



## OBJECTIF DU PROJET

J'envisage de mettre en place un projet englobant un certain nombre de technologies et matières abordées durant le cursus de formation **Data Analyste AURA 2021 avec M2I**.

Ce projet consiste à mettre en place un environnement de supervision journalisé, en déduire une analyse prédictive et ainsi définir les axes d'amélioration à apporter en termes de sécurité et d'infrastructure.

Vous trouverez dans le document ci-dessous une description de l'existant ainsi que la solution cible à mettre en place.

**La date de remise du projet est : à définir en fonction de son avancée et de l'entrée où non en Stage Entreprise**

## PRESENTATION DU CONTEXTE

Dans un contexte économique et industriel, le volume de données récoltées par bon nombre d'applications ; deviens de plus en plus croissant. La migration vers le cloud devient pour les entreprises comme pour le grand public un facteur indispensable.

A ceci la confidentialité des données qui pour les entreprises reste un enjeu financier et pour chaque individu peut représenter une atteinte à la vie privée, est un second facteur à ne pas négliger.

Dans cet optique, j'ai mis en place un certain nombre d'indicateurs et d'outils de supervision afin de préserver la sécurité et la protection de mes données quant à l'utilisation de mon matériel informatique (tant pour une utilisation personnelle que professionnelle).

Les fichiers journaux (système / Réseau / logs) constitue la base des données sur lesquelles je souhaite effectuer un audit, un traitement, une analyse et des prédictions sur les axes d'améliorer en termes de confidentialité.

## DESCRIPTION DE LA MISSION

Pour ce projet il faut à présent intégrer les données récoltées par le biais des journaux des différents outils de supervision, les intégrer dans un système prédictif Azure Machine Learning, générer des Dashboard avec Power BI et en déduire les améliorations à apporter à l'infra.

Une fois, les rapports, établis. Il faudra mettre en place le correctif et les automatiser à l'aide de Ansible (par exemple).

Cette démarche sera suivie d'une démonstration Teams expliquant les différentes technologies utilisées ainsi que le matériel et les logiciels, respectivement mis en place et configurés pour sa réalisation.

Les différents auditeurs qui participeront à cette démonstration, jugerons de la viabilité de mon projet.

# Cahier des charges de la mise en place d'un environnement de supervision avec analyses prédictives

---

J'ai déjà mis en place pour ce projet :

Un serveur CentOS 7 avec Docker, un environnement distribué Hadoop :

- Kafka manager
- Zookeeper
- Kafka & réplica
- Elasticsearch
- Kibana

Un Cloud provider Azure (à développer) :

- Création du compte et configuration de la sécurité
- Déploiement environnement via VScode source Python
- Machine learning (avec instance & cluster)

Installation et Configuration d'outils de supervision et de sécurité :

- Firewall
- VPN
- Outil Nagios
- Outil Webmin
- Outils d'audits

Pour la gestion de Projet et l'intégration continue :

- Suivi de l'évolution des tickets et tâches dans Jira
- Intégration continue avec la suite Git
- Mise à jour régulières par SSH et avec Ansible