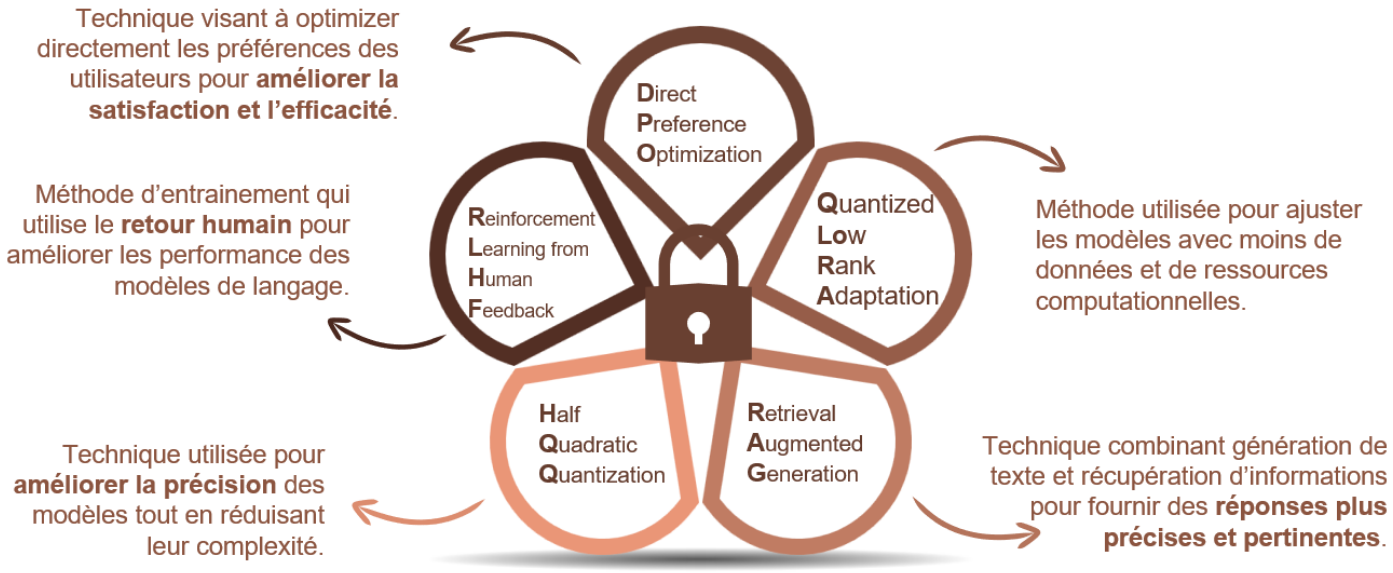
Sécurité

Imperfections

Les IA génératives suscitent un paradoxe : nous attendons d'elles qu'elles inventent et créent, tout en étant précises et fiables. Or, les IA peuvent créer des réponses incorrectes ou incohérentes en cherchant à créer (hallucinations). Le grounding, qui consiste à ancrer les réponses dans des données réelles et contextuelles, permet de de réduire les risques d'erreurs.
Le risque de sécurité lié aux IA génératives est le risque le moins souvent cité par les coopératives lors de notre enquête, l'IA générative étant actuellement moins présente que les autres types d'IA dans les coopératives de nutrition animale.
Nous pouvons citer quelques types d'imperfections :
→ L'injection indirecte de prompt peut être une source d'erreur pour l'assistant virtuel. Il répondra à des instructions
cachées dans des documents analysés.
→ Le piratage dans un contexte de récompense, les modèles de langage peuvent exploiter des failles dans les
spécifications pour obtenir des scores élevés sans répondre à la véritable intention de l'utilisateur.
→ Glitch tokens, une requête avec un mot mal orthographié conduit l'assistant virtuel à répondre de manière
incohérente en lien avec une erreur non détectée dans l'algorithme.

Pour aller plus loin, OWASP a publié le top 10 des risques pour les LLM et application d'IA générative. Des solutions ont été développées pour remédier à ces imperfections (Cf schéma ci-dessous).



Malveillance

L'hyper connectivité, leur niveau d'automatisation en environnement OT (technologique-opérationnel) expose les coopératives de nutrition animale à la cybercriminalité, au cyberterrorisme, à la biosécurité, à la désinformation.
Lors de l'enquête menée auprès des coopératives de nutrition animale en mai 2025, cyberattaque, fuite des données, perte de confidentialité, blocage du fonctionnement, perte de maîtrise de l'outil, malveillance, sont les risques liés à l'IA très majoritairement cités. La sécurité est bien perçue comme un facteur à prendre en compte lors de la mise en place de solutions IA.
Si l'utilisation de l'IA peut augmenter l'exposition aux algorithmes malveillants, l'IA grâce à sa capacité d'entraînement et de traitement de données massives permet en contrepartie d'apporter des solutions à la cybersécurité industrielle, analyse des risques, détection proactive, adaptation aux modèles d'attaque. De multiples solutions ont été développées par un panel d'acteurs pour trouver des solutions de défense adaptées aux besoins.
L'**Agence Nationale de Sécurité des Systèmes d'Information (ANSSI)** a publié en mai 2025 un **article** sur les nouveaux risques inhérents à l'introduction de nouvelles technologies dans le domaine des systèmes industriels à travers l'exemple de l'Industrial Internet of Things (IIoT). Il propose un modèle d'architecture de passerelle d'interconnexion sécurisée entre les mondes industriel et de gestion sans nécessiter la refonte de l'architecture du système industriel en profondeur. L'article vise un **public de DSI**.

L'European Union Agency for cybersecurity (ENISA) propose également des **référentiels de management du risque**.

> « Giskard s'est donné comme mission de combler le fossé entre la technologie complexe de l'intelligence artificielle (IA) et son application dans des scénarios réels, en garantissant la robustesse et la transparence. Giskard est une solution en open source, dédiée à l'amélioration de la qualité de l'IA. En minimisant les erreurs de l'IA telles que les biais éthiques et les risques de sécurité, Giskard permet la création de modèles d'IA supérieurs pouvant être déployés en production. » (Alex Combessie, co-fondateur de Giskard).

Souveraineté

A travers l'enquête menée en mai 2025, les coopératives de nutrition animale expriment leur besoin de souveraineté, au sens contrôle de leurs données, sécurisation de leurs infrastructures, maîtrise de leurs algorithmes.
Le besoin de maîtrise au niveau du territoire national ou au sein de d'un espace de confiance comme l'union Européenne est perceptible. Les fournisseurs de solutions, d'algorithme, d'espaces de stockage, les API, les sources et partages de données des modèles sont à prendre en compte pour répondre à ce besoin de maîtrise.

> Mistral AI, est une startup française pionnière en intelligence artificielle fondée en avril 2023. Elle s’est fixée comme mission : démocratiser l’intelligence artificielle grâce à des modèles, produits et solutions IA open-source, efficaces et innovants.

Les fuites de données sont parfois liées à l’utilisation de l’IA générative par les personnes avec leurs outils de travail.
Elles peuvent être tentées d’utiliser l’IA générative en libre accès, s’ils leur semblent que la mise en place est trop lente dans leur structure. Cette Shadow IA peut présenter un danger si la coopérative n’a pas mis en place les sécurités nécessaires. Une charte d’utilisation de l’IA peut s’avérer utile.

> **Points clés**
> **→ Trouver le bon équilibre entre besoin de créativité et nécessité d’avoir des données sûres et fiables,**
> **→ Intégrer dans les algorithmes des solutions pour lutter contre les imperfections,**
> **→ Analyser les risques pour la cybersécurité, intégrer l’IA pour la cyberdéfense,**
> **→ Faire des choix technologiques éclairés en lien avec la volonté de souveraineté de la coopérative.**

Révision #1
Créé 1 juin 2026 15:27:59 par Céline Ravel
Mis à jour 1 juin 2026 15:39:31 par Céline Ravel