aquesta ha començat a implementar-se per a automatitzar els ajuts públics (BOSCO), per detectar possibles casos de frau a l'Estat (SyRI, Robodebt), per a ajudar a la policia (VioGén, VeriPol, PredPol), per la prevenció de delictes (RisCanvi, COMPAS, LSI-R, OASys), per a fomentar l'ocupació, per accedir a la universitat (Ofqual) i per a la detecció d'infants i adolescents en situació de vulnerabilitat des de Serveis Socials (Gladsaxe, Allegheny), entre altres funcions⁹.

Gràfic 6.



Font: Elaboració pròpia

Més concretament, d'acord amb la informació recollida al *Joint Research Centre* de la Comisió Europea¹⁰, les principals funcions a les quals s'ha començat a treballar amb Intel·ligència Artificial és en l'àmbit del transport; en òrgans executius i legislatius, afers financers i fiscals, i afers exteriors; en serveis públics generals; en serveis de salut pública; per a Serveis de

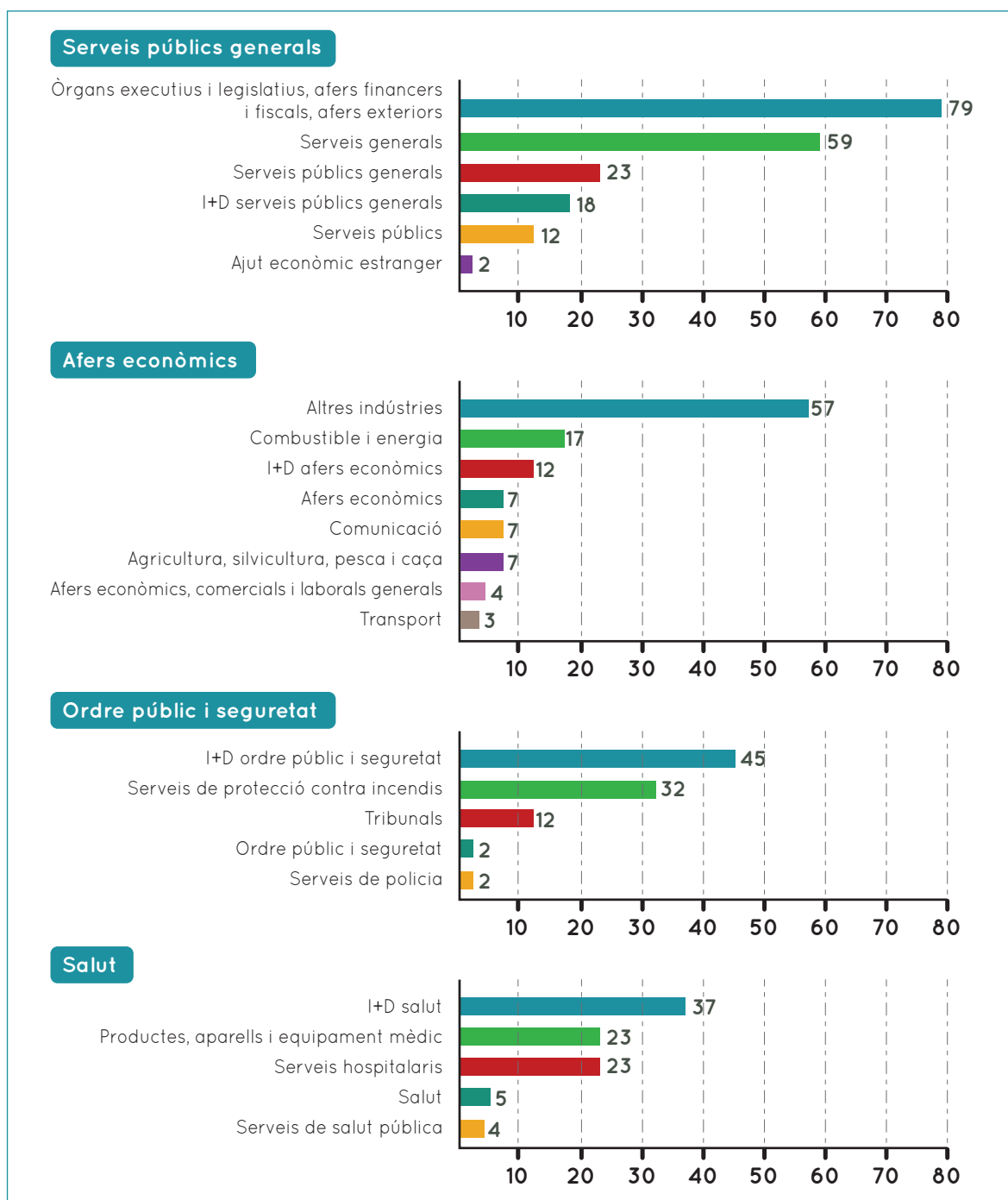
⁹ A continuació es detallen el país que ha implementat cadascun dels algoritmes mencionats:

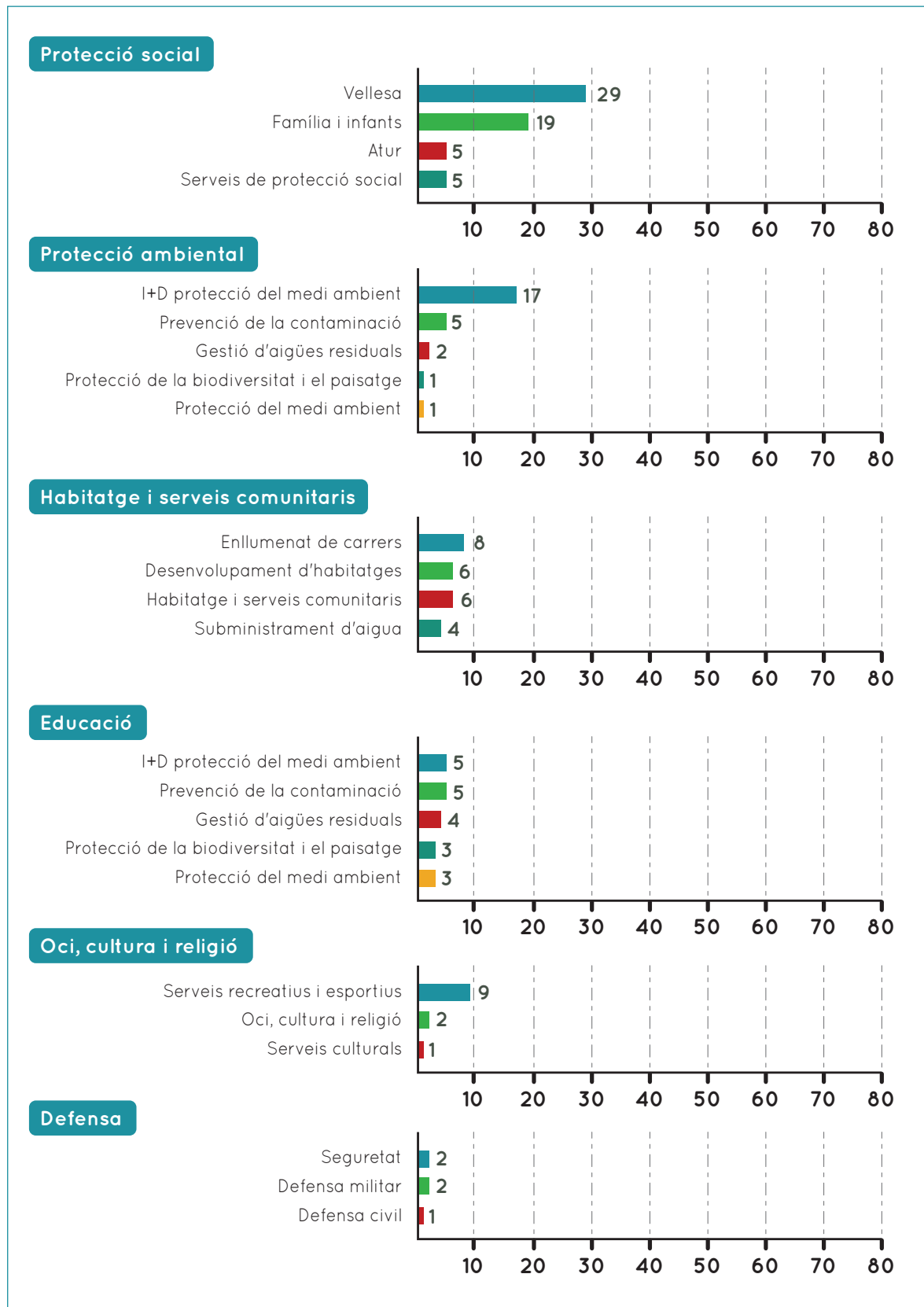
Espanyols: BOSCO, VioGén, VeriPol, RisCanvi; Holanda: SyRI; Austràlia: Robodebt; Estats Units d'Amèrica: PredPol, COMPAS, LSI-R, Allegheny; Canadà: LSI-R; Regne Unit: LSI-R, OASys, Ofqual; Dinamarca: Gladsaxe.

¹⁰ Disponible a: <https://data.jrc.ec.europa.eu/dataset/7342ea15-fd4f-4184-9603-98bd87d8239a>

Polícia; per a l'ordre públic i seguretat; i finalment, en l'àmbit de la protecció social. Tanmateix, tal com es pot observar en el Gràfic 1, els àmbits on actualment hi ha una major presència d'Intel·ligència Artificial són per als ajuts econòmics estrangers, per als serveis públics, en afers econòmics d'altres indústries i per a l'I+D en ordre públic i seguretat.

Gràfic 7. Nombre d'algoritmes d'Intel·ligència Artificial que són utilitzats a la Unió Europea, per funcions.



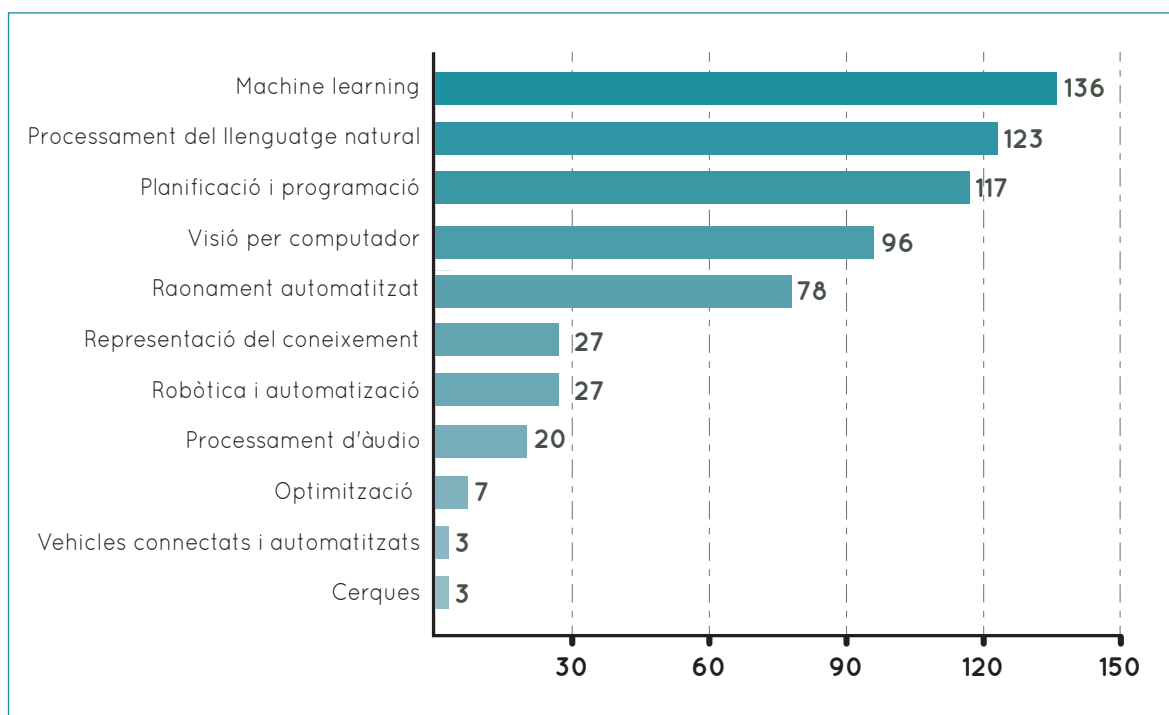


Font: elaboració pròpia a partir de les dades del Joint Research Centre.

Pel que fa al tipus d'IA més freqüent dins del sector públic es troba el *Machine Learning*, seguit del Processament de llenguatge natural i algoritmes per a la planificació i programació (veure Gràfic 2).

A la vegada, aquests algoritmes majoritàriament són introduïts en l'àmbit nacional, (371 d'un total de 686, que representa un 54,1%), un 26,7% són introduïts en l'àmbit local, i només un 10,1% i un 9,2%, són aplicats per a funcions regionals o entre diferents països, respectivament.

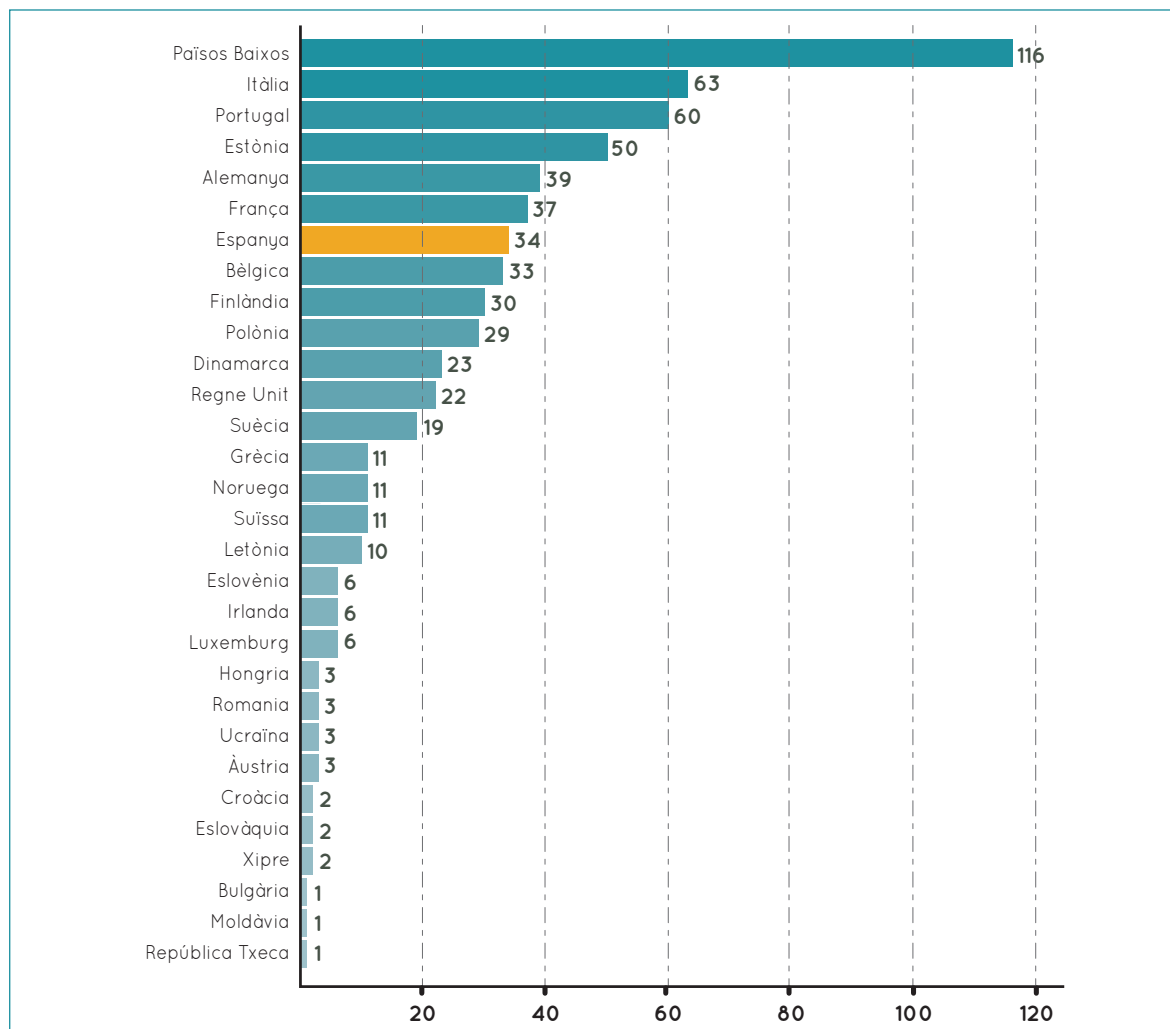
Gràfic 8. Nombre d'algoritmes d'Intel·ligència Artificial que són utilitzats al Sector Públic dins la Unió Europea, per tipus.



Font: elaboració pròpia a partir de les dades del Joint Research Centre.

A més, el següent gràfic mostra com els tres països on hi ha més algoritmes introduïts, en desenvolupament o en fase de prova pilot en el sector públic són Holanda (amb 116 algoritmes), Itàlia (amb 63 algoritmes) i Portugal (amb 60 algoritmes). Espanya se situa en el sisè lloc molt per sobre de la mitjana.

Gràfic 9. Nombre d'algoritmes d'Intel·ligència Artificial que són utilitzats a la Unió Europea, per país.



Font: elaboració pròpia a partir de les dades del Joint Research Centre.

No obstant això, quan es consideren solament els algoritmes actualment implementats, el rànquing varia. Així doncs, tenint en compte només els 260 algoritmes implementats al llarg del territori de la Unió Europea, els països que ara com ara treballen amb més algoritmes són Itàlia (amb 38 algoritmes), Estònia (amb 27) i Holanda (amb 26).

D'aquesta manera, la IA pot ser implementada en àmbits específics com la salut, el transport públic, la defensa nacional, l'educació, l'administració de la justícia o en els Serveis Socials (veure Secció 6.4). Així, la IA té un gran potencial per a millorar el treball de les Administracions Públiques, capacitat que haurà de demostrar-se amb els resultats de les aplicacions dutes a terme en el present i el futur. Tanmateix, i tal com s'exposarà amb més detall a la Secció 7, la seva introducció també està associada a diferents reptes que el Sector Públic ha d'intentar minimitzar per tal de poder ma-

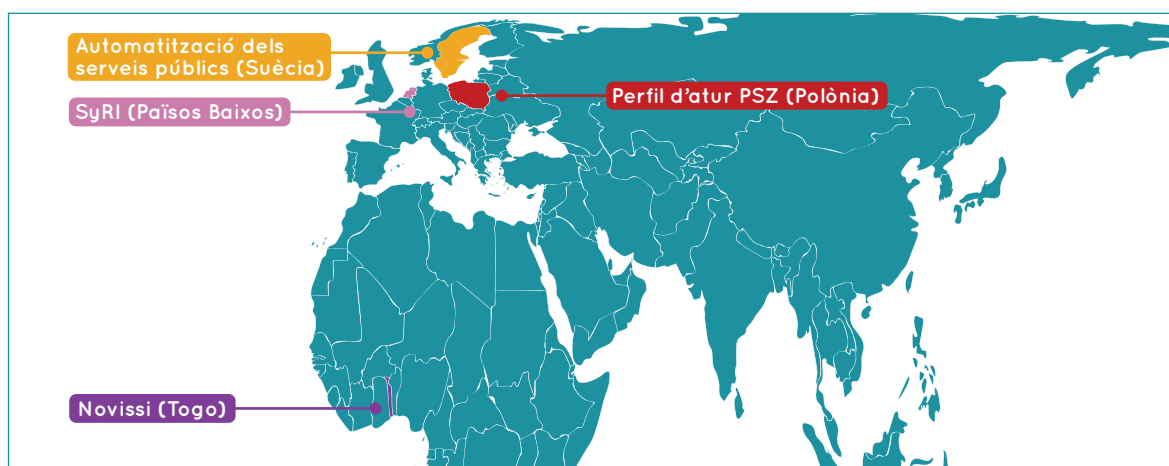
ximitzar-ne els beneficis. D'aquí se'n desprèn **la importància d'establir un pas previ a la implementació de la IA que serveixi per determinar si per a aquell aspecte concret la IA és la millor solució, mitjançant una anàlisi que valori tant els potencials beneficis com els riscos que pot suposar.**

Per mostrar tant els beneficis com els riscos de la IA dins de les Administracions Públiques, a continuació s'aporten casos d'experiències internacionals i nacionals que serveixen d'exemple de bones i males pràctiques. A la vegada, aquests casos mostren la necessitat de realitzar un esforç addicional per tal de garantir que la IA sigui utilitzada de manera fiable, ètica i equitativa, i d'aquesta manera, no impactar negativament sobre la confiança de la ciutadania sobre el govern. De fet, una investigació del Boston Consulting Group assenjala que el suport de la IA governamental està correlacionat amb la confiança en el govern, i que 'la confiança en les institucions és essencial perquè els governs tinguin el suport necessari per a desplegar les capacitats de la IA'¹¹. Aquesta confiança no depèn únicament de l'eficiència i l'eficàcia del Servei públic, sinó també de la satisfacció i de la justícia percebuda en el seu funcionament.

6.3 Casos internacionals

Per exemplificar com s'està implementant sistemes d'IA en processos i serveis de l'Administració Pública, a continuació s'exposaran quatre casos diferents que s'han seleccionat donada la seva popularitat i el volum d'informació disponible. A més, es mostren dos casos amb resultats molt positius (Suècia i Togo) i altres dos on han sorgit problemàtiques (Polònia i Països Baixos).


Gràfic 10.



Font: Elaboració pròpia

¹¹ Carrasco et al., 2019, citat a Berryhill, J., 2019

Suècia (Automated Public Services)

 Automatització dels Serveis Públics a Trelleborg	
País	Suècia
Nom	Automatització dels Serveis Públics
Administració Pública que ho ha implementat	Protecció social, a nivell local
Impacte esperat	Reducció del temps d'espera, augment de l'eficiència, millora de l'experiència de la ciutadania
Descripció de la innovació	Des del 2016, els serveis socials del municipi de Trelleborg (Suècia) compten amb un programa que utilitza IA per a l'automatització de diversos ajuts socials. Davant l'escenari on els ciutadans havien d'esperar una mitjana de 8 dies (que a vegades arribaven a 20), i que acabaven sobrecarregant els professionals que havien de respondre a les seves consultes i gestionar les peticions manualment, l'automatització robòtica de processos (RPA, per les seves sigles en anglès) es va plantejar com la millor solució per a reduir temps d'espera i retards en els pagaments als ciutadans. Tot i així, la supervisió d'un professional s'ha mantingut per a aquells casos on es rebutja la sol·licitud. L'any 2020, el sistema ja era capaç de processar sol·licituds d'atenció domiciliària, prestacions de malaltia, prestacions d'atur i impostos.
Resultat i impacte	<p>Aquest sistema automatitzat ha permès reduir el temps d'espera en les sol·licituds d'ajuts significativament, s'han reduït el temps de gestió de casos amb vulnerabilitat econòmica, i totes les decisions d'assistència financera es prenen en un termini de 24 hores.</p> <p>A més, un estudi sobre l'impacte en els professionals mostra un efecte positiu atès que la Intel·ligència Artificial els ajudava a fer la seva feina de manera més eficaç i els aportava seguretat jurídica. Així doncs, es constata el potencial de la col·laboració tecnologia i humana per a redefinir les pràctiques d'assistència social. No obstant això, en un altre informe s'ha assenyalat la importància de fer que aquesta tecnologia sigui rebuda amb confiança pels professionals, ja que si no aquests es mostren reticents als canvis.</p> <p>Per tant, aquest cas mostra com l'automatització de processos i els sistemes d'Intel·ligència Artificial poden ser molt beneficiosos, sempre que no es descuidi la relació dels professionals amb aquests sistemes.</p>

Togo (Novissi)

 Novissi	
País	Togo
Nom	Novissi
Administració Pública que ho ha implementat	Govern de Togo, nivell estatal
Impacte esperat	Accés a mesures d'emergències del govern, transferències monetàries
Descripció de la innovació	<p>Davant la situació d'urgència marcada per la Covid-19, el Govern de Togo tenia la voluntat de desplegar ajuts econòmics per a aquelles famílies amb més necessitats i així, reduir l'impacte negatiu de la pandèmia. No obstant això, el Govern de Togo no disposava d'un registre social complet que els permetés identificar les persones més pobres, i la pandèmia feia impossible la recollida d'un nou registre. La manca d'informació augmentava l'exclusió de persones elegibles per als ajuts. Per aquest motiu van haver de buscar una alternativa. Aquest alternativa, anomenada Novissi, es va desenvolupar i implementar mitjançant algorismes d'aprenentatge automàtic, dades satèl·lits i xarxes de telefonia mòbil. Inicialment, els ajuts es van distribuir a famílies que complissin amb els següents tres requisits: 1) Estar registrat a la plataforma de Novissi i haver facilitat informació bàsica des del seu telèfon mòbil, 2) Estar registrat per votar a una regió específica (Greater Lomé), 3) Autodeclarar treballar en una ocupació informal en el seu registre de votants. Més endavant, es va voler estendre aquest servei a les persones en situació de pobresa a les zones rurals del país.</p>
Resultat i impacte	<p>L'avaluació d'aquest programa mostra el potencial d'utilitzar noves fonts de dades complementàries a les fonts tradicionals per a arribar a un major volum de persones elegibles, sobretot en contextos de crisi on les dades tradicionals poden faltar o no estar actualitzades. D'aquesta manera, posa de relleu el potencial que té el <i>Machine Learning</i> per a processar grans quantitats de dades i compaginar diverses fonts de dades, que no seria possible d'una altra manera. En aquest cas concret, per tant, l'avaluació mostra com l'agilitat del sistema de <i>Machine Learning</i> va permetre reduir errors i arribar a més persones que necessitaven l'ajuda. A més, va reduir el nombre de persones que haurien estat excloses de manera incorrecta de l'assistència social en comparació amb l'opció geogràfica.</p>

Països Baixos (SyRI)

 SyRI	
País	Holanda
Nom	SyRI (Systeem Risico Indicatie)
Administració Pública que ho ha implementat	Protecció Social, a nivell central i municipal
Impacte esperat	Capacitats d'inspecció millorades, millora del benestar dels nens, reducció del mal ús dels fons públics
Descripció de la innovació	L'any 2012, l'Agència Tributària d'Holanda va començar a utilitzar algoritmes d'autoaprenentatge per crear perfils de risc de cometre frau per tal d'evitar el frau en les ajudes destinades per a la cura d'infants.
Resultat i impacte	<p>Després d'uns anys d'estar en funcionament, aquest sistema es va retirar per tenir conseqüències clarament negatives. L'algoritme s'havia desenvolupat de tal manera que categoritzava com a deutors a famílies que havien omplert incorrectament els documents de sol·licitud. Alhora, tenir una doble nacionalitat també influïa en aquest perfilat, així com provenir d'un nivell socioeconòmic baix, ser immigrants o pertànyer a una minoria ètnica eren característiques que conduïen a l'algoritme a penalitzar desproporcionadament sobre aquests grups poblacionals. A més, els professionals encarregats de supervisar els casos van deixar passar moltes falses alarmes i van continuar amb la sol·licitud de retorn dels diners.</p> <p>Com a resultat, més de 10.000 persones van entrar en situació de pobresa, altres van morir per suïcidi després de rebre factures de deutes amb unes quanties impossibles, i fins i tot, més de 1.100 infants van ser separats de les seves famílies i es van posar en centres d'acollida. Un total de 30.000 famílies es van veure afectades per aquest algoritme.</p> <p>Davant aquest escenari, es va iniciar un procés judicial per a investigar l'algoritme, que va conduir a una multa de 2,75 milions d'euros a l'Agència Tributària i el Govern en funcions va dimitir.</p>

Polònia (Unemployed Profiling)

**Perfil d'atur PSZ (Publiczne Służby Zatrudnienia)**

País	Polònia
Nom	PSZ (Publiczne Służby Zatrudnienia)
Administració Pública que ho ha implementat	Afers econòmics, nivell central i municipal
Impacte esperat	Serveis públics personalitzats, reducció de l'atur, millora eficiència
Descripció de la innovació	L'any 2012, el Ministeri de Treball i Política Social de Polònia tenia la voluntat de reformar les oficines laborals que en aquell moment tenien poc personal i les tasques desenvolupades eren poc eficients. No obstant això, el Govern de Polònia no disposava de suficients recursos econòmics per ampliar la plantilla. D'aquesta manera, es va recórrer a un sistema automatitzat de perfilat per a gestionar més eficientment els casos de persones aturades. El sistema es va dissenyar amb l'objectiu de situar a les persones aturades en una de les tres categories, tenint en compte característiques individuals. Cada una d'aquestes tres categories està associada al tipus de programa que aquesta persona pot optar, és a dir, defineix l'elegibilitat de les persones davant del catàleg de serveis (com per exemple, programes d'inserció laboral, de formació professional, i inclús del subsidi d'atur). Les dades -com l'edat, el sexe, i la durada de l'atur- es recollien durant una entrevista inicial, que després servien per categoritzar a cada individu. Per tant, aquest sistema d'assignació definia el nivell de suport i la càrrega de recursos a destinar a cada persona, on una de les categories era l'absència d'ajuda estatal.
Resultat i impacte	Després de 7 anys, l'any 2019 es va dictaminar l'eliminació d'aquest sistema de categorització, ja que es va trobar que era ineficient i es van identificar patrons de discriminació: ser mare soltera, tenir una discapacitat o viure en una zona rural implicaven una pitjor categoria i, per tant, rebre menys assistència per part dels serveis professionals estatals. A més, un 44% del personal que treballava en les oficines deien que no els era útil per a reduir la càrrega de treball del dia a dia, i un 80% afirmava que el sistema s'havia de canviar.

6.4 Aplicacions en l'àmbit social

Tal com s'ha vist anteriorment, existeixen nombrosos casos on s'ha començat a experimentar i introduir la Intel·ligència Artificial dins de les Administracions Públiques, per al desenvolupament de tasques de caràcter molt divers. Tanmateix, quan aquestes innovacions es desenvolupen en l'àmbit social -i més concretament, en el de la protecció social- desperten un major interès i una necessitat d'avançar amb prudència, atès que poden impactar més notòriament als ciutadans. Precisament per la seva naturalesa, aquest àmbit és especialment sensible, donat que les persones a qui va dirigida aquesta protecció solen trobar-se en situacions de vulnerabilitat socioeconòmica. De fet, la protecció social és 'la cara més visible dels Governos', de manera que la introducció de la IA s'ha de fer amb més cura, ja que els riscos poden ser molt grans si no s'implementa bé.

Així doncs, amb una bona implementació de la IA (entesa tal com està definida per la Unió Europea, en tant que està centrada en les persones i és fiable), les Administracions Públiques de l'àmbit social poden utilitzar aquesta eina per a avaluar l'elegibilitat i les necessitats, per a la presa de decisions d'inscripció, per a l'assignació de prestacions així com per a fer-ne seguiment de les mateixes. A la vegada, la IA pot permetre que la sol·licitud d'una cita amb un professional sigui amb aquell qui pugui respondre millor les necessitats del ciutadà. També, pot ajudar a distribuir millor els usuaris al llarg del territori, gestionar les llistes d'espera o conèixer l'evolució de la demanda en temps real.

Es recomana llegir l'informe Intel·ligència Artificial: Decisions Automatitzades a Catalunya de l'Autoritat Catalana de Protecció de Dades. Tot i que exposa exemples d'aplicació d'algoritmes en diversos camps, com ara la salut, el sistema judicial o l'educació, també enumera fins a set aplicacions en l'àmbit social. Encara que no s'entrarà a comentar-les totes, sí que se'n farà menció. L'Ajuntament de Barcelona atén 50.000 persones de mitjana a l'any a través dels seus 40 centres de serveis socials. Aquestes persones tenen problemàtiques molt diverses que són ateses per una plantilla de 700 persones. L'Ajuntament va optar per explotar la informació emmagatzemada a les 300 mil entrevistes que s'havien realitzat, i aplicant tècniques de *Machine Learning*, han desenvolupat un model capaç d'analitzar els nous casos, classificar les demandes i aportar possibles respostes.

Finalment, per al seguiment de les intervencions, la IA permet que l'atenció per a cada usuari sigui més individualitzada i integral (com per exemple amb la Història Social Única Electrònica del 'Sistema Cohessiona' de

la Junta d'Andalusia¹²). També, permet anticipar futures demandes i, per tant, actuar de manera preventiva. A continuació, s'introdueixen dos casos d'aplicacions d'algoritmes en el territori espanyol que han tingut ressò a escala estatal a causa de la sensibilitat en l'àmbit on aquests operen.

Espanya (VioGén)

	
País	Espanya
Nom	VioGén
Administració Pública que ho ha implementat	Ministeri d'Interior, a nivell estatal
Impacte esperat	Poder fer seguiment i protegir a les dones víctimes de violència de gènere i dels seus fills i filles en qualsevol punt del territori nacional.
Descripció de la innovació	L'any 2007, el Ministeri d'Interior del Govern d'Espanya, va desenvolupar el Sistema de Seguiment Integral de casos de Violència de Gènere (VioGén). VioGén va ser dissenyat amb l'objectiu de determinar el nivell de perill de patir noves agressions, d'escollir els protocols més adequats per a protegir a les víctimes i als seus fills, i establir-ne un seguiment per a determinar i ajustar les mesures de seguretat. No obstant això, la decisió final és supervisada per un professional, que pot pujar el nivell de risc, però mai baixar-lo.
Resultat i impacte	Des de la seva implementació fins al maig de 2023, el sistema ha estat utilitzat per valorar 6.047.700 casos, amb 77.213 casos actius. A més, el sistema s'ha modificat 5 cops per tal d'adaptar-lo millor a les noves situacions. Tanmateix, aquest sistema no és inequívoc i continuen havent-hi casos de dones assassinades per les seves parelles o exparelles, inclús després d'haver denunciat. A la vegada, l'algoritme també és opac i no es coneix quin és el sistema de puntuació de risc que utilitza.

12 Veure més informació sobre el programa [aquí](#).

Espanya (BOSCO)

	
País	Espanya
Nom	BOSCO
Administració Pública que ho ha implementat	Serveis Socials, a nivell estatal
Impacte esperat	Oferir una millor regulació de l'atorgament de bons socials per al pagament de la factura de la llum
Descripció de la innovació	L'algoritme BOSCO va ser introduït pel Govern Espanyol i implementat per les elèctriques amb la finalitat de regular el bo social sobre el pagament de la factura de la llum. Aquest bo permet l'accés a un descompte a la factura de la llum.
Resultat i impacte	Tot i ser totalment opac per a la població, fet que ha estat denunciat per la Fundació Civio, aquest algoritme també sembla ser ineficient en tant que, persones elegibles que l'han sol·licitat, finalment no l'han obtingut. D'aquesta manera, tot i la seva opacitat, s'ha manifestat la preocupació envers la possibilitat que aquest algoritme estigui discriminant a grups poblacionals que tenen el dret de percebre'l, com serien els jubilats i les vídues.

7. ELS BENEFICIS I RISCOS DE L'APLICACIÓ DE LA INTEL·LIGÈNCIA ARTIFICIAL A LES ADMINISTRACIONS PÚBLIQUES EN L'ÀMBIT SOCIAL

Al llarg d'aquest informe hem anat veient quins han estat els avenços per part dels governs – amb especial èmfasi a Espanya– per tal de garantir que les Administracions Públiques i la societat en conjunt puguin maximitzar els beneficis i minimitzar els riscos que pot comportar un futur amb Intel·ligència Artificial. Si bé s'han anat exemplificant alguns d'aquests, encara no hem aprofundit sobre quins beneficis i riscos ens estem referint quan parlem d'IA. Per tant, en aquesta secció es presenten els beneficis i riscos que han estat identificats fins al moment.

Tal com s'ha esmentat, la introducció de la Intel·ligència Artificial a les Administracions Públiques comporta un seguit de beneficis i riscos associats que cal tenir presents, ja que poden impactar directament sobre l'exercici dels drets socials.

Beneficis

Pel que fa als **beneficis**, aquesta tecnologia ofereix nombroses oportunitats al Sector Públic per a dissenyar millors polítiques i millorar la presa de decisions, millorar la comunicació i el compromís amb els ciutadans, millorar la velocitat i la qualitat en la provisió dels serveis públics, així com substituir tasques rutinàries realitzades pels funcionaris públics i permetre que aquests realitzin tasques de més valor afegit. D'aquesta manera, la IA permet millorar el disseny i la provisió de serveis públics, així com millorar la gestió interna de les institucions estatals.

Respecte al primer punt, de **millorar el disseny i la provisió de serveis públics**, la Intel·ligència Artificial pot ajudar les Administracions Públiques a identificar interessos, preocupacions i percepcions de diferents actors per tal que siguin inclosos en l'agenda del govern. A més, pot ajudar a detectar problemes en l'accés a determinats serveis, a comprendre millor el comportament dels ciutadans i veure l'impacte de diferents entitats públiques sobre determinats grups poblacionals.

Ahora, aquesta tecnologia permet gestionar grans quantitats de dades que generen els ciutadans a través de la seva interacció amb dispositius mòbils i xarxes socials. Mitjançant aquest tractament, els professionals de

l'Administració Pública poden dissenyar serveis més personalitzats a la realitat de la ciutadania i, d'aquesta manera, proveir una millor atenció a les entitats amb un major estalvi de temps i recursos.

A la vegada, la IA també ofereix la possibilitat de fer que les interaccions amb els ciutadans siguin més eficients. A través de robots conversacionals, els ciutadans poden resoldre dubtes o sol·licituds de manera àgil, fet que afavoreix a una millor satisfacció d'aquests, sempre que aquests coneguin que hi ha un algoritme al darrere.

Respecte al segon punt, de **millorar la gestió interna de les institucions estatals**, la Intel·ligència Artificial disposa de tècniques de predicció, optimització o control que poden utilitzar-se per ajudar a l'assignació i gestió de recursos econòmics, a la detecció i prevenció de frau, i a evitar usos ineficients d'aquests recursos. A la vegada, pot ajudar a la generació de nova regulació, la seva actualització i per a garantir-ne el seu compliment. Gràcies a la seva capacitat d'automatització, aquesta tecnologia també pot utilitzar-se per a tasques repetitives i rutinàries, permetent als professionals de les Administracions Públiques realitzar activitats més complexes que requereixen majors habilitats emocionals, de creativitat i de perspectiva humana. Així doncs, un dels exemples on es pot automatitzar el servei públic és en l'entrada de dades a través del reconeixement automàtic de l'escriptura manual, el reconeixement de veu o el processament del llenguatge natural.

D'aquesta manera, en reduir la càrrega de treball eliminant les tasques repetitives i mecàniques, l'administració guanya temps que podrà dedicar a fer una atenció més ràpida, eficient i personalitzada, i a millorar altres processos que ho requereixin o proveir nous serveis.

Finalment, les capacitats de precisió i predicció poden ser usades per les Administracions Públiques per fer diagnòstics o definir línies d'actuació davant d'una emergència.

Riscos

Pel que fa als **riscos** d'aplicar IA en el Sector Públic es troba en primer lloc la **falta d'accountability** (rendició de comptes en català), **de transparència i explicabilitat dels algoritmes**.

La OCDE defineix els principis d'*accountability*, de transparència i explicabilitat de la següent manera:

- ▶ **Principi d'*accountability*:** 'els actors de la IA han de garantir que els seus sistemes d'IA són fiables'.
- ▶ **Principi de transparència i explicabilitat:** 'els actors han de comprometre's amb la transparència i la divulgació responsable dels sistemes d'IA. Per a això, han de proporcionar informació significativa adequada al context i coherent amb l'estat de la qüestió'.

Les principals crítiques d'aquesta tecnologia giren entorn de la seva **opacitat per a ser entesa pels ciutadans**, on la complexitat del seu procés juga un paper rellevant. Diversos casos internacionals han mostrat com, sovint, aquests algoritmes d'Intel·ligència Artificial no són de codi obert (i, per tant, no són transparents). En conseqüència, la falta de transparència dificulta que es puguin auditar els algoritmes i, per tant, que es pugui complir amb el principi d'*accountability*.

A la vegada, diversos algoritmes tampoc són fàcilment comprensibles per part dels ciutadans, incomplint amb el principi de transparència i explicabilitat. D'acord amb aquest segon principi, **les persones que utilitzen aquestes solucions han de poder entendre com s'ha arribat a aquell resultat, i tant aquestes com les persones que no les utilitzen però es veuen afectades, puguin exercir el seu dret a reclamar si no estan conforme**. Tanmateix, l'opacitat en el codi com de les caixes negres impedeixen que es pugui complir amb aquest principi. A més, aquest aspecte està estretament vinculat amb el principi d'*accountability*, ja que sense una garantia de transparència i explicabilitat, demostrar que un sistema IA és fiable resulta un exercici molt més complex.

Aquests principis són essencials per tal de garantir que els algoritmes no van en contra del bé comú i que, per tant, no s'estan vulnerant els drets socials de la ciutadania. Precisament, entendre com s'ha obtingut un resultat és essencial per a identificar biaixos o patrons de discriminació. D'aquesta manera, quan els actors públics no garanteixen un compliment d'aquests principis, el control públic queda reduït i afecta especialment els ciutadans pertanyents a minories i a grups desfavorits, en tant que no poden exigir la garantia dels seus drets socials. Per tal d'evitar-ho, estan apareixent iniciatives basades en la creació de registres d'algoritmes a ciutats com Bar-

celona¹³, amb l'objectiu de dotar de més transparència i fiabilitat a aquests sistemes.

A part de la manca de compliment d'aquests principis, un altre problema associat a l'**amplificació de biaixos i discriminació** és la falta de dades de bona qualitat. Sovint, la generació d'aquests patrons de discriminació sorgeix d'un aprenentatge que realitzen els algoritmes a partir de les dades amb les quals s'entrenen i aprenen. D'aquesta manera, si les dades introduïdes no són de qualitat i contenen errors i discriminacions ja existents dins de la societat, aquests algoritmes simplement seran una rèplica d'aquests biaixos. Per tant, d'aquí se'n deriva la gran importància de disposar de dades de qualitat i que no comportin a la perpetuació de desigualtats del '*Matthew effect*' on els privilegiats obtenen avantatges, mentre que els que ja es troben en una situació de desavantatge, encara pateixen més¹⁴. Alhora, aquests biaixos poden fomentar un procés de polarització social i adoptar comportaments radicals sobre determinats grups poblacionals.

Un altre risc a tenir en compte és la **seguretat i privacitat de la informació utilitzada**. La informació utilitzada per les Administracions Públiques és extremadament sensible i la fuga d'informació o els atacs cibernètics podrien comportar greus conseqüències, tant per les entitats públiques com pels propis ciutadans. En aquest sentit, també cal assenyalar que la Llei de Protecció de Dades¹⁵ pot suposar un impediment per a l'entrenament de certs models si no es garanteix la seguretat i privacitat de la informació utilitzada.

Com a conseqüència, tots aquests riscos poden acabar afectant a l'efectivitat de les polítiques públiques atès que les possibles situacions d'exclusió de població diana podrien reduir l'impacte d'aquestes. Al capdavant, tots aquests riscos impacten negativament sobre la confiança dipositada en les Administracions Públiques per part dels ciutadans. El conjunt d'errors i perpetuació de desigualtats acaba desgastant la confiança de la població sobre l'eficàcia de l'actuació de les Administracions Públiques, amb implicacions directes sobre la voluntat de participació ciutadana i, finalment, sobre l'estat de la democràcia.

En resum, entre els principals beneficis de la IA es troben **una millora en l'eficiència dels recursos** (tant econòmics com de personal), **una millor provisió d'un servei públic i una millor adequació de les polítiques públiques i programes de l'Agenda**. En canvi, els riscos es vinculen principalment amb

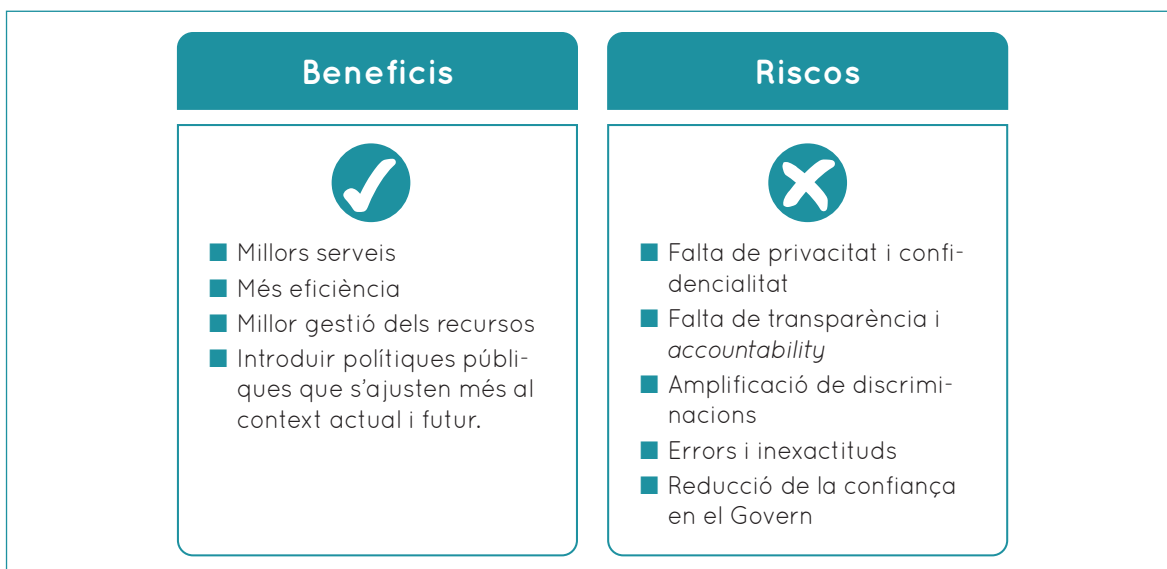
¹³ Veure més informació [aquí](#).

¹⁴ Herzog, 2021, citat a OCDE, 2023, p.8

¹⁵ Llei Orgànica 3/2018, de 5 de desembre, de Protecció de Dades Personals i garantia dels drets digitals.

l'actual **falta de transparència d'aquests processos, la falta de dades adequades per evitar situacions de biaixos i discriminació, la sensibilitat de les dades tractades per tal de garantir-ne la privacitat, les situacions de falsos positius i errors que condueixen a situacions greus per a les persones afectades** i, com a conseqüència, finalment, **la pèrdua de la confiança dels ciutadans sobre el treball realitzat pel Sector Públic.**

Gràfic 11.



8. CONCLUSIÓ

La Intel·ligència Artificial té potencial per transformar significativament la nostra societat. Cada vegada més empreses i administracions públiques integraran en els seus productes i processos, sistemes que utilitzin Intel·ligència Artificial buscant ser més eficients, efectius o beneficiosos.

La Unió Europea i els governs nacionals fa anys que treballen en una regulació que constrenyi les possibilitats de desenvolupament amb l'esperança de minimitzar i prevenir els riscos sense afectar la innovació i l'avenç en el camp de la IA. Una regulació que a priori serà definitiva a finals del 2023.

Davant aquest context, de present i futur, és fonamental que les entitats del tercer sector fixin una posició i discurs comú i ho traslladin tant a les esferes polítiques com públiques. En els pròxims anys, probablement es promouran la major part de les regulacions, però també es consolidarà la cosmovisió sobre la Intel·ligència Artificial. Per tant, **és fonamental que el Tercer Sector sigui un actor amb veu (i vot) si vol participar en la conformació de la societat futura.**

Les entitats del tercer sector actualment es troben en un moment clau per adoptar una estratègia proactiva i fer escoltar la seva veu durant l'elaboració de les següents legislacions. Alhora, les entitats han de poder exercir de pont entre la ciutadania i les Administracions Públiques per a conscienciar-les dels avenços que tenen a la seva disposició, i informar-les sobre com es poden beneficiar mentre eviten tots els riscos.

Addicionalment, comptar amb una visió comuna com a Tercer Sector, enforteix l'acció desenvolupada en el control d'aquestes aplicacions en l'àmbit dels drets socials i, a la vegada, permet impulsar nous acords amb diferents agents per tal de poder aprofitar les oportunitats d'aquesta tecnologia.

En definitiva, **el Tercer Sector es troba davant una oportunitat única perquè, a través d'una posició comuna, enforteixi la seva veu i continuï treballant en la garantia dels drets socials en aquest nou entorn canviant.**

REFERÈNCIES

- Ahn, M. J., & Chen, Y.-C. (2020). Artificial Intelligence in Government: Potentials, Challenges, and the Future. In S.-J. Eom & J. Lee (Eds.), Dg.o '20: The 21st Annual International Conference on Digital Government Research (pp. 243–252). Association for Computing Machinery. <https://doi.org/10.1145/3396956.3398260>
- AI Watch (2022). Spain AI Strategy Report. Comisión Europea. Consultado el día 21/06/2023. Disponible en: https://ai-watch.ec.europa.eu/countries/spain/spain-ai-strategy-report_en
- ADB (2020). AI in Social Protection - Exploring Opportunities and Mitigating Risks. Recuperat de: <https://www.adb.org/publications/ai-social-protection-exploring-opportunities-mitigating-risks>
- Abu Elyounes, D. (10 de febrer de 2021). Why the resignation of the dutch government is a good reminder of how important it is to monitor and regulate algorithms. Medium. Recuperat el 25 de juliol de 2023, de <https://medium.com/berkman-klein-center/why-the-resignation-of-the-dutch-government-is-a-good-reminder-of-how-important-it-is-to-monitor-2c599c1e0100>
- Algoritm Watch; Bertelsmann Stiftung. (Gener de 2019). Bertelsmann Stiftung. Automating Society. Taking Stock of Automated Decision-Making in the EU. Berlín: AW AlgorithmWatch gGmbH, gener de 2019. Recuperat de: https://algorithmwatch.org/de/wp-content/uploads/2019/02/Automating_Society_Report_2019.pdf
- Belinchón, F. (2023). La primera ley europea sobre inteligencia artificial se amplía para incluir a ChatGPT y entra en su recta final. El País. Disponible en: <https://cincodias.elpais.com/economia/2023-06-14/la-primer-ley-europea-sobre-inteligencia-artificial-se-amplia-para-incluir-a-chatgpt-y-entra-en-su-recta-final.html>
- Berryhill, J. et al. (2019), 'Hello, World: Artificial intelligence and its use in the public sector', OECD Working Papers on Public Governance, No. 36, OECD Publishing, Paris, <https://dx.doi.org/10.1787/726fd39d-en>.
- Big Data Magazine (2022). Algoritmos de la Administración Pública española: poco conocidos, opacos y sin auditar. Recuperat el 25 de juliol de 2023, de: <https://bigdatamagazine.es/algoritmos-de-la-administracion-publica-espanola-poco-conocidos-opacos-y-sin-auditar>

- Blázquez, L. (20 de febrer de 2023). ¿Qué es VioGén y cómo funciona?. Rtve. Recuperat el 25 de juliol, de: <https://www.rtve.es/noticias/20230220/objetivo-igualdad-funcionamiento-viogen/2424326.shtml>
- Boucher, P. N. (2020). Artificial intelligence: How does it work, why does it matter, and what can we do about it?. Scientific Foresight Unit (STOA), European Parliament.
- Carrasco, M., S. Mills, A. Whybrew y A. Jura (2019), 'The citizen's perspective on the use of AI in government', BCG, 1 de marzo, www.bcg.com/publications/2019/citizen-perspective-use-artificial-intelligence-government-digital-benchmarking.aspx
- De Schutter, O. (2022). La no percepción de derechos en el contexto de la protección social. Consejo de Derechos Humanos. Obtingut de: <https://documents-ddsny.un.org/doc/UNDOC/GEN/G22/322/20/PDF/G2232220.pdf>
- European Commission (2019). Building Trust in Human-Centric Artificial Intelligence. Communication From The Commission To The European Parliament, The Council, The European Economic And Social Committee And The Committee Of The Regions. COM(2019) 168. Brussels. Disponible en: <https://digital-strategy.ec.europa.eu/en/library/communication-building-trust-human-centric-artificial-intelligence>
- Felton, J. (11 de maig de 2022). The ai that led to children being re-homed and the fall of an elected government. IFL Science. Recuperat el 25 de juliol de 2023, de <https://www.iflscience.com/the-ai-that-led-to-children-being-rehomed-and-the-fall-of-an-elected-government-63622>
- Gleb Papyshv & Masaru Yarime (2023) The state's role in governing artificial intelligence: development, control, and promotion through national strategies, Policy Design and Practice, 6:1, 79-102, DOI: 10.1080/25741292.2022.2162252
- Heikkilä, M. (29 de març de 2022). Dutch scandal serves as a warning for europe over risks of using algorithms. Politico. Recuperado el 25 de juliol de 2023, de <https://www.politico.eu/article/dutch-scandal-serves-as-a-warning-for-europe-over-risks-of-using-algorithms/>
- High-Level Expert Group on Artificial Intelligence (2019). Ethics Guidelines for Trustworthy AI. European Commission. Disponible en: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>

- Jiménez Arandia, P. (2023). Govern Obert: Transparencia algorítmica en el sector públic. (rep.) Vol.9. Generalitat de Catalunya. Generalitat de Catalunya. Recuperat de: https://governobert.gencat.cat/web/.content/01_Que_es/04_Publicacions/colleccio_govern_obert/GovernObert_9/Govern-Obert-9-cast.pdf
- KOSA AI (2021). The importance of AI governance and 5 key principles for its guidance. Medium. Disponible en: <https://kosa-ai.medium.com/the-importance-of-ai-governance-and-5-key-principles-for-its-guidance-219798c8f407>
- Kuziemski, M., & Misuraca, G. (2020). AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings. *Telecommunications policy*, 44(6), 101976.
- La Moncloa (19 de maig de 2023). El Sistema VioGén supera los seis millones de valoraciones de riesgo a víctimas de violencia de género. Recuperat el 25 de juliol de 2023, de: <https://www.lamoncloa.gob.es/serviciosdeprensa/notasprensa/interior/Paginas/2023/190523-sistema-viogen-victimas-violencia-genero.aspx>
- Lokshin, M., & Umapathi, N. (23 de febrero de 2022). AI for social protection: mind the people. Brookings Institution. Recuperado el 28 de junio de 2023, de <https://www.brookings.edu/articles/ai-for-social-protection-mind-the-people/>
- Marilín Gonzalo (2023). Ley de IA en la Unión Europea: estas son las claves de la propuesta. Newtral. Disponible en: <https://www.newtral.es/ley-ia-union-europea-inteligencia-artificial-europa/20230520/>
- Minguíjon J. y Serrano-Martínez C. (2022). La Inteligencia Artificial en los Servicios Sociales: estado de la cuestión y posibles desarrollos futuros. *Cuadernos de Trabajo Social*, 35(2), 319-329. <https://doi.org/10.5209/cuts.78747>
- Misuraca, G., & Van Noordt, C. (2020). AI Watch-Artificial Intelligence in public services: Overview of the use and impact of AI in public services in the EU. JRC Research Reports, (JRC120399).
- Misuraca, G., i van Noordt, C. (2020). Overview of the use and impact of AI in public services in the EU, EUR 30255 EN, Publications Office of the European Union, Luxembourg, 2020, ISBN 978-92-76-19540-5, doi:10.2760/039619, JRC120399
- Ministerio de Asuntos Económicos y Transformación Digital (2020). Estrategia Nacional de Inteligencia Artificial. Disponible en: <https://portal.mineco.gob.es/es-es/digitalizacionIA/Paginas/ENIA.aspx>

- Niklas, J. (16 d'abril de 2019). Poland: Government to scrap controversial unemployment scoring System. Algorithm Watch. <https://algorithmwatch.org/en/poland-government-to-scrap-controversial-unemployment-scoring-system/>
- NL Times. (17 de desembre de 2020). Parents faced 'unprecedented injustice' for years in childcare subsidy scandal. NL Times. Recuperat el 25 de juliol de 2023, de <https://nltimes.nl/2020/12/17/parents-faced-unprecedented-injustice-years-childcare-subsidy-scandal>
- OASI (s/d). Discriminación algorítmica. Recuperat de: <https://eticasfoundation.org/oasi/social-impact-algorithms/>
- Observatorio Nacional de Tecnología y Sociedad (2023). Uso de inteligencia artificial y big data en las empresas españolas. Secretaría de Estado de Digitalización e Inteligencia Artificial, Ministerio de Asuntos Económicos y Transformación Digital. Disponible en: <https://www.ontsi.es/es/publicaciones/uso-de-inteligencia-artificial-y-big-data-en-las-empresas-espanolas>
- OECD (2023). Global Trends in Government Innovation 2023, OECD Public Governance Reviews, OECD Publishing, Paris, <https://doi.org/10.1787/0655b570-en>.
- OECD (2021). 'Tools for trustworthy AI: A framework to compare implementation tools for trustworthy AI systems', OECD Digital Economy Papers, No. 312, OECD Publishing, Paris, <https://doi.org/10.1787/008232ec-en>.
- OECD (s/d). Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449
- ONTSI (2022). Uso de inteligencia artificial y big data en las empresas españolas. Disponible en: <https://www.ontsi.es/es/publicaciones/uso-de-inteligencia-artificial-y-big-data-en-las-empresas-espanolas>
- Our World in Data (2022). Artificial intelligence is transforming our world — it is on all of us to make sure that it goes well. Disponible en: <https://ourworldindata.org/ai-impact>
- Our World in Data (2022). The brief history of artificial intelligence: The world has changed fast – what might be next?. Disponible en: <https://ourworldindata.org/brief-history-of-ai>
- Ranerup, A., & Henriksen, H. Z. (2022). Digital Discretion: Unpacking Human and Technological Agency in Automated Decision Making in Sweden's Social Services. Social Science Computer Review, 40(2), 445–461. <https://doi.org/10.1177/0894439320980434>

- Roser, M. (2022a). Artificial intelligence is transforming our world — it is on all of us to make sure that it goes well. Our World in Data. Disponible en: <https://ourworldindata.org/ai-impact>
- Roser, M. (2022b). The brief history of artificial intelligence: The world has changed fast - what might be next?. Our World in Data. Disponible en: <https://ourworldindata.org/brief-history-of-ai>
- Sun, T. Q., & Medaglia, R. (2019). Mapping the challenges of Artificial Intelligence in the public sector: Evidence from public healthcare. *Government Information Quarterly*, 36(2), 368–383. <https://doi.org/10.1016/j.giq.2018.09.008>
- Tangi L., van Noordt C., Combetto M., Gattwinkel D., Pignatelli F. (2022). AI Watch. European Landscape on the Use of Artificial Intelligence by the Public Sector, EUR 31088 EN, Publications Office of the European Union, Luxembourg , ISBN 978-92- 76-53058-9, doi:10.2760/39336, JRC129301.
- Tangi, L., van Noordt, C., Combetto, M., Gattwinkel, D., & Pignatelli, F. (2022). AI Watch. European landscape on the use of Artificial Intelligence by the Public Sector (No. JRC129301). Joint Research Centre.
- Turkish Policy Quarterly. (7 de març de 2022). AI and democracy: Towards better public oversight. *Turkish Policy Quarterly*. <http://turkish-policy.com/article/1108/ai-and-democracy-towards-better-public-oversight>
- Vélez, M. I., Gómez Santamaría, C., & Osorio Sanabria, M. A. (2022). Conceptos fundamentales y uso responsable de la inteligencia artificial en el sector público. Informe 2. Caracas: CAF. Recuperado de: <https://scioteca.caf.com/handle/123456789/1921>
- Veale, M., Matus, K., & Gorwa, R. (2023). AI and Global Governance: Modalities, Rationales, Tensions. *Annual Review of Law and Social Science*, 19.
- Wirtz, B. W., Weyerer, J. C., & Geyer, C. (2019). Artificial Intelligence and the Public Sector—Applications and Challenges. *International Journal of Public Administration*, 42(7), 596–615. <https://doi.org/10.1080/01900692.2018.1498103>

DOCUMENT 2

**RADAR D'ALGORITMES
I PROCESSOS
DE DECISIÓ
AUTOMATITZADA
PER A L'ACCÉS ALS
DRETS SOCIALS DE LA
CIUTADANIA**

DESEMBRE DE 2023

1. INTRODUCCIÓ

El primer document d'aquest projecte es centrava en l'elaboració de l'informe 'La Intel·ligència Artificial en l'Administració Pública'. Aquest informe va sorgir amb l'objectiu de produir un document que ajudés les entitats del tercer sector a comprendre millor el fenomen de la Intel·ligència Artificial (IA), tant en el seu vessant històric, tècnic i de governança com d'aplicació actual, especialment en l'àmbit públic i social. Per tant, es va tractar d'un informe divulgatiu dirigit especialment al tercer sector o a un públic no especialitzat.

En aquest segon document del projecte, l'objectiu prioritari ha estat la identificació de sistemes d'intel·ligència artificial i processos automatitzats en l'àmbit social, que estiguessin sent utilitzats, o en fase pilot, per part de les Administracions Públiques catalanes.

La IA està al centre del debat actual. Els últims llançaments, i els que estan per arribar, entreveuen una potencial transformació de les relacions socials i el mercat laboral d'abast mundial. Sense cap mena de dubte, **des del 2022 estem vivint en un avenç de publicacions de models i eines sorprenents, sense gaire temps per adaptar-nos-hi ni reflexionar sobre les seves conseqüències.**

L'objectiu d'aquest projecte és la realització d'un recull - també anomenat radar o repositori- de tots aquells algoritmes, models d'intel·ligència artificial i/o processos de decisions automatitzades que s'estiguin utilitzant a Catalunya en l'àmbit de l'accés als drets socials.

El Radar s'adreçarà a ciutadans, administracions públiques i entitats del tercer sector amb la finalitat de proporcionar informació clara i completa sobre l'impacte d'aquesta tecnologia i l'accés als drets socials a Catalunya. Això és molt rellevant atès que **l'ús que es pretén fer del repositori determinarà decisions com quina informació cal recollir o quin llenguatge utilitzar, entre d'altres.**

La creació d'aquest **Radar** es va realitzar en tres passos. En primer lloc, es va dissenyar el radar en si mateix. És a dir, quines variables inclouria i, per tant, quina informació es recolliria per a cada algoritme. Aquest exercici es va realitzar prenent com a referències altres repositoris internacionals.

En segon lloc, es va dur a terme el treball de camp. Després de dissenyar un qüestionari per a una entrevista i implementar un formulari en línia, es va contactar directament amb responsables de les administracions públi-

ques que es considerava que podrien comptar amb informació. A aquestes persones se'ls va donar l'opció d'emplenar directament el formulari o fer una entrevista. Per a ampliar les possibilitats de respostes, finalment es va decidir fer públic el formulari i difondre'l a través de les xarxes socials. En tercer lloc, després de la recollida de respostes, es va processar, netejar i completar la informació per a configurar el Radar d'Algoritmes.

La principal dificultat de la creació d'aquest radar és el desconeixement de l'univers poblacional. En altres paraules, **es desconeix el nombre total de sistemes d'IA actualment implementats o en procés de desenvolupament que estan sent utilitzats per les administracions públiques catalanes**. No obstant això, en les entrevistes amb responsables de diverses administracions es va expressar amb claredat el fet que, en l'àmbit dels drets socials, existeixen pocs casos d'ús d'aquests algoritmes. **En total s'han mapat fins a 12 algoritmes de 4 administracions diferents**.

Tanmateix, **el Radar d'Algoritmes és la primera fase d'un projecte més ampli**. En aquesta fase l'objectiu primordial ha estat mapar tots els algoritmes que estaven sent utilitzats.

Així doncs, aquest document s'estructura en tres apartats:

- (1)** Explicar la metodologia seguida per recaptar tota la informació sobre els algoritmes utilitzats per l'Administració Pública de Catalunya.
- (2)** Explicar les variables que integren el Radar, és a dir, els camps en l'àmbit social per als quals es va recaptar informació.
- (3)** Presentar les fitxes elaborades per a cadascun dels 12 algoritmes recollits.

2. METODOLOGIA

Aquesta fase ha consistit en el mapatge dels diferents algoritmes d'IA i processos de decisió automatitzada que són emprats per les Administracions Públiques catalanes. Atès que són sistemes prou recents i poc coneguts, per a realitzar aquest mapatge s'ha hagut de recórrer a una metodologia combinada. Concretament, l'equip investigador ha creat un formulari on s'inclouen tots els camps i variables d'interès, alhora que s'han realitzat entrevistes a persones identificades com rellevants.

Així doncs, en primer lloc, s'han identificat un total de 28 persones. Aquestes persones es consideren rellevants per complir dos requisits: en primer lloc, treballar en Administracions Públiques catalanes; i, en segon lloc, posseir coneixement directe o haver participat en el disseny o implementació d'un algoritme que va estar o està en ús en l'Administració Pública. Per tant, aquestes persones són la font d'informació principal del Radar.

Aquestes persones van ser identificades en bona part gràcies al coneixement de camp de la Taula d'entitats del Tercer Sector Social de Catalunya, juntament amb la qual es va elaborar un llistat de persones que potencialment podrien tenir coneixement de casos d'ús. Algunes d'aquestes persones van aportar contactes sobre altres responsables que prèviament no havien estat identificats. Sota aquesta dinàmica de bola de neu es va acabar per configurar el llistat definitiu de persones contactades. Tal com es pot observar en el següent gràfic, la majoria d'aquests perfils treballen a l'Ajuntament de Barcelona (un 28%), a la Generalitat de Catalunya (un 21%), i a diferents fundacions (un 21%).

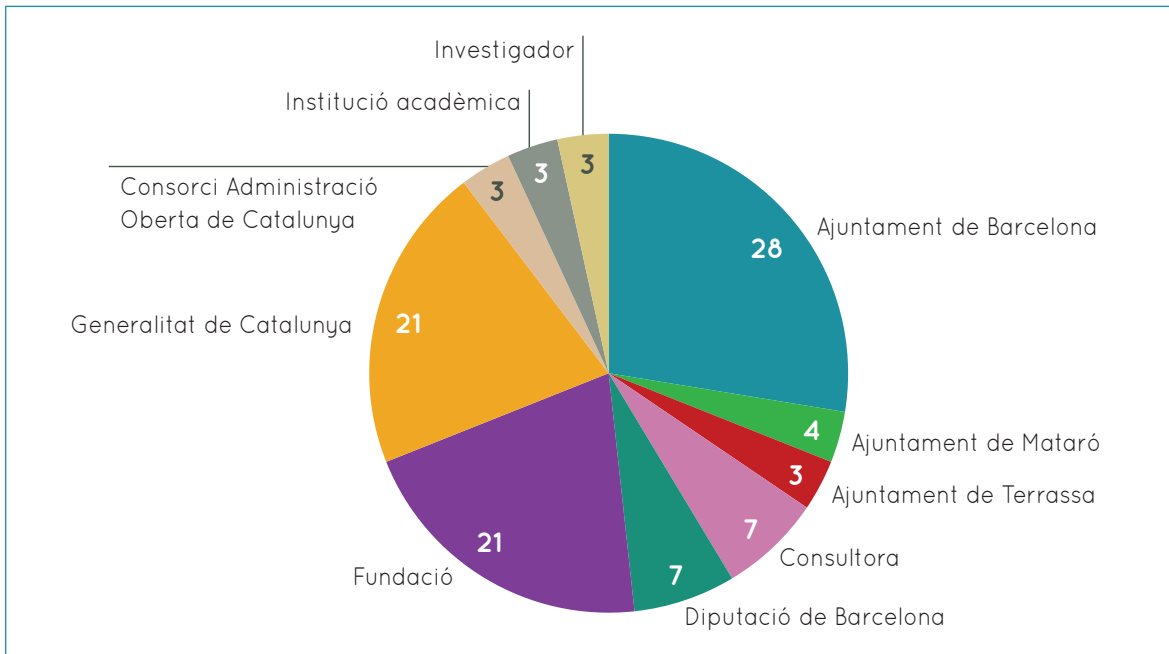
Un cop identificats aquests perfils, se'ls ha contactat per tal d'obtenir la seva col·laboració. En aquest contacte, s'ha ofert dues opcions per facilitar la informació a l'equip investigador: omplint el formulari i/o concertant una entrevista per a resoldre dubtes o aprofundir en qüestions concretes vinculades a la informació a recollir.

Paral·lelament a la difusió a aquest grup de perfils identificats com a rellevants, s'ha ampliat la difusió del formulari per les xarxes socials amb l'objectiu de maximitzar el nombre d'algoritmes recollits i, per tant, de fer un repositori el més complet possible.

En total s'ha obtingut un total de 13 respostes al formulari, i s'han realitzat diverses entrevistes amb diferents persones i institucions. En concret, s'han organitzat entrevistes amb professionals que treballen a l'Ajuntament de Barcelona, L'Institut Municipal de Serveis Socials de Barcelona, la Genera-

litat de Catalunya, l'Administració Oberta de Catalunya (AOC) i la Fundació TIC Salut Social.

Gràfic 12. Organitzacions on treballen les persones contactades



La informació recollida en el Radar ha estat proporcionada de manera majoritària per les persones contactades que van emplenar el formulari. No obstant això, el sistema DPR va ser registrat per tres experts diferents, per la qual cosa les seves respostes van ser combinades en una de sola. El registre de RisCanvi era molt breu, per la qual cosa es va refer per complet a partir de la informació pública trobada. Per a la resta de sistemes es van fer modificacions o ajustos menors en la redacció.

3. EL DISSENY DEL RADAR D'ALGORITMES

El radar o repositori s'ha dissenyat tenint en compte tant referències internacionals com decisions de l'equip investigador de KSNET. Entre les primeres, posseeixen una gran rellevància el repositori Casos d'IA seleccionades del sector públic d'AI-WATCH del Joint Research Center la Comissió Europea¹⁶ i OASI de la Fundació Ètica¹⁷. Encara que en menor mesura, també s'han tingut en compte els registres d'IA d'Amsterdam i Hèlsinki¹⁸.

Els dos primers casos mencionats consisteixen en bases de dades per a registrar la informació recollida dels diversos algoritmes, on el primer –l'AI-WATCH del Joint Research Center la Comissió Europea– d'aquests es centra en el territori europeu i el segon –l'OASI de la Fundació Ètica–, inclou casos internacionals més enllà del territori europeu.

D'aquesta manera, per a la selecció de les variables del Radar d'algoritmes s'ha buscat la comparabilitat amb aquestes dues bases de dades, incorporant camps comuns com són els d'objectiu, estatus i beneficis (Casos d'IA) i riscos i auditoria (OASI).

A la següent taula es recullen els camps inclosos en el radar, que van ser omplerts pels experts contactats a través del formulari en línia, o per KSNET a partir de la informació aportada pels experts durant l'entrevista.

3.1 Camps del Radar d'algoritmes

A continuació, es detallen els camps del Radar. La informació per a cada camp es va recollir mitjançant una pregunta específica del qüestionari.

Bloc 1. Informació Bàsica

1. Nom

En cas de tenir-ho, nom intern o públic de l'algoritme.

2. Descripció

Context en el qual es desenvolupa i s'aplica aquest procés automatitzat/algoritme/sistema d'IA. És a dir, quina problemàtica s'havia detectat, quina va ser la solució per la qual es va optar, com ha estat la implementació i en què consisteix exactament, i quins resultats ha donat fins ara.

¹⁶ Enllaç: <https://data.jrc.ec.europa.eu/dataset/7342ea15-fd4f-4184-9603-98bd87d8239a>.

¹⁷ Enllaç: <https://eticas.tech/oasi>.

¹⁸ Amsterdam: <https://algoritmeregister.amsterdam.nl/>. Hèlsinki: <https://ai.hel.fi/en/ai-register/>.

3. Objectiu

Quin és l'objectiu del procés automatitzat/algoritme/sistema d'IA? Per a quina funció ha estat desenvolupat?

3.1 objectiu

Seleccionar objectius que compleix l'algoritme. Opcions extretes de *Selected AI cases in the public sector* del Joint Research Centre de la Comissió Europea.

Automatització de tasques	Realitzar de forma automatitzada un conjunt de tasques que a una persona portaria molt més temps realitzar.
Recull de dades personals	Recopilar de forma sistemàtica o predeterminada dades sobre individus i/o grups per a fins públicament coneguts o desconeguts i basant-se en criteris públicament coneguts o desconeguts.
Avaluació del comportament humà	Generar avaluacions de la manera com es comporten els individus i/o grups en base a criteris públicament coneguts o desconeguts aplicats a dades públicament conegudes o desconegudes.
Reconeixement de trets facials	Identificar trets facials particulars en imatges de persones, com la forma dels ulls mentre una persona somriu, basant-se en criteris públicament coneguts o desconeguts aplicats a dades públicament conegudes o desconegudes.
Identificació d'imatges de cares	Comparar imatges de cares de persones individuals amb imatges de rostres registrades amb anterioritat en una base de dades basant-se en criteris públicament coneguts o desconeguts aplicats a dades públicament conegudes o desconegudes.
Predicció del comportament humà	Generar possibles escenaris futurs on individus i/o grups puguin comportar-se en base a criteris públicament coneguts o desconeguts aplicats a dades públicament conegudes o desconegudes.
Elaboració de perfils i classificació de persones	Generar perfils d'individus i/o grups i classificar-los i ordenar-los en funció de criteris públicament coneguts o desconeguts aplicats a dades públicament conegudes o desconegudes.
Simulació de la parla humana	Generar una parla que s'assembli molt a la manera com les persones parlen per a propòsits públicament coneguts o desconeguts.
Reconeixement d'imatges	Identificar el contingut d'imatges digitals, per exemple, si és una imatge d'un gat o d'un gos, basant-se en criteris públicament coneguts o desconeguts aplicats a dades públicament conegudes o desconegudes.
Generació de traduccions automàtiques	Traduir automàticament textos o discursos escrits d'un idioma a un altre/s.
Generació de resultats de cerca online	Produir una llista ordenada de llocs web o altres recursos en línia en resposta a una consulta de cerca, generalment com a termes de cerca escrits o parlats.

Reconeixement de sons	Identificar el contingut de la parla o altres sons, per exemple, si es tracta d'una persona parlant o d'un animal o objecte en particular, basant-se en criteris públicament coneguts o desconeguts aplicats a dades públicament conegudes o desconegudes.
Xatbot conversacional	Proveir de forma automatitzada respostes a les consultes ciutadanes.

BLOC 2. DESENVOLUPAMENT I IMPLEMENTACIÓ

4. Administració responsable

Administració pública que està utilitzant o utilitzarà l'algoritme.

5. Entitat desenvolupadora

Entitat que ha desenvolupat l'algoritme a nivell tècnic. Per entitat desenvolupadora es fa referència a l'empresa o persona que ha creat el sistema d'ia o procés automatitzat.

6. Estatus

En quina situació es troba la implementació?

- ▶ En definició: s'ha identificat una necessitat i s'està definint la funció (i disseny) de l'algoritme
- ▶ Planificat: el sistema ja està dissenyat, però encara no s'està desenvolupant
- ▶ En desenvolupament: el sistema s'està desenvolupant
- ▶ Pilot: el sistema està sent provat
- ▶ Implementat: el sistema està implementat i en ús
- ▶ Ja no està en ús
- ▶ Un altre: ...

7. Data de començament

Si l'algoritme s'està utilitzant, any en què es va començar a utilitzar. Si esteu en fase pilot, moment en què es va començar la prova. Si està planificat o en desenvolupament, data estimada en què es començarà a utilitzar.

7.1 Data de finalització

En cas que l'algoritme estigui en desús, en quin any va deixar d'utilitzar-se? Si està sota prova pilot, quan finalitza la prova?

BLOC 3. COL·LECTIU, BENEFICIS I RISCOS

8. Col·lectiu

Col·lectiu a qui està dirigit o sobre el que té un major impacte directe l'algoritme. Per exemple, l'algoritme podria estar implementat en un procés que únicament concerneix a menors d'edat.

9. Beneficis

Principals beneficis, tant per a l'administració com per al ciutadà, obtinguts gràcies a la implementació d'aquest algoritme/sistema d'ia.

9.1 Beneficis

Seleccionar els beneficis aportats. Les opcions estan agrupades en 3 categories i són extretes de *Selected AI cases in the public sector* del Joint Research Centre de la Comissió Europea.

[a] Millora dels serveis públics

- ▶ Serveis personalitzats (l'oferta dels serveis s'ajusta millor a les necessitats de la població)
- ▶ Serveis públics (centrats al ciutadà: major accessibilitat del ciutadà als serveis públics)
- ▶ Incrementar la qualitat de la informació i dels serveis públics
- ▶ Serveis públics més sensibles, eficients i rendibles
- ▶ Nous serveis o canals

[b] Millora de l'eficiència administrativa

- ▶ Reducció de costos
- ▶ Capacitat de resposta de l'operació governamental
- ▶ Millora de la gestió dels recursos públics
- ▶ Augment de la qualitat de processos i sistemes
- ▶ Millor col·laboració i millor comunicació
- ▶ Reduir o eliminar el risc de corrupció i abús de la llei per part dels funcionaris públics
- ▶ Permetre una major equitat, honestat i igualtat

[c] Capacitats de govern obert

- ▶ Més transparència de les operacions del sector públic
- ▶ Augment de la participació pública en les accions governamentals i en la formulació de polítiques
- ▶ Millora del control públic i la influència en les accions i polítiques governamentals

10. Riscos

En la seva opinió, quins són els principals riscos, tant per a l'administració com per al ciutadà, de la implementació d'aquest algoritme/sistema d'ia?

10.1 Riscos

Seleccionar els riscos que presenta o podria comportar l'ús d'aquest algoritme/sistema d'IA. Categories extretes del repositori OASI de Eticas.

- ▶ Discriminació de gènere
- ▶ Discriminació racial
- ▶ Discriminació religiosa
- ▶ Discriminació socioeconòmica
- ▶ Un altre tipus de discriminació
- ▶ Polarització social / radicalització
- ▶ Vigilància estatal
- ▶ Amenaça a la privacitat
- ▶ Generar addicció
- ▶ Manipulació / canvis en el comportament
- ▶ Difondre informació errònia
- ▶ Un altre: ...

BLOC 4. SUPERVISIÓ I TRANSPARÈNCIA

11. Intervenció humana

Quin rol juga la supervisió humana?

- ▶ El resultat de l'algoritme és supervisat per una persona que té la decisió final
- ▶ El resultat de l'algoritme és consultat per una persona, però aquesta no pot modificar la decisió

- ▶ El resultat de l'algoritme és supervisat/consultat per una persona, però aquesta no té el coneixement/capacitat per avaluar-lo
- ▶ El resultat no es supervisa/consulta
- ▶ Una altra: ...

11.1 Intervenció humana (detalls)

En cas de respondre la tercera opció, explicar per què no es supervisa. Si ha respost amb una opció alternativa, aportar més detalls sobre el grau d'intervenció humana.

12. Auditoria

L'algoritme ha superat un sistema d'auditoria? Seleccionar una opció si ha passat una auditoria:

- ▶ Es va realitzar o està passant actualment una auditoria interna
- ▶ Es va realitzar o està passant actualment una auditoria externa
- ▶ Pròximament passarà una auditoria interna (ja prevista)
- ▶ Pròximament passarà una auditoria externa (ja prevista)
- ▶ Una altra: ...

13. Transparència

Grau d'accessibilitat a la informació d'aquest algoritme. Existeix algun document públic i/o pàgina web on s'exposi la següent informació? (Marqui les que sí)

- ▶ dades utilitzades per entrenar l'algoritme
- ▶ tipus d'algoritme utilitzat
- ▶ problemàtica a la qual es busca donar solució
- ▶ com s'està implementant
- ▶ objectius i població que interactua amb l'algoritme
- ▶ resultats obtinguts fins ara
- ▶ una altra: ...

14. Visibilitat del ciutadà

El ciutadà afectat directament o indirectament per la decisió/resultat de l'algoritme és conscient que s'està usant aquest algoritme? En altres paraules, no implica saber si el ciutadà efectivament el coneix, sinó sobre si s'ha fet pública la informació, aquesta és fàcilment accessible i està clarament visible.

- ▶ S'informa directament el ciutadà (comunicació activa)
- ▶ La informació és pública i accessible (comunicació passiva)
- ▶ La informació és pública però poc accessible (depèn de la proactivitat del ciutadà per a informar-se)
- ▶ No s'informa

15. Enllaços i fonts

Existeix algun document públic i/o pàgina web on s'inclogui informació sobre la utilització d'aquest algoritme?

4. RESULTATS

A continuació, es mostren les fitxes dels casos, és a dir, tota la informació recollida per als **12 sistemes d'intel·ligència artificial identificats**. La informació de cada sistema està sintetitzada en una taula on cada fila representa un dels camps explicats anteriorment.

En la següent taula es recullen els noms dels algoritmes juntament amb l'administració que els ha posat en funcionament:

Algoritme	Administració
1 Sistema d'identificació de demandes, problemes i respostes (DPR)	Ajuntament de Barcelona i Institut Municipal de Serveis Socials (pertanyent a l'Ajuntament)
2 Sistema de suport a la classificació d'incidències - MARIO	
3 Processador de casos IRIS	
4 Sistemes conversacionals, infoCanal Alzheimer i Fons 0-16	Ajuntament de Barcelona
5 Transcripció d'informes d'emergència social (CUESB)	
6 RisCanvi	
7 Motor de Regles de Prestacions Socials	Generalitat de Catalunya - Àrea de Tecnologies de la Informació i les Comunicacions
8 Recomanador de Prestacions	
9 Identificació necessitats en la Intervenció Social	
10 Xatbots conversacionals	
11 Videoidentificació per a l'obtenció de l'idCat Mòbil	Administració Oberta de Catalunya (AOC)
12 Servei d'automatització dels informes de pobresa energètica	

Dels 12 sistemes d'IA mapats, 6 es troben en el nivell local i 6 en el nivell autonòmic. L'únic municipi amb sistemes d'IA és Barcelona. A més, 7 ja estan implementats i els restants, en fase pilot.

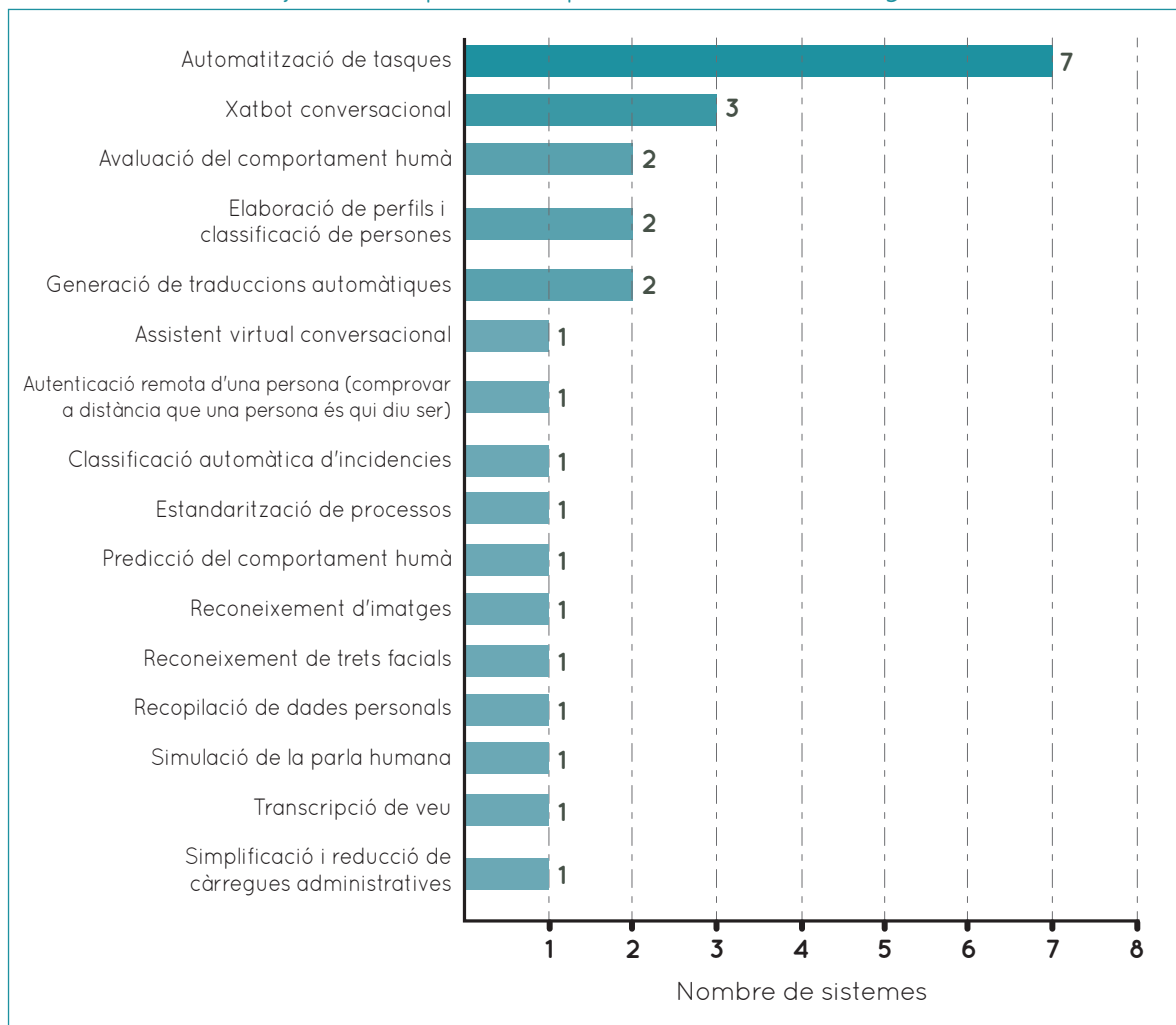
Pel que fa al grau d'intervenció humana, només 1 sistema no té cap mena de supervisió humana (sistema MARIO), però tenint en compte que es limita a recomanar a l'usuari com classificar la seva incidència en una sèrie de temàtiques, és raonable que no requereixi supervisió humana. En la resta de

casos, en funció de l'ús, o bé se supervisen els resultats periòdicament o bé és una persona qui té la decisió final. Així mateix, fins a 8 dels sistemes ja han passat o passaran recentment una auditoria, i també en 9 casos existeix informació pública sobre les principals característiques dels sistemes.

Per tant, en termes generals, i malgrat trobar-nos en una fase molt recent de desenvolupament, implementació i ús d'aquestes tecnologies en l'àmbit de l'Administració Pública catalana, **el grau de seguretat (entès com el control del sistema mitjançant auditories i supervisió, així com per la difusió que fan d'aquestes eines) és considerable, la qual cosa mostra un punt de partida positiu.**

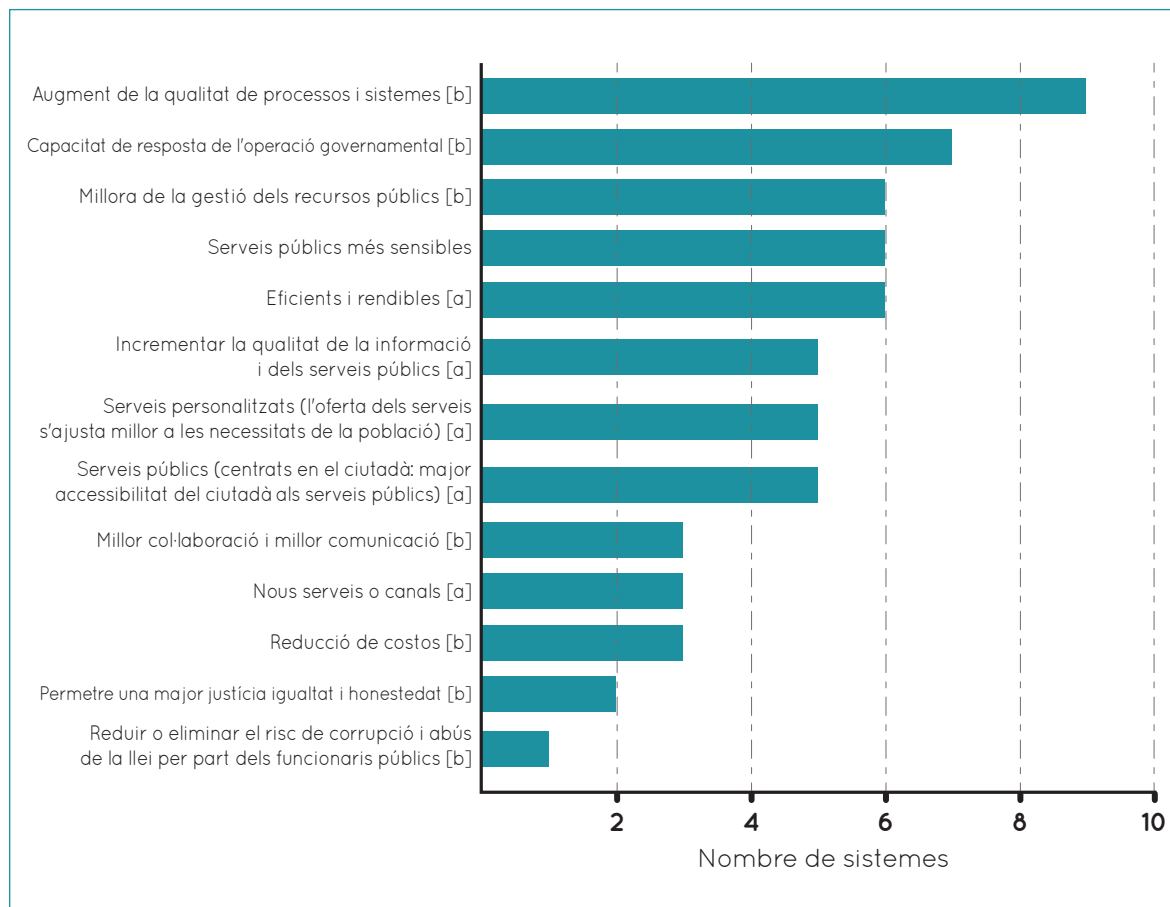
Els objectius que compleixen aquests sistemes d'IA, classificats segons els ítems utilitzats pel Joint Research Center (JRC), són diferents. No obstant això, destaca que 7 dels 12 sistemes recollits, tenen per objectiu l'automatització de tasques. Per tant, els guanys en eficiència i alliberament del temps semblen tenir un pes important en la decisió d'implementar aquests sistemes.

Gràfic 13. Els objectius que compleixen els d'Intel·ligència Artificial



Pel que fa als beneficis, i d'acord també amb les categories extretes del JRC, la major part dels experts han destacat l'augment de la qualitat dels processos i sistemes en 9 dels 12 algoritmes estudiats, seguit per la capacitat de resposta de l'operació governamental.

Gràfic 14. Els beneficis que aporten els sistemes d'Intel·ligència Artificial



Per acabar, els riscos identificats pels experts són pocs. En aquest punt també és important mencionar que, tal com es va destacar en més d'una entrevista, qualsevol procés implementat per les AA.PP. ha de ser garant. Això fa que es redueixi notòriament el ventall de sistemes que es poden utilitzar. Només en 3 dels sistemes recollits s'ha considerat que podrien suposar una amenaça a la privacitat, mentre que dos sistemes podrien comportar un risc de difondre informació errònia. Les discriminacions per raó de raça, nivell socioeconòmic o altres classes també han estat identificades en un o dos sistemes.

Risc identificat per l'expert	Nombre de sistemes d'IA que ho compleixen
Amenaça a la privacitat	3
Difondre informació errònia	2
Altres tipus de discriminació	2
Discriminació socioeconòmica	1
Discriminació racial	1

Fitxes informatives

A continuació, es mostren les fitxes informatives dels 12 sistemes identificats.

1. Intel·ligència col·lectiva - Sistema d'identificació de demandes, problemes i respostes (DPR automàtic)

CAMP	RESPOSTA
1. Nom	<p>Intel·ligència col·lectiva - Sistema d'identificació de demandes, problemes i respostes (DPR automàtic)</p>
2. Descripció	<p>El projecte de DPR automàtic s'emmarca en una iniciativa més àmplia d'implementació d'un sistema de suport a les decisions dels i les professionals basat en eines d'intel·ligència artificial avançada, que, recollint l'experiència passada (intel·ligència col·lectiva), realitzi propostes de diagnosi i de recursos. Es tracta d'un projecte que utilitza processament del llenguatge natural (dades no estructurades) per donar resposta a la necessitat d'agilitzar el temps en el registre de dades requerides al sistema d'informació utilitzat pels professionals d'atenció social.</p> <p>L'Àrea de Drets Socials de l'Ajuntament de Barcelona atén una mitjana de cinquanta mil primeres visites a l'any. Les persones que acudeixen als quaranta centres de serveis socials repartits per la ciutat tenen problemes econòmics, de dependència, per malaltia mental, d'alcoholisme, poden necessitar ajuda psicològica, d'adaptació, poden patir una situació de violència de gènere, etc. Aquestes problemàtiques tan diverses són ateses per una plantilla de més de set-cents professionals, entre treballadors/ores socials, psicòlegs/òlogues i educadors/ores socials.</p> <p>Quan la persona arriba al centre, se l'atén en unes cabines privades. El treballador o treballadora social enregistra la conversa i, en acabar, transcriu la problemàtica, així com l'ajuda o servei al qual se l'ha derivat. En el sistema intern, la informació s'ordena amb tres lletres: demanda (D), problema (P), recurs (R). Actualment, l'Ajuntament disposa de centenars de milers d'entrevistes, moltes de les quals acaben sent repetitives perquè els problemes s'assemblen força.</p> <p>El DPR automàtic dóna suport al procés de registre del DPR: Demandes, Problemes i Respostes (allò que demana la persona atesa, els problemes que identifiquen els i les professionals i les respostes que es proposen per a fer-hi front).</p> <p>La codificació d'aquest DPR és viscut per les professionals com quelcom que es fa pel fet de ser obligatori, però que aporta poc valor afegit a nivell d'atenció. Tanmateix, aquestes dades són de gran utilitat per a l'organització (per a l'extracció d'indicadors); tot i que també podrien servir per a aportar informació als i les professionals (per veure evolució del cas, en traspàs de casos...).</p>

CAMP

RESPOSTA

El projecte es va iniciar amb una prova pilot al 2018 a tres Centres de Serveis Socials. La prova va ser un èxit, tant a nivell qualitatiu (adequació de les propostes ofertes pel sistema) com quantitatiu (ús que se'n va fer): la seva eficàcia i el seu nivell de concordança va ser molt elevat (entre el 75% i el 95% en funció de la categoria) i es va detectar un major volum de multi proposta (més d'una D, P, R per cas) que per mitjà del registre manual.

Un cop finalitzat el període de prova, es va decidir mantenir la funcionalitat a aquests 3 centres de la ciutat a l'espera de fer extensiva l'eina a la resta dels territoris.

L'octubre del 2020 es va produir un canvi en la codificació dels diccionaris de problemàtiques i recursos emprats pels centres de serveis socials. Això va implicar un nou entrenament pel sistema, donat que havia de fer noves associacions. Sota l'ús dels nous diccionaris, el juny de 2021 es va demanar als centres que enregistressin de forma obligatòria tant les demandes, com les problemàtiques i els recursos durant un període de temps, i es va tornar a entrenar el sistema.

Arran de la revaloració positiva de la qualitat d'encert de la proposta automàtica, es va desplegar a l'octubre de 2021 aquesta nova funcionalitat al conjunt de centres de la ciutat de Barcelona, així com per al servei d'atenció a la dependència i l'Equip de Demandes d'Estudi d'Infància i Adolescència en Risc.

Durant l'any 2022 es va dur a terme un procés d'internalització del Sistema d'intel·ligència col·lectiva dins la infraestructura corporativa de l'Ajuntament. En aquest procés es van definir processos de mesurament de la qualitat i de reentrenament automàtic i manual que, a més, permeten monitoritzar el grau d'èxit de l'algoritme.

Actualment l'ús de la funcionalitat del DPR automàtic forma ja part del dia a dia de la feina dels equips de professionals, considerant-se ja com una eina consolidada en el sistema i part del journey de treball dels professionals d'atenció.

2. Descripció

El DPR automàtic és una eina integrada dins el sistema d'informació dels centres de serveis socials que serveix per automatitzar el procés de codificació dins el sistema de les demandes rebudes, les problemàtiques detectades i les propostes de prescripció de recursos que fan els professionals. Per tant, l'objectiu principal és servir de suport a la decisió del treballador/a social.

3. Objectiu I

El sistema ha estat entrenat amb un repositori de tres-centes mil entrevistes amb tècniques d'aprenentatge automàtic, bàsicament els registres de les conclusions de les cites realitzades pels professionals. A partir de les conclusions de l'entrevista introduïdes per un professional, el sistema, que ha estat entrenat prèviament, analitza tot el seu contingut i busca concordançes. A partir d'aquestes concordançes busca què havien informat els i les professionals en situacions similars i proposa Demandes, Problemes i Respostes.

A partir d'aquí el professional valida aquestes propostes, podent, després de validar o informar de nou la D, demanar al sistema que afini P i R altra vegada.

CAMP	RESPOSTA
	D'aquesta manera, els i les professionals s'estalvien la tasca de buscar entre els diccionaris de demandes, problemes i respostes definits, ja que el sistema els fa una proposta a partir de la interpretació de les conclusions de les cites i l'històric d'informació introduïda al sistema.
3. Objectiu I	Aquest projecte converteix una acció que fins al moment s'havia de fer de forma manual, en una tasca automàtica, requerint únicament la validació dels i les professionals. Amb aquesta eina es redueixen les hores destinades a la gestió per part dels equips de professionals, per tal de poder augmentar el temps dedicat a l'atenció social, i alhora s'aconsegueix un major registre i sistematització de la informació.
4. Objectiu II	<ul style="list-style-type: none"> ▶ Automatització de tasques ▶ Avaluació del comportament humà ▶ Elaboració de perfils i classificació de persones
5. Adm. Responsable	Institut Municipal de Serveis Socials - Ajuntament de Barcelona Àrea de Drets Socials - Ajuntament de Barcelona
6. Entitat desenvolupadora	El projecte va estar liderat en la seva fase pilot per la Direcció d'Innovació de l'àrea de Drets Socials, en col·laboració amb l'Institut Municipal de Serveis Socials i executat per l'empresa INNOVA. La fase d'interiorització de l'algoritme va ser gestionada per l'Institut Municipal d'Informàtica en col·laboració amb l'empresa Accenture.
7. Estatus	Implementat: el sistema està implementat i en ús
8. Data començament	2018 en format pilot, 2019 a l'àrea de drets socials i l'octubre de 2021 de forma definitiva per tots els centres de Barcelona
9. Data finalització	
10. Col·lectiu	Treballadors/es socials
11. Benerificis I	<ol style="list-style-type: none"> 1) Ajudar als treballadors/es socials a categoritzar les entrevistes amb els ciutadans 2) Sistematització dels serveis/prestacions prescrits a partir d'una visió neutral i objectiva de les demandes. El sistema ofereix ajuda guiant al professional amb propostes basades en l'experiència anterior (intel·ligència col·lectiva), de manera que el DPR aporta un valor afegit al procés d'atenció 3) Eficiència de recursos destinats: facilita el procés de registre de DPR, redueix el temps destinat a aquesta tasca i facilita el multi-registre.

CAMP	RESPOSTA
12. Benerificis II	<ul style="list-style-type: none"> ▶ Incrementar la qualitat de la informació i dels serveis públics [a] ▶ Serveis personalitzats (l'oferta dels serveis s'ajusta millor a les necessitats de la població) [a] ▶ Serveis públics (centrats en el ciutadà: major accessibilitat del ciutadà als serveis públics) [a] ▶ Serveis públics més sensibles, eficients i rendibles [a] ▶ Millora de la gestió dels recursos públics [b] ▶ Augment de la qualitat de processos i sistemes [b] ▶ Capacitat de resposta de l'operació governamental [b]
13. Riscos I	<p>Pot haver-hi algun biaix, ja que els treballadors socials són humans que tenen el seu propi biaix, però es minimitza gràcies a un gran volum d'entrevistes i al seu poc impacte en les persones.</p> <p>Les entrevistes formen part del sistema corporatiu municipal/estàndards de ciberseguretat. El model està allotjat en un servidor extern, però no les dades.</p> <p>Atès que el sistema entra en joc en una 2a fase -després que ja s'ha donat una resposta- el risc es pot associar a la mala gestió del servei/procés, no tant amb el sistema algorítmic. El procés en si, i la manera com s'utilitza el sistema, ja mitiguen els riscos.</p>
14. Riscos II	
15. Intervenció humana	El resultat de l'algorisme és supervisat per una persona que té la decisió final
16. Intervenció humana (detalls)	<p>El sistema ofereix una proposta al professional de Demandes, Problemes i Respostes.</p> <p>A partir d'aquí el professional pot validar aquesta proposta o informar de nou la D, demanar al sistema que afini P i R altra vegada. La decisió està per tant sempre en mans del professional.</p>
17. Auditoria	Se li va realitzar o està passant actualment una auditoria externa
18. Transparència	<ul style="list-style-type: none"> ▶ Problemàtica a la qual es busca donar solució ▶ Objectius i població que interactua amb l'algorisme ▶ Com s'està implementant ▶ Resultats obtinguts fins ara
19. Visibilitat	La informació és pública però poc accessible (depèn de la proactivitat del ciutadà per a informar-se)
20. Enllaços	<p>https://ajuntament.barcelona.cat/mesames/noticia/posem-en-marxa-una-prova-pilot-dintel%C2%B7ligencia-col%C2%B7lectiva-a-tres-centres-de-serveis-socials/</p> <p>https://ajuntament.barcelona.cat/dretsocials/es/innovacion-social/inteligencia-colectiva</p> <p>https://ajuntament.barcelona.cat/premsa/wp-content/uploads/2021/02/Mesura-de-Govern-dInnovacio-Social.pdf</p>

2. Processador de casos IRIS

CAMP	RESPOSTA
1. Nom	Processador de casos IRIS
2. Descripció	Eina de gestió interna que permet accelerar el procés de categorització en temàtiques de les queixes, incidències i suggeriments que arriben a l'Àrea de Drets Socials a través de la plataforma IRIS. El sistema també ofereix tendències de quines són les temàtiques més recurrents segons diversos criteris.
3. Objectiu I	L'objectiu és millorar l'extracció de dades rellevants a partir de les consultes/queixes que arriben a través del sistema IRIS, bé sigui utilitzant sistemes de classificació automatitzats, o creant consultes amb llenguatge natural.
4. Objectiu II	Automatització de tasques
5. Adm. Responsable	Ajuntament de Barcelona - Àrea de Drets Socials
6. Entitat desenvolupadora	Secretaria Tècnica d'Innovació Social - Àrea de Drets Socials
7. Estatus	Pilot: el sistema està sent provat
8. Data començament	2022
9. Data finalització	En curs
10. Col·lectiu	Gestors de l'Àrea de Drets Socials
11. Benerificis I	Millorar en l'eficiència, donar respostes més concretes i tenir un coneixement més profund a l'organització de les consultes/queixes que arriben
12. Benerificis II	<ul style="list-style-type: none"> ▶ Serveis públics més sensibles, eficients i rendibles [a] ▶ Capacitat de resposta de l'operació governamental [b] ▶ Augment de la qualitat de processos i sistemes [b]
13. Riscos I	No hi ha riscos derivats ja que el sistema de classificació és d'ús intern i per facilitar l'explotació de dades.
14. Riscos II	
15. Intervenció humana	El resultat de l'algoritme és supervisat per una persona que té la decisió final
16. Intervenció humana (detalls)	L'aplicació fa un suggeriment de temes tractats a la queixa, incidència o suggeriment enviada pel ciutadà. Poden ser editats o rebutjats.
17. Auditoria	Se li va realitzar o està passant actualment una auditoria interna

CAMP	RESPOSTA
18. Transparència	
19. Visibilitat	No s'informa
20. Enllaços	De moment no hi ha cap informació pública ja que s'està en fase de pilot.

3. Sistema de suport a la classificació d'incidències - MARIO

CAMP	RESPOSTA
1. Nom	Sistema de suport a la classificació d'incidències - MARIO
2. Descripció	<p>L'IRIS és el servei que permet a la ciutadania comunicar incidències o enviar reclamacions a l'Ajuntament de Barcelona perquè les resolgui. A través de l'IRIS, la ciutadania pot fer arribar informacions i consultes, però també queixes i suggeriments de millora. En aquest procés, el ciutadà o ciutadana que avisa de la incidència ha de classificar-la utilitzant un arbre de temàtiques que se li ofereix a través de l'aplicació informàtica. Aquesta classificació és important perquè s'utilitza per dirigir la incidència directament al departament responsable, i agilitzar així el procés de resposta. Els errors en la classificació temàtica ocasionen respostes inadequades i retards en la resolució de les incidències, afectant així a la qualitat del servei ofert. Dins del projecte d'actualització del servei IRIS, s'ha desenvolupat un mòdul anomenat MARIO, basat en algoritmes d'aprenentatge automàtic —<i>Machine Learning</i>— i processament del llenguatge natural (una de les tecnologies emmarcades en la IA) per simplificar al ciutadà o ciutadana el procés de classificació de les incidències. A partir de l'anàlisi del text lliure que descriu la incidència, MARIO suggereix al ciutadà o ciutadana les categories més probables on la incidència hi té cabuda perquè aquest/a triï la categoria més adequada.</p>
3. Objectiu I	Minimitzar la taxa d'error en la classificació inicial de les incidències i així reduir els processos de classificació manual
4. Objectiu II	Classificació automàtica d'incidències
5. Adm. Responsable	Direcció de serveis d'informació i atenció ciutadana - Ajuntament de Barcelona
6. Entitat desenvolupadora	
7. Estatus	Implementat: el sistema està implementat i en ús

CAMP	RESPOSTA
8. Data començament	Abril 2021
9. Data finalització	
10. Col·lectiu	Ciutadania amb accés a mitjans digitals
11. Benerificis I	Estalviar la tasca manual del personal funcionari reclassificant incidències i agilitzar la resposta a de l'administració.
12. Benerificis II	Incrementar la qualitat de la informació i dels serveis públics [a], Augment de la qualitat de processos i sistemes [b]
13. Riscos I	
14. Riscos II	
15. Intervenció humana	El resultat no es supervisa/consulta
16. Intervenció humana (detalls)	No es supervisa donat que, en principi, l'algoritme ja orienta al ciutadà cap a una classificació adequada per la incidència que està tramitant.
17. Auditoria	
18. Transparència	Problemàtica a la qual es busca donar solució
19. Visibilitat	La informació és pública i accessible (comunicació passiva)
20. Enllaços	https://media-edg.barcelona.cat/wp-content/uploads/2023/08/03111619/drets-digitals.pdf

4. Sistemes conversacionals, infoCanal Alzheimer i Fons 0-16

CAMP	RESPOSTA
1. Nom	Sistemes conversacionals, infoCanal Alzheimer i Fons 0-16
2. Descripció	És un canal d'informació nou present en les webs municipals per tal de donar informació a la ciutadania sobre l'Alzheimer i sobre el Fons 0-16. En ambdós casos les respostes dels sistemes conversacionals es construeixen a partir de fonts validades i pròpies.
3. Objectiu I	Oferir a les persones informació seleccionada a través de converses en llenguatge natural.
4. Objectiu II	Xatbot conversacional
5. Adm. Responsable	Ajuntament de Barcelona - Àrea de Drets Socials

CAMP	RESPOSTA
6. Entitat desenvolupadora	Secretaria Tècnica d'Innovació Social - Àrea de Drets Socials
7. Estatus	Implementat: el sistema està implementat i en ús
8. Data començament	2022, l'InfoCanal i 2023 el fons 0-16
9. Data finalització	En curs
10. Col·lectiu	Ciutadania en general
11. Benerificis I	Ampliar els canals d'informació i poder oferir respostes de forma més fàcil
12. Benerificis II	Incrementar la qualitat de la informació i dels serveis públics [a]
13. Riscos i	El principal risc és que el sistema conversacional no doni respostes adequades.
14. Riscos II	Difondre informació errònia
15. Intervenció humana	Totes les interaccions de les persones i les respostes del sistema conversacional es registren i poden ser avaluades.
16. Intervenció humana (detalls)	
17. Auditoria	Es va realitzar o està passant actualment una auditoria interna
18. Transparència	S'informa en els entorns on estan implementats.
19. Visibilitat	La informació és pública i accessible (comunicació passiva)
20. Enllaços	InfoCanal Alzheimer: https://ajuntament.barcelona.cat/sanitatisalut/ca/canal/barcelona-lalzheimer Fons 0-16: https://ajuntament.barcelona.cat/serveissocials/ca/canal/ajuts-durgencia-social-families-amb-infants-de-0-16-anys

5. Transcripció d'informes d'emergència social (CUESB)

CAMP	RESPOSTA
1. Nom	Transcripció d'informes d'emergència social (CUESB)
2. Descripció	En la cerca de solucions per agilitzar tasques de baix valor afegit per part de les professionals de l'Àrea de drets Socials s'identifiquen els processos de transcripció automatitzats mitjançant IA com un element rellevant. Els serveis d'emergència del CUESB realitzen informes després de cada intervenció que a dia d'avui obliga al personal a retornar a l'oficina i omplir formularis no sistematitzats. Amb aquest pilot es busca que les professionals puguin dictar els informes 'in situ' o en qualsevol moment de la intervenció i que aquest es transcrigui.
3. Objectiu I	Agilitzar el procés de creació dels informes a través de sistemes de transcripció.
4. Objectiu II	Transcripció de veu
5. Adm. Responsable	Ajuntament de Barcelona - Àrea de Drets Socials
6. Entitat desenvolupadora	Secretaria Tècnica d'Innovació Social - Àrea de Drets Socials
7. Estatus	En desenvolupament: el sistema s'està desenvolupant
8. Data començament	2024
9. Data finalització	El pilot s'executarà durant 6 mesos i s'avaluarà
10. Col·lectiu	Treballadores del servei d'emergències socials de l'Ajuntament de Barcelona
11. Benerificis I	Reducció de temps d'elaboració dels informes, Agilització de tasques de baix valor afegit
12. Benerificis II	Augment de la qualitat de processos i sistemes [b]
13. Riscos I	El sistema no té riscos associats ja que la transcripció s'ofereix a l'usuari com a complement i és revisable i editable.
14. Riscos II	
15. Intervenció humana	El resultat de l'algoritme és supervisat per una persona que té la decisió final
16. Intervenció humana (detalls)	
17. Auditoria	Pròximament passarà una auditoria interna (ja prevista)
18. Transparència	Per la transcripció s'utilitza un servei extern. La informació és pública a la seva pàgina web.

CAMP	RESPOSTA
19. Visibilitat	No s'informa
20. Enllaços	Servei extern: https://www.speechmatics.com/

6. RisCanvi

CAMP	RESPOSTA
1. Nom	RisCanvi
2. Descripció	<p>El projecte RisCanvi d'avaluació i valoració del risc té els seus orígens fundacionals a finals de l'any 2007, a partir de les recomanacions de la Comissió per l'estudi de les mesures de prevenció de la reincidència en delictes greus (Resolució JUS/2363/200, DOGC n. 4937, de 24 de juliol de 2007). El primer punt de les recomanacions, esmentava '(...) l'establiment d'un procediment tècnic específic de valoració del grau de risc que sigui compartit per totes les institucions competents – administració de justícia, serveis penitenciaris i cossos policials'</p> <p>L'avaluació RisCanvi és universal i continua: s'ha d'avaluar el risc de tots els interns i internes de tots els centres penitenciaris i en llibertat condicional, i aquesta avaluació s'ha d'actualitzar periòdicament. Com a norma general, les avaluacions tenen una vigència màxima de 6 mesos.</p> <p>Les escales RisCanvi (RisCanvi-Screening i RisCanvi-Completa) han estat dissenyades per predir el risc d'aparició de 5 tipus de comportaments: violència auto-dirigida, violència intra-institucional, reincidència general, reincidència violenta i trencament de condemna.</p> <p>Des de la seva implementació el 2008 s'han avaluat un total de 30.832 interns i internes diferents amb un total de 116.680 escales avaluades. Actualment es fa servir la tercera versió de l'algoritme: RisCanvi v3.</p>

CAMP	RESPOSTA
	<p>Els objectius del protocol de valoració RisCanvi es divideixen en objectius generals i específics.</p> <p>Generals:</p> <ol style="list-style-type: none"> 1. Millorar les prediccions individualitzades del risc de violència futura 2. Generalitzar l'aplicació d'eines de valoració del risc com a procediment de treball dels i de les professionals penitenciaris/es 3. Introduir la gestió del risc com a eina d'intervenció 4. Potenciar la coordinació institucional en la detecció i seguiment dels casos <p>Específics:</p> <ol style="list-style-type: none"> 1. Fer prediccions dinàmiques ajustades als canvis dels interns i internes. 2. Identificar interns i internes en risc, amb independència del delicte comès. 3. Introduir tècniques de cribratge i detecció d'un risc de violència de forma ràpida i senzilla. 4. Generalitzar l'avaluació del risc a tota la població penitenciària. 5. Avaluar l'efecte de la intervenció amb programes específics, en relació a la gestió del risc. 6. Homogeneïtzar criteris entre els/les professionals i qualificar la seva capacitat tècnica. 7. Organitzar procediments sistemàtics de valoració de riscos específics. 8. Fer anàlisis retrospectius i de qualitat de les decisions ja realitzades, per tal de millorar el procés de valoració.
3. Objectiu I	
4. Objectiu II	<ul style="list-style-type: none"> ▶ Avaluació del comportament humà ▶ Predicció del comportament humà ▶ Elaboració de perfils i classificació de persones
5. Adm. Responsable	Secretaria de Mesures Penals, Reinserció i Atenció a la Víctima. Departament de Justícia de la Generalitat de Catalunya
6. Entitat desenvolupadora	
7. Estatus	Implementat: el sistema està implementat i en ús
8. Data començament	2008
9. Data finalització	
10. Col·lectiu	La població reclusa
11. Benerificis I	

CAMP	RESPOSTA
12. Beneficis II	<ul style="list-style-type: none"> ▶ Serveis públics més sensibles, eficients i rendibles [a] ▶ Millora de la gestió dels recursos públics [b]
13. Riscos I	
14. Riscos II	<ul style="list-style-type: none"> ▶ Discriminació socioeconòmica ▶ Discriminació racial ▶ Altres tipus de discriminació
15. Intervenció humana	El resultat de l'algoritme és supervisat per una persona que té la decisió final
16. Intervenció humana (detalls)	L'algoritme assigna un nivell de risc que pot ser corregit per la persona.
17. Auditoria	Es va realitzar o està passant actualment una auditoria interna
18. Transparència	<ul style="list-style-type: none"> ▶ Problemàtica a la qual es busca donar solució ▶ Com s'està implementant ▶ Objectius i població que interactua amb l'algoritme ▶ Resultats obtinguts fins ara
19. Visibilitat	La informació és pública però poc accessible (depèn de la proactivitat del ciutadà per a informar-se)
20. Enllaços	<p>https://justicia.gencat.cat/web/.content/home/ambits/reinsercio_i_serveis_peni/manual_aplicacio_protocol_avaluacio_riscanvi.pdf</p> <p>https://digitalfuturesociety.com/es/podcasts/capitulo-1-riscanvi-i-el-algoritmo-de-la-carcel/</p> <p>https://eapc.gencat.cat/web/.content/home/biblioteca/fons-coleccions/docs_eapc/tic/2018/03_bones_practiques_adm_electronica/10179_eRis-Canvi-EAP-marc-2018-PRE-1.pdf</p>

7. Motor de Regles de Prestacions Socials

CAMP	RESPOSTA
1. Nom	Motor de Regles de Prestacions Socials
2. Descripció	El motor de regles és una de les peces de la nova plataforma eSocial (nova plataforma tecnològica per donar resposta a les necessitats de gestió del Departament de Drets Socials) i està pensada per aplicar els criteris pels quals s'atorga una prestació social durant la seva tramitació.
3. Objectiu I	Per verificar si un ciutadà o entitat compleix el criteris per rebre una prestació social. Actualment hi ha 5 algoritmes desenvolupats: Necessitats Bàsiques, Despeses de la Llar, PNC Jubilació, Complement de la PNC Jubilació i Prestació per naixement, acolliment i adopció (famílies)
4. Objectiu II	Automatització de tasques
5. Adm. Responsable	Generalitat de Catalunya - Departament de Drets Socials
6. Entitat desenvolupadora	Centre de Telecomunicacions i Tecnologies de la Informació (CTTI)
7. Estatus	Implementat: el sistema està implementat i en ús
8. Data començament	Desembre de 2018
9. Data finalització	
10. Col·lectiu	Estan desenvolupats 5 algoritmes: Necessitats Bàsiques, Despeses de la Llar, PNC Jubilació, Complement de la PNC Jubilació i Prestació per naixement, acolliment i adopció (famílies). I es beneficien tots els ciutadans que necessiten aquesta prestació, és a dir, gent gran, famílies o ciutadans en situació d'exclusió social.
11. Beneficis I	Eficiència en la tramitació. Reducció dels terminis de tramitació.
12. Beneficis II	<ul style="list-style-type: none"> ▶ Serveis públics (centrats en el ciutadà: major accessibilitat del ciutadà als serveis públics) [a] ▶ Capacitat de resposta de l'operació governamental [b] ▶ Millora de la gestió dels recursos públics [b] ▶ Augment de la qualitat de processos i sistemes [b]
13. Riscos I	No s'han detectat
14. Riscos II	
15. Intervenció humana	Es supervisa el bon funcionament de l'algoritme de manera periòdica
16. Intervenció humana (detalls)	
17. Auditoria	Pròximament passarà una auditoria interna (ja prevista)

CAMP	RESPOSTA
18. Transparència	<ul style="list-style-type: none"> ▶ Tipus d'algoritme utilitzat ▶ Problemàtica a la qual es busca donar solució ▶ Com s'està implementant, objectius i població que interactua amb l'algoritme ▶ Resultats obtinguts fins ara
19. Visibilitat	
20. Enllaços	https://ctti.gencat.cat/ca/detalls/detallnoticia/Pla_Transformacio_Digital_Drets_Socials_noticia_2023

8. Recomanador de Prestacions

CAMP	RESPOSTA
1. Nom	Recomanador de Prestacions
2. Descripció	Facilitar l'accés a la informació de prestacions socials
3. Objectiu I	Per implementar un xatbot que ajudi a la ciutadana a conèixer quin tipus d'ajut pot rebre de l'administració.
4. Objectiu II	<ul style="list-style-type: none"> ▶ Simulació de la parla humana ▶ Generació de traduccions automàtiques ▶ Xatbot conversacional
5. Adm. Responsable	Generalitat de Catalunya - Departament de Drets Socials
6. Entitat desenvolupadora	Centre de Telecomunicacions i Tecnologies de la Informació (CTTI)
7. Estatus	Pilot: el sistema està sent provat
8. Data començament	Juliol 2023
9. Data finalització	
10. Col·lectiu	A tots els ciutadans en general
11. Beneficis I	Per l'administració és crear un altre canal de comunicació amb la ciutadania.

CAMP	RESPOSTA
12. Benerificis II	<ul style="list-style-type: none"> ▶ Serveis personalitzats (l'oferta dels serveis s'ajusta millor a les necessitats de la població) [a] ▶ Incrementar la qualitat de la informació i dels serveis públics [a] ▶ Nous serveis o canals [a] ▶ Capacitat de resposta de l'operació governamental [b]
13. Riscos i	Compartició de dades personals
14. Riscos II	
15. Intervenció humana	El resultat de l'algoritme es supervisat periòdicament
16. Intervenció humana (detalls)	
17. Auditoria	Pròximament passarà una auditoria interna (ja prevista)
18. Transparència	
19. Visibilitat	
20. Enllaços	

9. Identificació necessitats en la Intervenció Social

CAMP	RESPOSTA
1. Nom	Identificació necessitats en la Intervenció Social
2. Descripció	Durant la intervenció social, l'empleat públic, amb el consentiment del ciutadà, pot activar una funcionalitat de la eina d'intervenció social eSocial, que a través de la transcripció de la part essencial de la conversa amb el ciutadà identifica de manera automàtica les situacions de necessitat.
3. Objectiu I	Facilitar la feina del treballador social.
4. Objectiu II	<ul style="list-style-type: none"> ▶ Automatització de tasques ▶ Recopilació de dades personals ▶ Generació de traduccions automàtiques
5. Adm. Responsable	Generalitat de Catalunya - Departament de Drets Socials
6. Entitat desenvolupadora	Centre de Telecomunicacions i Tecnologies de la Informació (CTTI)
7. Estatus	Pilot: el sistema està sent provat

CAMP	RESPOSTA
8. Data començament	Desembre de 2023
9. Data finalització	
10. Col·lectiu	A tota la ciutadania
11. Benerificis I	Eficiència administrativa, Homogeneïtzació del criteri de l'administració.
12. Benerificis II	<ul style="list-style-type: none"> ▶ Serveis personalitzats (l'oferta dels serveis s'ajusta millor a les necessitats de la població) [a] ▶ Serveis públics (centrats en el ciutadà: major accessibilitat del ciutadà als serveis públics) [a] ▶ Capacitat de resposta de l'operació governamental [b] ▶ Augment de la qualitat de processos i sistemes [b] ▶ Millor col·laboració i millor comunicació [b]
13. Riscos I	
14. Riscos II	
15. Intervenció humana	El resultat de l'algoritme és supervisat per una persona que té la decisió final
16. Intervenció humana (detalls)	
17. Auditoria	
18. Transparència	
19. Visibilitat	S'informa directament el ciutadà (comunicació activa)
20. Enllaços	

10. Xatbots conversacionals de l'Administració Oberta de Catalunya (AOC)

CAMP	RESPOSTA
1. Nom	<p>Xatbots conversacionals de l'Administració Oberta de Catalunya (AOC)</p>
2. Descripció	<p>Problemàtica detectada:</p> <p>Necessitat d'assistència constant: Els empleats públics i la ciutadania requerien suport i informació continua les 24 hores del dia, els 7 dies de la setmana, però l'atenció presencial o telefònica no podia oferir aquesta disponibilitat total.</p> <p>Augment en les consultes i incidències: Amb el creixement dels tràmits digitals i l'ús del portal de transparència, hi havia un augment en la demanda d'informació i assistència.</p> <p>Solució optada i implementació:</p> <p>Desenvolupament d'assistents conversacionals: Per abordar aquesta problemàtica, es va optar per crear xatbots amb capacitat per interpretar i respondre preguntes mitjançant l'ús d'algoritmes d'IA.</p> <p>Entrenament amb dades existents: Els algoritmes van ser entrenats amb informació ja existent, com les preguntes freqüents del portal de suport de l'AOC, estructurades en preguntes-respostes i fluxos guiats de conversa.</p> <p>Implementació a diverses àrees específiques: Els xatbots van ser integrats en diferents àrees relacionades amb el procés d'identificació digital, obtenció, ús i renovació de l'idCat Certificat i navegació pel portal de Transparència i e-TRAM 2.0.</p> <p>Resultats fins ara:</p> <p>Millora en l'experiència de l'usuari: Els usuaris poden accedir a informació i resoldre dubtes de forma més àgil i immediata, reduint la dependència del suport presencial o telefònic.</p> <p>Increment de l'autoresolució: Els xatbots han permès als usuaris resoldre consultes i incidències sense la necessitat d'interaccions humanes, incrementant l'eficiència i alliberant recursos.</p> <p>Supervisió i millora contínua: L'anàlisi dels registres de conversa i els feedbacks dels usuaris han contribuït a millorar els algoritmes, afegir nous fluxos de conversa i ampliar el contingut per adaptar-se millor a les necessitats dels usuaris.</p>

CAMP	RESPOSTA
3. Objectiu I	<p>El sistema té com a objectiu principal facilitar els processos d'identificació digital i l'accés a la informació del portal de transparència, millorar l'experiència dels usuaris i incrementar l'autoresolució de les incidències i consultes dels ciutadans i empleats públics. El sistema està desenvolupat per a 3 funcions específiques:</p> <ol style="list-style-type: none"> 1. Assistència en identificació: es dona suport als usuaris en els processos d'identificació i validació de la identitat necessaris per fer un tràmit digital. 2. Suport a la sol·licitud i gestió de certs documents: ajuda a resoldre els dubtes i preguntes més comuns relacionats amb l'obtenció, ús i renovació de l'idCAT Certificat. 3. Aplicació d'assistents conversacionals per àrees específiques, com el portal de Transparència i e-TRAM 2.0, proporcionant informació específica i facilitant la navegació en aquestes àrees.
4. Objectiu II	<ul style="list-style-type: none"> ▶ Automatització de tasques ▶ Xatbot conversacional ▶ Assistent virtual conversacional
5. Adm. Responsable	Administració Oberta de Catalunya (AOC)
6. Entitat desenvolupadora	ONE MILLION BOT, S.L.
7. Estatus	Implementat: el sistema està implementat i en ús
8. Data començament	21 de gener de 2021
9. Data finalització	
10. Col·lectiu	És un servei de suport a empleats públics i ciutadania
11. Beneficis I	<p>Beneficis per a l'Administració:</p> <p>Optimització de recursos: Reducció de càrrega de treball per als empleats públics, ja que els xatbots gestionen consultes freqüents i repetitives, alliberant-los per a tasques més complexes i especialitzades.</p> <p>Millora de l'eficiència: Augment de l'eficiència en la gestió d'informació i resolució de consultes, permetent una resposta immediata i precisa a un gran volum de consultes simultànies.</p> <p>Anàlisi de dades: L'anàlisi dels registres de converses ofereix una valuosa font de dades per a l'Administració, permetent identificar patrons, àrees de millora i necessitats dels usuaris per adaptar millor els serveis futurs.</p> <p>Reducció de costos: Reducció dels costos associats amb l'atenció al client presencial o telefònica, ja que els xatbots proporcionen assistència digital sense necessitat de recursos humans.</p>

CAMP	RESPOSTA
<p>11. Beneficis I</p>	<p>Beneficis per al ciutadà:</p> <p>Accés 24x7 a la informació: Disponibilitat constant per accedir a informació i assistència sense restriccions d'hores o dies, millorant la comoditat i l'experiència de l'usuari.</p> <p>Resposta immediata: Obtenció de respostes ràpides i immediates a consultes i dubtes, evitant temps d'espera i agilitzant els tràmits o processos que requereixen informació específica.</p> <p>Autoresolució de consultes: Capacitat d'autoresoldre consultes o incidències sense la necessitat de recórrer a personal d'atenció als usuaris (CAU), donant als ciutadans més autonomia en el seu procés d'interacció amb l'Administració.</p> <p>Millora de l'experiència de l'usuari: A través de la navegació guiada i la resposta precisa a preguntes específiques, s'ha millorat l'experiència dels usuaris en la interacció amb els serveis públics, facilitant la comprensió i la resolució de dubtes.</p>
<p>12. Beneficis II</p>	<ul style="list-style-type: none"> ▶ Serveis personalitzats (l'oferta dels serveis s'ajusta millor a les necessitats de la població) [a] ▶ Serveis públics (centrats en el ciutadà: major accessibilitat del ciutadà als serveis públics) [a] ▶ Incrementar la qualitat de la informació i dels serveis públics [a] ▶ Serveis públics més sensibles, eficients i rendibles [a] ▶ Nous serveis o canals [a] ▶ Reducció de costos [b] ▶ Millora de la gestió dels recursos públics [b] ▶ Augment de la qualitat de processos i sistemes [b] ▶ Millor col·laboració i millor comunicació [b]
<p>13. Riscos I</p>	<p>Riscos per a l'Administració:</p> <p>Dependència tecnològica i interrupcions en el servei: Si els xatbots cometen errades o hi ha interrupcions en el funcionament del sistema, això pot provocar una interrupció en l'atenció als usuaris i generar desconfiança en la fiabilitat del servei.</p> <p>Seguretat de les dades i privacitat: Malgrat les mesures de seguretat preses, existeix el risc de vulnerabilitats en la protecció de dades, especialment pel registre de converses, que podrien posar en perill la privacitat dels usuaris (això pot succeir si l'usuari introdueix dades personals innecessàries)</p> <p>Qualitat de les respostes: Si no es gestiona adequadament el material d'entrenament i el procés de millora contínua, els xatbots poden proporcionar respostes inadequades, afectant la confiança en la informació proporcionada i la percepció de l'eficàcia del sistema.</p>

CAMP	RESPOSTA
13. Riscos I	<p>Riscos per als ciutadans:</p> <p>Confiança i comprensió de l'assistència: Els ciutadans poden perdre la confiança en el sistema si perceben que les respostes no són satisfactòries o no són adients per a la seva assistència, especialment en casos complexos o situacions específiques.</p> <p>Protecció de dades i gestió de la informació personal: Hi ha la preocupació que les dades personals dels usuaris puguin ser vulnerables si el sistema no gestiona adequadament la informació recollida o si es produeixen errors en els protocols de seguretat.</p> <p>Comunicació i limitacions en l'assistència: Els xatbots poden tenir limitacions en la comprensió del llenguatge natural o en la capacitat de proporcionar respostes adequades per a totes les situacions, podent afectar la qualitat de l'assistència i la satisfacció de l'usuari.</p>
14. Riscos II	<ul style="list-style-type: none"> ▶ Amenaça a la privacitat ▶ Difondre informació errònia
15. Intervenció humana	El resultat de l'algoritme és consultat per especialistes del CAU (de primer o segon nivell) i per experts en seguiment de l'algoritme, però no poden modificar la resposta donada
16. Intervenció humana (detalls)	Els especialistes del CAU (de primer o segon nivell) reben les qüestions que no s'han pogut respondre i desenvolupen els fluxos de conversa que permetran donar la resposta adequada la propera vegada. Els experts en seguiment supervisen periòdicament els registres de comentaris i converses i analitzen les estadístiques d'ús del servei. A partir d'aquestes dades, els experts identifiquen les qüestions que interessin als usuaris i desenvolupen nous fluxos de conversa i amplien el corpus conversacional amb noves paraules clau
17. Auditoria	
18. Transparència	<ul style="list-style-type: none"> ▶ Dades utilitzades per a entrenar l'algoritme ▶ Tipus d'algoritme utilitzat ▶ Problemàtica a la qual es busca donar solució ▶ Com s'està implementant ▶ Objectius i població que interactua amb l'algoritme ▶ Resultats obtinguts fins ara ▶ Gestió de riscos, informació de contacte per a consultes i del desenvolupador de l'algoritme
19. Visibilitat	S'informa directament el ciutadà (comunicació activa)

CAMP	RESPOSTA
<p>20. Enllaços</p>	<p>La iniciativa de publicar la fitxa de transparència algorísmica dels xatbots conversacionals de l'AOC al Portal de Transparència pretén ser una de les actuacions recomanables a què es refereix la proposta de reglament d'IA i les directrius de la UE per a una IA fiable. La motivació principal que ens ha portat a formular aquesta proposta és vetllar per la transparència dels algoritmes d'IA que apliquem als nostres serveis d'administració digital i perquè les persones estiguin informades i capacitades per utilitzar-los quan hi interactuïn.</p> <p>Fitxa de transparència algorísmica: https://www.aoc.cat/ia-transparencia-xatbot/</p>

11. Servei de videoidentificació per obtenir l'idCAT Mòbil

CAMP	RESPOSTA
<p>1. Nom</p> <p>2. Descripció</p>	<p>Servei de videoidentificació per obtenir l'idCAT Mòbil</p> <p>La problemàtica inicial rau en la dificultat per a les persones d'obtenir una identitat digital (eID) a Catalunya, especialment durant la pandèmia. Amb les restriccions de mobilitat i la necessitat creixent d'interactuar amb les administracions des de casa, l'absència d'una identificació digital generava obstacles i limitacions en les transaccions i comunicacions amb les institucions públiques.</p> <p>La solució va ser el desenvolupament d'un sistema d'IA de videoidentificació per a l'idCAT Mòbil. Aquesta solució permet a les persones obtenir una identitat digital amb validesa oficial per interactuar amb les administracions catalanes des de qualsevol lloc amb accés a Internet. Mitjançant algoritmes de reconeixement facial i la validació de documents oficials, aquest sistema assegura una autenticació remota fiable, equiparable a la presència física.</p> <p>La implementació consisteix en un procés senzill per a l'usuari, que pot realitzar una videoidentificació des del seu telèfon mòbil. El sistema verifica la coincidència entre el document oficial presentat i el vídeo 'selfie' de la persona, assegurant la integritat del procés. Aquesta verificació és posteriorment validada per operadors humans especialitzats per assegurar la correcció i l'eliminació de possibles riscos o problemàtiques.</p> <p>Fins ara, aquest sistema ha proporcionat més de 110.000 identitats digitals amb èxit, resolent la problemàtica inicial d'accés a una identitat digital durant situacions restrictives com la pandèmia. Els resultats mostren una alta eficiència i acceptació per part dels usuaris, oferint una solució viable i segura per a les seves necessitats d'interacció amb les administracions sense necessitat de desplaçament físic.</p>

CAMP	RESPOSTA
3. Objectiu I	<p>Objectiu del sistema d'IA:</p> <p>L'objectiu principal és permetre que qualsevol persona major de 16 anys pugui obtenir una identitat digital per interactuar amb les administracions catalanes des de casa seva, utilitzant només un telèfon mòbil i un document oficial amb fotografia, com un passaport o un document nacional d'identitat.</p> <p>Funció del sistema d'IA:</p> <p>La funció principal del sistema és proporcionar un servei de videoidentificació que, mitjançant l'ús d'algoritmes d'IA, permeti comprovar a distància que una persona és qui diu ser amb una elevada precisió i seguretat. Aquest procés té com a resultat l'alta de la persona en el registre del sistema d'identificació i signatura electrònica idCAT Mòbil, obtenint així una identitat digital per poder tramitar amb qualsevol administració catalana amb garanties de seguretat i privacitat.</p> <p>El sistema està dissenyat per ser accessible 24x7 amb un alt nivell de disponibilitat i ofereix beneficis en comoditat, d'inclusió i major seguretat, tot mantenint una supervisió humana per garantir la correcció del procés i complint amb les normatives legals de protecció de dades i seguretat.</p>
4. Objectiu II	<ul style="list-style-type: none"> ▶ Automatització de tasques ▶ Reconeixement de trets facials ▶ Reconeixement d'imatges ▶ Autenticació remota d'una persona (comprovar a distància que una persona és qui diu ser)
5. Adm. Responsable	Administració Oberta de Catalunya (AOC)
6. Entitat desenvolupadora	DELOITTE ADVISORY, S.L. que aporta la tecnologia Deloitte OBA integrada amb la solució de VERIDAS
7. Estatus	Implementat: el sistema està implementat i en ús
8. Data començament	Maig de 2020
9. Data finalització	
10. Col·lectiu	Ciutadania major de 16 anys

CAMP	RESPOSTA
<p>11. Beneficis I</p>	<p>Beneficis per a l'Administració:</p> <ul style="list-style-type: none"> ▶ Eficiència: Reducció significativa del temps i dels costos administratius associats a la verificació d'identitat i la gestió presencial. ▶ Seguretat: Augment substancial de la seguretat en el procés d'autenticació, amb una supervisió humana que redueix riscos de frau o errors. ▶ Compliment normatiu: Alineament amb les normatives de protecció de dades i certificacions legals per a l'emissió de certificats qualificats, assegurant la fiabilitat i legalitat dels documents emesos. Val a dir que no hi ha normativa específica aplicable a l'ide-Cat Mòbil i l'AOC ha decidit seguir voluntàriament el model dels certificats. <p>Beneficis per al ciutadà:</p> <ul style="list-style-type: none"> ▶ Comoditat: Accés fàcil i còmode a una identitat digital sense necessitat de desplaçaments, amb la possibilitat de completar el procés des de casa amb un telèfon mòbil. ▶ Inclusió: Accessibilitat millorada, amb un procés senzill i orientat a tot tipus de persones independentment de les seves habilitats digitals. ▶ Seguretat i confiança: Garantia d'un procés d'autenticació precís i fiable, amb un alt nivell de seguretat en la verificació d'identitat.
<p>12. Beneficis II</p>	<ul style="list-style-type: none"> ▶ Serveis personalitzats (l'oferta dels serveis s'ajusta millor a les necessitats de la població) [a] ▶ Serveis públics (centrats en el ciutadà: major accessibilitat del ciutadà als serveis públics) [a] ▶ Serveis públics més sensibles, eficients i rendibles [a] ▶ Nous serveis o canals [a] ▶ Reducció de costos [b] ▶ Capacitat de resposta de l'operació governamental [b] ▶ Millora de la gestió dels recursos públics [b] ▶ Augment de la qualitat de processos i sistemes [b] ▶ Permetre una major justícia, igualtat i honestat [b]
<p>13. Riscos I</p>	<p>Riscos per a l'Administració:</p> <ul style="list-style-type: none"> ▶ Protecció de dades sensibles: La gestió de dades biomètriques i personals durant la videoidentificació pot exposar a riscos de seguretat, especialment si hi ha interrupcions en el xifratge o si les dades són susceptibles de ser compromeses durant el procés. ▶ Supervisió i control humà: Malgrat la supervisió dels operadors, hi ha la possibilitat d'errors humans en la validació i verificació dels processos, que podrien resultar en identificacions incorrectes o en la no detecció d'intents de frau. ▶ Suplantació d'identitat: Tot i els controls, existeix el risc que algunes persones puguin fer servir mètodes per enganyar el sistema i suplantar la identitat d'altres individus. ▶ Dependència tecnològica: L'alta dependència d'aquest sistema podria crear vulnerabilitats en cas de fallades del sistema o interrupcions tecnològiques que afectin la continuïtat del servei d'identificació.

CAMP	RESPOSTA
13. Riscos I	<p>Riscos per al ciutadà:</p> <ul style="list-style-type: none"> ▶ Privacitat i seguretat de dades: L'ús de dades biomètriques en el procés pot generar inquietuds sobre la privacitat i la seguretat dels usuaris, especialment si aquestes dades són vulnerables a atacs cibernètics o mal ús per tercers. ▶ Possibles biaixos en el reconeixement facial: Malgrat la baixa taxa de rebuig, encara hi ha una possibilitat de biaixos en el reconeixement facial que podrien conduir a errors en la identificació o a casos de falsos positius o negatius, afectant certs grups de persones de manera desigual. ▶ Dependència de la connectivitat i tecnologia: El procés de videoidentificació depèn de la connexió a Internet i de la tecnologia del dispositiu, la qual cosa podria limitar l'accés a aquelles persones amb connexions deficientes o mancances tecnològiques.
14. Riscos II	<ul style="list-style-type: none"> ▶ Amenaça a la privacitat ▶ Altres tipus de discriminació
15. Intervenció humana	<p>El resultat de l'algoritme és supervisat per una persona que té la decisió final</p>
16. Intervenció humana (detalls)	<p>En concret, les accions que realitzen els operadors que supervisen les transaccions són:</p> <ul style="list-style-type: none"> a. Visualitzar el vídeo i comprovar que la persona interessada ha fet tot el procés sense coacció, que ha fet el procés segons els requeriments establerts (no porta el cap cobert, no porta ulleres de sol, se li veu la cara, etc.) i també revisar que es vegin les marques d'aigua i hologrames d'autenticitat del document mostrat. b. Revisar el resultat dels paràmetres d'avaluació de la documentació que el sistema ha avaluat de manera automàtica: percentatge de coincidència entre la foto del document identificador i la foto selfie, data de caducitat del document identificador, data de naixement, número del document i que sigui llegible i contingui les marques d'aigua.
17. Auditoria	<ol style="list-style-type: none"> 1) Per identificar els riscos associats a la disponibilitat i seguretat del sistema, l'AOC va realitzar, el 7/7/2020, un anàlisi de seguretat del servei de videoidentificació seguint les guies de l'Agència Catalana de Ciberseguretat i de l'Esquema Nacional de Seguretat (ENS). 2) Per identificar els riscos del sistema des del punt de vista dels drets fonamentals, els principis i valors subjacents, la Generalitat de Catalunya va realitzar, en data 17/12/2020, una avaluació d'impacte sobre la protecció de dades (AIPD) i una avaluació sobre la implementació del procés de videoidentificació per obtenir l'idCat Mòbil. 3) Per garantir la imparcialitat dels algoritmes aplicats es realitzen proves de qualitat periòdiques que permeten detectar imprecisions i evitar biaixos. • Prova de proveïdors de reconeixement facial del National Institute of Standards and Technology (NIST) (PDF darrer informe, 16/6/2023)

CAMP	RESPOSTA
18. Transparència	<ul style="list-style-type: none"> ▶ Dades utilitzades per a entrenar l'algoritme ▶ Tipus d'algoritme utilitzat ▶ Problemàtica a la qual es busca donar solució ▶ Com s'està implementant ▶ Objectius i població que interactua amb l'algoritme ▶ Resultats obtinguts fins ara
19. Visibilitat	S'informa directament el ciutadà (comunicació activa)
20. Enllaços	<p>Fitxa d'innovació 'Identificació remota dels ciutadans mitjançant videoidentificació': https://www.aoc.cat/projecte-innovacio/identificacio-remota-dels-ciutadans-mitjançant-videoidentificacio/</p> <p>Fitxa de transparència algorítmica es publicarà en breu al Portal de transparència de l'AOC. [https://www.seu-e.cat/ca/web/consorcio-oc/govern-obert-i-transparencia/accio-de-govern-i-normativa/normativa-plans-i-programes/transparencia-algorismica-234]</p>

12. Servei d'automatització dels informes de pobresa energètica

CAMP	RESPOSTA
1. Nom	Servei d'automatització dels informes de pobresa energètica
2. Descripció	<p>El servei d'automatització dels informes de pobresa energètica és una solució desenvolupada per automatitzar la generació dels informes socials de vulnerabilitat energètica. Aquests informes són essencials per garantir l'accés als subministraments bàsics a persones i unitats familiars en risc d'exclusió residencial, segons la Llei 24/2015, de mesures urgents per a afrontar l'emergència en l'àmbit de l'habitatge i la pobresa energètica (article 6).</p> <p>Per aquest motiu, mensualment, les diferents Àrees Bàsiques de Serveis Socials (ABSS) han de redactar els esmentats informes per acreditar la situació de vulnerabilitat econòmica davant dels talls que puguin acordar les companyies subministradores d'electricitat, gas i aigua potable (més de 50 companyies diferents arreu del territori) per manca de pagament.</p> <p>La problemàtica principal detectada era la complexitat i la dispersió de les dades necessàries per a l'elaboració dels informes de pobresa energètica. Amb l'existència de múltiples fonts de dades i formats heterogenis, això creava un procés manual propens a errors i ineficient en la verificació de la vulnerabilitat energètica de les persones i unitats familiars en risc d'exclusió residencial.</p>

CAMP

RESPOSTA

La solució adoptada va ser la implementació d'un sistema d'IA que utilitza algoritmes de processament de dades i tecnologies com RPA i BPM. Aquest sistema s'ha desenvolupat per automatitzar la recopiació i verificació de dades, la generació dels informes de pobresa energètica i la comunicació del resultat a les empreses subministradores d'energia. La seva implementació implica la integració amb múltiples sistemes, com els serveis d'interoperabilitat Via Oberta o el sistema d'informació de serveis socials Hèstia.

La implementació concreta consisteix en la creació d'una plataforma al núvol proporcionada per una empresa privada (Consultors-BPM), la qual està hostatjada a Microsoft Azure dins la Comunitat Europea. Aquesta plataforma permet carregar els fitxers dels proveïdors energètics, obtenir les dades socioeconòmiques dels titulars i altres passos necessaris per calcular els coeficients i generar automàticament els informes de vulnerabilitat. A més, es realitza la comunicació dels informes a les empreses subministradores via e-Notum.

Fins ara, els resultats mostren un percentatge d'èxit del 90%, tot i que es preveu que amb canvis legals i organitzatius aquesta xifra pugui arribar gairebé al 100%.

2. Descripció

La implementació s'ha posat a prova en sis municipis diferents: El Prat de Llobregat, Amposta, Granollers, Sant Boi de Llobregat, Palau Solità i Plegamans i Santa Coloma de Gramenet.

Els principals resultats obtinguts fins ara són:

1. Estalvi important de temps i tasques administratives: S'ha millorat la velocitat i l'eficiència en la generació dels informes de pobresa energètica, permetent oferir un millor servei a la ciutadania en situació de vulnerabilitat.
2. Major interoperabilitat i millora en la gestió de dades: S'ha aconseguit una millora en la interconnexió entre empreses, municipis i altres administracions, facilitant una visió més precisa de la situació de vulnerabilitat energètica. També una important reducció de carregues administratives dels ciutadans gràcies a la interoperabilitat.
3. Suport als municipis amb recursos limitats: El sistema ofereix suport als municipis que no disposen dels recursos necessaris per a grans projectes de transformació digital, permetent-los accedir a aquest servei a través d'un servei al núvol sense necessitat d'instal·lació local.

CAMP	RESPOSTA
3. Objectiu I	<p>El sistema d'IA té com a objectiu principal automatitzar i estandaritzar la generació dels informes de pobresa energètica per millorar la gestió, la verificació de dades i la presa de decisions en situacions de vulnerabilitat energètica, tot assegurant el compliment de la normativa de protecció de dades i garantint una supervisió humana per assegurar la precisió dels resultats.</p> <p>Pel que fa a la funció, aquest servei permet:</p> <ol style="list-style-type: none"> 1. Estandarditzar i automatitzar la generació dels informes de pobresa energètica. 2. Millorar la gestió de dades i la interoperabilitat entre múltiples fonts i sistemes. 3. Optimitzar els recursos dels municipis per a la verificació de situacions de vulnerabilitat energètica. 4. Assegurar el compliment de normatives de protecció de dades i la regulació vigent. <p>La supervisió humana és essencial per validar els resultats del sistema abans d'enviar-los a les empreses subministradores, assegurant la precisió i evitant possibles errors en casos específics que puguin escapar a la lògica del sistema.</p> <p>Així mateix, el sistema està dissenyat i implementat seguint els requisits de la normativa de protecció de dades, especialment el Reglament General de Protecció de Dades (RGPD) i altres regulacions locals. Complir amb aquests estàndards és crucial, ja que es treballa amb dades personals i s'ha de garantir la transparència, la limitació en la recopilació de dades i l'ús correcte del consentiment de les persones afectades.</p>
4. Objectiu II	<ul style="list-style-type: none"> ▶ Automatització de tasques ▶ Estandardització de processos, simplificació i reducció de càrregues administratives
5. Adm. Responsable	Administració Oberta de Catalunya (AOC)
6. Entitat desenvolupadora	Consultors BPM
7. Estatus	Pilot: el sistema està sent provat
8. Data començament	27 de gener de 2023
9. Data finalització	
10. Col·lectiu	Personal de les àrees bàsiques de serveis socials (ABSS)

CAMP

RESPOSTA

11. Beneficis I

Beneficis per a l'Administració:

Estalvi de temps i recursos: Reducció significativa del temps i les tasques administratives necessàries per generar els informes de pobresa energètica. Això permet una millor gestió dels recursos humans i econòmics de l'Administració, optimitzant el seu rendiment.

Estandardització i millora en la gestió de dades: A través de la integració de múltiples fonts de dades i la seva automatització, s'aconsegueix una visió més precisa i estandarditzada de la situació de vulnerabilitat energètica. Això millora la presa de decisions i facilita la comunicació entre empreses, municipis i altres administracions.

Suport a municipis amb recursos limitats: El servei al núvol ofereix la possibilitat d'accés a una tecnologia avançada sense la necessitat de grans inversions en infraestructura o recursos interns.

Beneficis per als Ciutadans:

Millor servei i agilitat en l'atenció: L'automatització i estandardització del procés permet una resposta més ràpida i precisa als casos de vulnerabilitat energètica, millorant la qualitat dels serveis socials i l'atenció a les persones en situació de risc.

Assegurar l'accés als subministraments bàsics: A través d'aquest sistema, s'agilitza la determinació de la situació de vulnerabilitat, assegurant que les persones i unitats familiars en risc rebin l'atenció i els subministraments bàsics necessaris.

Reducció de les càrregues administratives i millora de l'experiència de l'usuari: El procés automatitzat evita tràmits innecessaris i simplifica la interacció amb l'Administració, reduint la burocràcia i millorant l'experiència dels ciutadans en situacions de vulnerabilitat.

12. Beneficis II

- ▶ Serveis públics més sensibles, eficients i rendibles [a]
- ▶ Reducció de costos [b]
- ▶ Capacitat de resposta de l'operació governamental [b]
- ▶ Millora de la gestió dels recursos públics [b]
- ▶ Augment de la qualitat de processos i sistemes [b]
- ▶ Millor col·laboració i millor comunicació [b]
- ▶ Reduir o eliminar el risc de corrupció i abús de la Llei per part dels funcionaris públics [b]
- ▶ Permetre una major justícia, igualtat i honestedat [b]

CAMP	RESPOSTA
<p>13. Riscos I</p>	<p>La implementació d'aquest sistema d'IA, malgrat els seus avantatges, també pot presentar alguns riscos, tant per a l'Administració com per als ciutadans.</p> <p>Riscos per a l'Administració:</p> <ul style="list-style-type: none"> ▶ Dependència de la tecnologia: Si alguna cosa falla en la plataforma o en el procés automatitzat, podria provocar retards o errors en la generació dels informes, posant pressió addicional sobre els recursos humans per solucionar-los. ▶ Vulnerabilitat i seguretat de dades: La integració de múltiples fonts de dades i l'ús de serveis al núvol poden generar vulnerabilitats de seguretat. Si el sistema no està prou protegit, podria ser susceptible a atacs cibernètics o pèrdua de dades, compromentent la privacitat i la seguretat de la informació dels ciutadans. <p>Riscos per als ciutadans:</p> <ul style="list-style-type: none"> ▶ Possibles errors en l'avaluació de la vulnerabilitat i pèrdua de confiança: Malgrat la supervisió humana, hi ha possibilitat que el sistema cometi errors en casos específics. Això podria resultar en informes inexactes sobre la vulnerabilitat energètica de les persones o unitats familiars, afectant la seva capacitat d'accedir als subministraments bàsics i provocant una pèrdua de confiança en l'Administració. <p>La pèrdua de confiança dels ciutadans en l'Administració és un factor crític que pot sorgir si es produeixen errors freqüents en l'avaluació de la vulnerabilitat. Aquests errors poden generar desconfiança en la fiabilitat dels informes generats, afectar la percepció de transparència i justícia en la presa de decisions i tenir un impacte emocional negatiu en les persones afectades. Si es denega l'accés als subministraments bàsics a causa d'informes incorrectes, això pot provocar situacions de desconfiança i desesperació en aquells que ja es troben en situacions de vulnerabilitat. Per això, és crucial gestionar aquest risc mitjançant una supervisió minuciosa, processos transparents i una comunicació efectiva per no perdre la confiança dels ciutadans en els serveis que ofereix l'Administració en situacions de vulnerabilitat energètica.</p> <ul style="list-style-type: none"> ▶ Problemes de privacitat i consentiment: Treballar amb dades personals implica el risc de violació de la privacitat si no es gestionen adequadament. Cal garantir que s'obtingui i s'utilitzi el consentiment de les persones afectades de manera transparent i conforme a la normativa de protecció de dades.
<p>14. Riscos II</p>	<p>Amenaça a la privacitat</p>
<p>15. Intervenció humana</p>	<p>El resultat de l'algoritme és supervisat per una persona que té la decisió final</p>
<p>16. Intervenció humana (detalls)</p>	
<p>17. Auditoria</p>	

CAMP	RESPOSTA
18. Transparència	<ul style="list-style-type: none"> ▶ Problemàtica a la qual es busca donar solució ▶ Com s'està implementant ▶ Resultats obtinguts fins ara
19. Visibilitat	<p>La informació és pública però poc accessible (depèn de la proactivitat del ciutadà per a informar-se)</p>
20. Enllaços	<ul style="list-style-type: none"> ▶ Post 'Pilot per l'automatització de l'informe de pobresa energètica per evitar talls de subministraments' (27/1/2023, blog AOC): https://www.aoc.cat/blog/2023/pilot-pobresa-energetica/ ▶ Article 'Automating the energy poverty report in Catalunya' (10/10/2023, Plataforma Joinup de la Comissió Europea): https://joinup.ec.europa.eu/collection/public-sector-tech-watch/automating-energy-poverty-report-catalunya ▶ Presentació 'Avaluació automatitzada de la pobresa energètica' (18/10/2023, Semic 2023: Interoperable AI in the Age of AI): https://www.linkedin.com/posts/miquelestape_20231018-semic-2023-trustworthy-ai-for-activity-7122219700981256194-f9es/?utm_source=share&utm_medium=member_desktop ▶ La fitxa de transparència algorítmica es publicarà en breu al Portal de transparència de l'AOC. [https://www.seu-e.cat/ca/web/consorciaoc/govern-obert-i-transparencia/accio-de-govern-i-normativa/normativa-plans-i-programes/transparencia-algorismica-234]

5. VALORACIONS FINALS

En aquest segon document del projecte, **l'objectiu prioritari ha estat la identificació de sistemes d'intel·ligència artificial i processos automatitzats en l'àmbit social, que estiguessin sent utilitzats, o en fase pilot, per part de les Administracions Públiques catalanes.**

Per a aconseguir aquest objectiu va ser necessari, d'una banda, definir amb claredat quina informació volia recollir-se. És a dir, els camps del radar. D'altra banda, va ser necessari dur a terme un treball de camp minuciosos identificant tant persones com institucions i departaments que poguessin tenir coneixement o estar utilitzant aquests sistemes. Arran d'aquest esforç **es van aconseguir mapar fins a un total de 12 algoritmes**, que molt probablement són la major part (si no tots) els sistemes d'IA utilitzats actualment a Catalunya o que estan en fase pilot.

En aquest punt **és imprescindible reivindicar la col·laboració dels professionals de l'Administració Pública.** Els tècnics de la Generalitat de Catalunya, de l'Ajuntament de Barcelona, de l'Institut Municipal de Serveis Socials i de l'Administració Oberta de Catalunya van respondre el qüestionari proporcionant un notable grau de detall sobre els sistemes d'IA, la qual cosa ha permès construir el Radar.

A més, en les converses amb els tècnics van quedar patents les particularitats de l'àmbit dels drets socials. L'obligació de dissenyar processos amb garanties a l'hora de proveir serveis al ciutadà limita considerablement el ventall de tècniques d'intel·ligència artificial que poden aplicar-se. Per exemple, si s'utilitzés un sistema d'IA per decidir si donar o no una prestació a una persona, aquest sistema no hauria de tractar-se d'un algoritme de caixa negra (com una xarxa neuronal o un *random forest*) atès que, almenys avui dia, l'explicabilitat d'aquests models és baixa i, per tant, la capacitat d'explicar la decisió del model també ho és. Per a una aplicació d'aquest tipus, seria preferible recórrer a automatitzacions (estrictament basades en la regulació) o algoritmes explicables (com les regressions, habitualment utilitzades en economia o ciència política).

Els tècnics van transmetre en tot moment que l'obligació de donar respostes clares als ciutadans sobre el perquè d'un resultat condicionava adoptar o no aquest tipus de sistemes. Això queda clarament reflectit al Radar, on bona part dels algoritmes són de 'consum propi', és a dir, estan dirigits a automatitzar procediments o a ajudar als tècnics de l'administració, però no a prendre decisions que afectin directament el ciutadà. És raonable es-

perar que els tipus de solucions que adoptin les Administracions Públiques en l'àmbit social diferiran en molts casos de les quals es poden prendre en altres àrees.

A tall de reflexió final, molt probablement estem davant d'un canvi de paradigma. I en tot procés de transformació social profund, és habitual que part de la societat es divideixi entre la por per la incertesa futura, i un optimisme excessiu i acrític davant els canvis esdevenidors. No és senzill trobar punt mitjà que contribueixi a un desenvolupament efectiu i just. És a dir, una perspectiva que opti per analitzar cas per cas de manera rigorosa i sense apriorismes. **El valor fonamental d'aquest treball és que ha ajudat a obtenir aprenentatges importants a l'hora de comprendre com s'estan aplicant actualment els sistemes d'Intel·ligència Artificial en l'àmbit social.**

En particular, ha permès extreure una sèrie conclusions clau. En primer lloc, **el nombre total de sistemes identificats és reduït, un total de 12.** Diversos factors poden ajudar a entendre per què en aquest àmbit hi ha un menor nombre de sistemes que en d'altres com el de la salut. Probablement la disponibilitat de recursos en sigui un. No obstant això, la relació entre l'Administració Pública i el ciutadà en l'àmbit social segueix una lògica de garanties, per la qual cosa el ventall d'opcions on implementar aplicacions d'Intel·ligència Artificial es veu reduït.

En segon lloc, i derivat de l'anterior, **bona part d'aquests sistemes són de consum intern,** és a dir, qui interactua amb ells és el personal de l'administració i no el ciutadà. D'aquesta manera, aquests sistemes estan **més centrats a produir millores en l'eficiència dels processos administratius.**

En tercer lloc, **els sistemes identificats comporten nivells de risc molt reduït o inexistent.** Els algoritmes aplicats estan dirigits a resoldre problemàtiques que, atesa la seva naturalesa, amb prou feines comporten cap risc per al ciutadà.

En quart lloc, **la col·laboració dels responsables de l'Administració Pública ha estat positiva.** Això ha permès que la major part de la informació recollida provingui directament de persones que han participat en el disseny, gestió o aplicació d'aquests sistemes.

Finalment, **no es pot perdre de vista que la intel·ligència artificial és una eina més, i només és la més òptima, eficient o efectiva per a algunes problemàtiques determinades.** No és un fi en si mateix. Des de l'Administració Pública ha de valorar-se si, per a cada problema, existeix un sistema alternatiu a la IA que sigui millor. Aquesta concepció va ser percebuda amb claredat en les converses mantingudes amb les persones tècniques.



Taula d'entitats
del Tercer Sector Social
de Catalunya

m4Social

Amb el suport de:



Generalitat de Catalunya
Departament
de Drets Socials