

Jug Summer Camp

-enjoy it-



SI ON AIDAIT LE CAPITAINE
NÉMO À CLASSIFIER LES
MONSTRES MARINS ?

Eléa Petton & Stéphane Philippart



SERLI

OVHcloud

GitHub

Doctolib

Clever Connect

sonar

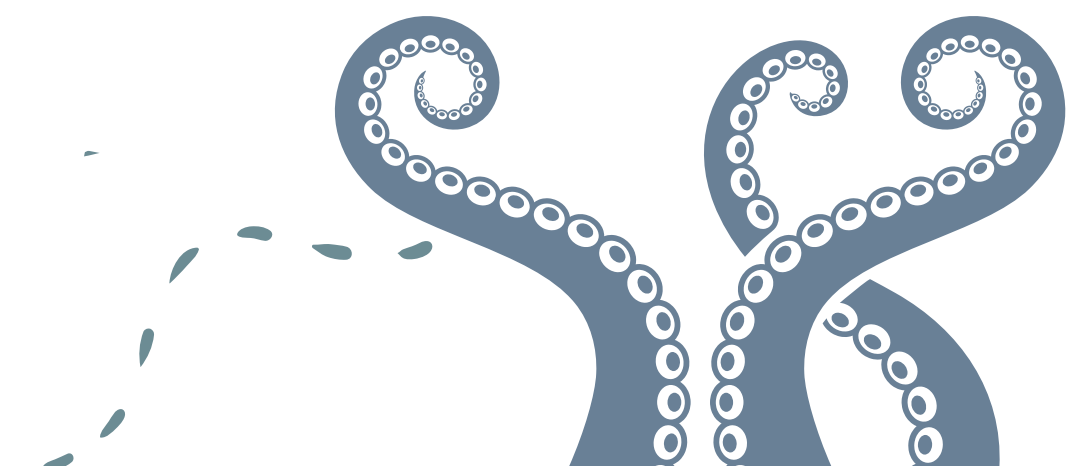
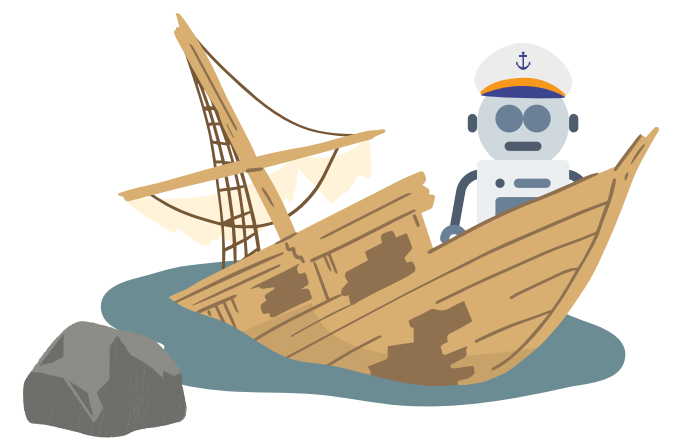
elastic

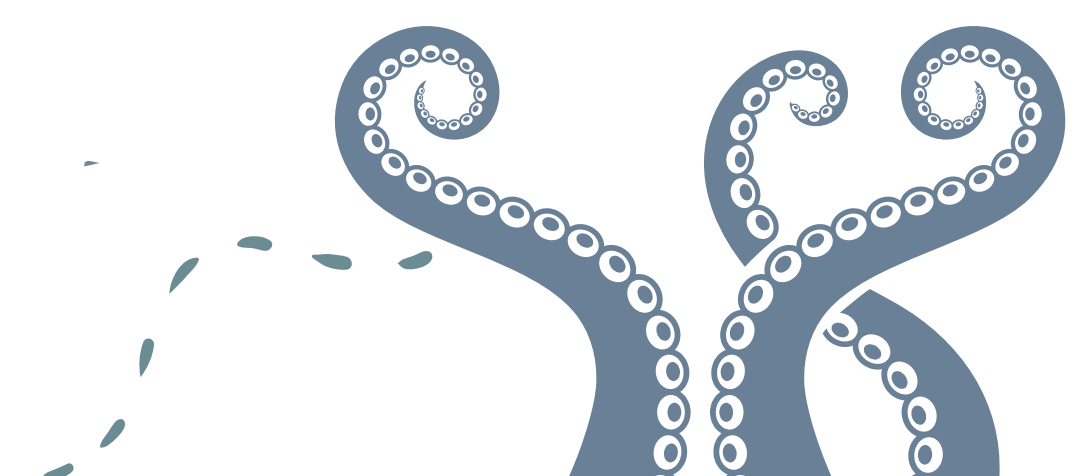
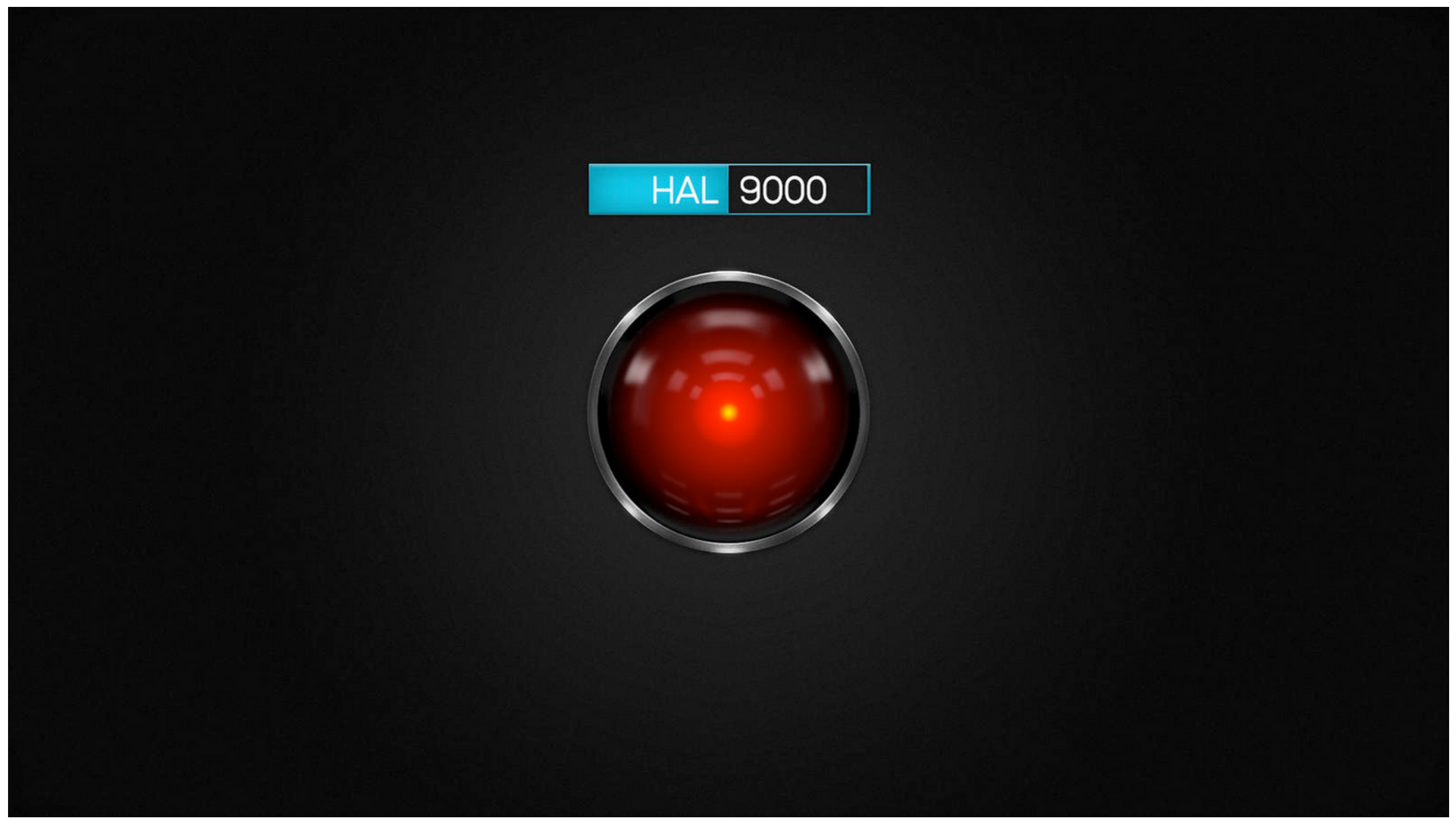
DARVA

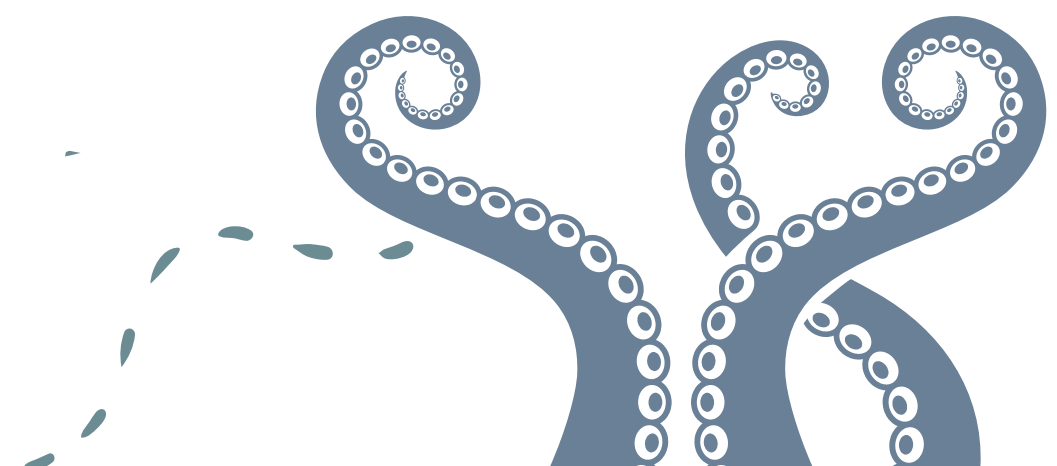
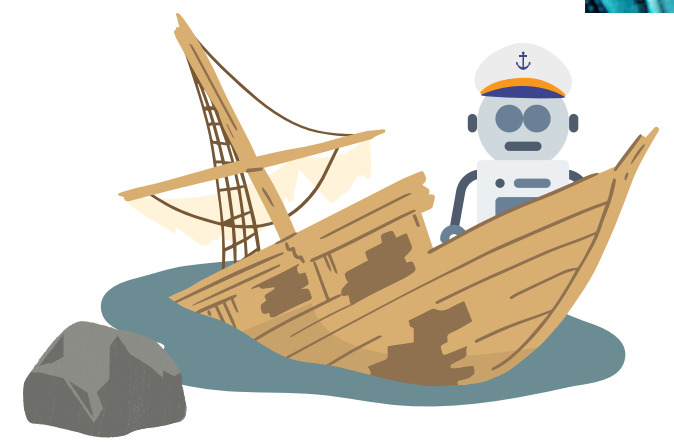
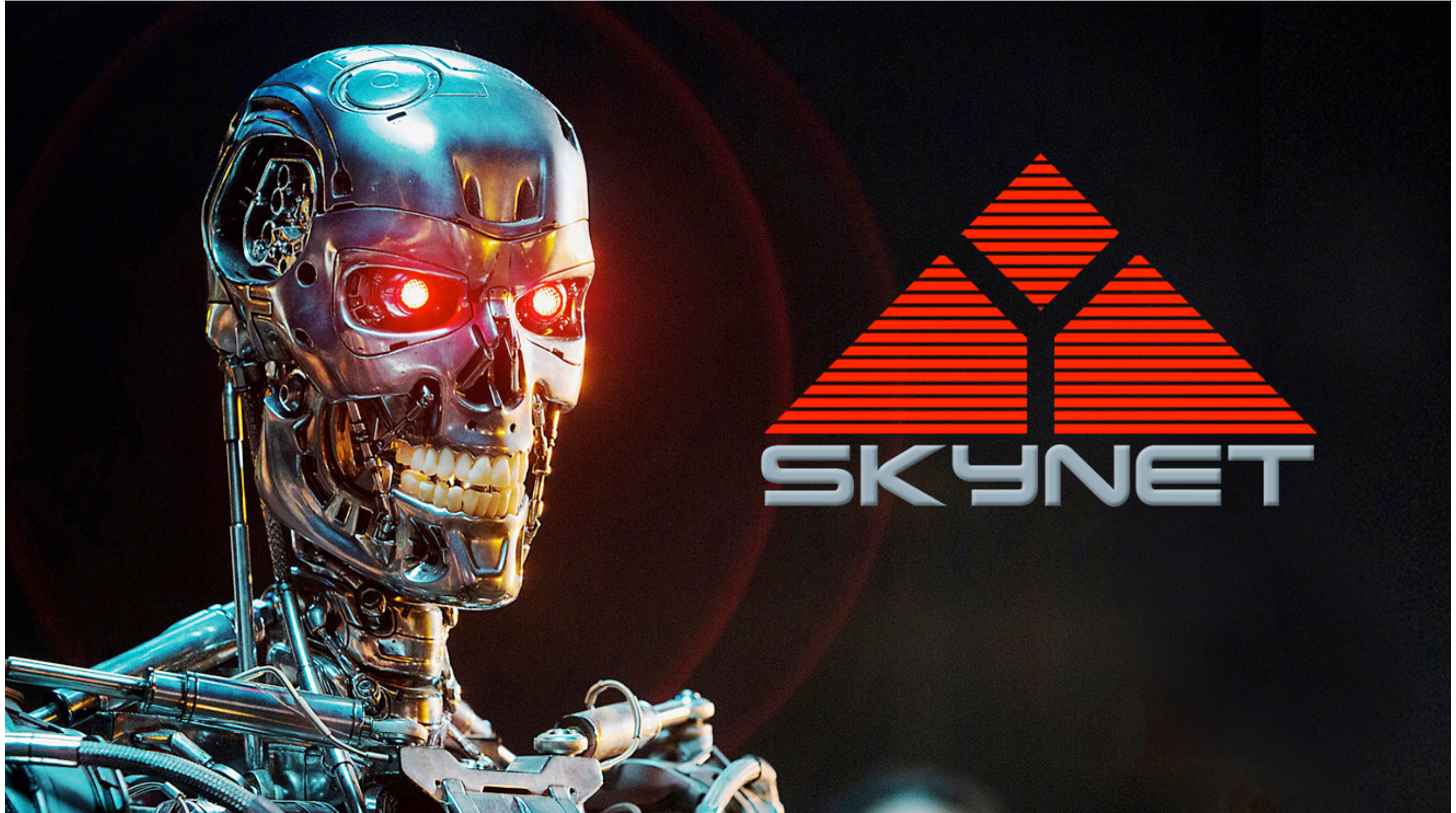




MAIS L'INTELLIGENCE ARTIFICIELLE C'EST QUOI AU FINAL ?







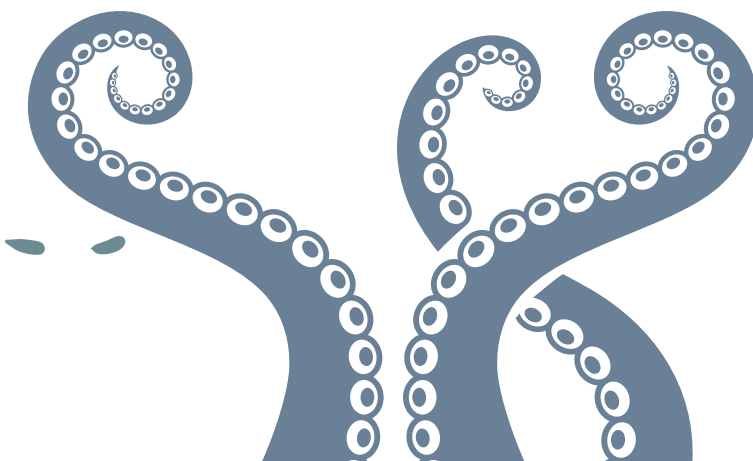
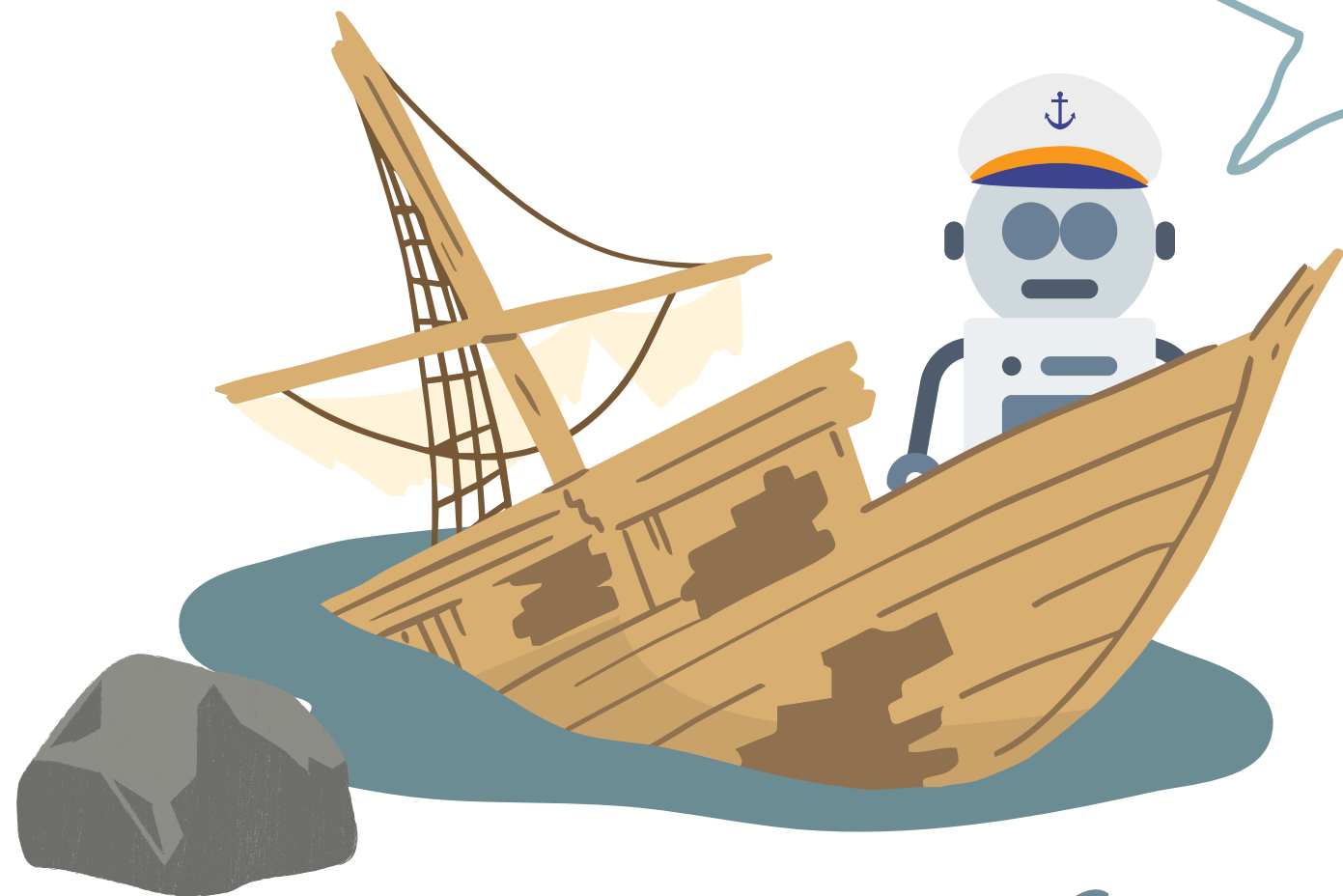


QU'EST-CE QUE L'IA ?

Toi humain tu sais différencier les animaux marins.

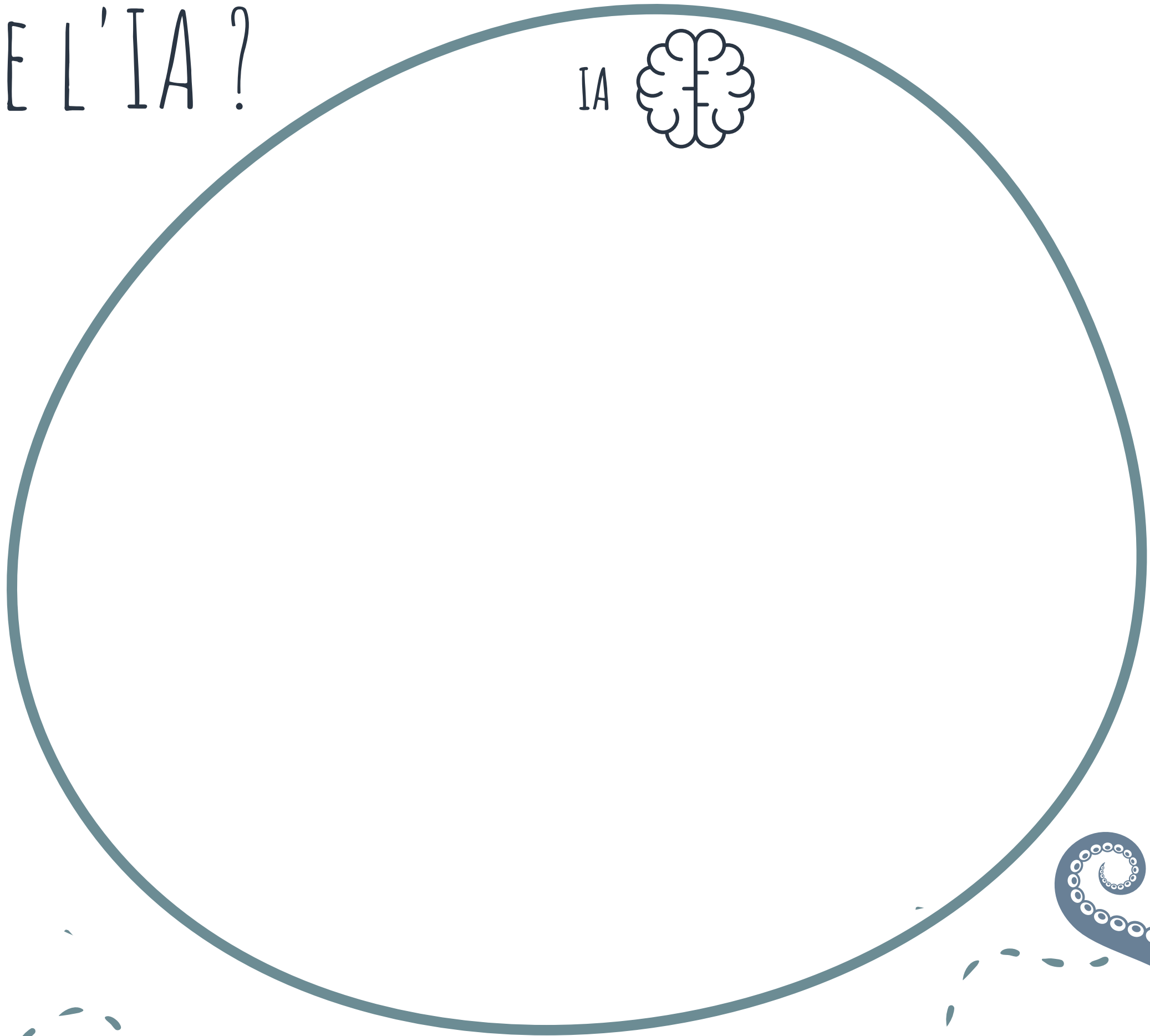
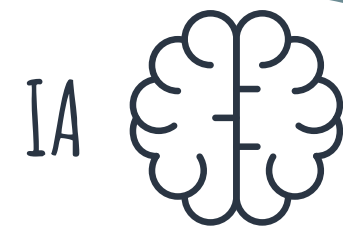
Comment faire pour l'apprendre ?

Attends, je vais te donner un algorithme qui va te permettre d'apprendre.

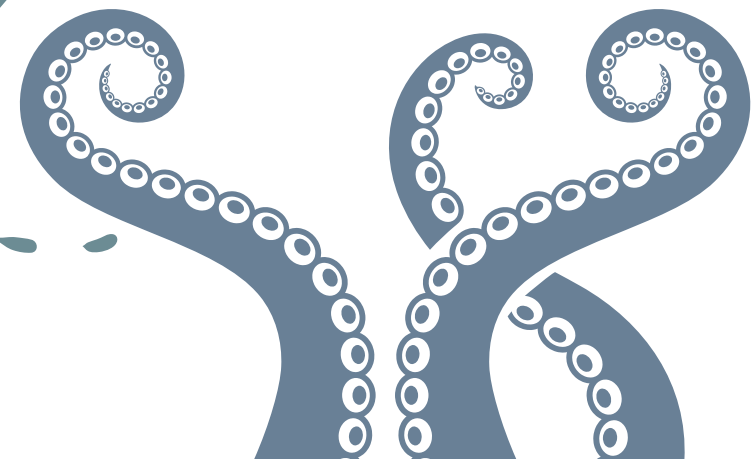
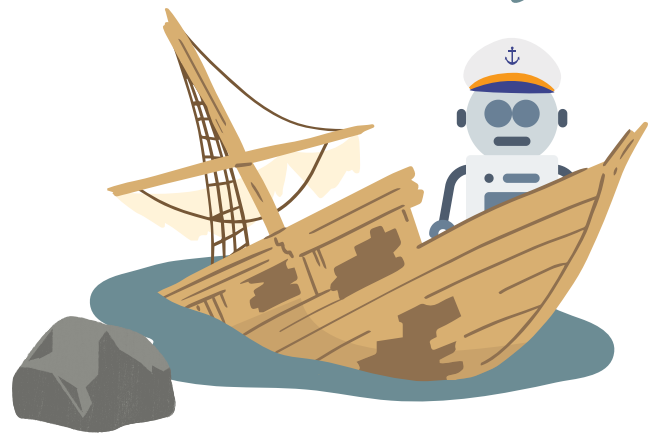




QU'EST-CE QUE L'IA ?

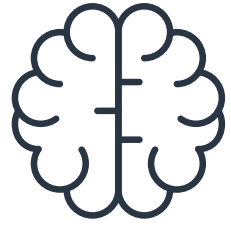


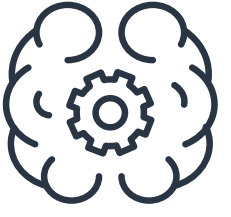
Je me demande
quelle est la
différence entre
tous ces termes ?



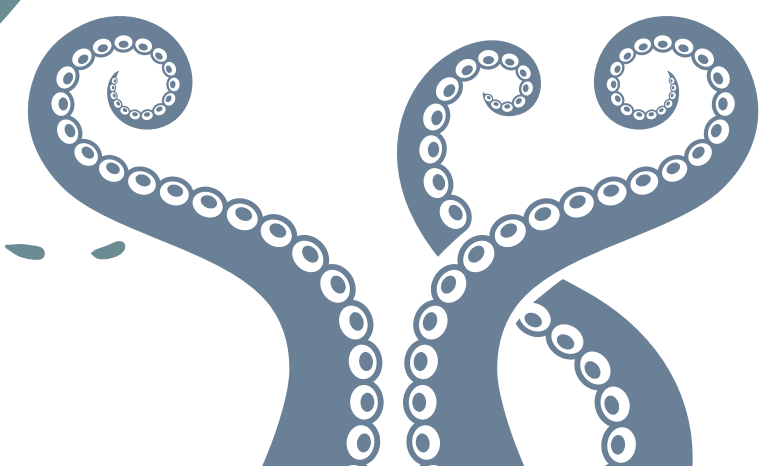
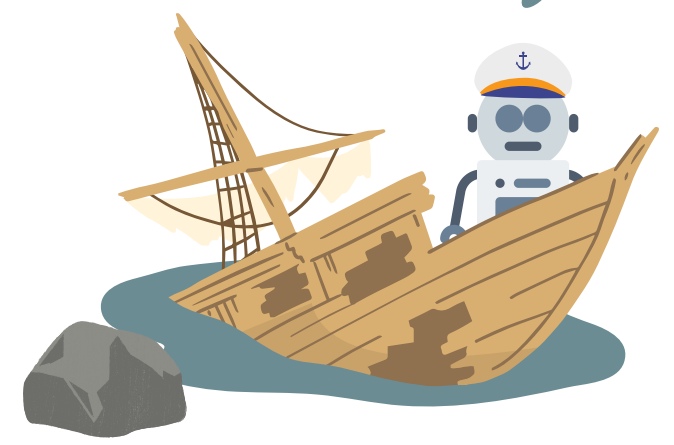


QU'EST-CE QUE L'IA ?

IA 

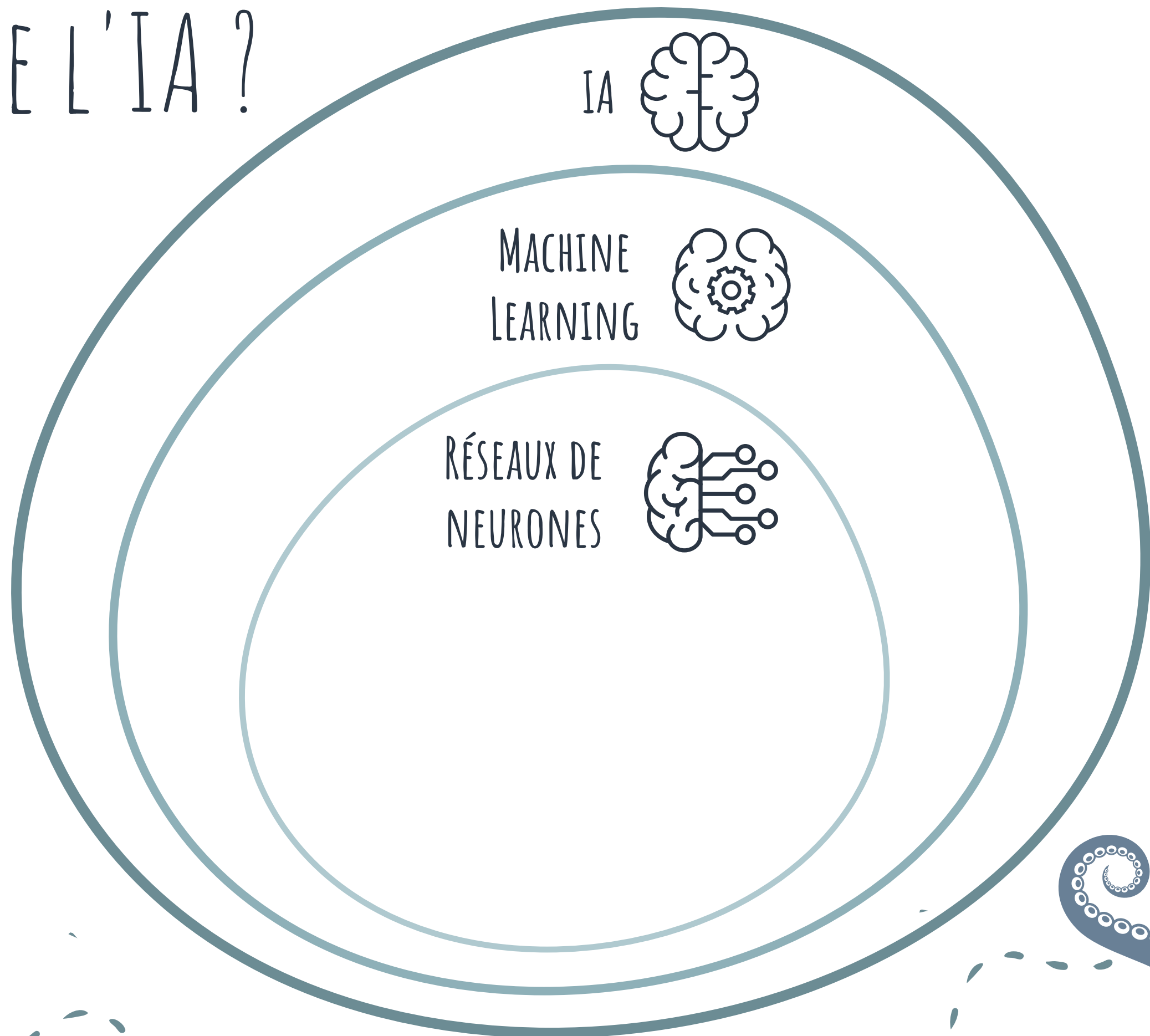
MACHINE
LEARNING 

Je me demande
quelle est la
différence entre
tous ces termes ?

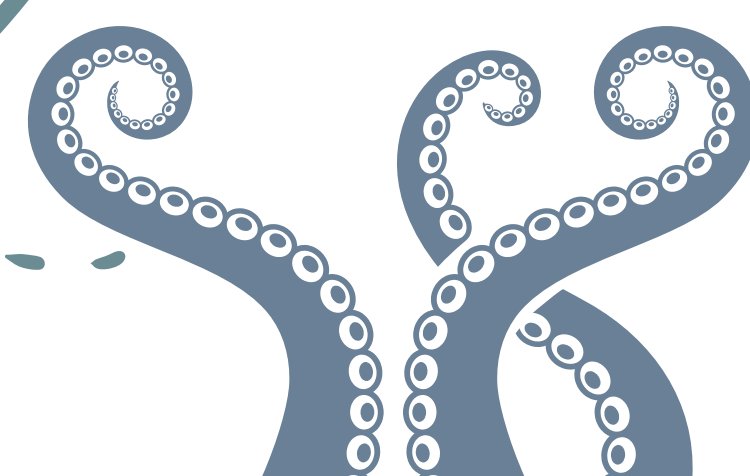
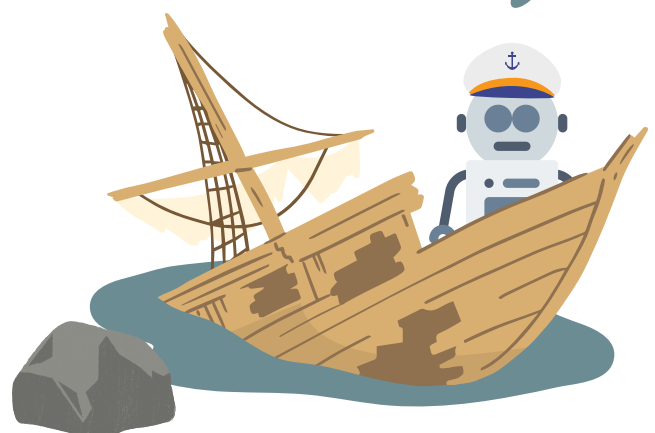




QU'EST-CE QUE L'IA ?



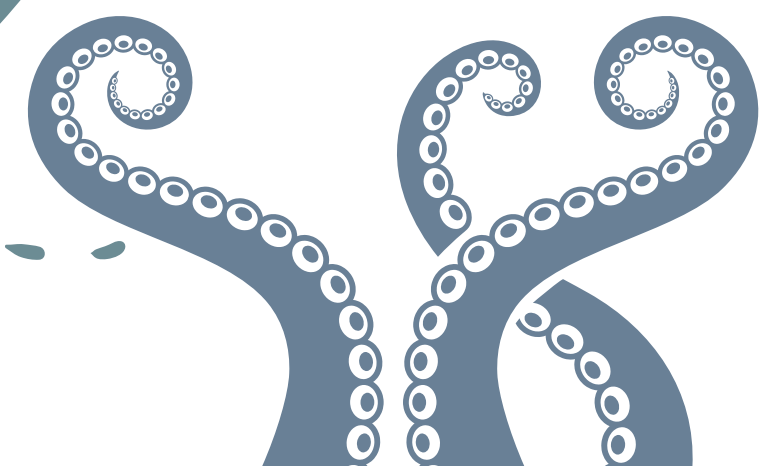
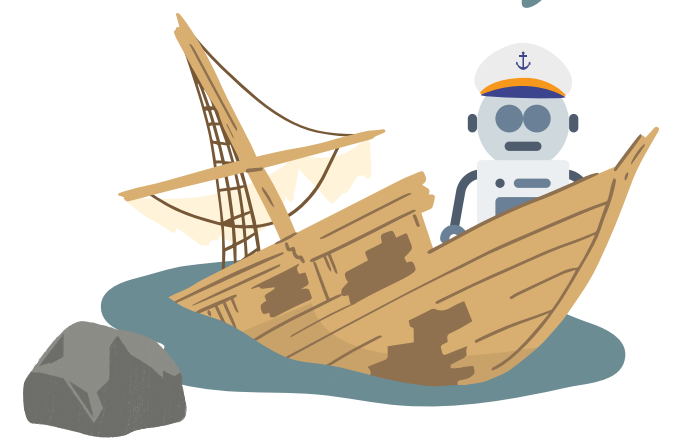
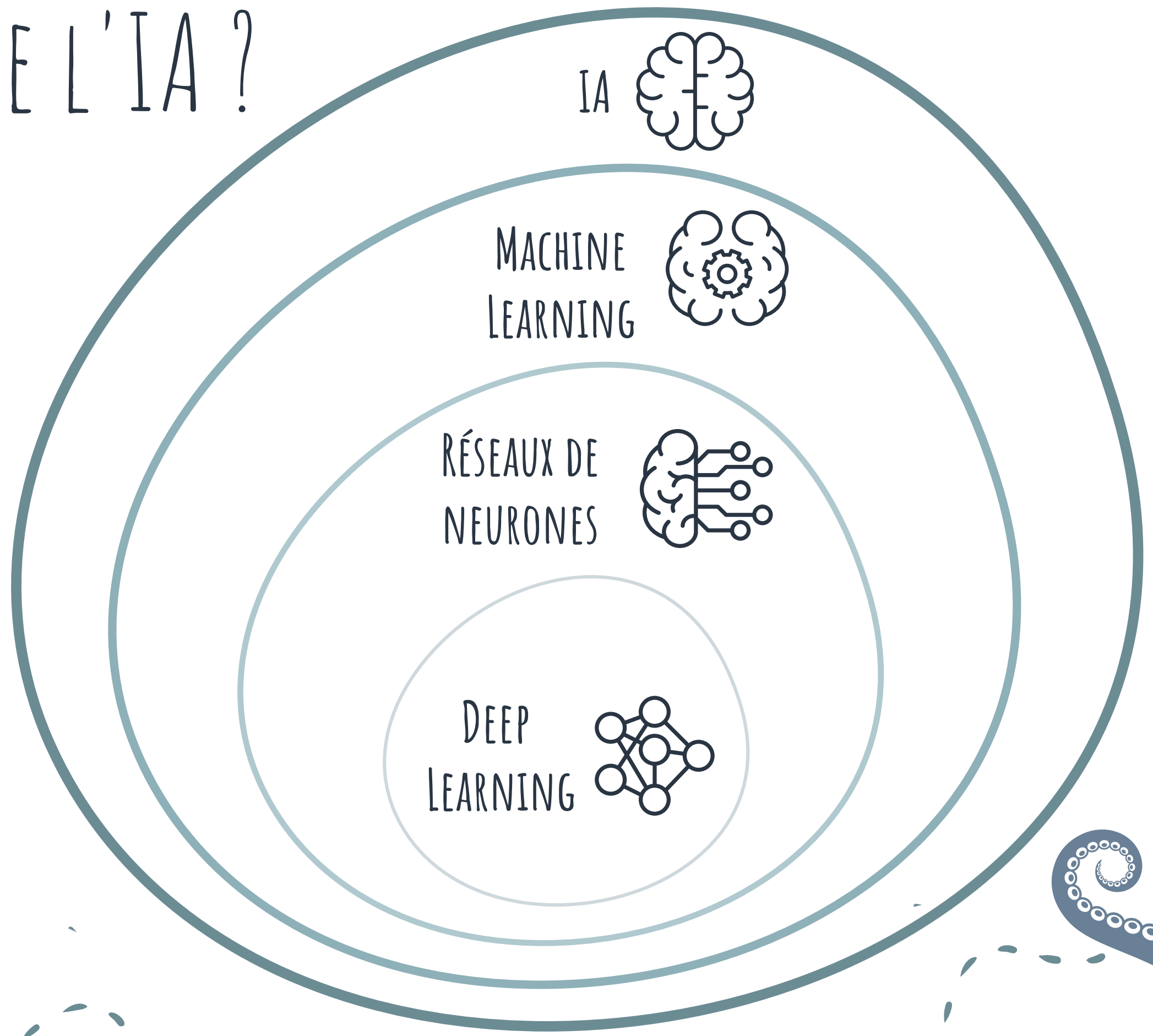
Je me demande
quelle est la
différence entre
tous ces termes ?





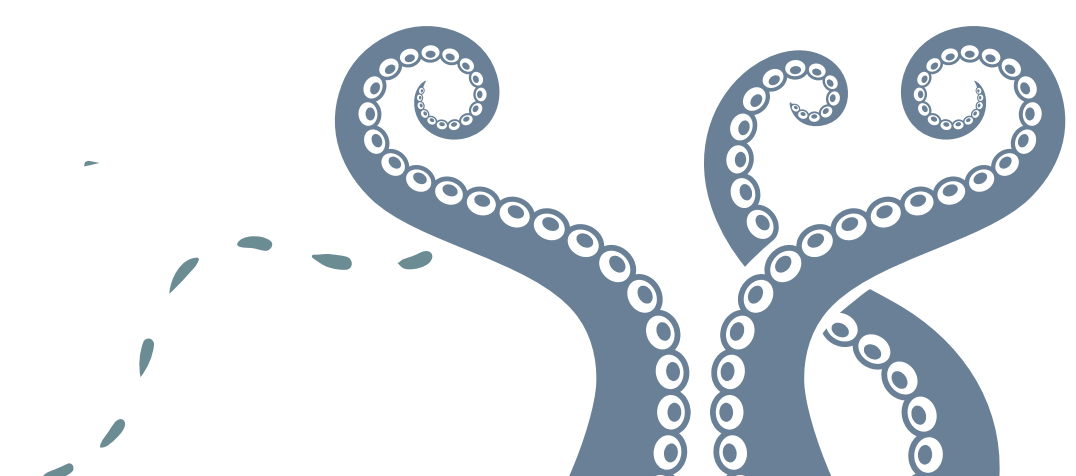
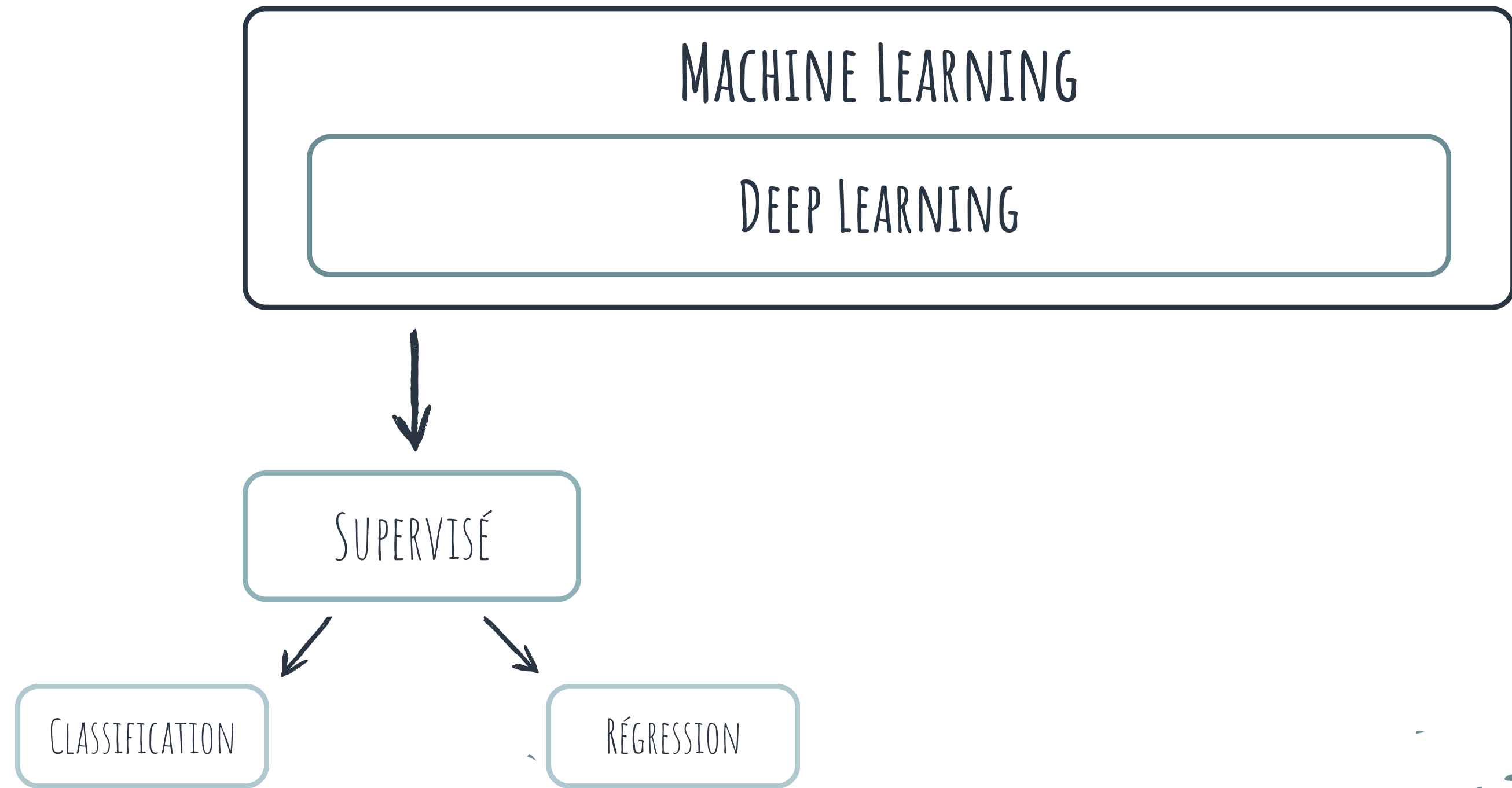
QU'EST-CE QUE L'IA ?

Je me demande
quelle est la
différence entre
tous ces termes ?



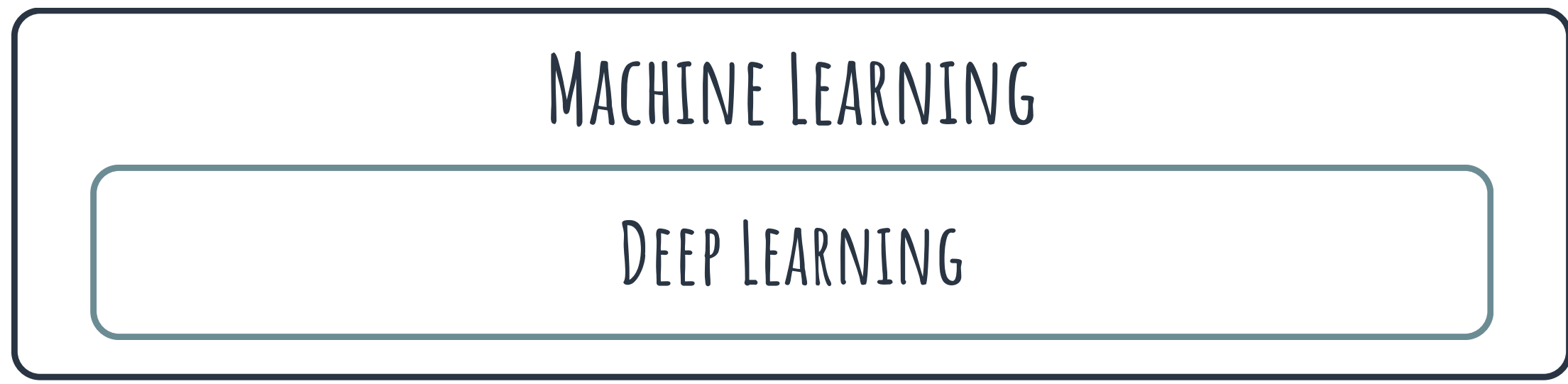


LE MACHINE LEARNING PLUS EN DÉTAIL





LE MACHINE LEARNING PLUS EN DÉTAIL

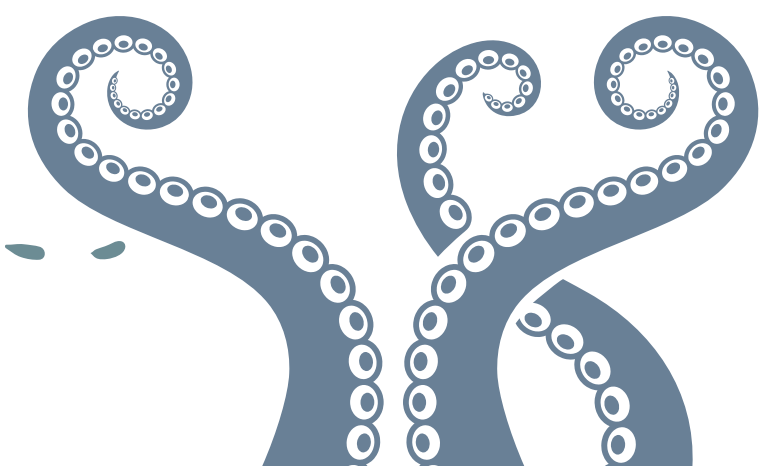


NON SUPERVISÉ



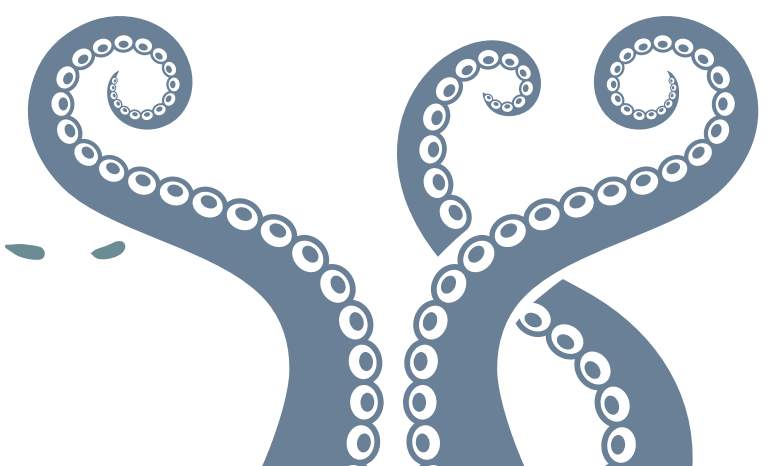
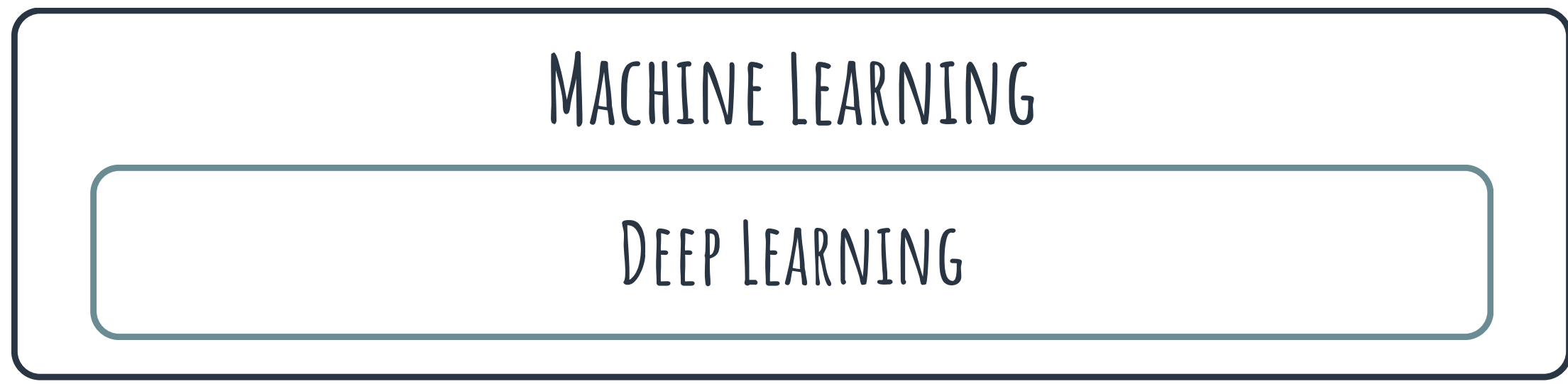
CLUSTERING

RÉDUCTION DE DIMENSION



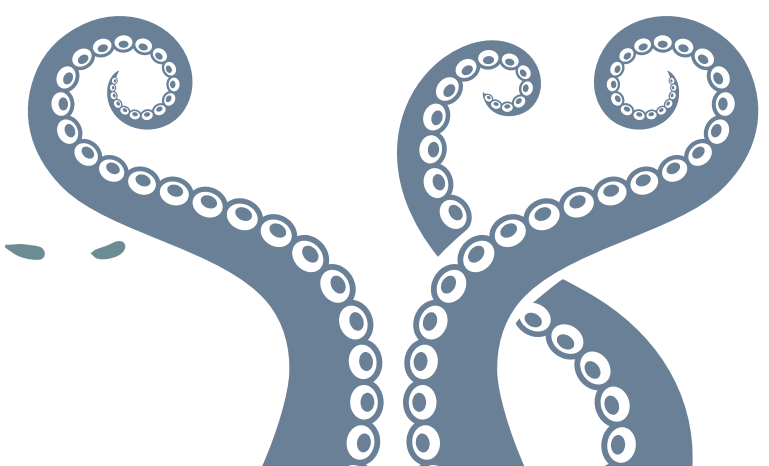
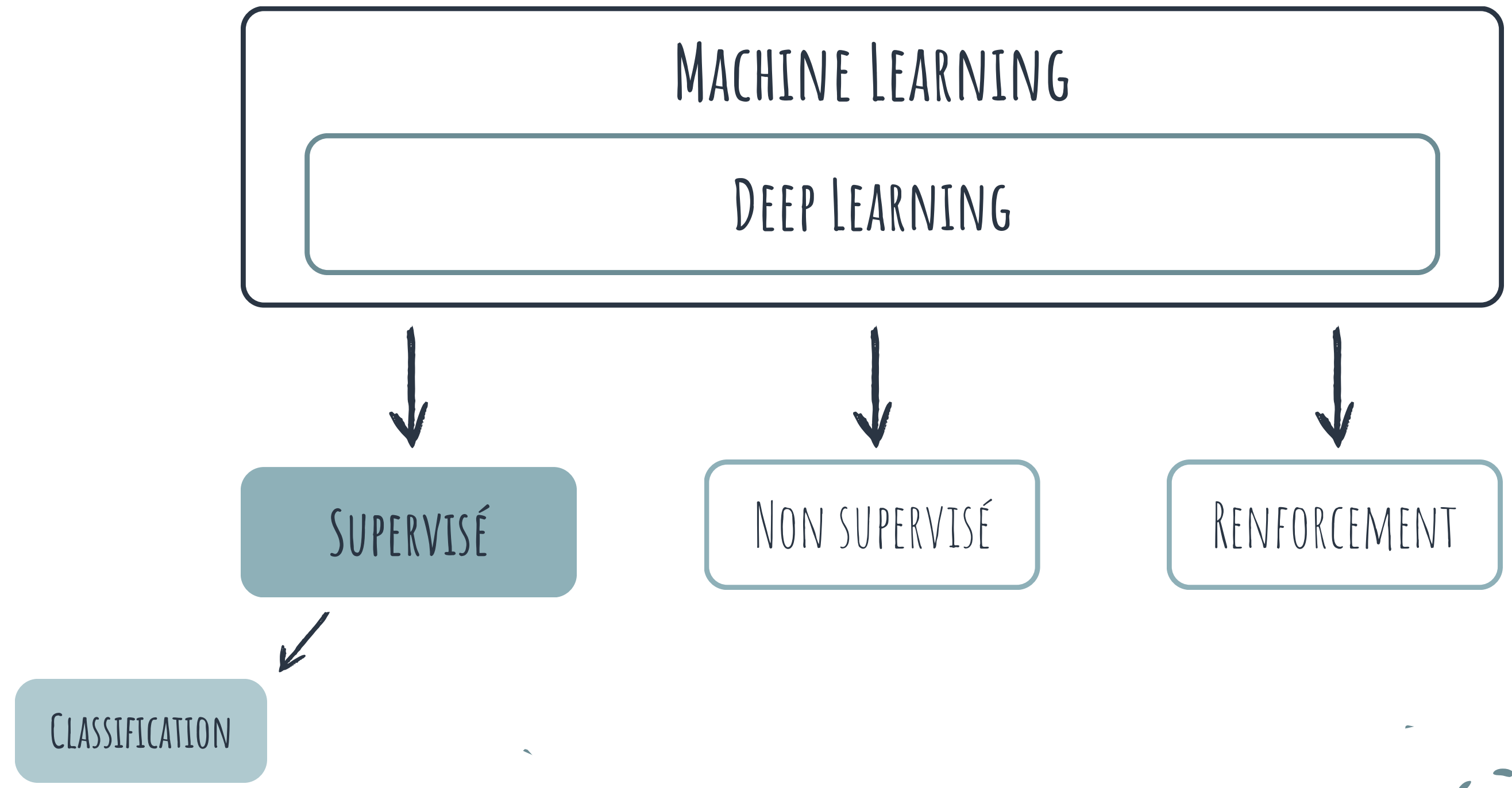


LE MACHINE LEARNING PLUS EN DÉTAIL





LE MACHINE LEARNING PLUS EN DÉTAIL

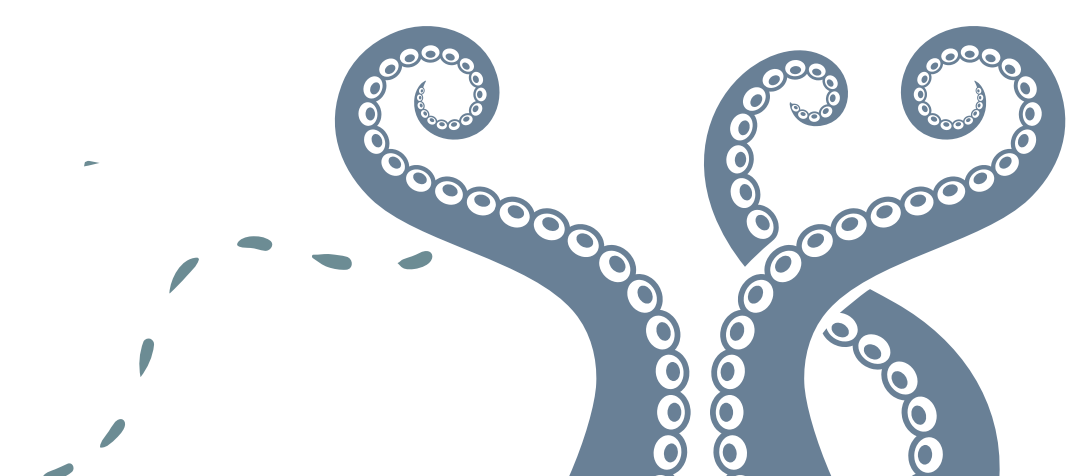
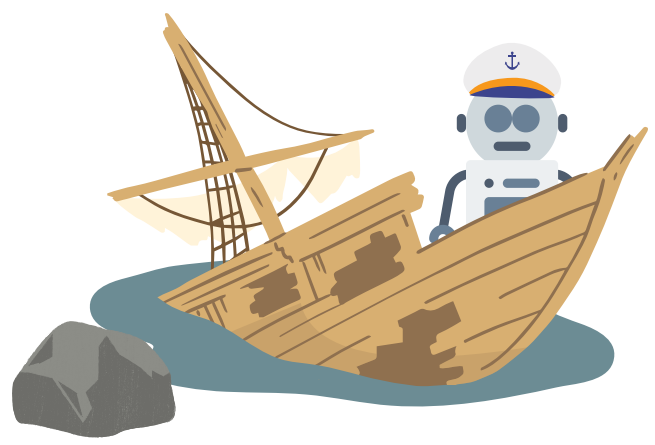




DIFFÉRENTES FAMILLES EN IA



NLP

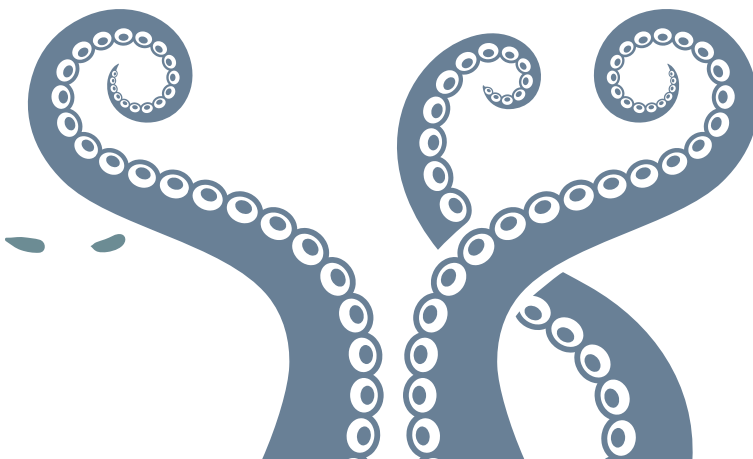
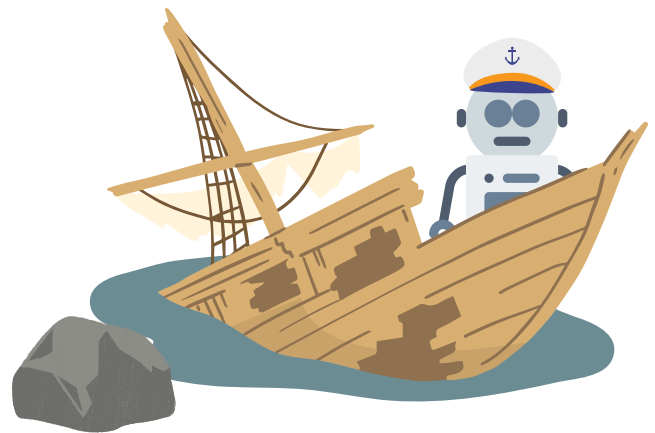




LE NLP, À QUOI ÇA SERT ?



Traduction





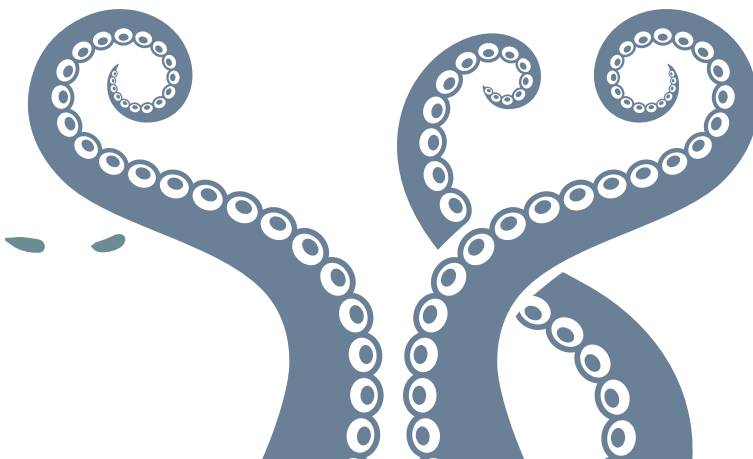
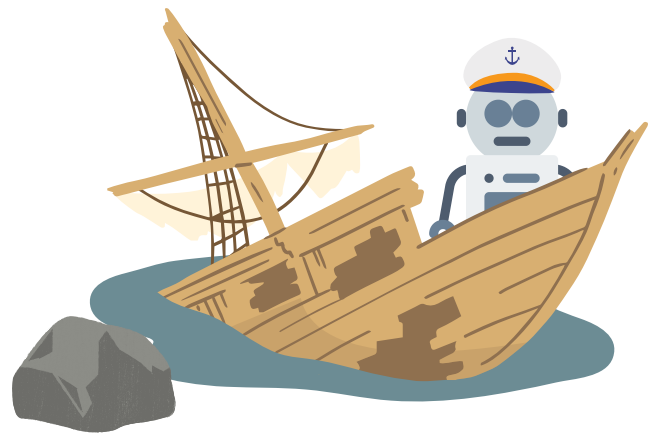
LE NLP, À QUOI ÇA SERT ?



Traduction



Classification de texte





LE NLP, À QUOI ÇA SERT ?



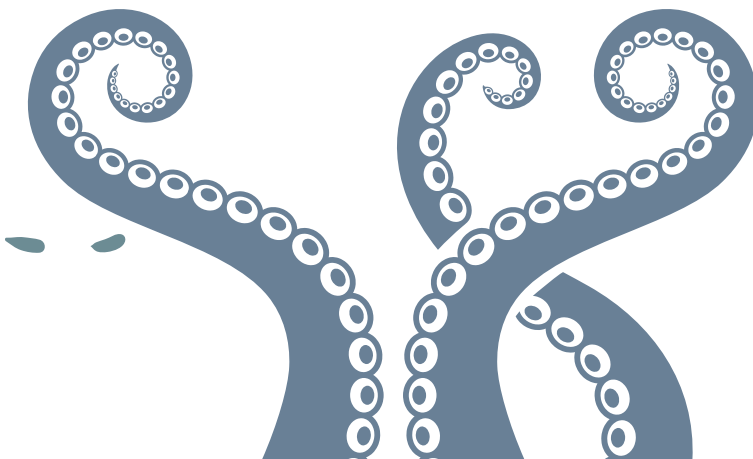
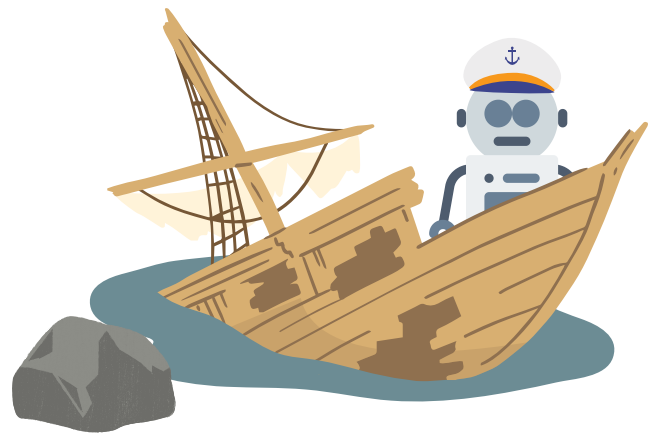
Traduction



Classification de texte



Résumés





LE NLP, À QUOI ÇA SERT ?



Traduction



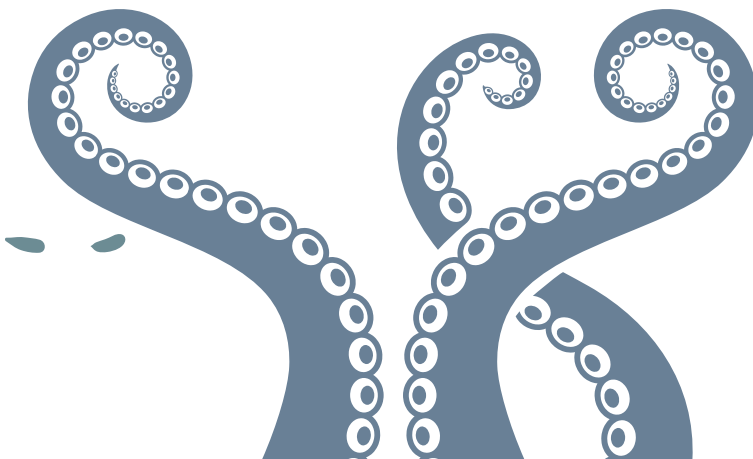
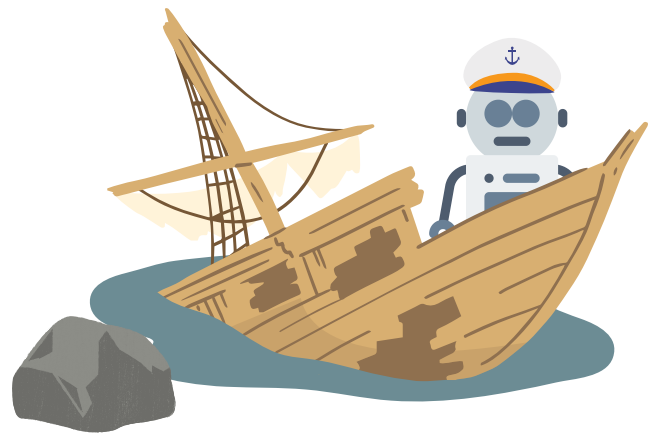
Classification de texte



Résumés



Génération de texte





LE NLP, À QUOI ÇA SERT ?



Traduction



Classification de texte



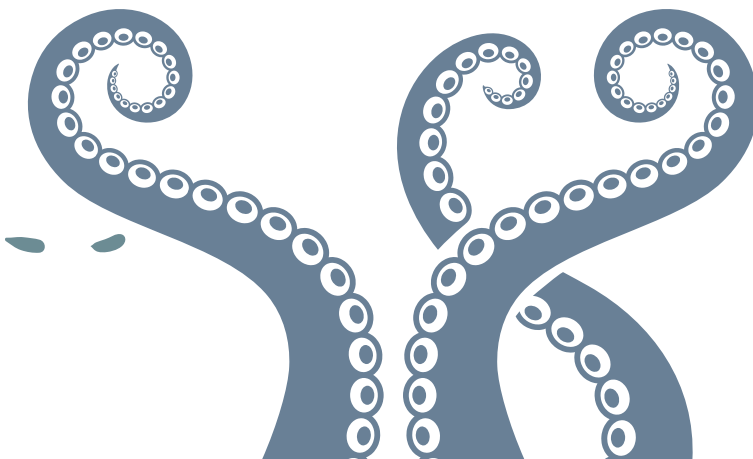
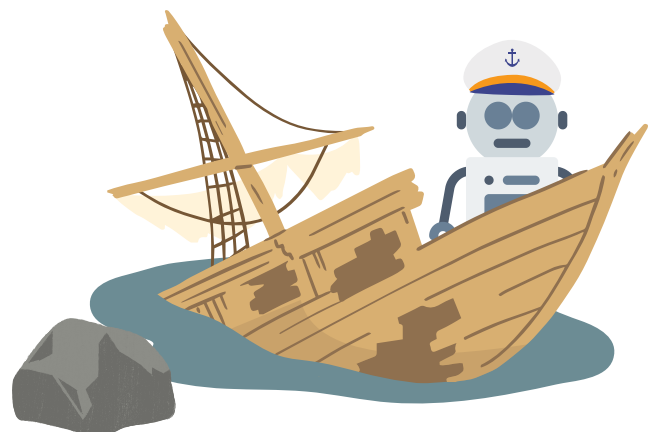
Résumés



Génération de texte



Réponses à des questions





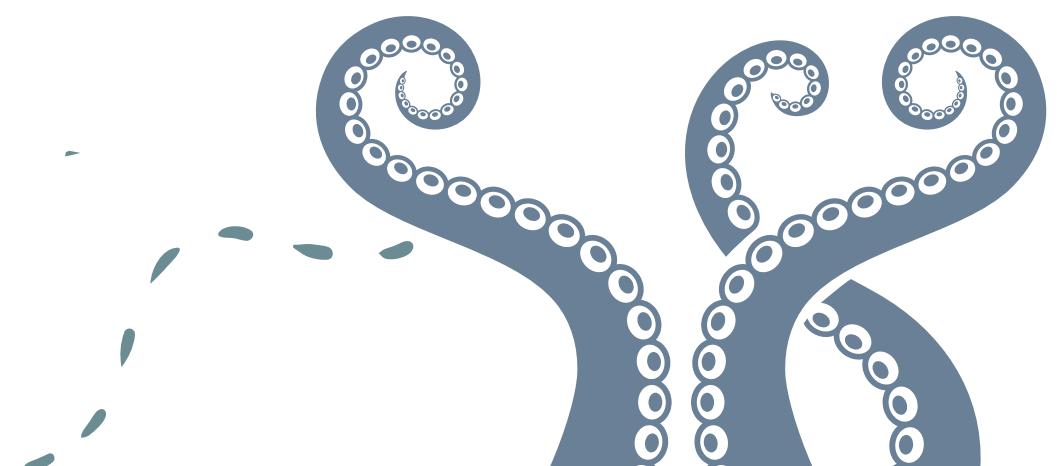
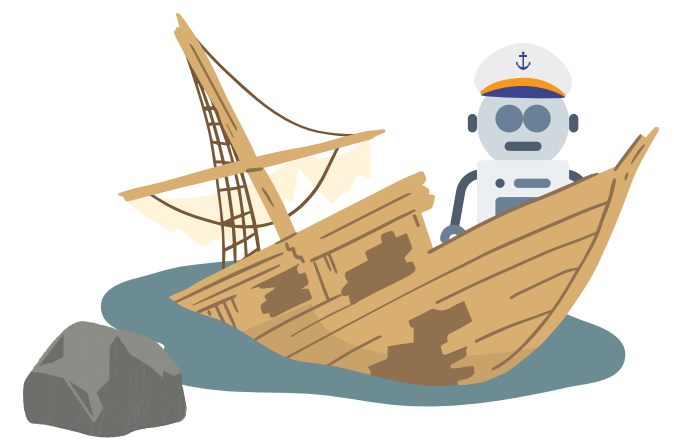
DIFFÉRENTES FAMILLES EN IA



NLP



Computer Vision





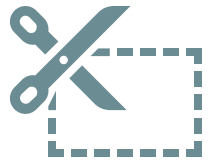
LA COMPUTER VISION, C'EST QUOI ?



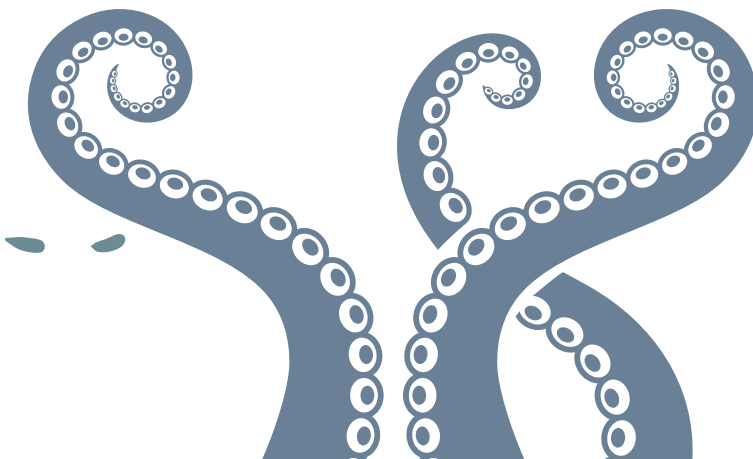
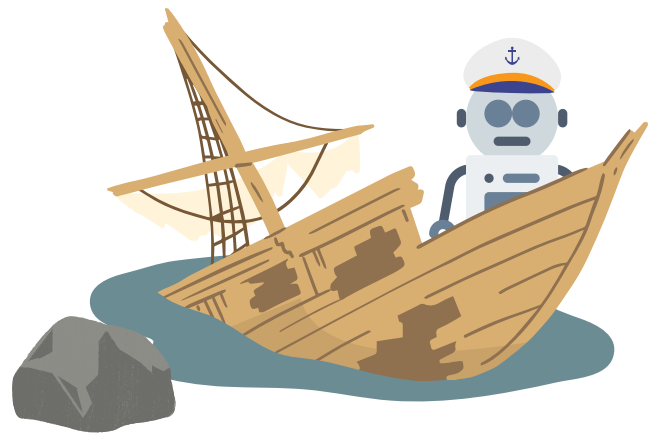
Classification d'images



Détection d'objets



Segmentation d'images





LA COMPUTER VISION, C'EST QUOI ?

Classification



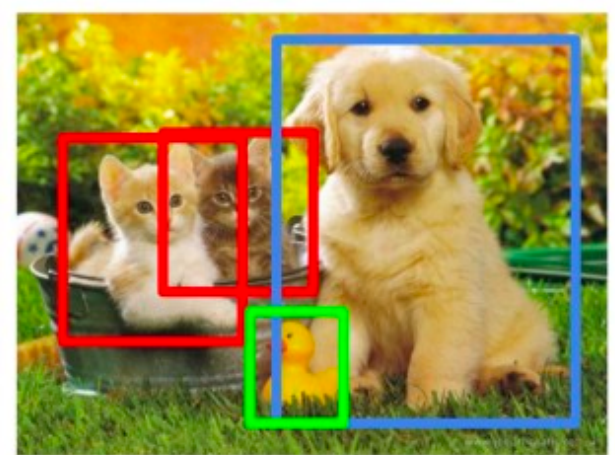
CAT

Classification + Localization



CAT

Object Detection



CAT, DOG, DUCK

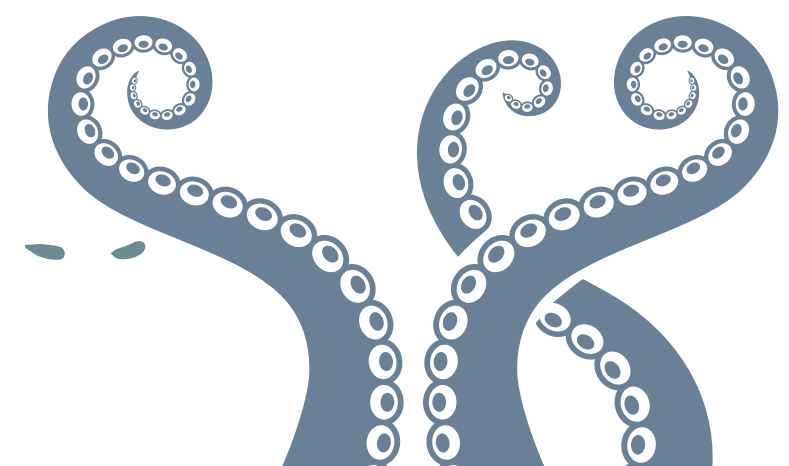
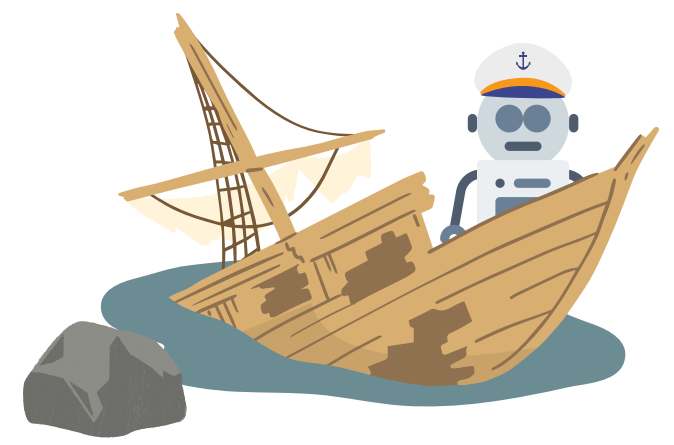
Instance Segmentation



CAT, DOG, DUCK

Single object

Multiple objects





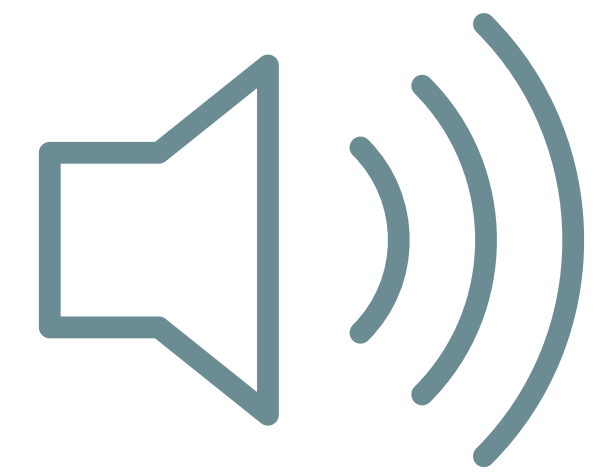
DIFFÉRENTES FAMILLES EN IA



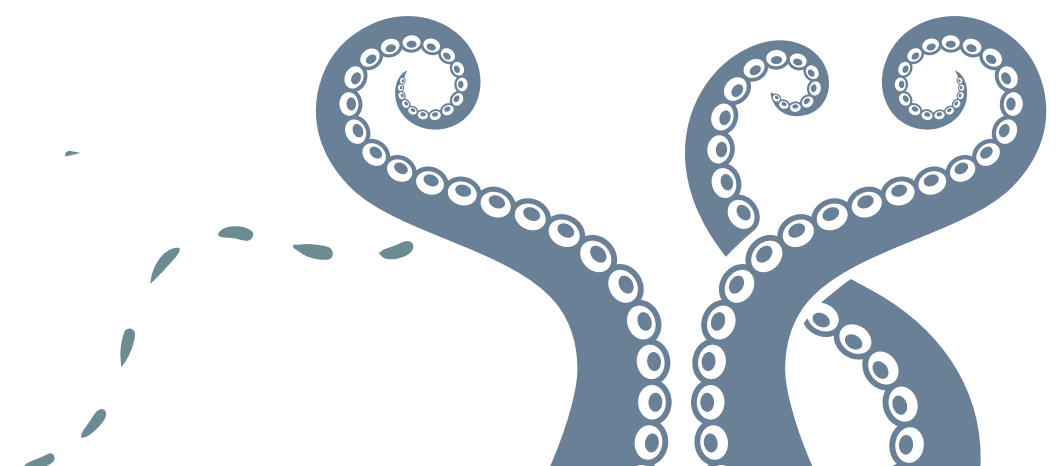
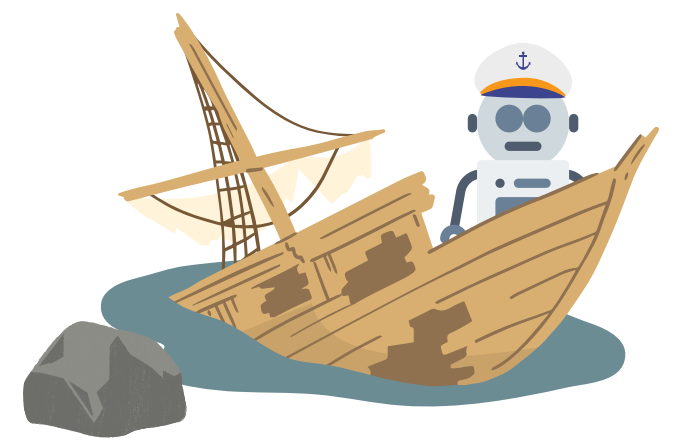
NLP



Computer Vision



Audio

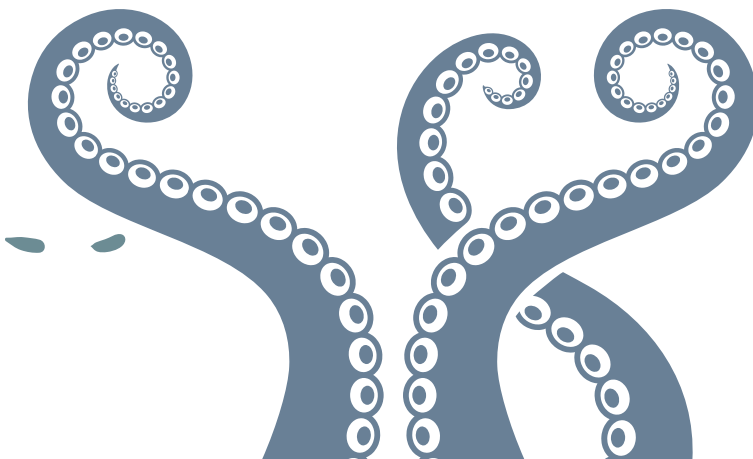
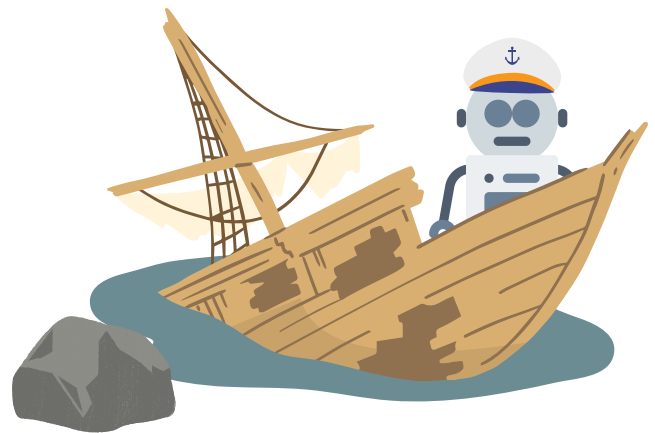




POURQUOI LE SON EN IA ?



Classification des données acoustiques





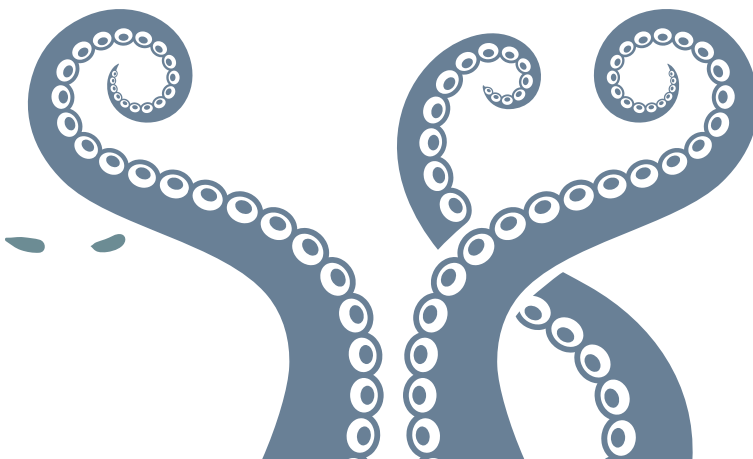
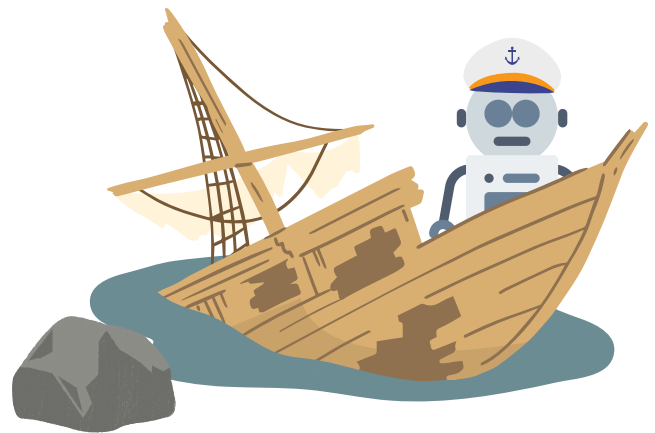
POURQUOI LE SON EN IA ?



Classification des données acoustiques



Classification des sons environnementaux





POURQUOI LE SON EN IA ?



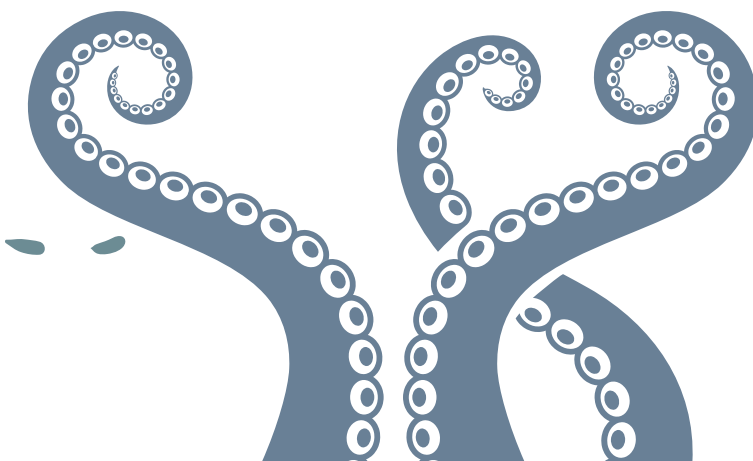
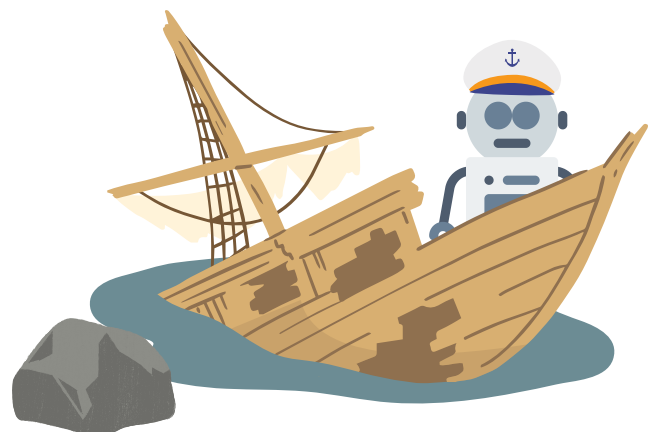
Classification des données acoustiques



Classification des sons environnementaux



Classification du langage naturel





POURQUOI LE SON EN IA ?



Classification des données acoustiques



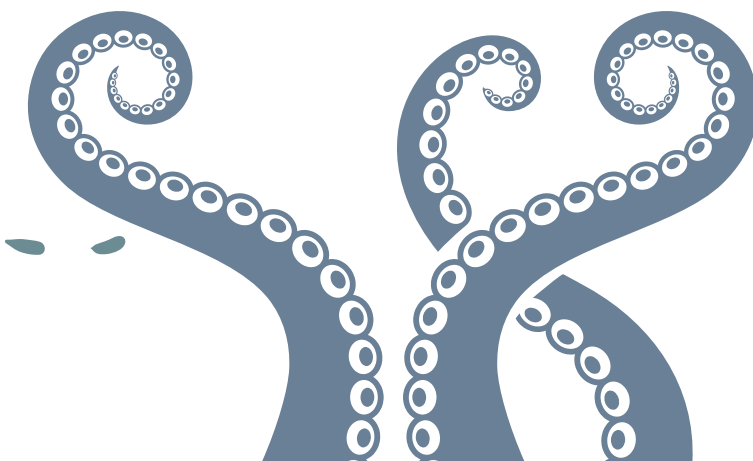
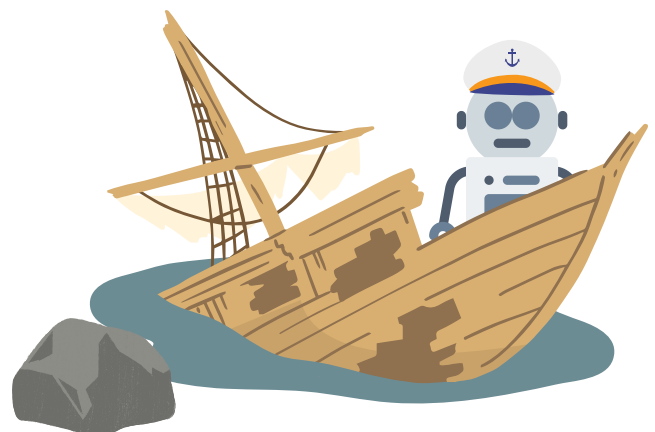
Classification des sons environnementaux



Classification du langage naturel



Classification de la musique

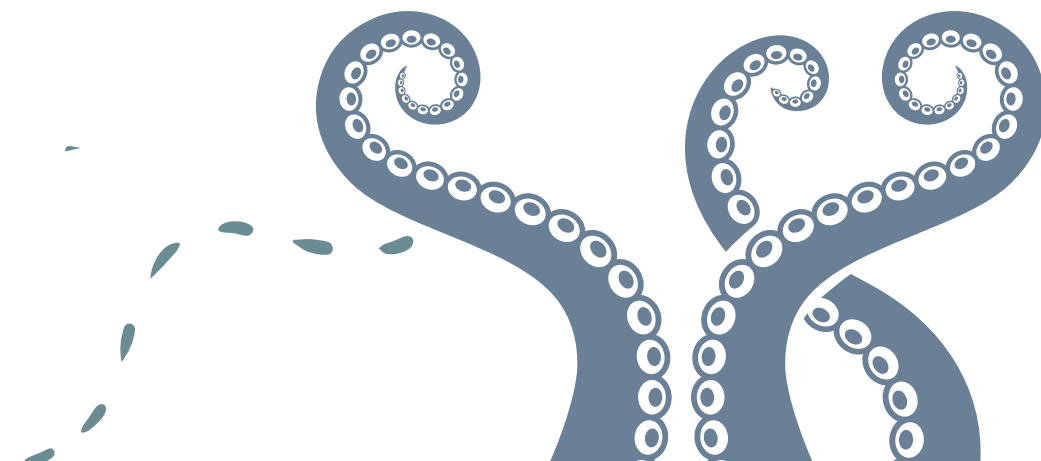
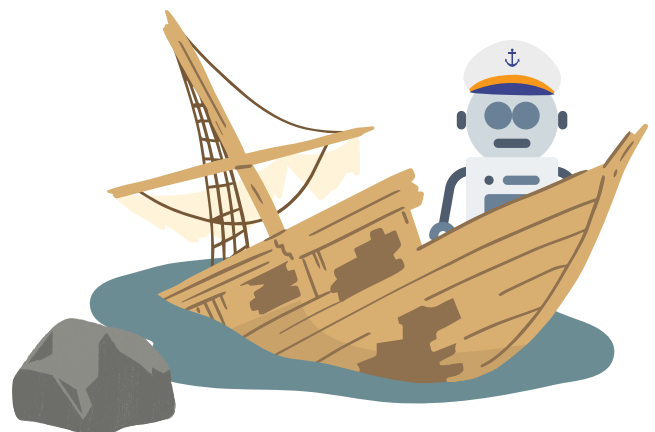




COMMENT CRÉER UNE APPLICATION INTELLIGENTE ?



Obtenir de la donnée





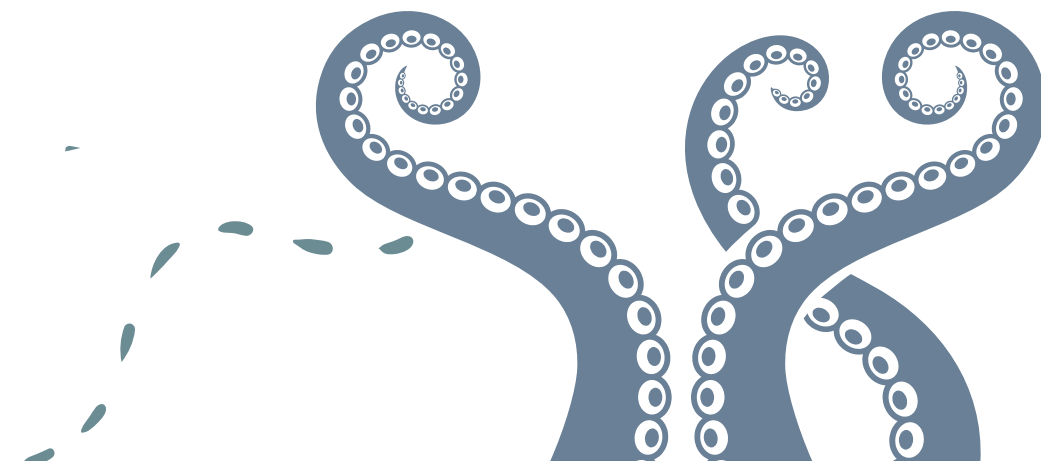
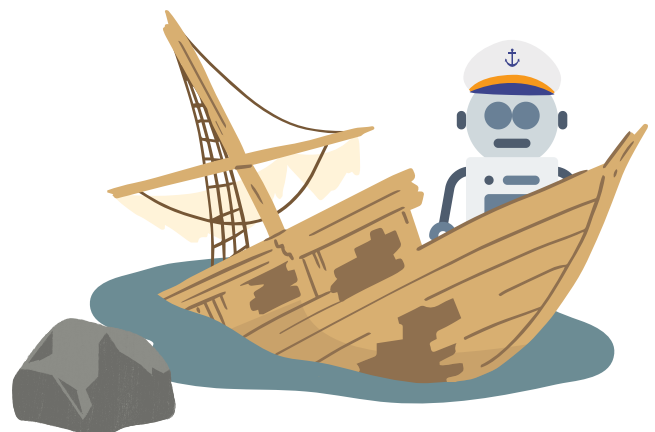
COMMENT CRÉER UNE APPLICATION INTELLIGENTE ?



Obtenir de la donnée



Traiter la donnée





COMMENT CRÉER UNE APPLICATION INTELLIGENTE ?



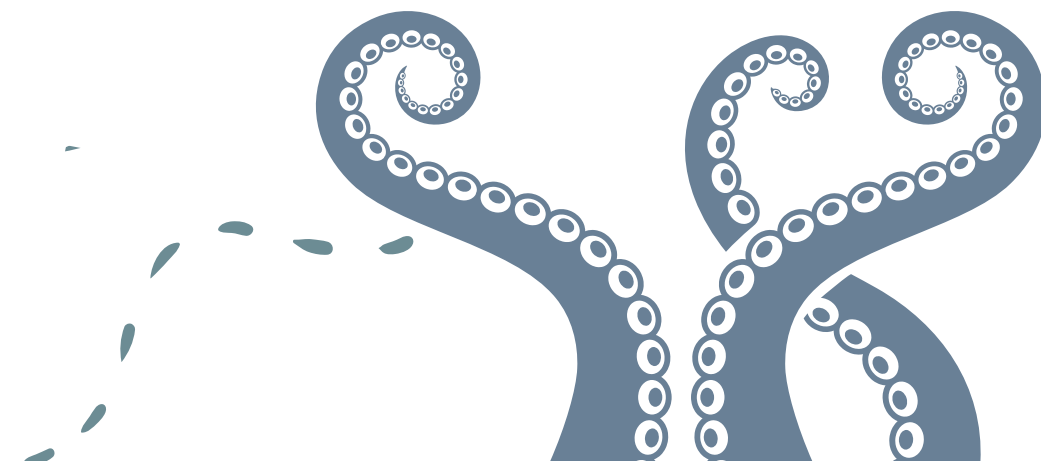
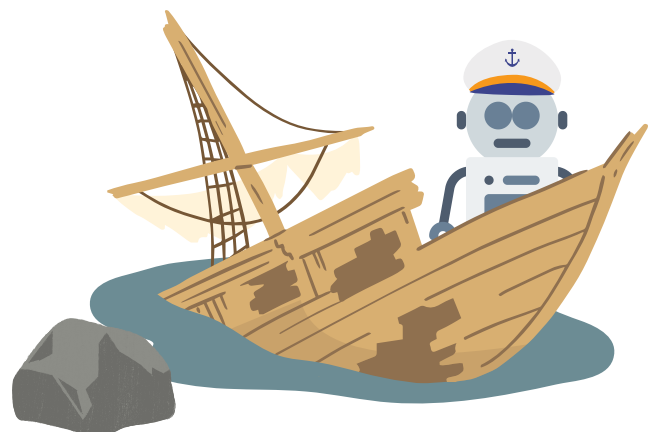
Obtenir de la donnée



Traiter la donnée



Créer ou optimiser un modèle existant





COMMENT CRÉER UNE APPLICATION INTELLIGENTE ?



Obtenir de la donnée



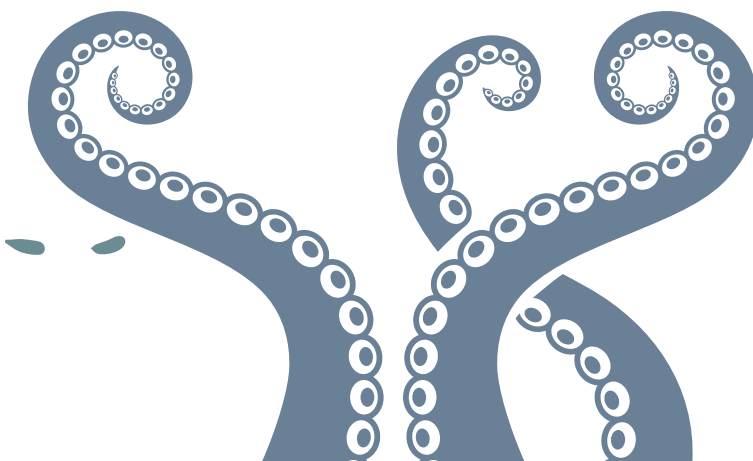
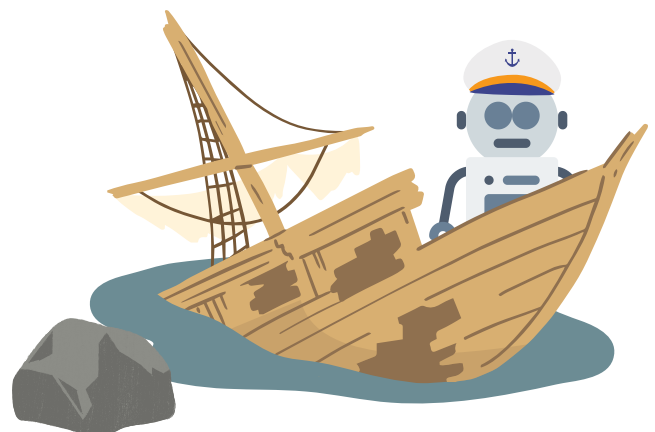
Traiter la donnée



Créer ou optimiser un modèle existant



Entraîner son modèle





COMMENT CRÉER UNE APPLICATION INTELLIGENTE ?



Obtenir de la donnée



Traiter la donnée



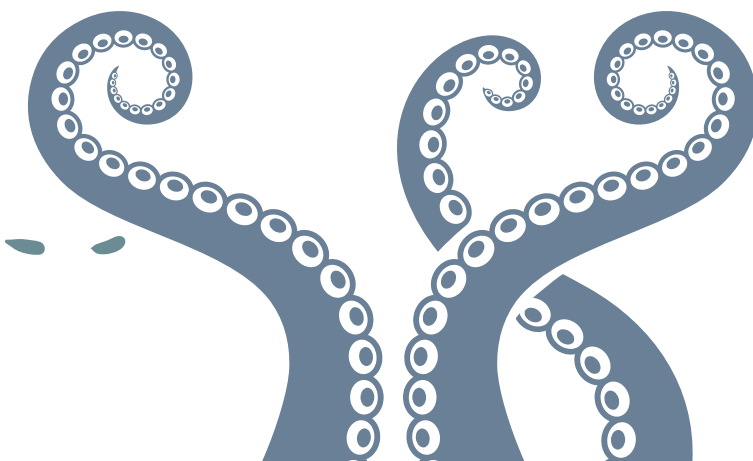
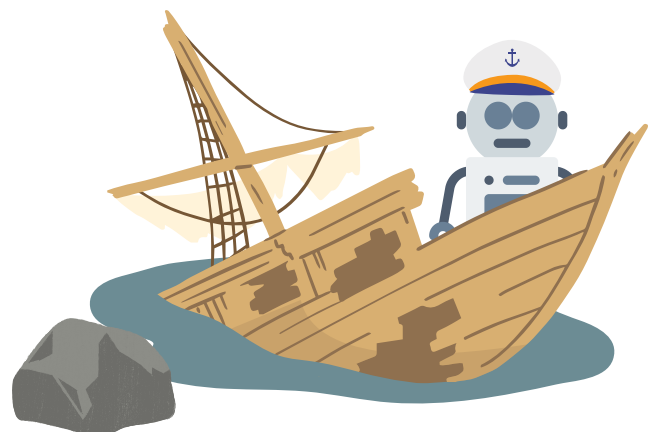
Créer ou optimiser un modèle existant



Entraîner son modèle



Utiliser son modèle





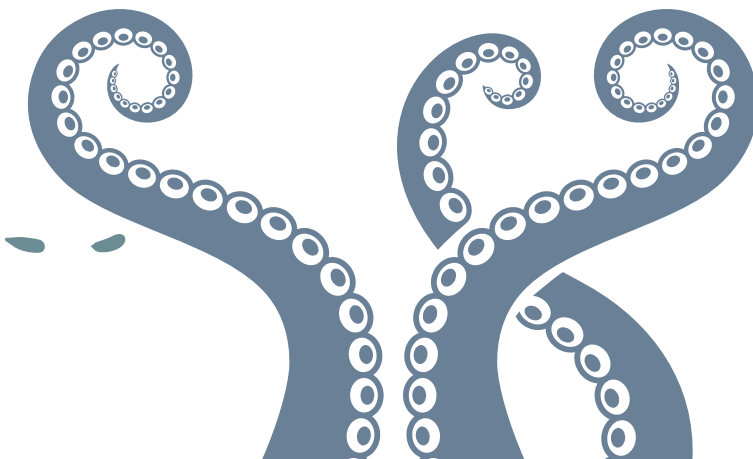
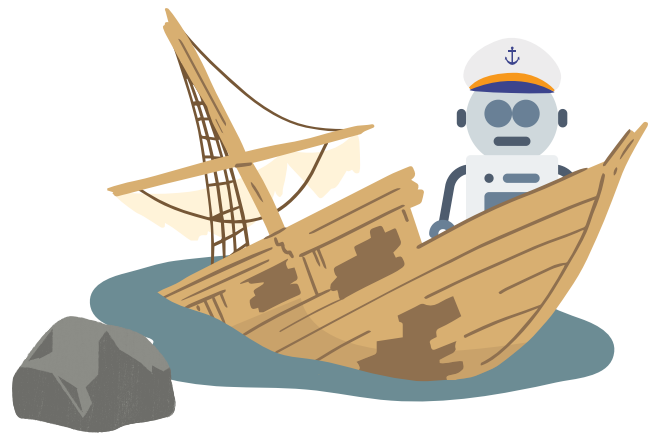
IMPORTANCE DU PRÉ-TRAITEMENT DES DONNÉES

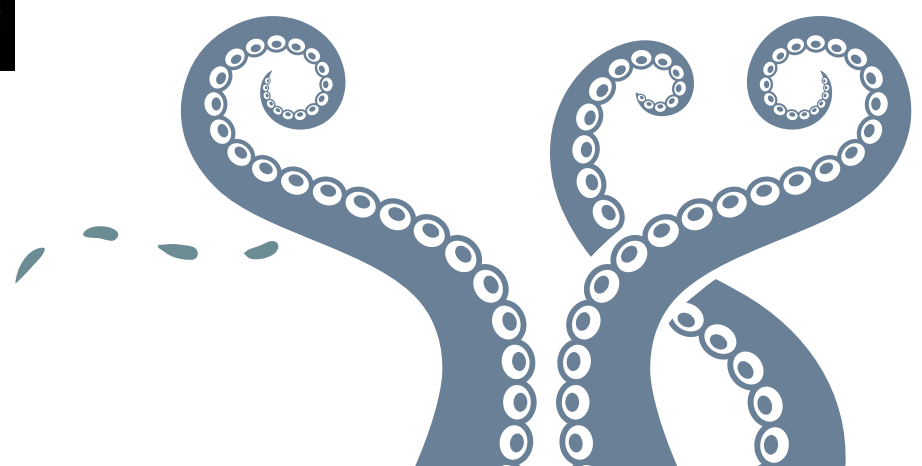
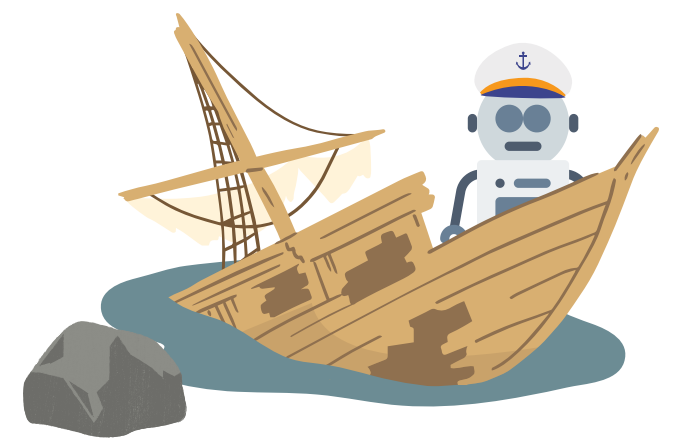
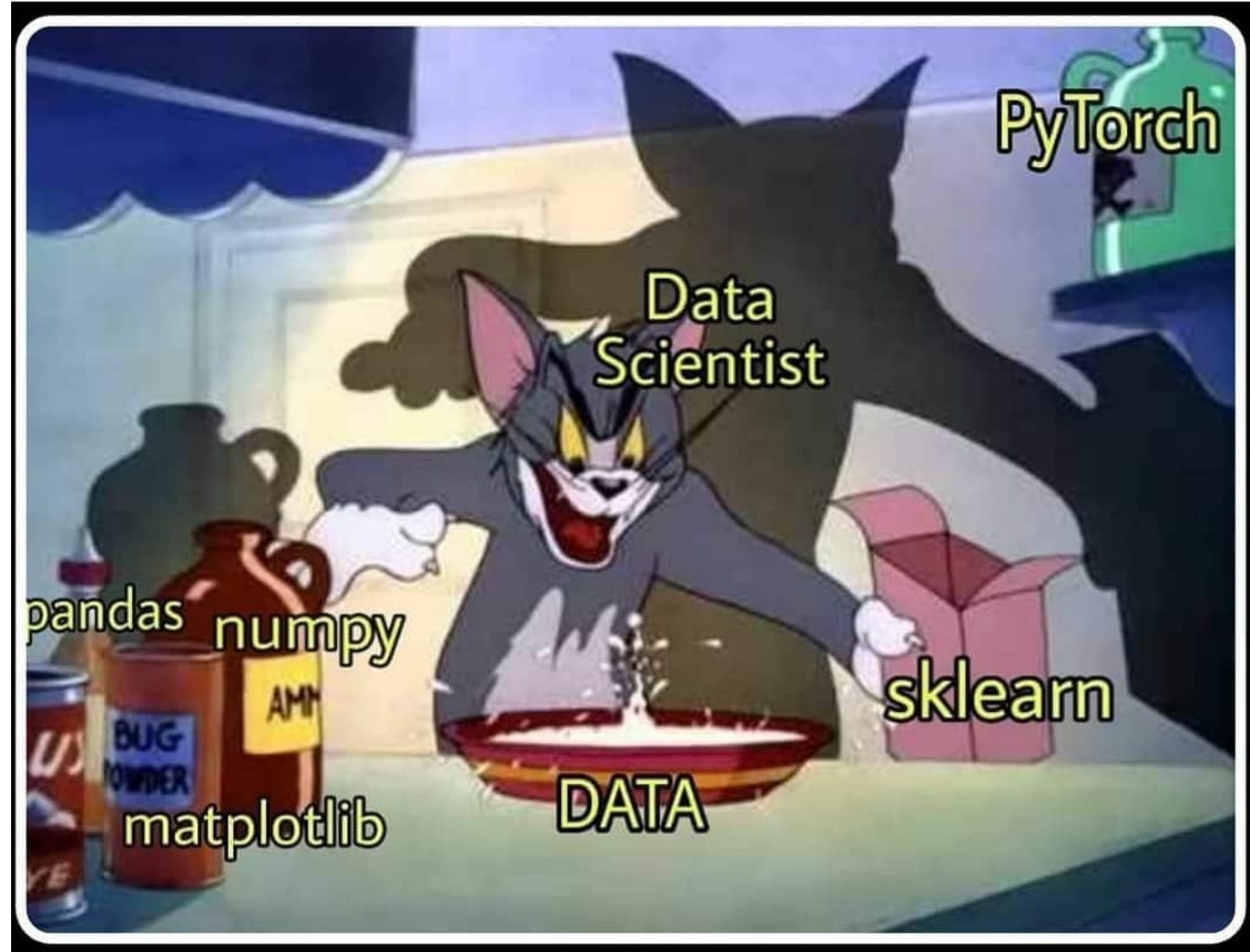


QUALITÉ DES DONNÉES



INFORMATIONS UTILES







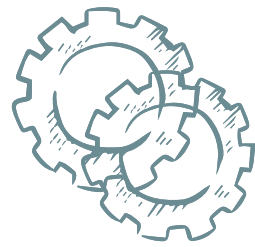
EXEMPLES DE PRÉ-TRAITEMENT POUR LES DONNÉES



Gestion des valeurs nulles / manquantes



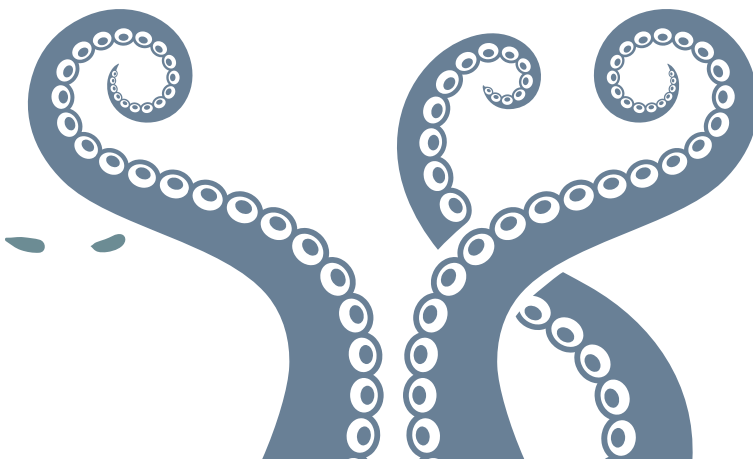
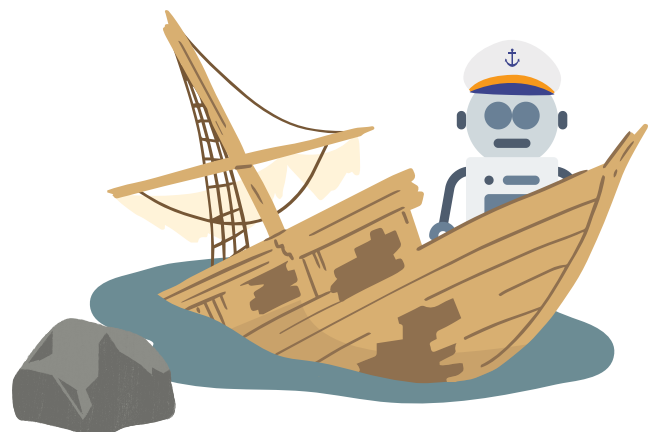
Normalisation



Gestion des variables qualitatives



Encodage à chaud

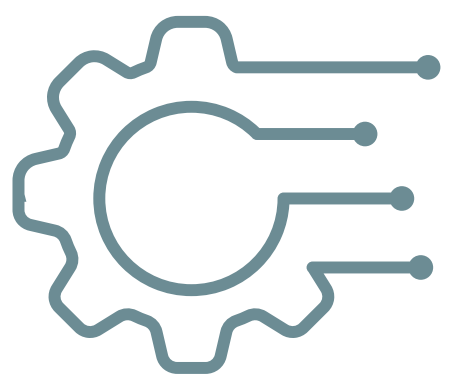




ET POUR LES DONNÉES AUDIOS ?



audios .wav



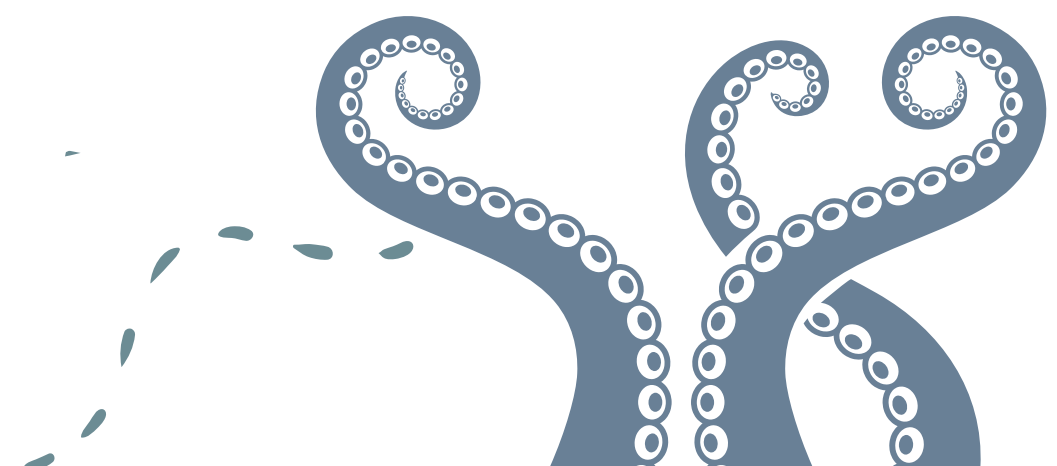
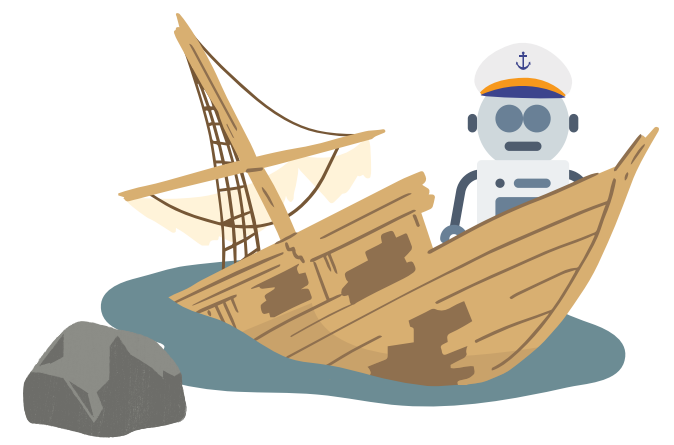
paramètres audio
Librosa



N paramètres audio Librosa
←-----→

name	length	name

fichier .csv

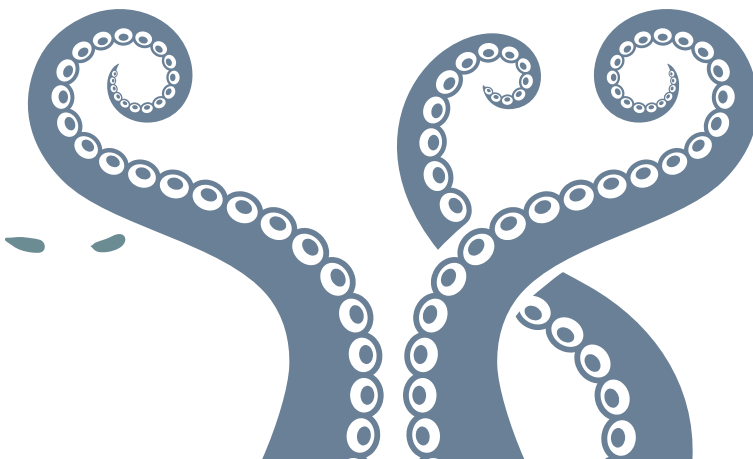
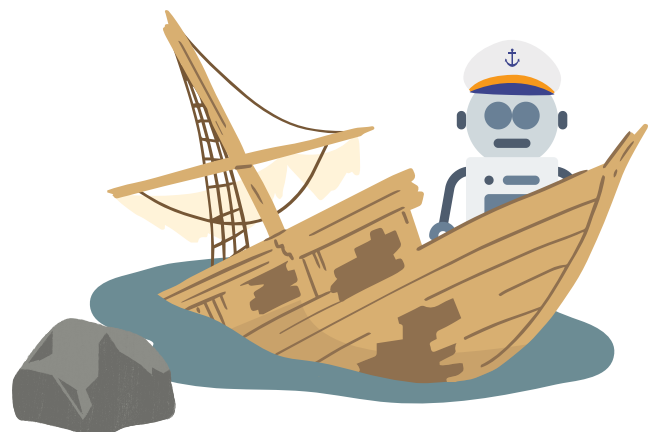




**me cleaning
the data**



**me building
the model**

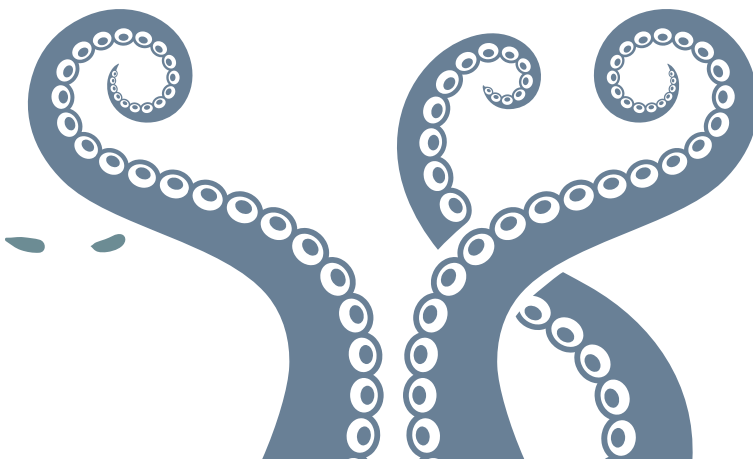
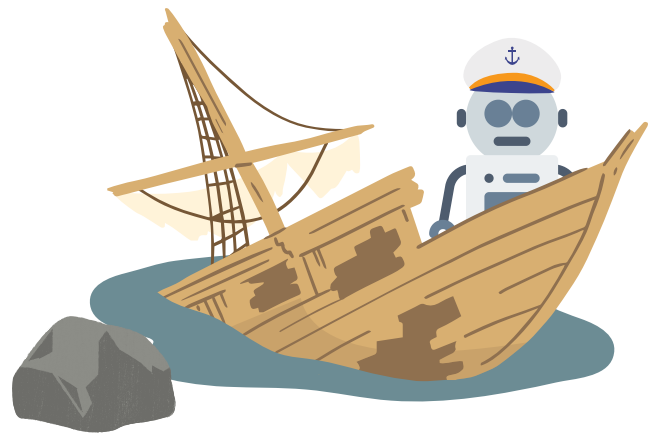




TECHNOS ET OUTILS



Le langage le plus utilisé : Python





TECHNOS ET OUTILS

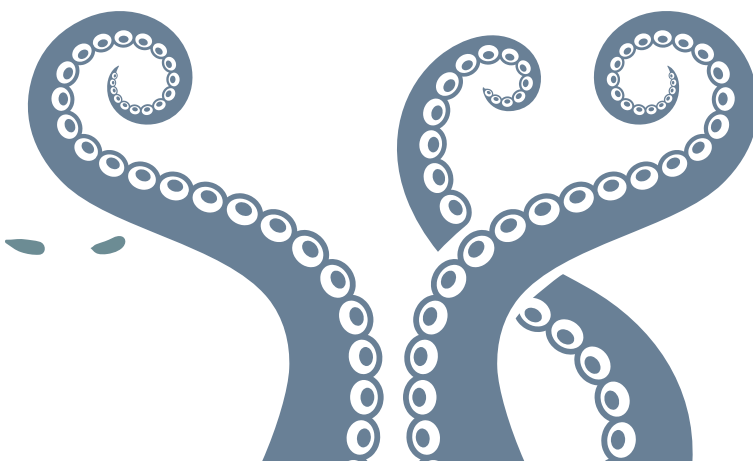
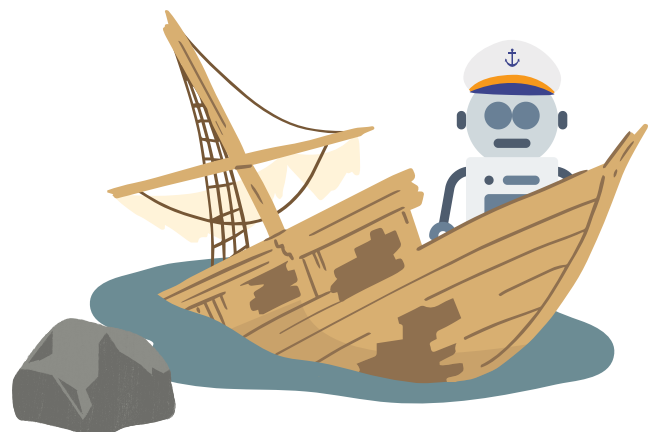


Le langage le plus utilisé : Python



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

Le cas à part de JupyterLab





TECHNOS ET OUTILS



Le langage le plus utilisé : Python



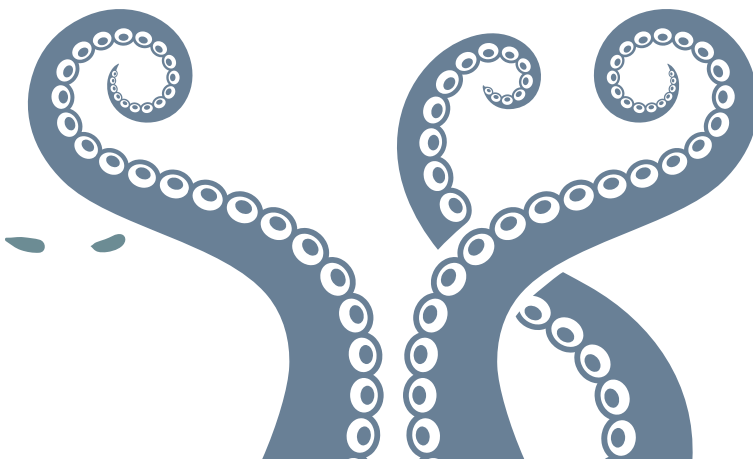
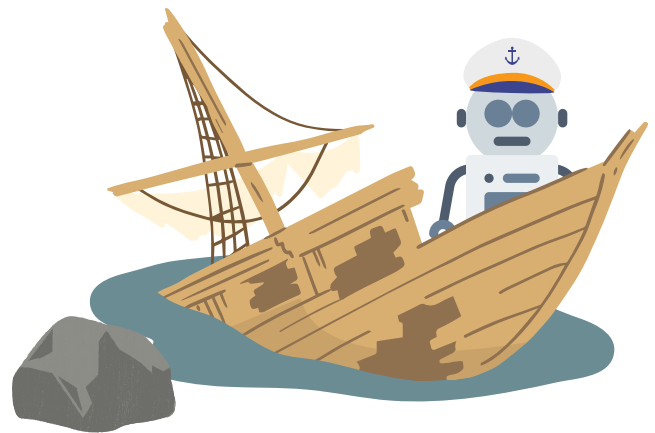
Beaucoup d'IDE (PyCharm, VSCode ...)

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 ...)

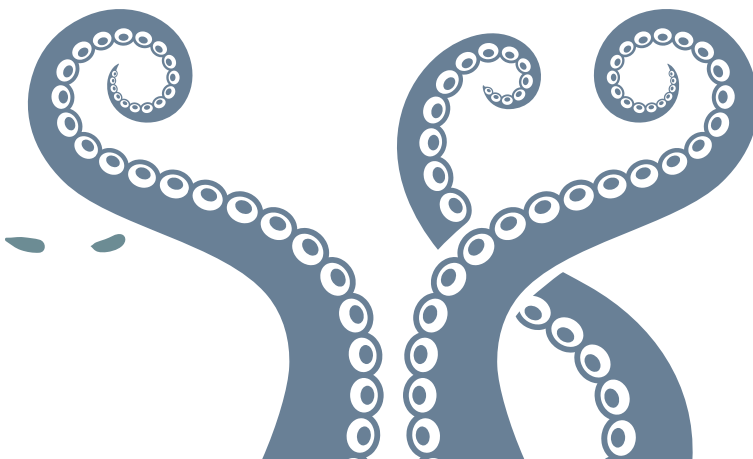
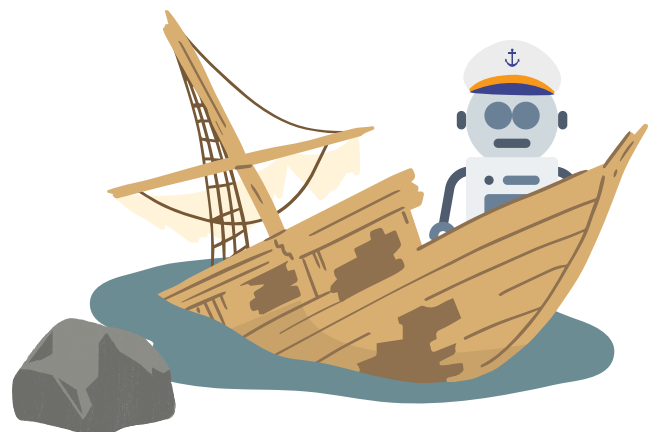
Le cas à part de JupyterLab



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

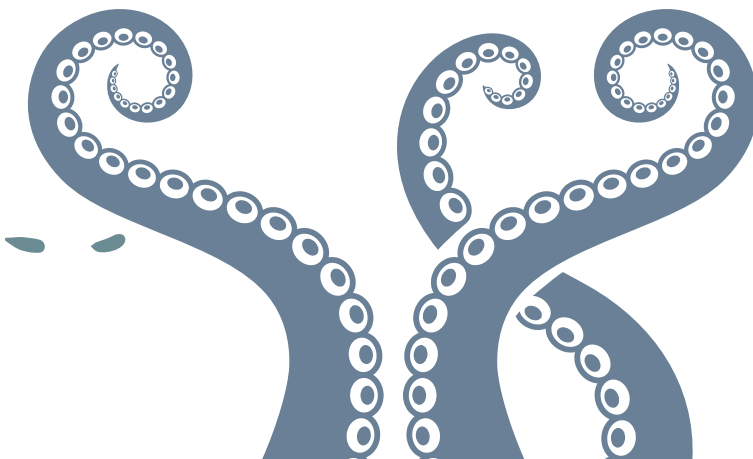
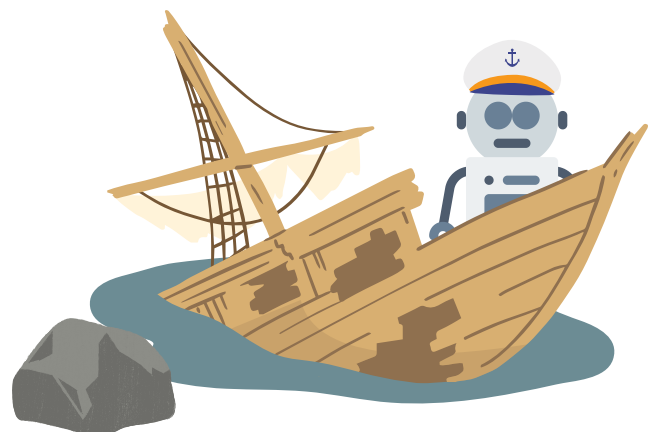


Les frameworks sont nombreux

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



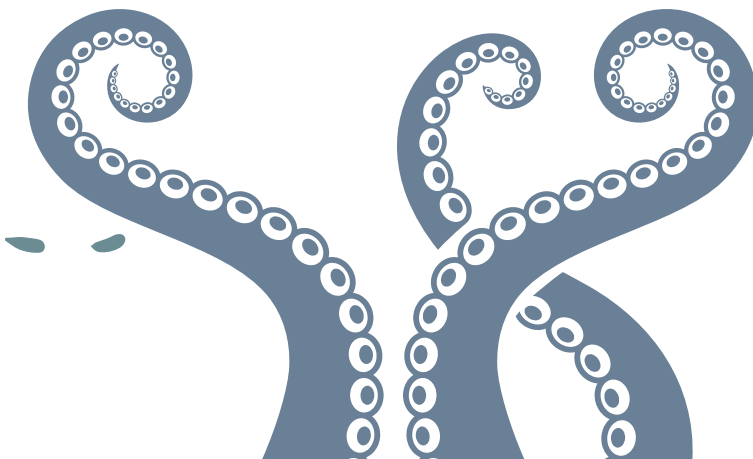
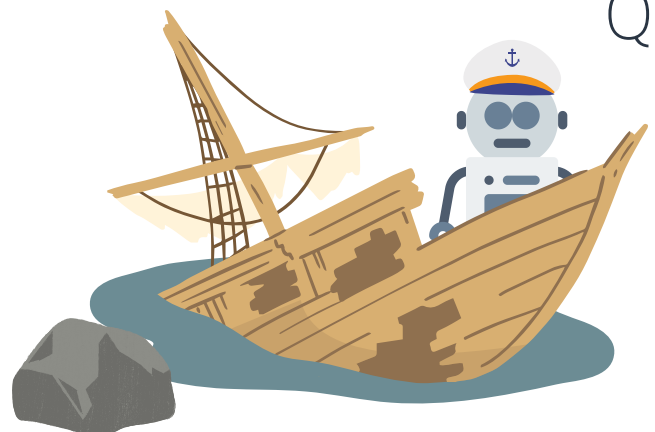
Les frameworks sont nombreux

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, ...

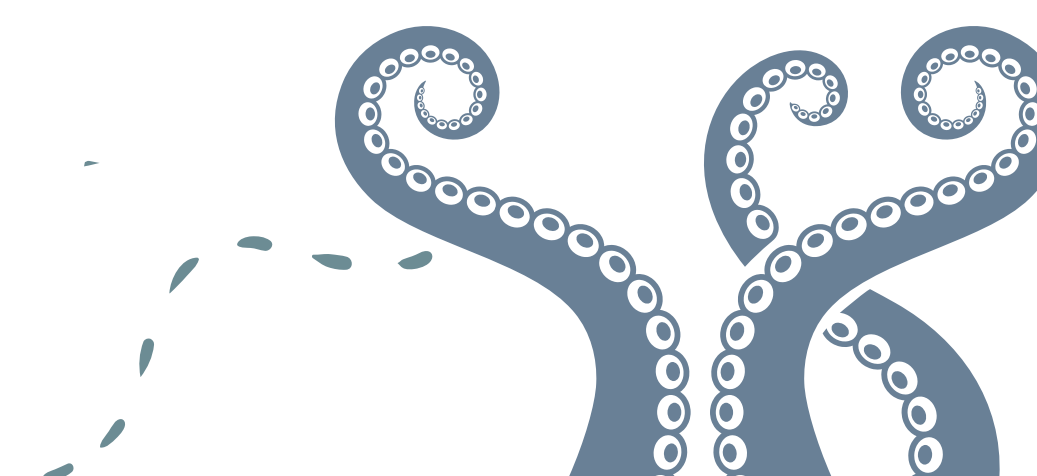
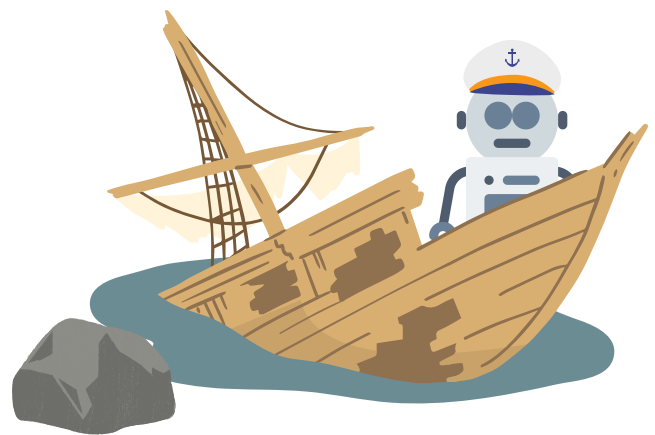




DÉMO TIME !

```
0 response = requests.get(url)
1 # checking response.status_code (if you get 502, try rerunning the code)
2 if response.status_code != 200:
3     print(f"Status: {response.status_code} - Try rerunning the code!")
4 else:
5     print(f"Status: {response.status_code}\n")
6
7 # using BeautifulSoup to parse the response object
8 soup = BeautifulSoup(response.content, "html.parser")
9
10 # finding Post images in the soup
11 images = soup.find_all("img", attrs={"alt": "Post image"})
12
13 # downloading images
14 for i in range(len(images)):
15     image = images[i]
16     url = image.get("src")
17     response = requests.get(url)
18     with open(f"image_{i}.jpg", "wb") as f:
19         f.write(response.content)
```

Artturi Jalli





IDENTIFIER UN MAMMIFÈRE PAR RAPPORT À SON CRI



Notebook

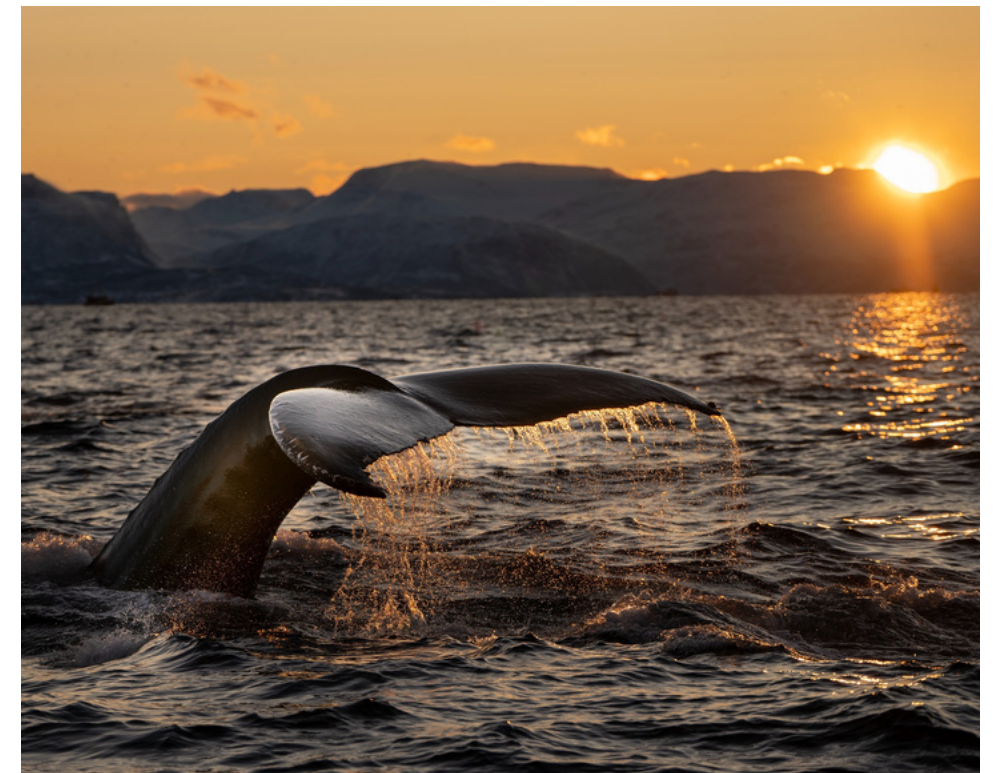
Analyse de l'audio

Pré-traitement des données

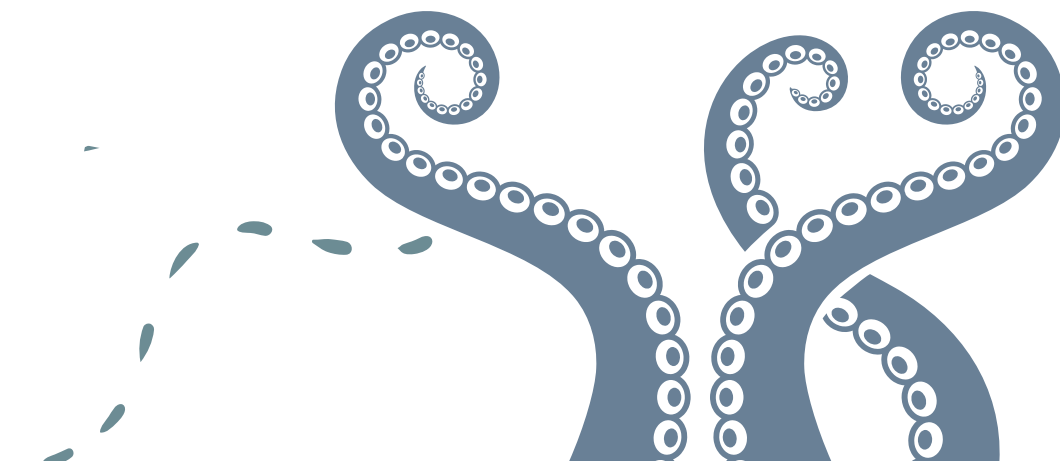
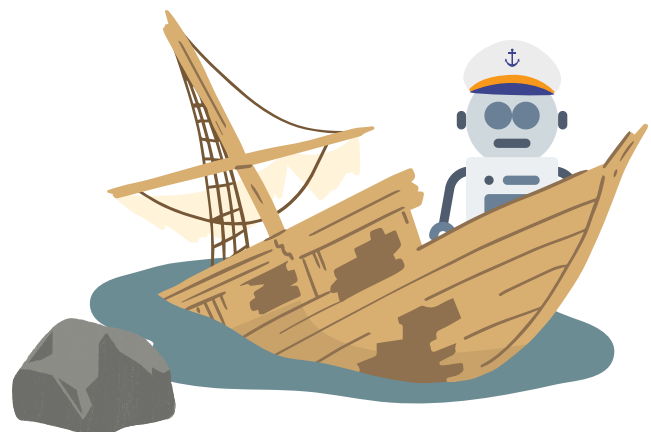
Construction du modèle

Entraînement et évaluation du modèle

Test et enregistrement modèle



Bart van meeke





IDENTIFIER UN MAMMIFÈRE PAR RAPPORT À SON CRI



Notebook

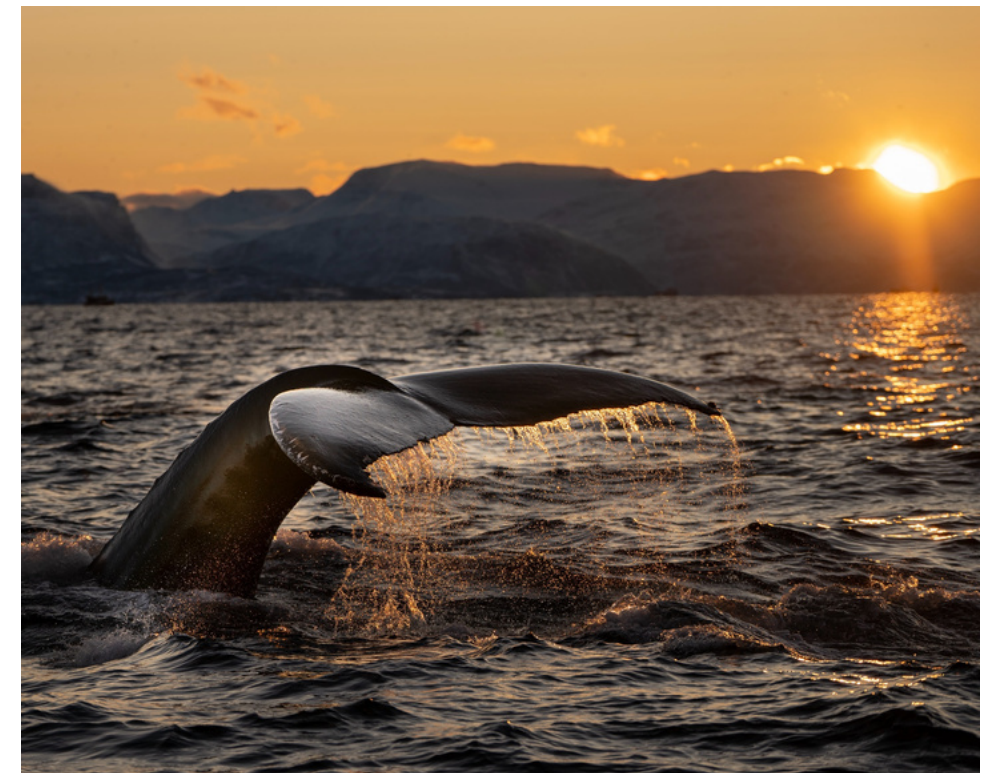
Analyse de l'audio

Pré-traitement des données

Construction du modèle

Entraînement et évaluation du modèle

Test et enregistrement modèle



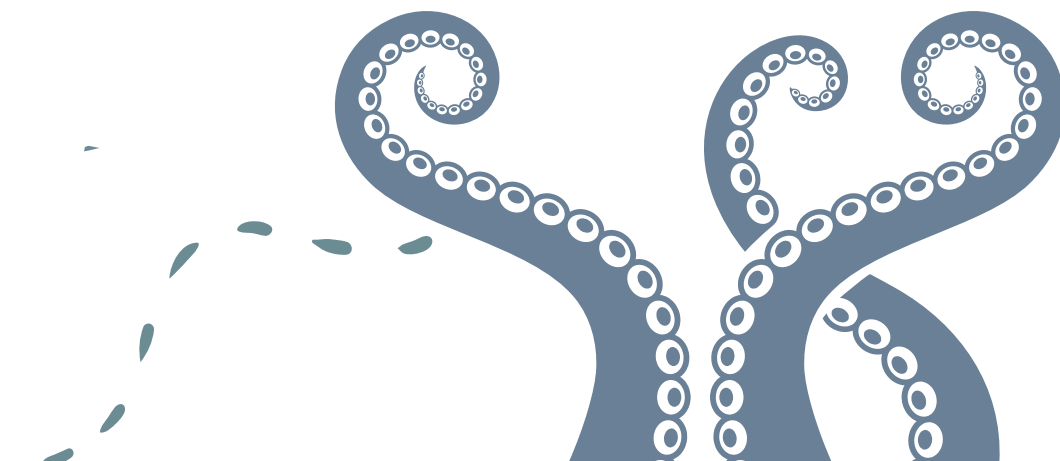
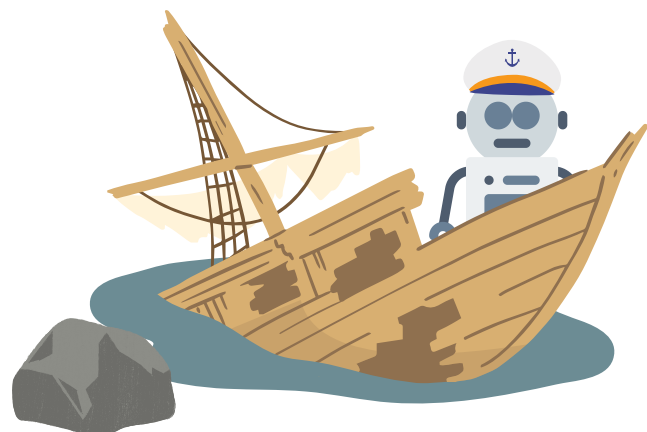
Bart van meeke



Applications

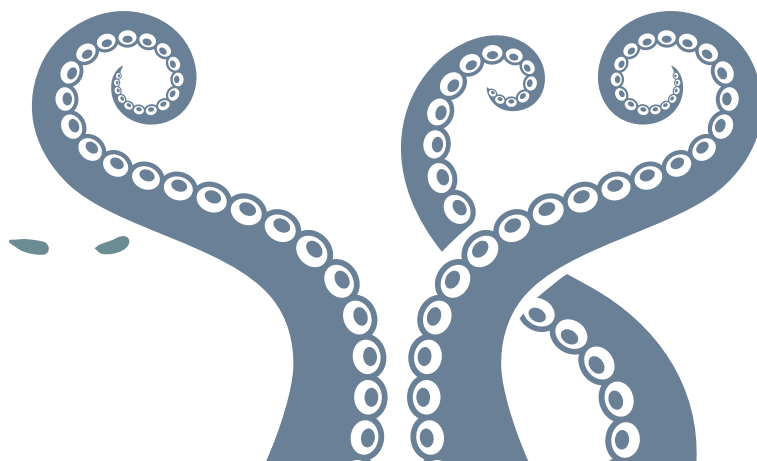
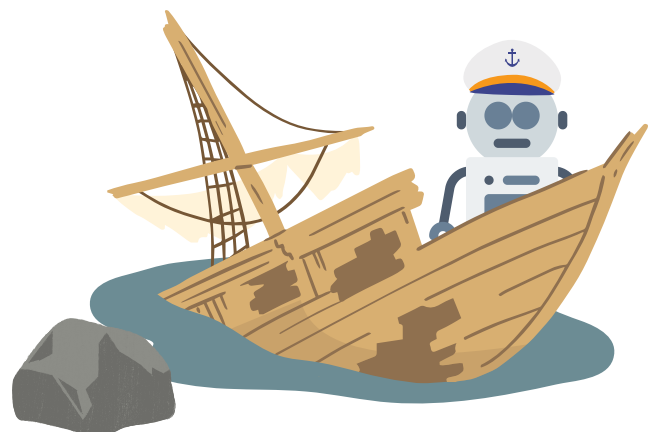
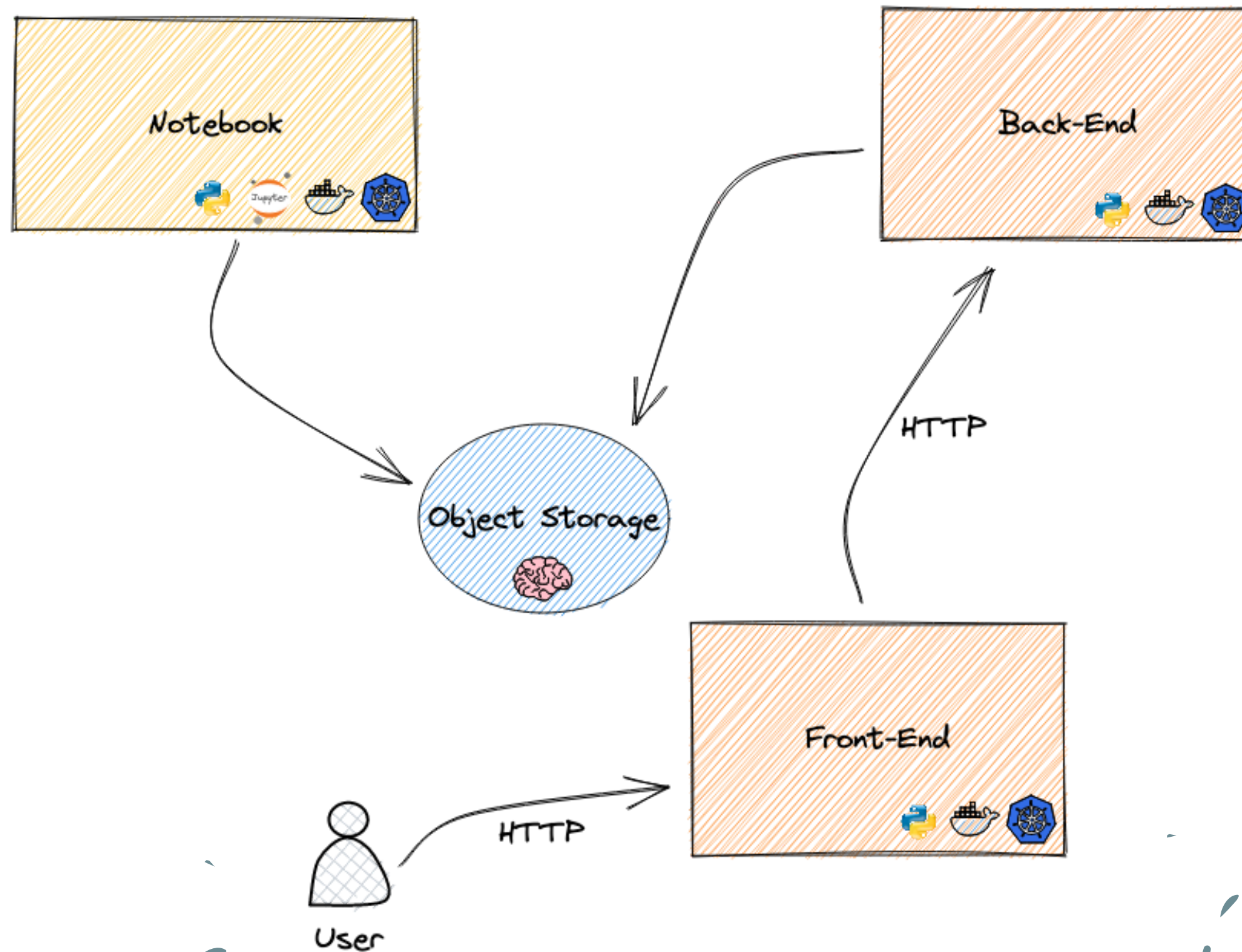
Back-End : exposition de l'API

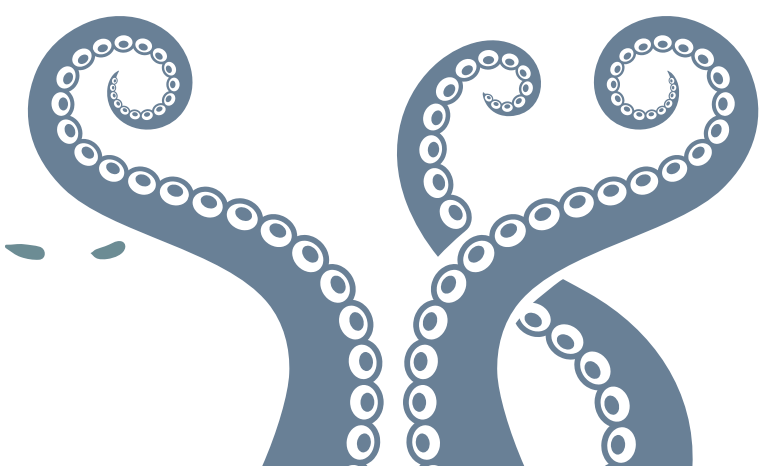
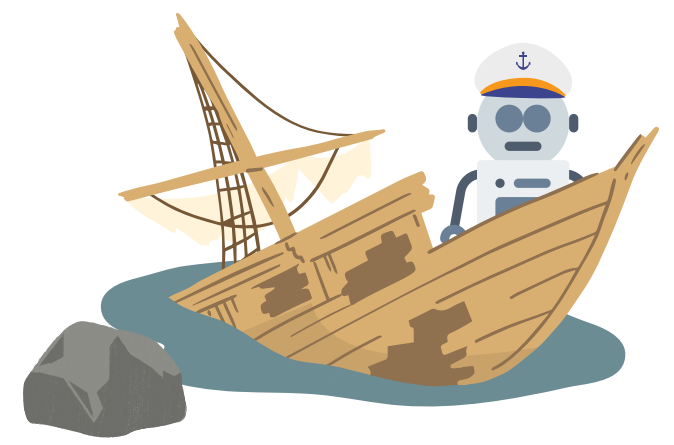
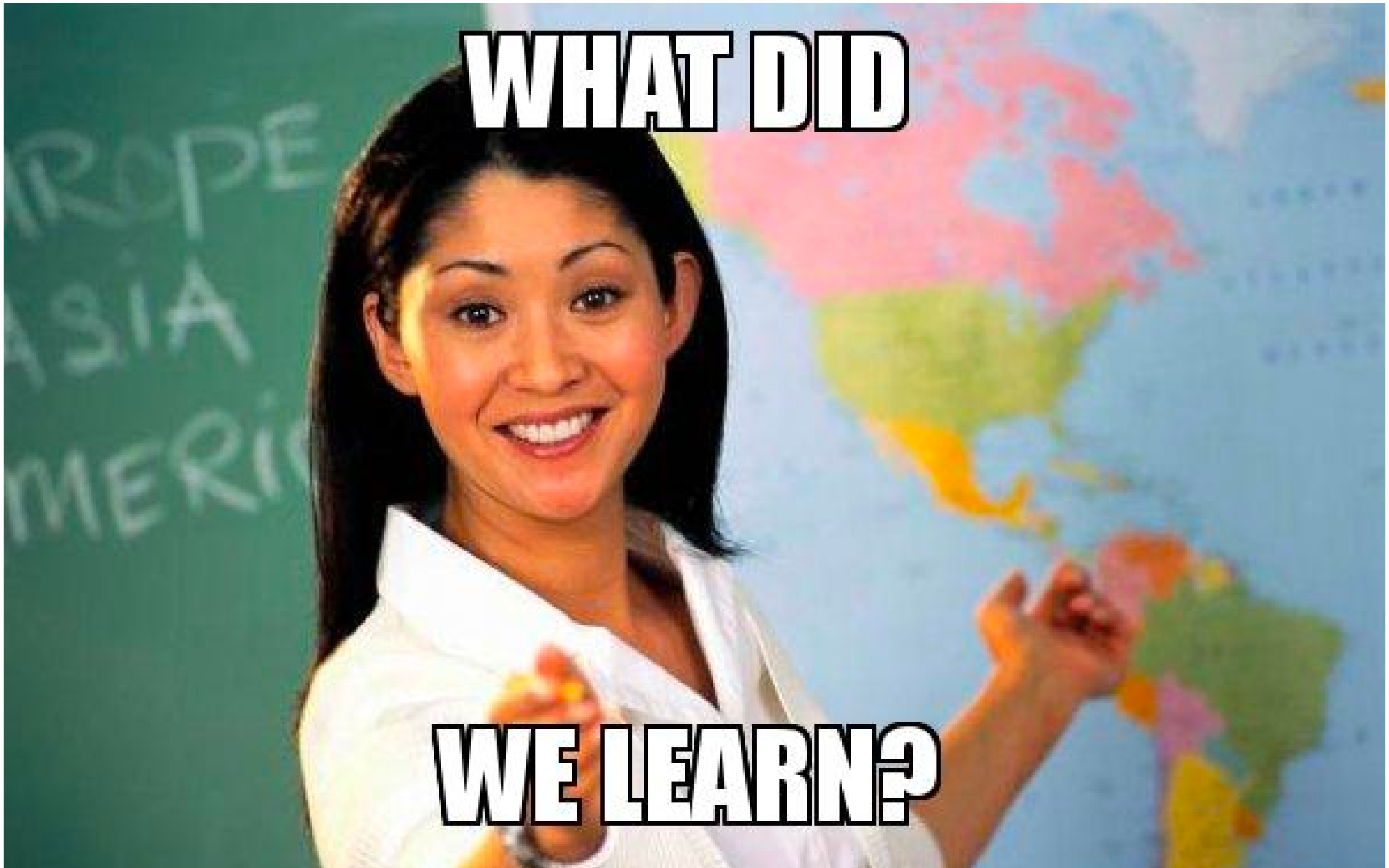
Front-End : analyse d'un son chargé ou enregistré





SCHEMA D'ARCHITECTURE GLOBALE



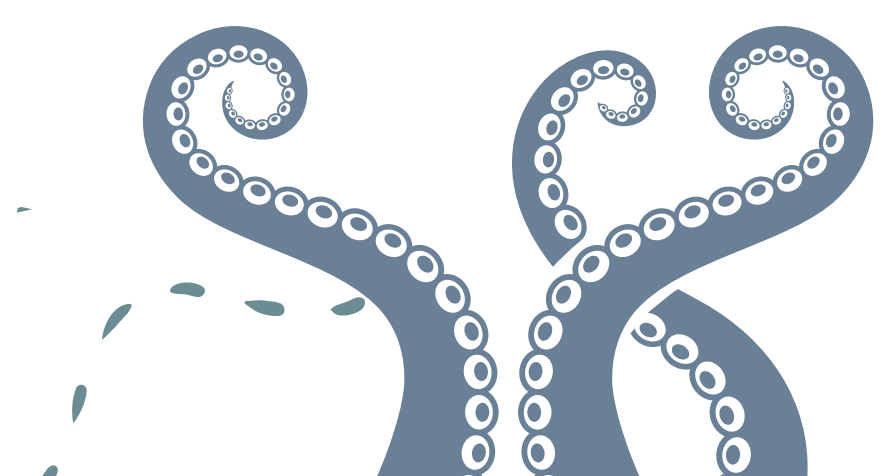




QU'AVONS NOUS APPRIS AUJOURD'HUI ?

NOTION D'IA

Datascience, Machine Learning, Deep Learning





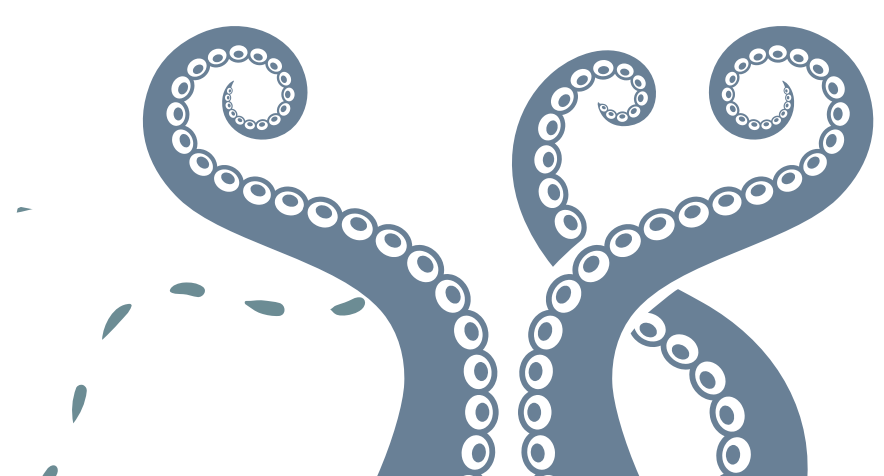
QU'AVONS NOUS APPRIS AUJOURD'HUI ?

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

AI NOTEBOOK

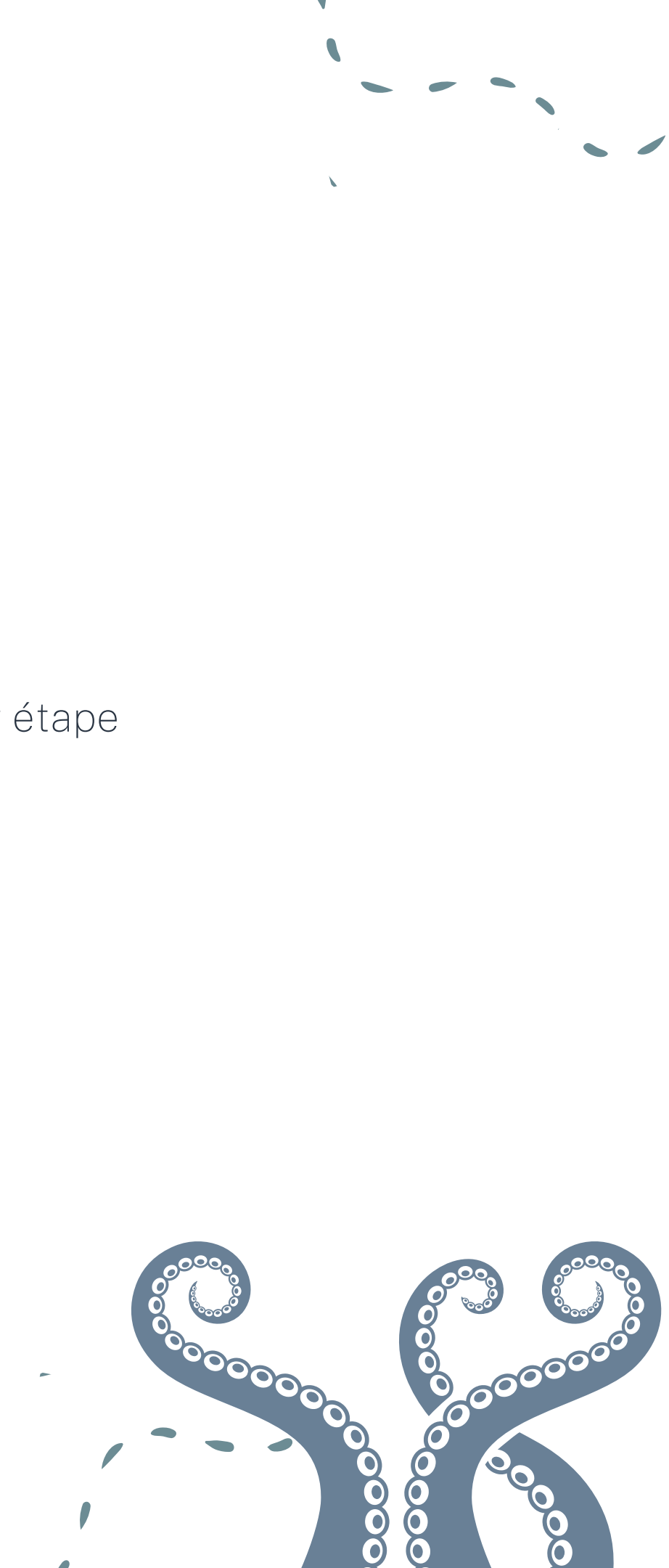
Développement de mon code étape par étape

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

AI NOTEBOOK

Développement de mon code étape par étape

PLACE DU SON EN IA

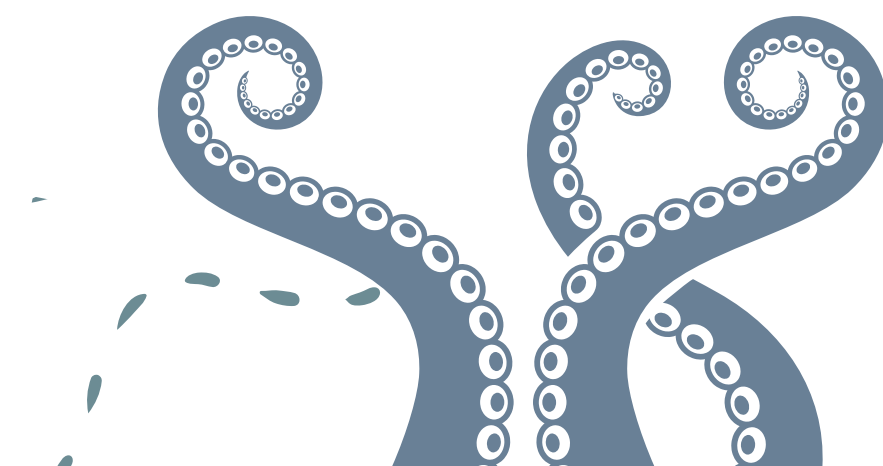
Environnement, musique, NLP

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

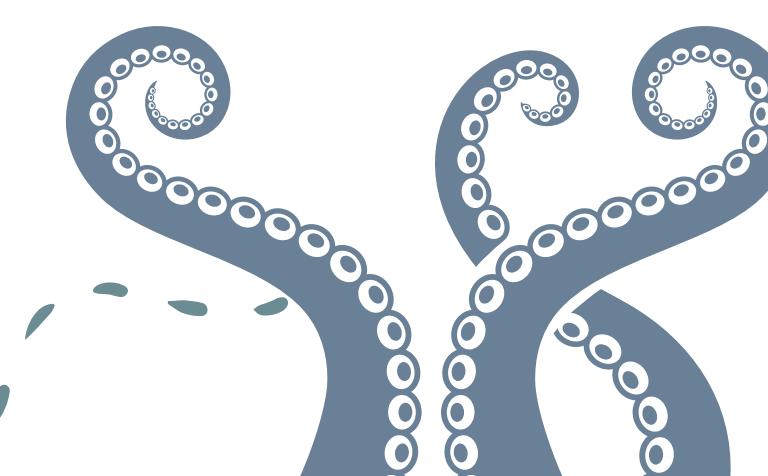
Réseau de neurones, exportation

PRÉTRAITEMENT DES DONNÉES

Répartition, transformation, labélisation

APP POUR IA

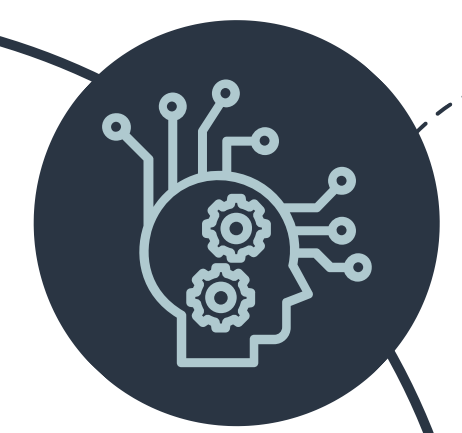
Object Storage, Streamlit, Docker, API



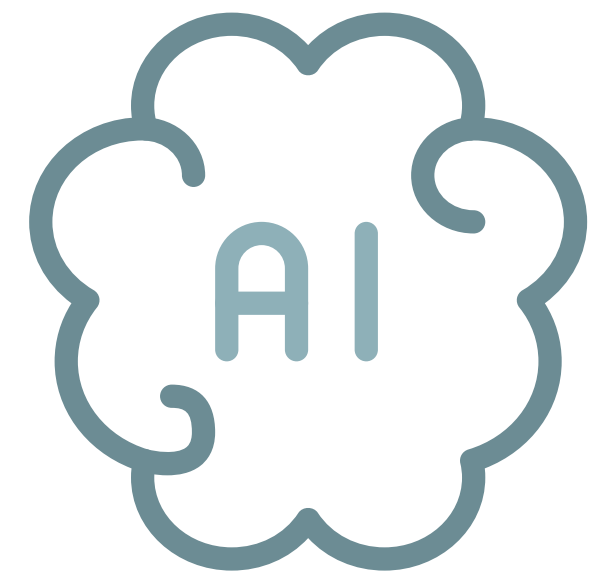
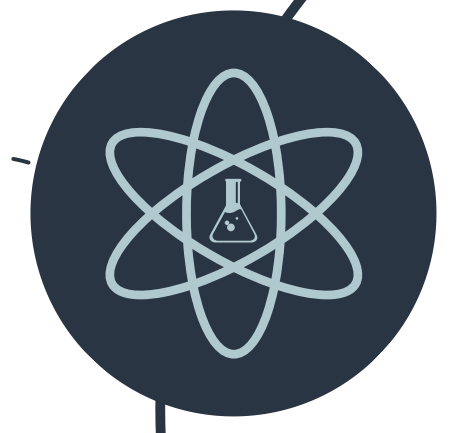


LES MÉTIERS DE L'IA

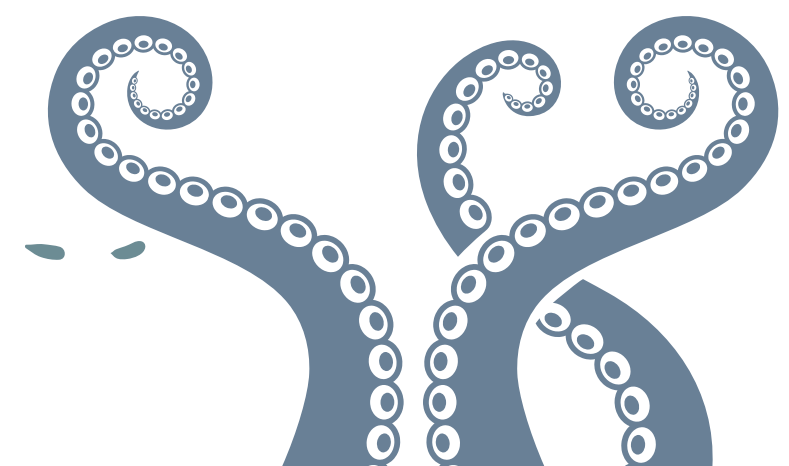
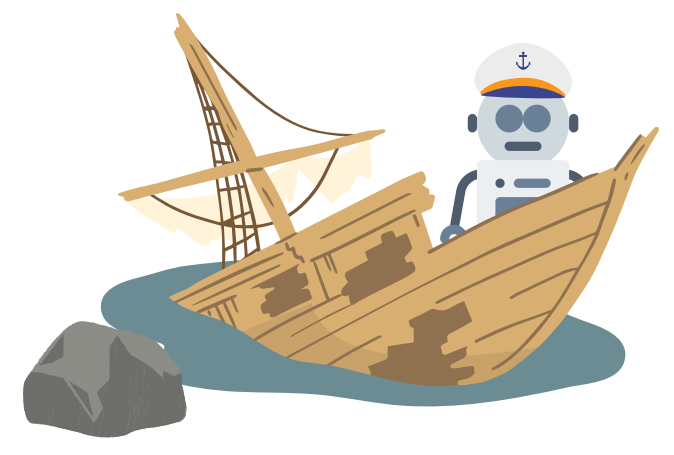
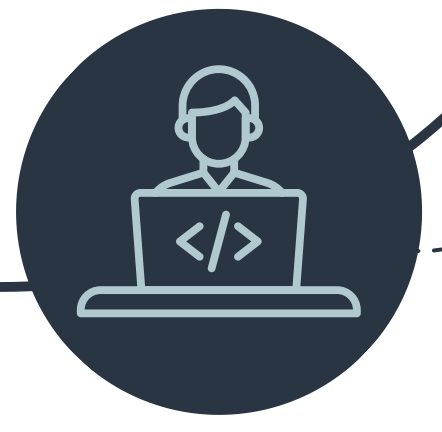
MACHINE LEARNING ENGINEER



DATA SCIENTIST



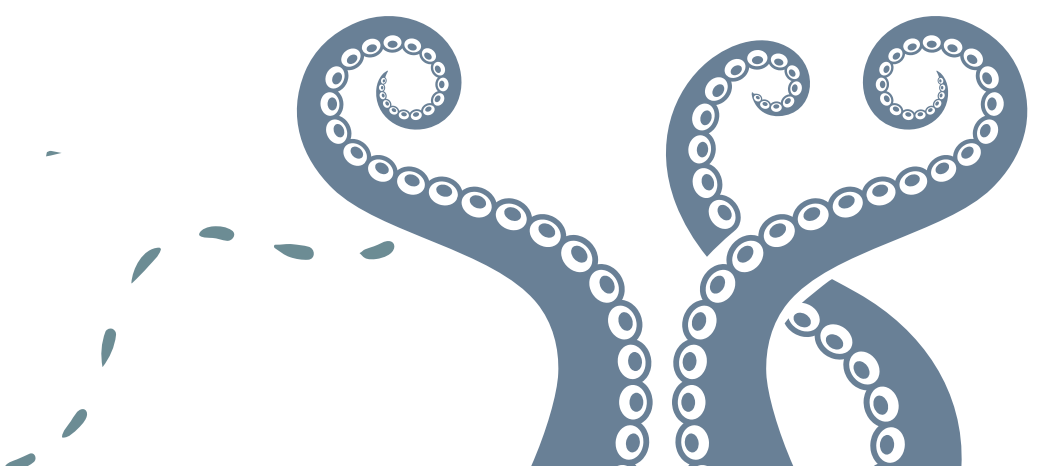
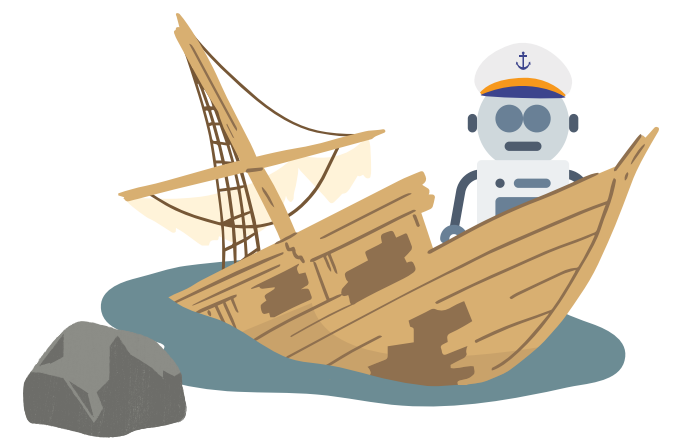
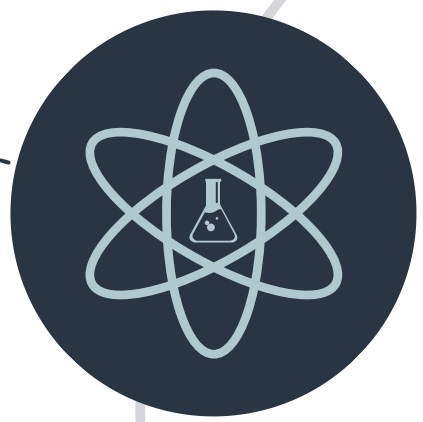
DEVELOPER





LES MÉTIERS DE L'IA


DATA SCIENTIST





LE DATA SCIENTIST

 Data Analysis

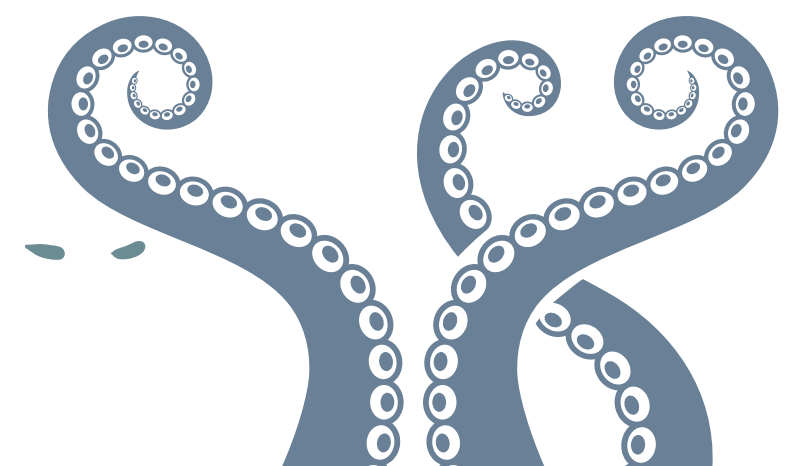
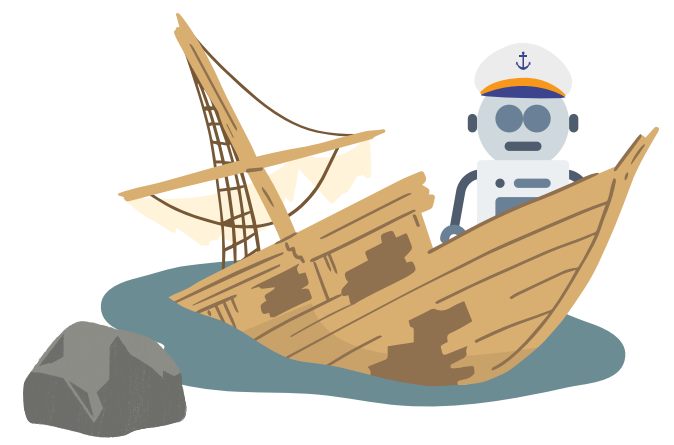
Machine Learning 

 Data Visuatization

Maths f_x

 Programmation

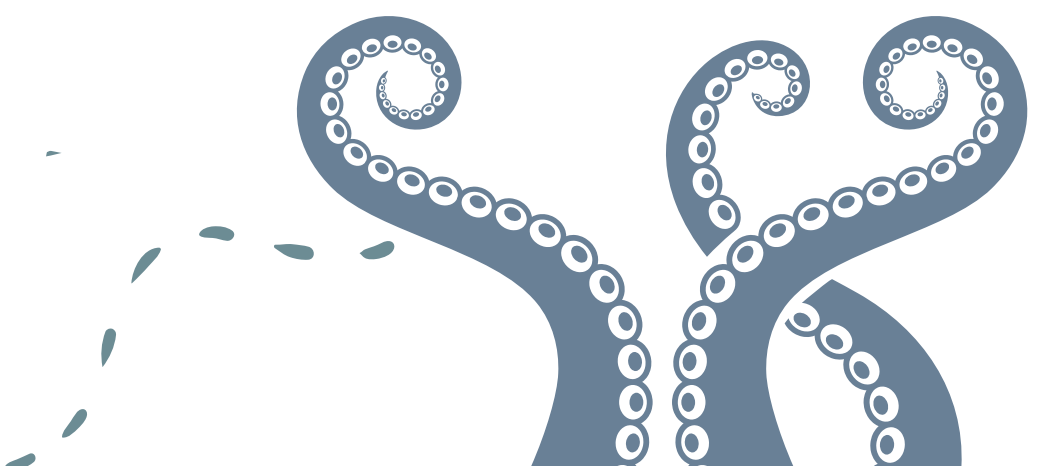
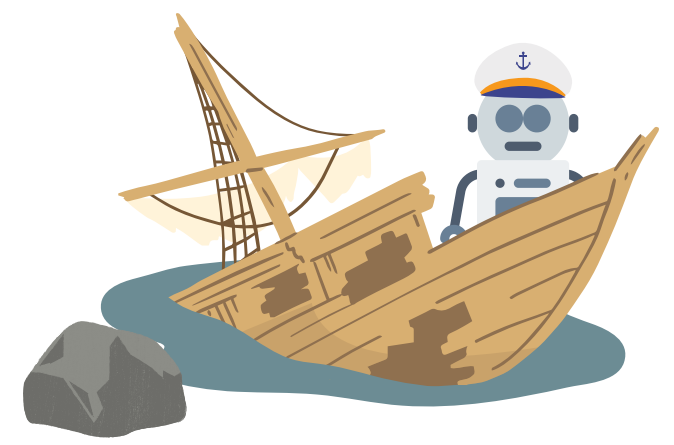
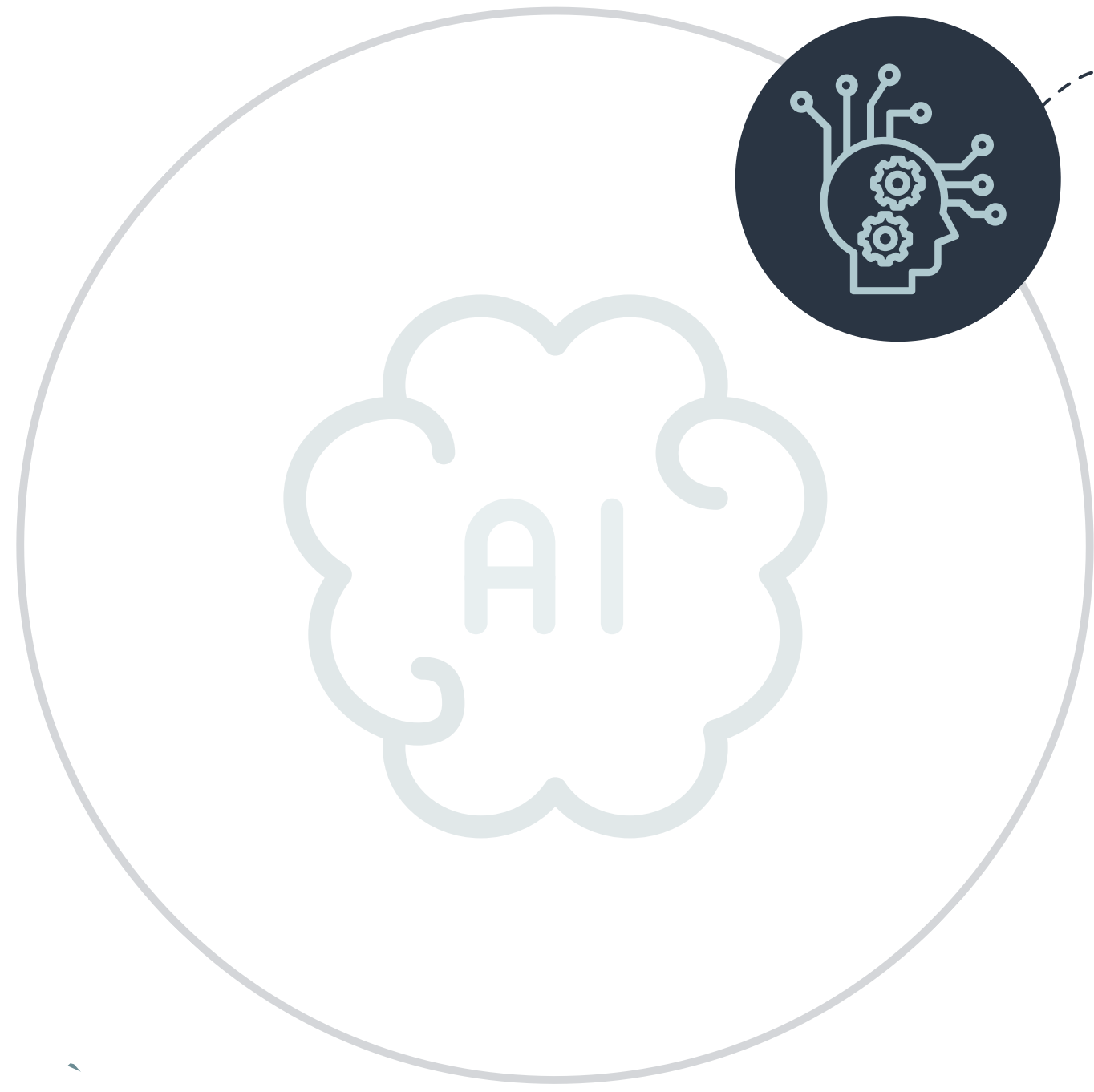
Statistics 





LES MÉTIERS DE L'IA

MACHINE LEARNING ENGINEER

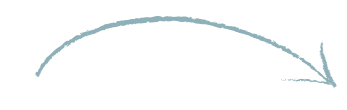




LE MACHINE LEARNING ENGINEER



Data Science

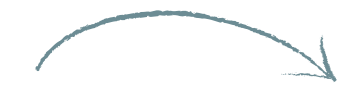


Frameworks



f_x

Maths & Statistics



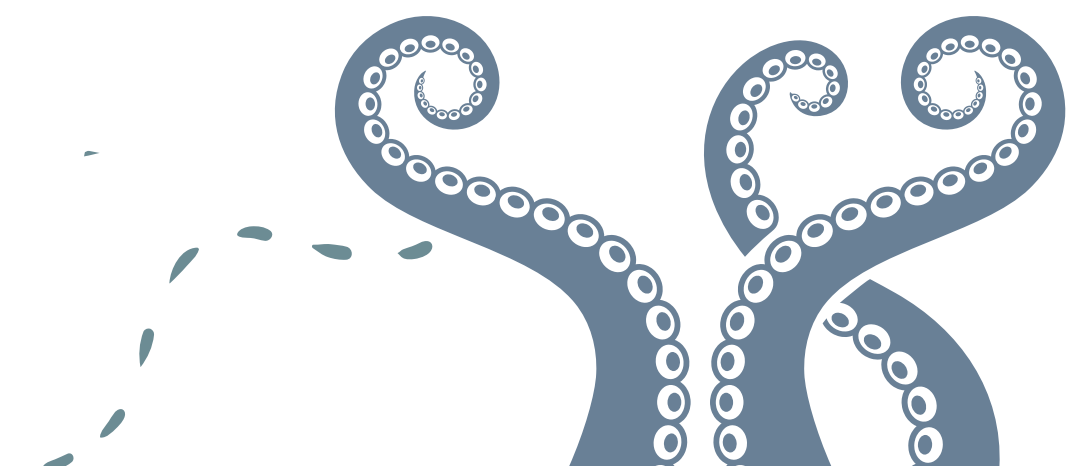
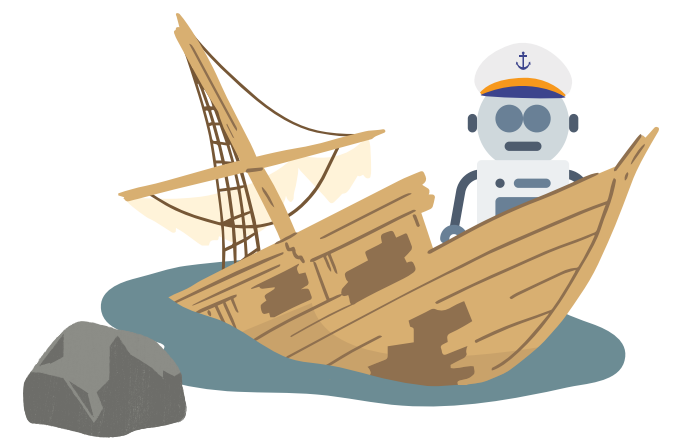
Environnement de prod



Programmation

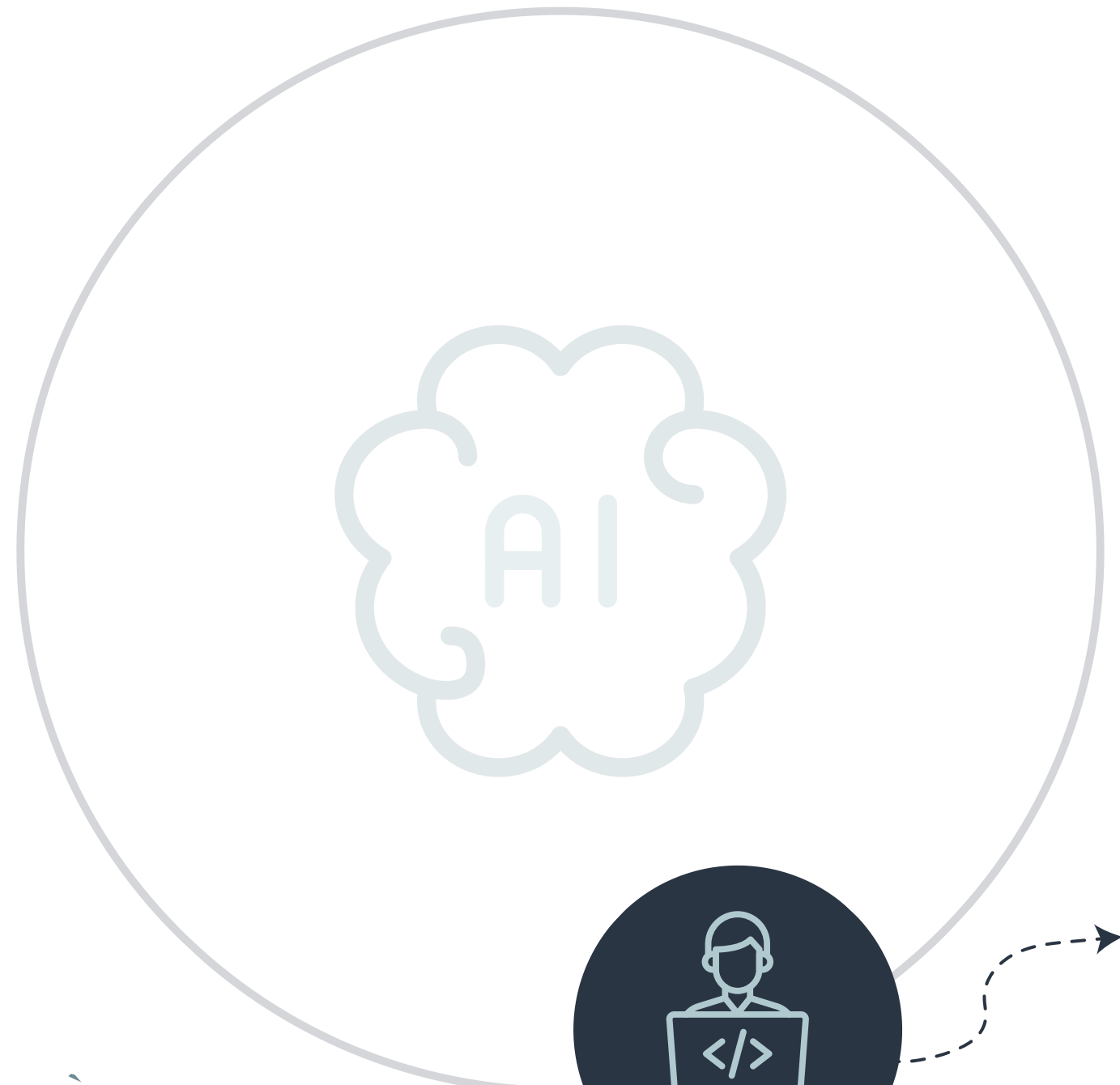


Services Cloud

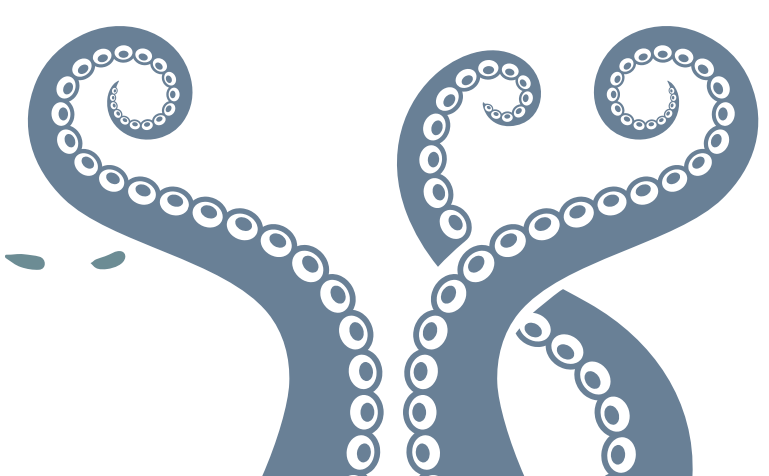
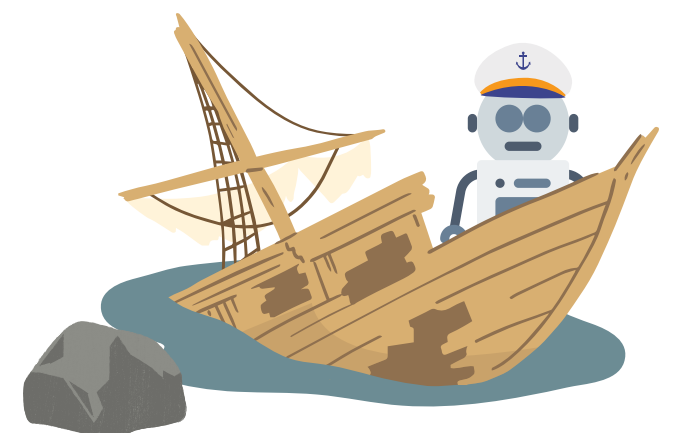




LES MÉTIERS DE L'IA



DEVELOPER





LE DÉVELOPPEUR

 Python 

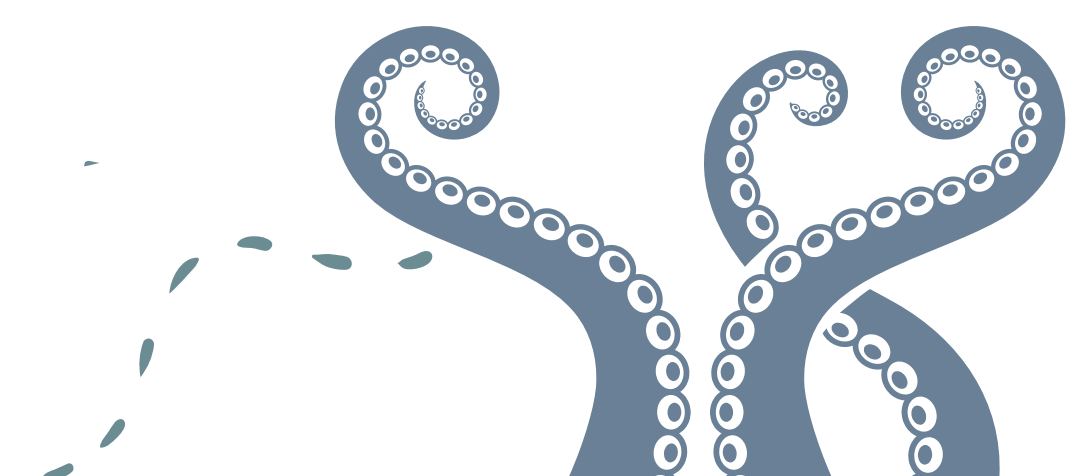
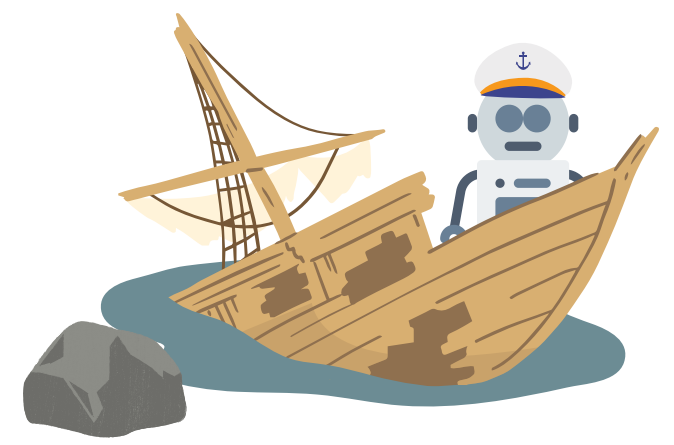
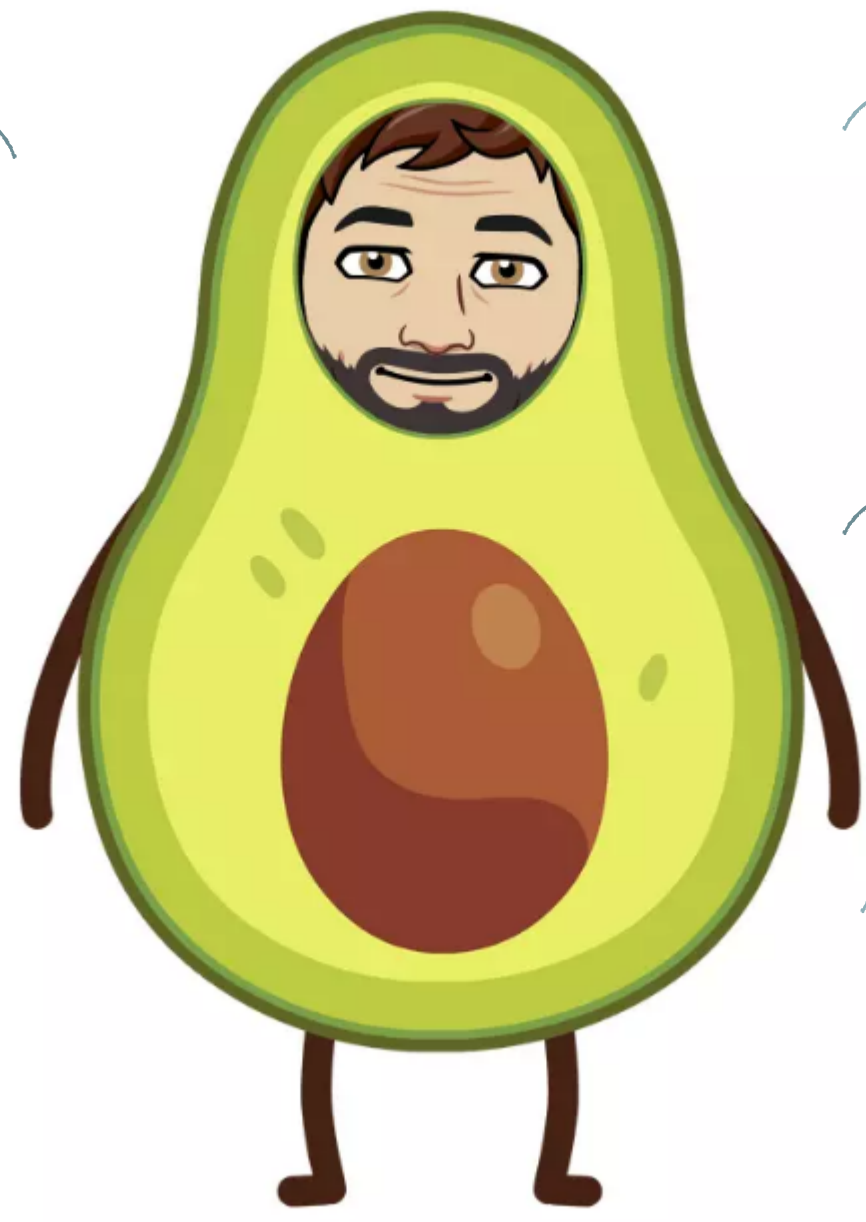
 Docker 

 Concepts IA 

 Frameworks 

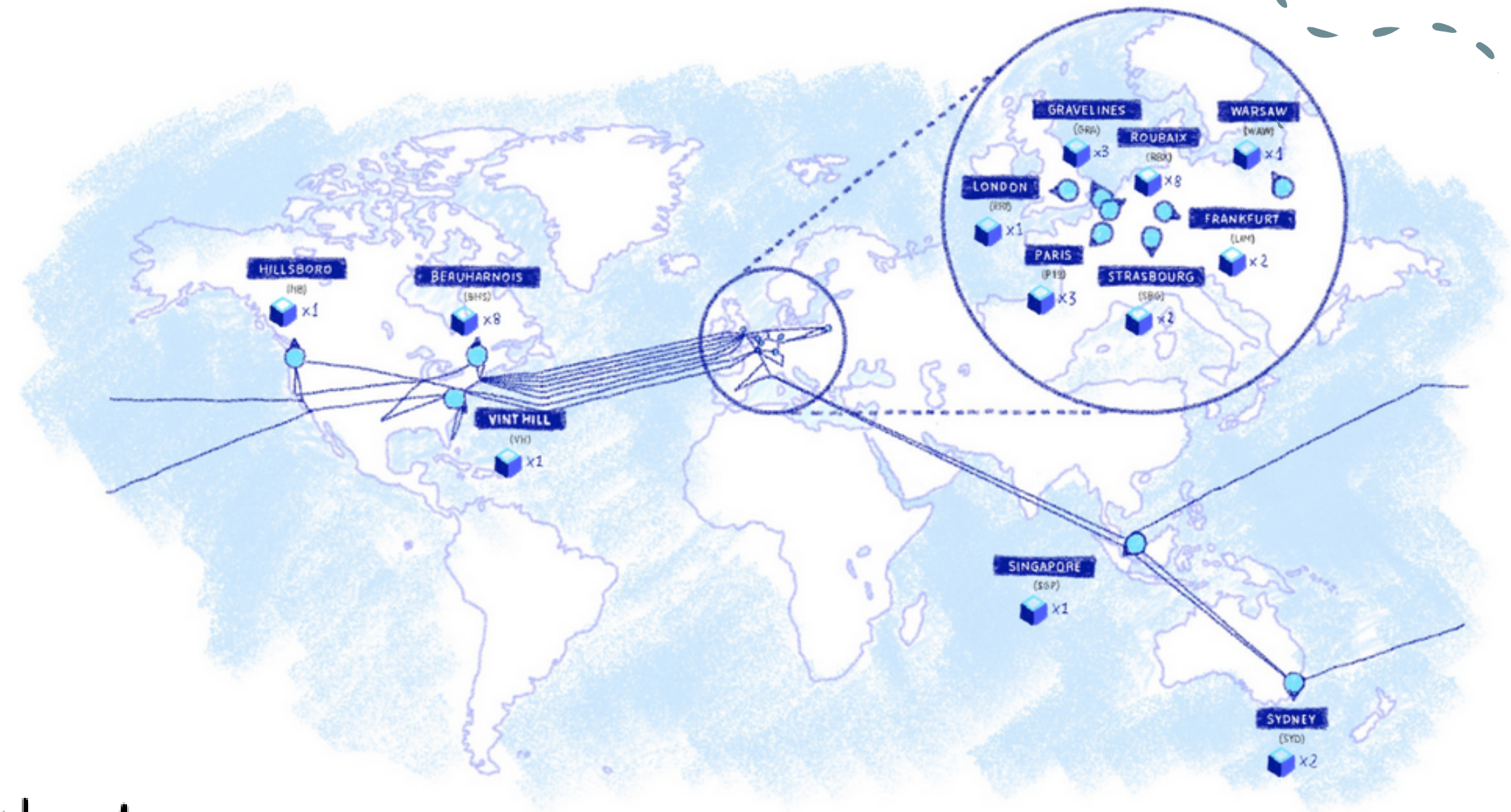
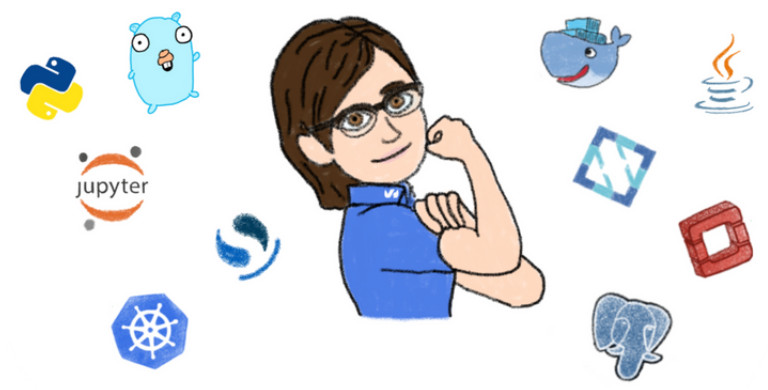
 Développement 

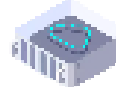

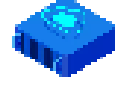
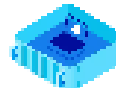

 Outillage 



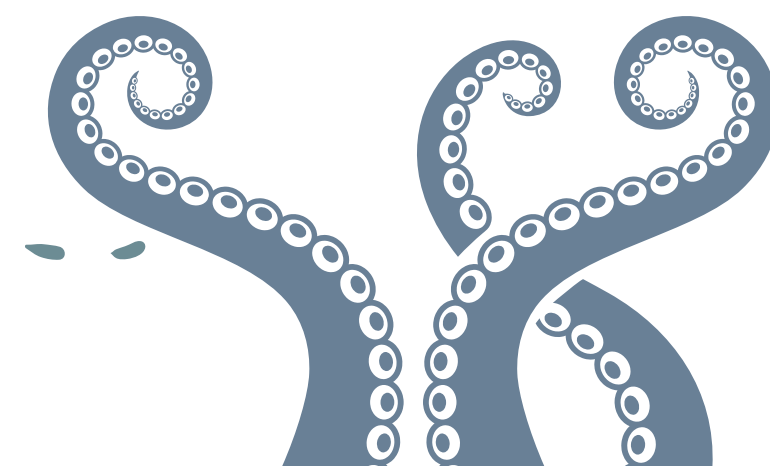
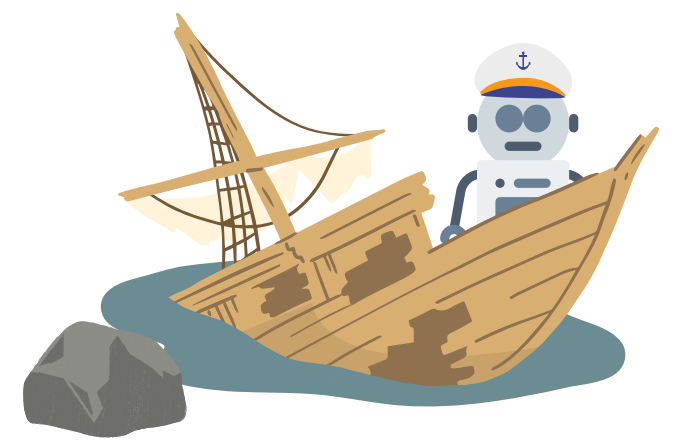
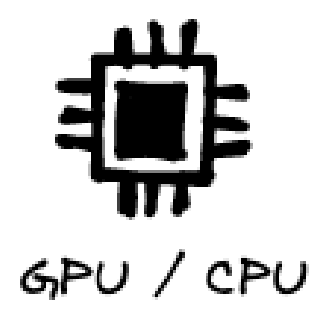


The Cloud with European values



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

- AI Notebooks
- AI Training
- AI Deploy

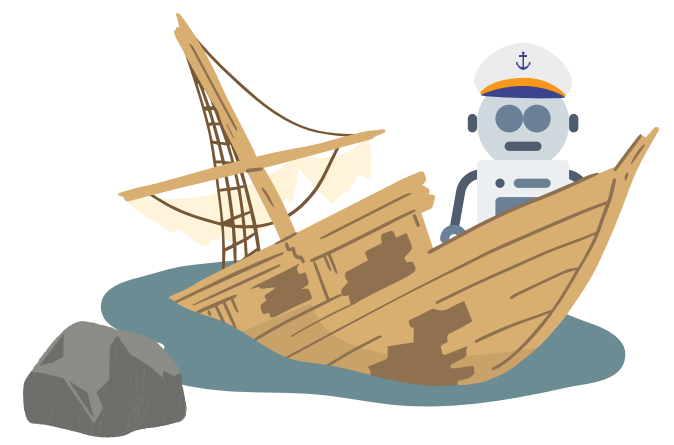




ELEA PETTON

 MACHINE LEARNING ENGINEER @OVHCLOUD

 ETUDIANTE @ISENBREST



@EleaPetton



eleaptn



Eléa PETTON

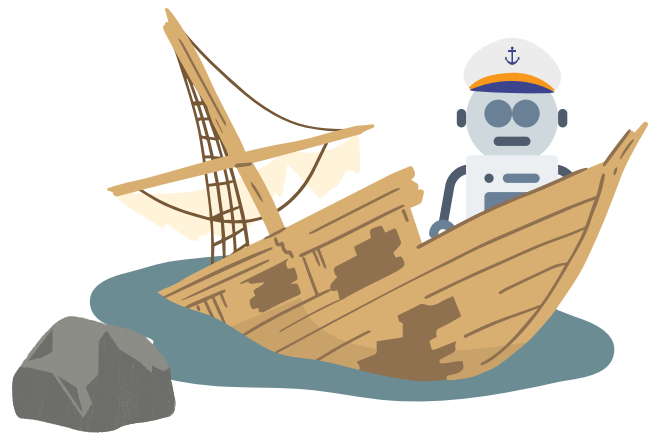


STÉPHANE PHILIPPART

DEVREL@OVHCLOUD

CO-CRÉATEUR DE TADX

[HTTPS://PHILIPPART-S.GITHUB.IO/BLOG](https://philippart-s.github.io/blog)



@wildagsx



philippart-s



philippartstephane

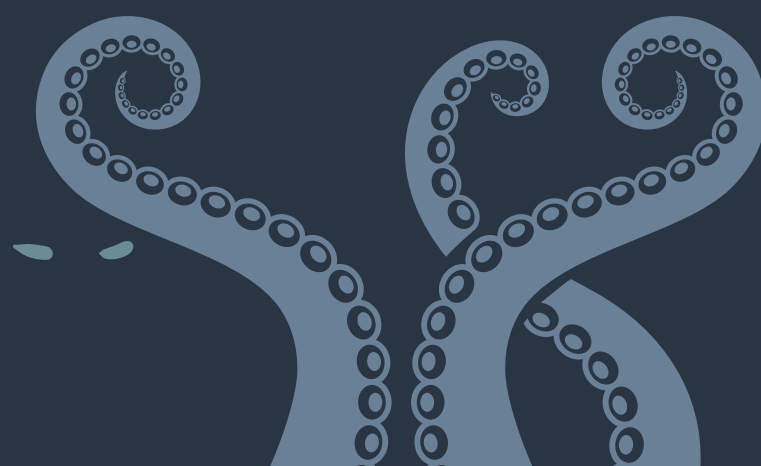


MERCI !



<https://tiny.one/JUG-NEMO>

Voucher pour essayer : T3R1-PLPH






POUR TESTER

 LE NOTEBOOK : <https://github.com/eleaptn/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>