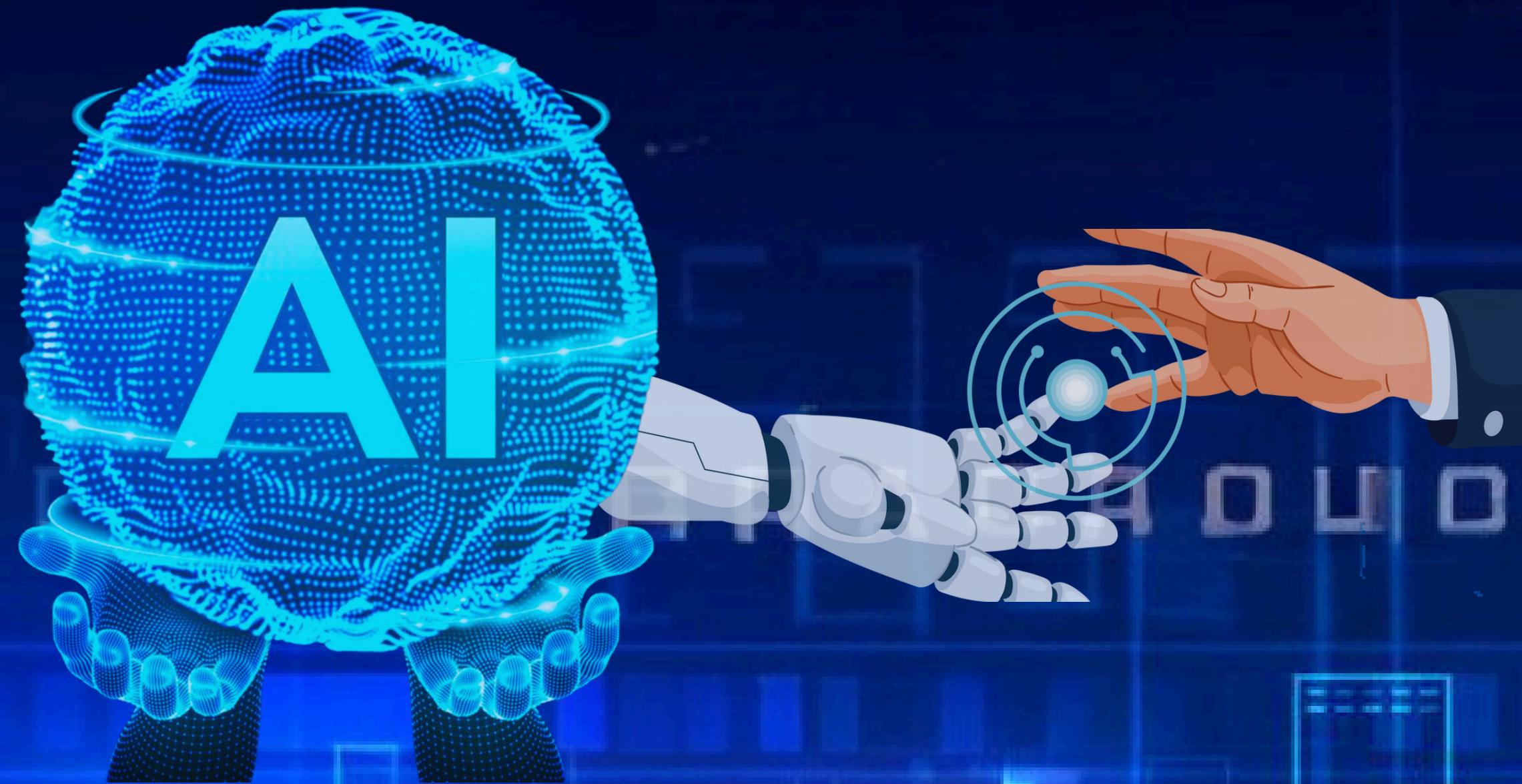


HACKATON



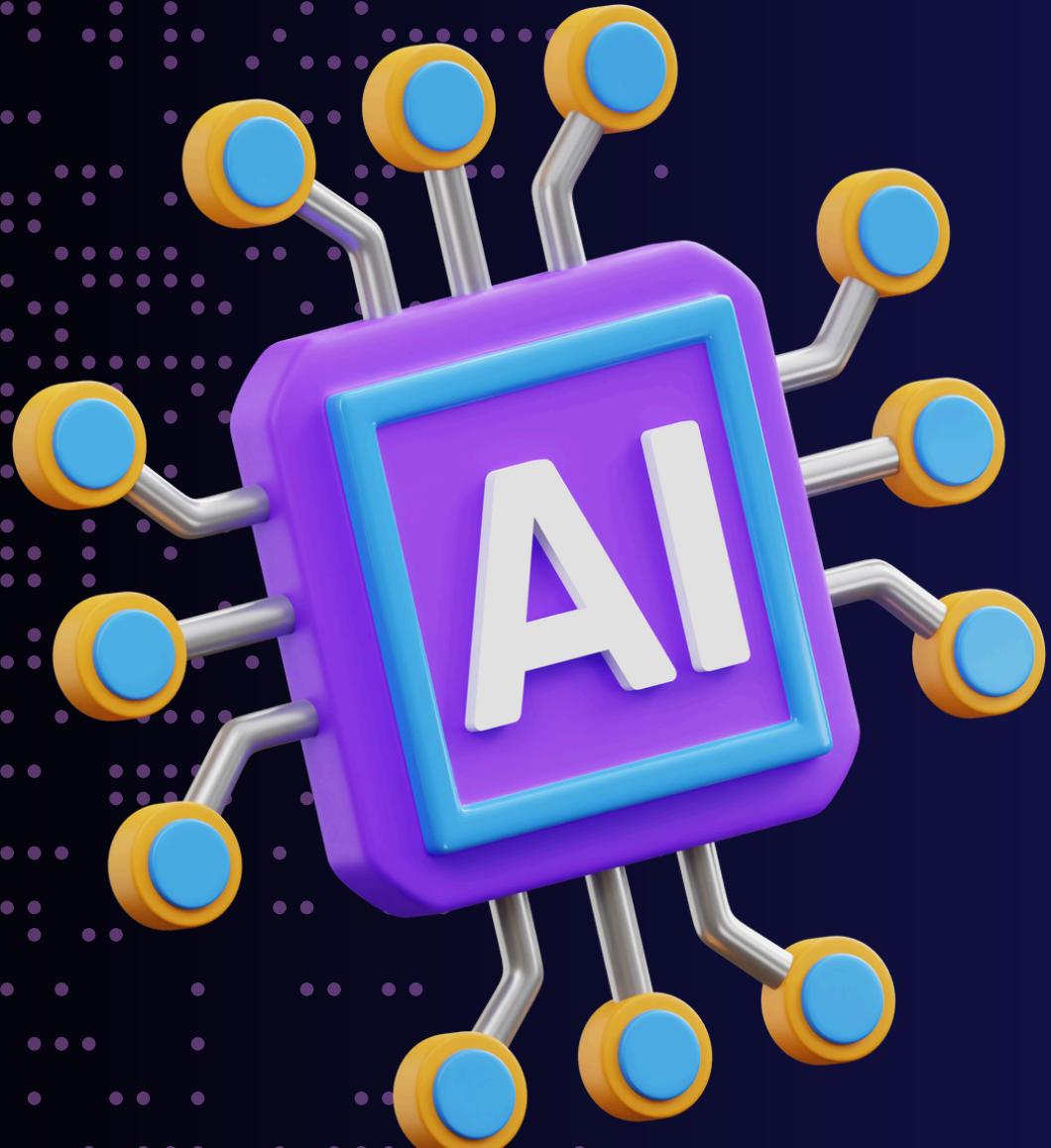
Oussama, Sophie, Nicolas



**MINISTÈRE
DES ARMÉES**

*Liberté
Égalité
Fraternité*

SOMMAIRE



- ① Notre Objectif
- ② Les différentes étapes
- ③ Présentation de notre Dataset

OBJECTIF

HACKATON



JOURNAL DE BORD SUR 3 JOURS

23 - 25 MAI 2024

Matin

Après-midi

CE QUE ON A FAIT AVANT-HIER

- Annothaton

CE QUE ON A FAIT HIER

- Choix des différents sites à scrapper
- Scrapping

CE QUE ON A FAIT AUJOURD'HUI

- Journal de bord / Présentation
- Création du fichier csv

- Répartition des tâches
- Choix des mots clefs à rechercher ("institution nationale", "FAQ ville de...")

- Scrapping
- Mise en commun du scrapping
- Structuration du Dataset

RESSOURCES UTILISÉES



- ① Request
- ② Re
- ③ Beautiful Soup
- ④ Pandas

RECHERCHES CLÉS



- ① Institut national FAQ
- ② FAQ ville de...
- ③ Questions/Réponses sur le droit...
- ④ Questions/Réponses culture générale

PROBLÈMES RENCONTRÉS



- ① Différences codes sources
- ② Problèmes d'accès
- ③ Duplications



SOURCES FINALES

- ① Site ville de Vincennes
- ② Site renseignement Droit
- ③ Site de culture générale pour enfant

DATA5

	Question	Réponse
0	Comment signaler une personne âgée isolée ?	Les personnes âgées isolées peuvent être inscr...
1	Comment se faire livrer des repas ?	Des associations peuvent vous aider en dispens...
2	Comment effectuer une demande de Pass nigo s...	Les démarches pour obtenir un Pass nigo seni...
3	Comment effectuer une demande de logement soci...	Il convient de déposer un dossier de demande d...
4	Comment effectuer une demande d'aide pour le m...	Depuis plusieurs années, Vincennes s'est donné...
...
1986	Qui sont les pompiers ?	Pour certains, c'est leur métier d'être pompiers
1987	Qu'est-ce qu'un ornithorynque ?	L'ornithorynque est un mammifère amphibia (qui...
1988	Vos questions : « Pourquoi l'escargot à ...	Eh oui
1989	Pourquoi les zèbres ont-ils des rayures ?	Certains scientifiques ont dit que les zèbres ...
1990	Où avaient lieu les jeux olympiques ?	Ils se déroulaient à Olympie, en l'honneur de ...
1991 rows x 2 columns		