



GUIA RÀPIDA:
GARANTIES I PRINCIPIS
DE LA INTEL·LIGÈNCIA
ARTIFICIAL GENERATIVA
EN L'ADMINISTRACIÓ
LOCAL

Localret

Agraïments

Aquest document és fruit d'un procés col·laboratiu d'anàlisi, reflexió i diàleg, al voltant de la intel·ligència artificial (IA) i la seva aplicació en l'administració pública. El grup de treball que ha treballat aquesta iniciativa està vinculat a l'Eix 1: Administració i Serveis Públics Intel·ligents de l'Agenda Digital dels Municipis de Catalunya, iniciativa que s'emmarca en l'estratègia del municipi digital i que el Consorci Localret du a terme amb la finalitat d'acompanyar els ajuntaments en la seva transformació digital.

El grup de treball Eix 1 -Intel·ligència Artificial ha treballat en 3 documents:

- Definició de metodologies de treball i protocols per a la implementació de sistemes algorítmics
- Guia ràpida: Garanties i principis de la intel·ligència artificial generativa en l'administració local
- Bones pràctiques i casos d'ús d'Intel·ligència Artificial

El present document ha estat elaborat pel Consorci Localret i ha comptat amb la participació i col·laboració de professionals que formen part de l'Ajuntament de Lleida, l'Ajuntament de Reus, l'Ajuntament de Mataró, l'Ajuntament de Viladecans, Ajuntament de Rubí i la Diputació de Barcelona.

També hi ha participat amb les seves indicacions i/o recomanacions l'Ajuntament de Barcelona.

Data d'elaboració: Octubre de 2023

Propòsit

L'aparició d'aplicacions com ChatGPT, Bard o Perplexity per a la generació de textos i altres com DALL-E o Midjourney per a la creació d'imatges ha obert molts interrogants sobre el seu ús en diversos àmbits. Aquestes eines es consideren part de la Intel·ligència Artificial Generativa, que comprèn un conjunt de tecnologies que apliquen tècniques d'aprenentatge automàtic a grans volums de dades per generar contingut basat en les entrades dels usuaris. La ràpida evolució d'aquestes eines, la comprensió del seu funcionament i les implicacions del seu ús en la societat continuen sent temes de debat i d'investigació.

La gran potència de les tecnologies d'Intel·ligència Artificial Generativa planteja qüestions fonamentals sobre la seva aplicació responsable, ètica i els seus efectes en àmbits com la comunicació i la creació de contingut. Per aquestes raons, és necessari un marc regulador que asseguri una aplicació adequada de la Intel·ligència Artificial en general. En aquesta línia, la [Declaració de Mont-Real per a un Desenvolupament Responsable de la IA](#) (2015) va establir una sèrie de principis fonamentals per guiar el desenvolupament dels sistemes d'intel·ligència artificial, SIA; i serveix com a referència per a futurs estàndards i normatives en aquest àmbit.

Aquests principis es centren en:

- **Benestar.** Els SIA han de contribuir en la millora de les condicions de vida, de salut i laboral de les persones
- **Autonomia.** Els SIA s'han de desenvolupar respectant l'autonomia de les persones per a que disposin d'un major control sobre la seva vida i allò que les envolta.
- **Protecció de la privacitat i la intimitat.** Es prendran mesures per evitar qualsevol intrusió del SIA (Sistemes d'Intel·ligència Artificial) o dels sistemes d'adquisició de dades.
- **Solidaritat.** Es preservaran els vincles entre les persones i no es substituiran en aquelles tasques que requereixin relacions humanes de qualitat.
- **Participació democràtica.** Les persones hem de ser coneixedores de l'ús dels SIA quan ens veiem afectats per les seves decisions.
- **Equitat.** Els SIA s'han de dissenyar i entrenar per a que no reproduïxin patrons de discriminació basats en diferències socials, ètniques, sexuals, culturals,...
- **Inclusió de la diversitat.** Els SIA han de ser inclusius i donar compte de la diversitat social evitant l'homogeneïtzació.

- **Prudència.** Tothom que treballi en IA haurà d'anticipar-se als possibles efectes adversos dels SIA
- **Responsabilitat.** Recau sobre les persones la responsabilitat de les recomanacions dels SIA i de les accions que se'n derivin.
- **Sostenibilitat.** El desenvolupament i ús dels SIA hauran de garantir al sostenibilitat del medi ambient.

Raó de ser del document

Aquest document té com a objectiu facilitar la introducció de la Intel·ligència Artificial Generativa a les nostres administracions, sempre adherint-se als principis i garantint les màximes salvaguardes. A través de preguntes formulades a eines d'IA, es proporcionen una sèrie d'indicacions i recomanacions que s'han de tenir en compte quan s'utilitzen aquests sistemes. Cal destacar que aquest no pretén ser un manual exhaustiu d'ús d'eines d'IA Generativa, ja que les seves possibilitats són moltes i en constant creixement.

Mentre esperem que futures normatives evolucionin i defineixin la forma d'aplicar aquestes tecnologies dins del context de les nostres administracions, aquestes recomanacions es presenten com un punt de partida per a l'ús d'un conjunt de tecnologies que ja tenen una presència significativa en el nostre quotidià.

Més enllà de Bard, Chat GPT i Perplexity, cada dia apareixen nous usos de la IA que faciliten el treball diari. Els principis establerts en aquest document també seran aplicables a aquestes eines i ens ajudaran a incorporar-les a les nostres administracions.

Aquest document és una adaptació del document: City of Boston Interim Guidelines for Using Generative AI¹

Principis bàsics per a l'ús d'IA generativa

Principi 1

- ✓ Revisar i validar sempre el contingut generat per la IA, especialment el que s'acabarà publicant d'alguna forma.

Els sistemes d'IA poden generar continguts amb informació errònia, parcial, inventada i/o discriminatòria per a determinats col·lectius. Els treballadors hauran de supervisar els resultats obtinguts i vetllar per a la seva veracitat i no discriminació. Caldrà posar especial atenció en aquells continguts utilitzats per a la presa de decisions.

Principi 2

- ✓ No utilitzar mai dades personals, ja que desconeixem l'ús que en pot fer la IA. En aquest sentit, l'[Agència de Protecció de Dades \(APDCAT\)](#)² afirma que l'ús de ChatGPT pot suposar un risc per a la privacitat i recomana no utilitzar-lo amb dades personals fins que el Comitè Europeu de Protecció de Dades es pronunciï o emeti recomanacions, directrius o bones pràctiques que garanteixin el respecte a la normativa de protecció de dades.

¹ Garcés, Santiago (2023). City of Boston Interim Guidelines for Using Generative AI. <https://www.boston.gov/sites/default/files/file/2023/05/Guidelines-for-Using-Generative-AI-2023.pdf>

² https://apdcat.gencat.cat/web/.content/01-autoritat/normativa/documentos/Recomanacio12023_ChatGPT.pdf

Podeu consultar el [següent document](#)³ de l'Escola d'Administració Pública de Catalunya sobre el que s'entén per dades personals

Principi 3

✓ No introduir dades sensibles

La introducció de dades sensibles en un sistema d'IA comporta un risc de pèrdua de privadesa si aquestes dades es fan públiques o es comparteixen sense autorització. També hi ha un risc de seguretat si les dades són interceptades o vulnerades de qualsevol manera.

Principi 4

✓ Indicar que s'ha fet ús d'IA en els continguts generats

Manifestar que s'ha fet ús d'IA generativa en els continguts genera confiança mitjançant la transparència i pot ajudar els altres a detectar errors.

Es recomana: incloure una referència com a peu de pàgina indicant model i versió utilitzada.

Exemples:

- Descripció generada per ChatGPT v3.5 i editada per Joan Garcia.
- Resum generat per Google Bard v1.0.1.

Exemples d'aplicació de la IA generativa

La Intel·ligència Artificial Generativa utilitza el terme "PROMPT" per referir-se a les instruccions, preguntes o ordres amb les que s'interactua amb els SIA. Tot seguit es proposen diferents usos de la IA i els prompts que es poden utilitzar a Bard, Perplexity o Chat GPT⁴.

1. Generar contingut

La IA pot generar contingut de text de manera automàtica, com ara notes de premsa, informes, correus electrònic, etc.

`Prompt: Redacta un escrit adreçat als ciutadans d'un municipi sobre les conseqüències dels talls en el subministrament elèctric`

Com més detallat sigui el prompt, més rellevant serà la resposta. Es recomana no fer ús de la IA en matèries especialment sensibles.

Aplicació dels principis:

- Principi 1: revisió del text
 - Comprovar que el text s'ajusta al que s'ha demanat a la IA.
 - Vetllar per la inclusió i la diversitat:
 - Cal utilitzar altres llengües?
 - Cal diferenciar per grups d'edat?
 - Hi ha col·lectius especialment vulnerables a tenir en compte en el text? Gent gran, malalts, ...
 - Fer correccions ortogràfiques i sintàctiques.

³ <https://formaciooberta.eapc.gencat.cat/contingutsdels cursos/abpd/020/inici.html>

⁴ Chat GPT 3.5 no disposa d'informació posterior a setembre de 2021. Tampoc té accés a recursos disponibles on-line a diferència de Bard v1.0.1.

- Principi 2: No utilitzar noms, cognoms o dades de caràcter personal en el prompt, en tot cas afegir-los un cop el text s'hagi generat per la IA.
- Principi 4: Informar que s'ha fet ús d'IA en la redacció del text.

2. Localitzar fonts d'informació

La IA retorna resultats més rellevants que els cercadors tradicionals en la cerca de fonts d'informació o bibliografia sobre qualsevol tema.

Prompt (Bard): on puc trobar plecs per a la contractació de serveis de neteja viària d'un municipi?

Aplicació dels principis:

En aquest cas no serien d'aplicació els principis, ja que obeeix a un ús intern i no es genera cap text per ser publicat.

3. Elaborar resums i extractes amb condicions específiques.

La intel·ligència artificial és capaç de processar una gran quantitat de textos, el que li permet conèixer el contingut d'un text.

Prompt 1: Explica en un paràgraf l'Estratègia d'Intel·ligència Artificial de Catalunya.

Prompt 2: Implicacions de la Llei 19/2014 de transparència per a un ajuntament petit.

Prompt 3 (Bard): Resumeix:

https://apdcat.gencat.cat/web/.content/03-documentacio/Reglament_general_de_proteccio_de_dades/3132.pdf

Aplicació dels principis:

- Principi 1: revisió del text
 - Comprovar que el text s'ajusta al que s'ha demanat a la IA.
 - Fer correccions ortogràfiques i sintàctiques.
 - No confiar en la IA si el que es vol són respostes molt acurades.
- Principi 4: Informar que s'ha fet ús d'IA en la redacció del text si s'utilitza en algun tipus de publicació.

4. Adaptar textos a diferents audiències.

La intel·ligència artificial pot adaptar textos gràcies al seu model de llenguatge

Prompt: explica les funcions d'un ajuntament mitjà a alumnes de 2n de primària

Aplicació dels principis:

- Principi 1: revisió del text
 - Comprovar que el text s'ajusta al que s'ha demanat a la IA.
 - Fer correccions ortogràfiques i sintàctiques.
- Principi 4: Informar que s'ha fet ús d'IA en la redacció del text si s'utilitza en algun tipus de publicació.

5. Traducció de textos

Fent ús d'algoritmes de processament del llenguatge natural la IA pot analitzar el contingut del text i produir traduccions fiables.

Prompt (Bard): tradueix al francès

<https://www.localret.cat/municipi-digital/agenda-digital/>

Aplicació dels principis:

- Principi 1: revisió del text
 - Revisar la traducció per garantir que s'ajusta a l'original.
 - Fer correccions ortogràfiques i sintàctiques.
- Principi 4: Informar que s'ha fet ús d'IA en la redacció del text si s'utilitza en algun tipus de publicació.

6. Llenguatges de programació

La Intel·ligència Artificial Generativa té la capacitat d'interpretar i generar codi basat en instruccions específiques. Es pot utilitzar per generar, revisar o documentar codi de programació.

Prompt 1: crea el codi javascript per extreure els carrers d'un municipi continguts en un fitxer excel i ordenar-los per la seva longitud

Prompt 2: avalua el següent codi:

```
fileInput = document.getElementById('fileInput');
```

Aplicació dels principis:

- Principi 1: revisió del text/codi generat
 - Comprovar que el text/codi s'ajusta al que s'ha demanat a la IA
 - Entendre el que fa el codi abans d'executar-lo.
- Principi 3: es desconeixen les implicacions sobre drets d'autor del codi enviat a la IA per a la seva avaluació. Per tant, es podria tractar de dades sensibles en alguns casos.

7. Generació d'imatges, vídeo i àudio

La IA generativa també pot crear imatges, vídeo i àudio a partir dels prompts. Eines com DALL-E o Canva permeten la creació d'imatges i vídeos, en base a prompts molt detallats. A l'apartat de referències s'indiquen repositoris actualitzats amb recursos i eines basats en IA.

Exemples:

Prompt 1 (DALL-E/imatge): fotografia de nens de 8 anys de diferents ètnies anant a l'escola en un entorn rural

Prompt 2 (Canva/video): gent passejant per una zona verda en una ciutat mitjana

Prompt 3 (DALL-E/imatge): logo de piscina que mostri un trampolí

En general, la creació d'imatges o vídeos segueix un procés iteratiu fins trobar el prompt que obté millors resultats.

La revisió de les imatges generades ha de tenir presents el principis basics traslladant-los als nous formats, per exemple:

- Garantir que els continguts generats son respectuosos i no discriminatoris envers cap col·lectiu.
- Preservar la confidencialitat de persones, de les seves veus, d'espais privats,... i en tot allò que no es disposa de consentiment,...

En algunes organitzacions les imatges/vídeos es consideraran esborranys fins l'aprovació dels serveis encarregats de vetllar per la imatge corporativa.

Referències

ChatGPT (OpenAI)

- URL: <https://chat.openai.com/>
- Descripció: model de llenguatge generatiu de gran format desenvolupat per OpenAI.
- Idioma: Multi idioma

Bard (Google AI)

- URL: <https://bard.google.com/chat?hl=es>
- Descripció: model de llenguatge generatiu de gran format desenvolupat per Google AI.
- Idioma: Multi idioma

Perplexity

- URL: <https://www.perplexity.ai>
- Descripció: utilitza ChatGPT i afegeix un cercador d'internet.
- Idioma: Multi idioma

DALL-E (Open AI)

- URL: <https://openai.com/dall-e-3/>
- Descripció: model d'aprenentatge automàtic que pot generar imatges a partir de descripcions en text.
- Idioma Multi idioma

Chat GPT 4 i DALL-E 3 (Open AI) – Accés gratuït limitat via Bing de Microsoft

- URL: <https://www.bing.com/search?q=Bing+AI&showconv=1&FORM=hpcodx>
- Descripció: pagina de Microsoft que incorpora ChatGPT 4 i DALL-E 3 de forma integrada
- Idioma Multi idioma

Portals de serveis basats en IA

AiFindy

- URL: <https://aifindy.com/categorias>
- Idioma: Castellà

EasyWithAI

- URL: <https://easywithai.com/>
- Idioma: Anglès

Localret

<https://localret.cat>