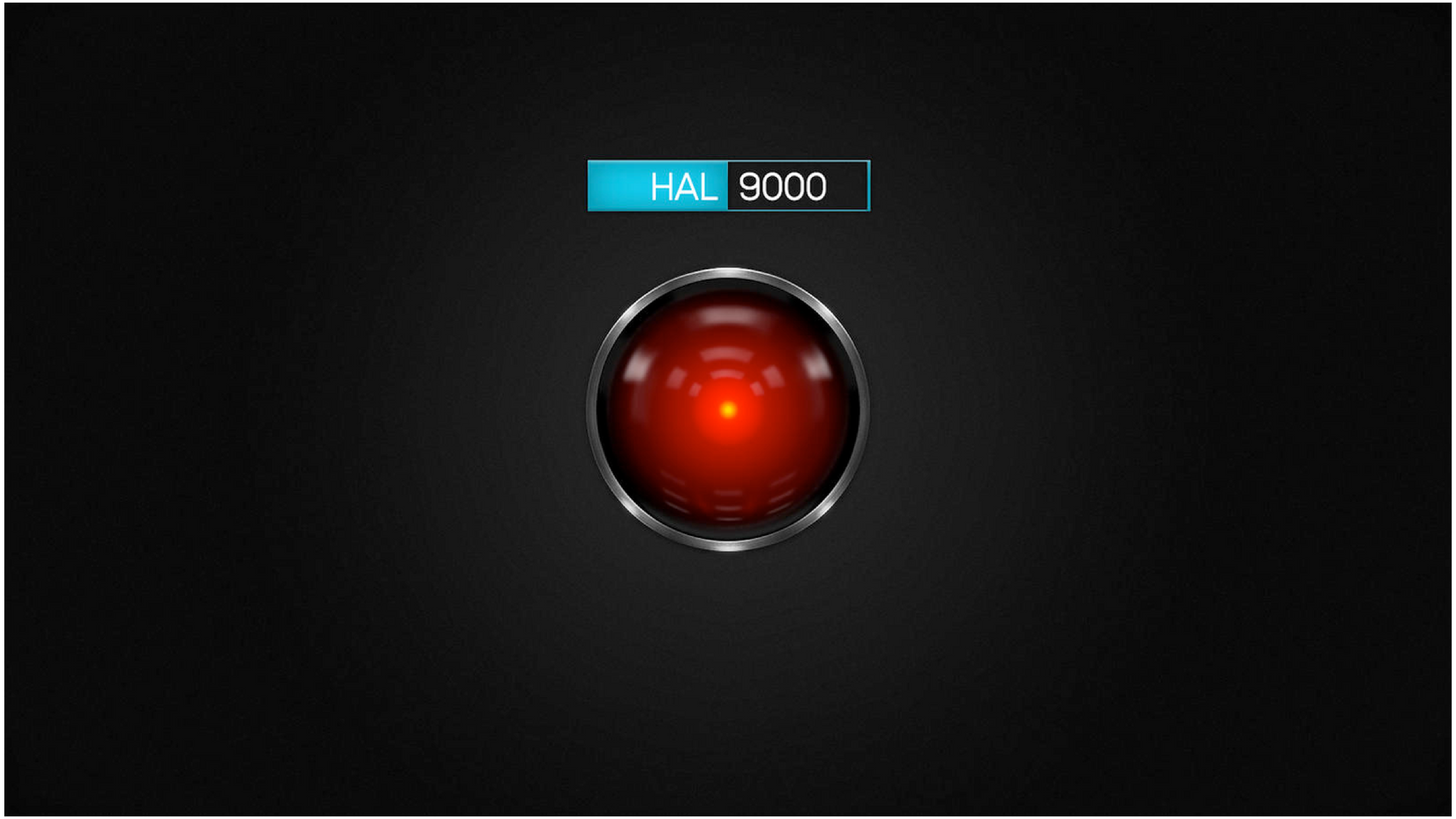
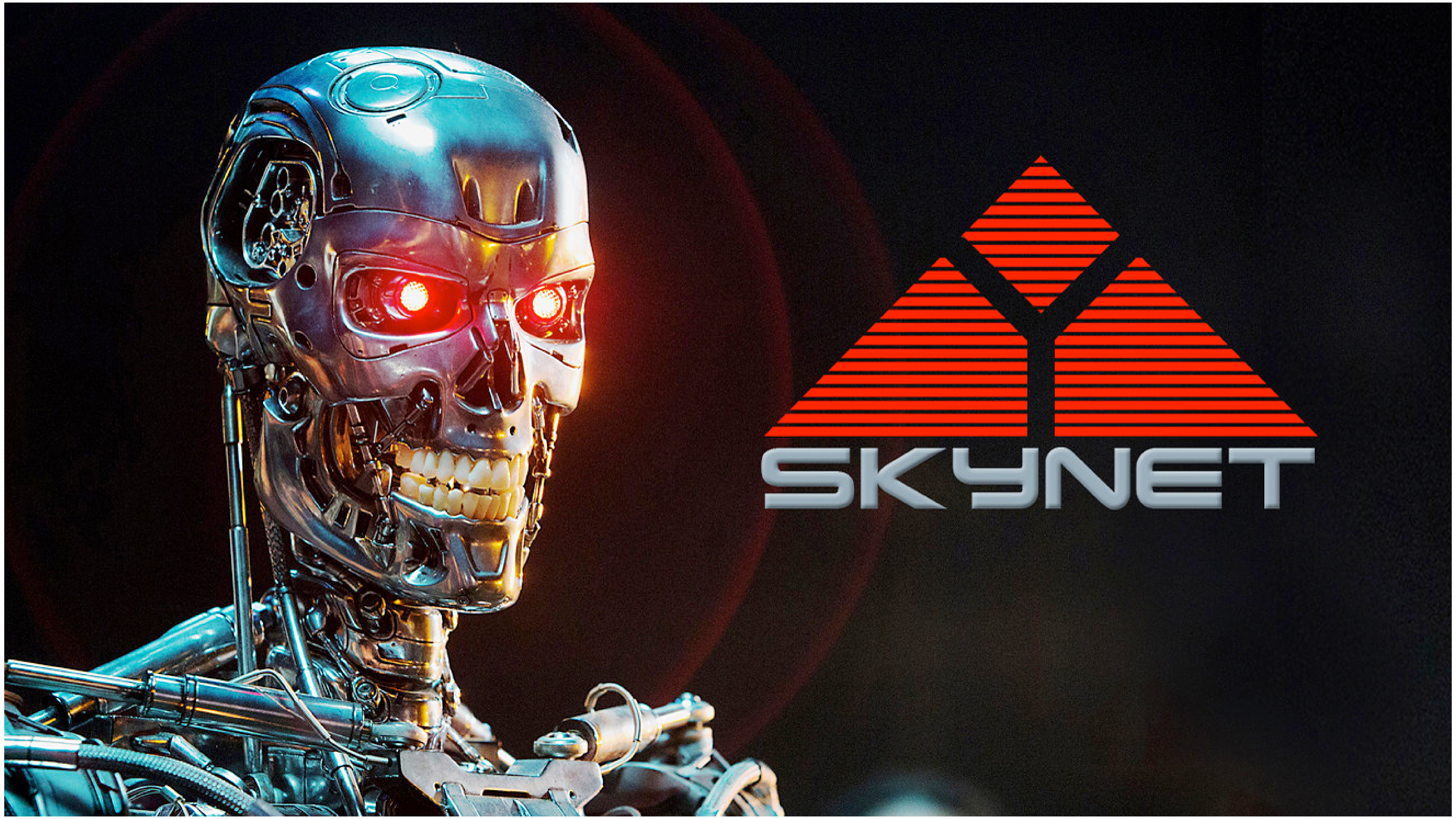
Si on aidait le capitaine Némo à classifier les monstres marins ?

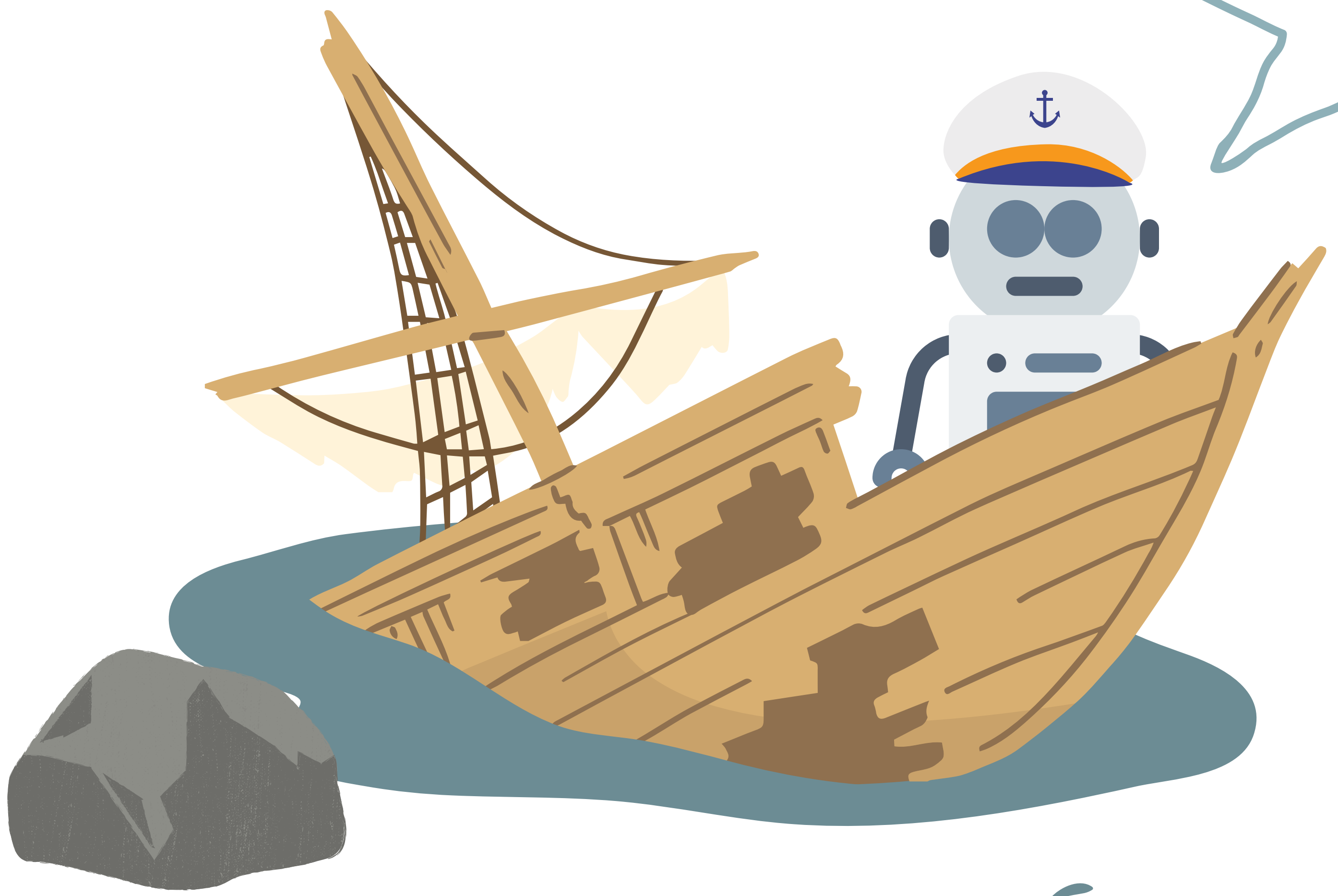
Eléa Petton & Stéphane Philippart
@EleaPetton & @wildagsx

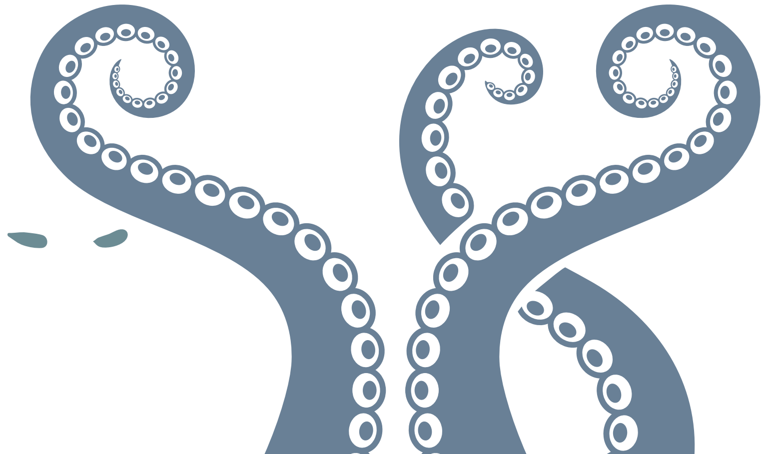
Mais l'Intelligence Artificielle c'est quoi au final ?


HAL 9000


SKYNET

QU'EST-CE QUE L'IA ?

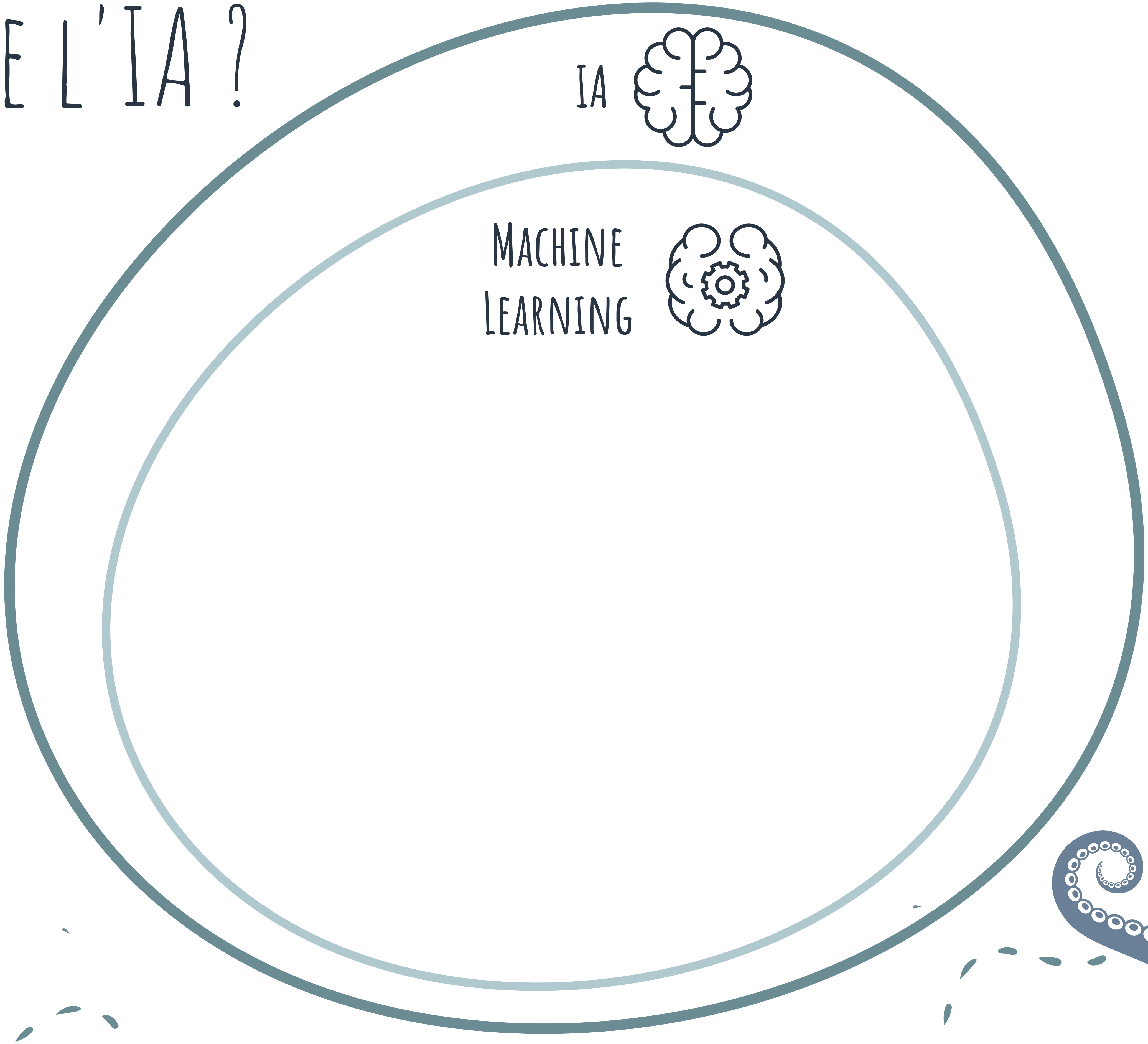




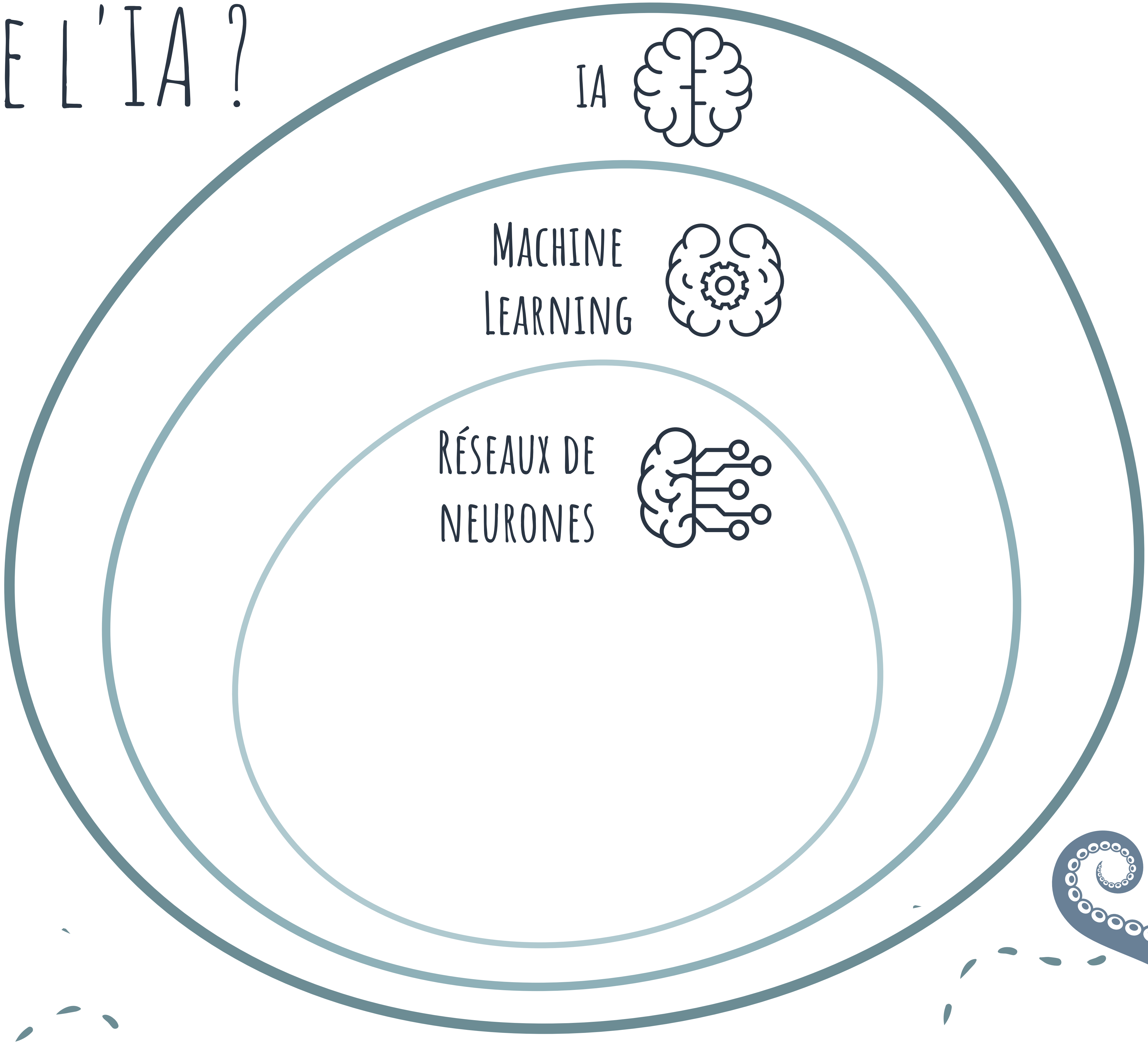
QU'EST-CE QUE L'IA ?

IA

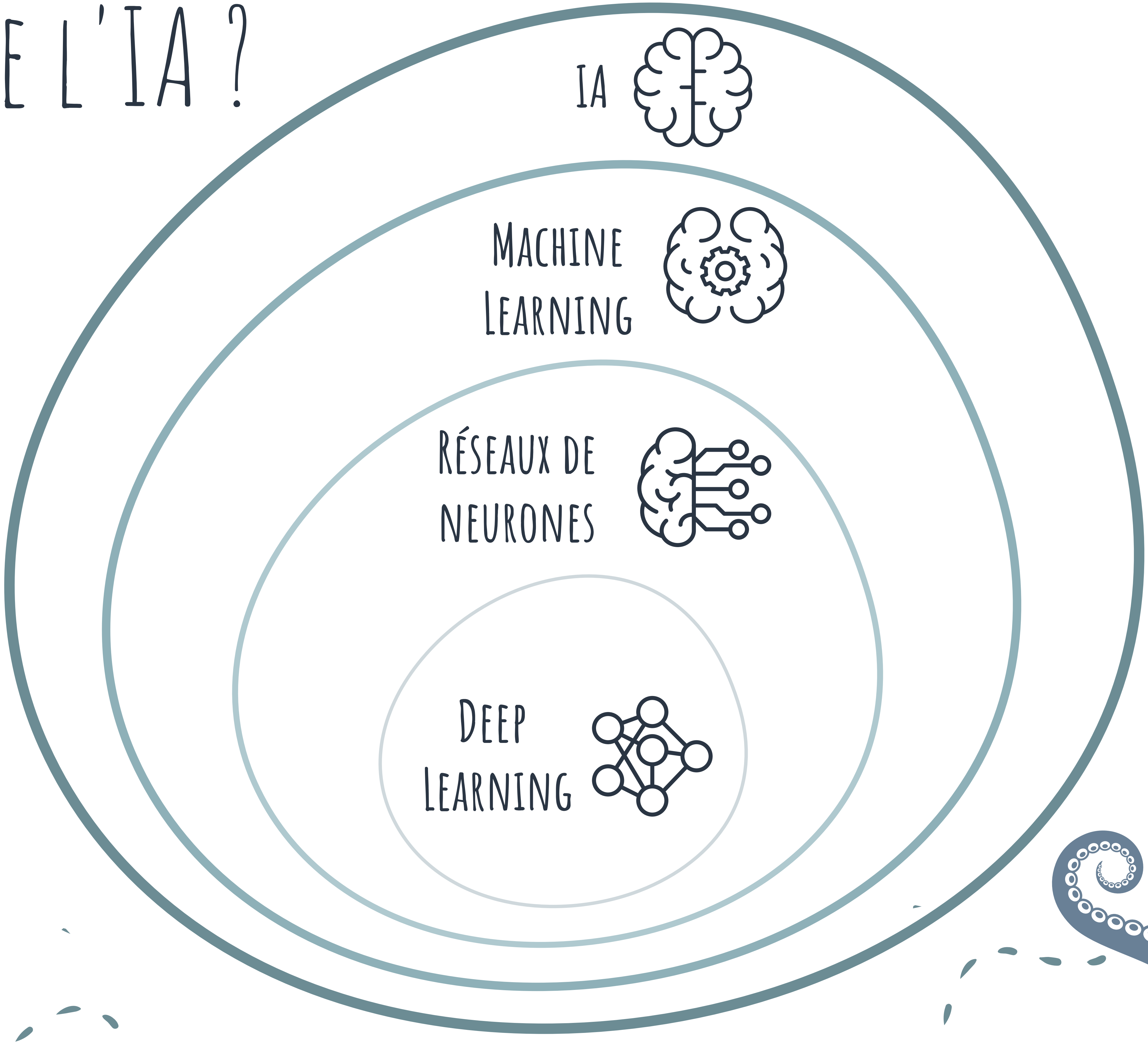
Qu'est-ce que l'IA ?



QU'EST-CE QUE L'IA ?



Qu'est-ce que l'IA ?



Le Machine Learning plus en détail

Machine Learning

Deep Learning

Supervisé

Classification

Régression

Le Machine Learning plus en détail

Machine Learning

Deep Learning

Non supervisé

Clustering

Réduction de dimension

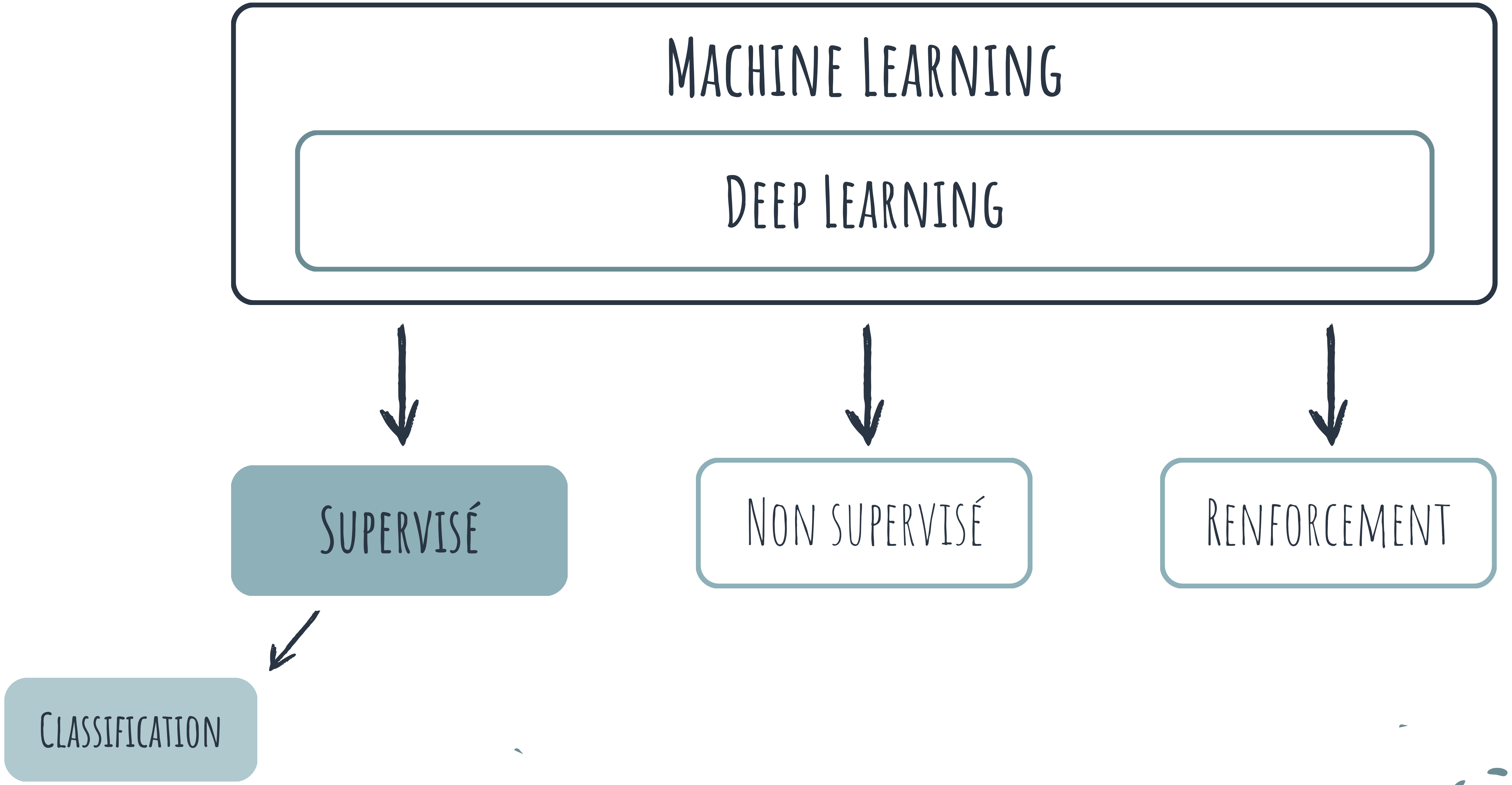
Le Machine Learning plus en détail

Machine Learning

Deep Learning

Renforcement

Le Machine Learning plus en détail

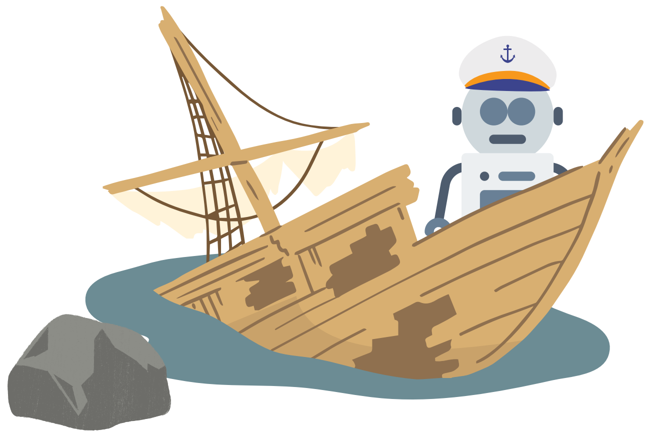


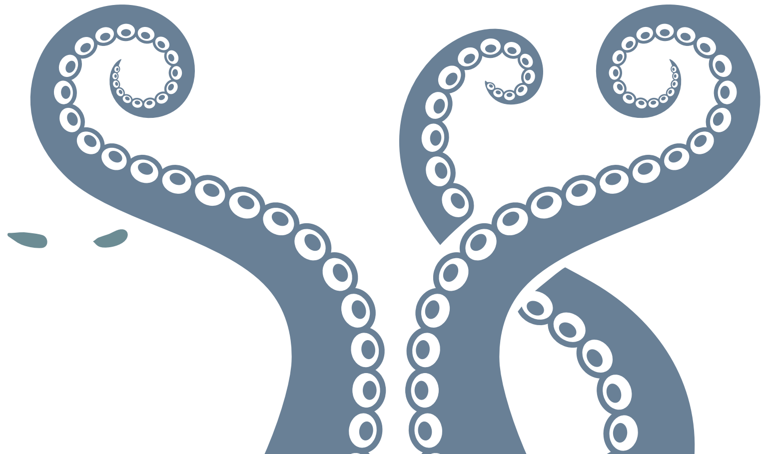
Différentes familles en IA

NLP

Le NLP, à quoi ça sert ?

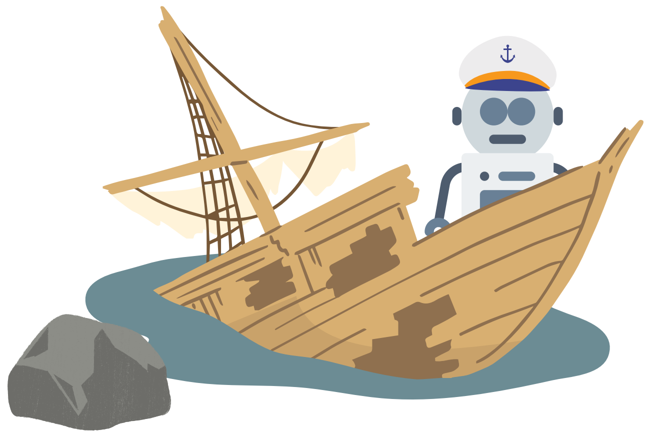
Traduction

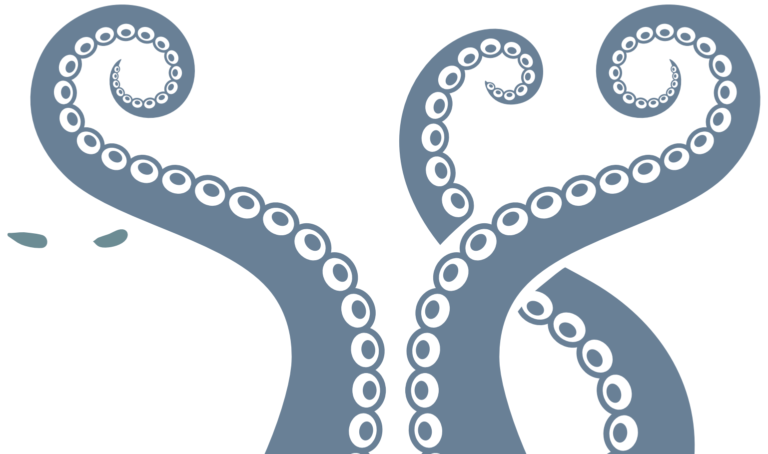




Le NLP, à quoi ça sert ?

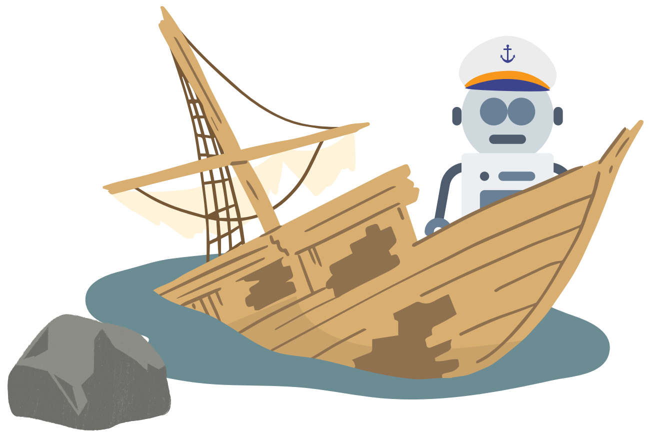
- Traduction
- Classification de texte

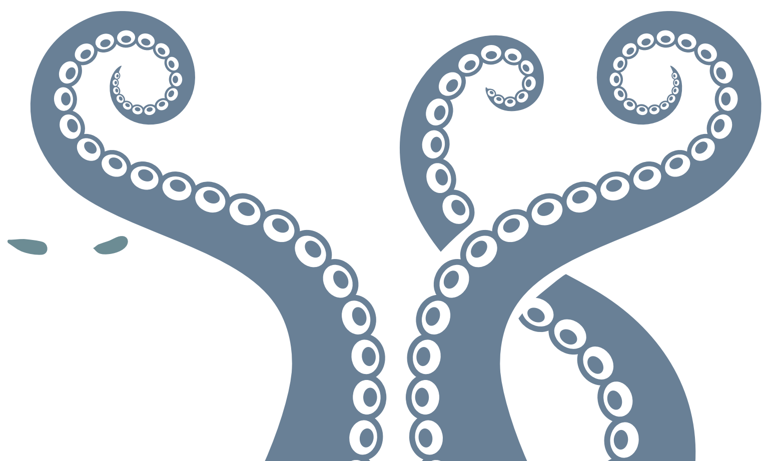




Le NLP, à quoi ça sert ?

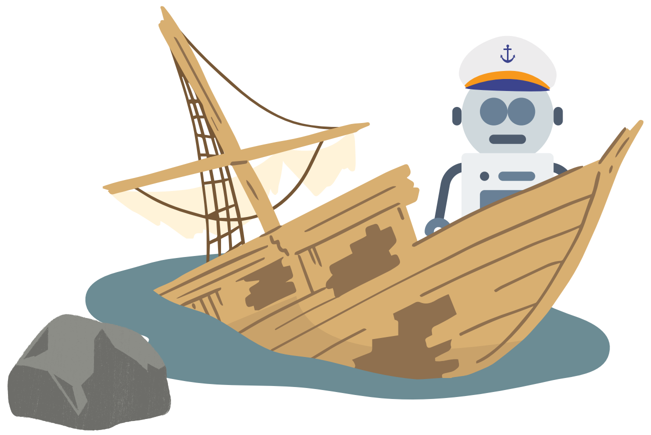
- Traduction
- Classification de texte
- Résumés

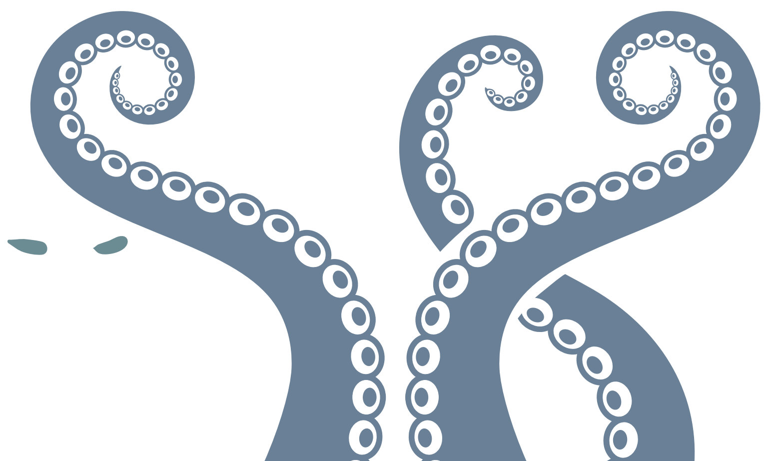




Le NLP, à quoi ça sert ?

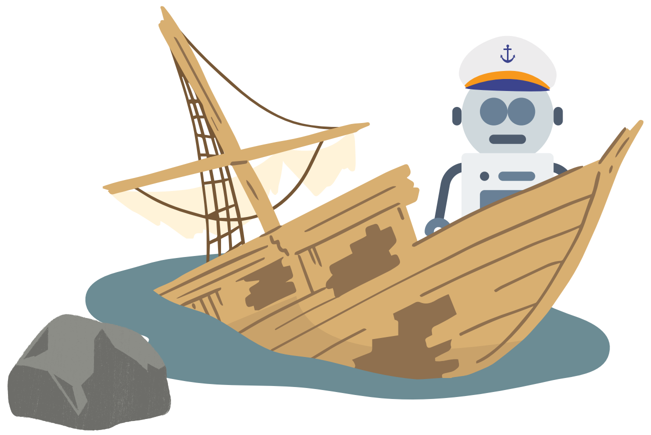
- Traduction
- Classification de texte
- Résumés
- Génération de texte

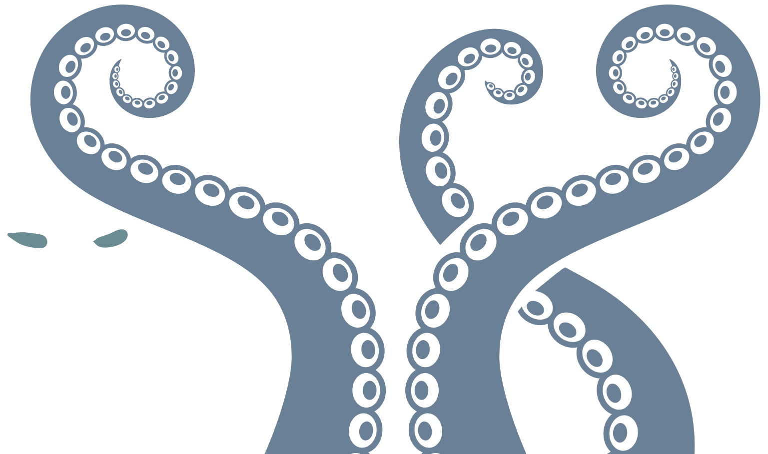




Le NLP, à quoi ça sert ?

- Traduction
- Classification de texte
- Résumés
- Génération de texte
- Réponses à des questions





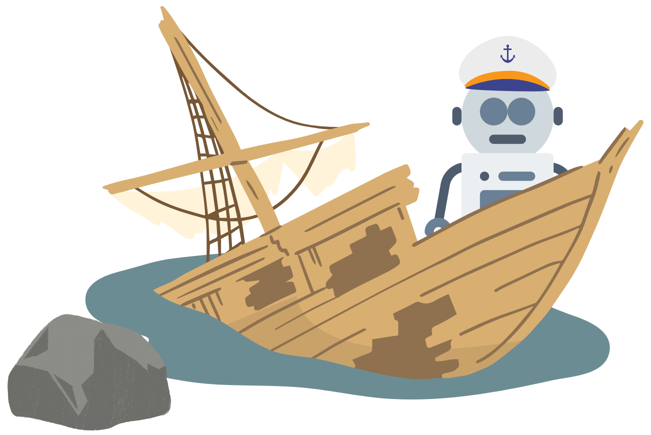
Différentes familles en IA

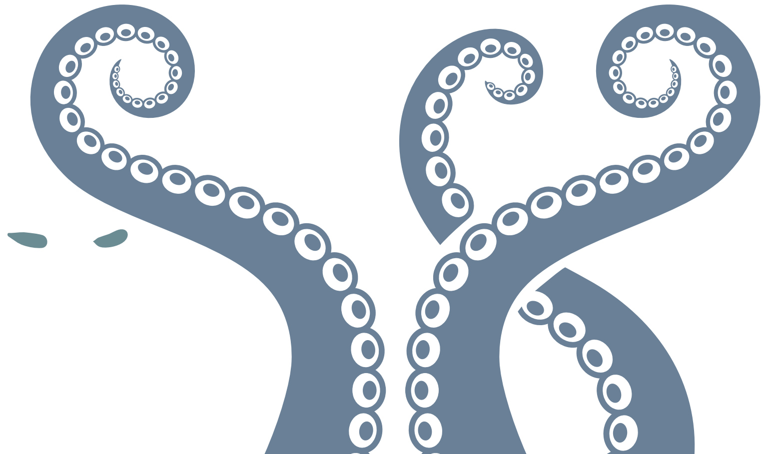
NLP

Computer Vision

La Computer Vision, c'est quoi ?

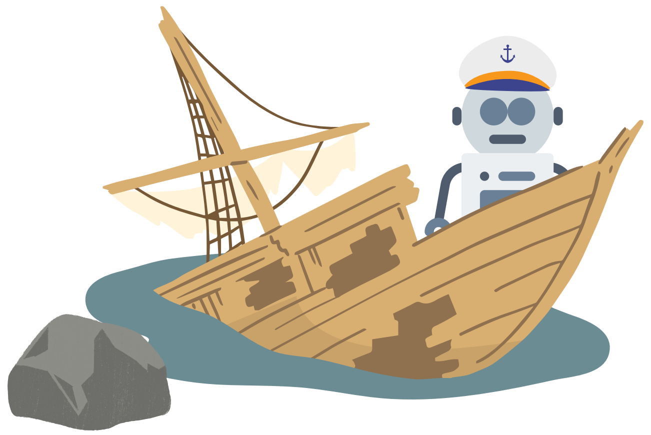
- Classification d'images

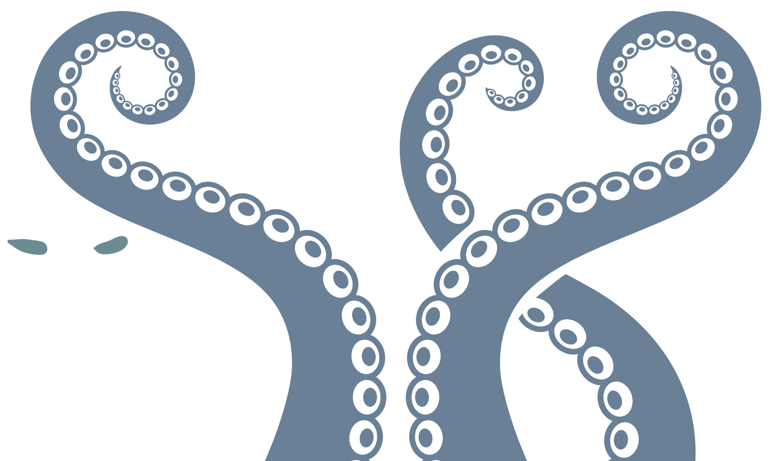




La Computer Vision, c'est quoi ?

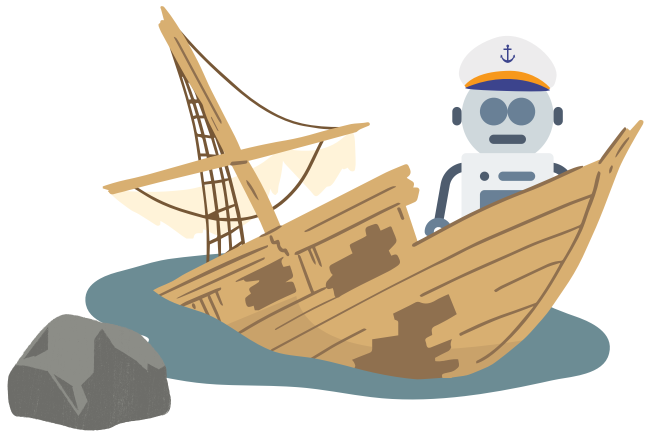
- Classification d'images
- Détection d'objets

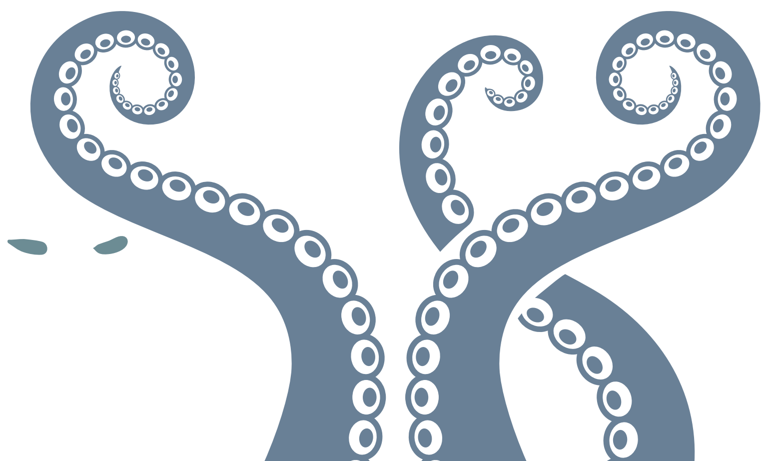




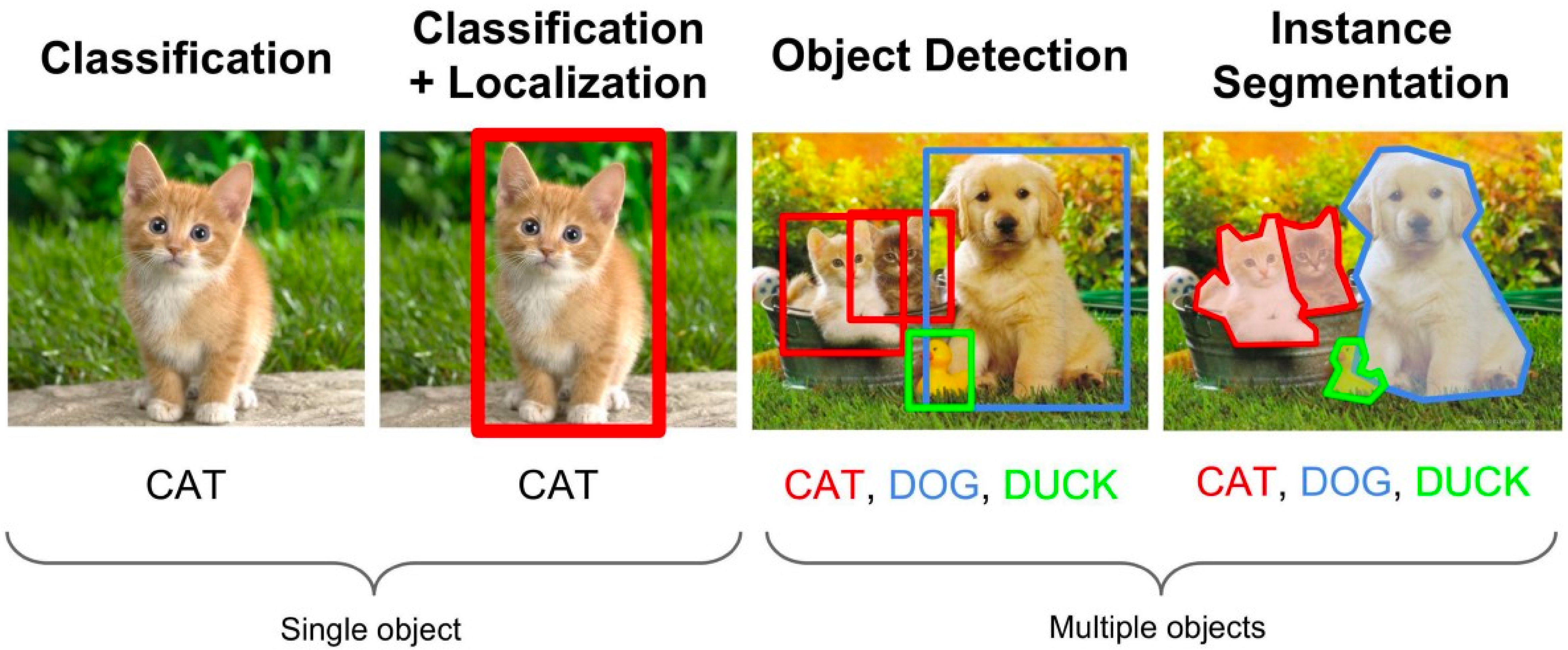
La Computer Vision, c'est quoi ?

- Classification d'images
- Détection d'objets
- Segmentation d'images





La Computer Vision, c'est quoi ?



Différentes familles en IA

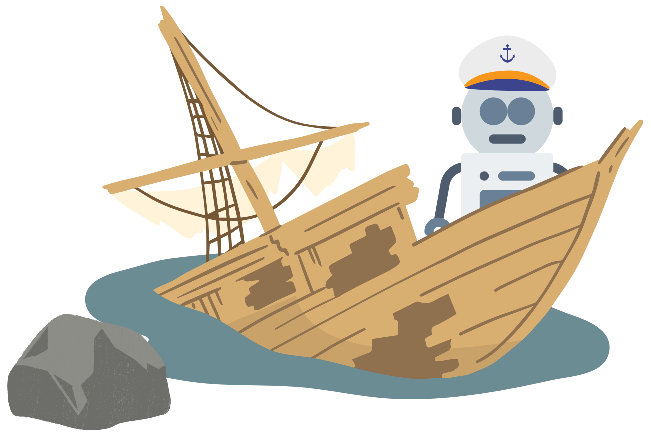
NLP

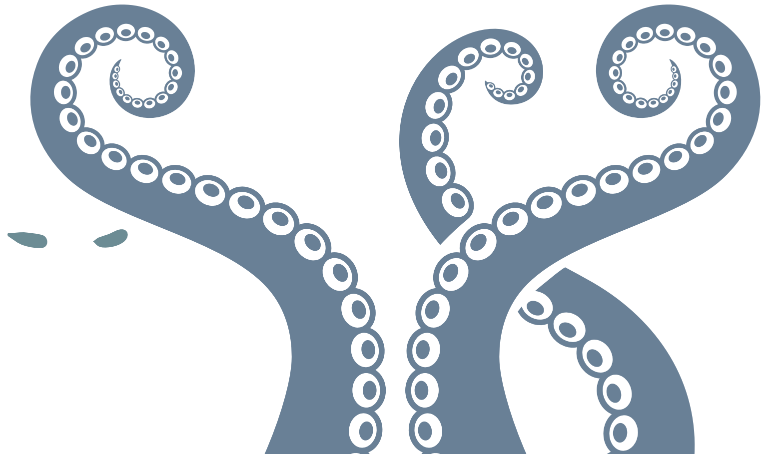
Computer Vision

Audio

Pourquoi le son en IA ?

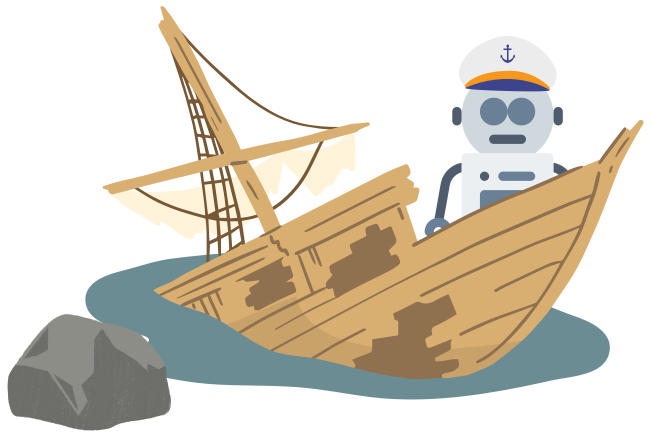
Classification des données acoustiques

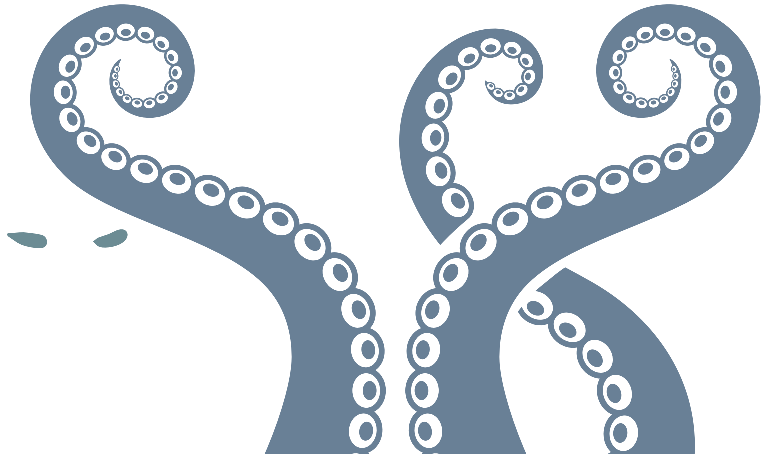




Pourquoi le son en IA ?

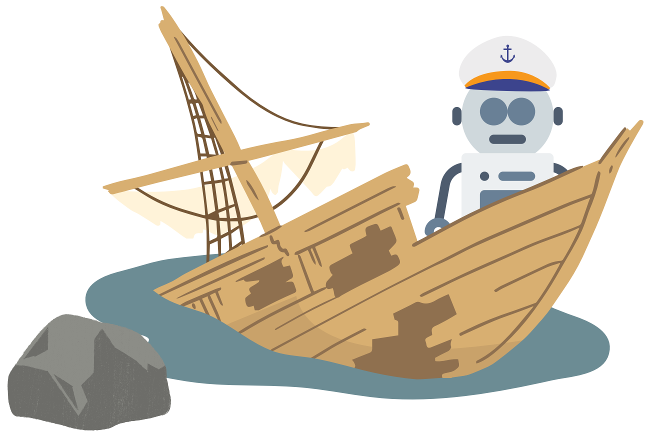
- Classification des données acoustiques
- Classification des sons environnementaux

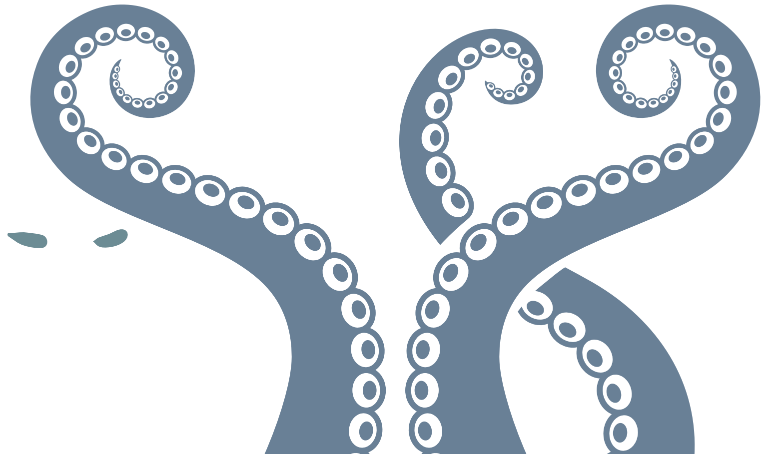




Pourquoi le son en IA ?

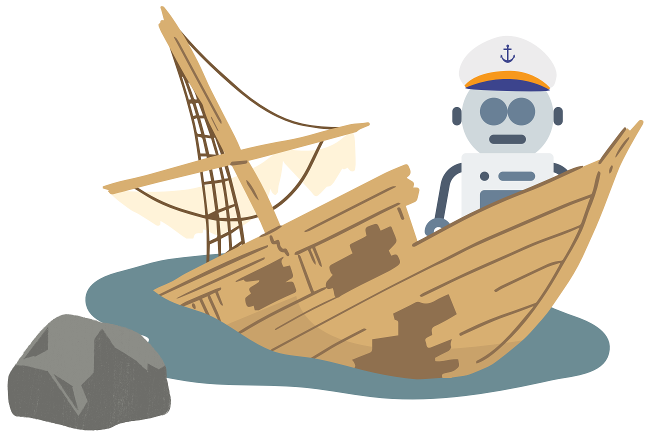
- Classification des données acoustiques
- Classification des sons environnementaux
- Classification du langage naturel

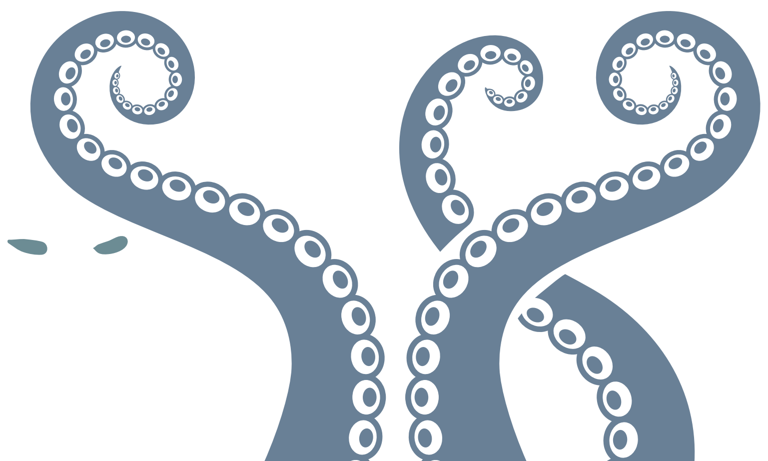




Pourquoi le son en IA ?

- Classification des données acoustiques
- Classification des sons environnementaux
- Classification du langage naturel
- Classification de la musique





Comment créer une application intelligente ?

- Obtenir de la donnée

Comment créer une application intelligente ?

- Obtenir de la donnée
- Traiter la donnée

Comment créer une application intelligente ?

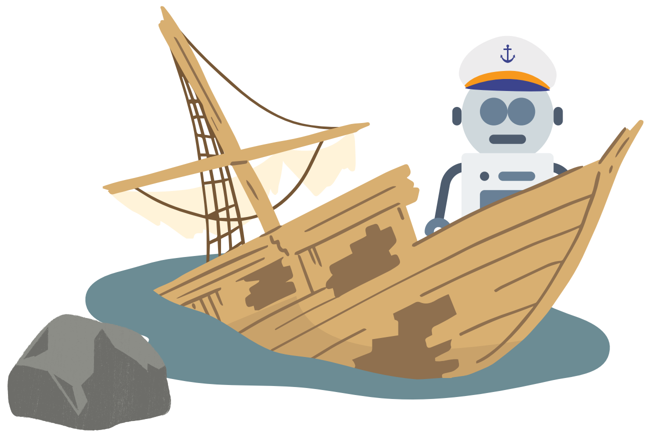
- Obtenir de la donnée
- Traiter la donnée
- Créer ou optimiser un modèle existant

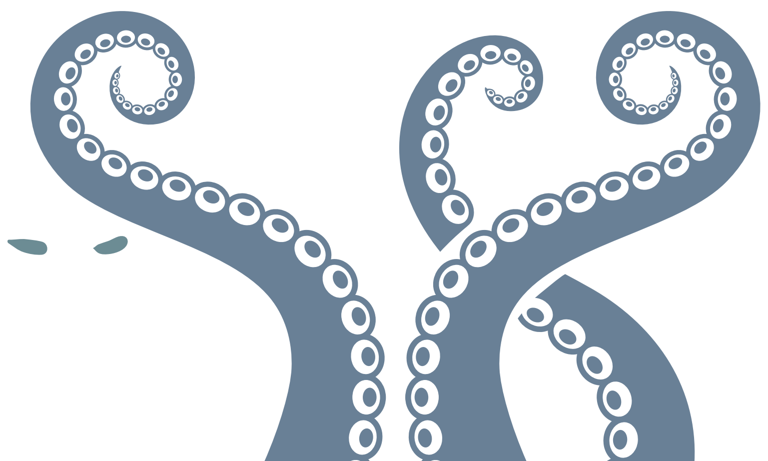
Comment créer une application intelligente ?

- Obtenir de la donnée
- Traiter la donnée
- Créer ou optimiser un modèle existant
- Entraîner son modèle

Comment créer une application intelligente ?

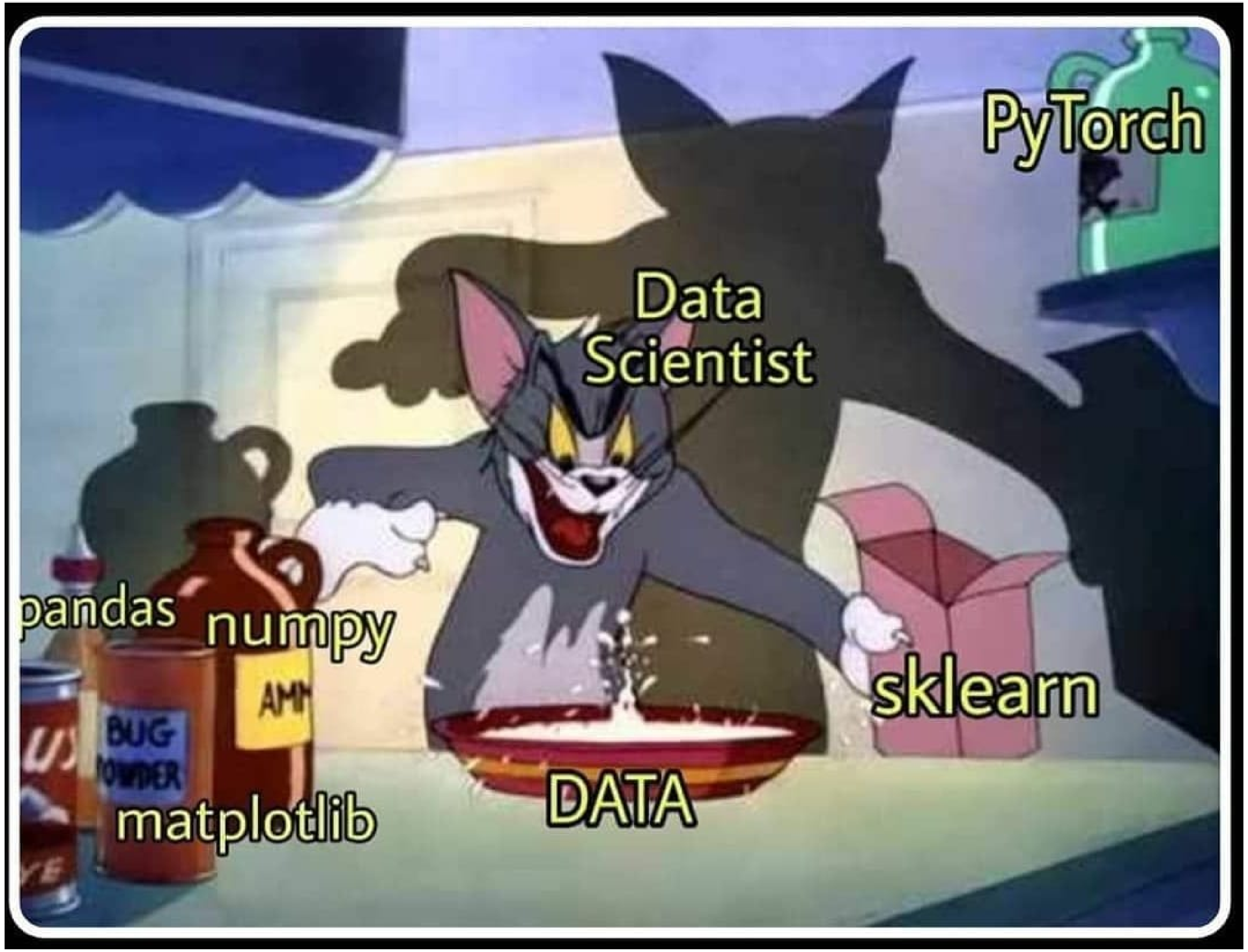
- Obtenir de la donnée
- Traiter la donnée
- Créer ou optimiser un modèle existant
- Entraîner son modèle
- Utiliser son modèle





Importance du pré-traitement des données

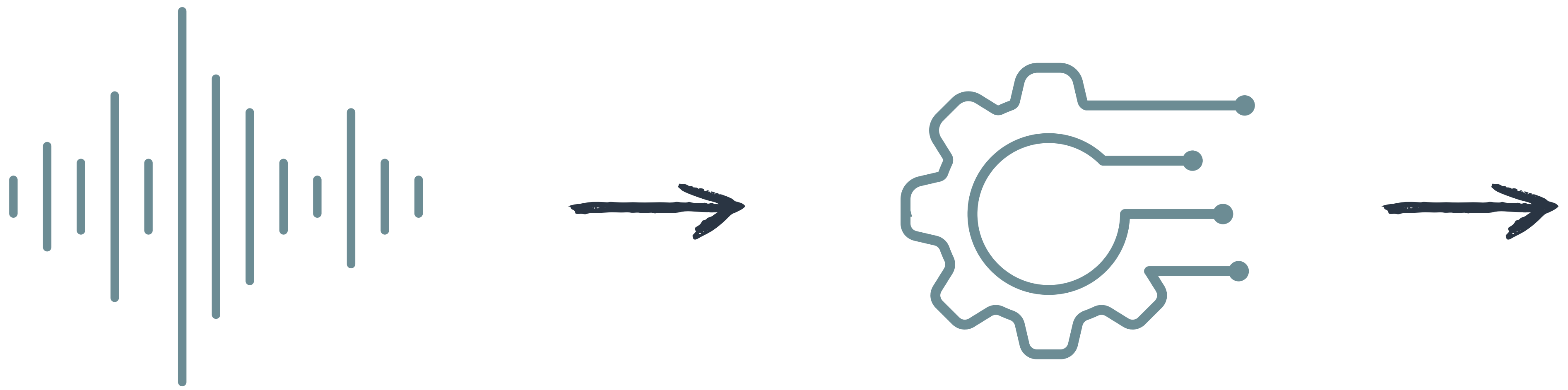
- Qualité des données
- Informations utiles


PyTorch
Data Scientist
pandas
numpy
AMM
BUG POWDER
matplotlib
DATA
sklearn

Exemples de pré-traitement pour les données

- Gestion des valeurs nulles / manquantes
- Normalisation
- Gestion des variables qualitatives
- Encodage à chaud

Et pour les données audios ?



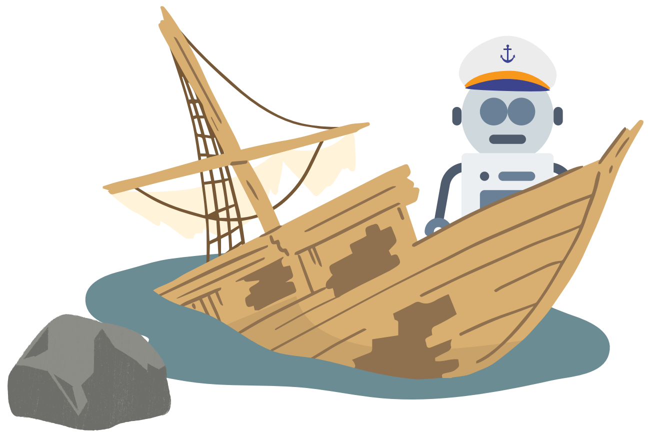
audios .wav

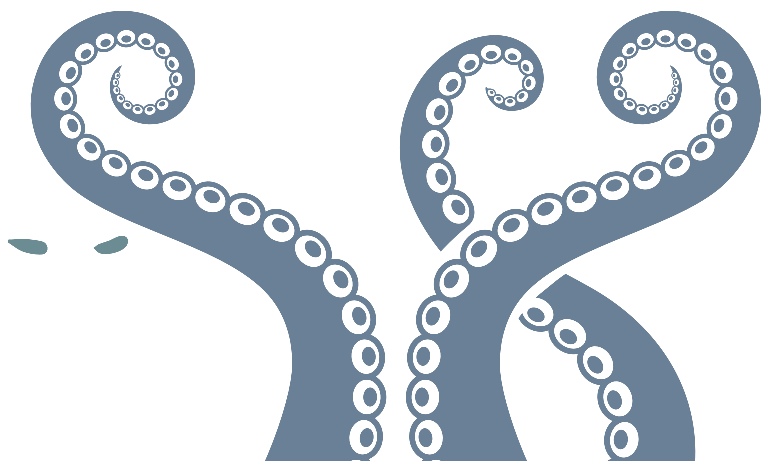
paramètres audio Librosa

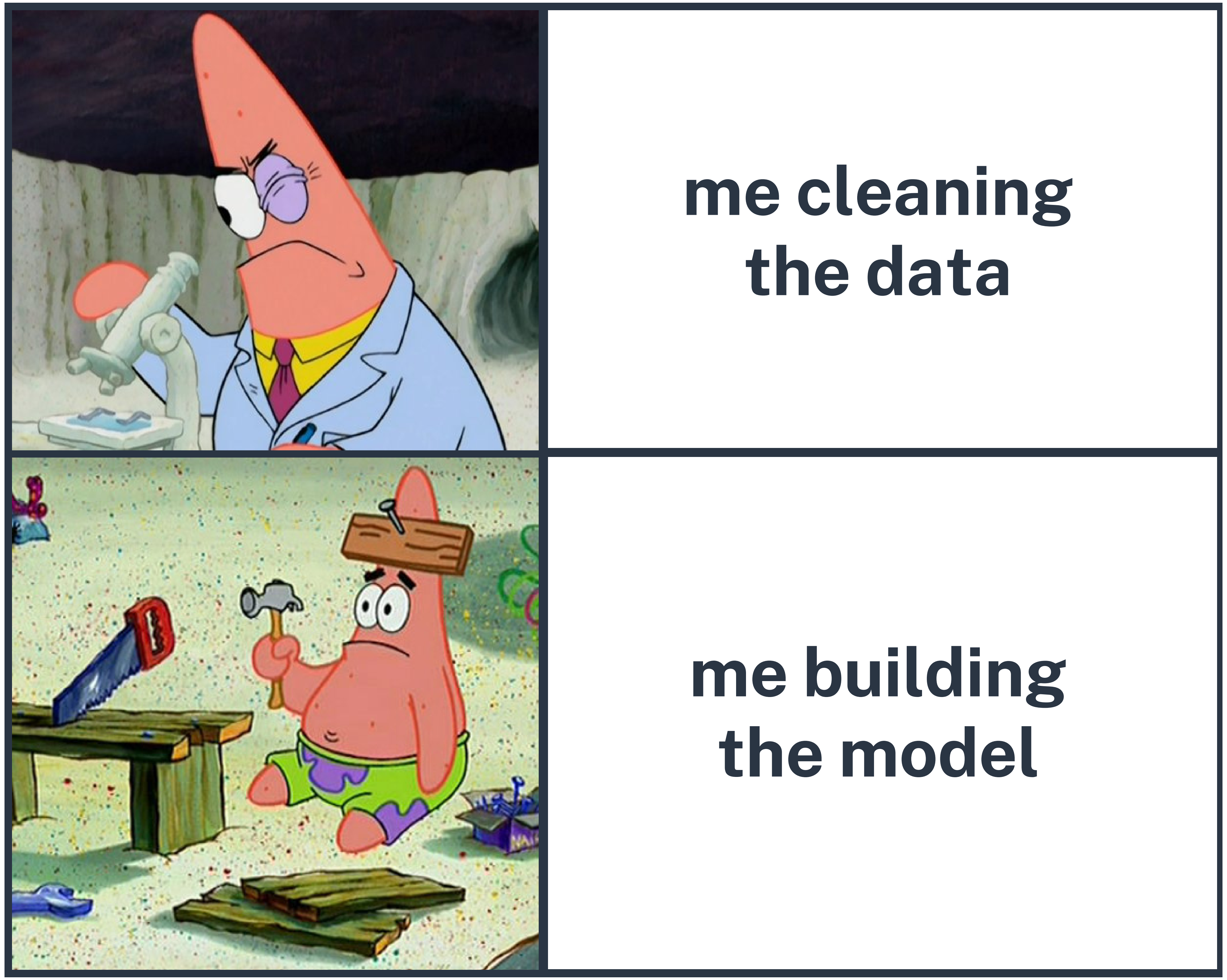
N paramètres audio Librosa

name	length	...	...	...	name

fichier .csv






me cleaning the data
me building the model

Technos et Outils

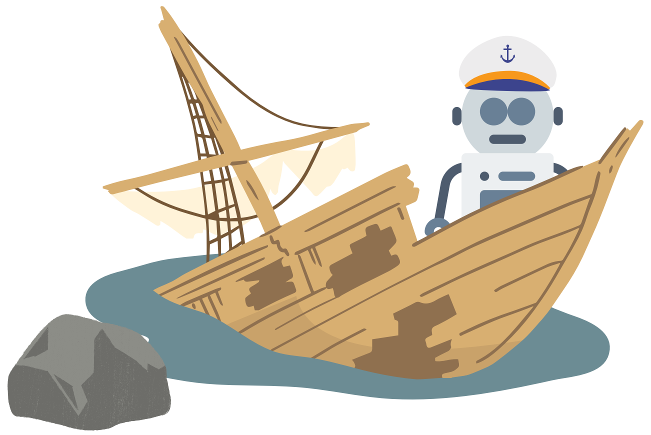
- Le langage le plus utilisé : Python

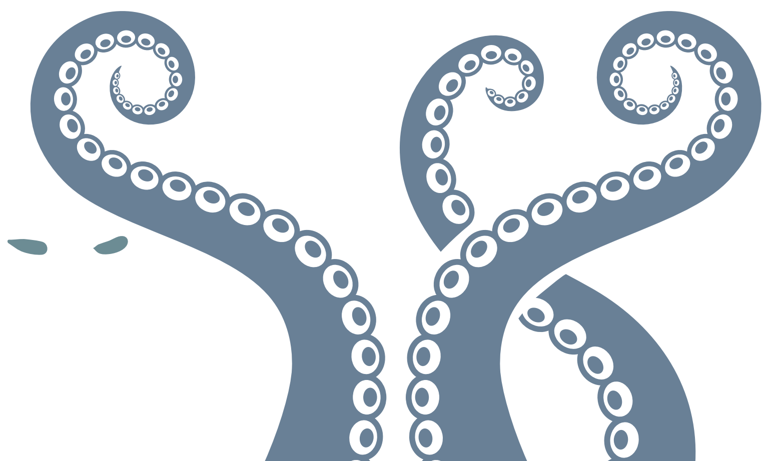
Technos et Outils

- Le langage le plus utilisé : Python
- Beaucoup d'IDE (PyCharm, VSCode …)
  - Le cas à part de JupyterLab

Technos et Outils

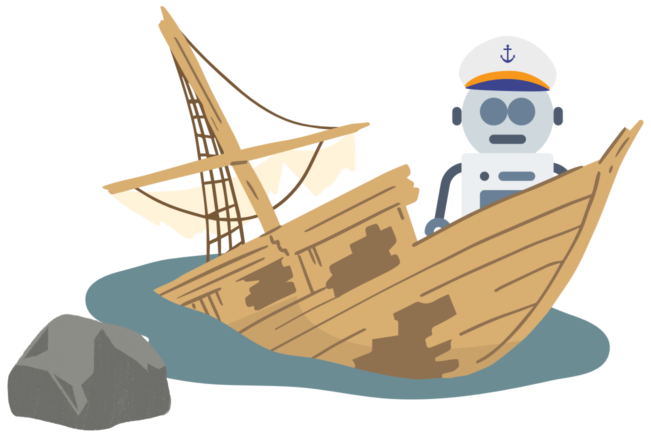
- Le langage le plus utilisé : Python
- Beaucoup d'IDE (PyCharm, VSCode …)
  - Le cas à part de JupyterLab
- Les frameworks sont nombreux
  - Peut dépendre du type de modèle ou donnée manipulée

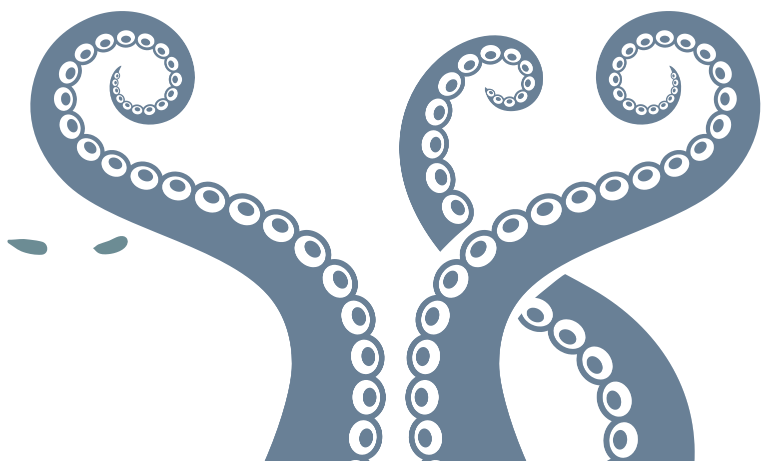




Technos et Outils

- Le langage le plus utilisé : Python
- Beaucoup d'IDE (PyCharm, VSCode …)
  - Le cas à part de JupyterLab
- Les frameworks sont nombreux
  - Peut dépendre du type de modèle ou donnée manipulée
  - Compatibilité GPU (surtout pour la partie traitement de la donnée)





Technos et Outils

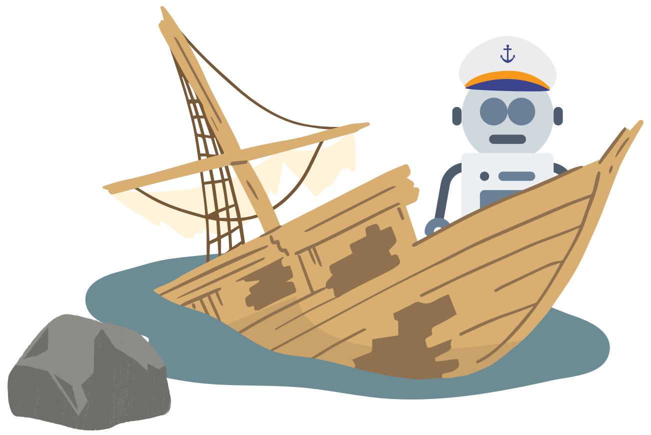
Le langage le plus utilisé : Python

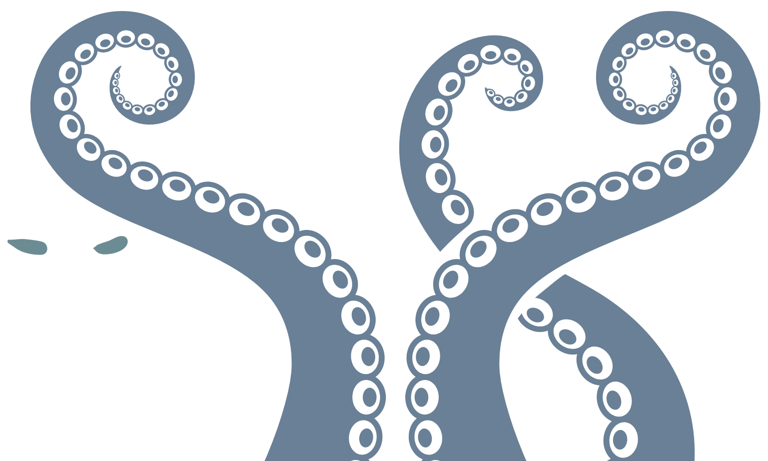
Beaucoup d'IDE (PyCharm, VSCode ...)

- Le cas à part de JupyterLab

Les frameworks sont nombreux

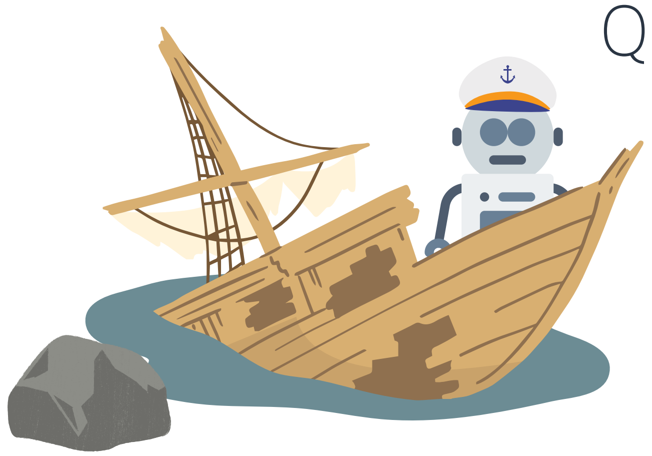
- Peut dépendre du type de modèle ou donnée manipulée
- Compatibilité GPU (surtout pour la partie traitement de la donnée)
- Construction de modèles, Apprentissage, Front, Back, ...



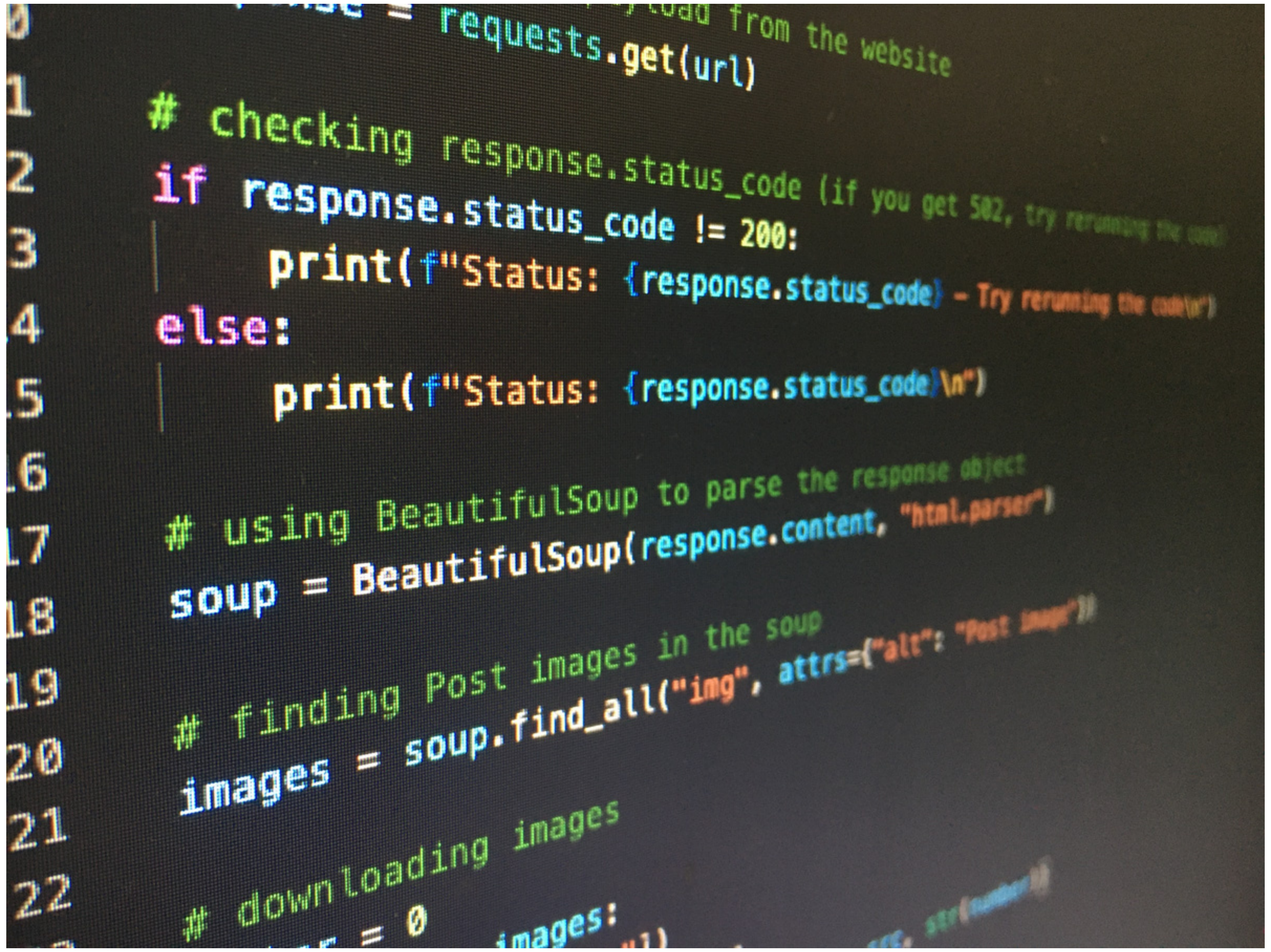


Technos et Outils

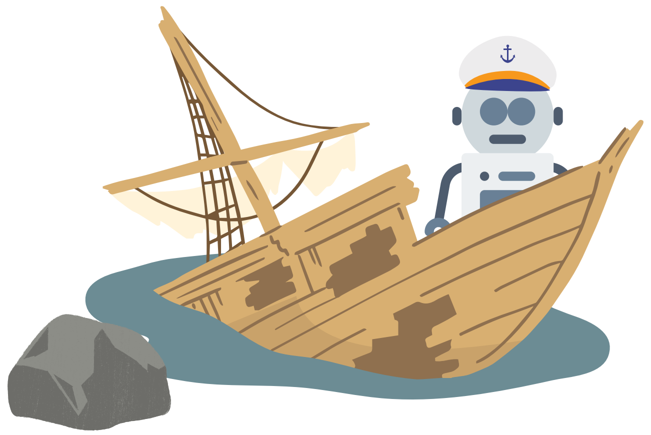
- Le langage le plus utilisé : Python
- Beaucoup d'IDE (PyCharm, VSCode …)
  - Le cas à part de JupyterLab
- Les frameworks sont nombreux
  - Peut dépendre du type de modèle ou donnée manipulée
  - Compatibilité GPU (surtout pour la partie traitement de la donnée)
  - Construction de modèles, Apprentissage, Front, Back, …
  - Quelques frameworks connus : Pytorch, TensorFlow, Streamlit, …



DÉMO TIME !



Artturi Jalli



Identifier un mammifère par rapport à son cri

Notebook

Analyse de l'audio
Pré-traitement des données
Construction du modèle
Entraînement et évaluation du modèle
Test et enregistrement modèle

Bart van meele

Identifier un mammifère par rapport à son cri

Notebook

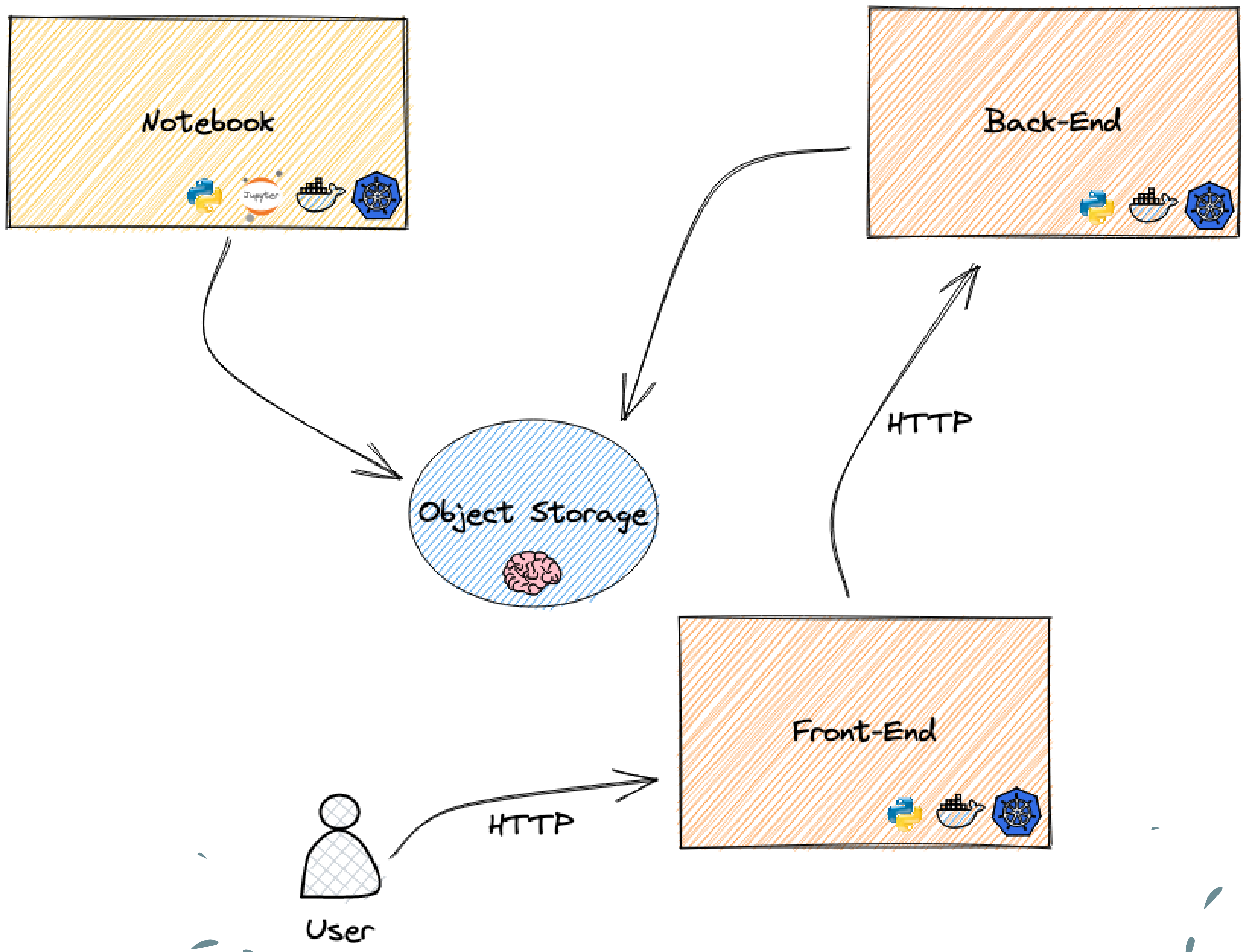
- Analyse de l'audio
- Pré-traitement des données
- Construction du modèle
- Entraînement et évaluation du modèle
- Test et enregistrement modèle

Bart van meele

Applications

- Back-End : exposition de l'API
- Front-End : analyse d'un son chargé ou enregistré

Schéma d'architechture globale




WHAT DID
WE LEARN?

Qu'avons nous appris aujourd'hui ?

Notion d'IA

Datascience, Machine Learning, Deep Learning

Qu'avons nous appris aujourd'hui ?

Notion d'IA

Datascience, Machine Learning, Deep Learning

Place du son en IA

Environnement, musique, NLP

Qu'avons nous appris aujourd'hui ?

Notion d'IA

Datascience, Machine Learning, Deep Learning

Place du son en IA

Environnement, musique, NLP

Prétraitement des données

Répartition, transformation, labélisation

Qu'avons nous appris aujourd'hui ?

Notion d'IA

Datascience, Machine Learning, Deep Learning

Place du son en IA

Environnement, musique, NLP

Prétraitement des données

Répartition, transformation, labélisation

AI Notebook

Développement de mon code étape par étape

Qu'avons nous appris aujourd'hui ?

Notion d'IA

Datascience, Machine Learning, Deep Learning

AI Notebook

Développement de mon code étape par étape

Place du son en IA

Environnement, musique, NLP

Entraînement d'une IA

Réseau de neurones, exportation

Prétraitement des données

Répartition, transformation, labélisation

Qu'avons nous appris aujourd'hui ?

Notion d'IA

Datascience, Machine Learning, Deep Learning

AI Notebook

Développement de mon code étape par étape

Place du son en IA

Environnement, musique, NLP

Entraînement d'une IA

Réseau de neurones, exportation

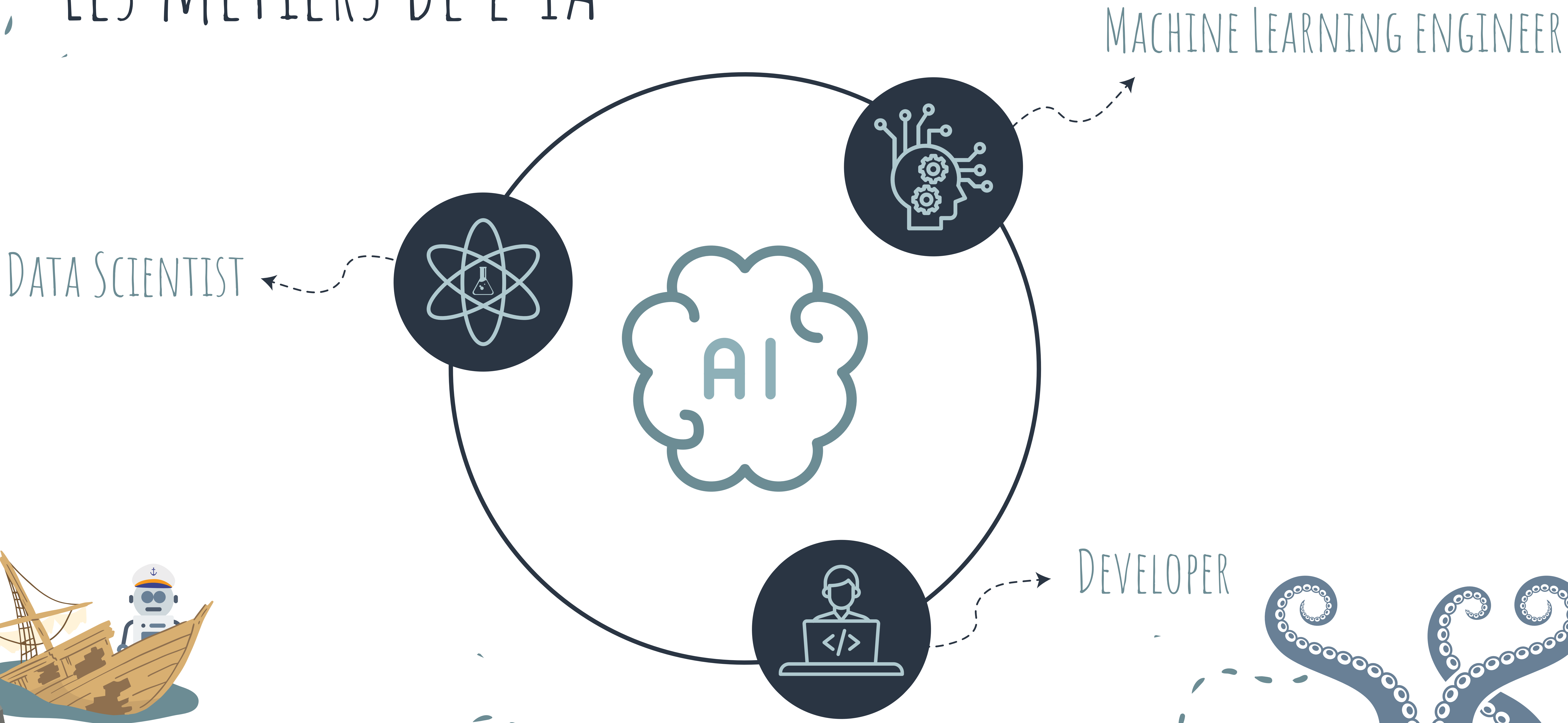
Prétraitement des données

Répartition, transformation, labélisation

App pour IA

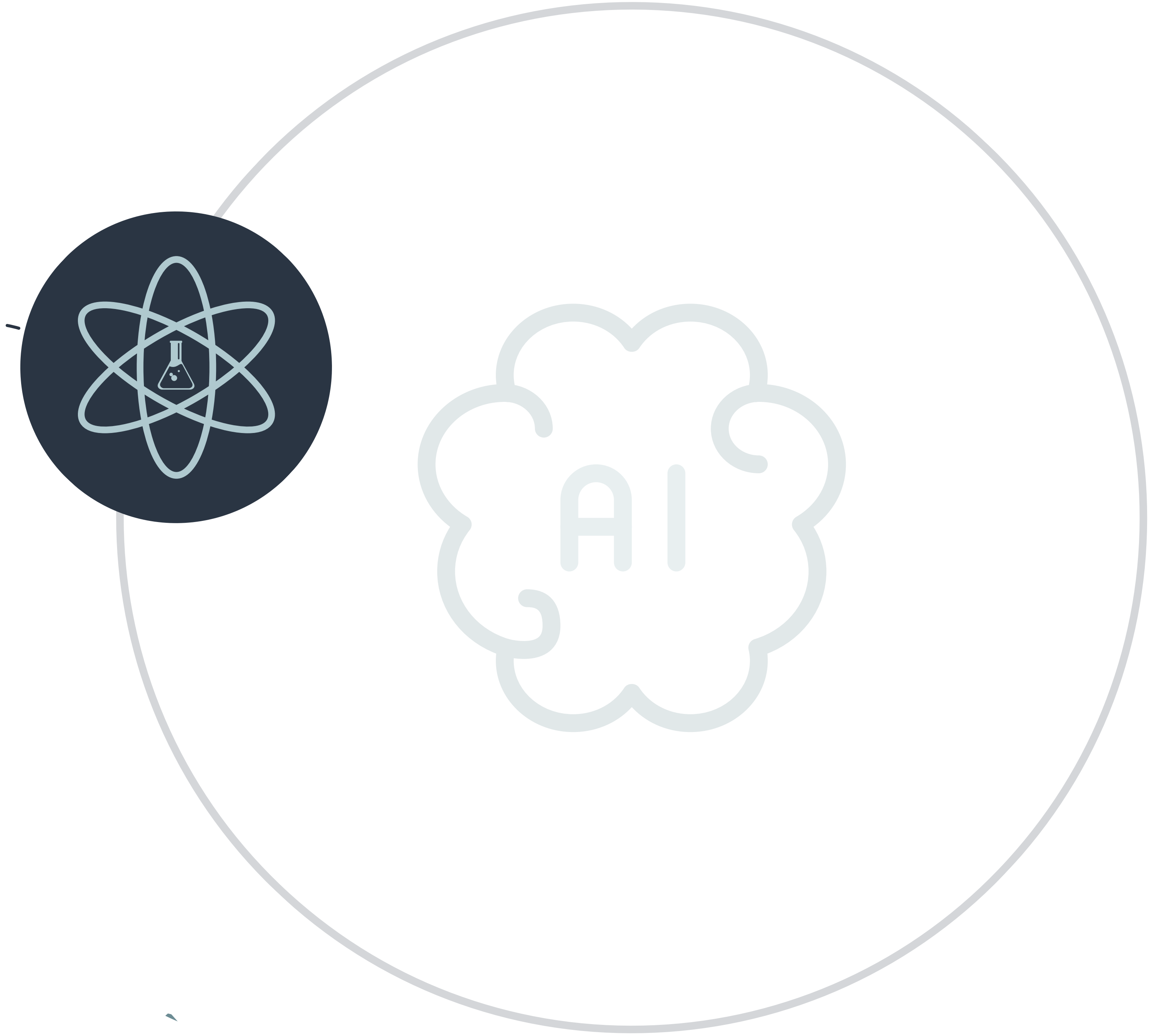
Object Storage, Streamlit, Docker, API

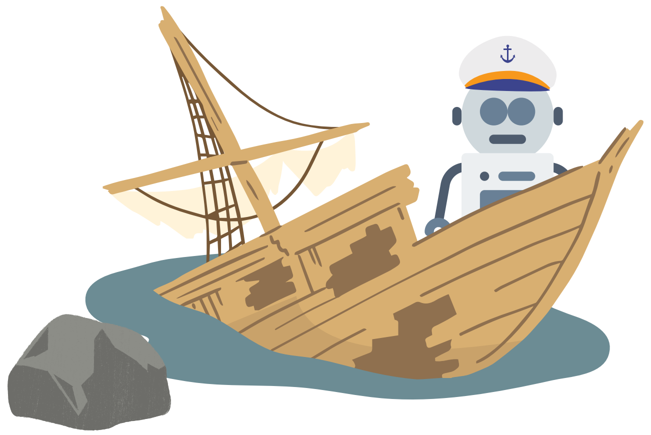
Les métiers de l'IA

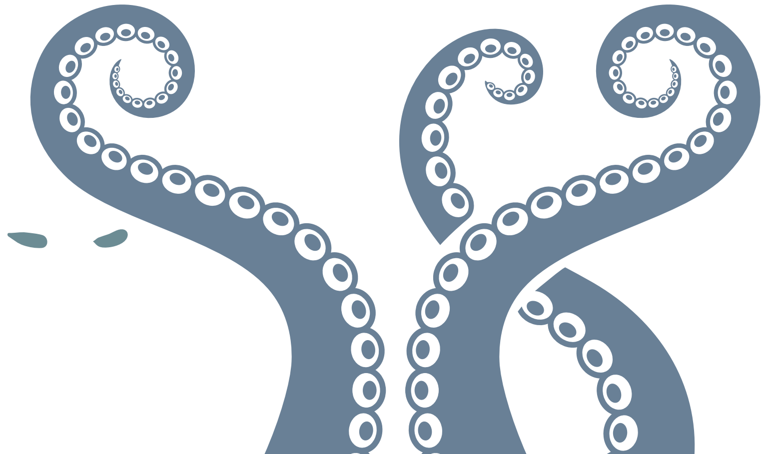


Les métiers de l'IA

Data Scientist







Le Data Scientist

Data Analysis

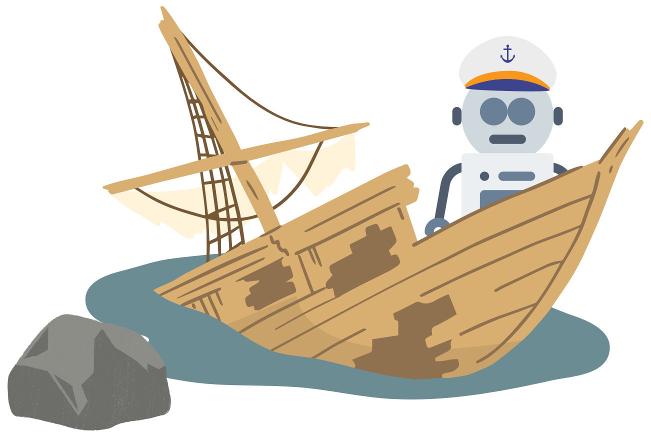
Data Visualization

Programmation

Machine Learning

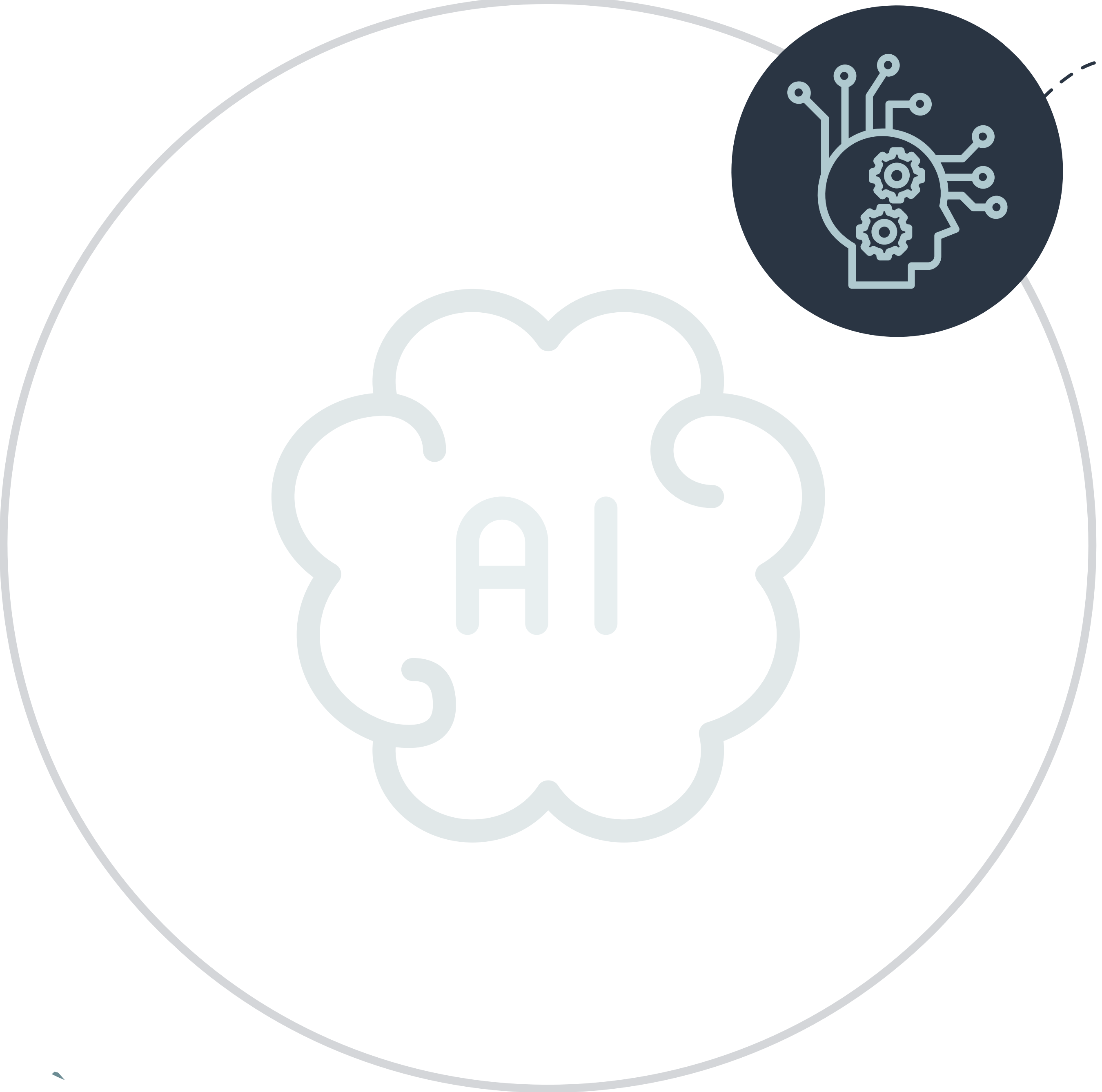
Maths

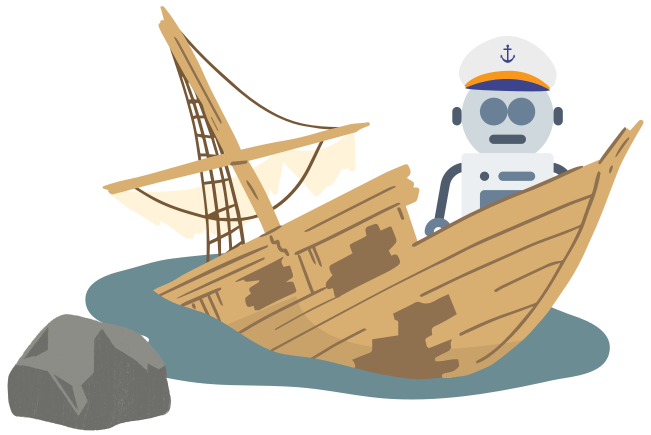
Statistics

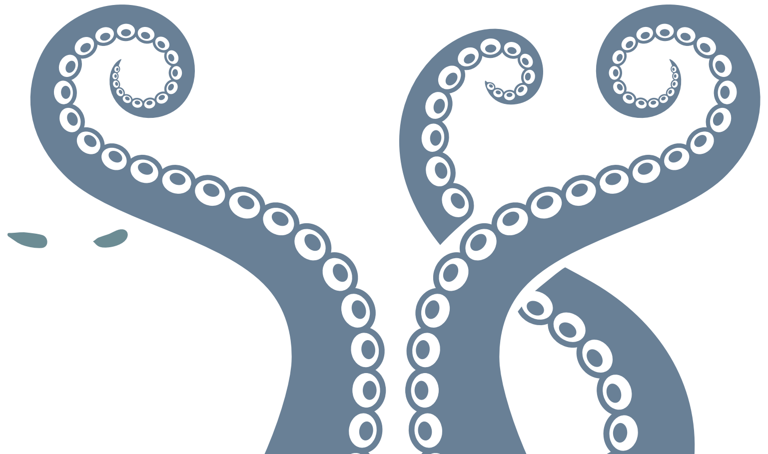


Les métiers de l'IA

Machine Learning engineer







Le Machine Learning Engineer

Data Science

Maths & Statistics

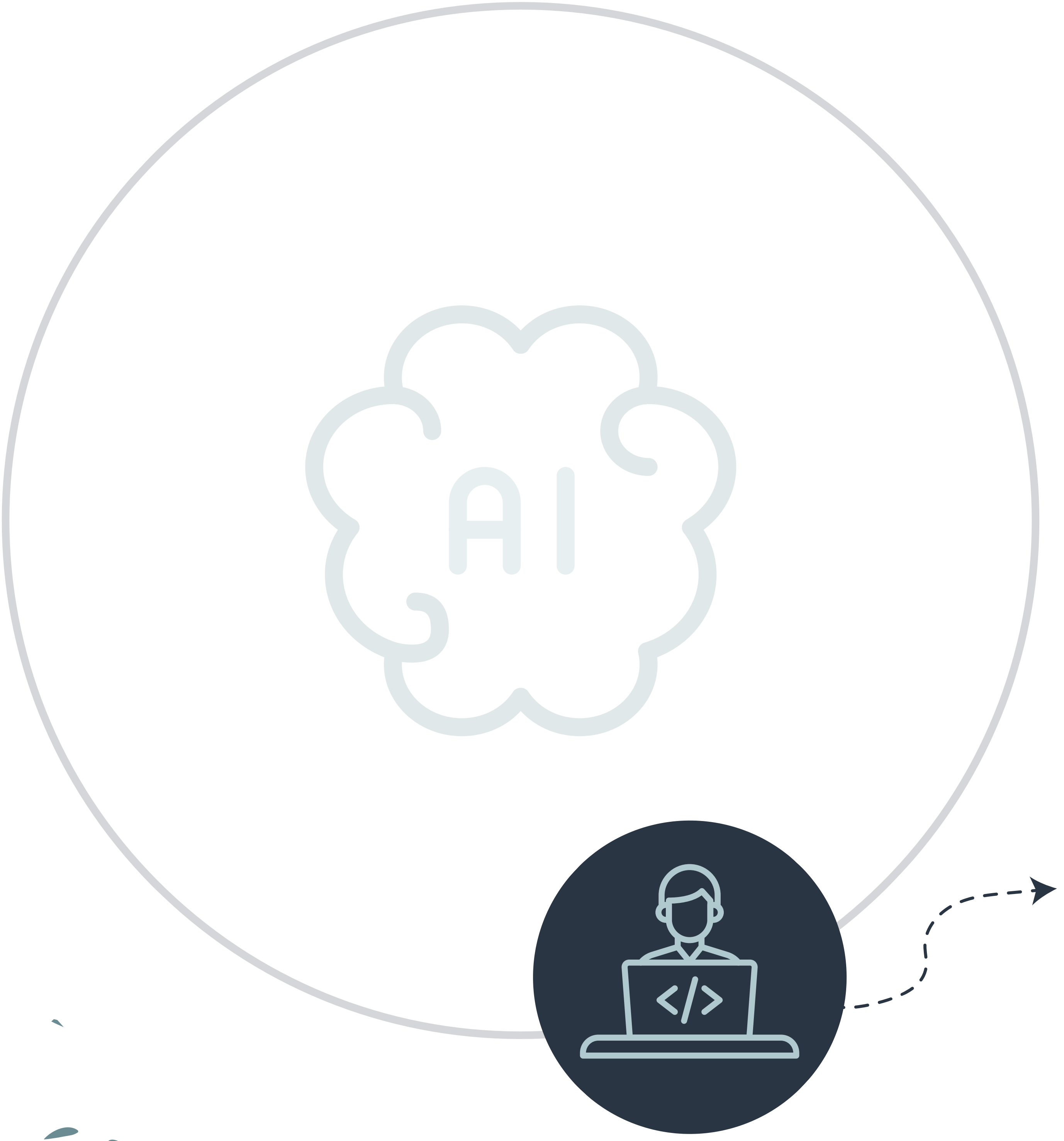
Programmation

Frameworks

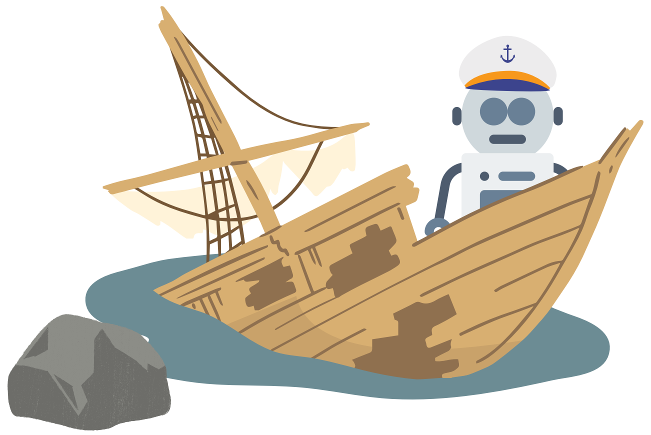
Environnement de prod

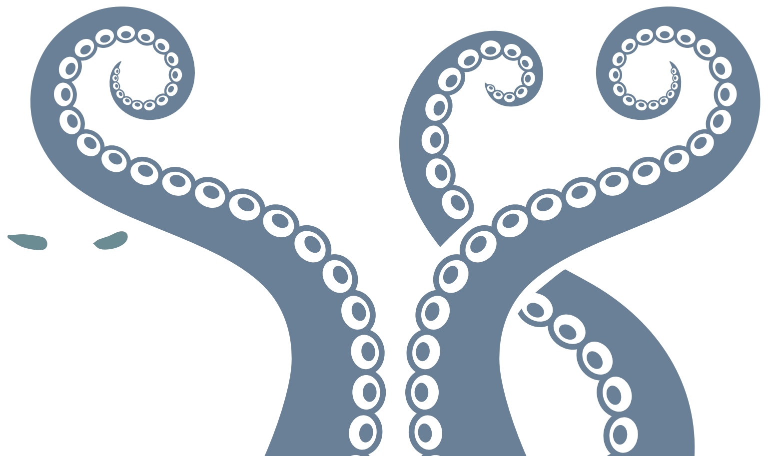
Services Cloud

Les métiers de l'IA



Developer



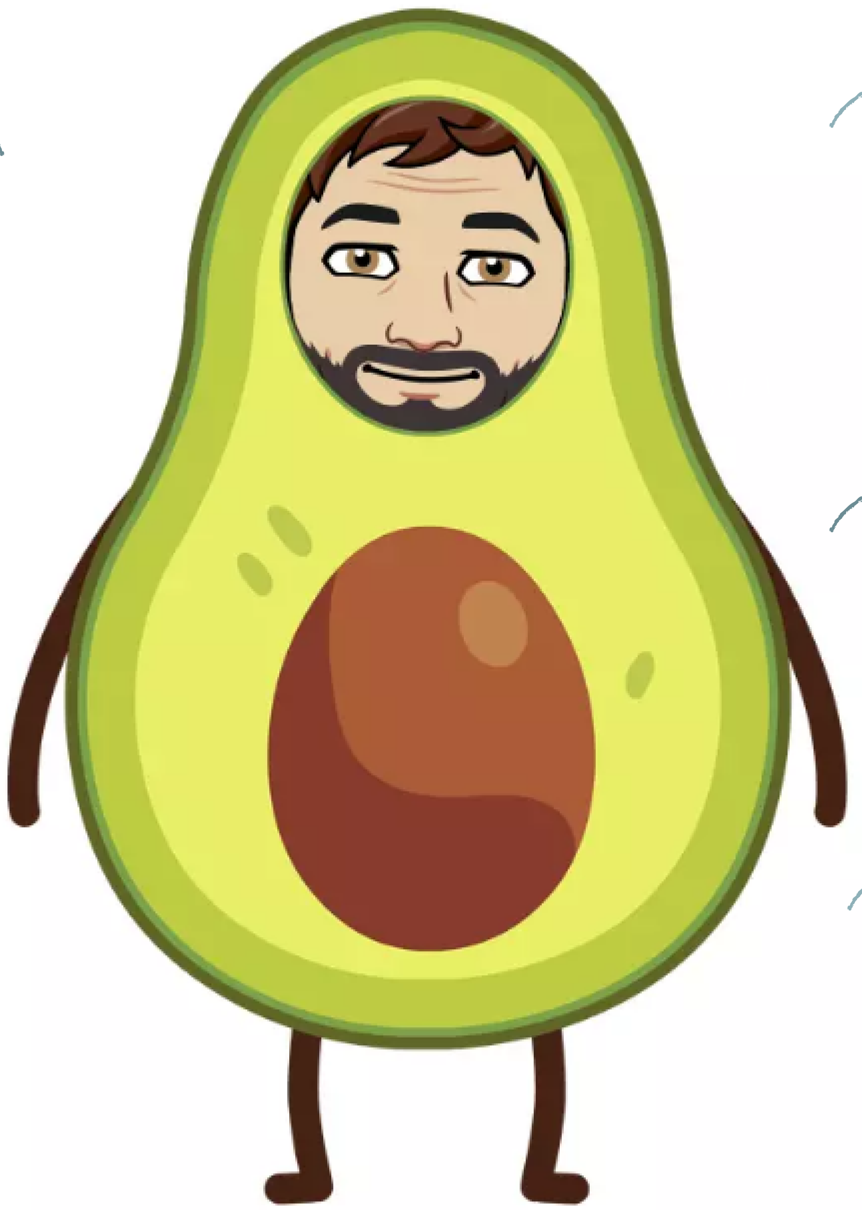


Le Développeur

Python

Concepts IA

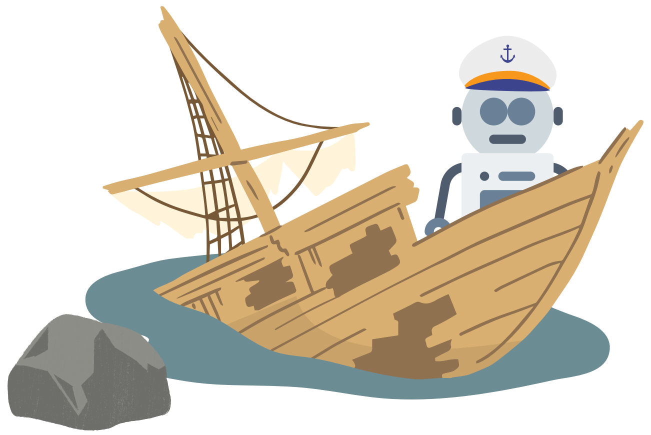
Développement



Docker

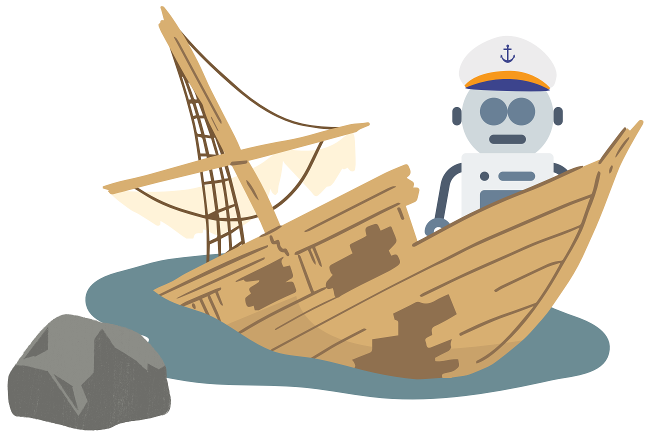
Frameworks

Outillage



Eléa Petton

Machine Learning Engineer @OVHcloud



@EleaPetton

eleapttn

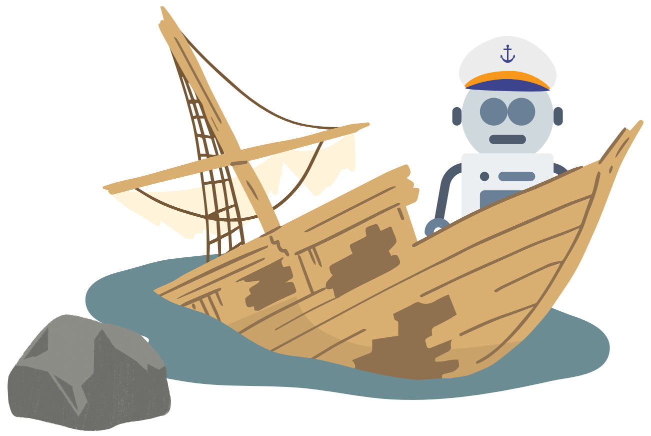
Eléa PETTON

Stéphane Philippart

DevRel@OVHcloud

Co-Créateur de TADx

https://philippart-s.github.io/blog

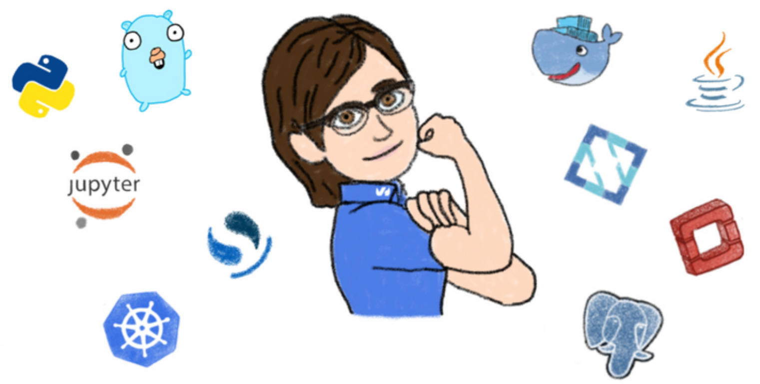


@wildagsx

philippart-s

philippartstephane

The Cloud with
European values



OVHcloud

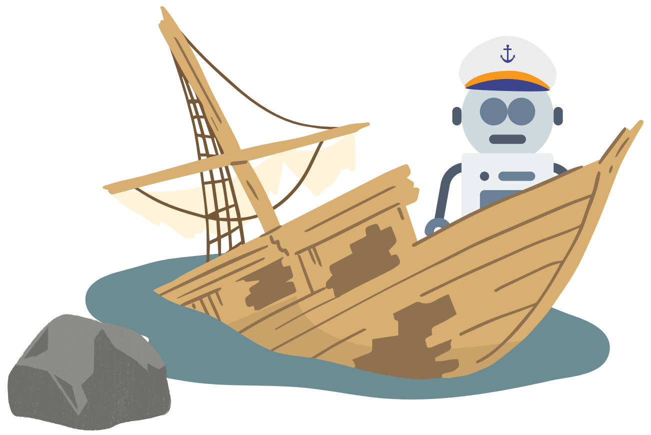
- Bare Metal Cloud
- Hosted Private Cloud
- Public Cloud
- Web Cloud
- Télécom

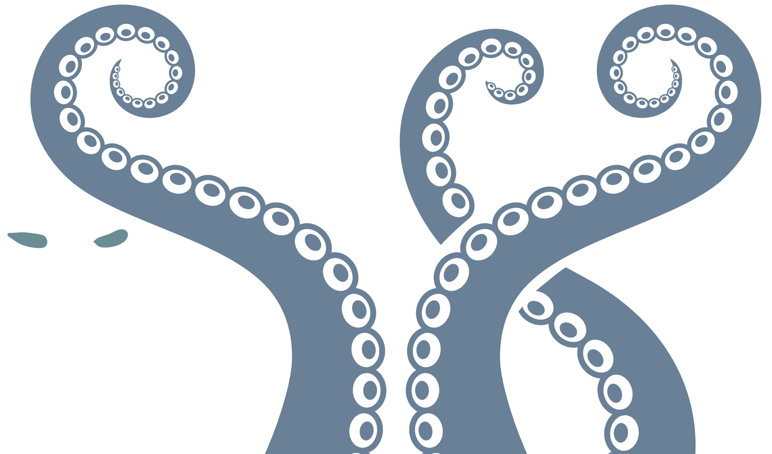
AI Notebooks

AI Training

AI Deploy

GPU / CPU



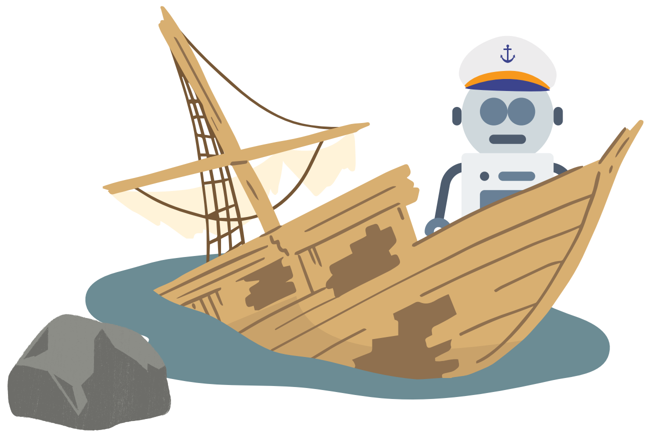


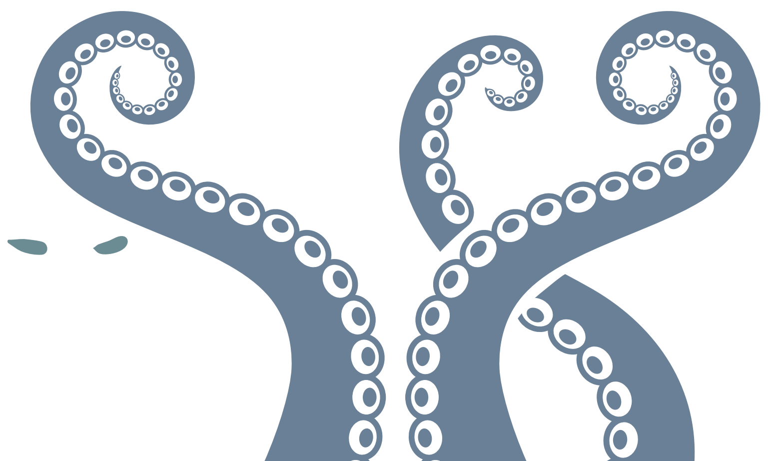
MERCI !

HTTPS://OVH.TO/8sTVj1C

Pour tester

- Le notebook : https://github.com/eleapttn/nemo-notebook
- L'app : https://github.com/philippart-s/nemo-streamlit-front
  https://github.com/philippart-s/nemo-flask-back
- Quelques docs : https://docs.ovh.com/gb/en/publiccloud/ai/
  https://blog.ovhcloud.com/tag/ai/
- Quelques Exemples : https://github.com/ovh/ai-training-examples/





Références

- **TensorFlow :** https://www.tensorflow.org/guide?hl=fr
- **JupyterLab :** https://jupyter.org/
- **Pandas :** https://pandas.pydata.org/
- **Sklearn :** https://scikit-learn.org/stable/
- **Numpy :** https://numpy.org/
- **Matplotlib :** https://matplotlib.org/
- **Dataset des sons de mammifères marins :** https://www.kaggle.com/datasets/shreyj1729/best-of-watkins-marine-mammal-sound-database
- **Informations sur la dataset :** https://cis.whoi.edu/science/B/whalesounds/index.cfm
- **Classification de musique avec un CNN :** https://blog.clairvoyantsoft.com/music-genre-classification-using-cnn-ef9461553726
- **Classification des genre musicaux :** https://towardsdatascience.com/music-genre-classification-with-python-c714d032f0d8
- **Pré-traitement des données en Machine Learning :** https://towardsdatascience.com/introduction-to-data-preprocessing-in-machine-learning-a9fa83a5dc9d