

Creació d'un xatbot aplicat en l'institut

NAME

2023-2024

TUTOR NAME

1r - 2n Batxillerat

xatbot

“Un xatbot és una aplicació de programari o una interície web dissenyada per imitar la conversa humana mitjançant interaccions de text o de veu.”

4 EDUCACIÓN
DE CALIDAD



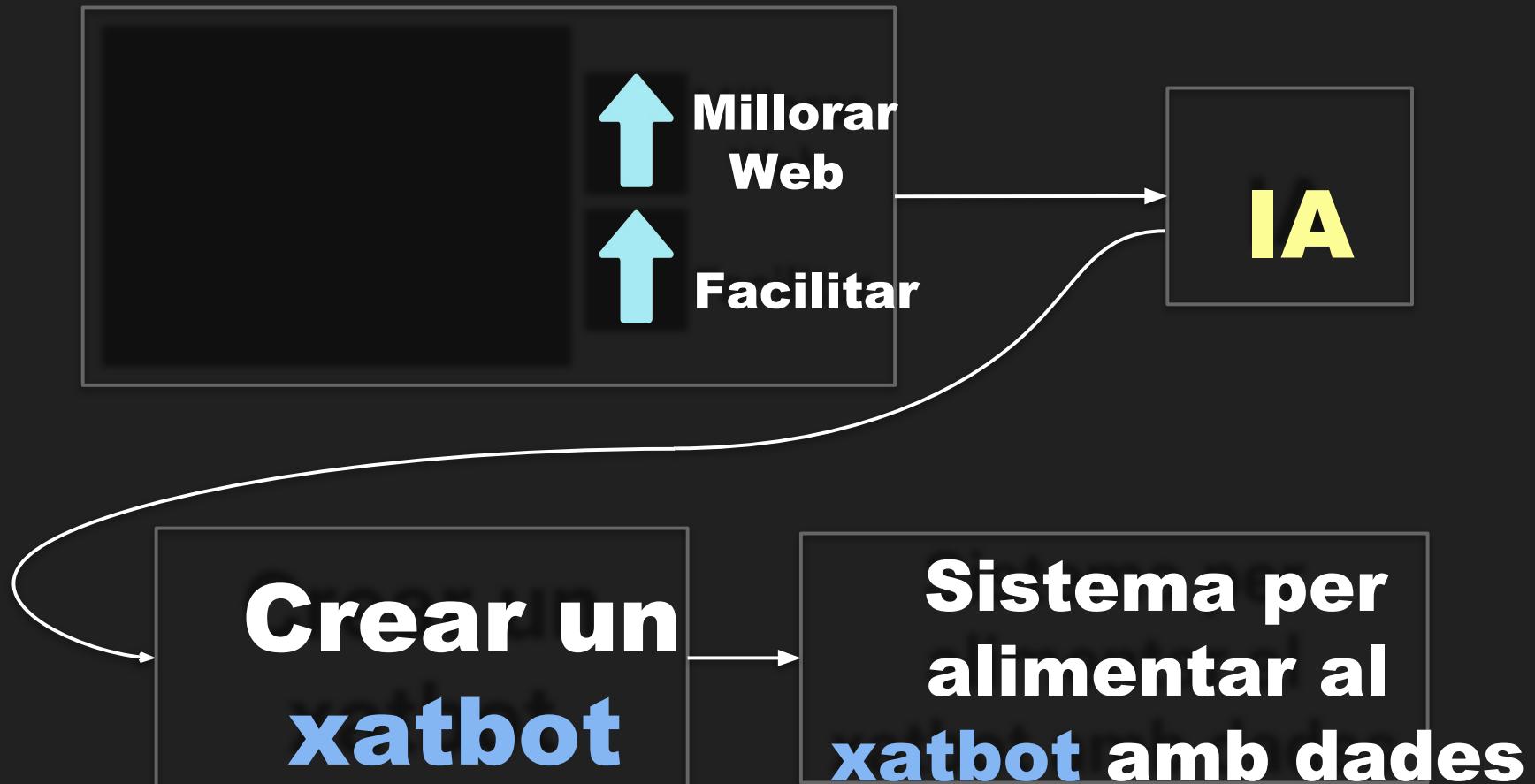
9 INDUSTRIA,
INNOVACIÓN E
INFRAESTRUCTURA



10 REDUCCIÓN DE LAS
DESIGUALDADES



Objectiu i plantejament del TR



Per què vaig decidir fer aquest?

Millorar
Web

Facilitar

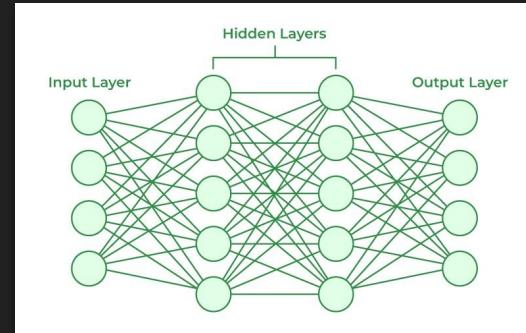
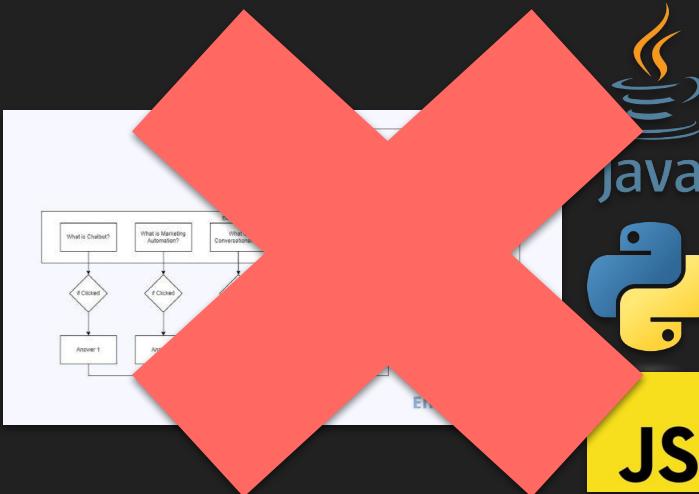
IA

L'objectiu

Millorar l'accessibilitat i l'experiència d'usuari en l'ús de la pàgina web de l'institut.

Crear un xatbot

Sistema que faré servir

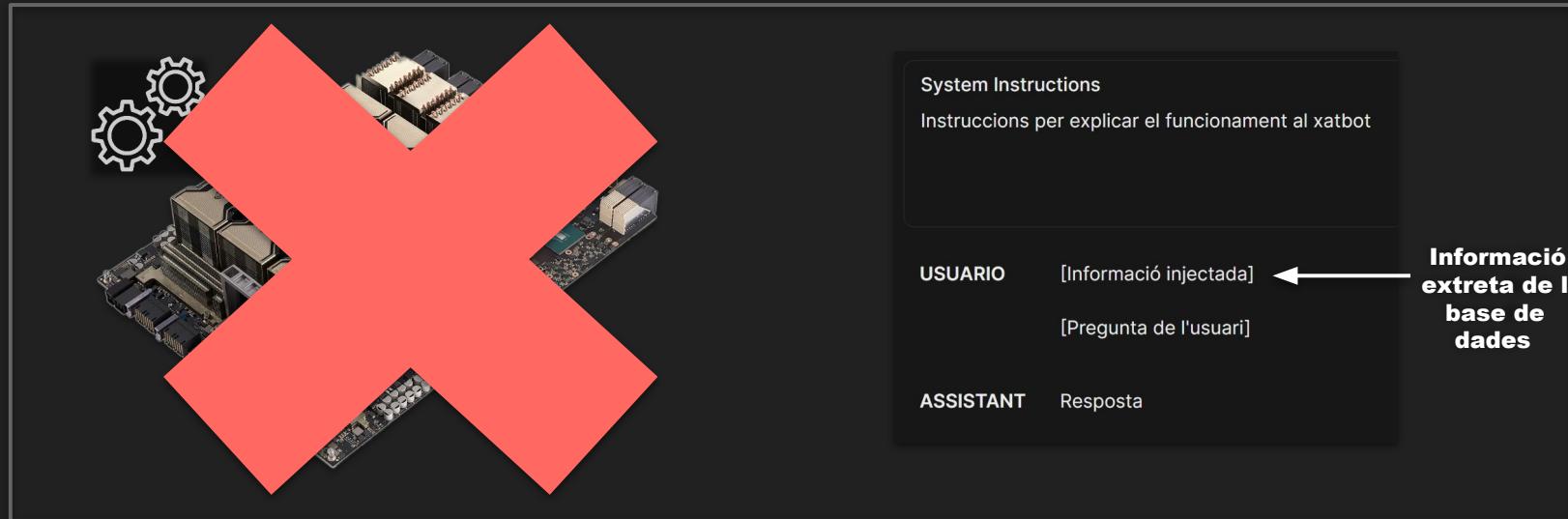


?

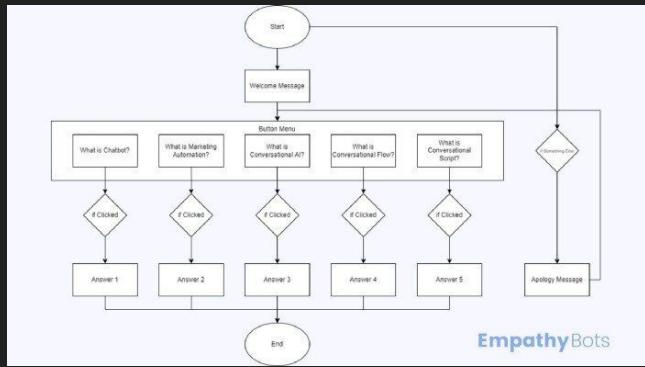
No tenen informació

Sistema per
alimentar al **xatbot**
amb dades

Com aconsegueixo que en tinguin informació



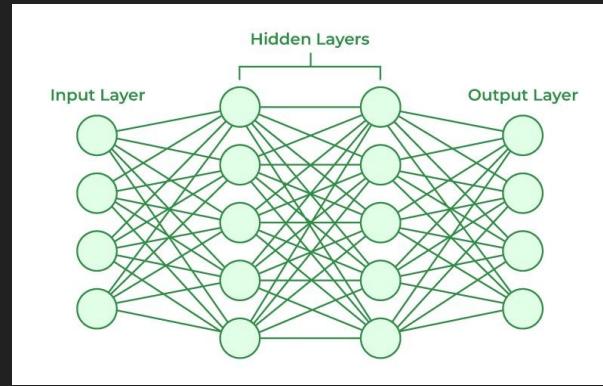
Com ho faré?



Java



JS



Problema:
ChatGPT, Claude, Gemini i Donar-li la informació
altres: No tenen informació directament

Solució:

No viable:

- Entrenar tot un model des de Falta de dades i de poder computacional.
- Donar tota la informació com a entrada al ~~xatbot~~ El model no pot rebre tanta informació com a entrada.

Solució:

Sistema de cerca (RAG)

→ Realitzar cerques especialitzades per extreure de tota pàgina web només la informació rellevant.



Amb codi →

MarkDown

Text gran

Sub apartat

Text text text...

```
<h1>Text gran</h1>
<h3>Sub apartat</h3>
<p>Text text text...</p>
```

Text gran

Sub apartat

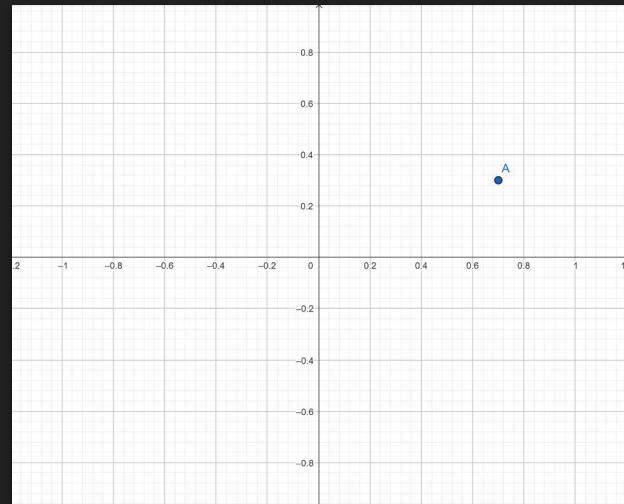
Text text text

```
# Text gran
### Sub apartat
Text text text
```

Text → Embedding

```
# Text gran  
### Sub apartat  
Text text text
```

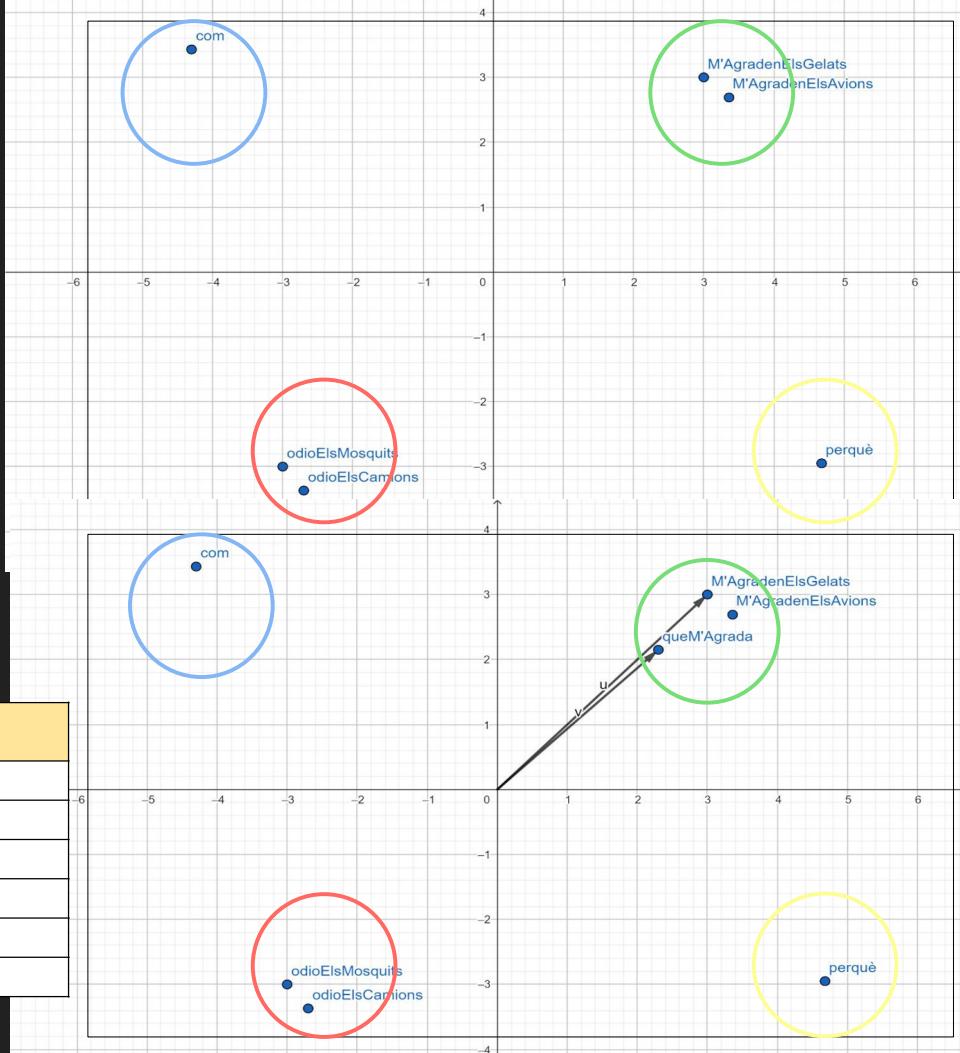
→ [0.7, 0.3, 0.5, 0.1, ...]

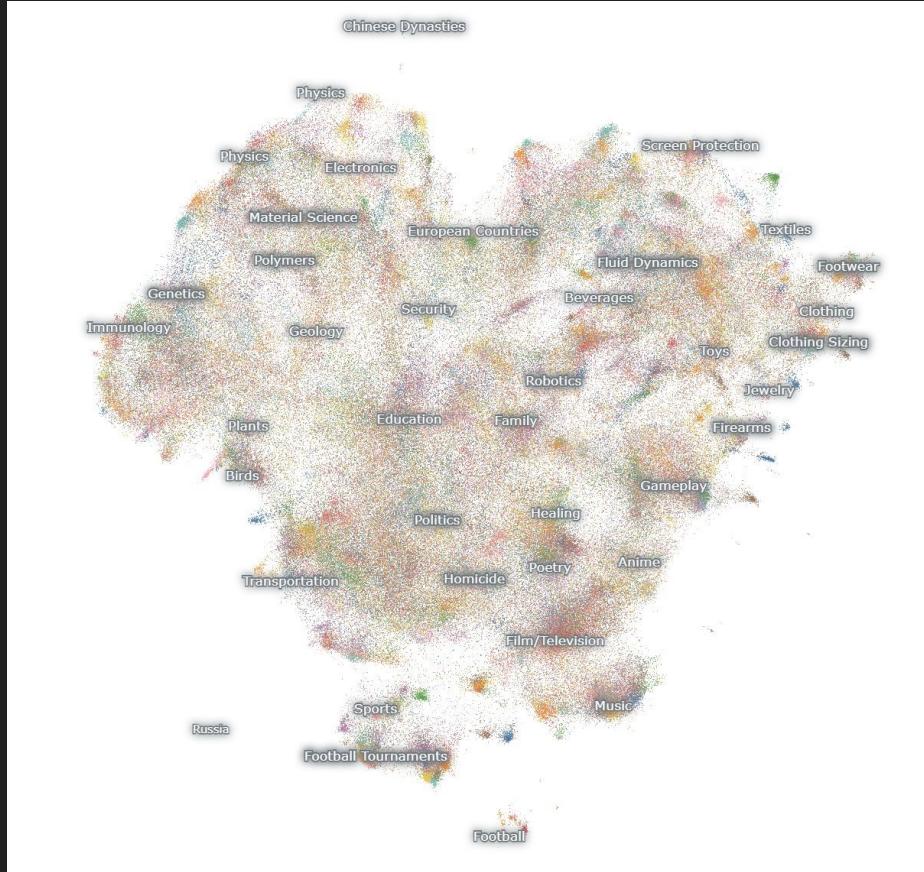


- **M'agraden els gelats**
- **M'agraden els avions**
- **Odio els camions**
- **Odio els mosquits**
- **com**
- **perquè**

Què m'agrada?

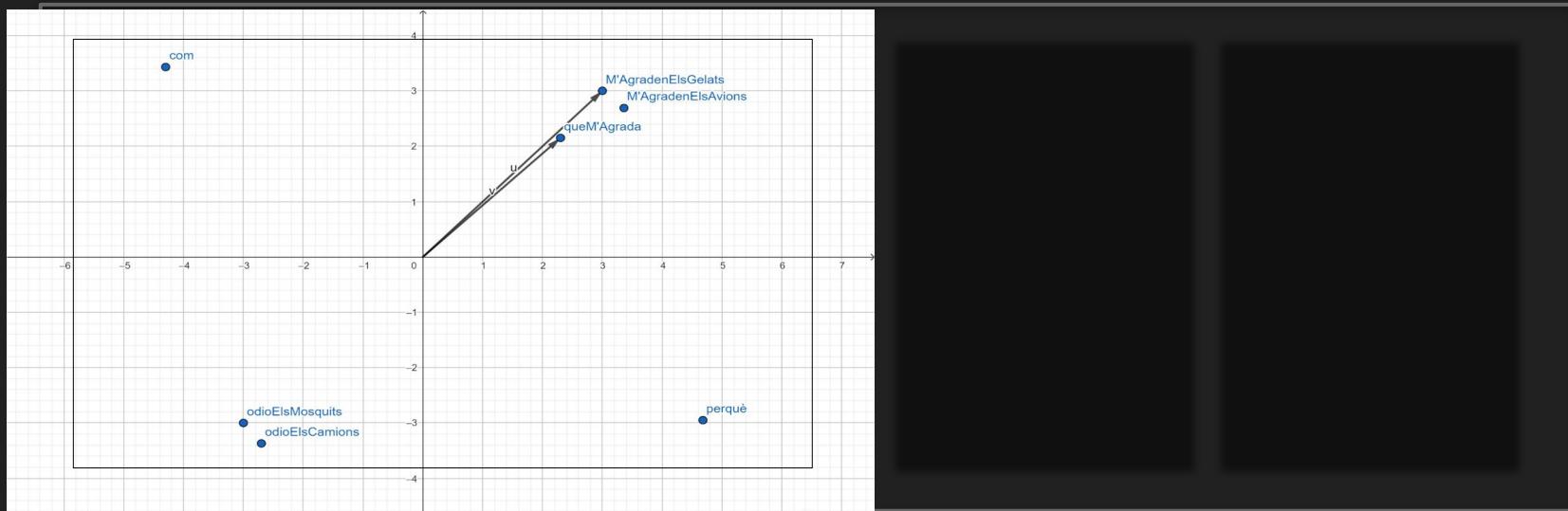
Vector	(s) Semblança respecte a v
M'agraden els gelats	0.989274
M'agraden els avions	0.975617
Perquè	0.581697
Com	0.453599
Odio els camions	0.045718
Odio els mosquits	0.010725





<https://atlas.nomic.ai/map/nomic-text-embed-v1-5m-sample>

Base de dades de la pàgina de l'institut



chatbot-tr-2024-app Private

The front-end of the TR

JavaScript ⭐ 1 Updated 3 weeks ago

Starred



chatbot-tr-2024-api Private

JavaScript ⭐ 1 MIT License Updated 3 weeks ago

Starred



chatbot-tr-2024-rag Private

JavaScript ⭐ 1 MIT License Updated last month

Starred

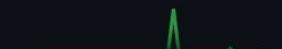


chatbot-tr-2024-webdataextract Private

Extract the data to use in the RAG

JavaScript ⭐ 1 MIT License Updated last month

Starred



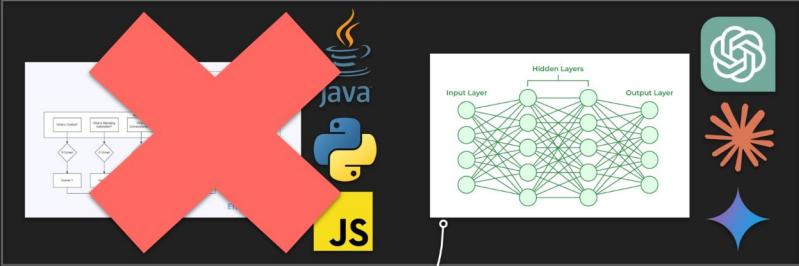
~2222 Línes de codi

Punts de millora:

- **Embedding** Millorar la precisió de cerca en més idiomes.
- **RAG** Millorar la precisió de les dades.
- **Finestra de context** Realitzar converses més llargues.
- **Fer servir un model més gran** Millor ús de la informació.
- **Imatges, Canva, PowerPoint** Té tota la informació.

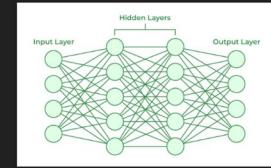
☆☆ <http://PRIVATE SERVER:8080/>

Sistema que faré servir



?

No tenen informació



Sistema per alimentar al xatbot amb dades

