

FEVRIER 2026

SYSTÈME DE SURVEILLANCE DES DONNÉES DE PRESSION ARTÉRIELLE PATIENT

avec Kafka, Elasticsearch et Kibana

Presented by

Hamrouni Jinene
Bourbia Khoubeib
Hakimi Tasnime



TABLE of contents

01. Introduction

02. Problématique & Objectif

03. Architecture du système

04. Génération & Streaming des données

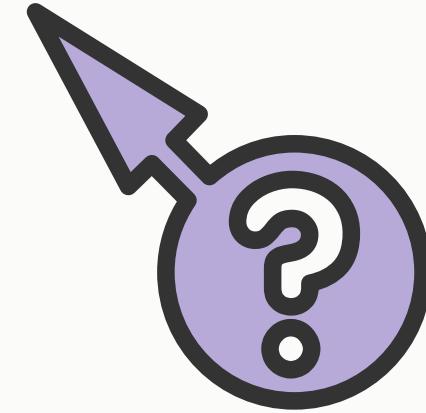
05. Visualisation avec Kibana

06. Machine Learning

07. Visualisation avec Flask

08. Conclusion

Contexte général & besoin

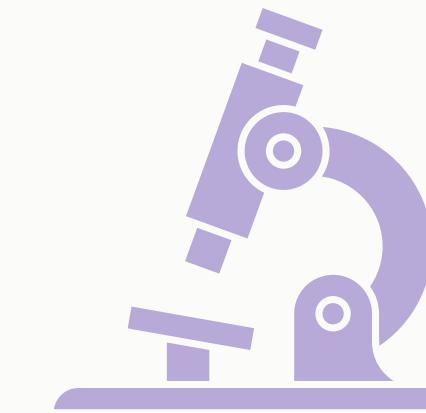


Contexte

Digitalisation de la santé

Objets connectés

Données physiologiques continues



Constats

Volume massif de données

Flux en temps réel

Analyse manuelle inefficace



Besoin

Traitement automatisé

Détection d'anomalies

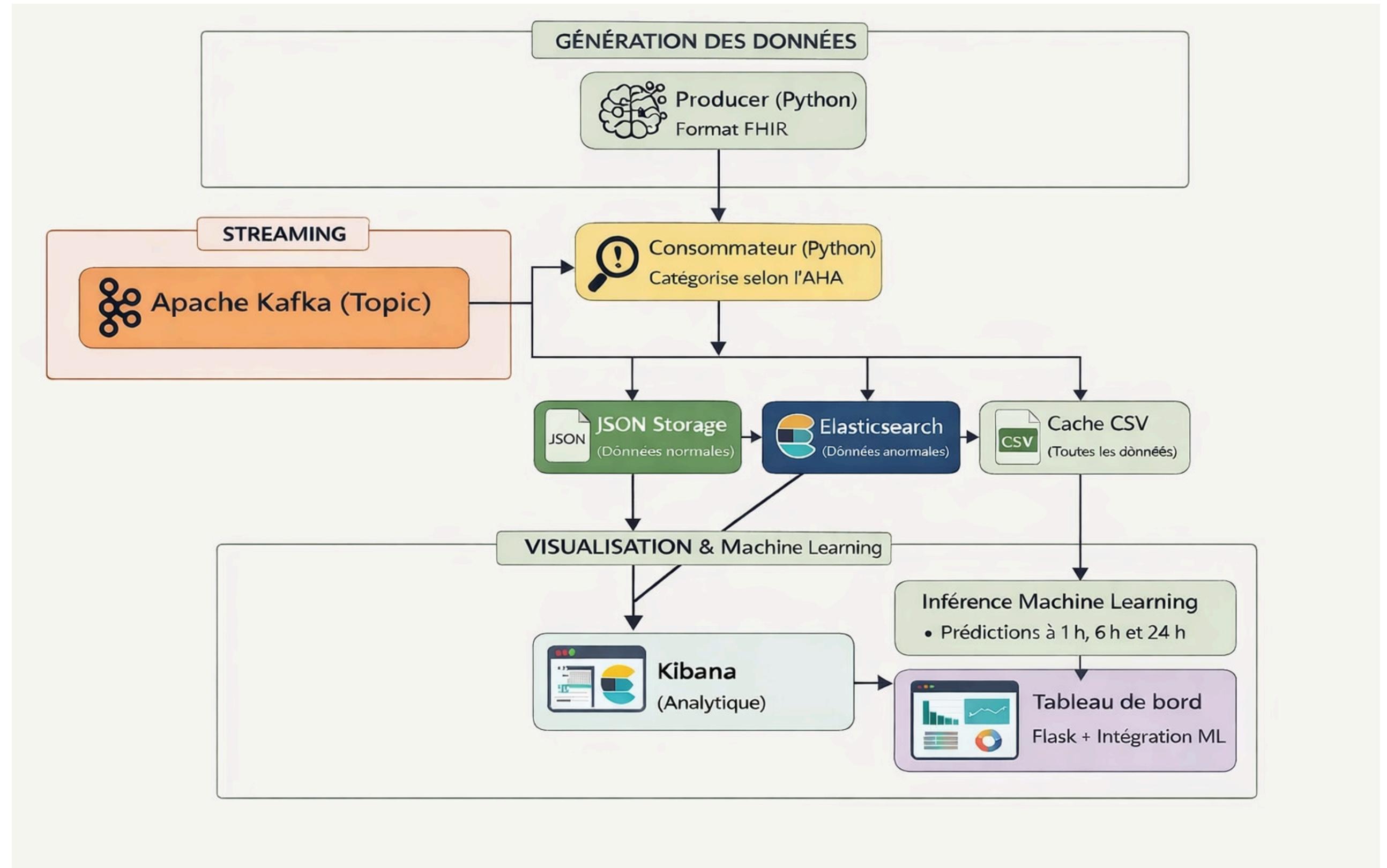
Détection d'anomalies

Problématique

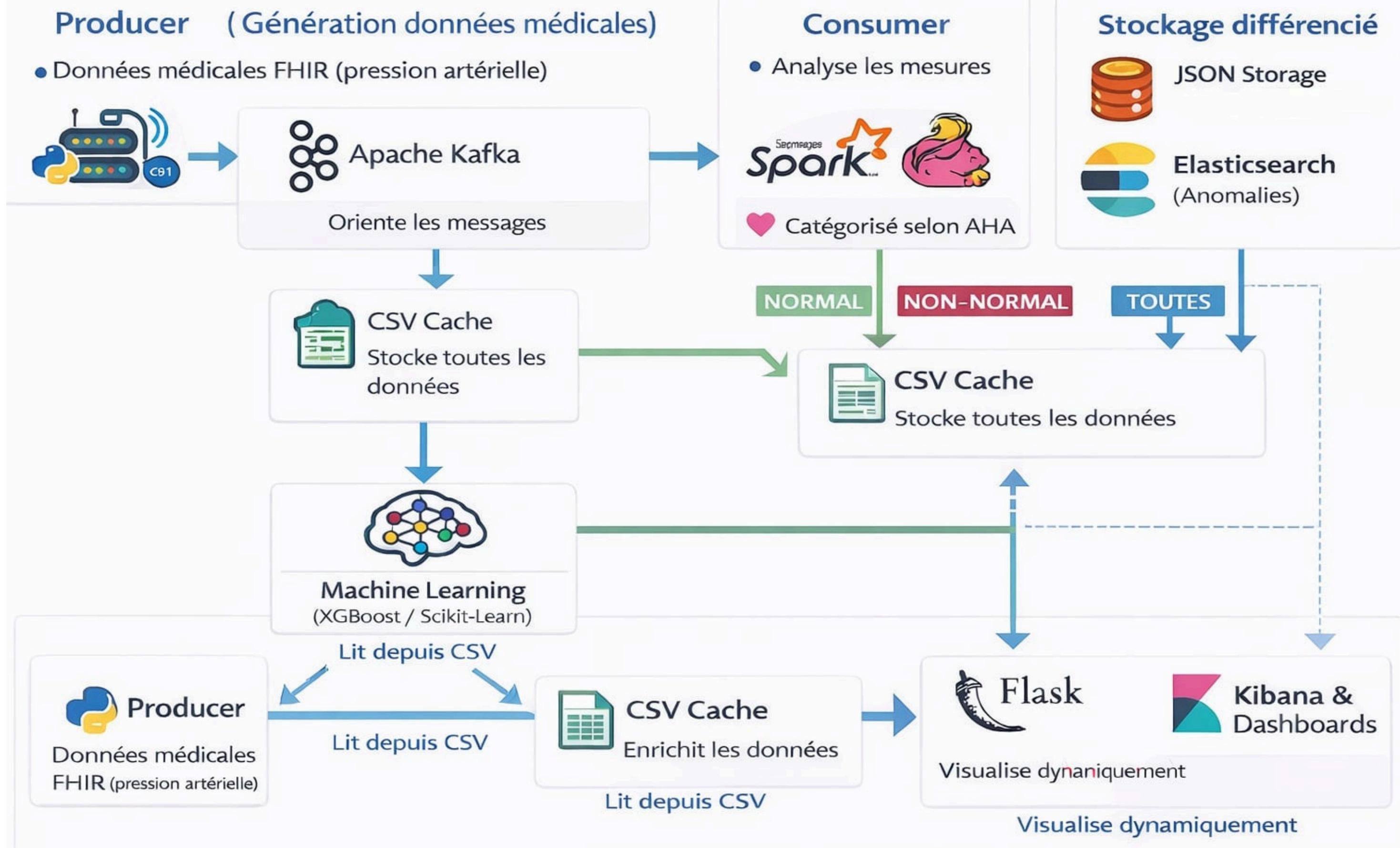
- ? Comment gérer un flux massif de données médicales ?
- ? Comment détecter automatiquement les situations à risque ?
- ? Comment organiser le stockage selon la criticité ?
- ? Comment aider les médecins à prendre des décisions rapides ?



Architecture Globale



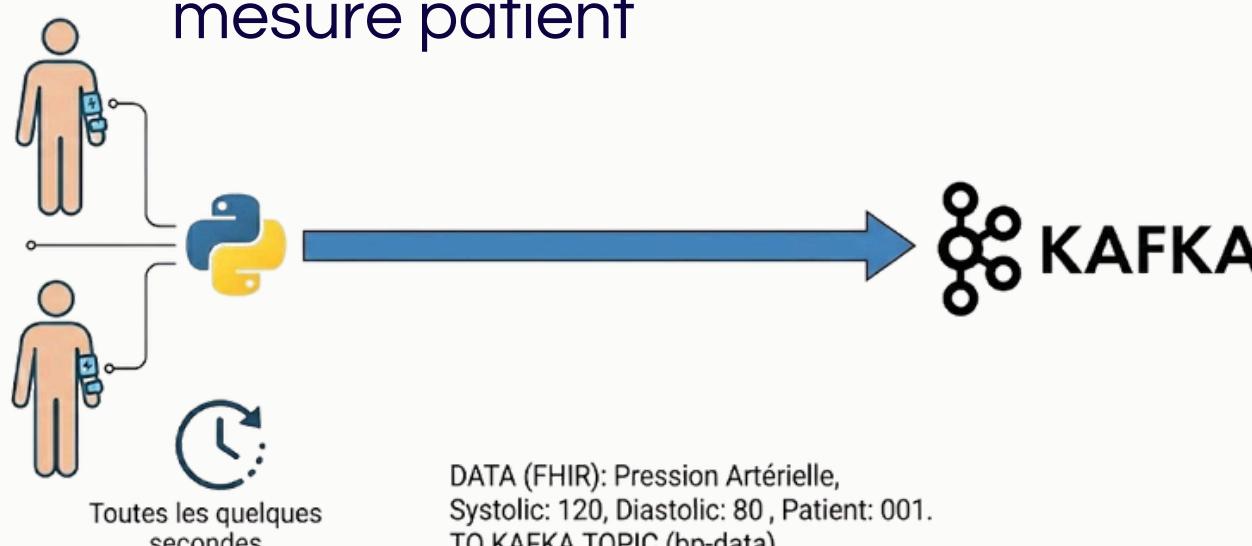
RELATIONS ENTRE LES COMPOSANTS



Génération & Streaming des Données

2. Kafka Producer

- Producteur Kafka en Python
- Envoi des messages vers un topic Kafka
- Envoi à intervalles réguliers
- Chaque message = une nouvelle mesure patient



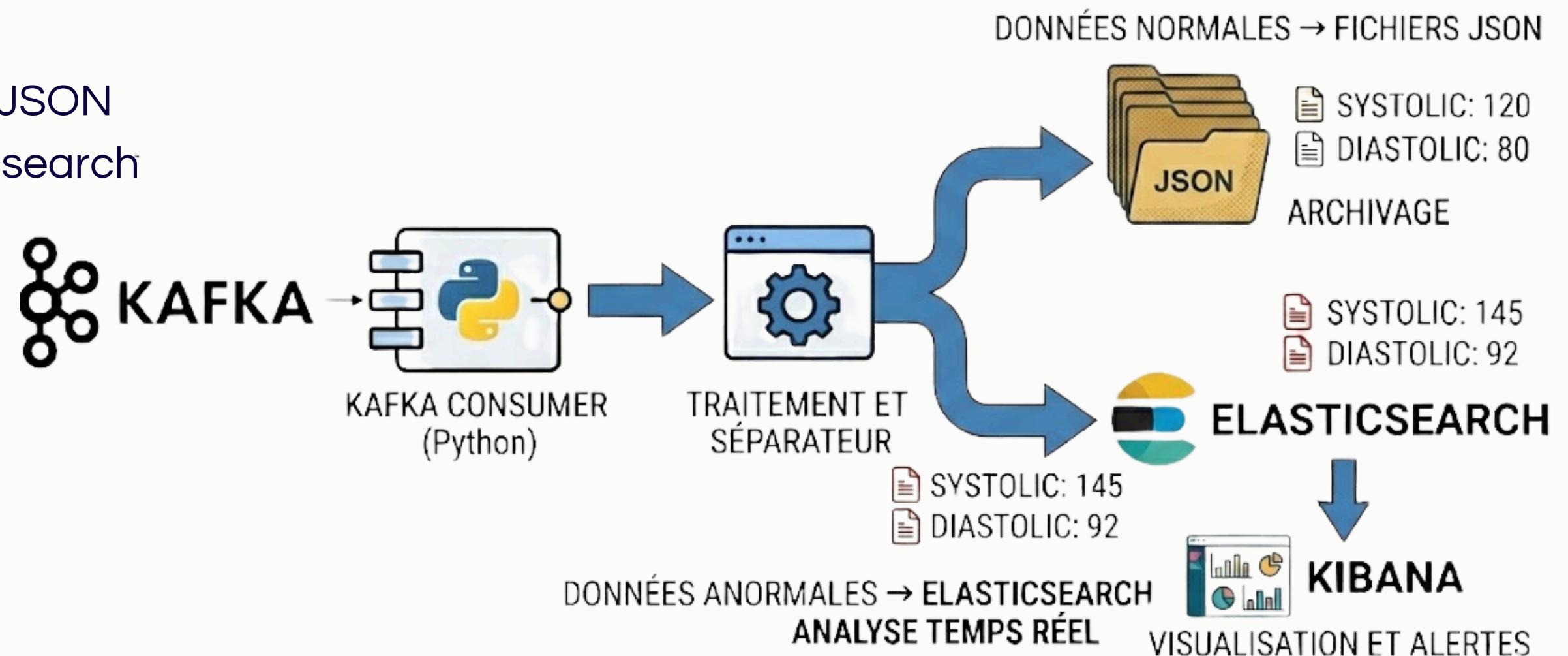
3. Kafka Consumer & Détection

- Lecture des messages depuis Kafka
- Extraction des valeurs médicales
- Détection d'anomalies :
 - Systolique > 140 ou < 90
 - Diastolique > 90 ou < 60
- Classification :
 - Normal
 - À risque

Génération & Streaming des Données

4. Stockage différencié

- Données normales → fichiers JSON
- Données anormales → Elasticsearch
- Objectif :
 - conserver l'historique
 - prioriser les cas critiques





Génération & Streaming des Données

1. Simulation des données

- 50 patients simulés
- Chaque patient possède :
 - un identifiant unique
 - pression systolique
 - pression diastolique
 - rythme cardiaque
 - horodatage
- Format standard médical : FHIR (JSON)
- Mise à jour continue des mêmes patients

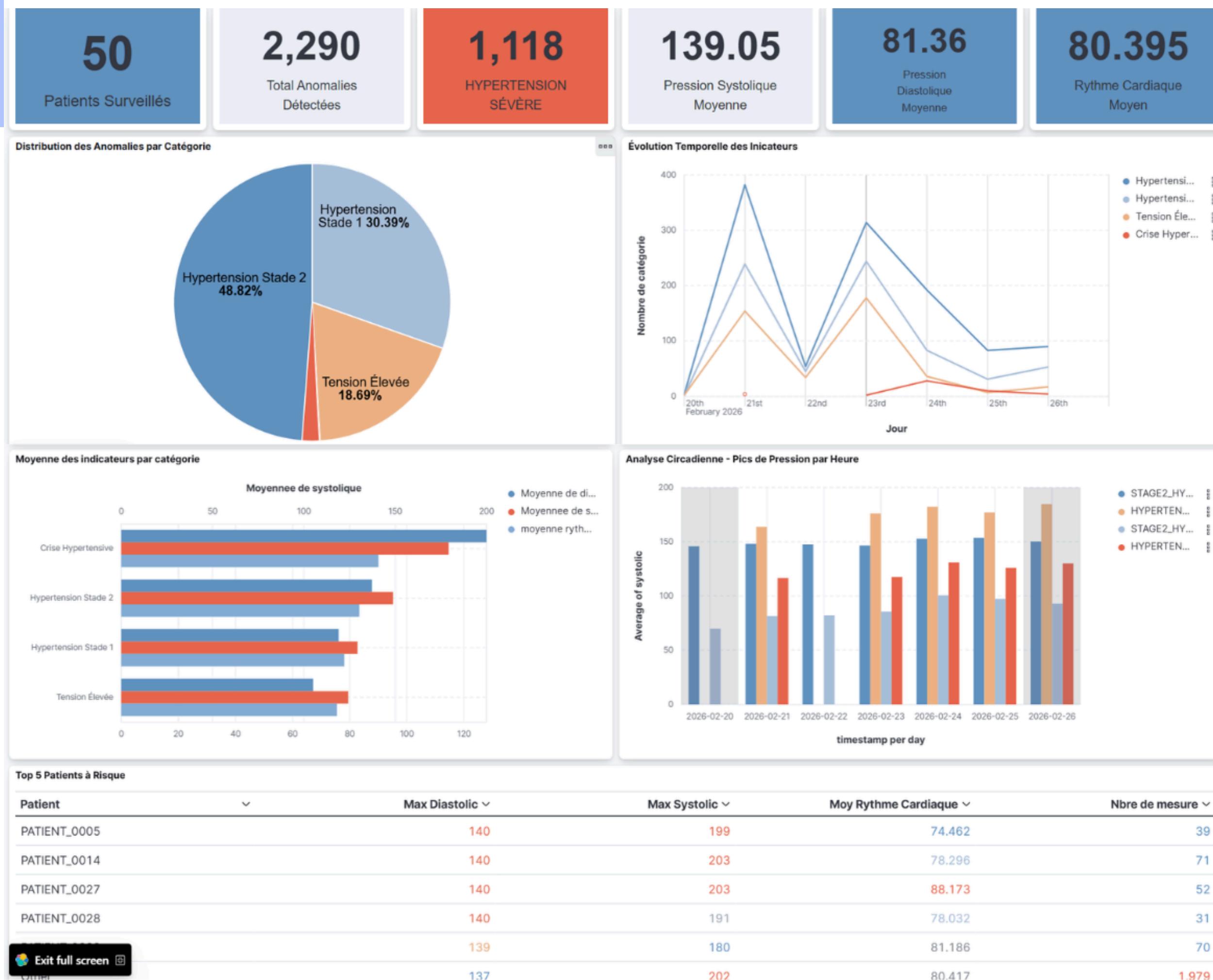
```
{  
  "observation_id": "508878d5-7067-45d6-a595-6b84713e0607",  
  "timestamp": "2026-02-20T22:27:02.320173",  
  "patient_id": "PATIENT_0002",  
  "systolic": 100,  
  "diastolic": 62,  
  "heart_rate": 55,  
  "category": "NORMAL",  
  "device_id": "DEVICE_3"  
}
```

Visualisation



kibana

- Supervision en temps réel
- Détection rapide des patients à risque
- Aide à la décision médicale
- Mise à jour toutes les 10 secondes



INTELLIGENCE ARTIFICIELLE POUR LA PRÉDICTION DES RISQUES

- Objectif:

Prédire la probabilité de crise à 1h, 6h, 24h

Surveillance proactive

- Dataset:

5000 séquences × 30 observations

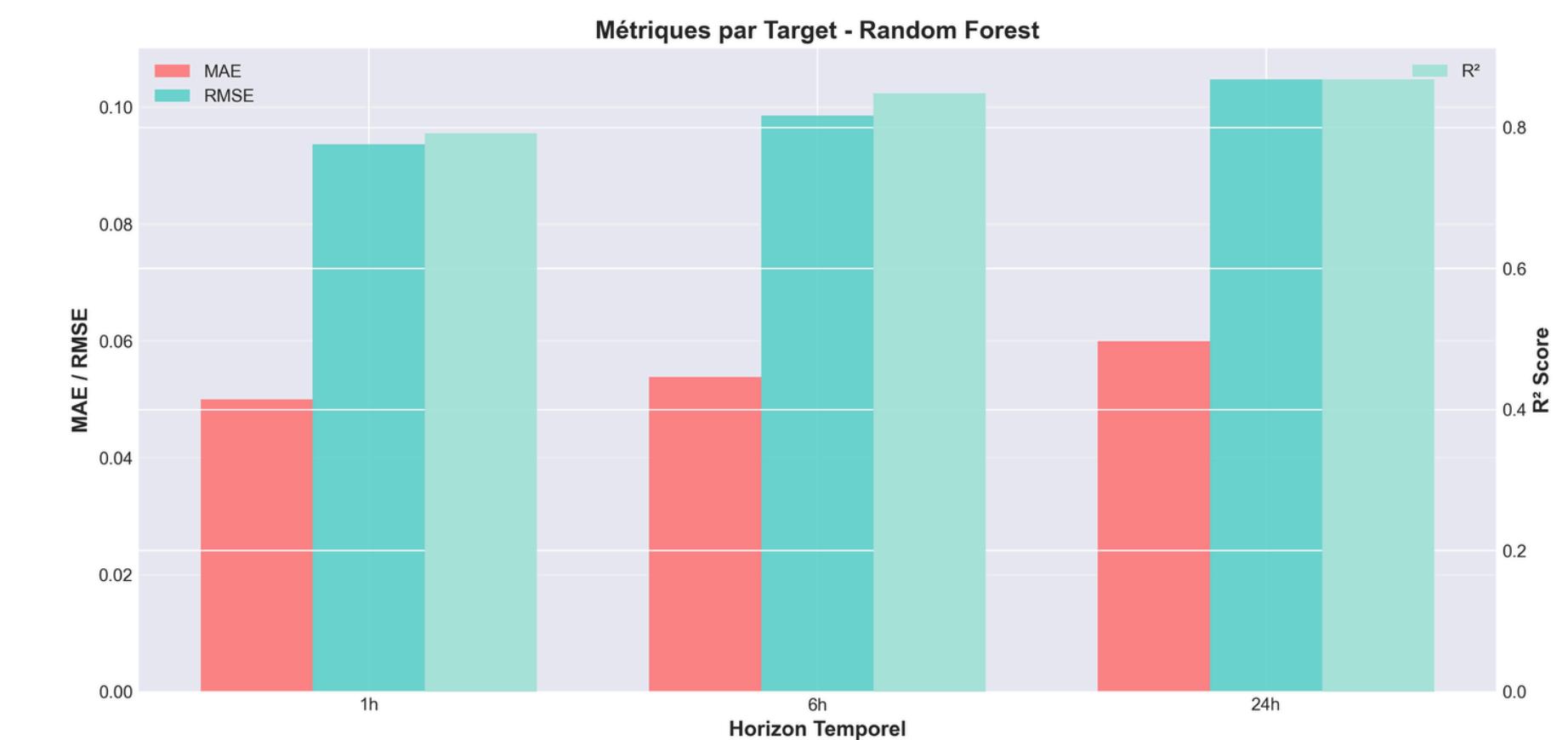
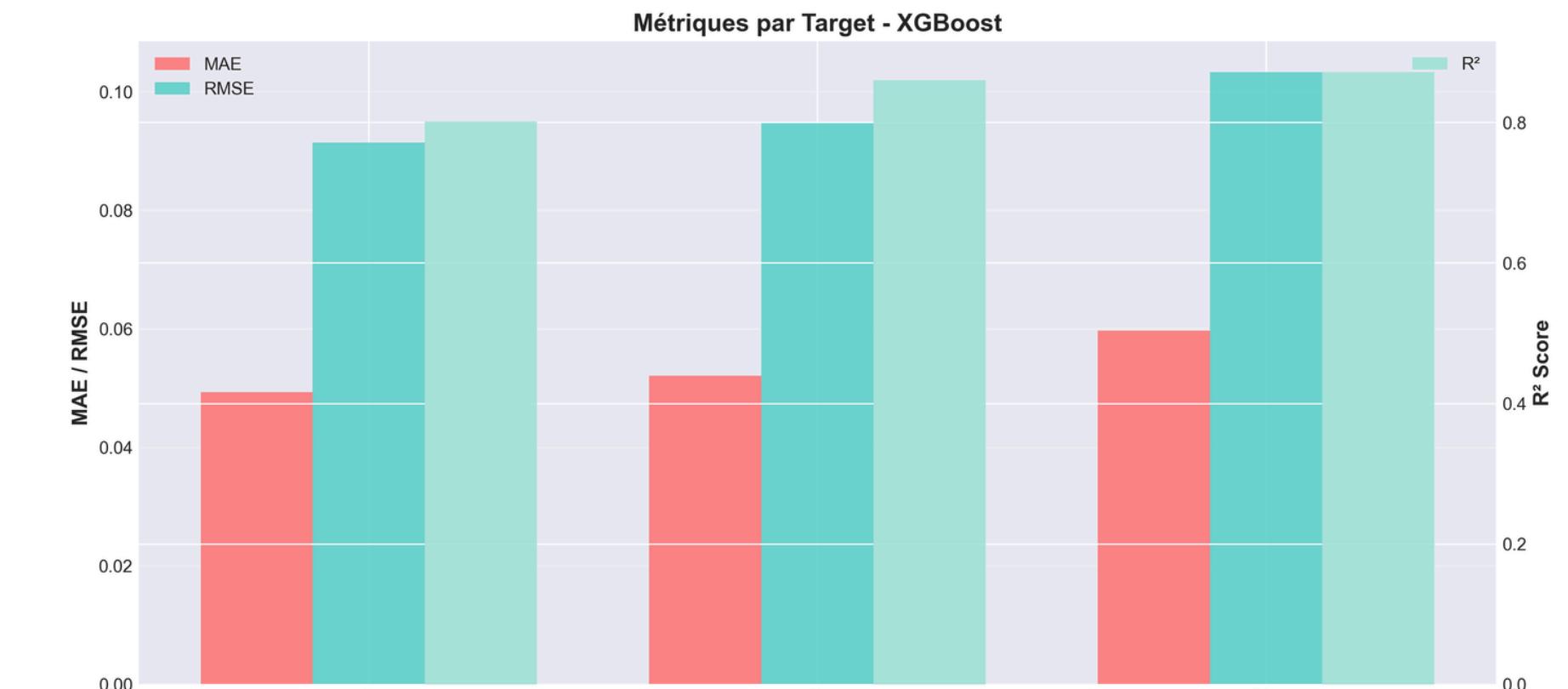
120 features : 30 obs × 4 variables

(systolic, diastolic, heart_rate, hour)

- Modèles comparés:

XGBoost (modèle choisi)

Random Forest



DASHBOARD FLASK – MONITORING TEMPS RÉEL

- Vue Monitoring

Statistiques globales (moyennes, répartition AHA)

Tableau des dernières mesures (Kafka temps réel)

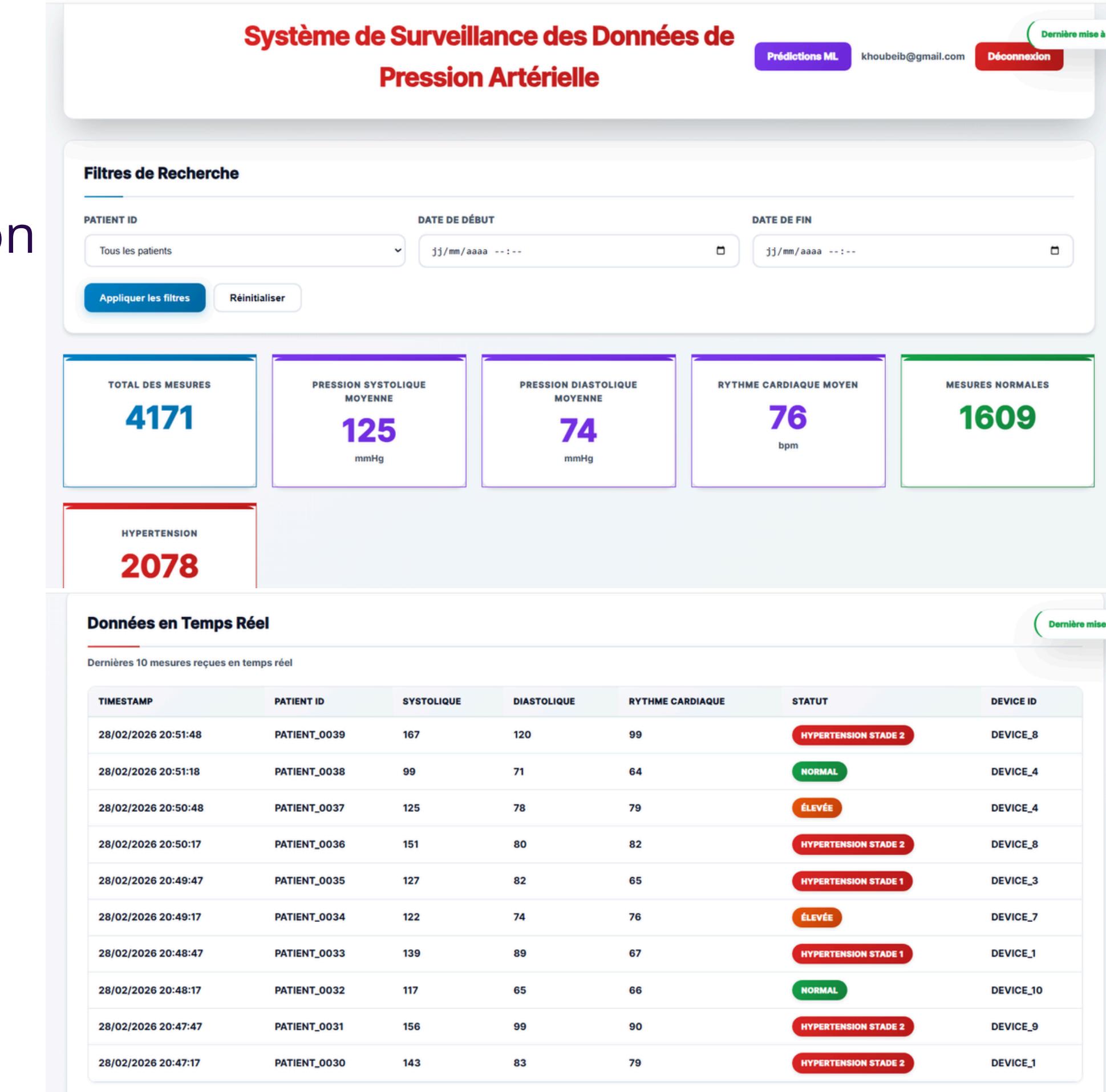
Filtres (patient, dates)

- Fonctionnalités

Authentification

Mise à jour automatique

Code couleur par catégorie médicale



DASHBOARD FLASK – PRÉDICTIONS MACHINE LEARNING

• Vue Prédiction ML

Probabilités 1h, 6h, 24h par patient

Niveaux d'alerte (CRITICAL, HIGH, MEDIUM, LOW)

Section alertes automatiques (CRITICAL + HIGH)

• Fonctionnalités

Guide des prédictions

Filtres par patient

Notifications visuelles pour cas critiques

Prédictions ML - Risque de Crise Hypertensive

[Dashboard Principal](#)

Guide des Prédictions ML

Risque 1h Probabilité (0-100%) qu'une crise hypertensive survienne dans la prochaine 1 heure. Une crise hypertensive est définie par une pression artérielle > 180/120 mmHg.	Risque 6h Probabilité (0-100%) qu'une crise hypertensive survienne dans les prochaines 6 heures. C'est l'indicateur principal utilisé pour déterminer le niveau d'alerte.	Risque 24h Probabilité (0-100%) qu'une crise hypertensive survienne dans les prochaines 24 heures. Permet d'anticiper les risques à moyen terme.	Niveau d'Alerte Niveau de priorité basé sur le Risque 6h: <ul style="list-style-type: none">CRITICAL: ≥ 60% - Intervention immédiateHIGH: ≥ 40% - Surveillance renforcéeMEDIUM: ≥ 20% - Surveillance normaleLOW: < 20% - Surveillance standard
--	---	--	--

Observations Utilisées

Le modèle utilise les 30 dernières observations du patient (environ 7.5 heures d'historique) pour faire ses prédictions. Chaque observation contient: pression systolique, pression diastolique, rythme cardiaque et heure du jour.

TOTAL PATIENTS
50

ALERTE
17

Prédictions par Patient

PATIENT	RISQUE 1H	RISQUE 6H	RISQUE 24H	NIVEAU ALERTE	OBSERVATIONS
PATIENT_0001	49.8%	56.5%	53.2%	HIGH	98
PATIENT_0002	38.3%	36.8%	32.9%	MEDIUM	98
PATIENT_0003	48.5%	47.3%	48.5%	HIGH	96
PATIENT_0004	45.0%	51.3%	51.5%	HIGH	94
PATIENT_0005	43.5%	45.7%	44.8%	HIGH	93
PATIENT_0006	41.4%	35.5%	32.3%	MEDIUM	91
PATIENT_0007	36.8%	37.0%	30.8%	MEDIUM	90
PATIENT_0008	37.9%	30.7%	36.5%	MEDIUM	90

Conclusions

Architecture Big Data complète

- Pipeline de streaming temps réel (Kafka)
- Stockage multi-niveaux optimisé
- Visualisation adaptée (Kibana analytique + Flask décisionnel)

Intelligence Artificielle intégrée

- Modèle XGBoost performant
- Prédictions proactives (1h, 6h, 24h)
- Alertes automatiques pour priorisation

Résultats

- Système fonctionnel et déployable
- Surveillance proactive des patients
- Aide à la décision médicale

Perspectives

- Prototype évolutif et scalable
- Convergence Big Data + IA pour la santé
- Base pour solutions cliniques modernes

