

Criteris d'ús de la intel·ligència artificial generativa en l'àmbit de l'Administració de la Generalitat de Catalunya i el seu sector públic

Aprovats per la sessió plenària de la Comissió d'Intel·ligència Artificial de 14 d'abril de 2025, a proposta del Comitè d'Ètica de les Dades

Índex

1. Antecedents.....	2
1.1. Objecte	2
1.2. Objectius	2
1.3. Abast	2
1.4. Marc de referència.....	2
1.5. Estructura del document.....	5
2. Què és la IA generativa?	5
2.1. Com s'emmarquen aquests criteris en l'Administració pública	6
2.2. Quines tasques ofereix la IA generativa.....	6
2.3. Què no pot fer la IA generativa	7
2.4. Riscos associats.....	7
2.5. Valors i cultura ètica	9
3. Criteris per a la utilització ètica i responsable de l'Administració de la Generalitat i el seu sector públic institucional, i dels empleats públics.....	10
3.1. Introducció	10
3.2. Directrius	11
3.2.1. Què no pot fer un empleat públic amb la IA generativa	11
3.2.2. Fase d'avaluació prèvia	11
3.2.3. Criteris d'ús de la informació inclosa en la indicació o <i>prompt</i>	15
3.2.4. Criteris d'ús del resultat	17
3.3. Qüestions pràctiques	20
3.3.1. Redacció de <i>prompts</i> o indicacions	20
3.3.2. Material de suport	20
3.3.3. Conclusió	20

1. Antecedents

La Comissió d'Intel·ligència Artificial, en la sessió plenària de l'11 de desembre de 2024, va demanar al Comitè d'Ètica de les Dades una proposta de criteris d'ús d'eines d'intel·ligència artificial (en endavant, IA) generativa adreçada a l'Administració de la Generalitat i el seu sector públic institucional, i als seus empleats públics. Aquest document n'és la concreció.

1.1. Objecte

L'objecte d'aquesta proposta és l'establiment de criteris per a l'ús ètic i responsable d'eines d'IA generativa en l'exercici de les funcions atribuïdes als ens que formen l'Administració de la Generalitat i als empleats públics de l'Administració de la Generalitat. Aquests criteris han d'estar plenament alineats amb el marc de principis ètics en l'ús de la IA corporatius.

1.2. Objectius

Els objectius d'aquesta proposta són:

- Contribuir a la cultura ètica i de les dades de l'Administració de la Generalitat.
- Contribuir a l'alfabetització de la IA de l'organització.
- Contribuir a generar confiança pública.
- Donar suport a l'organització per avaluar els riscos i les limitacions de l'ús de la IA generativa.
- Establir les condicions per al seu ús segur i fiable en l'exercici de les funcions de l'organització.

1.3. Abast

El Reglament europeu d'intel·ligència artificial (RIA) dona suport a l'objectiu de promoure l'enfocament europeu de la IA centrat en l'ésser humà i té la voluntat d'actuar com a referent mundial a establir un marc per al desenvolupament d'una IA segura, digna, de confiança i ètica.

En conseqüència, els principis i criteris ètics en l'ús de les dades i els usos massius són aplicables en tot el cicle de vida dels sistemes d'IA i responen a l'objectiu d'aplicar l'ètica des del disseny. Aquest esperit implica que l'aplicació pràctica d'aquests criteris abasten:

- Els sistemes d'IA generativa disponibles públicament en línia i disponibles a internet.
- El desenvolupament i la utilització d'aquelles eines internes d'ús d'IA generativa especialitzades i adaptades a àmbits específics de l'organització i fetes sota petició, sens perjudici de les directrius que específicament s'hi estableixin de conformitat amb el RIA.

1.4. Marc de referència

Aquest document s'elabora de conformitat amb:

- La [Instrucció 8/2020, de 24 de novembre](#), sobre l'ús de les tecnologies de la informació i la comunicació a l'Administració de la Generalitat de Catalunya.
- La [Recomanació 1/2023, de 21 d'abril](#), de l'Autoritat Catalana de Protecció de Dades en relació amb la utilització de ChatGPT i el seu impacte en la protecció de dades personals.
- El [Codi ètic del servei públic de Catalunya](#), aprovat per l'[Acord GOV/164/2021, de 26 d'octubre](#), emprat com a codi marc. De forma singular, s'han pres com a referència els valors inherents al servei públic i els vinculats a les dimensions relacionals dels servidors públics.
- L'[Agenda 2030 i els objectius de desenvolupament sostenible](#) (ODS).
- Les [Directrius ètiques per a una IA fiable de la Comissió Europea](#).
- Les [Recomanacions sobre l'ètica de la IA de la UNESCO](#).
- El [Programa de les Nacions Unides per al medi ambient](#) (PNUMA), que és la principal autoritat mundial en matèria de medi ambient.
- L'[Institut Nord-americà de Recerca en Intel·ligència Artificial Distribuïda](#).
- El [Diàleg de l'Organització Mundial de la Propietat Intel·lectual i Tecnologies d'Avantguarda. IA generativa](#).
- El [Manifest de la cultura ètica davant la IA. Pel respecte dels drets fonamentals dels i les professionals de la cultura davant la irrupció de la intel·ligència artificial generativa](#), de 2 d'octubre de 2024, promogut per diverses associacions de professionals de la indústria cultural.
- La [Directiva \(UE\) 2019/790 del Parlament Europeu i del Consell, de 17 d'abril de 2019](#), sobre els drets d'autor i drets afins en el mercat únic digital i per la qual es modifiquen les directives 96/9/CE i 2001/29/CE.
- El [Reglament \(UE\) 2024/1689 del Parlament Europeu i del Consell, de 13 de juny de 2024](#), pel qual s'estableixen normes harmonitzades en matèria d'intel·ligència artificial (RIA).
- El [Reial decret llei 24/2021, de 2 de novembre](#), de transposició de directives de la Unió Europea en les matèries de bons garantits, distribució transfronterera d'organismes d'inversió col·lectiva, dades obertes i reutilització de la informació del sector públic, exercici de drets d'autor i drets afins aplicables a determinades transmissions en línia i a les retransmissions de programes de ràdio i televisió, exempcions temporals a determinades importacions i subministraments, de persones consumidores i per a la promoció de vehicles de transport per carretera nets i energèticament eficients.
- La [Llei 39/2015, d'1 d'octubre](#), del procediment administratiu comú de les administracions públiques.

- La Llei 26/2010, de 3 d'agost, de règim jurídic i de procediment de les administracions públiques de Catalunya.
- La Llei 29/2010, de 3 d'agost, de l'ús dels mitjans electrònics al sector públic de Catalunya.
- El Reial decret 203/2021, de 30 de març, pel qual s'aprova el Reglament d'actuació i funcionament del sector públic per mitjans electrònics.
- Els principis generals de l'Administració digital, que preveu el marc del Decret 76/2020, de 4 d'agost, que són:
 - Proactivitat i personalització en la prestació de serveis públics digitals amb l'objectiu de situar l'experiència de les persones en el centre del disseny de serveis.
 - Eficàcia i eficiència en la gestió dels recursos públics necessaris tant per posar en pràctica les polítiques adients per a l'Administració digital com per generar resultats positius, així com per utilitzar els recursos públics de manera responsable i evitar tant com sigui possible traslladar les càrregues del present al futur.
 - Transparència per exposar, de manera clara i intel·ligible, les dades i informacions que siguin rellevants per conèixer, entendre, controlar i avaluar la seva acció.
 - Innovació i adaptació contínua dels serveis públics tenint en compte l'aprofitament del potencial de la tecnologia en cada moment.
 - Confiança en l'Administració digital mitjançant el desenvolupament d'un model de ciberseguretat, especialment en relació amb la protecció de dades de caràcter personal.
 - Legalitat, que suposa el manteniment de la integritat de les garanties jurídiques de la ciutadania davant les administracions públiques, d'acord amb la normativa que regula el règim jurídic i el procediment de les administracions públiques.
- Els principis i criteris ètics en l'ús de les dades i la IA proposats pel Comitè d'Ètica de les Dades i aprovats per la Comissió d'IA en data 11 de desembre de 2024, que són:
 - Principi de justícia, equitat i no-discriminació.
 - Acció i supervisió humana.
 - Confidencialitat, privacitat i seguretat.
 - Governança de les dades avançades i de sistemes d'IA.
 - Governança dinàmica de l'ètica.
 - Transparència i retiment de comptes.
 - Desenvolupament sostenible i protecció del medi ambient.

1.5. Estructura del document

Els criteris proposats s'emmarquen en la cultura ètica i de les dades de l'organització i de la innovació responsable per a la transformació digital.

Per facilitar la comprensió dels criteris i situar en context, aquesta proposta s'inicia amb una descripció didàctica de la IA generativa d'acord amb les definicions establertes i el context normatiu del RIA.

A continuació, es descriuen possibles escenaris que preveuen una relació de situacions sobre què es pot fer i què no amb la IA generativa. Aquest apartat pretén únicament situar els empleats públics en context, en tant que aquesta regulació substantiva, així com les guies i orientacions tècniques que se'n derivin, formen part de les polítiques corporatives en matèria d'IA i de governança de les dades.

Així mateix, cal tenir en compte que aquests criteris poden evolucionar en funció de l'estat de la tècnica, de la governança de les dades i de la IA, de la normativa del procediment administratiu, i l'àmbit regulatori de la responsabilitat de l'Administració i dels seus empleats públics que es puguï desplegar.

2. Què és la IA generativa?

És un tipus d'intel·ligència artificial destinada a crear models capaços de generar contingut nou i original com ara text, imatges, vídeo i àudio.

La IA generativa permet treballar grans volums de dades i fer-ho en contextos complexos, i aprèn els patrons i l'estructura de les dades a partir de dades d'entrenament per generar noves dades que tenen característiques similars. Els models de llenguatge es poden entrenar amb dades que provenen de diverses fonts, moltes de públiques i disponibles a internet, aspecte que repercuteix en la seva accessibilitat global i el seu impacte. Tot i que es poden trobar conjunts de dades estructurades, codis de programació o biblioteques digitals que ofereixen un grau alt de fiabilitat, es pot donar el cas que algunes d'aquestes fonts no siguin del tot representatives i siguin esbiaixades en la seva resposta.

També pot passar que les dades que utilitzin models de la IA generativa incloguin, sovint, contingut protegit per drets d'autor, per aprendre i generar nous continguts. Això pot suposar una infracció dels drets d'autor si s'utilitza material sense el permís dels titulars.

L'impacte de la IA generativa pot resultar significatiu en els drets fonamentals, en la protecció de les dades de caràcter personal, en els valors democràtics i en la propietat intel·lectual o industrial, i pot fomentar la desprotecció d'artistes, col·lectius de dissenyadors i il·lustradors, escriptors, traductors, acadèmics i altres professionals científics i treballadors culturals davant la IA generativa.

La IA generativa interactua amb l'usuari mitjançant els *prompts* o indicacions, que són una sol·licitud, una instrucció, una pregunta o un text perquè dugui a terme una tasca i creï una resposta. Un cop fet això, la tecnologia utilitza els patrons apresos per generar contingut nou.

El resultat obtingut, encara que respongui a les indicacions de l'empleat, sempre exigeix revisió i adaptació humana.

El marc de definició és l'establert en el context regulatori del RIA i cal situar-lo en l'àmbit de l'actuació administrativa.

2.1. Com s'emmarquen aquests criteris en l'Administració pública

Els criteris proposats s'han d'emmarcar en els principis de l'actuació administrativa, que estableix la normativa bàsica de procediment administratiu i de règim jurídic.

En aquest sentit, l'Administració serveix amb objectivitat els interessos generals i actua d'acord amb els principis d'eficàcia, jerarquia, descentralització, desconcentració i coordinació amb sotmetiment ple a la llei i al dret. Per això, tota actuació administrativa ha d'estar orientada i guiada pels principis esmentats.

Aquesta coherència en l'aplicació dels principis a tot el marc normatiu és la base de tots els models d'actuació que regulen el funcionament de les diferents administracions públiques. En especial, dels que estableix el Reial decret 203/2021, de 30 de març, pel qual s'aprova el Reglament d'actuació i funcionament del sector públic per mitjans electrònics, en la mesura en què vinculen l'actuació i el funcionament de l'Administració i el funcionament electrònic del sector públic.

L'aplicació dels principis d'actuació administrativa faciliten l'adaptabilitat que han de tenir les administracions públiques al progrés continuat de la tecnologia i als sistemes de comunicació electrònica, i fonamenta el desenvolupament dels principis i criteris ètics en l'ús de les dades i la IA. En essència, milloren l'eficiència administrativa, incrementen la transparència i la participació, garanteixen que els serveis digitals siguin fàcilment utilitzables, i milloren la seguretat jurídica.

Els principis i criteris ètics en l'ús de les dades i la IA que ha proposat el Comitè d'Ètica de les dades contribueixen a assegurar una implantació de la tecnologia justa i centrada en les persones a l'Administració de la Generalitat de Catalunya i guarden coherència amb els principis que regeixen el dret de l'Administració digital. Els principis i criteris ètics esdevenen una base ferma en un entorn regulatori complex i incipient com és el de la IA.

Cal destacar el principi d'accessibilitat, estretament vinculat al que es predica en aquests principis i criteris, en tant que és el conjunt de principis i tècniques que cal respectar a l'hora de dissenyar, construir, mantenir i actualitzar els serveis electrònics per garantir la igualtat i la no-discriminació en l'accés de les persones com a destinatàries finals dels serveis, aspecte també cabdal vinculat a la lluita contra la bretxa digital.

2.2. Quines tasques ofereix la IA generativa

D'altra banda, l'ús de la IA generativa a l'Administració pública ha de contribuir a millorar la interacció amb la ciutadania mitjançant assistents virtuals, així com a automatitzar processos, optimitzar la gestió de recursos i augmentar la productivitat dels empleats públics.

La IA generativa ofereix la possibilitat de dur a terme tasques automatitzades com ara:

- Generació de contingut: redacció i edició de textos, imatges, vídeo i àudio.
- Revisió i descobriment de literatura.
- Resums i anàlisi de documents i informació.
- Visualització de dades.
- Traducció de textos.
- Revisió i aplicació de criteris d'accessibilitat.
- Programació bàsica.
- Estructuració i anàlisi de dades.

La IA generativa també permet millorar l'eficiència i la precisió en les cerques més avançades i en l'anàlisi de grans volums de dades, i permet identificar tendències que poden passar desapercebudes amb una revisió manual.

2.3. Què no pot fer la IA generativa

Cal tenir present que la IA generativa no pot:

- Garantir la fiabilitat, l'expertesa ni el raonament que aporta una persona, per la qual cosa el seu resultat pot ser erroni, incorporar al·lucinacions o accentuar els biaixos que pot tenir una persona humana.
- Crear informació, en el format que sigui, amb informació de qualitat.
- Comprendre de la mateixa forma que una persona humana, i, per això, no pot prendre ni substituir la decisió, la diligència ni el capteniment de l'empleat públic, sinó únicament donar suport en aquesta presa de decisions.

2.4. Riscos associats

L'ús d'IA generativa planteja, en la societat i també a l'Administració, un debat ètic des de les perspectives següents:

- **La perspectiva social:** diverses qüestions de transcendència social tenen un impacte en el funcionament i el resultat dels sistemes, aspecte que genera debat sobre la sostenibilitat social.

D'entrada, l'origen de les dades es basa molts cops en l'explotació humana de persones del sud global que han estat els responsables de la introducció prèvia i el tractament de la informació. Aquesta situació planteja desafiaments ètics i morals sobre la sostenibilitat social d'aquests paradigmes basats en explotació de les persones. L'ètica aplicada a la IA no només ha de referir-se a accions vinculades a la transparència, el tractament de dades i la desinformació, sinó que cal ampliar la reflexió cap al sosteniment de les asimetries que presenten l'ús d'aquests sistemes. D'entrada, una possibilitat per combatre aquest aspecte pot resultar de l'aplicació d'estratègies

corporatives, com ara l'Estratègia catalana de millora de la compra pública ([eix 4, "Contractar amb integritat i transparència"](#)).

D'altra banda, la qualitat de les dades és clau perquè no es produeixi una informació esbiaixada, poc representativa o que discrimini o provoqui una desigualtat de tracte tant en l'entrada com en la sortida. Això inclou l'ús d'informació sensible de forma inadequada, que s'ha d'evitar.

Els sistemes d'IA generativa no poden substituir el raonament humà i, per tant, no poden distingir el que és correcte, ètic o just.

- **La perspectiva ambiental:** aquests sistemes necessiten un ús massiu d'energia per la utilització que fan d'infraestructures com ara centres de dades que consumeixen molta energia, aigua i recursos minerals escassos, extrets en alguns casos en condicions que posen en dubte el respecte a la sostenibilitat. Aquests centres generen emissions i residus electrònics que contribueixen a l'augment de la contaminació i al canvi climàtic. Tot això accelera una triple crisi ambiental: l'augment de residus, la pèrdua de biodiversitat i el canvi climàtic, que ja és una realitat. Per això, abans de fer-la servir, cal avaluar el cost mediambiental en comparació amb el benefici obtingut i fer-la servir només en els casos adequats. Cal fer-ne un ús responsable.
- **La perspectiva legal:** l'accés i el processament sense control d'informació i dades confidencials, inclosos els secrets comercials, professionals o empresarials, pot vulnerar drets fonamentals —en especial, de la protecció de dades de caràcter personal— o d'altres drets, com ara la protecció de drets de propietat intel·lectual de tercers. La compartició d'aquesta mena d'informació pot causar bretxes en la seguretat de les dades i dels sistemes d'informació i comprometre la integritat de les dades. A més, el RIA estableix una sèrie d'obligacions per facilitar el control dels proveïdors i responsables del desplegament de sistemes. Alhora, no es pot entendre aquest desplegament sense prendre en consideració la regulació de protecció de dades personals.¹ Ambdues regulacions són l'eix vertebrador del desplegament de la IA perquè precisament distingeix i singularitza l'enfocament europeu de la resta de regulacions internacionals.
- **La perspectiva de l'impacte en la creativitat humana:** la IA generativa està transformant les indústries creatives i innovadores, impulsant el progrés econòmic i generant debats sobre els seus efectes en la creativitat humana. Un tema clau és si els continguts generats per IA han de ser objecte de protecció de drets d'autor. Això inclou discussions sobre la intervenció humana necessària i la relació entre els models d'IA. Alhora, planteja desafiaments en la protecció dels drets d'autor.

¹ Essencialment, el [Reglament \(UE\) 2016/679 del Parlament i del Consell, de 27 d'abril de 2016](#) (RGPD), i la [Llei orgànica 3/2018, de 5 de desembre, de protecció de dades personals i garantia dels drets digitals](#) (LOPDGDD).

- **La perspectiva de la veracitat:** la utilització de la IA pot generar riscos d'informació errònia, desinformació i discurs d'odi, així com causar danys per l'ús indegut de les dades personals. A més, pot contribuir a l'aparició de noves formes de relat social, la discriminació, la llibertat d'expressió, la privacitat i l'alfabetització mediàtica i informacional, entre d'altres. Per això, abordar aquest risc i el potencial perill que suposa exigeix un increment proporcional de l'educació en matèria de dades i de l'alfabetització mediàtica i informacional, així com l'accés a fonts d'informació independents, pluralistes i fidedignes.
- **La perspectiva de l'Administració pública com a garant de les salvaguardes dels drets i valors públics:** existeix un debat en l'àmbit de les administracions públiques i l'acadèmia sobre la creixent participació del sector privat en el desenvolupament d'IA per a l'Administració pública, a causa de la falta de mitjans propis, la qual cosa genera desafiaments reguladors.

La IA en cap cas no hauria d'alterar la configuració del procediment administratiu ni reduir garanties de les persones.

Més enllà de la motivació de l'acte administratiu, es planteja la necessitat d'assegurar-ne l'explicabilitat per reforçar la transparència i la confiança i també perquè, com a òrgan decisor, ha de ser capaç de retre comptes. L'ús de la IA en l'actuació administrativa s'ha de tractar i valorar d'acord amb les potencialitats del sistema emprat i les seves característiques específiques, i no serà igual una actuació administrativa automatitzada sense IA que una que sí que l'aplica (per exemple, una IA predictiva).

La manca d'un corpus regulador substantiu en la matèria fa que sigui recurrent recórrer a la sistemàtica remissió als principis generals de l'activitat administrativa. Aquesta forma sistemàtica respon a l'alta complexitat tècnica de la regulació i a la ràpida evolució de la tecnologia que fa difícil establir una regulació amb caràcter de permanència quan en poc temps pot quedar obsoleta.

En conseqüència, un ús de les eines incontrolat, inconscient, irresponsable o que no respecti aspectes de la normativa de protecció de dades o altres normatives d'aplicació al cas davant d'aquestes eines podria causar un greu perjudici. Per tant, cal tenir en compte el seu impacte en els valors del respecte, la protecció i la promoció dels drets humans, la diversitat, la igualtat de tracte, la no-discriminació i la sostenibilitat del medi ambient.

2.5. Valors i cultura ètica

L'ètica és un element clau perquè la IA pugui desplegar-se i fer-se servir de forma fiable i respectuosa amb els drets fonamentals, així com de manera beneficiosa per a la ciutadania, per als empleats públics i, en definitiva, per a la societat en general.

Només un ús ètic, conscient i responsable de la IA, en totes les fases del seu cicle de vida, contribueix a aconseguir nivells alts d'èxit, de confiança i de benefici a curt i llarg termini. Per tant, tenint en compte totes les vessants (social, legal, ambiental i de veracitat) que ofereix la

perspectiva ètica s'assegura la innovació sostenible a l'Administració i es fomenta la confiança pública. Per aquest motiu, l'elecció dels criteris es basa en els principis ètics corporatius en l'ús de la IA.

3. Criteris per a la utilització ètica i responsable de l'Administració de la Generalitat i el seu sector públic institucional, i dels empleats públics

3.1. Introducció

Abans de fer ús de la IA, penseu, reflexioneu i actueu en conseqüència de forma ètica i responsable. L'ètica té relació directa amb la presa de decisions i amb els criteris i els valors que s'incorporen en aquest procés.

L'ètica és un concepte dinàmic que implica una valoració prèvia a un comportament o a una decisió, i configura el nostre caràcter i la nostra forma d'actuar. Es tradueix de forma planera en "fer bé les coses, fins i tot quan ningú et vegi". Aquest aspecte està lligat al deure d'imparcialitat i objectivitat de tot servidor públic i constitueix la base de la integritat professional.

El respecte és essencial per orientar la forma en què utilitzem la IA i les dades, així com l'impacte dels seus usos. Actuar amb respecte envers les persones i la integritat professional implica esforçar-se per aconseguir uns resultats justos, equitatius i confiables.

També és necessari prendre consciència respecte de l'impacte d'ús de la IA generativa en els valors públics que, com a institució, ens representen.

En resum, implica que, com a organització i com a empleat públic, es valori el comportament en dos aspectes: acció i decisió. Per això, es distingeix entre:

- Fase d'anàlisi de les finalitats: identificar si el seu ús s'alinea amb els principis d'actuació administrativa.
- Fase d'avaluació prèvia: identificar si cal aplicar la IA generativa per resoldre la vostra tasca i pensar en les conseqüències d'allò que fareu i la seva dimensió ètica a l'hora d'introduir les dades, interpretar-les i distribuir-les.
- Fase de verificació i adequació del resultat.

Si feu ús de la IA generativa com a consumidor, heu de saber quin tipus d'IA esteu utilitzant, ser conscients de la seva precisió limitada, i avaluar-ne l'aptitud per al propòsit que us ocupa. Fins i tot, si l'Administració actua com a proveïdora, responsable d'una eina a mida o responsable del desplegament d'un sistema sota la seva autoritat, perquè s'han de complir els requeriments que preveu el RIA.

Finalment, en la fase de verificació del resultat, heu de garantir que, d'acord amb el vostre nivell de coneixement en la matèria, podeu avalar la certesa i la validesa de la resposta obtinguda.

A través d'aquest document, es vol ajudar a identificar els possibles riscos ètics a través d'aquestes pautes d'avaluació prèvia perquè tingueu una comprensió clara dels perills i del que suposa no respectar els principis ètics, i perquè vetlleu per mantenir-ne l'alineament.

3.2. Directrius

Amb caràcter previ, s'escau assenyalar, per la seva criticitat, aspectes que es consideren no aptes per aplicar la IA generativa.

3.2.1. Què no pot fer un empleat públic amb la IA generativa

Els empleats públics no han de fer ús de la IA generativa per a:

- Processos crítics o urgents que afectin la presa de decisió sobre competències pròpies de l'Administració de la Generalitat sense supervisió humana.
- Fer-ne un ús intencionadament indegut i maliciós, com ara la creació d'informació falsa o la manipulació d'imatges, entre d'altres.
- Utilitzar el resultat directament sense la supervisió humana.
- Fer ús d'informació confidencial, protegida, sensible o de caràcter personal, excepte que s'utilitzi amb eines internes en entorns segurs o aprovades, i que específicament es permeti aquest ús.

Aquestes directrius es vinculen amb els principis ètics i els valors al procés de decisió. Per tant, de forma sintètica, distingim:

- Fase d'avaluació prèvia (segons la finalitat, la pertinença del seu ús i el risc).
- Què he de tenir en compte a l'hora d'escriure el *prompt* o les indicacions.
- Com faig servir el resultat de la IA generativa.

3.2.2. Fase d'avaluació prèvia

Avalueu prèviament si realment és necessari l'ús de la IA generativa en la vostra tasca

La supervisió humana en l'ús de la IA vol prevenir o reduir al mínim els riscos per als drets fonamentals, la salut o la seguretat que puguin sorgir. Un ús indegut amb informació no fiable pot propagar la desinformació i generar resultats erronis amb un efecte directe sobre la ciutadania.

- Informeu-vos i formeu-vos adequadament abans d'utilitzar la IA generativa. Assegureu-vos de conèixer-ne el funcionament, les limitacions i els possibles riscos associats. Busqueu informació en fonts fiables, participeu en sessions de capacitat, i consulteu les directrius i recomanacions establertes per tota l'Administració.

- Considereu si és apropiat fer ús de la IA generativa. Per exemple, si podeu trobar una informació fàcilment a través d'un cercador o de la consulta directa d'una font fiable, no cal que utilitzeu una eina d'IA generativa, o si trigueu un minut a fer una tasca manualment, potser tampoc cal fer-ne ús.
- Valoreu si per a la tasca que heu de dur a terme, l'ús de la IA generativa és respectuós amb les instruccions o les directrius pròpies del vostre servei. En concret, valoreu si és apte per complir el propòsit de la vostra unitat i si està alineat amb la seva missió i el compliment de polítiques públiques i principis de l'actuació administrativa.
- Reflexioneu si teniu alternatives que permetin reduir la petjada ambiental. S'utilitza una gran quantitat d'energia i aigua per entrenar i fer funcionar les eines d'IA, per la qual cosa l'ús indiscriminat té conseqüències negatives per al medi ambient.
- Determineu el risc, avalueu el possible impacte que podria tenir una errada i valoreu si teniu elements de control per detectar l'errada.
- Davant del dubte, segurament l'ús no és adequat. Consulteu sempre el vostre superior o les directrius que s'estableixin prèviament.
- No useu la IA generativa per a tasques de les quals no pugueu comprovar la fiabilitat de la resposta. Heu de tenir el nivell d'expertesa adequat en l'àmbit per comprovar la fiabilitat i seguretat de la resposta, de tal manera que pugueu detectar els aspectes no fiables, els errors o la baixa qualitat.
- Actueu sempre amb transparència, integritat i honestedat. Actuar amb integritat significa reconèixer i mantenir la vostra responsabilitat en els usos de les dades, de manera coherent amb el Codi ètic del servei públic de Catalunya. Aquest compromís també implica garantir la integritat i precisió de les dades recopilades i utilitzades en l'ús de la IA, així com el compliment estricte de la normativa aplicable.

Per exemple, heu de determinar si una persona pot rebre un ajut d'habitatge. Esteu pensant en la possibilitat d'usar un bot de conversa d'IA generativa que heu vist per internet per valorar la informació rebuda, perquè és el primer cop que ho feu tot. La idea que teniu és dir-li al bot de conversa els requisits i passar-li la informació de cada persona, de tal manera que us digui si els compleixen o no. En aquest cas, ni el bot de conversa d'IA generativa ni la persona no tindran prou coneixement específic en l'àmbit i, a més, no s'hi poden introduir informació delicada i dades personals especialment protegides. Per tant, cal descartar l'ús d'aquell bot de conversa d'IA generativa i la millor opció seria demanar ajuda als companys amb més experiència, o bé suggerir a la unitat responsable que creï un bot de conversa a partir de fonts oficials.

Aquest criteri està relacionat amb els **principis**:

- Principi d'acció i supervisió humana: garantir la protecció dels drets fonamentals mitjançant la intervenció humana i l'ús responsable de les dades.

- Principi de desenvolupament sostenible i protecció del medi ambient: promoure pràctiques tecnològiques que tinguin en compte la sostenibilitat i protegeixin el medi ambient.

No useu la IA per a processos on es valoren les habilitats, les qualitats o la situació d'una persona

No és possible fer valoracions de diferents aspectes d'una persona a través de la introducció d'informació en sistemes d'IA generativa per garantir l'aplicació del principi d'igualtat i preservar una decisió no esbiaixada i presa amb raonament just i equitatiu. Aquests aspectes poden afectar la valoració professional o social, o qualsevol altre àmbit de l'esfera personal.

Per exemple, no es pot fer servir la IA generativa per elaborar documents de processos competitiu destinats a valorar les capacitats de les persones, com ara memòries per a processos selectius, tret que s'admeti explícitament en les bases. Aquesta praxi podria, per exemple, faltar a la igualtat de condicions i d'oportunitats.

Aquest criteri està relacionat amb els **principis**:

- Principi de transparència i retiment de comptes: fomentar la difusió amb caràcter permanent i actualitzat i la claredat i el retiment de comptes en l'ús de dades, IA i altres algorismes, així com la participació ciutadana i els processos de treball transparents.
- Principi de justícia, equitat i no-discriminació: assegurar l'ús equitatiu, just i respectuós de les dades (equitat, no-discriminació i respecte als drets humans), tenint en compte que cal assegurar que les dades siguin igual d'accessibles per a tothom i s'ajustin als requisits funcionals i d'interacció que garanteixin la utilització autònoma, segura i amb comoditat a totes les persones respectant la legislació establerta en matèria de protecció de dades, accessibilitat i propietat intel·lectual i intel·lectual. Els processos també han d'assegurar l'accés just i no discriminatori a les capacitats i els beneficis dels usos avançats i sistemes de la IA per a tots els que hi participen.

No lesioneu el dret a propietat intel·lectual

Amb caràcter previ, cal recordar que l'Administració de la Generalitat disposa d'uns criteris corporatius sobre l'edició. Segons aquests criteris, l'aposta en l'ús de les imatges, igual que amb la resta de dades, és la reutilització, per la qual cosa es recomana l'ús d'imatges lliures (**ús de les llicències i els drets d'autoria**).

Una conseqüència lògica de la integritat és també el respecte i la protecció dels drets de les persones i el seu treball o creacions pròpies, com podria ser la propietat intel·lectual.

La IA generativa pot haver fet servir imatges o contingut creatiu sense l'autorització dels seus autors. Per això, quan creeu un text, imatge o vídeo, podeu estar reproduint el contingut d'algú altre i podríeu vulnerar-ne el dret a la propietat intel·lectual.

No feu servir la IA generativa per generar contingut que pugui infringir drets. En cas de dubte sobre si podria infringir-los o no, com ara en la generació d'imatge o vídeo, no la feu servir.

Les eines d'IA no són autores dels seus resultats perquè no poden ser responsables de l'exactitud, la integritat i l'originalitat del treball o de la tasca. L'autoria i la responsabilitat dels treballs o les tasques han de recaure exclusivament en humans i no en màquines.

A més, reviseu que, en el cas de les respostes que contenen referències bibliogràfiques o normatives, aquestes siguin reals i estiguin citades correctament, i que la temàtica de l'article, el llibre o l'altra publicació sigui pertinent. Així mateix, s'han de respectar les llicències d'ús aplicables.

Per exemple en una licitació de creació de continguts creatius, expresseu explícitament que en permeteu l'ús, que cal etiquetar el contingut generat per IA generativa, i que si es pren com a base el treball d'un autor o autora, vetlleu per protegir i respectar els drets d'autor i la propietat intel·lectual de les persones autores, artistes i legítimes titulars dels drets sobre les obres artístiques i literàries.

A efectes pràctics, això implica tenir en compte algunes recomanacions, com ara:

- Obtenir l'autorització prèvia per a l'ús legal de les obres en qualsevol fase de construcció o disseny, desenvolupament i explotació dels models d'IA.
- Remunerar l'ús legal de les obres autoritzades.
- Els productes generats per la IA no poden considerar-se obres de creació intel·lectual ni poden estar protegits per la normativa de drets d'autor i propietat intel·lectual.
- Els productes obtinguts amb IA generativa s'han d'identificar com a tals.
- Exigiu transparència i informació accessibles als proveïdors d'IA; les persones autores, artistes i titulars de drets d'autor tenen dret a conèixer si les seves obres s'han utilitzat per a l'entrenament o qualsevol fase del cicle de vida, com el desenvolupament i la posada en producció dels models d'IA.
- No feu ús de les clàusules de consentiment tàcit per a l'ús de les obres per la IA per la mera utilització de plataformes i xarxes socials.

Aquest criteri està relacionat amb els **principis**:

- Principi de justícia, equitat i no-discriminació: assegurar l'ús equitatiu, just i respectuós de les dades (equitat, no-discriminació i respecte als drets humans), tenint en compte que cal assegurar que les dades siguin igual d'accessibles per a tothom, i s'ajustin als requisits funcionals i d'interacció que garanteixin la utilització autònoma, segura i amb comoditat a totes les persones respectant la legislació establerta en matèria de protecció de dades, accessibilitat i propietat industrial i intel·lectual. Els processos també han d'assegurar l'accés just i no discriminatori a les capacitats i els beneficis dels usos avançats i sistemes de la IA per a tots els que hi participen.

3.2.3. Criteris d'ús de la informació inclosa en la indicació o *prompt*

Vigileu si podeu fer ús o no de les dades personals o la informació confidencial

La informació o les dades que l'empleat públic usa al seu lloc de treball tenen caràcter confidencial, a excepció de les dades que ja tinguin la consideració de públiques.²

Per tant, amb **caràcter general, no podeu incorporar-hi dades personals, informació confidencial o dades sensibles.**

Cal tenir en compte que qualsevol entrada a un bot de conversa d'IA generativa en línia es transmet al proveïdor del bot de conversa, que, posteriorment, podria utilitzar aquesta informació per generar resultats futurs que es podrien divulgar al públic en general. Mitjançant l'ús d'eines externes d'IA generativa, hi ha el risc que el personal posi involuntàriament a disposició del públic informació només per a ús intern i organitzatiu de l'Administració o informació sensible, confidencial o protegida per drets de protecció de dades personals i de propietat intel·lectual.

- Feu consultes amb informació i dades tot respectant la legalitat (sense dades personals, privades o sensibles) i el propòsit per al qual es tracten.
- Utilitzeu la resposta de forma adequada i d'acord amb el propòsit de la consulta.
- Heu d'estar en disposició de garantir que les dades que es fan servir s'obtinguin de fonts confiables, acreditades i verificables.

Per exemple, si teniu un document administratiu amb dades personals i voleu passar-lo a comunicació clara, no ho feu a ChatGPT, però, en canvi, sí que ho podeu fer al Copilot corporatiu, sempre que no siguin dades especialment protegides.

Per això, si heu d'utilitzar dades de caràcter personal o confidencial, **únicament heu d'utilitzar una eina autoritzada i homologada prèviament per l'organització i ha de ser perquè està previst en les vostres funcions.**

L'Administració pot desenvolupar eines internes d'IA generativa especialitzades i adaptades a àmbits específics i fetes sota petició, amb nivells alts de configuració de privacitat i seguretat i sense que les dades surtin del seu entorn. En alguns casos, i només quan s'estableixi expressament, es podria autoritzar l'ús de dades personals en aquestes aplicacions, d'acord amb les directrius i els protocols que específicament s'estableixin.

Per a una major comprensió, distingim:

Dades personals³

². [Instrucció 8/2020, de 24 de novembre](#), sobre l'ús de les tecnologies de la informació i la comunicació a l'Administració de la Generalitat.

³. [Línies bàsiques de la protecció de dades personals al sector públic i RGPD](#).

Les dades personals són totes les dades que permeten identificar, directament o indirecta, una persona física. Algunes dades són útils exclusivament per identificar les persones; d'altres permeten conèixer aspectes de la seva personalitat, avaluar la persona o conèixer aspectes de la seva història i circumstàncies (estudis, família, vida laboral, salut, etc.).

Són exemples de dades personals:

- Dades identificadores: nom, adreça, fotografia o DNI.
- Dades de característiques personals: lloc de naixement, nacionalitat o sexe.
- Dades de circumstàncies socials: aficions, estils de vida, situació familiar, opinions, valors, creences, ideologia, etc.
- Dades acadèmiques i professionals: historial acadèmic i experiència professional.
- Dades econòmicofinanceres: dades bancàries, assegurances, subsidis, targetes de crèdit, etc.
- Dades d'ocupació laboral: cos, categoria, lloc de treball, historial laboral, etc.
- Dades relatives a la comissió d'infraccions i delictes.

Dades personals de categoria especial

Les categories especials de dades demanen una protecció més elevada i, amb caràcter general, la normativa en prohibeix el tractament. En concret, són les dades que revelen l'origen ètnic o racial, les opinions polítiques, les conviccions religioses o filosòfiques, l'afiliació sindical, les dades genètiques, les dades biomètriques adreçades a identificar de manera unívoca una persona física, les dades relatives a la salut o a la vida sexual o les orientacions sexuals d'una persona física.

Dades confidencials no personals

Les dades confidencials són les que, sense ser dades personals, tenen un accés restringit, és a dir, que es garanteix que la informació és accessible només per a aquells autoritzats a tenir-hi accés, per exemple, per qüestions de propietat intel·lectual o perquè hi ha alguna normativa que ho determini.

En qualsevol cas, segueix sempre les directrius corporatives establertes de protecció de dades i de seguretat.

Aquest criteri està relacionat amb els **principis**:

- Principi ètic de confidencialitat, privacitat i seguretat: protegir la informació personal, garantir la seguretat digital i custodiar la informació confidencial. En especial, la confidencialitat de les dades en la recollida, el tractament i la difusió de resultats i les dades personals.
- Principi de justícia, equitat i no-discriminació: assegurar l'ús equitatiu, just i respectuós de les dades (equitat, no-discriminació i respecte als drets humans), en tant que

estableix com a criteri protegir els interessos dels titulars de les dades i vetllar per evitar que hi hagi efectes nocius derivats del seu ús.

- Principi d'acció i supervisió humana: garantir la protecció dels drets fonamentals mitjançant la intervenció humana i l'ús responsable de les dades.

3.2.4. Criteris d'ús del resultat

No perpetueu la discriminació o la desigualtat de tracte ni amplifiqueu els biaixos o els estereotips

Els models de llenguatge es poden entrenar amb dades d'internet que no s'han supervisat ni són representatives, o que, tot i ser representatives, són esbiaixades i, per tant, la seva resposta serà també esbiaixada. Cal recordar que les eines les han creat persones que poden tenir els seus biaixos i, en conseqüència, les eines també en poden tenir.

Això pot comportar respostes excloents o ofensives per a certs individus, grups o drets protegits, fenomen conegut com a invisibilització i discriminació algorítmica. Un ús tenint en compte les limitacions de l'eina permet evitar reproduir resultats erronis o esbiaixats.

Heu de ser conscients que existeixen desigualtats sistèmiques a partir de factors socials, com ara el gènere, l'ètnia o la classe social, entre d'altres. En conseqüència, tant els desavantatges com els avantatges que té una persona en un moment i un lloc determinats no poden entendre's examinant de forma aïllada els diversos elements de la seva realitat i identitat. Per contra, heu de prestar atenció a la realitat en conjunt.

Tenir aquesta mirada us permetrà ser conscients i visibilitzar les diferències entre dones, homes o qualsevol altre grup de persones i, alhora, us permetrà tenir en compte els efectes dels sistemes de discriminació, com ara el sexisme, l'edatisme, el racisme o el classisme, entre d'altres.

En aquest context, es recomana que, més enllà del cas de l'ús de la IA generativa, tots els processos d'elaboració i anàlisi de dades adoptin la perspectiva de gènere i l'enfocament intersectorial de la realitat de les persones per incloure les experiències i les realitats de persones i grups de població que poden presentar vulnerabilitats específiques.

A l'hora d'escriure la consulta a la IA generativa, reviseu que la vostra indicació no sigui tendenciosa i derivi en respostes que puguin ser discriminatòries o ofensives amb la identitat de gènere, l'educació, la religió, la discapacitat, l'heterosexisme o la llengua, entre d'altres. En conseqüència, una primera recomanació en aquest sentit és evitar l'ús de dades no representatives, no contrastades o esbiaixades en l'anàlisi de dades en el si de l'Administració pública.

Per exemple, una eina d'IA generativa fa una proposta de possibles llocs de feina i recomana als homes i les dones feines que respectivament perpetuen els rols de gènere, sense tenir-ne en compte la trajectòria professional.

Un altre ús discriminatori seria que una eina d'IA generativa determini, només perquè detecta el nom d'una persona com a nom d'origen estranger, que aquella persona estigui en una situació de pobresa, ja que això estaria basat en estereotips discriminatoris.

Aquests són exemples de males praxis i cal recordar que quan es genera i publica contingut amb IA generativa, l'usuari que en fa ús té una responsabilitat semblant a les referències d'internet o dels llibres o documents. És necessari seleccionar les fonts amb cura, elaborades per entitats reconegudes, consensuades i independents, i comprovar que els continguts són ètics, fiables i de qualitat i que no contradiguin el marc legal, igual que amb la resta de fonts.

Aquest criteri està relacionat amb els **principis**:

- Principi de justícia, equitat i no-discriminació: assegurar l'ús equitatiu, just i respectuós de les dades (equitat, no-discriminació i respecte als drets humans). De forma singular, estableix el criteri de vetllar per l'equitat de representació, de recursos, d'accés i de resultats.
- Principi d'acció i supervisió humana: garantir la protecció dels drets fonamentals mitjançant la intervenció humana i l'ús responsable de les dades. De forma singular, estableix el criteri de tenir en compte la relació amb l'entorn sociocultural i econòmic, i de ser conscients de l'impacte de les accions i decisions humanes en relació amb l'ús de dades.
- Principi de governança dinàmica de l'ètica: fomentar una avaluació constant de l'ètica mitjançant l'ús de dades i tecnologies d'IA, i garantir l'actualització del desenvolupament ètic de la IA, la seva implementació i el seu foment.

No feu passar com a vostra una informació obtinguda per IA generativa

Les eines d'IA no són autores dels seus resultats perquè no poden ser responsables de l'exactitud, la integritat i l'originalitat del treball o de la tasca. L'autoria i la responsabilitat dels treballs o les tasques han de recaure exclusivament en humans i no en màquines.

El plagi també pot ser una forma de tractament deshonest de les dades i de la informació, aspecte que pot comportar una forma de frau.

Com a derivada de la integritat, en el context de la veracitat i l'honestedat, si heu obtingut informació que l'ha generat una eina d'IA, no us n'atribuiu l'autoria. Podeu fer-ho si hi aporteiu el vostre coneixement, milloreu el resultat obtingut o si aporteiu valor a la qualitat del seu treball.

Per exemple, si milloreu el codi d'una anàlisi o una presentació gràfica a través de l'ús de sistemes d'IA generativa, no us podeu apropiat de la modificació excepte si canvieu alguna part del resultat de la consulta. De la mateixa manera, si demaneu a la IA generativa que us doni una anàlisi del risc d'exclusió social de la població que viu al barri del Raval de Barcelona, a part de revisar la informació, si no la modifiqueu, no us en podeu atribuir l'autoria.

Sigueu transparents en l'ús de la informació

La transparència és la pedra angular per fomentar la confiança pública en l'ús de les dades i en l'exercici dels drets de les persones. Aquest principi ha de considerar-se de manera significativa i contextualitzada adequadament al context que correspon i als destinataris últims.

Els empleats públics han de ser capaços d'explicar les decisions que prenen, i això inclou documentar adequadament quan han fet ús d'eines com la IA generativa. Aquest fet permet mantenir l'explicabilitat i contribuir a facilitar l'auditabilitat, en cas que sigui necessari. Per això, esmenteu específicament que en la tasca ha intervingut una eina d'IA generativa i proporcioneu la traçabilitat tant de les dades com del mateix model d'IA que s'ha utilitzat. A més, heu d'oferir informació actualitzada, objectiva, veraç, completa, útil i comprensible.⁴

És necessari que s'expliciti l'ús de la IA generativa en la interacció directa amb la ciutadania. En canvi, pel que fa a l'ús de la IA generativa en les actuacions administratives, és convenient que quedi traçabilitat del *prompt* o la indicació formulada, del resultat, de la data de la consulta, de l'eina i de la versió d'ús.

Per exemple, si una actuació administrativa com l'enviament d'una notificació per al pagament d'un impost s'ha generat amb intel·ligència artificial generativa, cal que ho indiqueu en el mateix document. Així mateix, és important que sapigueu i documenteu en quina de les parts heu fet ús d'aquesta tecnologia, i que reviseu com s'ha efectuat el càlcul de la base imposable, la selecció del tipus i la base que s'ha de liquidar.

Aquest criteri està relacionat amb els **principis**:

- Principi de transparència i retiment de comptes: de forma singular, està vinculat amb el criteri d'assegurar la transparència de les actuacions administratives per mitjà dels instruments de difusió que permetin un accés ampli i fàcil a les dades, i, així, facilitar a la ciutadania l'auditoria i promoure'n la capacitat de comprendre i conèixer les actuacions de l'Administració.
- Principi d'acció i supervisió humana: de forma singular, estableix el criteri que és necessari establir una responsabilitat ètica en les tasques de supervisió humana que vagi més enllà de les responsabilitats legals obligatòries i que s'apliqui de forma transversal, des de l'obligació prèvia de prevenció fins a l'eventual reparació de qualsevol dany.
- Principi de governança de les dades avançades i de sistemes d'IA: de forma singular, està vinculat al criteri d'incrementar la comunicació i la divulgació per facilitar la comprensió, la interpretació i la reutilització de les dades.

⁴ [Codi ètic del servei públic de Catalunya](#).

3.3. Qüestions pràctiques

3.3.1. Redacció de *prompts* o indicacions

1. Elaboreu indicacions amb contingut concís, lògic i explícit sobre el context, i que sigui adaptatiu i reflexiu. Com més clares i específiques siguin les indicacions a l'eina, més efectiva serà la resposta i, per tant, més rellevant i creativa.
2. Examineu que la indicació no contingui dades de caràcter personal ni informació sensible, confidencial o privada, excepte si feu ús d'una eina autoritzada i homologada prèviament per l'organització i està previst en les vostres funcions.
3. Prepareu les dades de la consulta o la indicació.
4. Utilitzeu identificadors genèrics per no individualitzar un determinat interessat en un procediment administratiu concret.
5. Verifiqueu el context de la consulta.
6. Confirmeu que no hi ha traçabilitat de les dades.

3.3.2. Material de suport

Abans de fer servir la IA generativa, feu ús dels recursos formatius disponibles tant a la vostra unitat com a l'administració en general per garantir-ne un ús ètic i responsable. Participeu en sessions de capacitació, consulteu les guies i protocols establerts i feu ús de materials actualitzats sobre bones pràctiques en la matèria.

En cas de dubte, recordeu als vostres responsables i als canals oficials de formació de l'Administració. Recordeu que la capacitació, el coneixement i la comprensió dels riscos i les oportunitats dels sistemes d'IA formen part d'una obligació legal que la Generalitat de Catalunya ha de complir.

3.3.3. Conclusió

En conclusió, tot i que es poden explorar models d'IA generativa disponibles en línia per millorar l'eficiència i la qualitat de la vostra feina, ho hauríeu de fer sempre tenint en compte que:

- No hauríeu de revelar informació privada o dades personals als sistemes d'IA generativa, excepte si feu ús d'una eina autoritzada i homologada prèviament per l'organització en entorns segurs i està previst en les vostres funcions.
- Heu de tenir cura de la qualitat de les dades usades per aconseguir les respostes de sistemes d'IA generativa. Recordeu que aquests sistemes tenen limitacions amb relació a la seva fiabilitat i qualitat, tant o més que quan es fa una cerca a internet.
- Sempre heu de parar atenció als riscos i les limitacions inherents dels sistemes d'IA generativa. Recordeu que tenim una responsabilitat ètica.

- Heu de comprovar qualsevol resultat en la resposta del sistema d'IA generativa, en especial, per vetllar per la qualitat de les dades obtingudes i per detectar-hi possibles biaixos, informació inexacta i infraccions dels drets d'autor o de tercers.

Al mateix temps, és important participar en sessions de formació contínua per conèixer les limitacions, els riscos i les aplicacions adequades d'aquesta tecnologia, per assegurar-ne un ús alineat amb els valors de l'Administració.