



## MLOps roadmap



# Data and Machine learning operationalization

[Building and deploying ml app]

Presented By:  
Généreux Akotenou



[github.com/genereux-akotenou](https://github.com/genereux-akotenou)



[genereux.akotenou@gmail.com](mailto:genereux.akotenou@gmail.com)

27 Avril 2024

# Mon parcours

## • Education



## • Entreprise:

MALAIKA GAME STUDIOS  
MIFY SARL  
University UC-Louvain  
MARKETSCC CRYPTOCURRENCIES  
CYBERSPECTOR

## • Intérêt:

NATURAL LANGUAGE PROCESSING  
MACHINE LEARNING/DEEP LEARNING  
INFORMATION RETIRVAL

# Plan

- 1- Motivation
- 2- Objectifs de déploiement
- 3- Exemple de déploiement
- 4 - Résumé
- 5 – Motivation Tendances



# 1- Motivation

---

Pourquoi le Data et le ML-Ops sont si importants ?

# 1 - Motivation

85 % des projets de big data échouent.

- [Gartner](#), 2017

87 % des projets de science des données ne parviennent jamais à la production.

- [VentureBeat](#), 2019

Qualité des données et difficulté de mise en œuvre des modèles dans les environnements de production.

- [Harvard Business Review](#), 2023

# 1 - Motivation

85 % des projets de big data échouent.

- [Gartner](#), 2017

87 % des projets de science des données ne parviennent jamais à la production.

- [VentureBeat](#), 2019

Qualité des données et difficulté de mise en œuvre des modèles dans les environnements de production.

- [Harvard Business Review](#), 2023

Pourquoi tant de projets échouent-ils?

# 1- Motivation

Pourquoi tant de projets échouent-ils ? (1)

## Absence de bonnes données

- Une enquête de 2020 menée par l' [International Data Corporation](#) révèle que "le manque de volumes adéquats et de qualités des données d'apprentissage reste un défi majeur en développement"

## Absence de bon talent

- En 2020, [QuantHub](#) a classé la "science des données" comme la deuxième compétence la plus difficile à trouver (après la cybersécurité)

## Échec du déploiement

- Selon un article de Forbes de 2020, seules 15 % des grandes entreprises ont déployé des produits d'IA en production.
- Il existe souvent des écarts techniques, de compétences et de motivation entre ceux qui développent les modèles et ceux qui les maintiennent.

# 1- Motivation

Pourquoi tant de projets échouent-ils ? (2)

## Penser que le déploiement est la dernière étape

- Cependant, le véritable travail ne s'arrête pas là. Si vous le pensez, imaginez...
- Échouer à s'adapter aux changements inévitables des données et du marché. Par exemple, quelle serait l'exactitude de votre modèle de prévision de la demande de vols pendant les périodes de COVID si vous aviez entraîné le modèle avec des données de 2019 ?
- Ne pas savoir que vos flux de données cessent d'envoyer des données ou commencent à envoyer des données corrompues.
- Échouer à effectuer des mises à jour critiques du système pour maintenir la disponibilité du système et la sécurité appropriée.



# 1- Motivation

Pourquoi tant de projets échouent-ils ? (3)

## **Résumé:**

- En réalité on traite les projets de science des données comme des projets de développement logiciels.
  - Au lieu de cela, les meilleures approches combinent le cycle de vie de la science des données avec un cadre de collaboration agile.

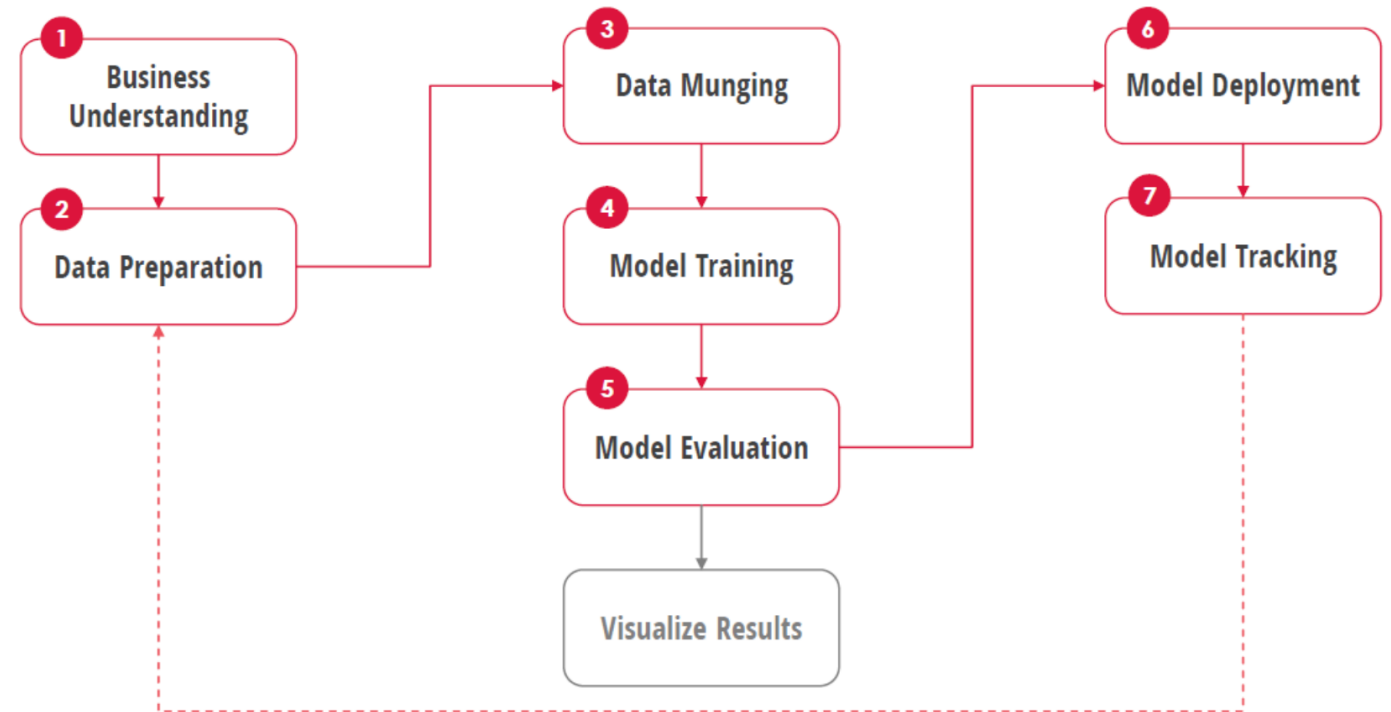
# 1- Motivation

Pourquoi tant de projets échouent-ils ? (3)

## Résumé:

- En réalité on traite les projets de science des données comme des projets de développement logiciels.
- Au lieu de cela, les meilleures approches combinent le cycle de vie de la science des données avec un cadre de collaboration agile.

Processus classique des projets en science de donnée (4)



**Image source:** MLOps and Cloud Native AI/ML, Prof. Fahd Kalloubi, Associate professor in data mining and Big Data

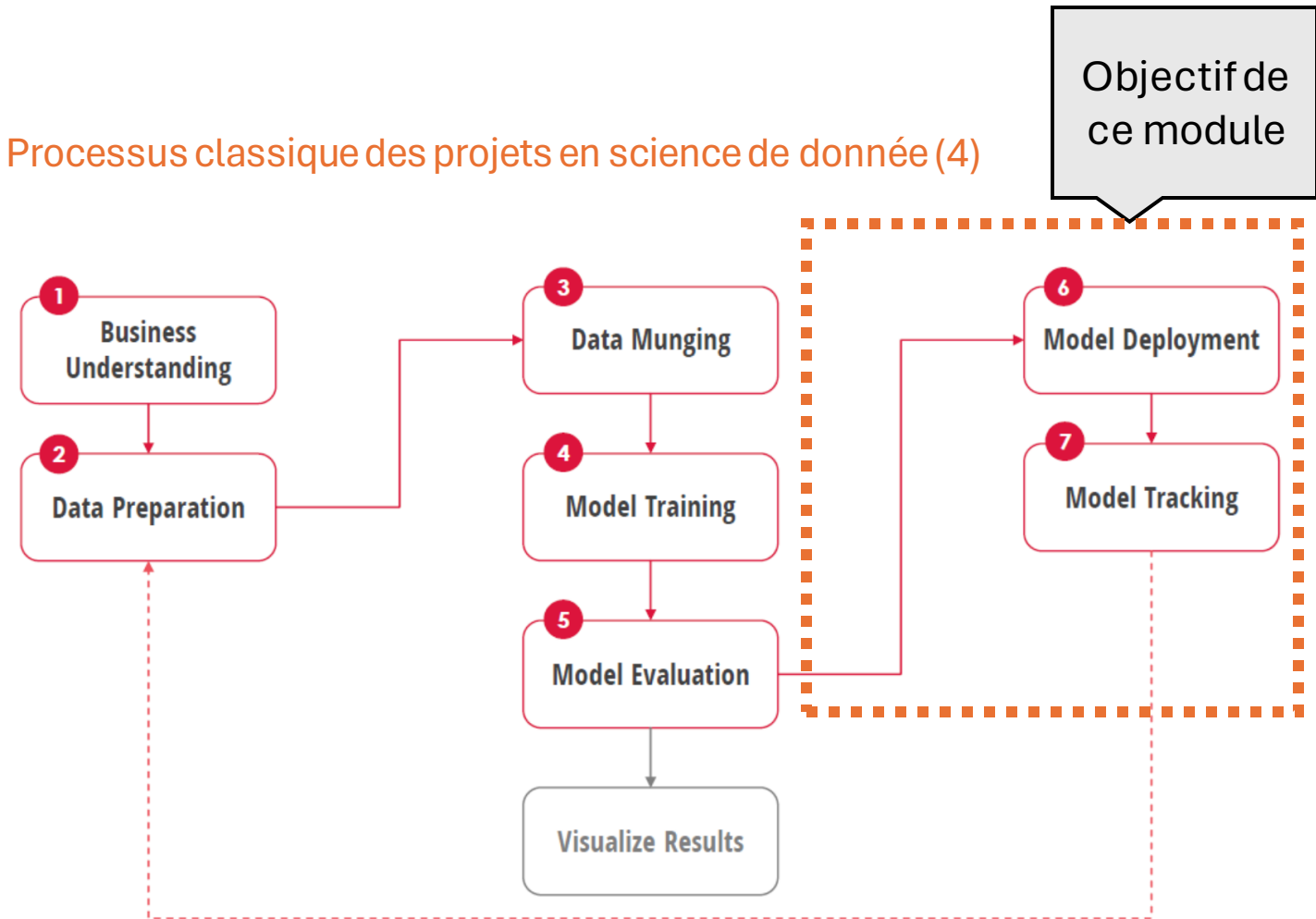
# 1- Motivation

Pourquoi tant de projets échouent-ils ? (3)

## Résumé:

- En réalité on traite les projets de science des données comme des projets de développement logiciels.
- Au lieu de cela, les meilleures approches combinent le cycle de vie de la science des données avec un cadre de collaboration agile.

Processus classique des projets en science de donnée (4)



**Image source:** MLOps and Cloud Native AI/ML, Prof. Fahd Kalloubi, Associate professor in data mining and Big Data

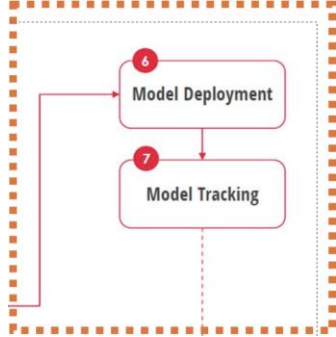
## 2 – Objectifs d'apprentissages

### Déploiement

- 1 - Déploiement d'un modèle dans une application web
- 2 – Déploiement d'un modèle comme un web service
- 3 – Déploiement conteneurisé sur le cloud

### Monitoring

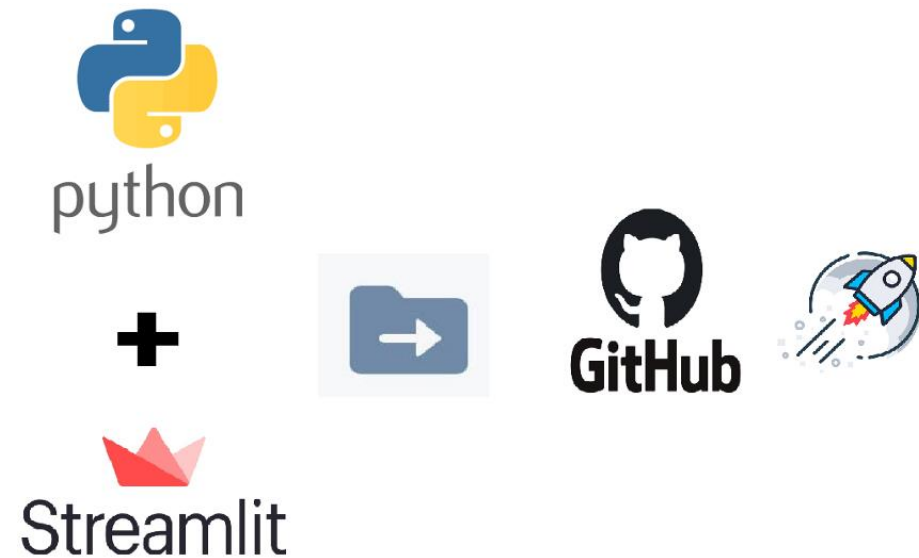
- 1 - Monitoring et maintenance
- 2 – Comment ajouter de nouvelles sources de données
- 3 – Version control



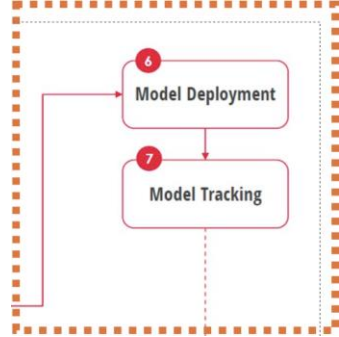
# 3– Exemple de déploiement

## Déploiement

1 - Déploiement d'un modèle dans une application web



[image source:](#) medium



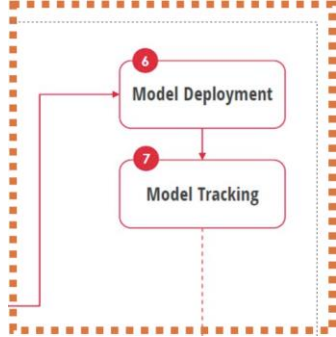
# 3– Exemple de déploiement

## Déploiement

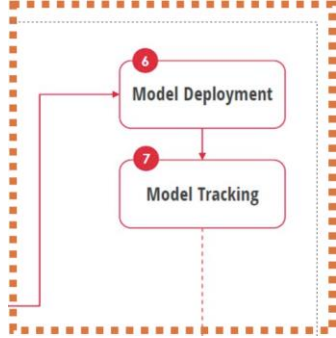
1 - Déploiement d'un modèle dans une application web



[image source:](#) medium



# 3– Exemple de déploiement



## Déploiement

1 - Déploiement d'un modèle dans une application web



Quelle est cette fleur?

[image source:](#) medium

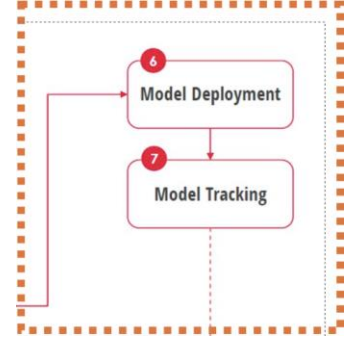
# 3– Exemple de déploiement

## Déploiement

1 - Déploiement d'un modèle dans une application web

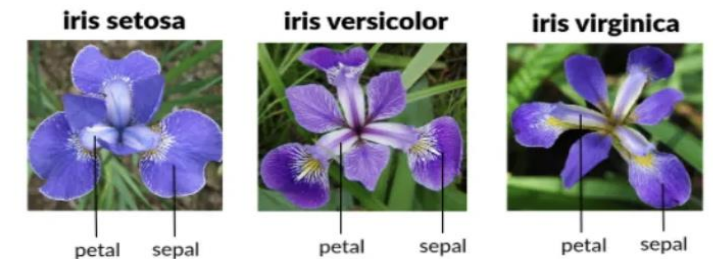


[image source:](#) medium



Lab 1: Iris prediction with Streamlit

## Iris Variety Prediction



Project 1 : Income Prediction ML App

## Income Predictor

The screenshot shows the interface of the 'Income Prediction ML App'. At the top, there is a dark green header with the text 'Income Prediction ML App' in white. Below the header, there are three input fields. The first field is labeled 'Age' and contains the value '0'. The second field is labeled 'Working Class' and has a dropdown menu with 'Federal-gov' selected. The third field is labeled 'Education' and is currently empty.



# 3– Exemple de déploiement

## Déploiement

### 2- Déploiement d'un model comme un web service

Un Web Service est une application qui permet d'échanger des données avec d'autres applications web. Même si ces dernières sont construites dans des langages de programmation différents.

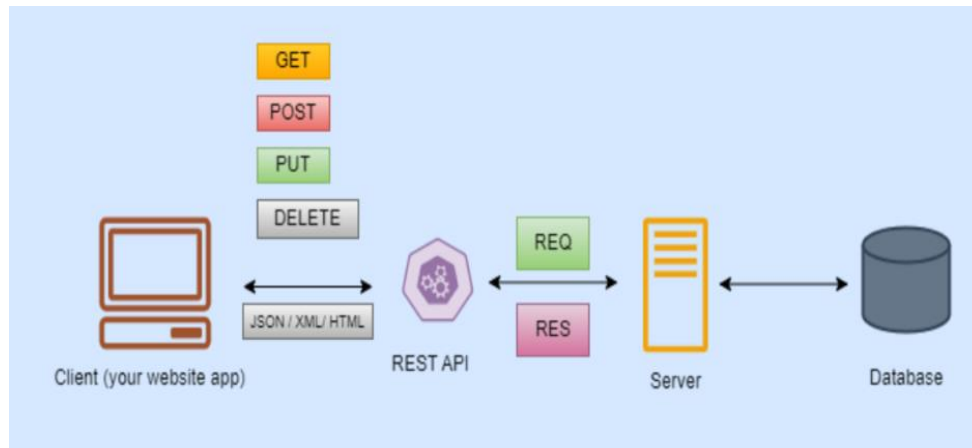
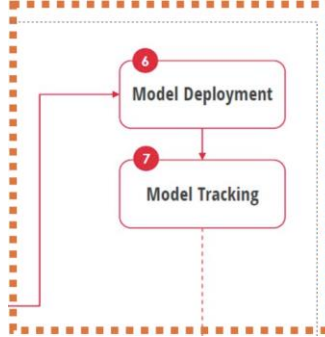


image source: analyticsvidhya



# 3– Exemple de déploiement

## Déploiement

### 2- Déploiement d'un modèle comme un web service

Un Web Service est une application qui permet d'échanger des données avec d'autres applications web. Même si ces dernières sont construites dans des langages de programmation différents.

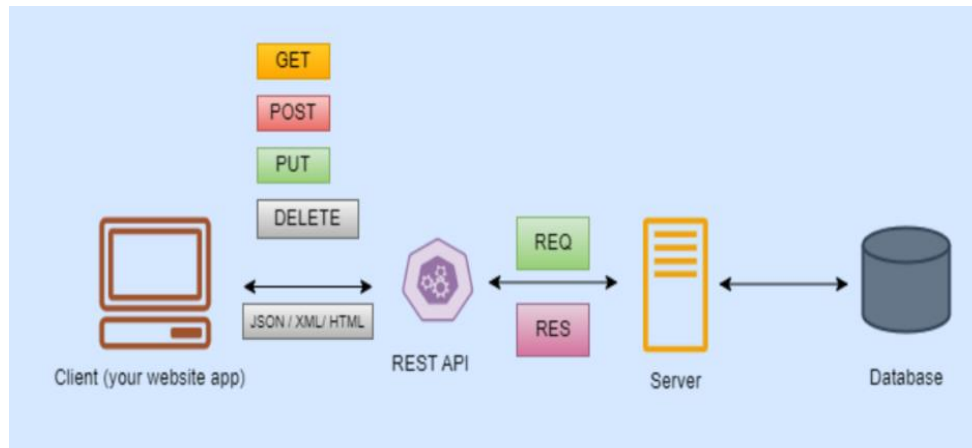
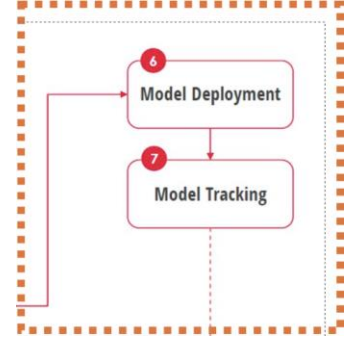
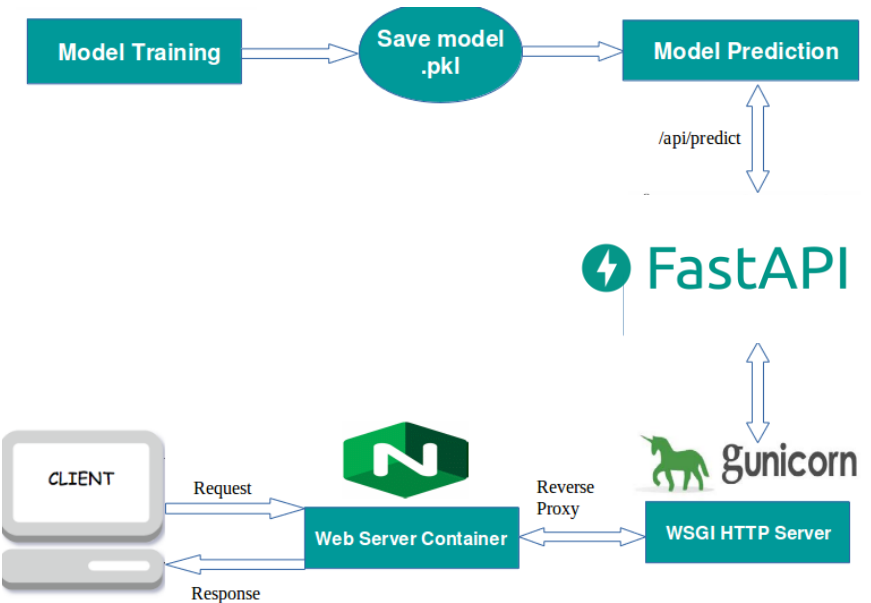


image source: analyticsvidhya



Projet:



# 3– Exemple de déploiement

## Déploiement

### 2- Déploiement d'un modèle comme un web service

Un Web Service est une application qui permet d'échanger des données avec d'autres applications web. Même si ces dernières sont construites dans des langages de programmation différents.

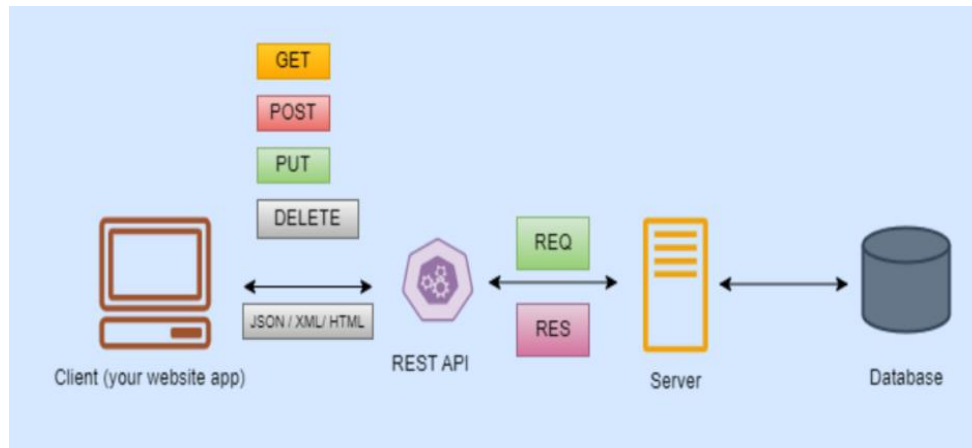


image source: analyticsvidhya

Projet:

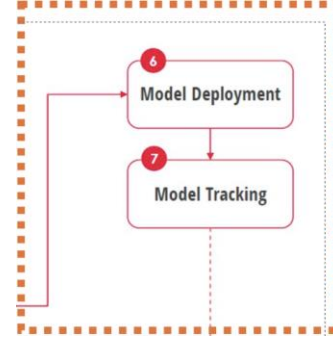


### Lab 2: Deploy ML model as API

Le Lab 2 vise à vous familiariser avec le déploiement d'un modèle d'apprentissage automatique en tant qu'API utilisant le framework FastAPI. Au cours de ce laboratoire, vous apprendrez à convertir un modèle d'apprentissage automatique pré-entraîné en une API web, accessible via des endpoints définis.

### Project 1 : ChurnPredictor

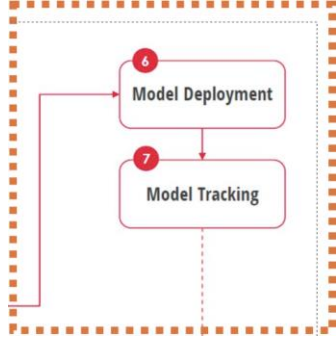
Nous construirons un service FastAPI pour servir ce model



# 3– Exemple de déploiement

## Déploiement

### 3- Déploiement conteneurisé



# 3– Exemple de déploiement

## Déploiement

### 3- Déploiement conteneurisé

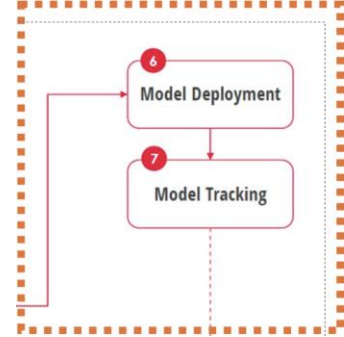


Objectif et compétences:

**1 - Création d'un service API FastAPI**

**2 - Conteneurisation avec Docker** : Encapsuler l'API FastAPI et le modèle d'apprentissage automatique dans des conteneurs Docker pour assurer la portabilité et la reproductibilité de l'environnement de déploiement.

**3 - Déploiement en tant que service conteneurisé** : Déployer le service conteneurisé sur une plateforme cloud ou sur une infrastructure locale, permettant ainsi un déploiement et une mise à l'échelle efficaces tout en maintenant l'isolation et la cohérence de l'environnement.



# 3– Exemple de déploiement

## Déploiement

### 3- Déploiement conteneurisé



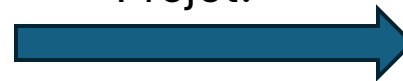
Objectif et compétences:

**1 - Création d'un service API FastAPI**

**2 - Conteneurisation avec Docker** : Encapsuler l'API FastAPI et le modèle d'apprentissage automatique dans des conteneurs Docker pour assurer la portabilité et la reproductibilité de l'environnement de déploiement.

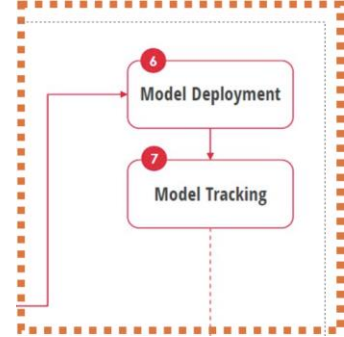
**3 - Déploiement en tant que service conteneurisé** : Déployer le service conteneurisé sur une plateforme cloud ou sur une infrastructure locale, permettant ainsi un déploiement et une mise à l'échelle efficaces tout en maintenant l'isolation et la cohérence de l'environnement.

Projet:



**Lab 2:** Deploy ML model as API + Docker

**Project 1 :** My LLAMA + Docker



# 3– Exemple de déploiement

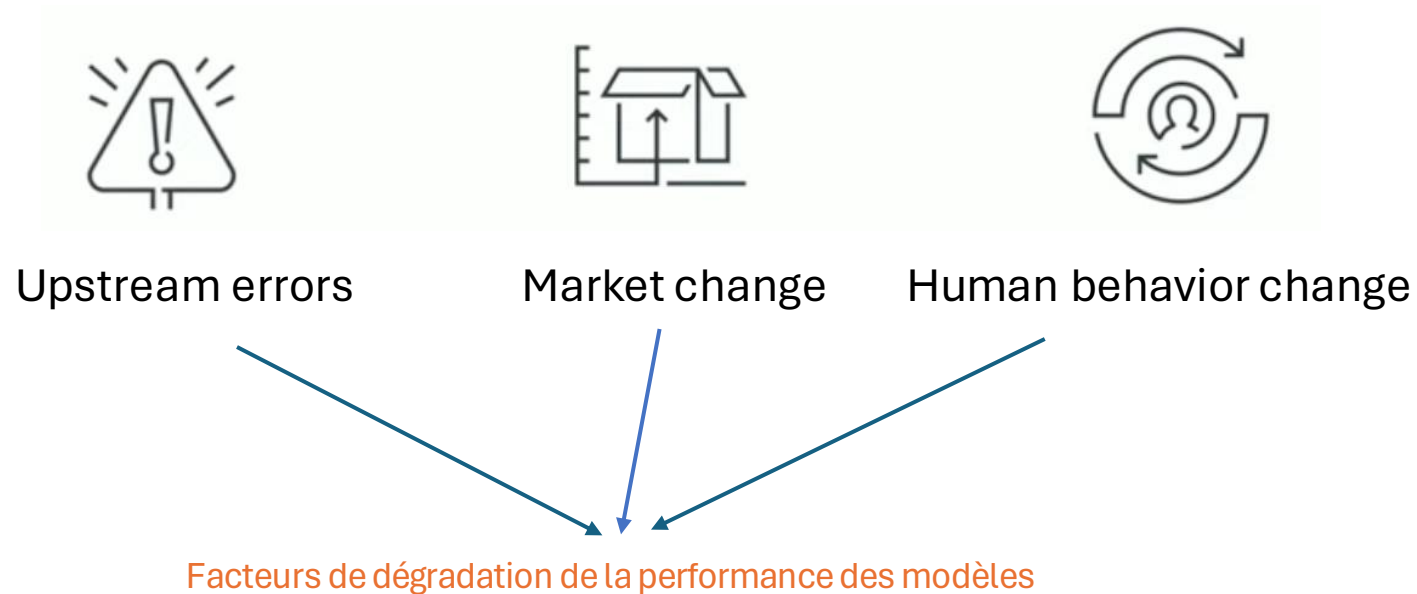
Déploiement & Monitoring

Rappel: "Ne pas penser que le déploiement est la dernière étape"

# 3– Exemple de déploiement

## Déploiement & Monitoring

- La distribution des données et des caractéristiques peut changer à travers le temps:



....

....



# 3– Exemple de déploiement

## Déploiement & Monitoring

- La distribution des données et des caractéristiques peut changer à travers le temps:

**Certitude:** La performance des modèles va se dégrader avec le temps

**Challenge:** Être au courant des que cela survient

# 3– Exemple de déploiement

## Déploiement & Monitoring

- La distribution des données et des caractéristiques peut changer à travers le temps:

**Certitude:** La performance des modèles va se dégrader avec le temps

**Challenge:** Être au courant des que cela survient

Ce changement on va l'appeler le "Model Drift":

....

# 3– Exemple de déploiement

## Déploiement & Monitoring

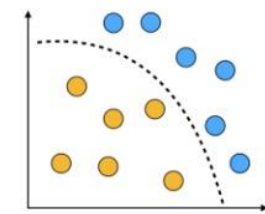
- La distribution des données et des caractéristiques peut changer à travers le temps:

**Certitude:** La performance des modèles va se dégrader avec le temps

**Challenge:** Être au courant des que cela survient

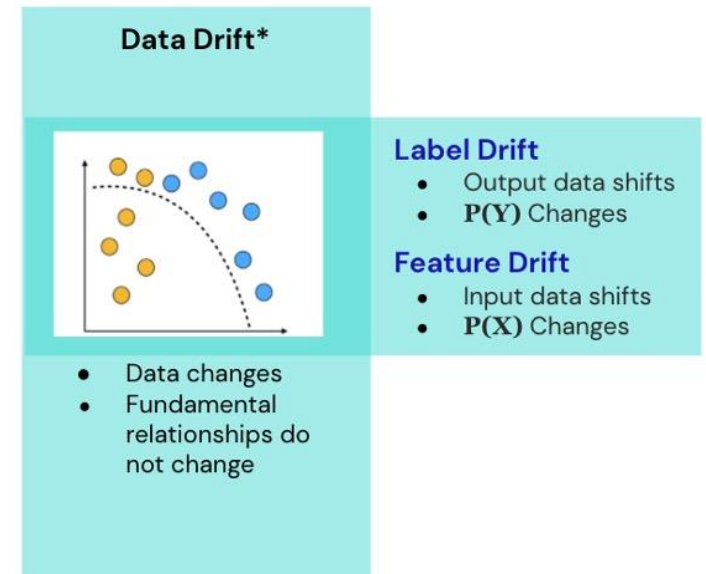
Ce changement on va l'appeler le "Model Drift":

1 - Label/features drift



Training data with  
decision boundary

$P(Y|X)$   
Probability of  
y output  
given x input



....

# 3– Exemple de déploiement

## Déploiement & Monitoring

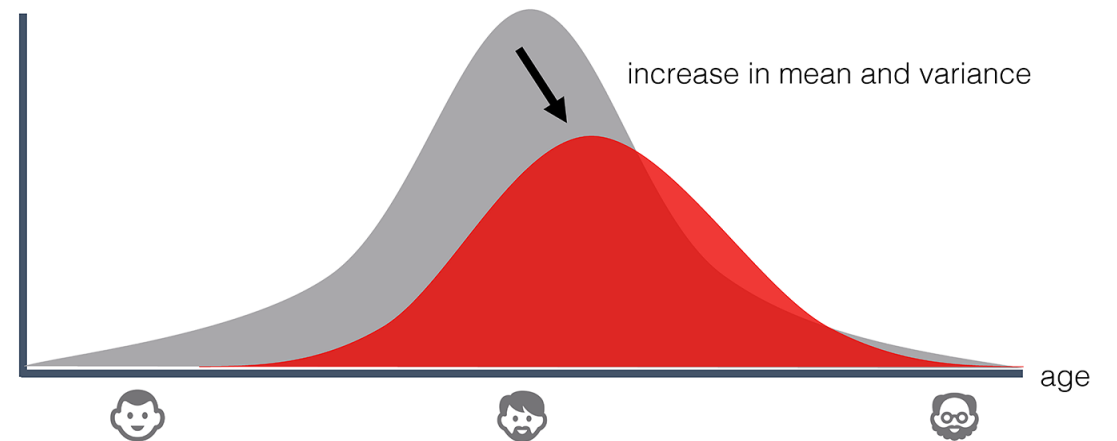
- La distribution des données et des caractéristiques peut changer à travers le temps:

**Certitude:** La performance des modèles va se dégrader avec le temps

**Challenge:** Être au courant des que cela survient

Ce changement on va l'appeler le "Model Drift":

1 - Label/features drift



....

# 3– Exemple de déploiement

## Déploiement & Monitoring

- La distribution des données et des caractéristiques peut changer à travers le temps:

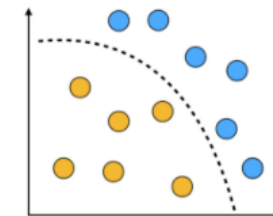
**Certitude:** La performance des modèles va se dégrader avec le temps

**Challenge:** Être au courant des que cela survient

Ce changement on va l'appeler le "Model Drift":

### 2 - Concept drift

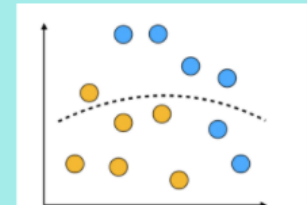
....



Training data with  
decision boundary

$P(Y|X)$   
Probability of  
y output  
given x input

#### Concept Drift $P(Y|X)$ Changes



- Reality/behavioral change
- Relationships change, not the input

# 3– Exemple de déploiement

## Déploiement & Monitoring

- La distribution des données et des caractéristiques peut changer à travers le temps:
- **Besoin:** Nous devons monitorer le model?
  - 1 – Résumé statistique de base des caractéristiques et de la cible
  - 2 - Distribution des caractéristiques et de la cible
  - 3 - Métriques de performance du modèle
  - 4 - Métriques métier

....

# 3– Exemple de déploiement

## Déploiement & Monitoring

- La distribution des données et des caractéristiques peut changer à travers le temps:
- **Besoin:** Nous devons monitorer le model?
  - 1 – Résumé statistique de base des caractéristiques et de la cible
  - 2 - Distribution des caractéristiques et de la cible
  - 3 - Métriques de performance du modèle
  - 4 - Métriques métier
- Comment effectuer ce monitoring? Quel outils?

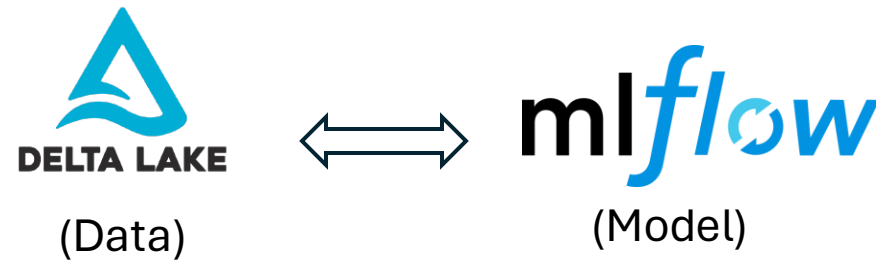
....

# 3– Exemple de déploiement

## Déploiement & Monitoring

- Comment effectuer ce monitoring? Quel outils?

1 - Logging and Versioning



2- Test statistiques

SciPy

....



# 3– Exemple de déploiement

## Déploiement & Monitoring

- Comment effectuer ce monitoring? Quel outils?

### 1 - Logging and Versioning



(Data)



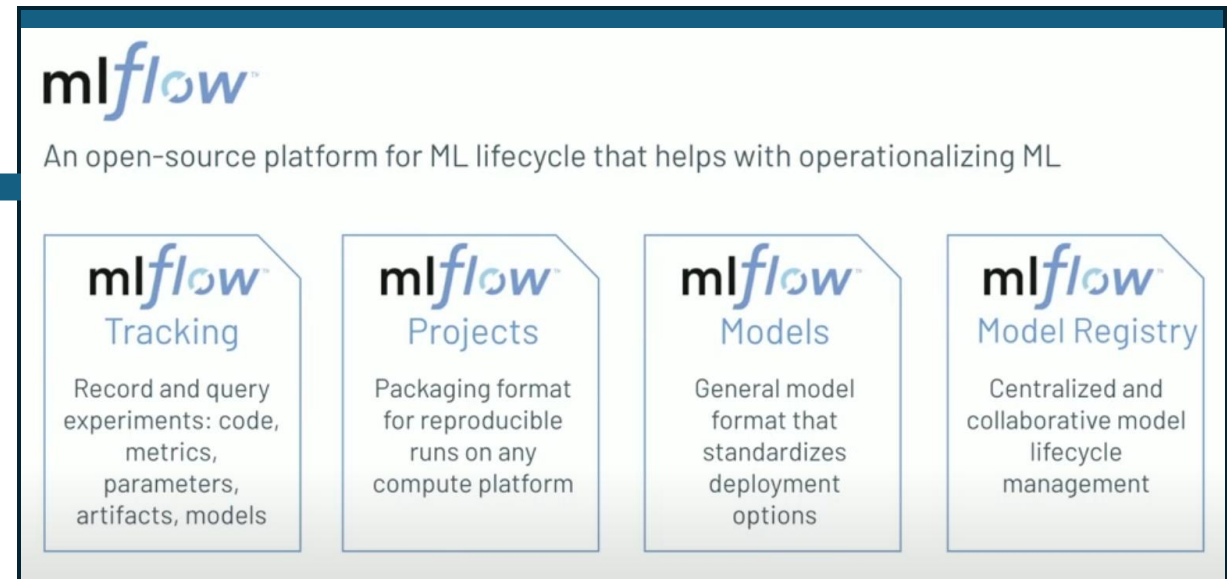
**mlflow**

(Model)

### 2- Test statistiques

SciPy

....



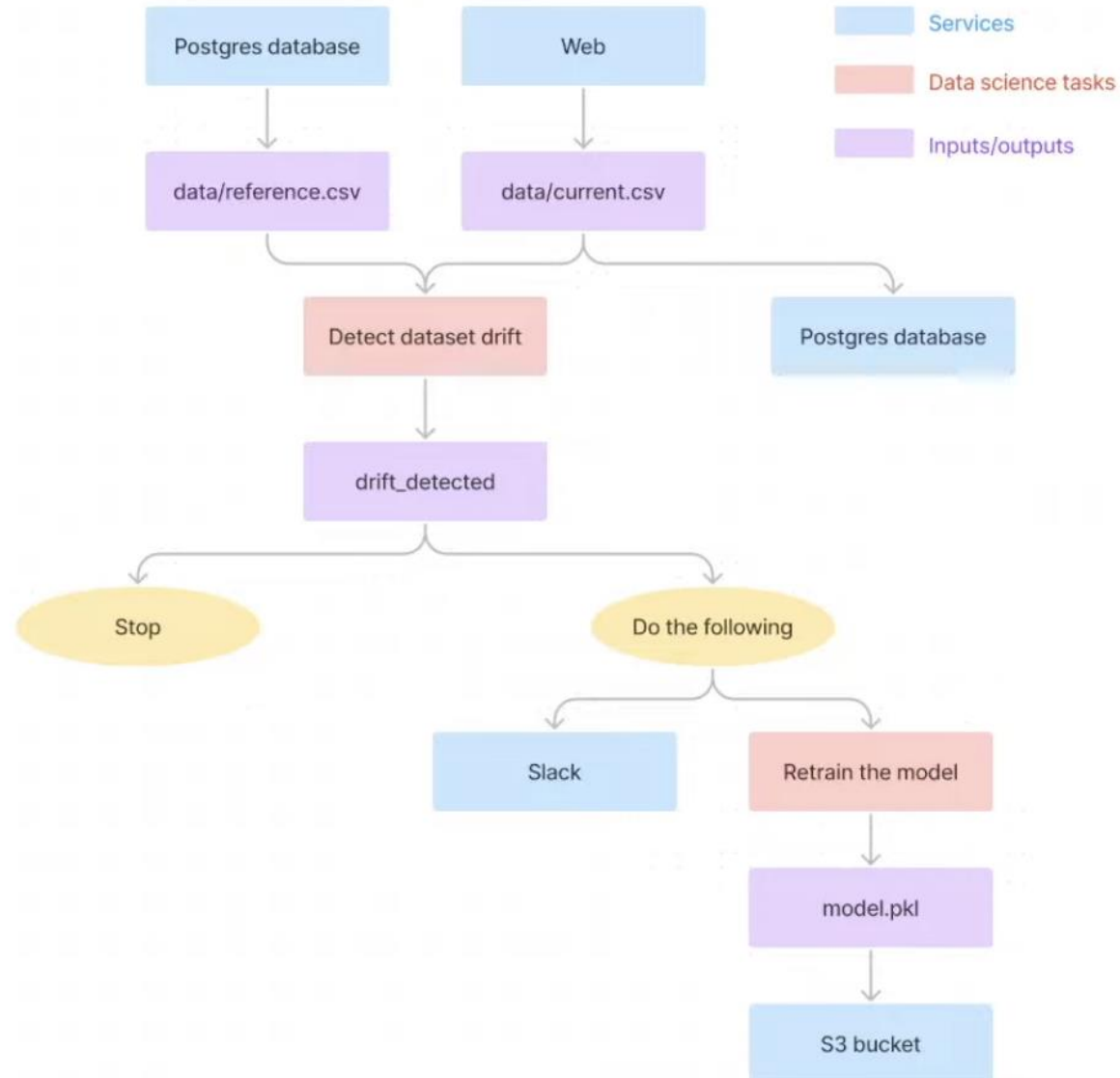
# 3– Exemple de déploiement

Déploiement & Monitoring

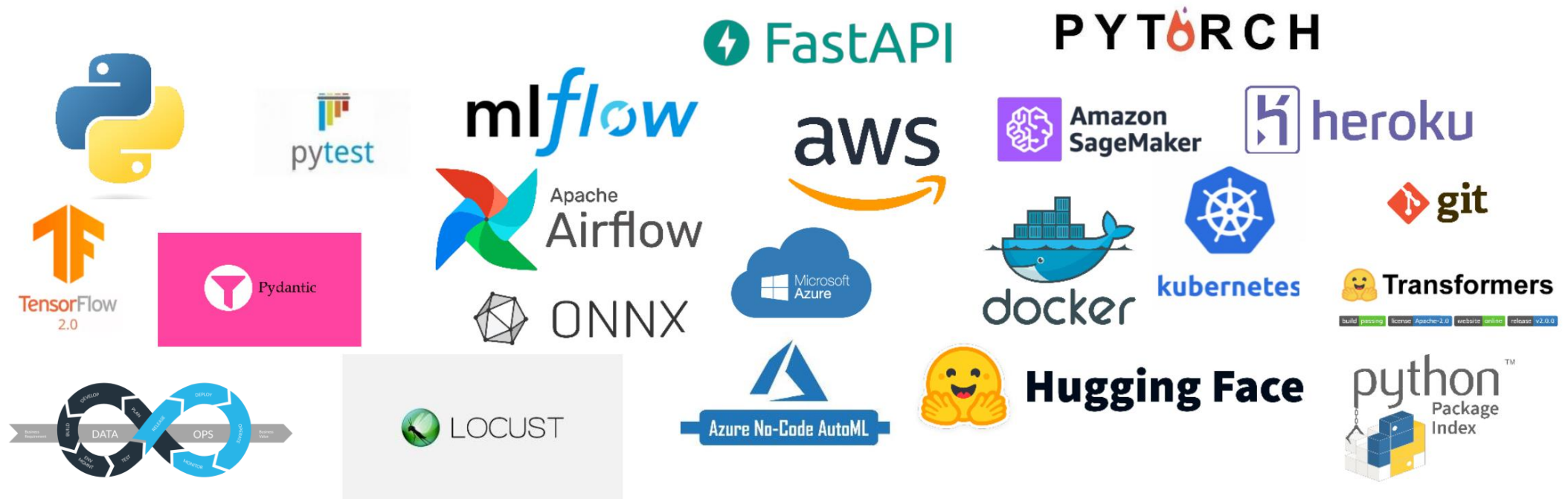


Lab 3 & Project 3

....



# 4 – Résumé



# 5 - Tendances

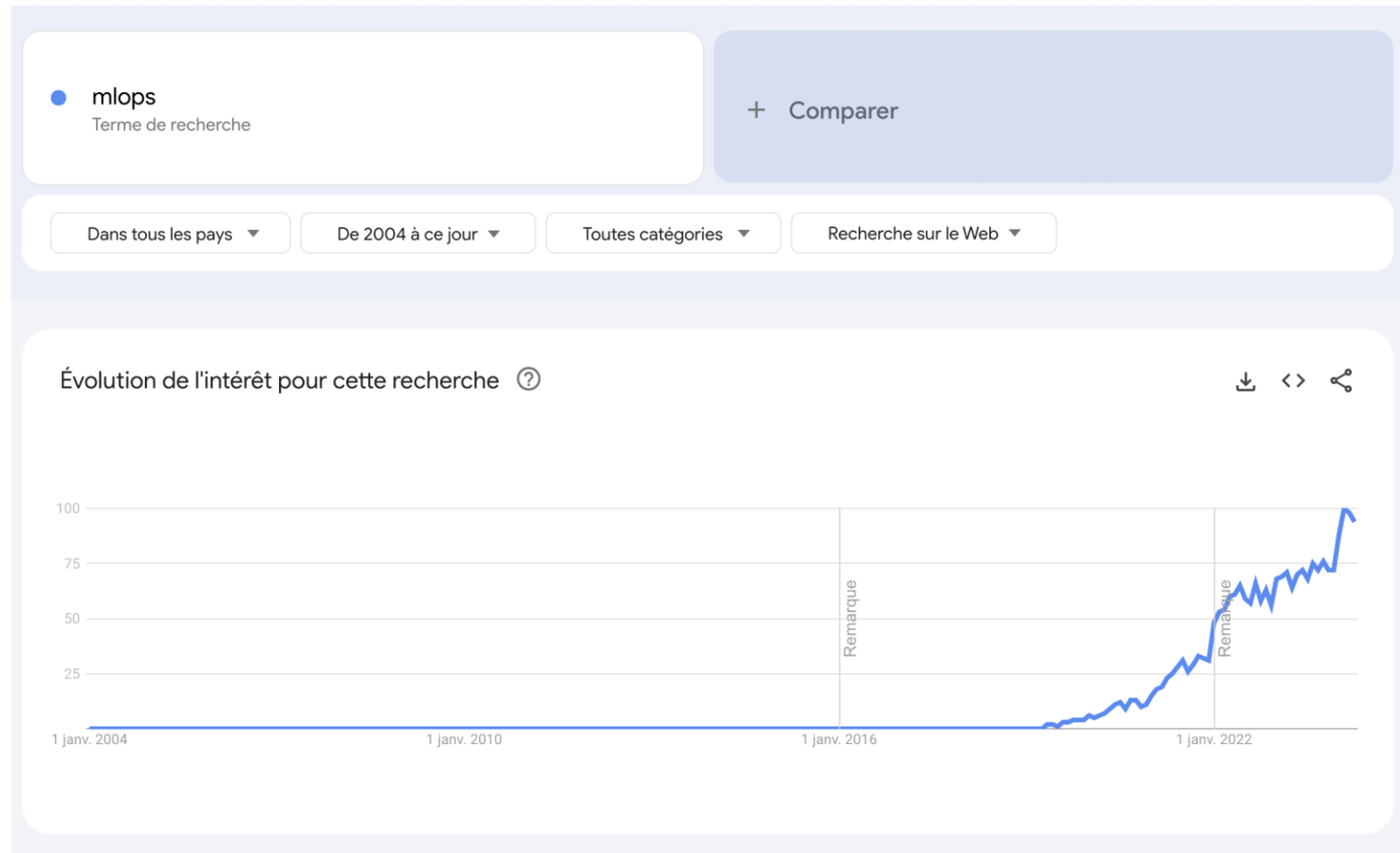
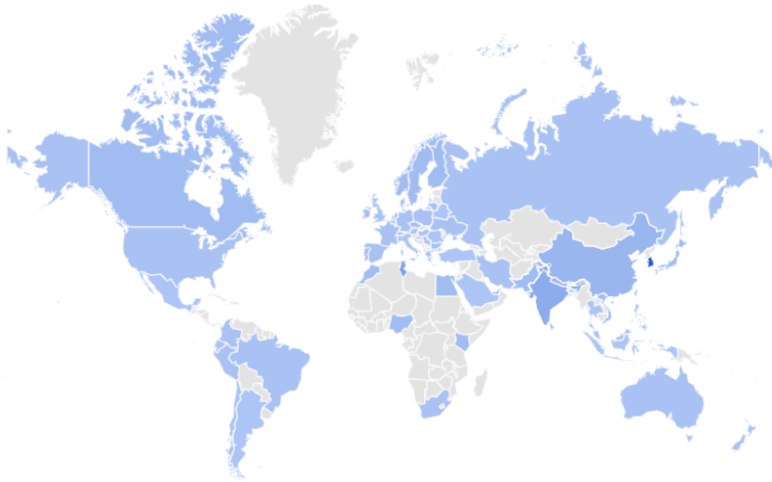


image source: GoogleTrends

## 5 - Tendances

Intérêt par région ?



☐ Inclure les régions à faible volume de recherche

Région ▼



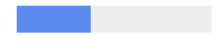
1 Corée du Sud

100



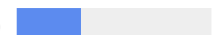
2 Singapour

38



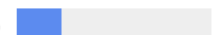
3 Tunisie

33



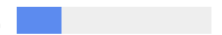
4 Inde

23



5 Israël

23



< Affichage des régions 1 à 5 sur 52 >

image source: GoogleTrends



Merci pour  
votre attention.

---





## MLOps roadmap



# Data and Machine learning operationalization

[Building and deploying ml app]

Presented By:  
Généreux Akotenou



[github.com/genereux-akotenou](https://github.com/genereux-akotenou)



[genereux.akotenou@gmail.com](mailto:genereux.akotenou@gmail.com)

27 Avril 2024