

# Speech to text

Une opportunité pour améliorer  
notre productivité?

Mathieu BUSQUET

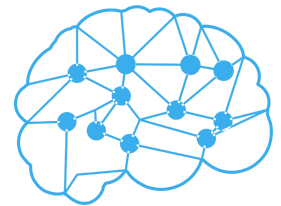


**Le Camping des Speakers**



**Le Camping des Speakers**

# Mathieu BUSQUET



Alternant Machine Learning Engineer



AI Solutions Team



@mat\_bsqt



MathieuBsqt



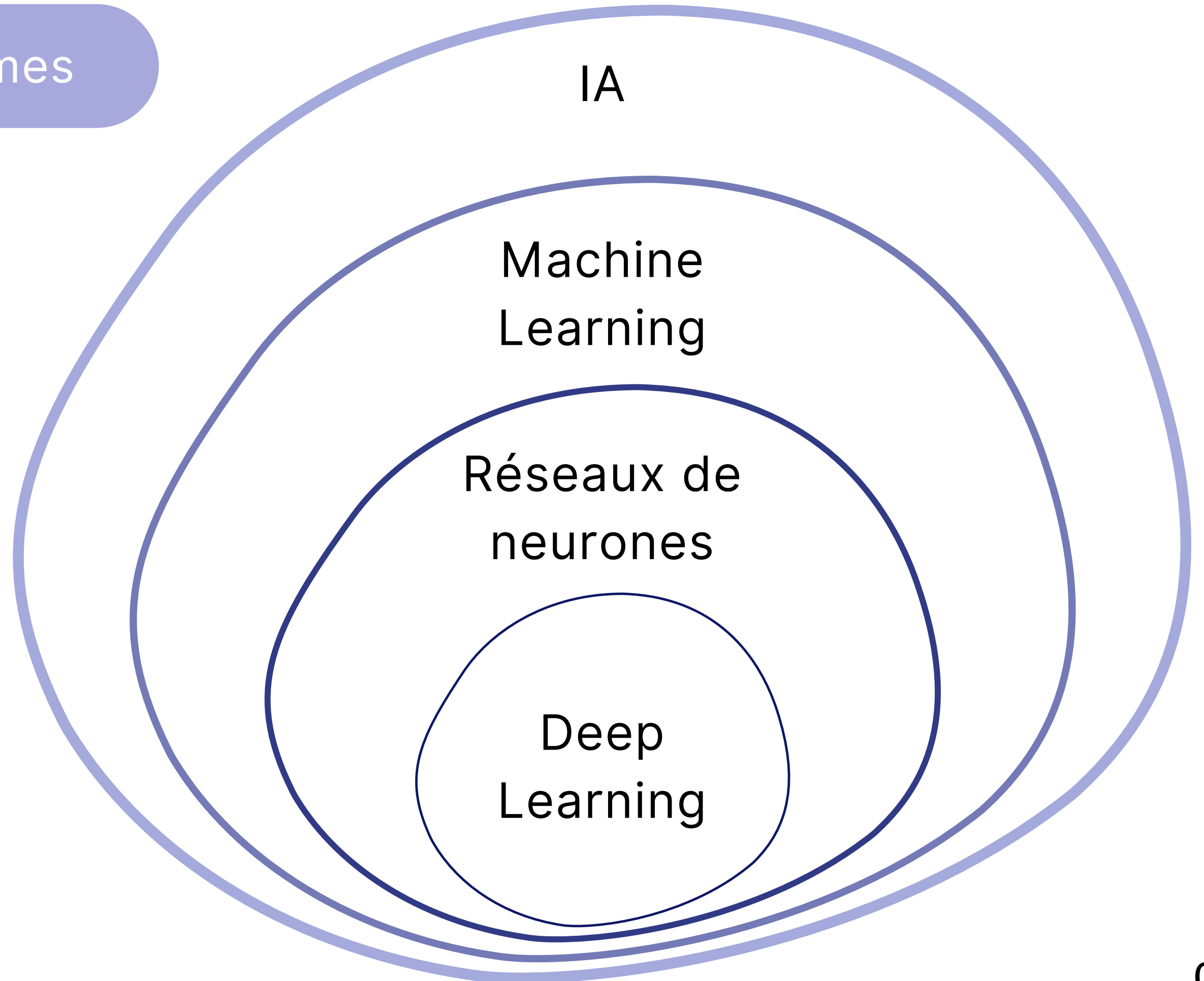
Mathieu BUSQUET

## Ce que nous allons voir aujourd'hui

- Rappeler ce qu'est l'IA
- Présenter le Speech to Text
- Découvrir comment le mettre en place facilement
- Comprendre comment on peut créer une application complète en ajoutant des fonctionnalités
- Comprendre comment le speech to text améliore notre productivité
- Montrer l'intérêt d'utiliser un GPU



# La différence entre ces termes



## Des exemples de tâches populaires



Des exemples de tâches populaires



Vélo    Voiture


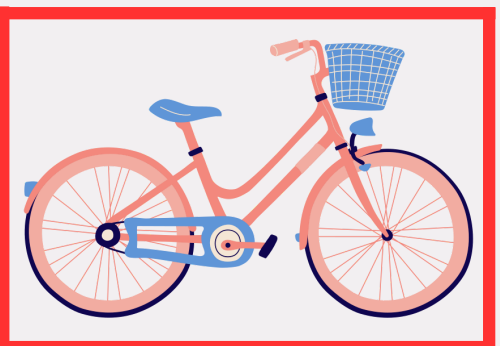
Personne

A diagram showing three items: a bicycle, a red car, and a black silhouette of a person. Each item is enclosed in a red rectangular box. The bicycle and car are positioned side-by-side at the top, and the person silhouette is centered below them. The labels 'Vélo', 'Voiture', and 'Personne' are placed above and below their respective boxes.


Des exemples de tâches populaires



Vélo      Voiture

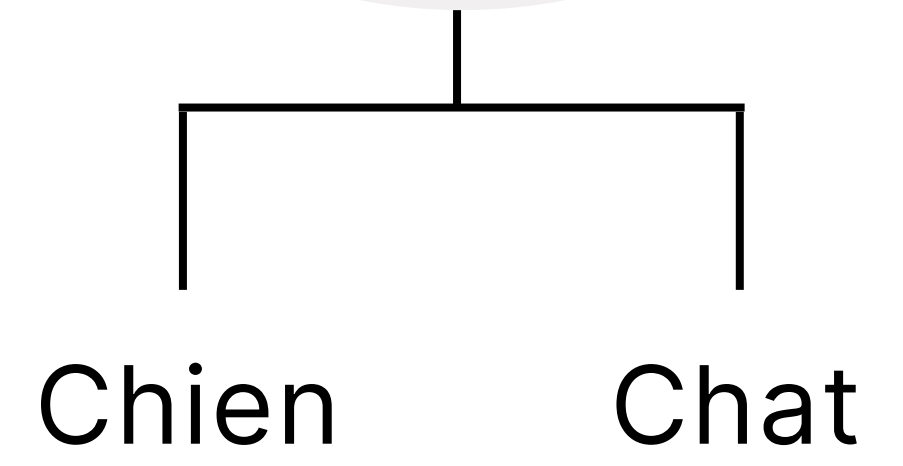
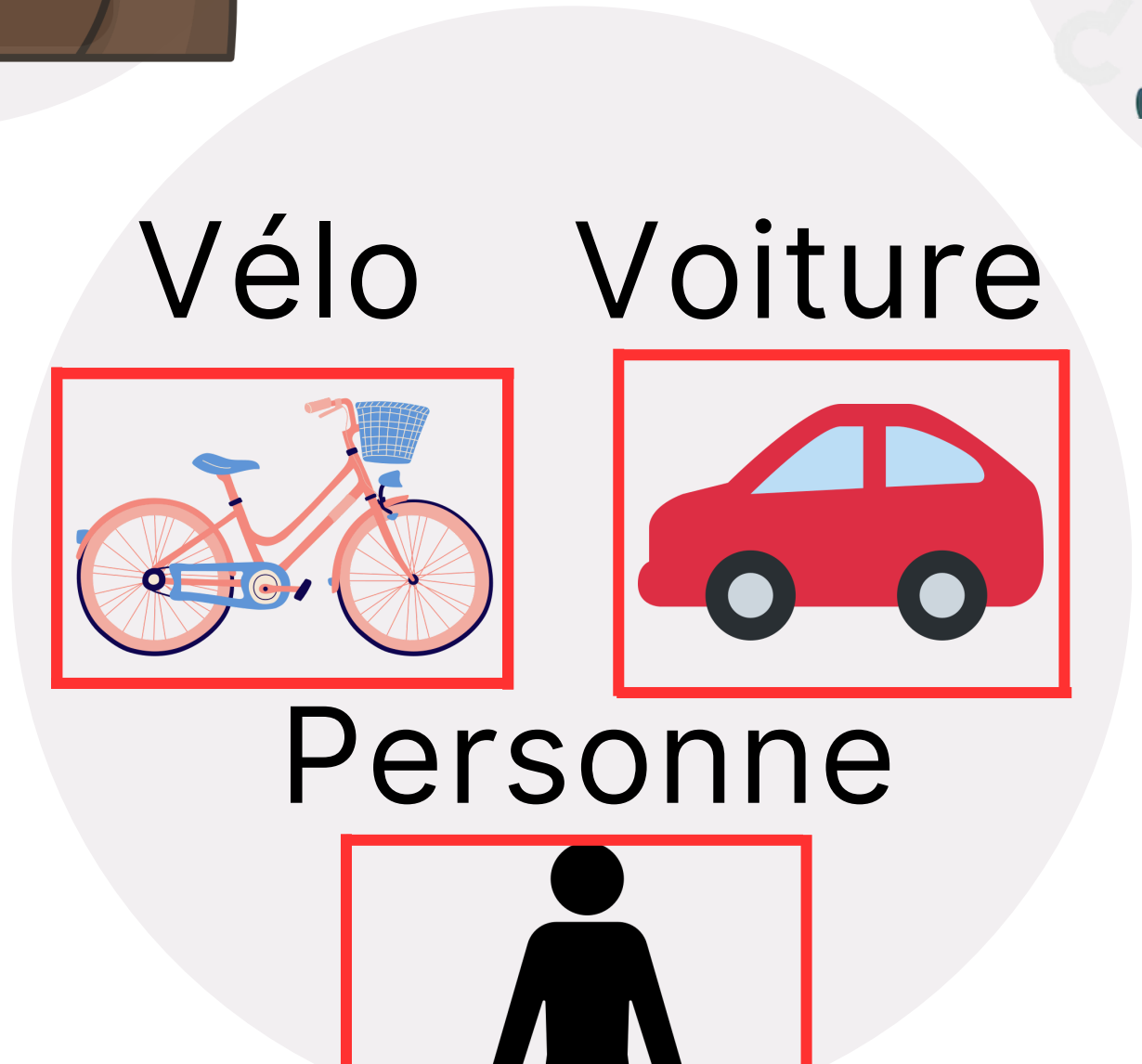
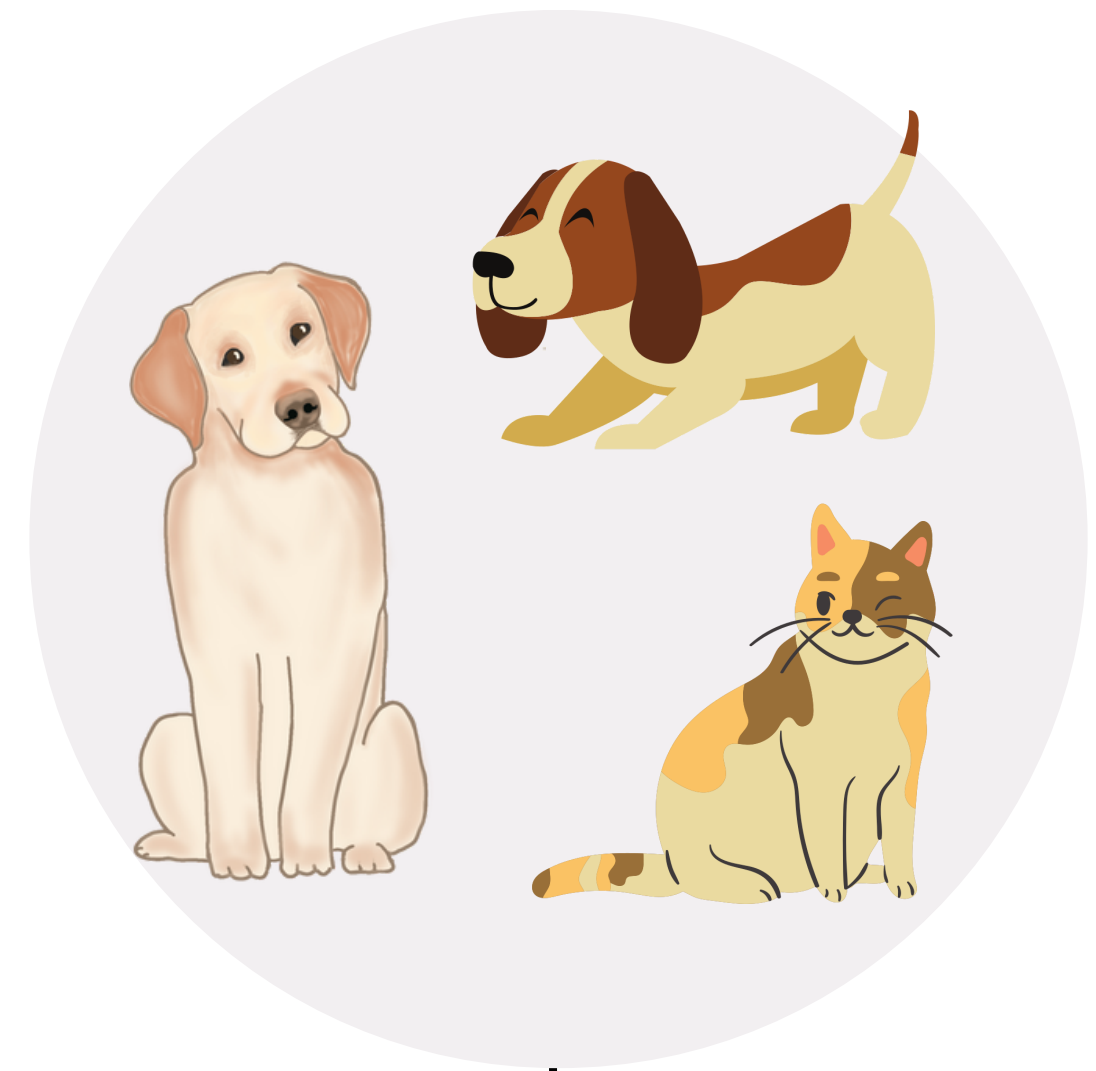


Personne



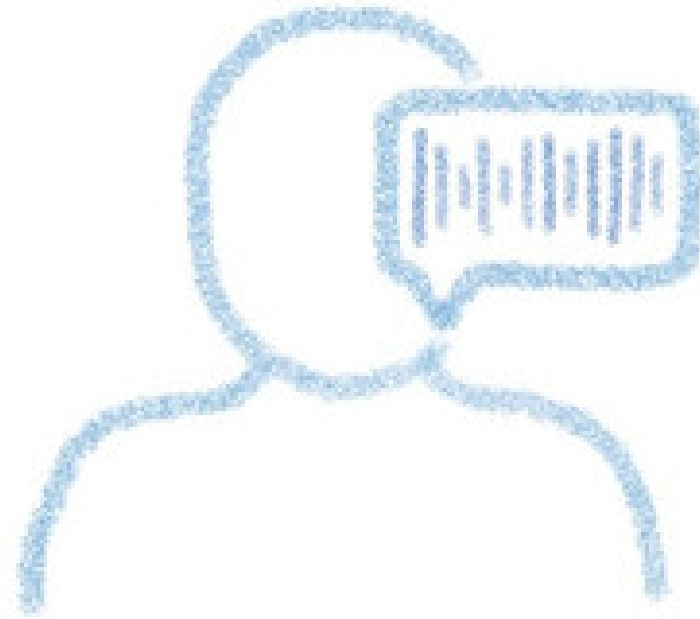


Des exemples de tâches populaires

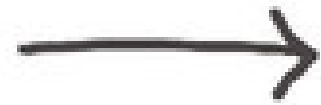




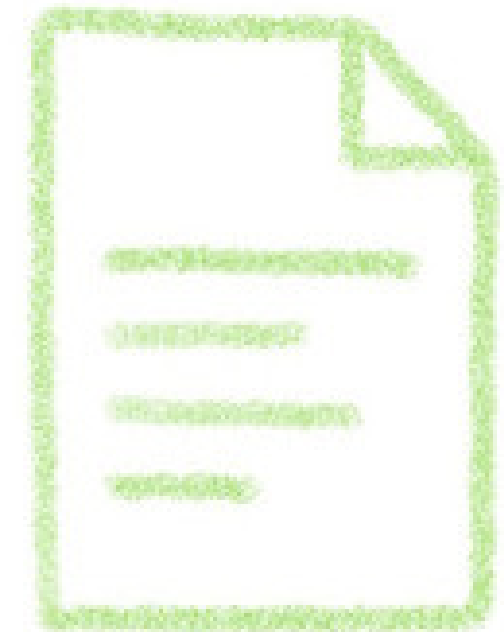
## Speech to text



Speech



Modèle IA



Transcription

## Cas d'usage



### Transcrire des enregistrements audio

interviews, réunions, conférences, discours, directs

Gain de temps



### Dictée vocale

Augmenter la productivité et l'efficacité d'un travail

Personnes ayant des difficultés à utiliser un clavier



### Service client

Améliorer la qualité du service en analysant les transcriptions des appels reçus

## Mise en place



APIs

## Mise en place



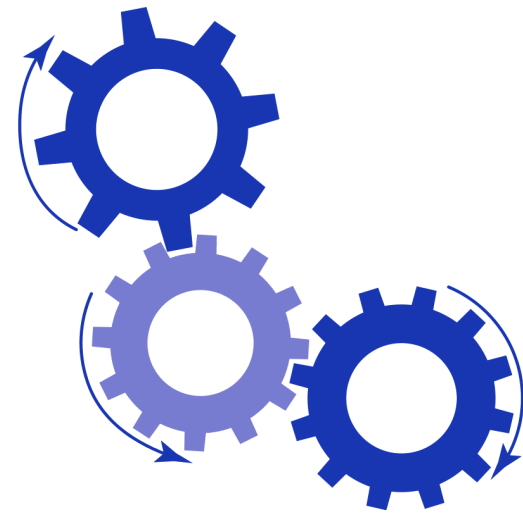
APIs

Confidentialité des données

## Mise en place



APIs



Construire et entraîner  
son modèle

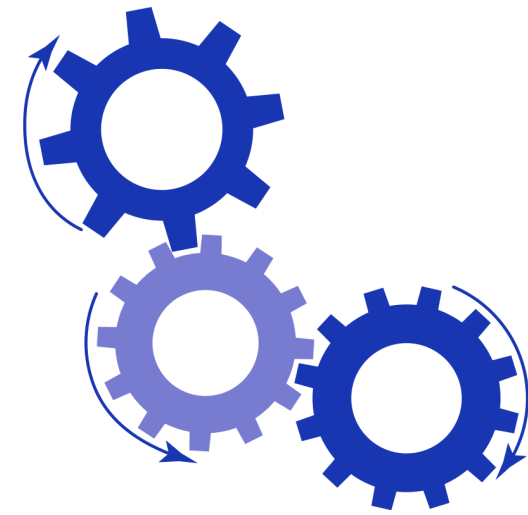
Confidentialité des données

## Mise en place



APIs

Confidentialité des données



**Construire et entraîner  
son modèle**

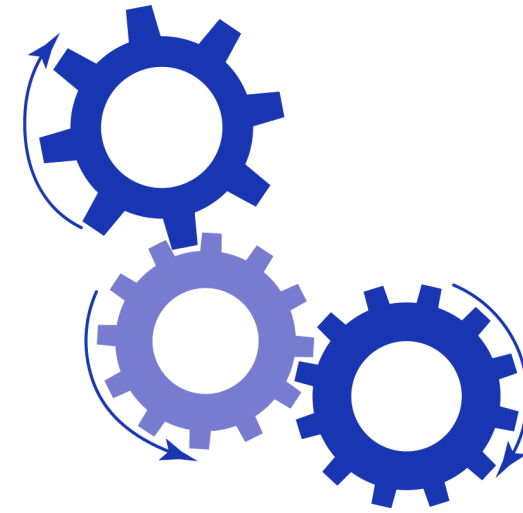
Données, temps, ressources,  
impact carbone

## Mise en place



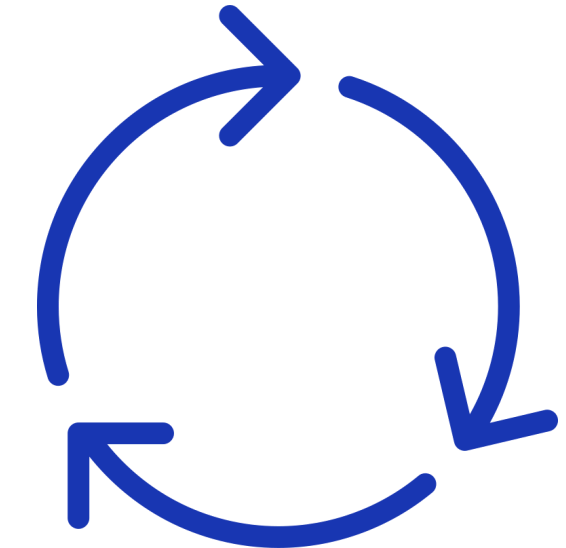
APIs

Confidentialité des données



Construire et entraîner  
son modèle

Données, temps, ressources,  
impact carbone



Utiliser un modèle  
pré-entraîné

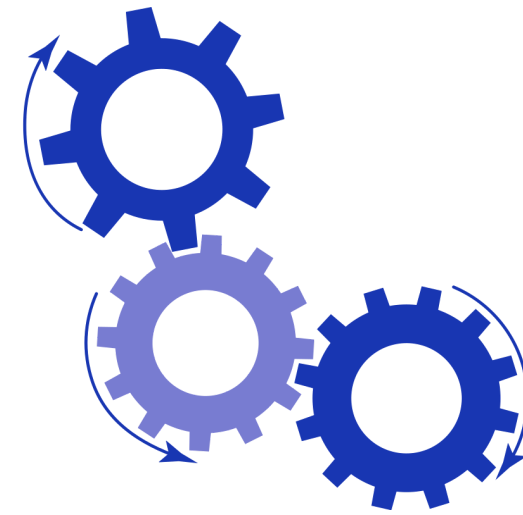


## Mise en place



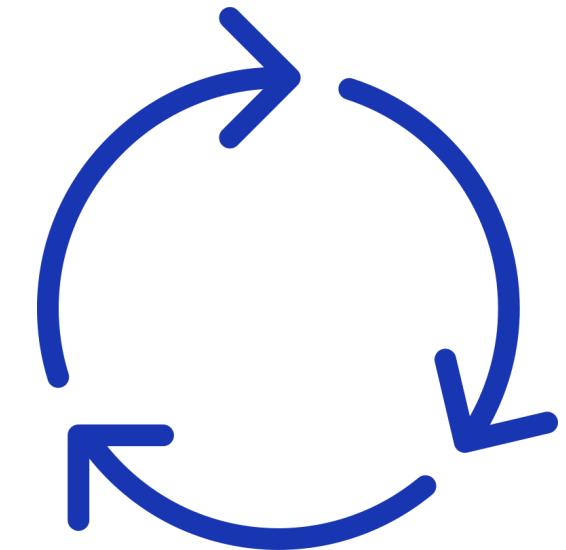
**APIs**

Confidentialité des données



**Construire et entraîner  
son modèle**

Données, temps, ressources,  
impact carbone



**Utiliser un modèle  
pré-entraîné**

Économies, gain de temps et  
accessibilité

# Transcrire un fichier audio

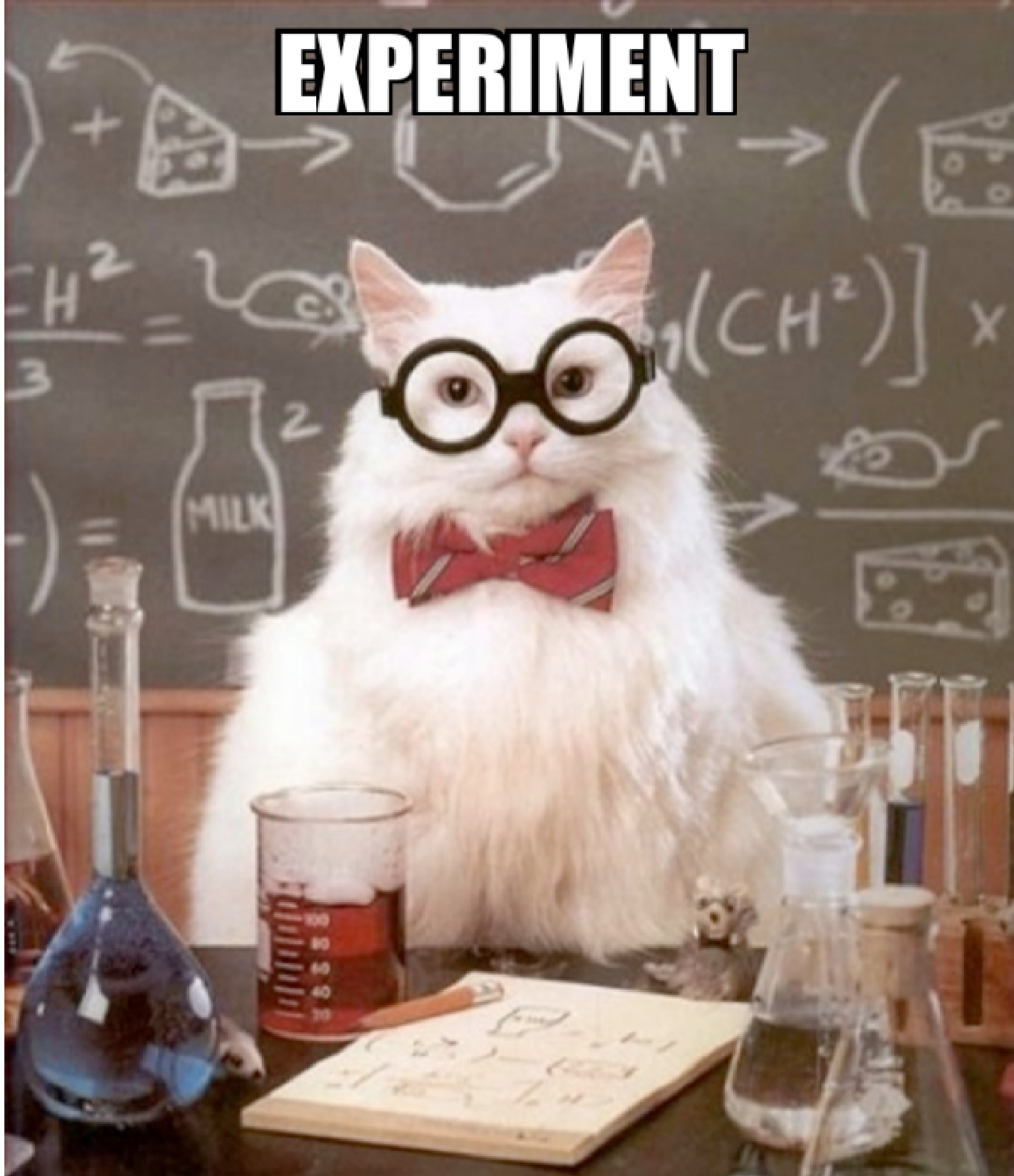


**HUGGING FACE**

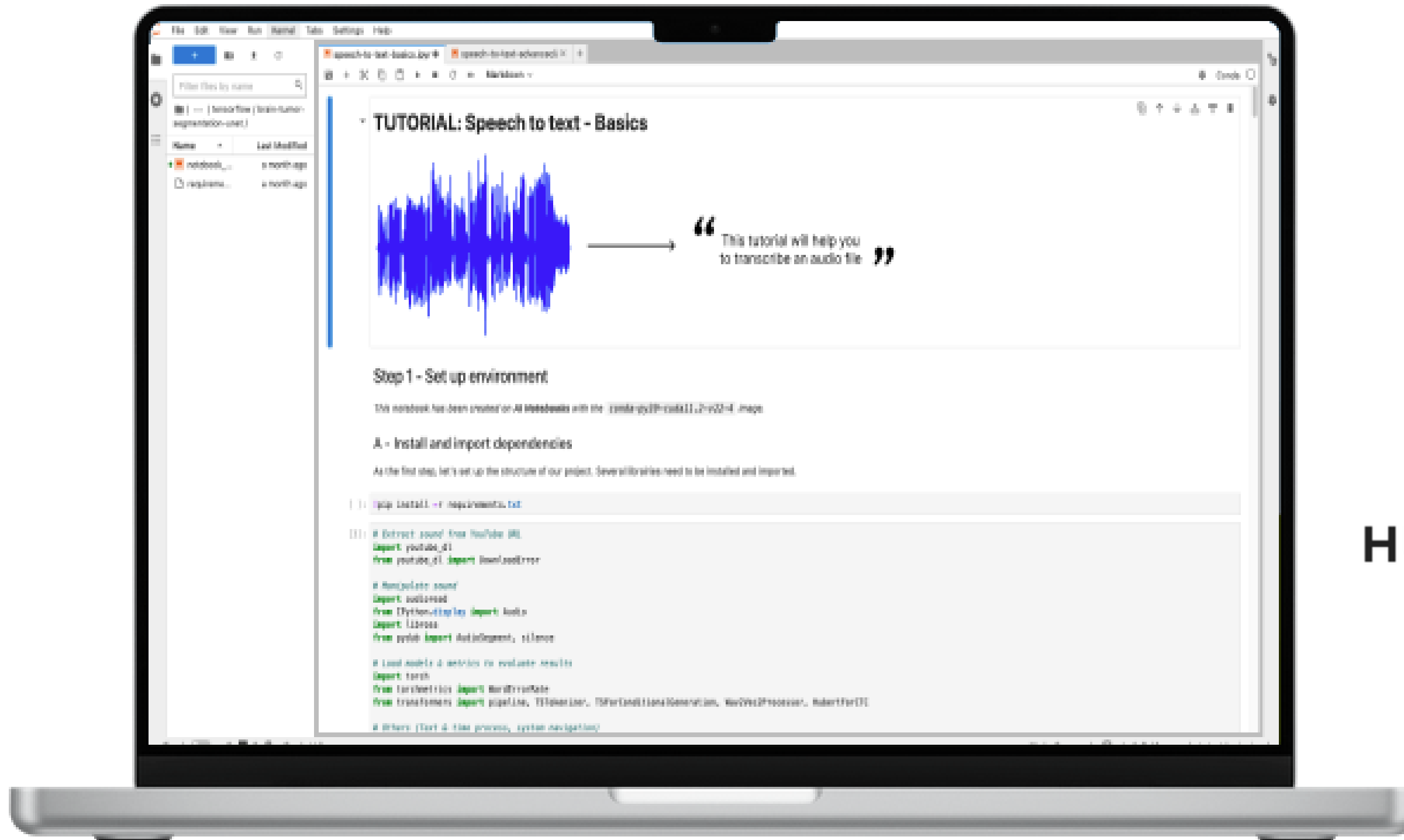
*Transformers*

Classification d'images  
Traduction de texte  
Détection d'objets  
**Transcription audio**  
Réponses automatiques  
Classification de sons  
Texte vers image

# EXPERIMENT

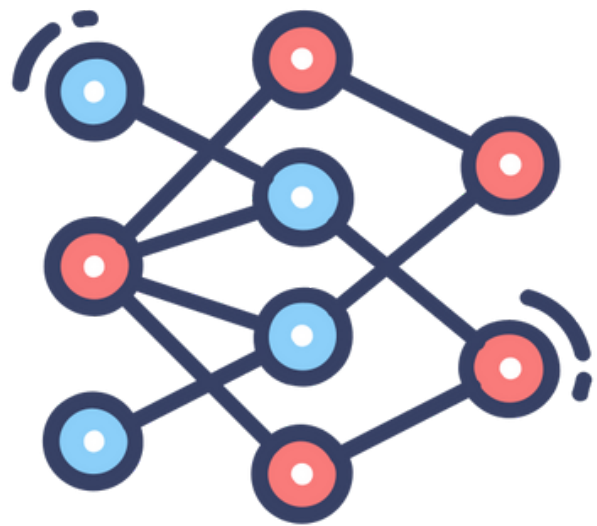


# Un Notebook pour effectuer mes tests



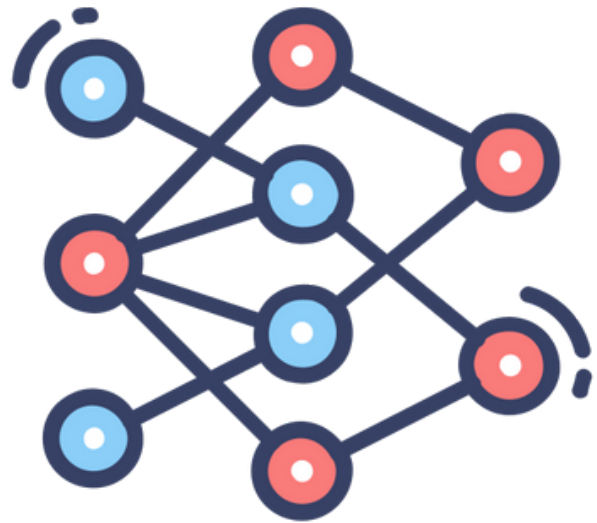
**HUGGING FACE**

# Création d'une application

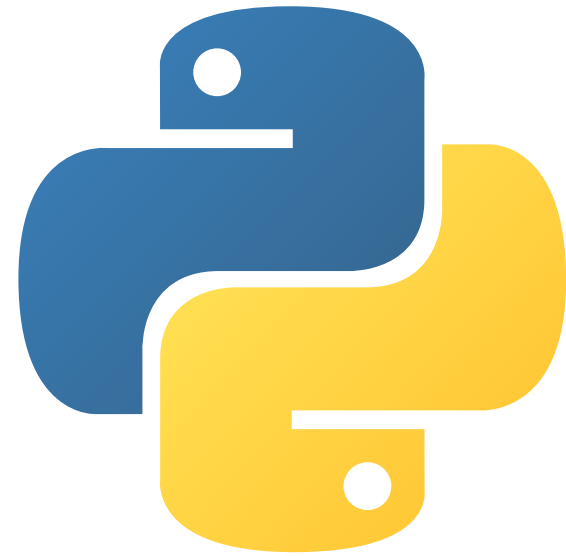


**Modèles d'IA**

# Création d'une application

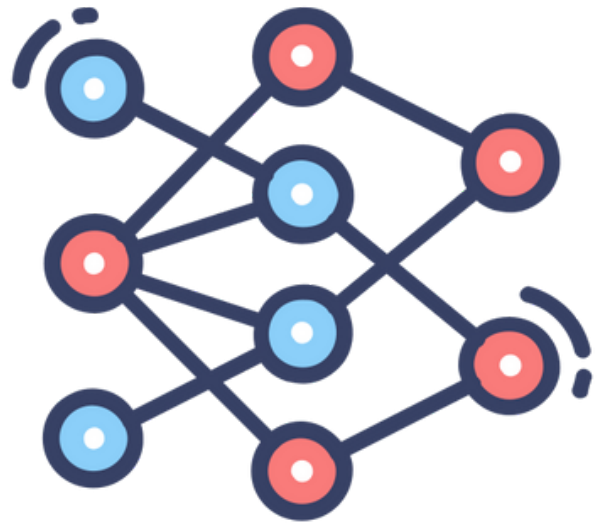


**Modèles d'IA**

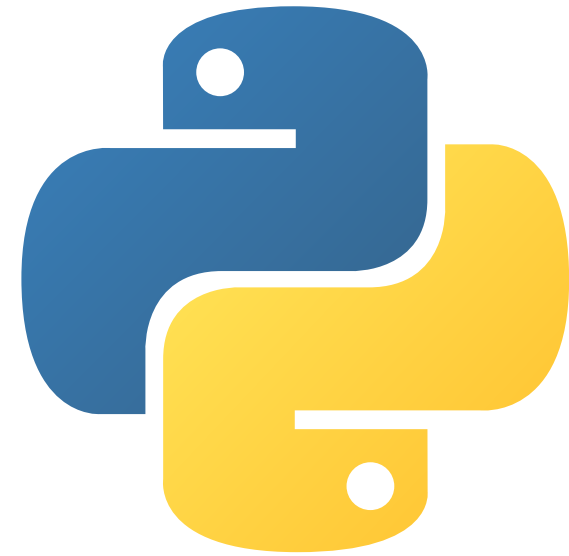


**Codes**

# Création d'une application



**Modèles d'IA**



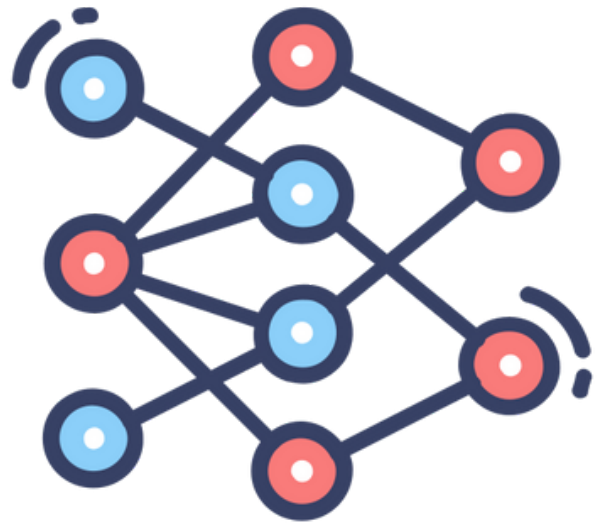
**Codes**



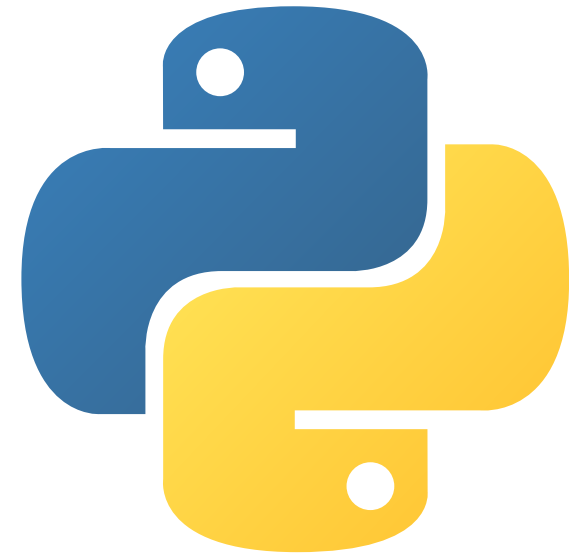
**Web app Framework**



# Création d'une application



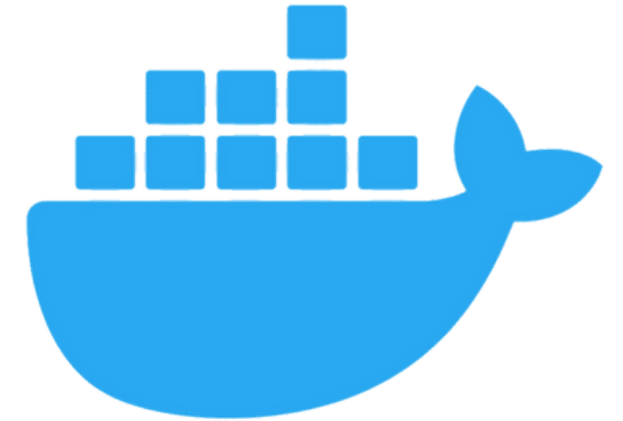
**Modèles d'IA**



**Codes**

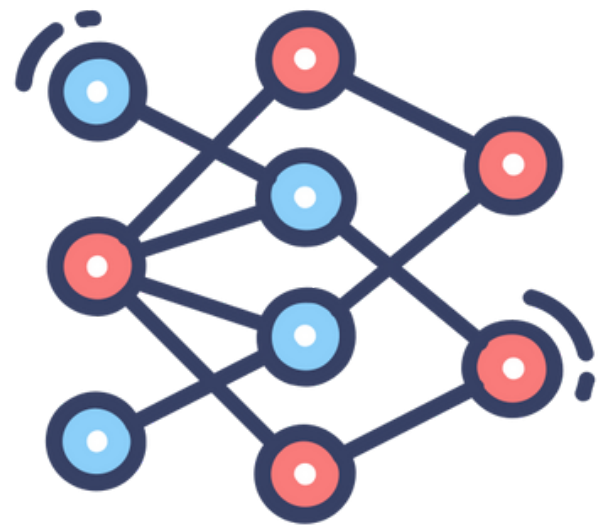


**Web app Framework**



**Dockerfile**

# Création d'une application



**Modèles d'IA**



**Codes**

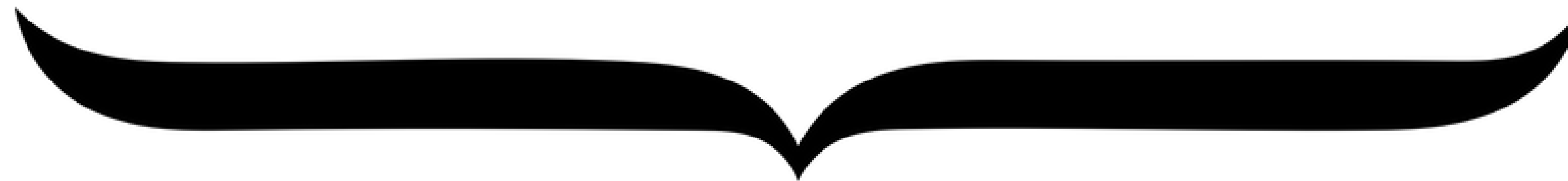


**Streamlit**

**Web app Framework**



**Dockerfile**



**Application déployable  
chez un Cloud Provider**

**AI Deploy**  
 **OVHcloud**

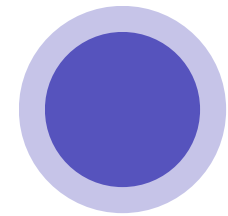


**SO...**

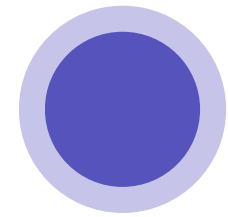
**WHAT DID WE LEARN?**

[meme-generator.net](http://meme-generator.net)

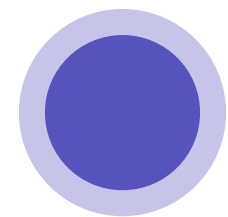
## Qu'avons nous appris ?



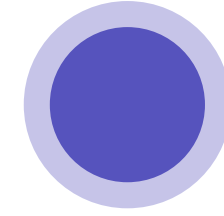
Notions d'IA



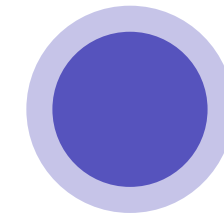
L'IA est accessible à toutes et tous



L'intérêt d'utiliser un GPU

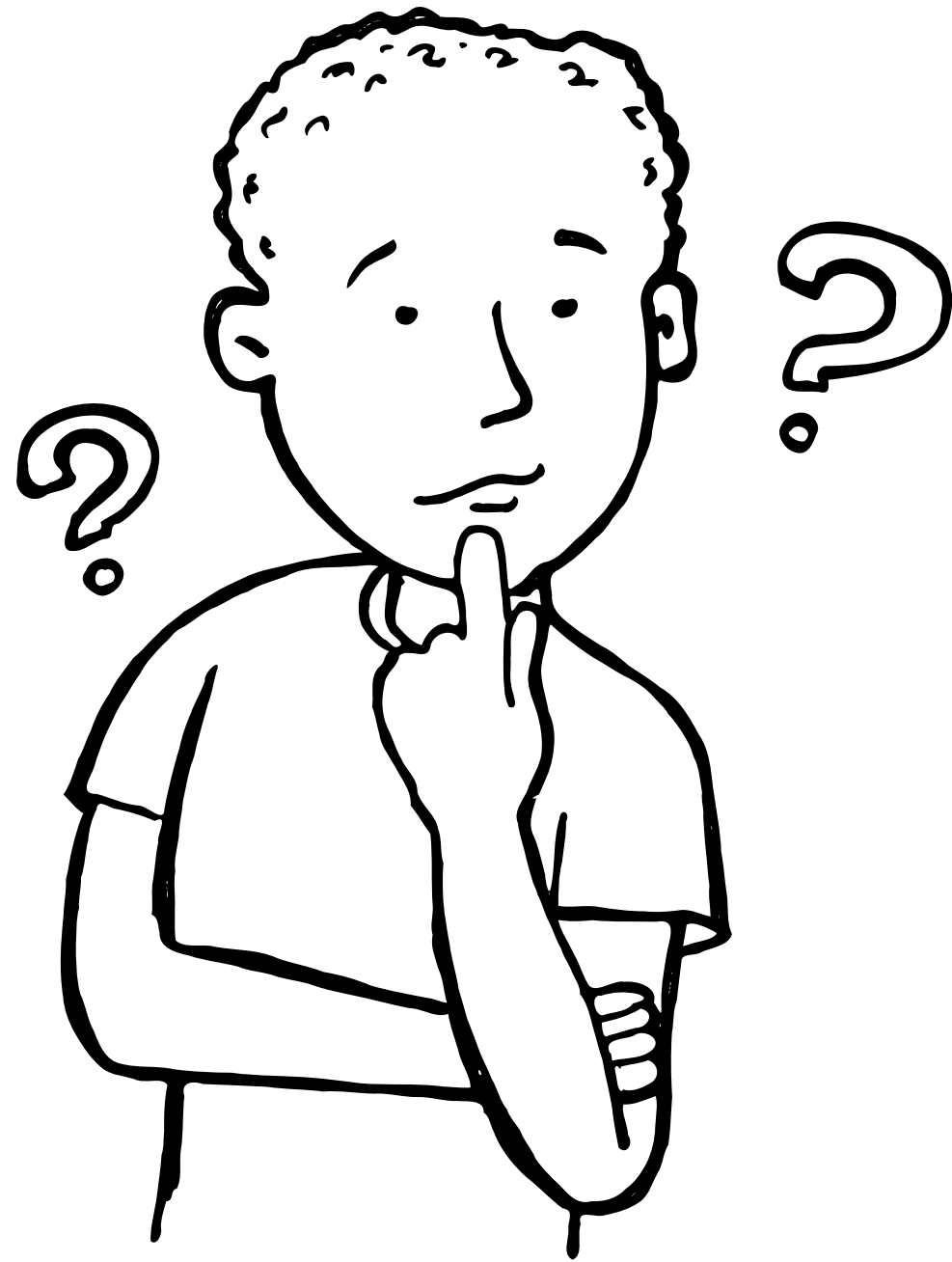


Le Speech to Text



Plusieurs modèles peuvent être mis en relation

# DES QUESTIONS ?



Slides

<https://ovh.to/ECDx57Q>



@mat\_bsqt



MathieuBsqt



Mathieu BUSQUET

# References

**Repo GitHub d'exemples:** <https://github.com/ovh/ai-training-examples/>

**Notebooks speech to text:** <https://github.com/ovh/ai-training-examples/tree/main/notebooks/natural-language-processing/speech-to-text/miniconda/>

**L'app de speech to text:** <https://github.com/ovh/ai-training-examples/tree/main/apps/streamlit/speech-to-text>

**Les docs pour utiliser les notebooks et l'app:** <https://docs.ovh.com/gb/en/publiccloud/ai/>

**Les articles de blogs:** <https://blog.ovhcloud.com/how-to-build-a-speech-to-text-application-with-python-1-3/>

**Streamlit:** <https://streamlit.io/>

**pyannote.audio:** <https://github.com/pyannote/pyannote-audio>