

The background image shows a forest fire at night. Bright orange flames and thick black smoke rise from the trees in the distance. In the foreground, two firefighters wearing helmets and gear are seen from behind, looking towards the fire. The overall scene is dark, with the fire providing the primary light source.

# TERRE-VENT-FEU-EAU DATA

***“Voir le feu avant qu’il ne commence”***

Romuald COURTOIS  
MSc 1 IA / Data Science



## CONTEXTE



7 août 2025

# GEO

+1 million d'hectares brûlés (1)

# Le Monde

Superficie brûlée : +250%  
par rapport à la moyenne des 18 années précédentes (2)

1. Geo. 22/08/2025. Incendies en Europe: 2025 est déjà une année record, avec plus d'un million d'hectares brûlés.

2. Léa Prati, William Audureau et Pierre Breteau. 08/08/2025. Le Monde. Visualisez le niveau exceptionnel des incendies en France en 2025.



## PROBLEMATIQUE



INSTITUT NATIONAL  
DE L'INFORMATION  
GÉOGRAPHIQUE  
ET FORESTIÈRE

Base de Données sur les Incendies de Forêt en France  
(BDIFF) (1)

- Fragmentation des données
- Absence de géolocalisation précise
- Analyse complexe
- 90 % des départs de feu d'origine humaine



Développement d'outils d'intelligence prédictive

- Améliorer la prévention
- Réduire l'impact des incendies

## OBJECTIFS

Développer un outil prédictif du risque d'incendie



Reconstituer un  
dataset unifié



Enrichir les  
données



Analyses multi-  
échelles



Bases d'une  
version future

- Base de données SQL consolidée
- Application Streamlit interactive
- Modèle de Machine Learning



## ACCES AU DONNEES

**Base de Données sur les Incendies de Forêts en France (BDIFF)**



1973

2006

2023

**Premiers  
enregistrement des  
feu de foret en France**

- Sud de la France
- Corse

**Création de la BDIFF**

- Centralisé
- Standardisé

**Nouvelle nomenclature  
des surfaces**

- Forêt
- Autres surfaces naturelles HF
- Surfaces agricoles
- Autres surfaces

## OUTILS ET TECHNOLOGIE

Gestion du projet et développement



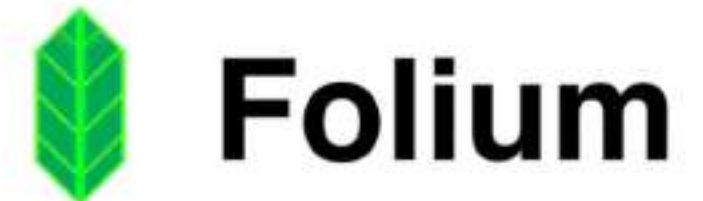
Stockage et gestion des données



Modèles de machine learning

- Scikit-learn
- XGBoost / LightGBM
- HDBSCAN

Visualisation et interface utilisateur





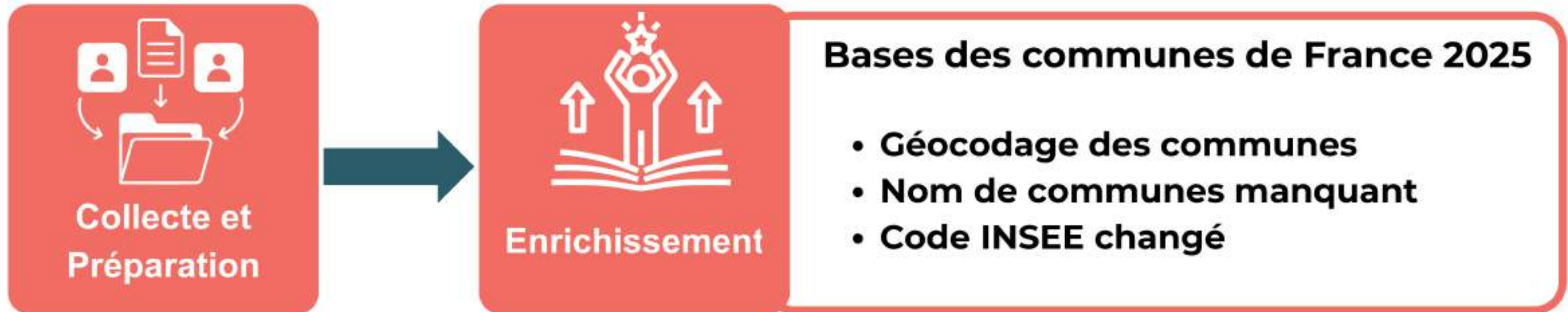
## ROADMAP



**Collecte et  
Préparation**

- **Extraction par paquet de 10 ans**
    - = 140 247 Feu
  - **Exclusion de colonnes non pertinentes**
    - Peu de données
    - Peu utiles
  - **annee ; numero ; code\_insee ; Nom de commune ;  
date\_premiere\_alerte ; Surface\_parcourue\_m2 ;  
type\_de\_peuplement ; precision\_de\_la\_donnee**
- 
- **Géolocalisation**
  - **Nom de ville**
  - **Doublons**

## ROADMAP





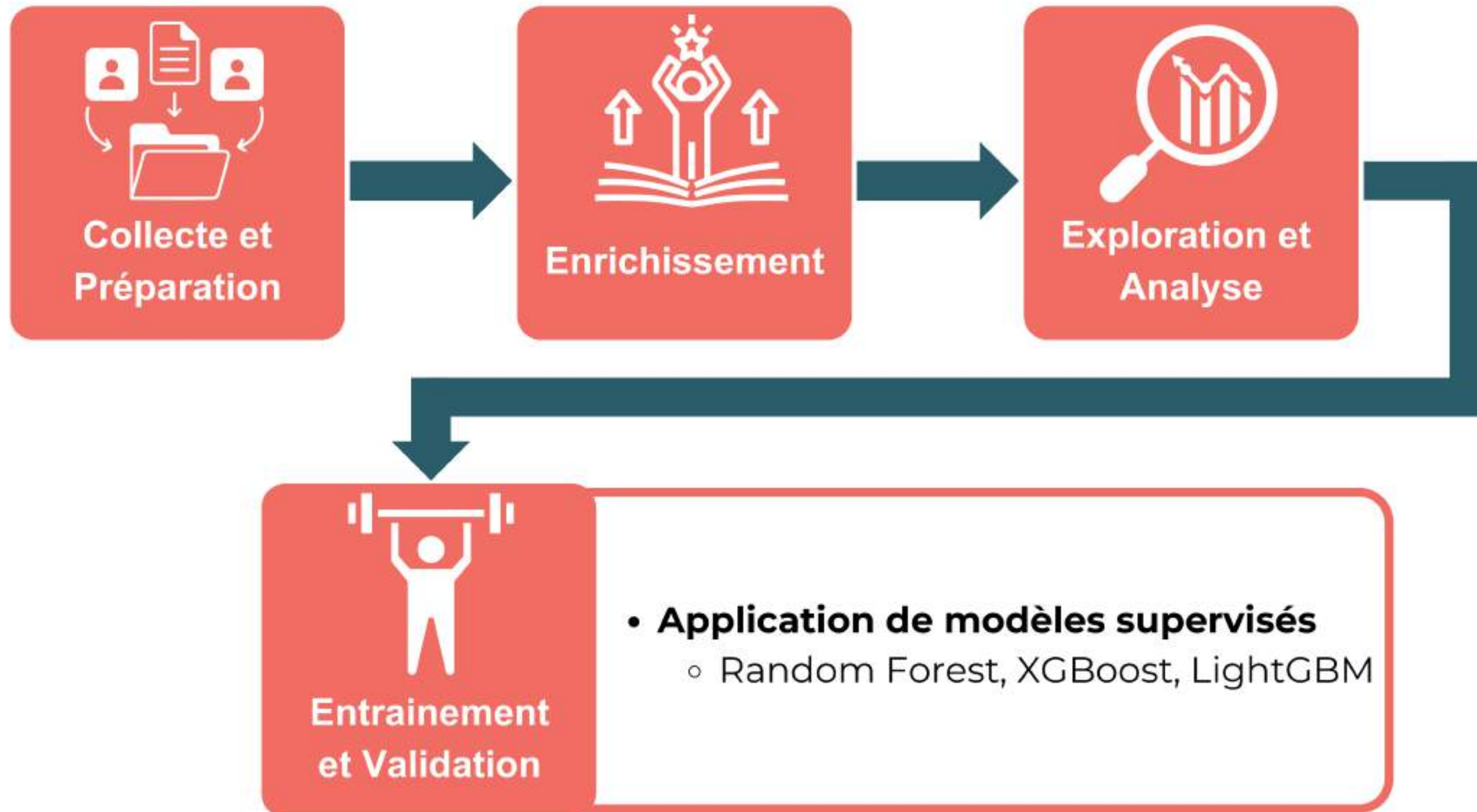
## ROADMAP



- **Suppression des valeurs manquantes**
- **Suppression des doublons**
  - 2.3 % = 3,219 feux
- **680 feu sans coordonnées**

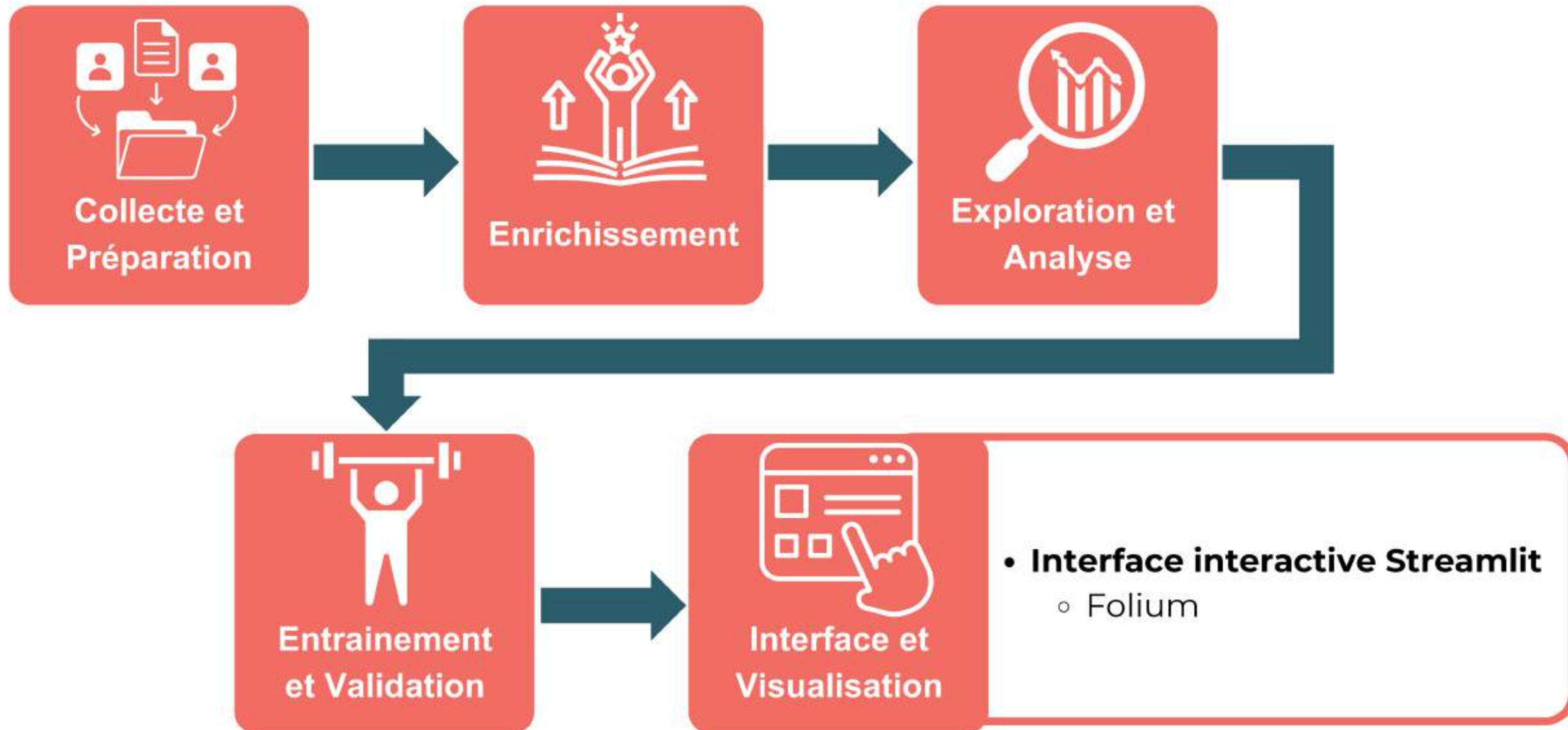
- **Identification des tendances spatiales et temporelles**
- **Détection de clusters**
  - HDBSCAN
- **Traitement des valeurs aberrantes**

## ROADMAP





## ROADMAP



## FIABILITE / LIMITES

### **Analyses statistiques**

- Descriptives
- Série temporelle
- Spatiales
- Corrélations

Dataset unifié

- **Biais temporel**
- **Biais de déclaration**
- **Colonnes exclues**
- **Limites méthodologiques**
  - **variables exogènes manquantes (météo)**



## PERSPECTIVES

- **Développement des modèles**
  - HDBSCAN
  - Machine learning
- **Création de l'application interactive**

- **Intégration des données météorologique**
- **Tester des modèles de deep learning**
  - RNN
  - LSTM



**Outil prédictif opérationnel pour les collectivités**



A dramatic night scene of a forest fire. In the background, a large fire burns brightly, sending thick smoke into the dark sky. The fire's glow illuminates the silhouettes of tall evergreen trees. In the foreground, two firefighters are seen from behind, wearing orange protective gear and helmets. They are looking towards the burning forest. The overall atmosphere is one of urgency and danger.

# TERRE-VENT-FEU-EAU DATA

***“Voir le feu avant qu’il ne commence”***

Romuald COURTOIS  
MSc 1 IA / Data Science