

# Intelligence Artificielle et Cybersécurité

Céline Blandin – Mars 2025

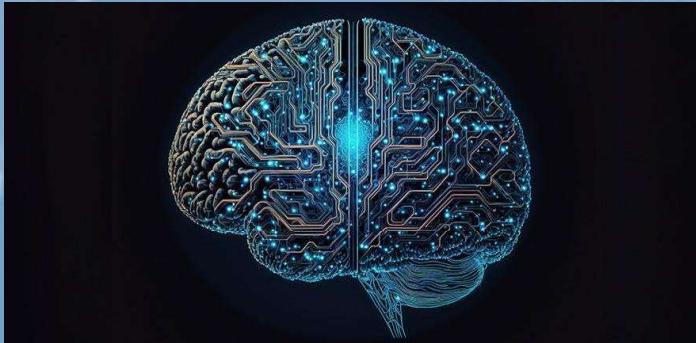
# Intelligence Artificielle et Cybersécurité

- 1. Comprendre l'Intelligence Artificielle**
  - Définition et fonctionnement
  - Exemples d'IA dans la vie courante
- 2. L'IA en Cybersécurité: Atout et Menace**
- 3. Bonnes pratiques, comment utiliser une IA?**

## 1. Comprendre l'Intelligence Artificielle

## ➤ Qu'est ce que l'IA

Un ensemble de technologies qui permettent aux machines de **simuler** une forme d'intelligence humaine.



→ Capacité à analyser des données, apprendre des schémas, et générer des réponses adaptées.



ensemble d'étapes mathématiques effectuées dans un ordinateur.

## 1. Comprendre l'Intelligence Artificielle

### ➤ Les différents types d'IA



### Aujourd'hui

#### 1) Les Machines réactives

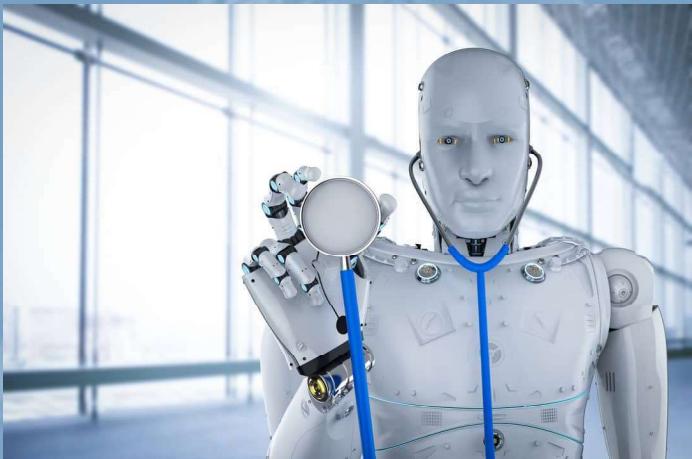
- Simple, agit en temps réel à des situations précises
- Pas de mémoire, pas d'apprentissage permanent
- Ex: Deep Blue (IA d'échec d'IBM)

#### 2) IA à mémoire limitée

- Peut stocker temporairement des données
- Modèles d'apprentissage automatique (Machine Learning)
- Ex: voitures autonomes qui analysent en temps réel les routes
- Ex: Chatbot IA: mémorisent le contexte pour répondre plus intelligemment dans une conversation (chatGPT)
- Ex: Systèmes de détection des fraudes bancaires

## 1. Comprendre l'Intelligence Artificielle

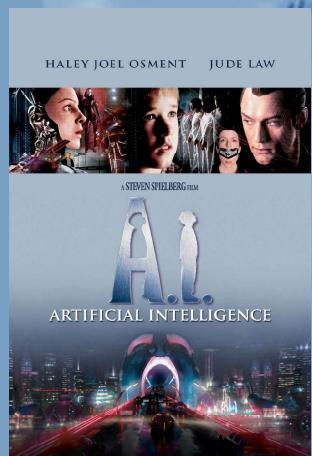
### ➤ Les différents types d'IA



#### Hypothétique

##### 3) L'IA a la théorie de l'esprit

- Comprendre les émotions humaines
- Raisonner sur les intentions des autres
- Ex théorique: Robots sociaux, IA thérapeutiques



##### 4) IA consciente

- Conscience d'elle-même
- Pourrait penser de manière autonome, avoir des désirs et prendre des décisions indépendamment de l'homme.

##### → HYPOTHESE FUTURISTE

## 1. Comprendre l'Intelligence Artificielle

### ➤ Comment fonctionne une IA?

#### Les 5 étapes du fonctionnement d'une IA

- 1) Collecte des données
- 2) Préparation et nettoyage des données
- 3) Entraînement du modèle d'IA
- 4) Test et validation du modèle
- 5) Déploiement et amélioration continue



- 3) A/ Apprentissage supervisé  
→ Apprend a partir d'exemples fournis : DATA
- 3) B/ Apprentissage non supervisé  
→ L'IA détecte elle-même des schémas dans les données.  
(tendances: statistiques)
- 3) C/ Apprentissage par renforcement  
→ L'IA apprend par essais et erreurs avec un système de « récompense/punition »  
→ (Reinforcement Learning from Human Feedback (RLHF))

## 1. Comprendre l'Intelligence Artificielle

- Comment fonctionne une IA?

### Notion de Machine Learning

Besoin de règles humaines 	Apprend tout seul 
Fonctionne bien avec peu de données	A besoin de beaucoup de données
Moins puissant, mais plus rapide 	Plus puissant, mais demande plus d'entraînement 
Exemple : une IA qui apprend grâce à des règles données par un humain.	Exemple : une IA qui apprend seule en regardant des milliers d'exemples. (comme cerveau humain avec neurones artificiels)



### Notion de Deep Learning



# Intelligence Artificielle et Cybersécurité

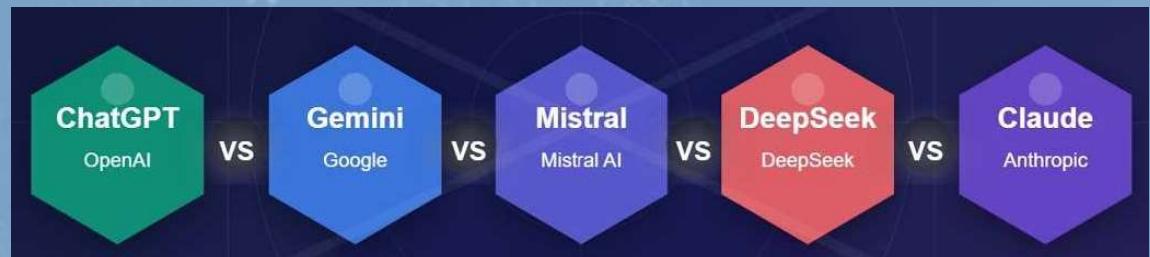
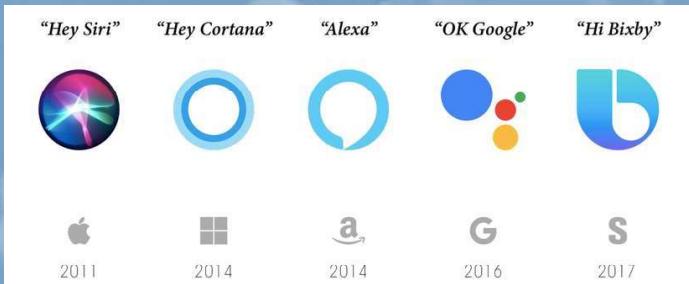
- 1. Comprendre l'Intelligence Artificielle**
  - Définition et fonctionnement
  - Exemples d'IA dans la vie courante
- 2. L'IA en Cybersécurité: Atout et Menace**
- 3. Bonnes pratiques, comment utiliser une IA?**

## 1. Comprendre l'Intelligence Artificielle

### ➤ Exemple d'IA dans la vie courante

- Chatbots et Assistant vocaux

Exemples: Google Assistant, Siri, Alexa, Meta AI, ChatGPT, DeepSeek, Gemini, Mistral...



Utilisation: Ces IA sont capables de répondre aux questions, de donner des recommandations, d'effectuer des tâches comme programmer un réveil ou envoyer un message.

Impact: Elles facilitent l'accès à l'information et automatisent certaines actions, mais elles ont encore des limites, comme la compréhension du contexte complexe.

## 1. Comprendre l'Intelligence Artificielle

### ➤ Exemple d'IA dans la vie courante

- Traduction automatique

Exemple: Google Translate, DeepL, Microsoft Translator



Utilisation: Traduire des textes instantanément, faciliter la communication entre personnes ne parlant pas la même langue, sous-titrer des vidéos.

Impact: Ces outils permettent de briser les barrières linguistiques, bien qu'ils ne soient pas encore parfaits pour les subtilités et les expressions culturelles.

## 1. Comprendre l'Intelligence Artificielle

### ➤ Exemple d'IA dans la vie courante

- Médecine et Diagnostic assisté par IA

Exemples: IA de détection des cancers (Google Health, IBM Watson), analyse d'IRM, assistants médicaux virtuels.



Utilisation: Aider les médecins à détecter des maladies plus rapidement et avec plus de précision, optimiser les traitements en fonction des données patients.

Impact: Amélioration des soins de santé, mais nécessite une supervision humaine pour éviter les erreurs.

## 1. Comprendre l'Intelligence Artificielle

### ➤ Exemple d'IA dans la vie courante

- Reconnaissance faciale

Exemples: Déverrouillage de smartphones (Face ID d'Apple), surveillance et sécurité, tri des photos dans Google Photos et Facebook.



Utilisation: Permet de sécuriser l'accès aux appareils, d'identifier des personnes dans des lieux publics, ou encore d'améliorer l'expérience utilisateur sur les réseaux sociaux.

Impact: Questions éthiques sur la vie privée et le respect des données personnelles.

## 1. Comprendre l'Intelligence Artificielle

### ➤ Exemple d'IA dans la vie courante

- Voitures autonomes

Exemples: Tesla (Autopilot), Waymo (Google), systèmes avancés d'assistance à la conduite (ADAS).



Utilisation: Ces IA analysent l'environnement (panneaux, piétons, autres véhicules) pour aider ou remplacer le conducteur.

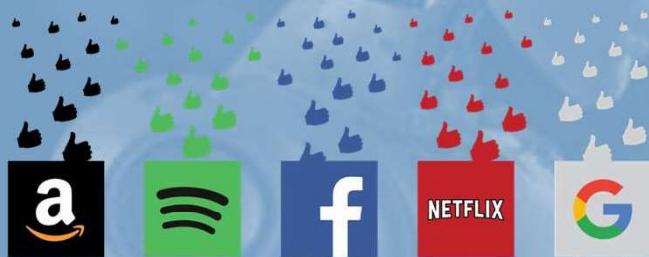
Impact: Elles promettent de réduire les accidents causés par l'erreur humaine, mais posent encore des défis de sécurité et de réglementation.

## 1. Comprendre l'Intelligence Artificielle

### ➤ Exemple d'IA dans la vie courante

- Recommandations personnalisées

Exemples: Algorithmes de Netflix, YouTube, Amazon, Spotify



Utilisation: Proposer du contenu basé sur les préférences et comportements de l'utilisateur (films, vidéos, musique, achats).

Impact: Expérience utilisateur améliorée, mais risque d'enfermement dans une bulle algorithmique (« filter bubble »).

## 1. Comprendre l'Intelligence Artificielle

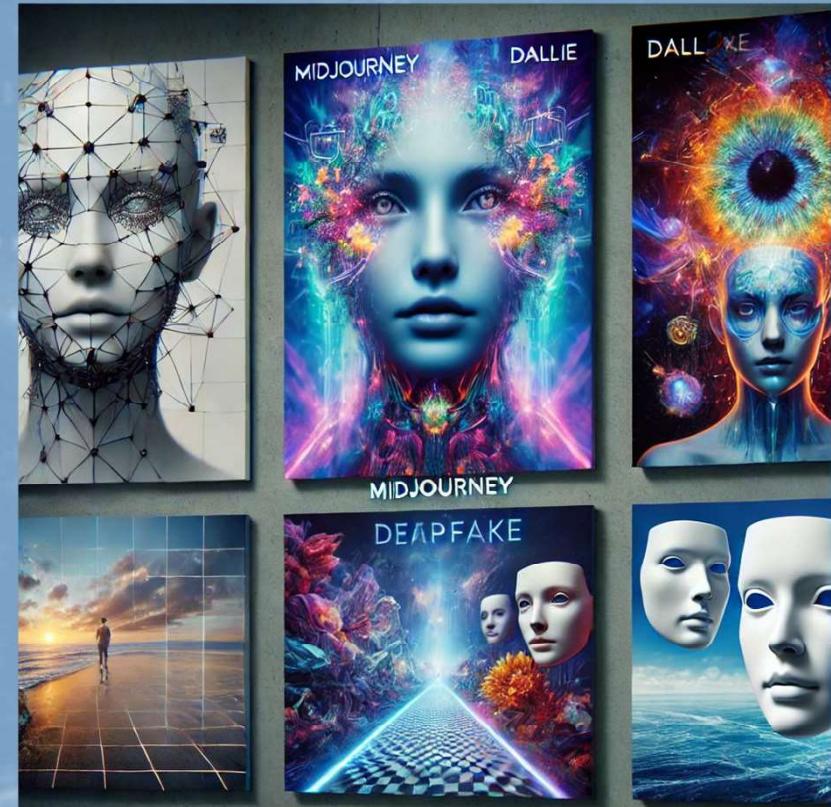
### ➤ Exemple d'IA dans la vie courante

- **Création d'images et de vidéos**

Exemples: MidJourney, DALL·E, RunwayML, deepfakes

Utilisation: Générer des images réalistes à partir d'un texte, modifier des visages dans des vidéos, créer du contenu artistique et publicitaire.

Impact: Certains usages, comme les deepfakes, soulèvent des inquiétudes sur la désinformation et la manipulation de l'image.



# Intelligence Artificielle et Cybersécurité

1. Comprendre l'Intelligence Artificielle
2. L'IA en Cybersécurité: Atout et Menace
  - Les avantages
  - Les risques
3. Bonnes pratiques, comment utiliser une IA?

## 2. L'IA en cybersécurité: Atout et Menace

### ➤ Les avantages

- Détection des menaces

Analyse de grandes quantités de données pour repérer les comportements suspects



Identification rapide de tentatives de phishing, de malware ou d'attaques réseau, analyses prédictives des cyberattaques

## 2. L'IA en cybersécurité: Atout et Menace

### ➤ Les avantages

- Automatisation de la réponse aux attaques

#### Réactions en temps réel:

Firewalls intelligents, Antivirus basés sur l'IA (Microsoft Defender ATP), Filtrage intelligent des spams et des tentatives de phishing



## 2. L'IA en cybersécurité: Atout et Menace

### ➤ Les avantages

- Amélioration de la gestion des accès

#### Authentification biométrique



Détection d'anomalies dans les connexions suspectes



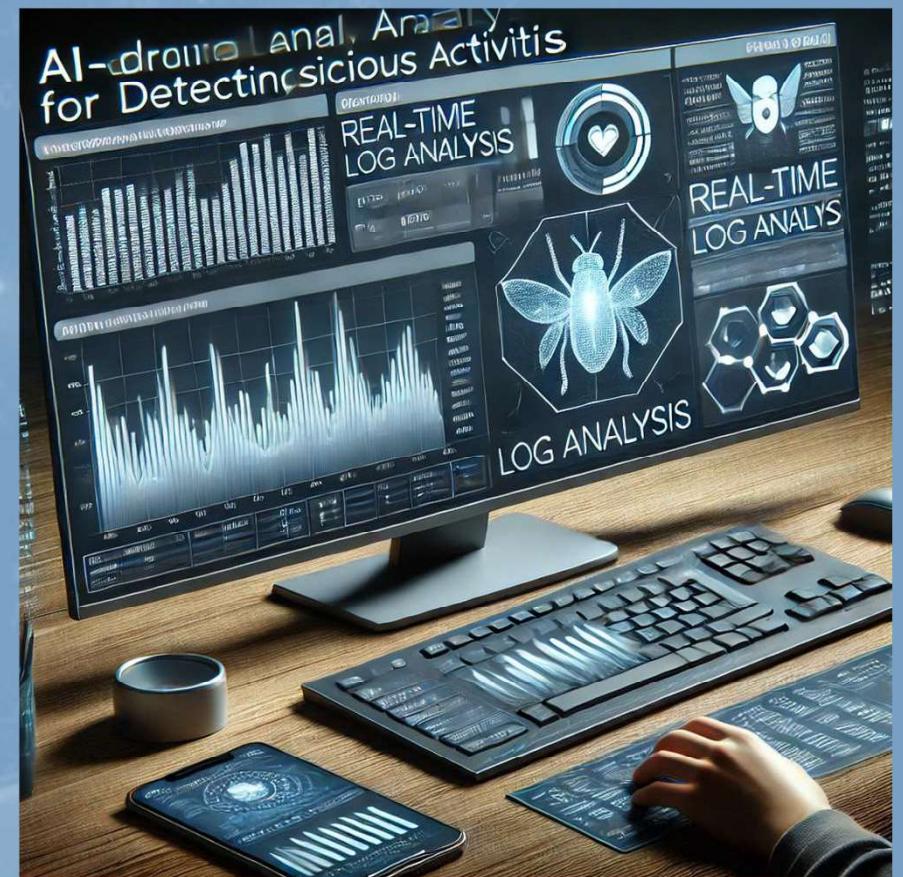
## 2. L'IA en cybersécurité: Atout et Menace

### ➤ Les avantages

- Amélioration de la gestion des logs et des audits

Rapide Tri et analyses de vastes volumes de journaux de connexion

Identifications plus rapide des activités suspectes



# Intelligence Artificielle et Cybersécurité

1. Comprendre l'Intelligence Artificielle
2. L'IA en Cybersécurité: Atout et Menace
  - Les avantages
  - Les risques
3. Bonnes pratiques, comment utiliser une IA?

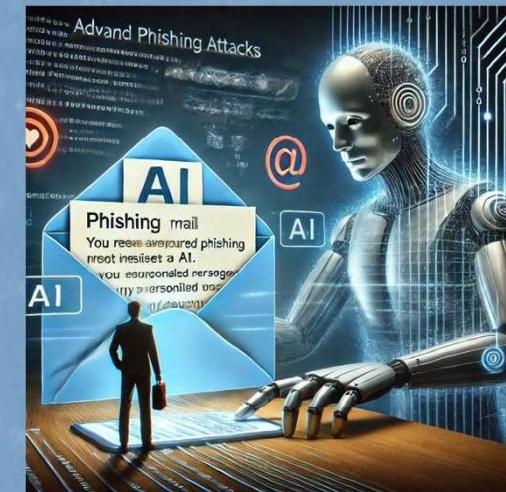
## 2. L'IA en cybersécurité: Atout et Menace

### ➤ Les risques

- Attaques facilitées par l'IA

Phishing amélioré (spear-phishing...)

Génération de logiciels malveillants (malwares qui peuvent évoluer et de s'adapter pour contourner les solutions de sécurité existantes: auto-modification )



Attaques adversariales  
(manipulation d'IA)



## 2. L'IA en cybersécurité: Atout et Menace

### ➤ Les risques

- L'IA et la désinformation: le cas des deepfakes

Un **deepfake** est une technologie d'IA qui permet de créer des vidéos ou des audios manipulés de manière tellement réaliste qu'il devient extrêmement difficile de distinguer ce qui est vrai de ce qui est faux. Ces vidéos peuvent montrer une personne en train de dire ou faire des choses qu'elle n'a jamais faites.

## 2. L'IA en cybersécurité: Atout et Menace

### ➤ Les risques

- L'IA et la désinformation: le cas des deepfakes

<https://www.youtube.com/watch?v=WE88bkPt7Uo>



## 2. L'IA en cybersécurité: Atout et Menace

### ➤ Les risques

- L'IA et la désinformation: le cas des deepfakes

<https://www.youtube.com/watch?v=KuwXJIKOOzo>



## 2. L'IA en cybersécurité: Atout et Menace

### ➤ Les risques

- L'IA et la désinformation: le cas des deepfakes

Un **deepfake** est une technologie d'IA qui permet de créer des vidéos ou des audios manipulés de manière tellement réaliste qu'il devient extrêmement difficile de distinguer ce qui est vrai de ce qui est faux. Ces vidéos peuvent montrer une personne en train de dire ou faire des choses qu'elle n'a jamais faites.

Exemples concrets :

- Faux discours de personnalités politiques
- Usurpation d'identité
- Manipulation d'opinions publiques

Impact en cybersécurité :

- Arnaques aux entreprises
- **Chantage et atteinte à la réputation**
- Difficulté à distinguer le vrai du faux

## 2. L'IA en cybersécurité: Atout et Menace

### ➤ Les risques

- IA entre de mauvaises mains, quand l'IA devient une menace

**PimEyes:** un moteur de recherche par reconnaissance faciale qui permet de retrouver des images d'une personne sur Internet à partir d'une photo.



### Dangers potentiels :

- **Atteinte à la vie privée** → Surveillance, harcèlement, absence de consentement
- **Usurpation d'identité** → Vol de données, exploitation d'images personnelles
- **Utilisation abusive** → Police, accès non réglementé, absence de contrôle
- **Exposition des enfants** → Photos de mineurs, risques d'exploitation
- **Applications dangereuses** → Identification en temps réel, espionnage, atteinte à la sécurité

## 2. L'IA en cybersécurité: Atout et Menace

### ➤ Les risques

- IA entre de mauvaises mains, quand l'IA devient une menace

**The Follower:** IA capable de suivre et d'analyser les activités en ligne des utilisateurs, notamment leurs interactions sur les réseaux sociaux, pour en déduire des informations personnelles et comportementales

#### Dangers potentiels :

- **Usurpation d'identité** → Utilisation frauduleuse des informations, création de faux profils, escroqueries
- **Atteinte réputation** → Diffusion d'informations sensibles, cyberharcèlement, chantage



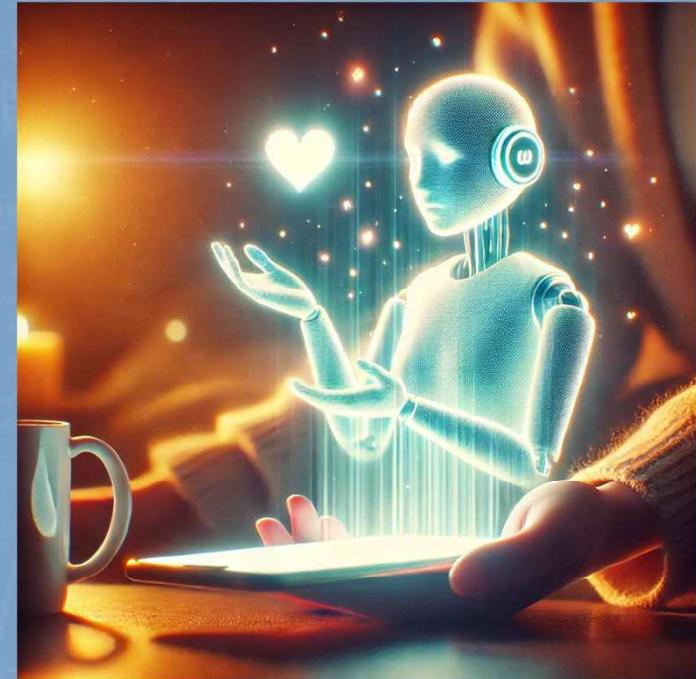
## 2. L'IA en cybersécurité: Atout et Menace

### ➤ Les risques

- IA entre de mauvaises mains, quand l'IA devient une menace

## Replika, Character.AI... :

applications de chatbot basée sur l'intelligence artificielle, conçue pour offrir une compagnie virtuelle aux utilisateurs en simulant des conversations humaines.



### Dangers potentiels :

- **Atteinte à la vie privée** → Collecte de données, stockage d'infos personnelles, risque de fuite
- **Contenu inapproprié** → Conversations sexuelles, interactions avec mineurs, modération insuffisante
- **Dépendance émotionnelle** → Attachement excessif, isolement social, influence psychologique
- **Influence comportement** → Encouragement à des actions nuisibles, manipulation inconsciente
- **Pratiques éthiques douteuses** → Failles de sécurité, exploitation des données, absence de transparence



# "CETTE IA PEUT REMPLACER TON MEC" : LE DANGER DES CHATBOTS

<https://www.youtube.com/watch?v=s8iEKDyUdDw>

IA et Cybersécurité - Céline Blandin - Mars 2025

## 2. L'IA en cybersécurité: Atout et Menace

### ➤ Les risques

- IA entre de mauvaises mains, quand l'IA devient une menace

## Clonage Vocal par IA:

technologie permettant de reproduire fidèlement la voix d'une personne à partir de quelques échantillons audio.

### Dangers potentiels :

- Arnaques téléphoniques
- Usurpation d'identité
- Désinformation
- Atteinte à la réputation
- Implications légales et éthiques

- Escroqueries, usurpation de proches, fraudes financières
- Accès frauduleux, contournement des sécurités vocales
- Faux discours, manipulation politique, deepfakes audio
- Fausse implication, chantage, destruction d'image publique
- Absence de consentement, réglementation insuffisante



# Intelligence Artificielle et Cybersécurité

1. Comprendre l'Intelligence Artificielle
2. L'IA en Cybersécurité: Atout et Menace
3. **Bonnes pratiques, comment utiliser une IA?**
  - Ne jamais partager des informations personnelles
  - Choisir une IA fiable
  - Toujours vérifier, une IA peut se tromper
  - Les limites de l'IA

### 3. Bonne pratique, comment utiliser une IA?

#### ➤ Ne jamais partager des informations personnelles

Les IA peuvent enregistrer, traiter et, dans certains cas, stocker les informations que vous leur fournissez. Même si certaines plateformes garantissent la confidentialité, il est toujours préférable de rester prudent.

Exemples de données sensibles à ne jamais partager :

- **Mots de passe**
- **Numéros de carte bancaire**
- **Adresses personnelles ou professionnelles**
- **Données médicales**
- **Identifiants d'accès (e-mail + mot de passe)**



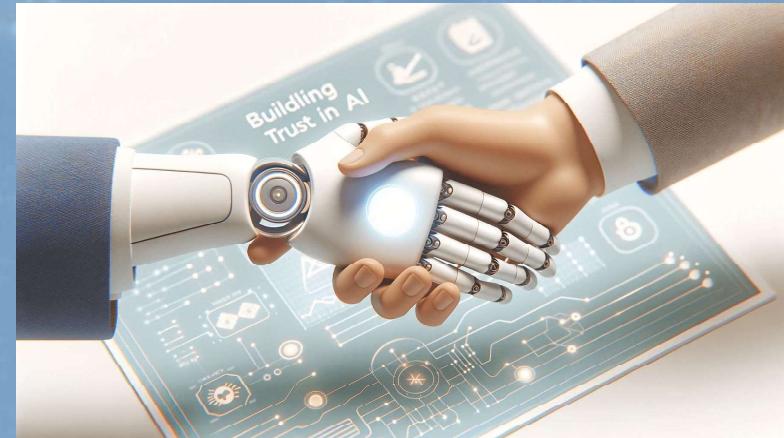
### 3. Bonne pratique, comment utiliser une IA?

#### ➤ Vérifier la fiabilité de l'IA utilisée

Toutes les IA ne sont pas conçues avec les mêmes standards de sécurité et d'éthique. Certaines plateformes peuvent collecter vos données sans transparence.

Comment reconnaître une IA fiable ?

- Si vous ne savez pas , privilégier les **IA développées par des entreprises reconnues** (OpenAI, Mistral)
- Vérifier les **politiques de confidentialité** avant d'utiliser un outil
- Se méfier des **applications IA gratuites** qui demandent un **accès excessif** à vos données personnelles
- Choisir l'**IA en fonction des besoins** (général, coding, photos...)



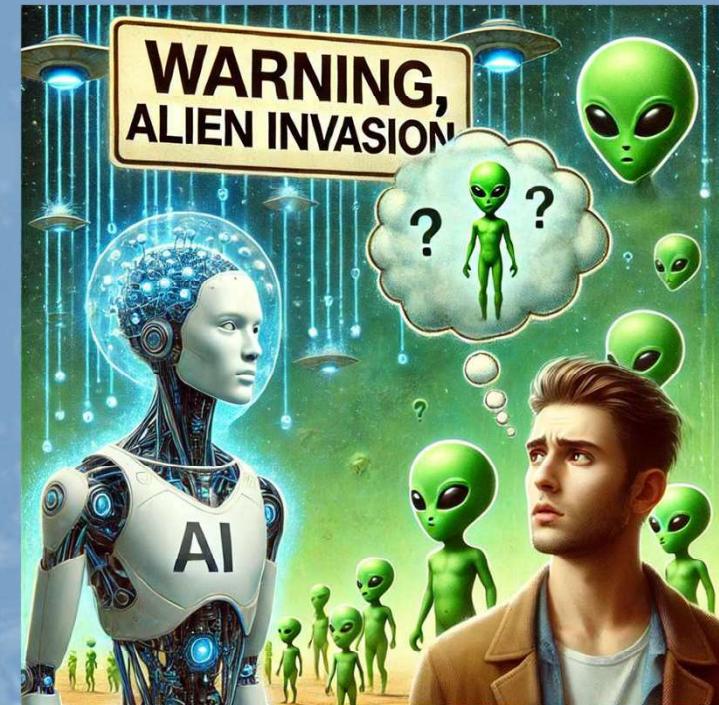
### 3. Bonne pratique, comment utiliser une IA?

#### ➤ Toujours vérifier les informations données par une IA

Une IA peut générer des erreurs, des biais ou des “hallucinations” (réponses fausses mais présentées comme vraies).

Exemples de risques :

- **Biais** : Certaines IA peuvent refléter des discriminations présentes dans les données sur lesquelles elles ont été entraînées
- **Désinformation** : Une IA peut inventer des faits ou mal interpréter une question
- **Manipulation** : Certains systèmes IA peuvent être exploités pour diffuser de fausses informations à grande échelle



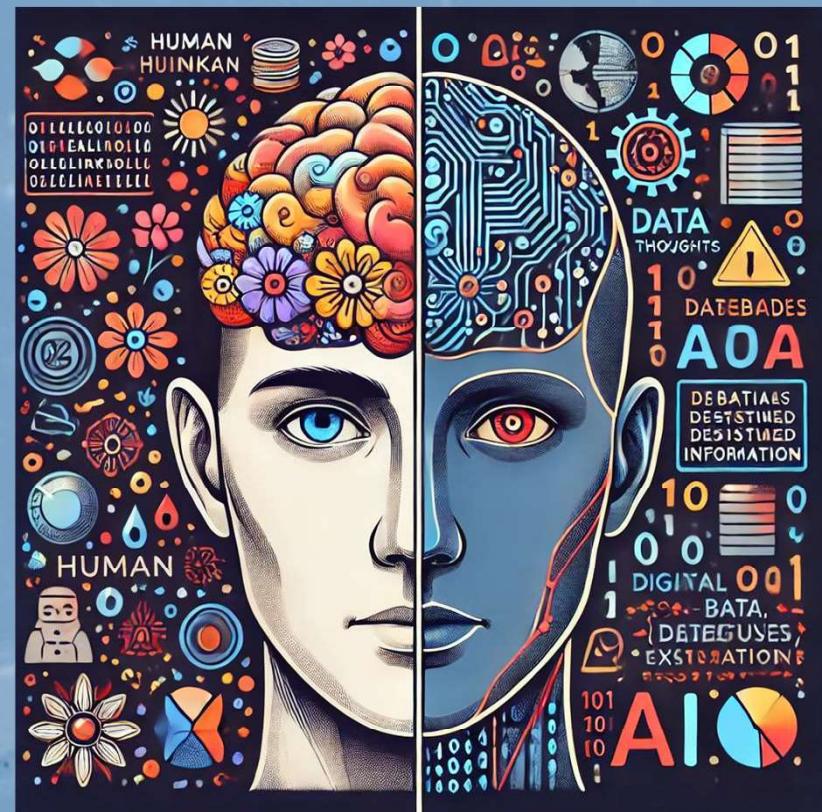
### 3. Bonne pratique, comment utiliser une IA?

#### ➤ Comprendre que l'IA a des limites

L'intelligence artificielle ne réfléchit pas comme un humain et n'a pas de véritable compréhension du monde.

Limitations courantes :

- **Dépendance aux données** : Une IA ne crée pas de nouvelles connaissances, elle se base sur des informations existantes
- **Absence d'émotions réelles** : Un chatbot IA peut sembler empathique, mais il ne ressent rien et ne peut pas remplacer un véritable soutien psychologique
- **Mauvaise compréhension du contexte** : Une IA peut mal interpréter une question complexe ou donner une réponse trop générale



# Intelligence Artificielle et Cybersécurité

## 1. Comprendre l'Intelligence Artificielle

- Définition et fonctionnement
  - Qu'est ce que l'IA?
  - Les différents types IA
  - Comment fonctionne une IA
- Exemple d'IA dans la vie courante
  - Chatbots et assistant vocaux
  - Traduction Automatique
  - Médecine et Diagnostic assisté par IA
  - Reconnaissance faciale
  - Voitures autonomes
  - Recommandations personnalisées
  - Creation d'images et de vidéos

## 2. Les L'IA en Cybersécurité: Atout et Menace

- Les Avantages
  - Détection des menaces
  - Automatisation de la réponse aux attaques
  - Amélioration de la gestion des accès
  - Amélioration des la gestion des logs et des audits
- Les Risques
  - Attaques facilitées par l'IA
  - L'IA et la désinformation: le cas des deepfakes
  - Automatisation des attaques, créations de malware via IA
  - IA entre de mauvaises mains, quand l'IA devient une menace

## 3. Prévention Bonnes pratiques, comment utiliser une IA?

- Ne jamais partager des informations personnelles
- Choisir une IA fiable
- Toujours vérifier, une IA peut se tromper
- Les limites de l'IA