Développement d'un système de DLP reposant sur l'IA

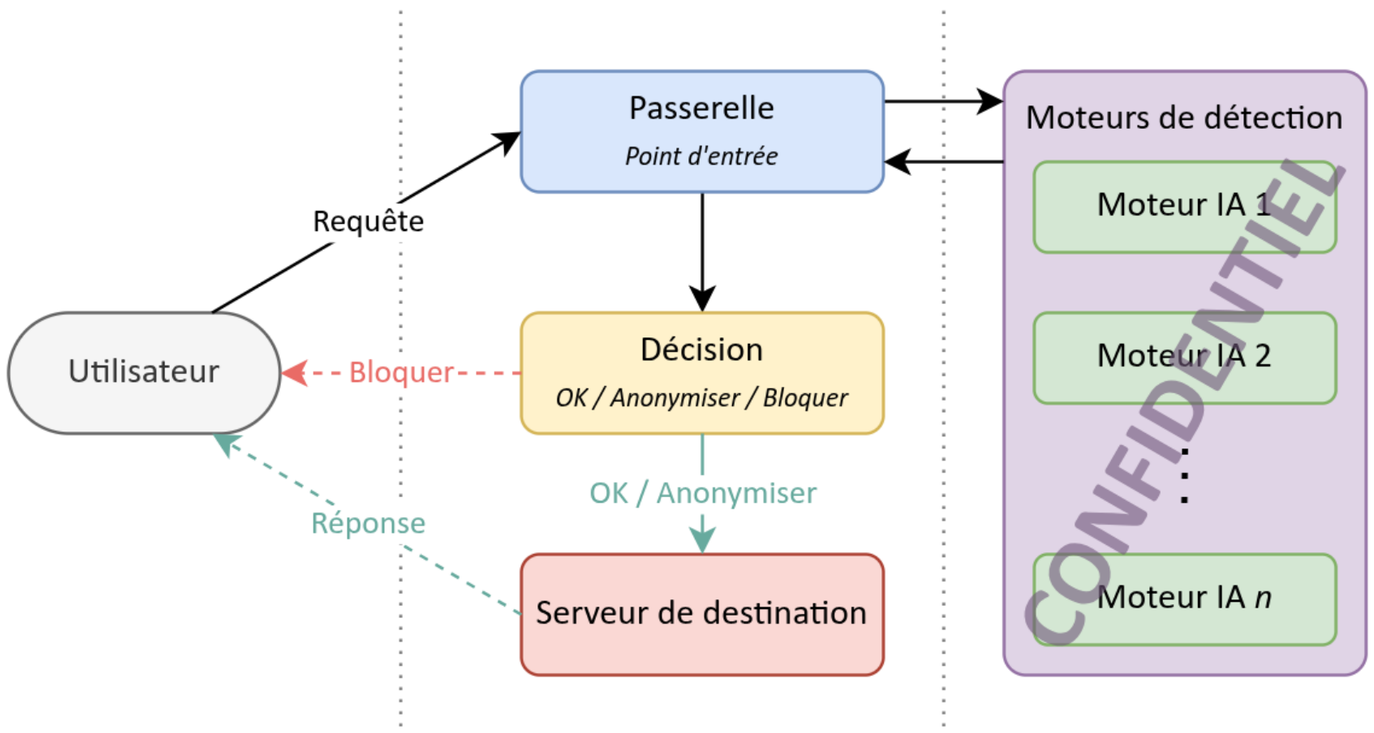
Auteurs : Chams ABDELWAHED - Lorentzo DONNÉ - Théo HENG

CONTEXTE ET OBJECTIF

INETI est une startup spécialisée en cybersécurité et intelligence artificielle, accompagnant les entreprises dans la protection de leurs données et leur conformité réglementaire (RGPD, NIS2, ISO 27001, ...).

Face à l'augmentation des risques d'exfiltration de données et aux limites des solutions de détection de fuites de données (DLP ou *DataLoss Protection*) traditionnelles, ce projet vise à concevoir une plateforme DLP souveraine et IA-native capable de détecter et bloquer en temps réel les comportements suspects et les fuites de données sensibles. L'objectif à la fin du projet est de fournir une preuve de concept (PoC) de notre système afin de détecter, masquer ou bloquer les données sensibles,

MÉTHODES ET DÉVELOPPEMENTS



Nous avons commencé par réaliser une veille concurrentiel puis un état de l'art sur les technologies existantes pour détecter des données sensibles à l'aide de l'IA.

La plateforme que nous avons proposé repose sur une architecture microservices orchestrée par une passerelle centrale, qui soumet chaque requête à l'ensemble des moteurs de détection IA qui sont exécutés en parallèle.

Les résultats sont fusionnés puis soumis à un moteur de décision qui applique les politiques de l'entreprise : laisser passer, alerter l'utilisateur, anonymiser ou bloquer.

L'ensemble est piloté via une interface d'administration permettant de configurer les politiques, les utilisateurs et de suivre les usages en temps réel. Chaque moteur est indépendant et activable à la demande, ce qui permet d'adapter le niveau de détection au contexte métier sans modifier l'architecture globale.

L'intégralité du traitement s'effectue en local, garantissant la souveraineté des données et la conformité aux exigences réglementaires (RGPD, NIS2).

RÉSULTATS ET CONCLUSION

Les moteurs de détection atteignent tous un score de performance respectant les spécifications initiales, démontrant une détection fiable des données sensibles sur des typologies variées (données personnelles et industriels). La combinaison de moteurs complémentaires s'avère plus efficace qu'une approche mono-modèle, notamment sur les cas limites où les règles statiques seules échouent. Ce projet pose les bases d'une solution souveraine et IA-native, adaptable aux exigences réglementaires des entreprises européennes.

MOTS-CLÉS : Cybersécurité · DLP · Protection des données · Multi-modèles d'IA · RGPD · Microservices