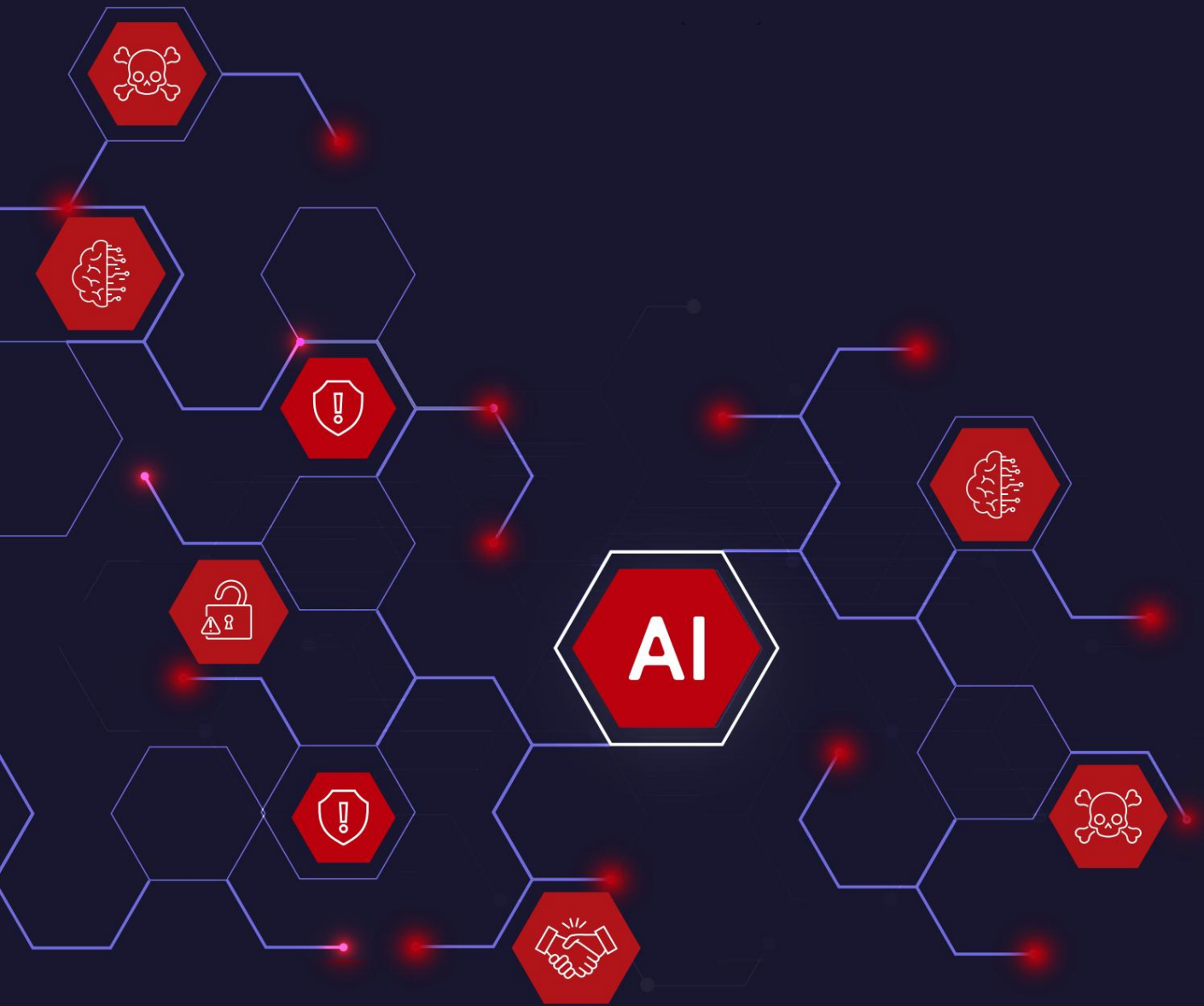
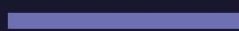




Garde profonde  
Sécurité native de l'IA



## SÉCURITÉ ET IA NATIVE FIABILITÉ



SÉCURISER GENAI AVEC GENAI

# AUTONOMISER L'IA/LLM AVEC SÉCURITÉ ET CONFIANCE

Les systèmes d'IA doivent être fiables, explicables, privés et sécurisés pour l'entreprise et ses clients, mais ces caractéristiques peuvent être compromises par des erreurs ou des activités malveillantes.

Parallèlement, la loi européenne sur l'IA, la charte américaine des droits de l'IA et d'autres cadres réglementaires des gouvernements locaux des États américains, du Japon, de la Corée et de l'Inde commencent déjà à définir des contrôles de conformité, exigeant l'évaluation, la surveillance et le contrôle de l'IA.

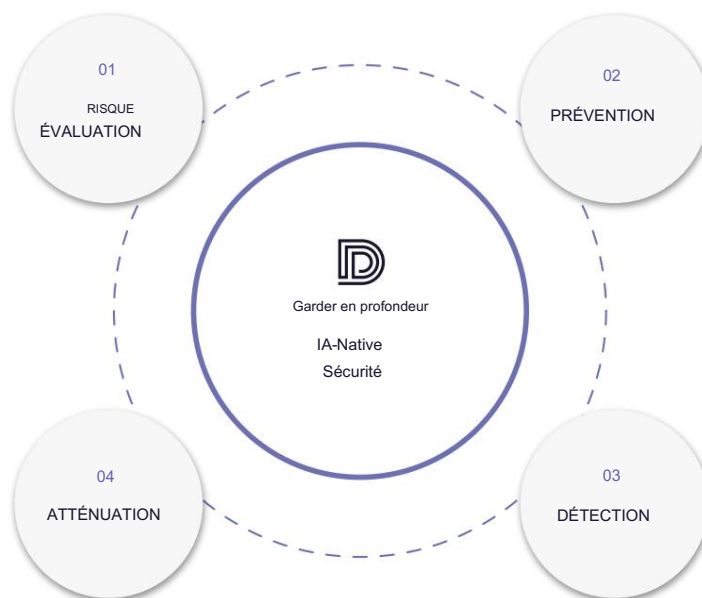


## PRÉ ET POST-DÉPLOIEMENT

- Tests de pénétration
- Vulnérabilité évaluation
- Empoisonnement et détection de porte dérobée
- Protection recommandations

## PRÉ-DÉPLOIEMENT DURCISSEMENT

- Prétraitement
- Post-traitement
- Réparation du modèle



## PARE-FEU IA

- Restriction d'accès
- Alerte en temps réel déclenchement
- Protection dynamique
- Centre d'opération et réponse

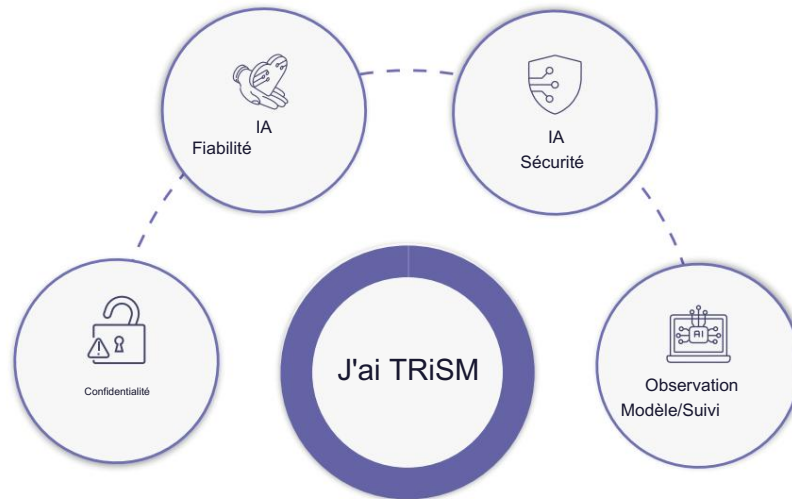
## DÉTECTION

- En ligne/hors ligne
- Avec état/sans état
- Détection d'anomalies
- Basé sur le contexte



# L'IA NATIVE DE DEEPCKEEP SOLUTION TRiSM

AI TRiSM - Trust, Risk and Security Management - est la principale tendance technologique stratégique de Gartner pour 2024. TRiSM est un ensemble de solutions qui identifient et atténuent de manière proactive les risques découlant des modèles et applications d'IA, ainsi que les risques liés à la fiabilité, à la fiabilité, à l'équité et à la sécurité.



La fiabilité protège contre les erreurs, les considérations éthiques et l'équité dans la prise de décision.



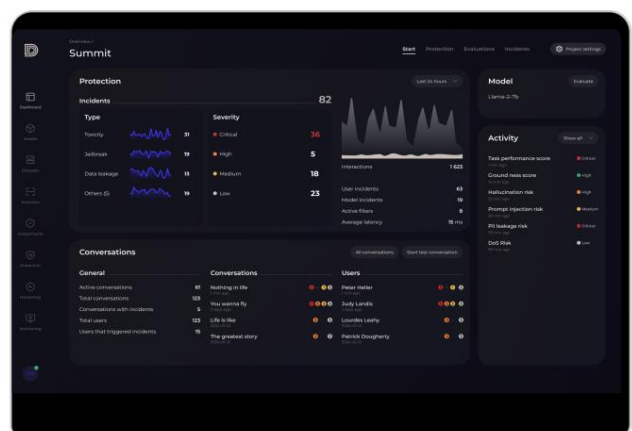
Le risque consiste à identifier les vulnérabilités potentielles et les menaces pesant sur la sécurité, la fiabilité et la confidentialité d'un système d'IA.



Gestion de la sécurité protège les modèles et les ensembles de données contre les attaques, les accès non autorisés et les manipulations.

DeepKeep garantit la santé et la robustesse des modèles ML pour protéger l'IA des erreurs et des menaces.

DeepKeep propose une solution complète de sécurité et de confiance pour l'ensemble du cycle de vie du modèle ML, couvrant toutes les étapes depuis la conservation des données, la formation du modèle, les évaluations des risques, la prévention, la détection, la surveillance et le déploiement jusqu'à l'atténuation.

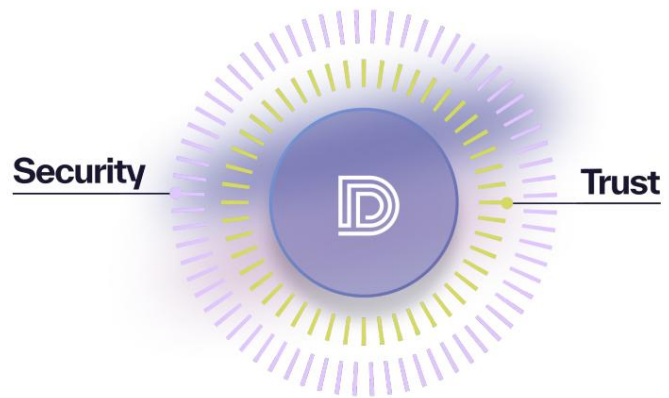


# PILERS INDISPENSABLES – SÉCURITÉ ET FIABILITÉ

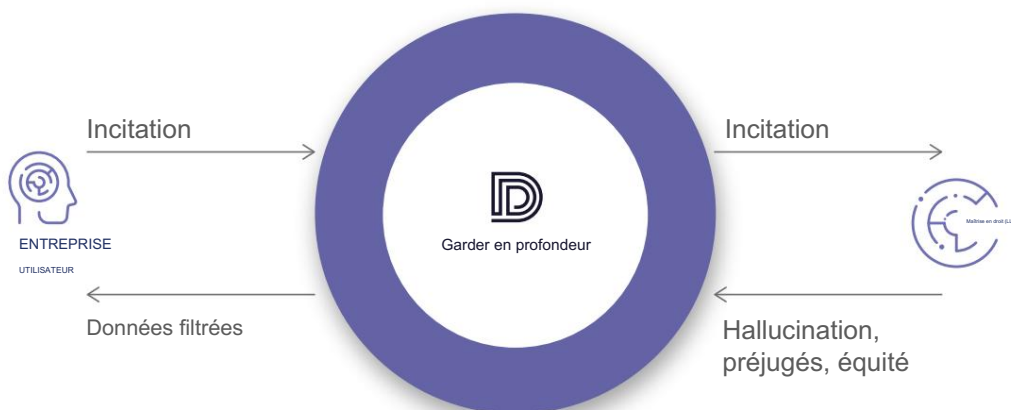
La fiabilité et la robustesse de l'IA sont complémentaires et indissociables : un modèle d'IA ne peut pas être fiable s'il n'est pas robuste et vice versa.

Lorsque les données utilisées pour former un modèle d'IA sont biaisées ou incomplètes, ce modèle peut apprendre à reproduire et à amplifier les erreurs, ce qui entraîne des résultats dangereux.

Lorsque les algorithmes utilisés pour développer un modèle d'IA sont défectueux ou incomplets, le modèle peut produire des résultats inexacts.



## LE GARDE PROFOND SOLUTION POUR LLM



- 1 | Protège contre les attaques LLM, y compris l'injection rapide, la manipulation adverse et les attaques sémantiques.
- 2 | Identifie et alerte contre les hallucinations à l'aide d'un système hiérarchique de sources de données, incluant des références internes et externes fiables.
- 3 | Mesures de protection contre les fuites de données, protégeant les données sensibles et les informations personnelles identifiables (PII).
- 4 | Détecte et supprime le langage toxique, offensant, nuisible, injuste, contraire à l'éthique ou discriminatoire.

## LA SOLUTION DE DEEPKEEP POUR VISION PAR ORDINATEUR

Les images encapsulent une multitude d'indices visuels, englobant des textures, des couleurs, des formes et, surtout, des éléments contextuels qui servent d'indicateurs cruciaux pour les modèles de détection d'objets, influençant leur capacité à détecter et à classer avec précision les objets.



Évalue l'intégrité des ensembles de données utilisés pour la formation des modèles ainsi que les performances, la fiabilité et la robustesse des modèles tout au long du parcours de l'IA, de la conservation des données et de la création du modèle jusqu'au déploiement dans les environnements de production.



Détecte et atténue les incidents malveillants et les problèmes de fiabilité en temps réel, notamment l'équité, les préjugés, les points faibles, la dérive des données, la non-distribution (OOD), l'empoisonnement, l'évasion et le déni de service.



Protège et surveille les modèles de classification d'images et de détection d'objets contre les attaques physiques et numériques, permettant des déploiements sécurisés et fiables.

Libérez la puissance de la confiance et de la sécurité dans l'IA avec DeepKeep.



# À PROPOS DE DEEPCKEEP

DeepKeep protège les pipelines d'apprentissage automatique contre les biais, les erreurs et les risques de cybersécurité, garantissant ainsi une IA fiable et digne de confiance.

La sécurité de l'IA de DeepKeep protège les pipelines d'apprentissage automatique, favorisant des solutions d'IA sécurisées, impartiales, sans erreur, explicables et fiables. Cela comprend les modèles de données de vision, les modèles LLM et les modèles tabulaires dans les évaluations des risques, la prévention, la détection, la surveillance et l'atténuation.

Seule la sécurité native de l'IA - construite elle-même avec l'IA générative - peut protéger des frontières illimitées et une génération de contenu sans fin dans divers domaines sources, modèles et ensembles de données.

La plate-forme logicielle d'entreprise de DeepKeep offre sécurité et fiabilité depuis la phase de recherche et développement des modèles d'apprentissage automatique jusqu'au déploiement et tout au long du cycle de vie du produit.



CONTACT

mettre ici  
ventes@deepkeep.ai  
www.deepkeep.ai

texte?



Garde profonde  
Sécurité native de l'IA