

Definició de metodologies de treball i protocols per a la implementació de sistemes algorítmics

ÍNDEX

1. Introducció	3
1.1. Antecedents i motivació	3
1.2. El cicle de vida d'un servei TIC a l'Ajuntament de Barcelona	5
1.3. El cicle de vida d'un sistema algorítmic	5
1.4. Governança	10
2. Concepció del servei	11
2.1. Divulgació d'informació	11
2.2. Assignació del risc	11
2.3. L'estudi d'impacte algorítmic (EIA)	12
2.4. Espais de transparència	13
2.5. Protecció de dades	14
3. Contractació	14
3.1. Qualitat de les dades	15
3.2. Drets sobre les dades	16
3.3. Qualitat del sistema algorítmic	16
3.4. Transparència del sistema algorítmic	17
3.5. Protecció de dades	18
4. Desenvolupament, producció i explotació	19
4.1. Estratègia de gestió de riscos en el desenvolupament del sistema algorítmic	19
4.2. Auditoria i altres tipologies d'inspecció	20
4.3. Despeses	20
4.4. Espais de transparència	20
4.5. Protecció de dades	21
5. Desmantellament	21
5.1. Destrucció de les dades	21
ANNEX 1. Definicions	23
ANNEX 2. Model de SOL·LICITUD	26

1. Introducció

1.1. Antecedents i motivació

L'ús de les tecnologies basades en sistemes algorítmics i intel·ligència artificial¹ ofereix oportunitats a l'Ajuntament de Barcelona per transformar els serveis municipals perquè siguin més eficients, proactius i adaptats a les necessitats de la ciutadania, així com també per agilitzar els processos interns i donar suport a la presa de decisions. No obstant això, el desplegament d'aquest tipus de sistemes duu associats múltiples riscos², per la qual cosa és imprescindible assegurar que la seva implementació introdueix mecanismes de salvaguarda i garantia dels drets fonamentals; que es duu a terme de manera proporcionada, supervisada i fonamentada —d'acord amb la legislació vigent i els estàndards ètics i tècnics adequats— i que el procés afavoreix la participació i deliberació constants amb la ciutadania i les comunitats impactades.

Aquest document és una de les accions previstes per la [“Mesura de Govern per l'Estratègia Municipal d'Algoritmes i Dades per a l'Impuls Ètic de la Intel·ligència Artificial”](#) publicada el 2021 per l'Ajuntament de Barcelona, i té com a finalitat crear un protocol i una metodologia de treball comuna a l'Ajuntament de Barcelona per a la contractació pública o l'admissió de proves pilot³ de sistemes algorítmics confiables. S'entenen com a confiables tots aquells sistemes que compleixin els principis rectoris definits a la Mesura de Govern: Acció i supervisió humana; Robustesa tècnica i seguretat; Privacitat i governança de les dades; Transparència; Diversitat, inclusió i equitat; Compromís social i ambiental i responsabilitat, rendició de comptes i control democràtic. Aquest protocol és aplicable a qualsevol sistema algorítmic impulsat⁴ per l'Ajuntament de Barcelona, independentment de si aquest ja ha estat desenvolupat pel proveïdor o si el proveïdor desenvolupa una iniciativa municipal des del principi.

La metodologia recollida en aquest document combina (1) el cicle de vida d'un servei TIC a l'Ajuntament de Barcelona, (2) el cicle de vida d'un sistema algorítmic —i les seves particularitats—, (3) els mecanismes existents a l'Ajuntament pel que fa a la protecció de dades i (4) els estàndards ètics definits per diverses administracions i organitzacions. El document fa accionables els principis, recomanacions i estàndards definits a la “Mesura de Govern per l'Estratègia Municipal d'Algoritmes i Dades per a l'Impuls Ètic de la Intel·ligència Artificial”, el Reglament General de Protecció de Dades (RGPD) i la Llei Orgànica de Protecció de Dades Personals i Garantia dels Drets Digitals (LOPDGDD), la Proposta de Reglament del Parlament Europeu i del Consell pel qual s'estableixen normes harmonitzades en matèria d'intel·ligència artificial, la Llei 15/2022, integral per la igualtat de

¹ La intel·ligència artificial fa servir conjuntament la recollida i anàlisi de dades massives (*big data*) amb algoritmes i l'entrenament de sistemes intel·ligents (*machine learning*) per fer prediccions o recomanacions més o menys automatitzades. Per a més definicions, vegeu l'Annex 1.

² Els riscos associats amb els sistemes algorítmics, sobretot en casos en què aquests s'apliquen en serveis sensibles que afecten drets fonamentals, tenen a veure amb la manca de transparència i explicabilitat de les decisions preses; el fet que sovint es prioritza l'optimització sense tenir en compte a qui afecta; la dificultat per atribuir responsabilitats i garantir el rendiment de comptes; la vulneració de la privacitat dels individus i amb el fet que tant les bases de dades com el codi dels algoritmes poden incorporar biaixos socials que afectin de manera discriminatòria col·lectius minoritzats.

³ L'Ajuntament de Barcelona, a través de BIT Habitat, impulsa el testat de pilots de solucions urbanes innovadores en fase pre-comercial per tal de validar-ne la utilitat i l'interès públic en l'entorn urbà real abans que aquestes siguin comercialitzades. En el marc d'aquest programa, l'Ajuntament de Barcelona testeja sistemes d'IA aplicats a la millora de serveis urbans.

⁴ Per “impulsar” entenem aquells sistemes algorítmics de licitats per l'Ajuntament de Barcelona, siguin aquests solucions ja existents o solucions desenvolupades expressament, o aquells que siguin testats en el marc del programa “Barcelona Innovat Lab”, de BIT Habitat (sobre això, vegeu la pàgina 6).

tractament i no discriminació, els de documents elaborats per altres ciutats⁵, organitzacions i el marc regulatori sobre intel·ligència artificial que ha proposat la Comissió Europea (AI Act)⁶. Tot i que la llei europea sobre intel·ligència artificial encara es troba en procés de tramitació legislativa, aquest protocol incorpora diversos elements que hi són presents, com la classificació de les aplicacions d'IA segons el risc, per tal d'anticipar-se i adaptar-se de manera proactiva als estàndards que seran d'obligatori compliment quan la llei entri en vigor. Quan les institucions europees aprovin el reglament sobre IA, aquest protocol s'adaptarà al seu contingut final.

Pel que fa als sistemes algorítmics que incloguin processament de dades personals, aquest protocol té en compte els tres criteris personals marcats pel RGPD de cara a la intel·ligència artificial: Requeriment que el processament sigui just (art. 5), Principi de minimització de dades (art. 5), Avaluacions d'impacte de la protecció de dades (art. 35).

La regulació d'actuacions automatitzades municipals l'Ajuntament de Barcelona té com a precedent el decret de Comissió de Govern pel qual es regulen les actuacions administratives automatitzades municipals aprovat en sessió de 22 de juny de 2022 i publicat al BOPB de 23 d'agost de 2022.⁷

D'aquesta combinació en sorgeix un protocol que defineix, pas a pas, els mecanismes de garantia i salvaguarda de drets que s'han d'introduir en cada moment de la implementació d'un sistema d'intel·ligència artificial per part de l'Ajuntament de Barcelona, els quals inclouen, entre d'altres, l'elaboració d'estudis d'impacte algorítmic, la introducció de Clàusules Tipus per la Contractació de Sistemes d'IA Fiables, l'obligació de dur a terme auditories algorítmiques o la creació d'espais de participació i de mecanismes de comunicació ciutadana i transparència, com els registres d'algoritmes.

Aquest protocol es basa en la classificació de riscos feta per la Comissió Europea en el marc de la AI Act⁸, i estableix diferents tipus de mecanismes de garantia en funció del nivell de risc del sistema que es licita. Els sistemes algorítmics de risc inacceptable seran rebutjats automàticament, mentre que per a la resta de sistemes s'establiran mecanismes en funció del nivell de risc, com explica aquest document. Per als sistemes algorítmics d'alt risc –de manera sintètica, aquells que tenen un impacte directe en els drets fonamentals de les persones–, aquest protocol preveu un seguit de mecanismes de garantia més estrictes que per als considerats de risc limitat o mínim.

Atesa la novetat del camp de la IA al sector públic i la ràpida evolució de les tecnologies emergents i dels riscos que duen associats, aquest protocol es concep com un document viu que anirà evolucionant i incorporant el feedback dels actors interns i externs implicats, les lliçons apreses en el desplegament de projectes per part de l'Ajuntament de Barcelona, així com els avenços normatius i les recomanacions fetes per experts.

⁵ Vegeu, per exemple, les Clàusules tipus per a la contractació de sistemes algorítmics d'Amsterdam

⁶ <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>

⁷ Aquest decret es circumscriu a les actuacions administratives automatitzades dins el procediment administratiu i té com a finalitat identificar l'ús de la tramitació electrònica automatitzada en el procediment administratiu i establir el protocol per a la seva autorització. Tot això en compliment de la normativa administrativa d'aplicació i com a desenvolupament d'allò establert a l'Ordenança reguladora de l'administració electrònica.

⁸ Trobareu la classificació de riscos a la pàgina 7.

1.2. El cicle de vida d'un servei TIC a l'Ajuntament de Barcelona

Per garantir que els sistemes algorítmics contractats per l'Ajuntament de Barcelona compleixen amb els principis rectorials definits en la "Mesura de Govern per l'Estratègia Municipal d'Algoritmes i Dades per a l'Impuls Ètic de la Intel·ligència Artificial", cal alinear el cicle de vida d'un servei TIC i els procediments de contractació pública amb les particularitats del cicle de vida dels sistemes algorítmics i introduir, en cadascuna d'aquestes fases, totes les salvaguardes necessàries.

Tot producte o servei TIC té un cicle de vida: introducció, creixement, maduresa i declivi. En concret, les fases del cicle de vida d'un producte o servei a l'Ajuntament de Barcelona inclouen:

- 1. Concepció del servei:** Fase en què es descobreix una nova necessitat o sorgeix la idea del projecte, i que pot incloure l'elaboració d'un avantprojecte o altres estudis previs.
- 2. Contractació:** Redacció dels plecs per l'adquisició del servei (de desenvolupament o d'una altra naturalesa).
- 3. Desenvolupament:** Creació del codi, documentació i d'altres artefactes, incloent l'establiment de la infraestructura necessària per a la construcció dels mateixos.
- 4. Producció:** Desplegament del servei, incloent la possible migració de dades i processos des d'un o més serveis propis.
- 5. Explotació:** Fase que s'entén durant tota la vida útil del sistema en producció, incloent les tasques d'operació i manteniment.
- 6. Desmantellament:** Fase que inclou totes les accions necessàries a realitzar un cop acabada la vida útil del producte o servei TIC.

Aquest cicle de vida, a l'Ajuntament de Barcelona, té dues particularitats:

- La posada en marxa i el manteniment de serveis TIC corren a càrrec de l'Institut Municipal d'Informàtica (IMI) de l'Ajuntament de Barcelona, organisme autònom que té per objectiu subministrar la majoria dels serveis TIC a l'Ajuntament i als òrgans i empreses públiques que depenen d'aquest.
- L'Ajuntament de Barcelona, com a administració pública de l'àmbit local es regeix per un procediments administratius regulats per llei. Aquests, especialment els procediments de contractació pública, necessiten integrar-se en el cicle de vida dels serveis TIC.

1.3. El cicle de vida d'un sistema algorítmic

La creació d'un sistema algorítmic també té unes fases definides:

1. Formulació del problema: En aquesta fase, cal definir clarament quina necessitat es vol abordar amb el sistema algorítmic i establir què vol dir que el model funcioni bé. També cal definir els *inputs* o valors d'entrada que s'introduiran al sistema i els *outputs* o resultats que s'espera que s'obtingran⁹.
2. Recollida i processament de les dades.
3. Selecció i construcció del model: els models codifiquen un conjunt de suposicions sobre la relació potencial entre l'*input* i l'*output*, i pot entendre's com la fórmula matemàtica que combina els inputs amb un conjunt de paràmetres o coeficients de ponderació ajustables per obtenir l'*output* predit¹⁰. En el cas que el sistema incorpori l'aprenentatge automàtic, aquesta fase inclouria l'entrenament del sistema.
4. Desenvolupament i monitoratge: Atesa la naturalesa dels sistemes algorítmics, que difereix de la d'altres sistemes de *software* pel fet que el model, les dades o els sistemes d'aprenentatge poden canviar amb el temps, és important fer un seguiment detallat i rigorós dels sistemes un cop han estat posats en funcionament.

En alguns casos, certes fases d'aquest procés es repeteixen de manera iterativa: pot ser que en algun moment del desenvolupament del sistema es recullin noves dades i el model es reestructuri, o que aquest incorpori noves dades de manera contínua (machine-learning).

Tenint en compte aquestes fases i les particularitats dels sistemes algorítmics, en aquest document es proposen els següents processos i la seva relació amb cada fase per establir un protocol que haurà de seguir qualsevol àrea de l'Ajuntament de Barcelona que vulgui licitar un sistema algorítmic. A les figures 1 i 2, se sintetitza gràficament el procés que s'explica de manera detallada en les properes pàgines.

Cal posar èmfasi en el fet que els sistemes algorítmics que es licitin a l'Ajuntament de Barcelona, en tant que serveis TIC, potencialment seran licitats per l'IMI, que centralitzarà totes les fases que inclou el cicle de vida d'un sistema algorítmic, en col·laboració estreta amb l'àrea impulsora.

La majoria d'aquests processos només tenen sentit en situacions on el sistema algorítmic pugui tenir un impacte negatiu en la ciutadania o en comunitats específiques, seguint la classificació proposada per la Comissió Europea pel que fa al risc dels sistemes d'intel·ligència artificial¹¹. En cap cas es tracta d'un protocol d'aplicació general a qualsevol sistema algorítmic, sinó que es desplegarà de manera sistemàtica en les aplicacions dels sistemes algorítmics d'alt risc, i els òrgans de governança de la IA ètica de l'Ajuntament podran decidir si l'apliquen, també, en aplicacions de risc limitat i mínim. Aquest risc serà determinat per l'Oficina Tècnica d'Intel·ligència Artificial de l'IMI.

⁹ Els *outputs* poden ser resultats binaris (sí, no); una classificació en categories o una probabilitat (en una escala del 0 al 100, o en percentatges). Pot ser que aquests *outputs* serveixin simplement per donar suport a una decisió humana o que la decisió estigui plenament automatitzada, o que el sistema estigui programat per tenir tres possibles resultats: sí, no i "potser", el qual implicaria la revisió del resultat per part d'un humà.

¹⁰ En la construcció dels models dels sistemes que fan servir Machine Learning es fan servir dades i un algoritme d'aprenentatge per establir els paràmetres o coeficients del model, començant amb valors aleatoris que es van ajustant en funció dels outputs.

¹¹ <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>

Fig.1. Relació de les fases de desenvolupament d'un sistema algorítmic LICITAT per l'Ajuntament de Barcelona amb els mecanismes definits en aquest protocol

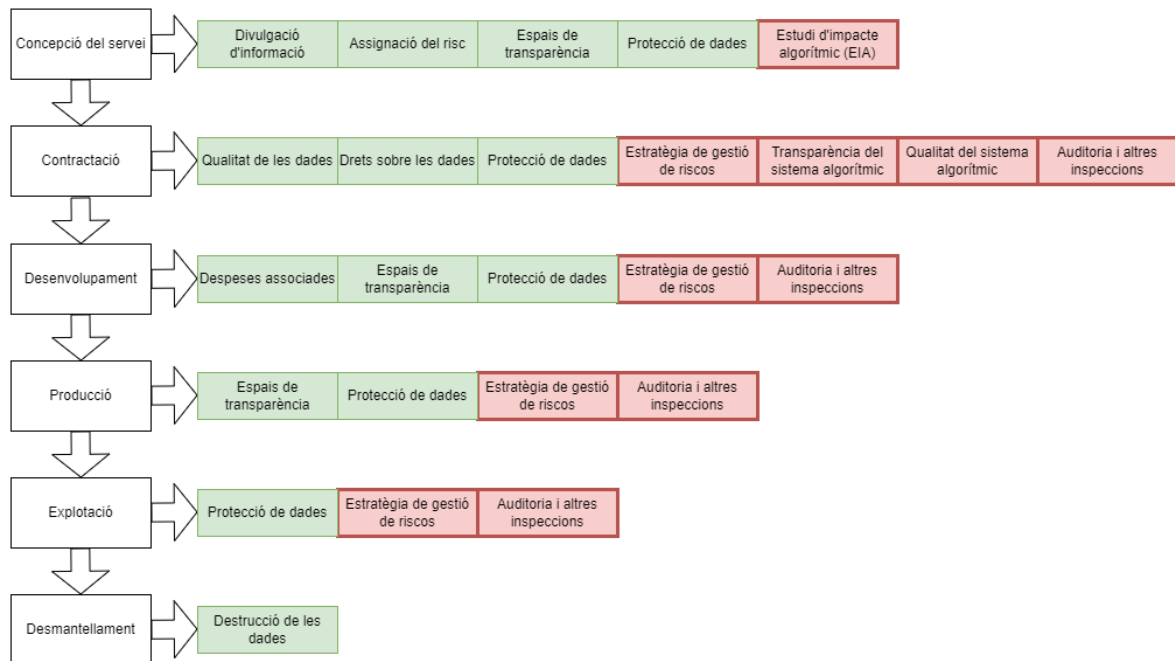
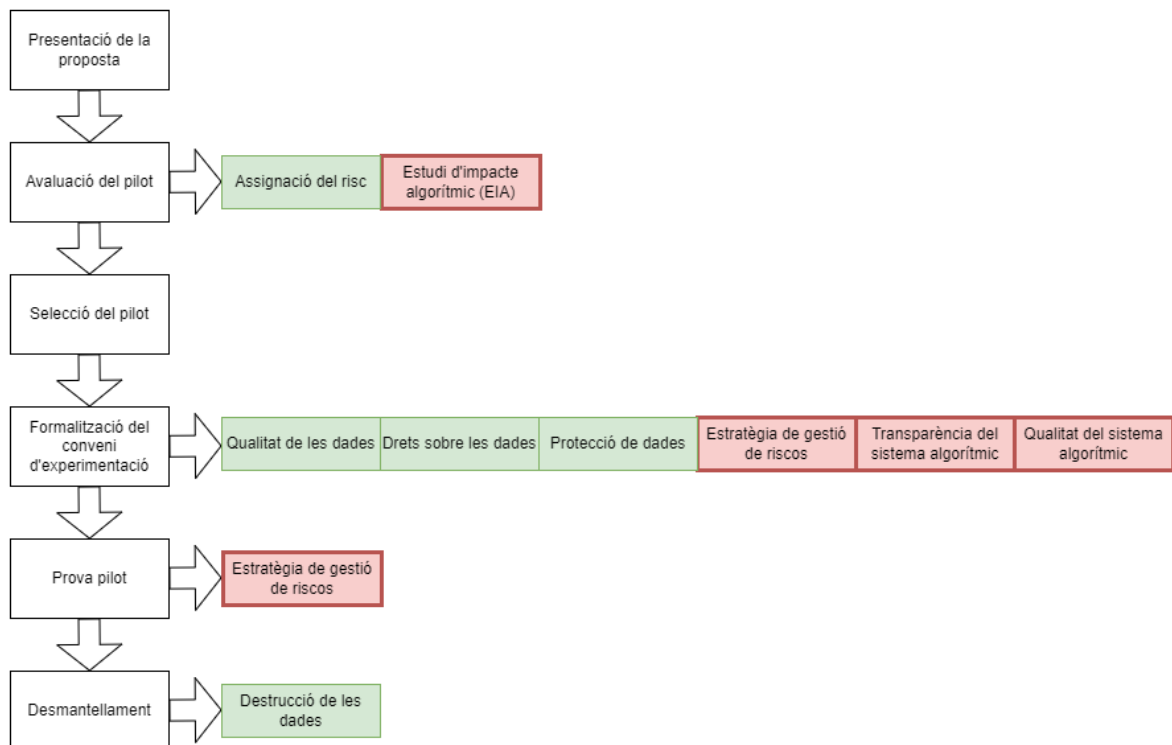


Fig. 2. Relació de les fases del desplegament d'un PROJECTE PILOT d'un sistema algorítmic impulsat per l'Ajuntament de Barcelona amb els mecanismes definits en aquest protocol



Testat de pilots d' intel·ligència artificial a l'Ajuntament de Barcelona

A banda dels casos en què l'Ajuntament de Barcelona licita un sistema algorítmic, una altra via d'entrada d'aquest tipus de sistemes a la ciutat és a través del programa del Barcelona Innova Lab impulsat per BIT Habitat i l'IMI. En el marc d'aquest programa, s'impulsa el testat de pilots de sistemes algorítmics aplicats a serveis urbans en fase pre-comercial per tal de validar-ne la utilitat i l'interès públic en l'entorn urbà real abans que aquests siguin comercialitzats. En aquest cas, l'Ajuntament de Barcelona no contracta el sistema algorítmic a través d'una licitació, sinó que facilita a l'empresa experimentadora les eines per testar el sistema (com per exemple, establint diàlegs amb les àrees municipals implicades per garantir l'accés a l'espai urbà, els actius municipals i la connectivitat 5G). En aquests casos, els mecanismes de salvaguarda i garantia de drets s'introdueixen mitjançant dos mecanismes:

- a) L'admissió dels pilots es duu a terme a través d'una convocatòria amb criteris estrictes d'elegibilitat. Aquests criteris, en el cas que el pilot involucri l'aplicació d'algoritmes de risc alt, poden incloure una valoració del Consell Assessor d'IA ètica.
- b) La introducció de les clàusules adequades al conveni d'experimentació que se signa amb les empreses experimentadores, seguint els criteris que s'apliquen a les Clàusules Tipus per la Contractació de Sistemes d'IA Fiables (vegeu apartat 3 d'aquest document).

Classificació de Riscos *proposada* per la Comissió Europea

La classificació de riscos de qualsevol sistema d'IA licitat o impulsat per l'Ajuntament de Barcelona serà la que determini el reglament europeu d'IA, en procés de tramitació. Serà l'Oficina Tècnica de l'IMI qui adjudicarà la categoria de risc a les sol·licituds de sistemes.

Aquests riscos fan referència al context d'aplicació del sistema i a la finalitat que se li dona. Els aspectes relacionats amb la protecció de dades personals i la ciberseguretat es regulen d'acord amb el previst a les lleis en vigor sobre el tema.

RISC INACCEPTABLE

Es prohibiran tots els sistemes d'IA que es considerin una clara amenaça per a la seguretat i els drets de les persones, des del "social scoring" dels governs fins a les joguines que fan servir assistència de veu i que fomenten comportaments perillosos. Es prohibiran també els sistemes d'identificació biomètrica remota en espais públicament accessibles.

RISC ALT

Aquests inclouen la tecnologia d'IA utilitzada en:

- Identificació biomètrica remota en temps real o en diferit de persones físiques;
- infraestructures crítiques (per exemple, el transport), que podrien posar en risc la vida i la salut de les persones
- la formació educativa o professional, que pot determinar l'accés a l'educació i el desenvolupament professional d'algú (per exemple, l'avaluació dels exàmens);
- els components de seguretat dels productes (per exemple, l'aplicació de la IA en la cirurgia assistida per robots);
- l'ocupació, la gestió dels treballadors i l'accés a l'autoocupació (per exemple, programari de classificació de CV per als procediments de contractació);
- serveis privats i públics essencials (per exemple, la qualificació de crèdits que nega als ciutadans l'oportunitat d'obtenir un préstec);
- aplicació de la llei que pot interferir amb els drets fonamentals de les persones (per exemple, avaluació de la fiabilitat de les proves);
- gestió de la migració, l'asil i el control de fronteres (per exemple, verificació de l'autenticitat dels documents de viatge);
- l'administració de justícia i els processos democràtics (per exemple, l'aplicació de la llei a un conjunt concret de fets).

Els sistemes de IA d'alt risc estaran subjectes a estrictes obligacions abans de poder comercialitzar-se o implementar-se:

- sistemes adequats d'avaluació i mitigació de riscos;
- alta qualitat dels conjunts de dades que alimenten el sistema per a minimitzar els riscos i els resultats discriminatoris;
- registre de l'activitat per a garantir la traçabilitat i transparència dels resultats;
- documentació detallada que proporcioni tota la informació necessària sobre el sistema i la seva finalitat perquè les autoritats puguin avaluar la seva conformitat;
- informació clara i adequada a l'usuari;
- mesures adequades de supervisió humana per a minimitzar el risc;
- alt nivell de solidesa, seguretat i precisió;
- Tots els sistemes d'identificació biomètrica a distància es consideren d'alt risc i estan subjectes a requisits estrictes. L'ús de la identificació biomètrica a distància en espais d'accés públic amb finalitats policials està, en principi, prohibit.

Es defineixen i regulen estrictament algunes excepcions, com quan és necessari buscar un nen desaparegut, prevenir una amenaça terrorista específica i imminent o detectar, localitzar, identificar o processar a un autor o sospitós d'un delictes greu.

Aquesta utilització està subjecta a l'autorització d'un òrgan judicial o un altre òrgan independent i a uns límits adequats quant a temps, abast geogràfic i bases de dades consultades.

RISC LIMITAT

El risc limitat es refereix als sistemes d'IA amb obligacions específiques de transparència. Quan s'utilitzen sistemes d'IA, com els chatbots, els usuaris han de ser conscients que estan interactuant amb una màquina per a poder prendre una decisió informada de continuar o retrocedir.

RISC MÍNIM O CAP RISC

La proposta permet l'ús lliure de la IA de risc mínim. Això inclou aplicacions com els videojocs amb IA o els filtres de spam. La gran majoria dels sistemes d'IA utilitzats actualment a la UE entren en aquesta categoria.

1.4. Governança

En el marc de la “Mesura de Govern per l’Estratègia Municipal d’Algoritmes i Dades per a l’Impuls Ètic de la Intel·ligència Artificial”, s’han creat òrgans de governança interns per garantir que els mecanismes definits en aquest protocol seran aplicats en la contractació o desenvolupament pilot de qualsevol sistema algorítmic per part de l’Ajuntament de Barcelona:

- **Comissió Transversal per a l’impuls de la IA ètica a l’Ajuntament de Barcelona.** En aquesta Comissió hi estan representades les àrees municipals i ens dependents relacionats amb la implementació de projectes d’intel·ligència artificial. Entre d’altres, té la funció de validar els sistemes basats en intel·ligència artificial d’aplicació a l’Ajuntament proposat per les àrees municipals i ens dependents, dirimir els conflictes que puguin sorgir en l’aplicació/ús d’algoritmes d’intel·ligència artificial en els processos interns i serveis públics prestats per l’Ajuntament i ens dependents i vetllar perquè la veu de la ciutadania sigui tinguda en compte en el procés de disseny i implantació de serveis basats en IA.
- **Oficina Tècnica d’IA de l’IMI.** Aquesta Oficina s’encarrega de definir, a nivell tècnic, els estàndards, eines, ús i implementació de tecnologies de IA sobre els projectes desenvolupats/implantats a l’Ajuntament de Barcelona; desenvolupar projectes pilot d’IA; donar suport per conceptualitzar les solucions a nivell tecnològic, metodològic i ètic i avaluar-ne els riscos.

Ahora, també en el marc del desplegament de la seva “Mesura de Govern per l’Estratègia Municipal d’Algoritmes i Dades per a l’Impuls Ètic de la Intel·ligència Artificial”, l’Ajuntament de Barcelona ha creat dos òrgans de participació per a que la ciutadania, el món acadèmic, les institucions i el món empresarial puguin participar en el procés de deliberació i exercir una supervisió externa sobre els sistemes algorítmics que es desenvolupen. (Vegeu l’Annex 4 per a la llista completa de membres de cada òrgan):

- **Consell assessor en Intel·ligència Artificial, Ètica i Drets Digitals.** Aquest òrgan estarà integrat per membres del món acadèmic, d’organitzacions de la societat civil i empreses tecnològiques i té, entre d’altres, la funció d’avaluar externament l’impacte dels sistemes algorítmics. Com es defineix més endavant, el Consell Assessor emetrà Estudis d’Impacte Algorítmic (EIA) preceptius i no vinculants per als algoritmes considerats d’alt risc.
- **Pacte de ciutat per a l’impuls de l’humanisme tecnològic.** Aquest Pacte, els actors signants del qual es reuneixen un cop l’any, funcionarà com a òrgan participatiu ampli d’entitats, empreses i institucions compromeses amb el desenvolupament ètic de la intel·ligència artificial i la protecció dels drets digitals.

Dins el Pacte es crearan grups de treball de temes específics, segons les necessitats, interessos i problemàtiques que vagin emergint pel que fa a la governança de les tecnologies digitals a la ciutat. En el marc d’aquest Pacte es

produiran diàlegs de caràcter més general (i no tan tècnic, com al Consell), entre d'altres temàtiques, sobre la idoneïtat i impacte social dels sistemes d'IA que s'impulsen des de l'Ajuntament de Barcelona.

2. Concepció del servei

El moment que hi ha entre que es detecta la necessitat de desenvolupar un sistema algorítmic a l'Ajuntament de Barcelona i el desenvolupament d'aquest (ja sigui a través d'un projecte pilot o dels procediments de contractació estàndards), és clau per garantir que la seva implementació es fa de manera transparent, segura, legítima i proporcionada i que s'inclou les persones afectades i les persones externes expertes en el procés de deliberació en aquelles aplicacions que tenen un risc més elevat.

2.1. Divulgació d'informació

Obligatori: **RISC ALT** **RISC MITJÀ** **RISC MÍNIM O CAP RISC**

Com a part de l'estudi previ a l'adquisició d'un sistema algorítmic, l'àrea municipal impulsora d'aquest omplirà una fitxa de sol·licitud de servei (en endavant, LA SOL-LICITUD, vegeu Annex 2 per a la fitxa completa), en la qual es detallarà:

- La seva definició interna de sistema de presa de decisions automatitzades.
- La informació sobre la necessitat que té, perquè es considera que la necessitat podria ser resolta amb un sistema algorítmic, com aquest podria respondre a aquesta necessitat i si s'han considerat alternatives a aquest.
- Els detalls de cada sistema automatitzat de presa de decisions que es planteja licitar, incloent detalls sobre la seva finalitat, abast, dades que es faran servir (si n'hi ha), política d'utilització i calendari d'aplicació.
- La definició preliminar de possibles riscos d'inexactitud, biaixos, o danys a comunitats específiques, definirà mesures de mitigació de riscos.
- La definició dels mecanismes per permetre l'accés dels experts externs que revisaran sistema una vegada hagi estat desplegat.

2.2. Assignació del risc

Obligatori: **RISC ALT** **RISC MITJÀ** **RISC MÍNIM O CAP RISC**

L'Oficina Tècnica d'IA de l'IMI rebrà i avaluarà LA SOL-LICITUD, i en un termini de 2 setmanes emetrà un informe que inclourà una extensió de la informació proporcionada per l'àrea municipal i una valoració del risc. De manera concreta, es detallarà:

- La informació sobre el sistema de presa de decisions automatitzades, incloent detalls sobre la seva finalitat, abast, política d'utilització i calendari d'aplicació.
- L'estudi de possibles riscos d'inexactitud, biaixos, o danys a comunitats afectades i definició de mesures de mitigació adreçades a minimitzar aquests riscos, incloent diàlegs proactius que fomentin el compromís amb els membres d'aquestes comunitats.

- L'explicació dels mecanismes adients per permetre l'accés a experts externs que s'ofereixin a revisar el sistema una vegada hagi estat desplegat.
- Determinació del risc (inacceptable/alt/limitat/mínim o cap risc).

En funció del risc determinat per l'Oficina Tècnica, es determinaran els següents passos:

- Risc inacceptable: LA SOL·LICITUD de l'àrea municipal serà denegada
- Risc alt: L'Oficina Tècnica informarà la Comissió Transversal i farà una petició d'estudi d'impacte algorítmic preceptiu no vinculant remunerat (a partir d'ara, EIA) al Consell Assessor en IA, que en un màxim de 3 setmanes¹² emetrà aquest estudi.
- Risc limitat i mínim o cap risc: No caldrà cap EIA i l'àrea municipal implicada podrà tirar endavant la redacció dels plecs.

2.3. L'estudi d'impacte algorítmic (EIA)

Obligatori: **RISC ALT**

Per als sistemes d'alt risc, el Consell Assessor emetrà un informe d'impacte algorítmic (EIA). Aquest tipus d'estudis descriuen el que el sistema algorítmic fa i qui és part responsable en cas que sorgeixi algun problema¹³. Es duu a terme en el moment pre-licitació, abans que l'Ajuntament de Barcelona hagi compromès la utilització d'un sistema algorítmic, per tal de conèixer-ne els possibles resultats indesitjables i insegurs que afecten els drets fonamentals de les persones i negociar i introduir les mesures necessàries per corregir-los.

Els EIA són crucials per donar a totes les parts interessades l'oportunitat de discutir i avaluar l'adopció del sistema, assegurar-ne la transparència i garantir l'accés de la ciutadania a la informació que l'afecta, així com per avaluar la seva idoneïtat tenint en compte el cas d'ús on es vol aplicar, tant abans de la posada en producció com durant la mateixa. En els casos d'alt risc, els EIA corresponents i les mesures de mitigació que s'introdueixin serviran de bones pràctiques tant per a l'Ajuntament de Barcelona com per a d'altres administracions locals. Aquest Protocol s'anirà actualitzant i millorant amb aquestes experiències.

Partint de l'informe "Algorithmic Impact Assessments: A Practical Framework For Public Agency Accountability"¹⁴, un EIA està dissenyat per assolir quatre objectius:

1. Respectar el dret de la ciutadania a saber quins sistemes afecten les seves vides, enumerant i descrivint públicament com son i de quina manera aquests sistemes afecten significativament les persones i comunitats.
2. Augmentar l'experiència i la capacitat interna dels organismes públics per avaluar els sistemes que construeixen o adquireixen i així poder-se anticipar als aspectes que poden suscitar preocupació, com els impactes negatius o les violacions de garanties processals.

¹²S'estableix un temps d'entrega de l'informe de tres setmanes per tal d'evitar l'alentiment del procés de contractació.

¹³ Les avaluacions d'impacte no són res nou. S'han aplicat en àmbits científics i polítics tan amplis com la protecció del medi ambient, els drets humans, la protecció de dades i la privacitat. L'estudi d'impacte algorítmic es basa en aquests marcs tot combinant-los amb la recerca dels experts científics i polítics que s'ha anat desenvolupant al voltant de la responsabilitat algorítmica.

¹⁴ <https://ainowinstitute.org/aiareport2018.pdf>

3. Garantir el retiment de comptes dels sistemes algorítmics i facilitar que experts externs revisin, auditin i avaluin aquests sistemes utilitzant mètodes que els permetin identificar i detectar falles o problemes.
4. Garantir a la ciutadania l'oportunitat de respondre i, si és necessari, impugnar l'ús d'un determinat sistema o l'enfocament realitzat sobre el retiment de comptes d'un sistema algorítmic determinat.

De manera orientativa i no restrictiva, els EIA preceptius i no vinculants contindran la següent informació relacionada amb el sistema algorítmic que es vol licitar:

- Descripció, propòsit, abast, política i calendari d'utilització
- Descripció del context d'aplicació
- Necessitat i proporcionalitat del sistema
- Identificació de parts implicades
- Revisió ètica: valors i conflictes (*trade-offs*)
- Impacte en els drets fonamentals de les persones i comunitats afectades
- Supervisió humana
- Definició de possibles riscos i de mesures de mitigació
- Recomanacions

En el cas de l'Ajuntament de Barcelona, i en línia amb els debats que s'han produït durant els darrers anys en el marc de la creació d'una regulació de la intel·ligència artificial a la Comissió Europea, els estudis d'impacte seran elaborats per experts externs a l'Ajuntament, reunits al Consell Assessor en intel·ligència artificial, Ètica i Drets Digitals, a fi de garantir la màxima imparcialitat sobre els sistemes algorítmics. Aquests EIA seran consultables per la ciutadania i qualsevol actor interessat al Registre Públic d'Algoritmes —juntament amb la resta d'informació pública sobre el sistema, com es definirà més endavant.

Cal destacar que els EIA no són la solució pels problemes que plantegen alguns sistemes algorítmics, i cal que aquests vagin acompanyats dels instruments necessaris per garantir-ne una implementació legítima, segura i transparent, com defineixen els següents capítols d'aquest document. Tot i així, són eines pràctiques per introduir el debat sobre la utilització d'aquests sistemes, proporcionar informació a la ciutadania i per determinar-ne la idoneïtat. Un cop licitat, l'impacte del sistema algorítmic es determinarà i revisarà a través de les auditories algorítmiques dutes a terme per una entitat externa a l'Ajuntament de Barcelona. El procés de contractació de les auditories el liderarà l'IMI.

2.4. Espais de transparència

Obligatori: **RISC ALT** **RISC MITJÀ** **RISC MÍNIM O CAP RISC**

Els sistemes algorítmics plantegen molts riscos i reptes: poden incloure els biaixos existents en les bases de dades (per falta de representativitat o per introducció de variables *proxy*, per exemple¹⁵); el procés pel qual arriben a una decisió pot ser opac o poc transparent i

¹⁵ Les variables *proxy* són variables aparentment no rellevants, però que actuen com una variable inobservable o inmesurable. Com assenyala l'[AI Primer de Nova York](#), n'és un exemple el codi postal, que als EUA sovint es relaciona de manera directa amb la racialització o el nivell d'ingressos i per tant, fer-lo servir com a variable pot conduir a decisions esbiaixades. Aquestes variables s'han de tractar de manera molt curosa per evitar comportaments inaccurats, inadequats i fins i tot perjudicials.

explicable, sobretot en els sistemes que incorporen l'aprenentatge automàtic o *machine-learning* —perquè no es poden explicar els pesos als coeficients que ponderen el resultat final a partir dels inputs donats—; sovint es prioritza l'optimització sense tenir en compte a qui afecta; entre d'altres. Tot això fa que l'Ajuntament de Barcelona es trobi davant la necessitat de vetllar per la legitimitat democràtica, respecte dels drets fonamentals, robustesa tècnica i seguretat, transparència i explicabilitat de qualsevol sistema algorítmic que es desenvolupi a nivell municipal.

De manera concreta, la ciutadania en general i les comunitats afectades disposaran de mecanismes de participació abans, durant i després de la compra i desplegament del sistema algorítmic. En el moment pre-licitació, s'establiran els següents mecanismes de participació i transparència:

- **Publicació de la informació pública** del servei que es vol licitar (en qualsevol tipus de sistema algorítmic, independentment del seu risc) al Registre d'Algoritmes
- **Revisió del sistema d'IA per part del Consell Assessor i elaboració de l'estudi d'impacte algorítmic pre-licitació.**
- Com preveu la instrucció en protecció de dades, hi haurà una fase d'informació al públic en el cas de supòsits presents a la llista d'alt Risc publicada per les Autoritats de control de protecció de dades.

2.5. Protecció de dades

Obligatori: **RISC ALT** **RISC MITJÀ** **RISC MÍNIM O CAP RISC**

En matèria de Protecció de dades, durant la fase de concepció es revisarà la inclusió de dades personals i la potencialitat d'obtenir decisions automatitzades, així com la naturalesa i adequació normativa d'aquestes, segons el que preveuen el RGPD i la normativa espanyola sobre les restriccions de les decisions automatitzades que puguin tenir efectes jurídics o equivalents per a les persones¹⁶.

Si el sistema inclou dades personals o les pot inferir i passa a la fase de contractació i explotació, s'haurà de registrar com a nou tractament al Registre d'Activitats de Tractament (RAT) municipal, d'acord al procediment previst a la Instrucció de protecció de dades.

3. Contractació

La classificació de riscos i els EIA tindran el següent efecte en la redacció dels plecs:

- *Risc inacceptable*: No es redactaran els plecs.
- *Risc alt*: L'àrea sol·licitant valorarà la idoneïtat de continuar amb el procés de licitació i, en cas afirmatiu, i valorarà la incorporació de les recomanacions emeses a l'EIA als plecs.
- *Risc limitat i mínim o cap risc*: No serà necessari un EIA i l'àrea municipal podrà tirar endavant la redacció dels plecs.

¹⁶ A tall d'exemple, en la mesura que tractem dades de caràcter personal o indirectament es pugui arribar a identificar persones, s'haurà d'identificar una base legal d'acord a l'article 6 de l'RGPD, tenint en compte que l'interès legítim no és aplicable a les Administracions Públiques. En el cas de decisions automatitzades que puguin afectar les persones, es complirà amb l'obligació normativa de transparència definida al RGPD.

En aquesta etapa és fonamental introduir els mecanismes necessaris per garantir que el sistema algorítmic contractat compleix amb els estàndards ètics d'intel·ligència artificial definits per l'Ajuntament de Barcelona. És per això que s'inclouran, als plecs de licitació de qualsevol sistema algorítmic contractat per l'IMI, les Clàusules Tipus per la Contractació de Sistemes d'IA Fiables¹⁷, en redacció per part dels serveis de contractació de l'Ajuntament de Barcelona.

Aquestes clàusules obliguen el contractista a desenvolupar el sistema algorítmic en qüestió sota les condicions que s'expliquen a continuació. Cal destacar que algunes de les clàusules s'apliquen només als sistemes algorítmics qualificats com a "risc alt". A l'inici de cada apartat, s'indica en quins casos s'aplicarà la condició: "independent del risc" vol dir que la condició s'aplicarà en qualsevol sistema algorítmic que es desenvolupi, i "alt risc" vol dir que la condició només s'aplicarà als sistemes d'alt risc.

3.1. Qualitat de les dades

Obligatori: **RISC ALT** **RISC MITJÀ** **RISC MÍNIM O CAP RISC**

La qualitat de les dades utilitzades en el desenvolupament d'un sistema algorítmic pot afectar de forma notòria la qualitat del sistema algorítmic en sí. Quan es recopilen dades, aquestes poden incorporar imprecisions i errors construïts socialment¹⁸. Les mesures que ha d'adoptar el contractista han d'estar focalitzades en evitar-ho. Per aquest motiu, cal establir uns requisits de qualitat de les dades.

Per altra banda, cal diferenciar entre els casos en què el contractista desenvolupa un sistema algorítmic amb les dades proporcionades per l'Ajuntament de Barcelona del cas en què és el contractista qui desenvolupa el sistema algorítmic a partir de les dades obtingudes a través d'altres fonts. El motiu és que, si el contractista rep les dades de l'Ajuntament de Barcelona, se'l pot considerar responsable de l'anàlisi, l'ordenació i l'edició de les dades, però no de la seva recollida. Al cap i a la fi, en aquest cas, és l'Ajuntament de Barcelona qui ho ha fet i, per tant, és lògic que la responsabilitat recaigui en el nom de l'Ajuntament de Barcelona.

Per assegurar la qualitat de les dades utilitzades en el desenvolupament d'un sistema algorítmic, l'Ajuntament de Barcelona exigirà la utilització d'un mètode fonamentat en la recollida (quan sigui el cas), l'anàlisi, l'organització i l'edició de les dades. Aquest mètode ha d'estar basat, quan sigui possible, en pautes científiques i altres normes existents en la pràctica real.

Un altre objectiu de l'exigència que la recollida (quan sigui el cas), l'anàlisi, l'organització i l'edició de les dades es duguin a terme seguint un mètode fonamentat és evitar, fins on sigui possible, biaixos en les dades. Què es considera biaix s'haurà d'avaluar en cada cas en concret. En qualsevol cas, es considera que existeix biaix quan la utilització de les dades

¹⁷ El document que conté les Clàusules es troba en construcció, i està basat en la proposta de [Clàusules estàndards de contractació de sistemes d'IA fiables](#) elaborades per l'Ajuntament d'Amsterdam.

¹⁸ Cal recordar que la recollida i classificació de les dades que entrenen i alimenten els sistemes algorítmics es duu a terme per humans, que poden tenir prejudicis -racistes o masclistes, per exemple- o visions esbiaixades que els poden conduir a crear etiquetes o categories per a l'ordenació de les dades que incloguin biaixos. A banda de la introducció de la subjectivitat humana, les bases de dades també poden conduir a imprecisions degut a la manca d'informació sobre col·lectius minoritzats que no queden representats de manera significativa. Un sistema algorítmic entrenat només amb dades d'un col·lectiu majoritari molt probablement oferirà resultats esbiaixats quan s'apliqui a d'altres col·lectius amb característiques diferents.

conduïx o pot conduir a una forma de discriminació protegida per les lleis i normatives aplicables, incloent l'article 14 del Conveni Europeu de Drets Humans¹⁹.

Els conjunts de dades utilitzats (tant per a entrenar els algoritmes com per a fer-los funcionar) poden presentar biaixos socials inadvertits, ja sigui per falta de representativitat de col·lectius, per introducció de variables "proxy", per llacunes o models de gestió incorrectes. Tot biaix demostrable i discriminatori ha de ser identificat i, sempre que sigui possible, el contractista ha d'eliminar-lo el més aviat possible. Si no és possible eliminar-lo, caldrà que adverteixi l'Ajuntament de Barcelona el més aviat possible i de forma clara. Per tal d'assegurar que el contractista compleix amb la seva responsabilitat de mantenir la qualitat de les dades, es durà a terme una auditoria algorítmica anual

3.2. Drets sobre les dades

Obligatori: **RISC ALT** **RISC MITJÀ** **RISC MÍNIM O CAP RISC**

El contractista serà l'encarregat del tractament, no el responsable de les dades, ja que qui definirà la finalitat del tractament i els mitjans a utilitzar és l'Ajuntament. Això serà aplicable tant en les situacions en què l'Ajuntament de Barcelona proporciona dades al contractista com en aquelles en les quals les dades es recullen en el context del contracte. En ambdues situacions, cal partir del principi base que els drets sobre les dades corresponen exclusivament a l'Ajuntament de Barcelona, que n'és el responsable —i no al contractista.

Els contractes de sistemes algorítmics inclouran acords clars respecte dels drets i obligacions sobre les dades, posant especial èmfasi en la sensibilitat per la matèria, en que l'adjudicatari resultant compleixi amb la normativa de protecció de dades²⁰. L'Ajuntament de Barcelona establirà acords contractuals específics amb cada proveïdor sobre els drets o les obligacions relatius a les dades que s'intercanvien.

Com defineixen les Clàusules Tipus per la Contractació de Sistemes d'IA Fiables creades per Amsterdam, de manera general s'acordarà que el contractista no té dret a utilitzar les ni les dades proporcionades per l'Ajuntament de Barcelona ni les dades generades durant el servei per a cap altre fi que no sigui l'execució del contracte. A petició de l'Ajuntament de Barcelona, i al final del contracte, el contractista destruirà les dades o les entregarà a l'Ajuntament de Barcelona, com defineixen les obligacions previstes per a encarregats de tractament de les clàusules dissenyades per a utilitzar a l'Ajuntament de Barcelona.

3.3. Qualitat del sistema algorítmic

Obligatori: **RISC ALT**
En la mesura del possible: **RISC MITJÀ** **RISC MÍNIM O CAP RISC**

¹⁹ https://www.echr.coe.int/Documents/Convention_CAT.pdf

²⁰ Cal recordar que no és possible posseir les dades en el sentit del dret de propietat, ja que el dret de propietat es limita als elements materials que poden ser objecte del control humà. Les dades tampoc solen estar subjectes a drets de propietat intel·lectual, ja que el desenvolupament de les dades no sol implicar decisions creatives i, per tant, les dades no solen estar subjectes a drets d'autor. En determinades circumstàncies, una base de dades que contingui dades pot estar subjecta a drets sobre les bases de dades. No obstant, el dret sobre les bases de dades només ofereix protecció davant la reutilització de tota la base de dades o una part substancial d'aquesta, no sobre la reutilització de dades específiques. Per a més informació sobre l'adequació del RGPD al tractament de dades que incorpora intel·ligència artificial, vegeu: <https://www.aepd.es/sites/default/files/2020-02/adequacion-rgpd-ia.pdf>

Quan en la fase pre-licitació s'hagi establert que un sistema algorítmic és d'alt risc caldrà afegir tots els mecanismes necessaris a les Clàusules Tipus per la Contractació de Sistemes d'IA Fiables per assegurar que aquest es desenvolupa de manera segura, transparent, explicable, legítima, respectant els drets fonamentals i que ho fa d'acord amb el marc legal aplicable. Per tant, serà necessari avaluar quina regulació és aplicable a cada cas en concret (p.e.: ús de tecnologies d'anonimització i GDPR) i establir l'obligació que el sistema algorítmic es desenvolupi seguint un mètode fonamentat, basat, quan sigui possible, el pautes científiques i altres normes existents en la pràctica real.

L'Ajuntament de Barcelona estableix un requisit de qualitat específic pel sistema algorítmic: el sistema algorítmic no pot presentar cap defecte i ha de proporcionar el resultat pel qual va ser concebut. Això significa, entre altres coses, que el sistema algorítmic ha de ser apte per a la seva utilització prevista, i ha de tenir la capacitat de realitzar decisions correctes (e.g., classificar correctament la informació en les categories adequades) o respecte a la seva capacitat de realitzar adequadament prediccions, formular recomanacions o prendre decisions correctes basades en dades o models. Això també implica que el sistema algorítmic ha de ser segur i fiable. Cal que el sistema algorítmic sigui robust des del punt de vista tècnic i el contractista ha de garantir que no es pugui destinar a usos malintencionats.

3.4. Transparència del sistema algorítmic

Obligatori: **RISC ALT**

En la mesura del possible: **RISC MITJÀ** **RISC MÍNIM O CAP RISC**

Com assenyalava un informe d'Algorithm Watch²¹, la transparència no assegura, per si mateixa, conformitat plena amb els requeriments ètics, però és una condició necessària per assolir aquesta conformitat. L'Ajuntament de Barcelona entén la transparència com una eina necessària per garantir que el sistema algorítmic es desenvolupa d'acord amb els estàndards ètics i de seguretat tècnica i que la ciutadania té accés a la informació sobre els sistemes que l'afecten.

L'Ajuntament de Barcelona diferencia entre transparència del procediment, transparència tècnica i explicabilitat, segons els diversos tipus d'informació que hi pot haver sobre un sistema algorítmic.

- **Transparència del procediment:** Es tracta del subministrament d'informació pública relativa al sistema algorítmic o al procés seguit en el desenvolupament i l'aplicació del mateix i de les dades utilitzades en aquest context. En qualsevol cas, la informació subministrada ha de ser suficient per garantir que, si sorgeixen danys, se'n pot rastrejar la causa, així com per garantir la comprensió de les opcions i dels supòsits assumits i les categories de dades utilitzades en el desenvolupament del sistema algorítmic. Ha d'incloure també la forma en què es dona la intervenció humana en el sistema algorítmic, la metodologia utilitzada per identificar els riscos, els riscos identificats i les mesures adoptades per mitigar-los, així com les parts que han participat en el desenvolupament del sistema algorítmic i les seves funcions.
- **Transparència tècnica:** Es tracta del subministrament d'informació per part del proveïdor a l'Ajuntament de Barcelona —respectant la propietat intel·lectual— que

²¹ Algorithm Watch (Juny 2021) "Automated Decision-Making Systems in the Public Sector: An Impact Assessment Tool for Public Authorities", Michele Loi, Anna Mätzener, Angela Müller, Matthias Spielkamp. Disponible online.

permeti comprendre el funcionament tècnic del sistema algorítmic. Aquesta informació ha d'incloure la divulgació del codi font del sistema, les especificacions tècniques i les dades utilitzades en el desenvolupament del sistema, la informació tècnica sobre com s'han obtingut les dades, el procés de desenvolupament utilitzat, la justificació de l'elecció del model concret i els seus paràmetres, així com la informació sobre el rendiment del sistema.

- **Explicabilitat:** Informació sobre perquè un sistema algorítmic porta a una decisió o resultat concret. Això inclourà una indicació clara dels factors clau que han portat a un sistema algorítmic a un resultat concret i els canvis que han de fer-se en les dades d'entrada de cara a obtenir un resultat diferent. Fer que un sistema algorítmic sigui explicable inclou proporcionar tota la informació tècnica i d'altra tipologia necessària per explicar als espais de participació o en procediments com els estudis d'impacte, les auditories algorítmiques o en procediments d'impugnació, de recurs o en altres procediments judicials, com s'ha arribat a una decisió i oferir a l'altra part o qualsevol part interessada l'accés a la informació i una protecció jurídica satisfactòria.

La transparència en el procediment i la transparència tècnica se centren en l'obtenció de la informació que l'Ajuntament de Barcelona considera necessària per avaluar el sistema algorítmic. A més, la transparència en el procediment també se centra en proporcionar a la ciutadania i altres parts interessades informació general sobre la utilització del sistema algorítmic.

Pel que fa a l'explicabilitat, tot i que de manera general s'obligarà al contractista a proveir la informació detallada sobre els factors clau que han portat un sistema algorítmic a un resultat particular o dels canvis que s'han de fer en les dades d'entrada per obtenir resultats diferents, podria donar-se el cas que en una situació específica això no fos tècnicament factible o necessari. En aquests casos, s'opta deliberadament per oferir a les parts la possibilitat d'establir una excepció en el contracte que es formalitzi.

L'Ajuntament de Barcelona serà lliure de divulgar la informació facilitada pel proveïdor relacionada amb la transparència del procediment. En canvi, la informació relacionada amb la transparència tècnica que facilitaran els proveïdors no serà de caràcter públic, i tindrà per objectiu que l'Ajuntament de Barcelona, en les auditories anuals corresponents (encarregades a una entitat tercera), pugui verificar si el contractista ha complert amb les seves obligacions. Aquesta pot incloure informació comercial confidencial i és per això que l'Ajuntament de Barcelona assumirà l'obligació de mantenir-la en secret i destruir-la un cop l'auditoria hagi finalitzat. En el cas que el contractista no vulgui proporcionar la informació relacionada amb la transparència tècnica a l'Ajuntament de Barcelona, s'estipularà al contracte que el contractista haurà de proporcionar-la directament a l'entitat tercera que dugui a terme l'auditoria.

3.5. Protecció de dades

Obligatori: **RISC ALT** **RISC MITJÀ** **RISC MÍNIM O CAP RISC**

En el disseny dels projectes d'IA, així com dels algorismes implicats, s'hauran d'aplicar els criteris de privacitat des del disseny i per defecte que marca el RGPD (art. 25), incloent-hi

aquelles funcionalitats que serveixin per salvaguardar els drets ARSOLP de la ciutadania, permetent la recuperació, modificació o cancel·lació de qualsevol dada personal, així com la oposició o limitació d'alguns dels seus processos.

Quan una àrea municipal tiri endavant la contractació d'un sistema algorítmic, en el cas que es tractin directament o indirecta dades de caràcter personal, per procedir al registre del tractament a la RAT de la Seu Electrònica, s'haurà d'omplir també el document DANTE (Document d'Avaluació de Nous Tractaments)

4. Desenvolupament, producció i explotació

Un cop licitat el servei, el proveïdor passarà a desenvolupar i posar en funcionament el sistema algorítmic. En aquestes fases del cicle de vida (desenvolupament, producció i explotació) caldrà garantir que el sistema funciona segons el seu objectiu inicial i que els riscos que emergeixen són mitigats. En el cas que el sistema incorpori l'aprenentatge automàtic, caldrà vetllar perquè les operacions dutes a terme per aquest continuïn sent explicables i el sistema no incorpori biaixos.

Les bones pràctiques i aprenentatges que emergeixin durant aquesta fase alimentaran aquest protocol i quedaran reflectides en aquest document viu per tal de millorar el procés de cara als projectes posteriors.

4.1. Estratègia de gestió de riscos en el desenvolupament del sistema algorítmic

Obligatori: **RISC ALT**

És important que qualsevol risc que pugui donar-se durant la utilització d'un sistema algorítmic s'identifiqui durant el desenvolupament del mateix i que es proposi una estratègia de mitigació de riscos adequada.

Les licitacions dels sistemes algorítmics de l'Ajuntament de Barcelona inclouran el disseny i l'execució d'una estratègia de gestió de riscos sistemàtica, que haurà d'anar sent actualitzada. Es proposa seguir la llista de comprovació proposada al document "Directrius Ètiques per una IA confiable" del grup d'experts d'alt nivell sobre IA facilitat per la Comissió Europea²².

L'estratègia de gestió de riscos que se segueixi ha d'estar en consonància amb la naturalesa del sistema algorítmic. Els sistemes algorítmics de risc alt seran sotmesos a una estratègia de gestió de riscos més intensa i estricta que la que s'aplicarà als sistemes algorítmics desplegat en casos de risc limitat o baix.

El resultat de l'execució d'una estratègia de gestió de riscos haurà de posar-se a disposició de l'Ajuntament de Barcelona mitjançant la sol·licitud de transparència del procediment. A més l'Ajuntament de Barcelona avaluarà, mitjançant auditories anuals, si s'ha dut a terme una estratègia de gestió de riscos correcta.

²² Independent High-Level Expert Group on Artificial Intelligence (2019). Ethics Guidelines For Trustworthy Ai, p. 26-31. Disponible en línia: <https://www.aepd.es/sites/default/files/2019-12/ai-ethics-guidelines.pdf>

En el cas en que s'utilitzi un sistema algorítmic ja desenvolupat pel contractista, el contractista ha de demostrar que ja s'ha dut a terme una estratègia de gestió de riscos.

4.2. Auditoria i altres tipologies d'inspecció

Obligatori: **RISC ALT**

L'Ajuntament de Barcelona ha de poder comprovar si el contractista compleix amb els requeriments especificats en el contracte i, amb aquesta finalitat, obligarà el contractista a col·laborar en aquestes comprovacions i a proporcionar totes les dades necessàries per a que una entitat externa dugui a terme una auditoria algorítmica. Cada sistema algorítmic que es liciti haurà de passar una auditoria anual durant tots els anys que estigui en funcionament.

Els resultats de l'auditoria es recolliran sempre en un informe. Si l'Ajuntament de Barcelona estableix que el contractista no compleix amb les obligacions contractuals, el contractista quedarà obligat a esmenar els incompliments identificats (i descrits a l'informe) en un termini raonable (també fixat per l'Ajuntament de Barcelona). Si el contractista no esmena els incompliments descrits per l'Ajuntament de Barcelona en el termini d'esmenes fixat, el contractista podrà ser penalitzat segons allò establert en el contracte.

L'Ajuntament de Barcelona tindrà el dret a difondre públicament les conclusions de l'informe al Registre d'Algoritmes.

4.3. Despeses

Obligatori: **RISC ALT** **RISC MITJÀ** **RISC MÍNIM O CAP RISC**

L'Ajuntament de Barcelona no haurà d'abonar un import addicional al contractista pels treballs que es derivin de les accions descrites anteriorment.

4.4. Espais de transparència

Obligatori: **RISC ALT** **RISC MITJÀ** **RISC MÍNIM O CAP RISC**

L'Ajuntament de Barcelona posarà a disposició de la ciutadania i de totes les parts interessades la informació pública relacionada amb els sistemes algorítmics que liciti al Registre Públic d'Algoritmes que es crearà a aquests efectes. De manera preliminar i no restrictiva²³, es preveu que es publiquin:

- Informació bàsica sobre el sistema algorítmic:
 - Finalitat del sistema algorítmic
 - Àrea municipal que l'impulsa
 - Actors implicats en les decisions del sistema
 - Abast del sistema algorítmic
 - Dades que es faran servir

²³ Atesa la novetat del camp de la IA al sector públic i el fet que encara no s'ha aprovat una regulació harmonitzada i comuna a nivell europeu, l'Ajuntament de Barcelona treballa de manera constant amb altres administracions, organitzacions internacionals, la Comissió Europea i experts independents per mantenir al dia i consolidar uns estàndards comuns i interoperables, entre d'altres temes, sobre el tipus d'informació que s'ha de posar a l'abast de la ciutadania per garantir una IA ètica, segura, robusta i confiable.

- Incorporació d'aprenentatge automàtic
- Política d'utilització
- Calendari d'aplicació
- Riscos que podrien aparèixer i estratègia de gestió d'aquests
- Legitimació de la finalitat en el cas de tractament de dades personals, prèviament alineat al RAT
- Els estudis d'impacte algorítmic duts a terme pel Consell Assessor (només dels sistemes que acabin sent licitats)
- El model de Clàusules Tipus per la Contractació de Sistemes d'IA Fiables
- Els informes d'auditoria algorítmica
- Informe de balanç final
- Informació clara i detallada dels mecanismes de comunicació ciutadana disponibles per tramitar peticions d'informació o consultes, incidències, queixes i suggeriments en relació al funcionament dels sistemes algorítmics

A banda del Registre d'Algoritmes, aquells sistemes que incorporin tractament de dades personals hauran d'incorporar-se al RAT de la Seu Electrònica.

A banda de la informació pública disponible al Registre d'Algoritmes, l'Ajuntament de Barcelona facilitarà la participació de la societat civil, empreses i entitats a través del Pacte de Ciutat per l'Humanisme Tecnològic, donant l'oportunitat a totes les parts implicades de debatre l'aplicació dels sistemes algorítmics i de conèixer-ne els resultats. A més, l'Ajuntament de Barcelona comptarà amb els experts del Consell Assessor en intel·ligència artificial, Ètica i Drets Digitals per fer un seguiment detallat dels sistemes d'IA que es trobin en funcionament. De manera concreta, s'afegirà un punt a l'ordre del dia de les reunions d'ambdós òrgans en què l'àrea municipal impulsora del sistema algorítmic i l'IMI hauran d'informar sobre el desenvolupament del sistema algorítmic i els membres dels òrgans podran fer preguntes sobre aquest i debatre'ls.

4.5. Protecció de dades

Obligatori: **RISC ALT** **RISC MITJÀ** **RISC MÍNIM O CAP RISC**

En el disseny dels projectes d'IA, així com dels algoritmes implicats, s'hauran d'aplicar els criteris de privacitat des del disseny i per defecte que marca el RGPD (art. 25), incloent-hi aquelles funcionalitats que serveixin per salvaguardar els drets ARSOLP de la ciutadania, permetent la recuperació, modificació o cancel·lació de qualsevol dada personal, així com la oposició o limitació d'alguns dels seus processos.

A la fase d'exploració, s'hauria de tenir molt clar el principi d'*accountability* exigint per l'RGPD, tant amb les autoritats de control com amb les persones afectades.

5. Desmantellament

5.1. Destrucció de les dades

Obligatori: **RISC ALT** **RISC MITJÀ** **RISC MÍNIM O CAP RISC**

Quan finalitzi el contracte amb el proveïdor, aquest estarà obligat a destruir, d'acord amb els requeriments del RGPD, totes les dades de caràcter personal relacionades amb el desenvolupament i explotació del sistema algorítmic. Per la resta de dades (no de caràcter personal), es fomentarà la reutilització d'aquestes, establint acords *ad hoc* amb cada proveïdor.

Pel que fa al codi font, en el cas que l'Ajuntament de Barcelona no n'hagi adquirit la propietat intel·lectual, el proveïdor podrà vendre'l a altres clients i seguir-lo fent servir, sempre que aquest no inclogui cap mena de dada recollida en el marc del contracte amb l'Ajuntament de Barcelona.

Un cop desmantellat el sistema algorítmic, l'àrea impulsora d'aquest, amb el suport de l'IMI, elaborarà un informe de balanç final, que inclourà una actualització de l'estudi d'impacte algorítmic, indicadors sobre el funcionament del sistema a nivell tècnic i social, qualsevol altra informació sobre la transparència tècnica, així com una explicació de com es reutilitzarà el coneixement adquirit en el servei i com s'integraran les bones pràctiques en futurs projectes.

ANNEX 1. Definicions

Algoritme, en un sentit molt ampli: des de models de regressió i arbres de decisió que poden fer prediccions i agilitzen processos, fins a sistemes més complexos, com ara les xarxes neuronals i els models bayesians, que funcionen amb aprenentatge automàtic a mesura que van fent càlculs i prediccions avançades.

Anàlisi de dades: Ha d'entendre's en un sentit ampli. L'anàlisi de dades pot incloure combinació, neteja, ordenació, classificació i derivació de dades.

Actuació Administrativa Automatitzada: D'acord amb la normativa administrativa aplicable, s'entén per actuació administrativa automatitzada, l'actuació realitzada íntegrament a través de mitjans electrònics, en el marc d'un procediment administratiu, en la qual no hi hagi intervingut de manera directa el personal al servei de l'Ajuntament i que utilitza algun dels sistemes de signatura establerts a la normativa administrativa d'aplicació, segell electrònic o codi segur de verificació.

Biaix algorítmic: Errors sistemàtics i repetibles en un sistema informàtic que creen resultats "injustos", com "privilegiar" una categoria sobre una altra de manera diferent de la funció prevista de l'algoritme.

Dades personals: tota informació sobre una persona física identificada o identificable («l'interessat»); es considerarà persona física identificable tota persona la identitat de la qual pugui determinar-se, directament o indirectament, en particular mitjançant un identificador, com per exemple un nom, un número d'identificació, dades de localització, un identificador en línia o un o diversos elements propis de la identitat física, fisiològica, genètica, psíquica, econòmica, cultural o social d'aquesta persona.

Destinatari: la persona física o jurídica, autoritat pública, servei o un altre organisme al qual es comuniquin dades personals, es tracti o no d'un tercer. No obstant això, no es consideraran destinataris les autoritats públiques que puguin rebre dades personals en el marc d'una recerca concreta de conformitat amb el dret de la Unió o dels Estats membres. El tractament de tals dades per aquestes autoritats públiques serà conforme amb les normes en matèria de protecció de dades aplicables als fins del tractament.

Elaboració de perfils: tota forma de tractament automatitzat de dades personals consistent a utilitzar dades personals per a avaluar determinats aspectes personals d'una persona física, en particular per a analitzar o predir aspectes relatius al rendiment professional, situació econòmica, salut, preferències personals, interessos, fiabilitat, comportament, ubicació o moviments d'aquesta persona física.

Encarregat del tractament o encarregat: la persona física o jurídica, autoritat pública, servei o un altre organisme que tracti dades personals per compte del responsable del tractament.

Explicable/explicable: Poder explicar a nivell individual per què un sistema algorítmic condueix a una decisió o resultat en particular. Tret que les Parts estiguin expressament

d'acord amb el contrari, això inclourà una indicació clara dels factors clau que han portat a un sistema algorítmic a un resultat particular i els canvis que s'han de fer per a que el sistema condueixi a un resultat diferent. Fer un sistema algorítmic explicable inclou la provisió de tota la informació tècnica per a explicar, en els procediments d'objecció, d'apel·lació o altres procediments judicials, com s'ha produït una decisió i oferir a les parts afectades o a qualsevol altra part interessada l'oportunitat d'avaluar la manera com s'ha adoptat una decisió, amb la finalitat única d'oferir una protecció jurídica realista.

Intel·ligència artificial: disciplina científica que inclou diversos enfocaments i tècniques com ara l'aprenentatge automàtic (del qual deriven l'aprenentatge profund i l'aprenentatge per reforç), el raonament automàtic (que inclou la planificació, la programació, la representació i el raonament de coneixements, la cerca i l'optimització), i la robòtica (que inclou el control, la percepció, els sensors i accionadors), així com la integració de totes les altres tècniques en sistemes "ciberfísics". Segons la definició del Consell d'Europa: "La IA és entesa com un conjunt de ciències, teories i tècniques, el propòsit del qual és reproduir mitjançant una màquina les capacitats cognitives d'un ésser humà".

Limitació del tractament: el marcat de les dades de caràcter personal conservats amb la finalitat de limitar el seu tractament en el futur.

Responsable del tractament o responsable: la persona física o jurídica, autoritat pública, servei o un altre organisme que, només o juntament amb uns altres, determini els fins i mitjans del tractament; si el dret de la Unió o dels Estats membres determina els fins i mitjans del tractament, el responsable del tractament o els criteris específics per al seu nomenament podrà establir-los el dret de la Unió o dels Estats membres.

Sistemes intel·ligents: algorismes de decisió automatitzada, capaços de prendre decisions sense intervenció humana. L'única exigència és que l'algoritme analitzi l'escenari donat a partir d'unes dades i prengui una decisió en funció d'aquesta anàlisi.

Tercer: persona física o jurídica, autoritat pública, servei o organisme diferent de l'interessat, del responsable del tractament, de l'encarregat del tractament i de les persones autoritzades per a tractar les dades personals sota l'autoritat directa del responsable o de l'encarregat.

Tractament: qualsevol operació o conjunt d'operacions realitzades sobre dades personals o conjunts de dades personals, ja sigui per procediments automatitzats o no, com la recollida, registre, organització, estructuració, conservació, adaptació o modificació, extracció, consulta, utilització, comunicació per transmissió, difusió o qualsevol altra forma d'habilitació d'accés, acarament o interconnexió, limitació, supressió o destrucció.

Transparència del procés: Informació sobre el propòsit del sistema algorítmic i el procés seguit en el desenvolupament i aplicació d'aquest i les dades utilitzades. S'ha d'incloure informació sobre les eleccions i suposicions assumides, els tipus de dades utilitzades en el desenvolupament del sistema, el mètode utilitzat per identificar els riscos, els riscos identificats i les mesures adoptades per mitigar els riscos, així com les parts implicades en el desenvolupament del sistema algorítmic i les seves funcions.

Transparència tècnica: Informació que permet que les parts afectades entenguin com funciona el sistema algorítmic de manera tècnica, que inclou la divulgació del codi font, les especificacions tècniques utilitzades, les dades utilitzades, la informació tècnica sobre l'obtenció i processament de les dades, la informació sobre el mètode de desenvolupament utilitzat i el procés de desenvolupament emprat, l'elecció d'un model particular i els seus paràmetres, i la informació sobre el rendiment del sistema algorítmic.



ANNEX 2. Model de SOL-LICITUD

INFORMACIÓ GENERAL

Àrea:

Definició de l'àrea _____ de "sistema de presa de decisions automatitzades":

Descripció de la necessitat i del sistema que es planteja implementar per solucionar-la:

Perquè el sistema algorítmic és la manera de respondre a aquesta necessitat?

S'han considerat alternatives a aquest? [SÍ/NO]

En el cas que s'hagi respòs SÍ, indiqueu quines alternatives:

DETALLS DEL SISTEMA QUE ES VOL LICITAR

Finalitat del sistema algorítmic:

Abast del sistema algorítmic:

Dades que es faran servir:

Política d'utilització:

Calendari d'aplicació:

Quins riscos d'inexactitud, biaixos o danys a comunitats específiques podrien aparèixer:

Quines mesures de mitigació de riscos s'empraran?:

ANNEX 3. Aprenentatges

- Cal pensar en models en els que participin els propis usuaris, com a propietaris de la futura eina, i en guies per propi desenvolupament del model amb auditories posteriors.
- Cal evitar que els mecanismes descrits en aquest protocol entorpeixin i facin més feixucs els processos.
- Caldrà afinar aquest protocol en funció de si el sistema que es licita ja ha estat desenvolupat o si aquest es desenvolupa específicament per a l'Ajuntament de Barcelona.
- Caldrà definir com es treballarà amb solucions i mètodes propietat de grans empreses (Google, OpenAI, etc).
- Valorar la introducció de nous criteris en les licitacions que, tal com passa amb la solvència tècnica i econòmica, introdueixin el concepte de “solvència ètica” dels proveïdors.