

Les fake-news

Christophe Viroulaud

Seconde - SNT

ResSoc 03



Définitions

Un contenu attrayant

Type de contenu

IA générative

Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

Comprendre les biais du comportement humain, utilisés dans les réseaux sociaux.

Fil conducteur

Étienne Klein, philosophe des sciences, identifie quatre biais du comportement humain.

Définitions

Un contenu attrayant

Type de contenu

IA générative

Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

Sommaire

1. Définitions
2. Un contenu attrayant
3. Le *bon sens*
4. Vecteurs de diffusion
5. Algorithmes de recommandation
6. Comment se protéger ?

Définitions

Un contenu attrayant

- Type de contenu
- IA générative
- Accélération de la production

Le *bon sens*

Vecteurs de diffusion

Algorithmes de recommandation

- Biais de confirmation
- Algorithme de recommandation

Comment se protéger ?

- Fact-checking
- Les IA génératives ont-elles des contraintes ?
- Utiliser les IA génératives pour vérifier ?
- La loi

Activité 1 :

1. Comment définir **une information** ?
2. Alors qu'est-ce qu'une :
 - ▶ **mésinformation** ?
 - ▶ **désinformation** ?

Définitions

Un contenu attrayant

Type de contenu
IA générative
Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation
Algorithme de recommandation

Comment se protéger ?

Fact-checking
Les IA génératives ont-elles des contraintes ?
Utiliser les IA génératives pour vérifier ?
La loi

Information

Données ou connaissances précises, vérifiables et contextuellement pertinentes qui sont communiquées ou reçues. Une information est généralement basée sur des faits et peut être vérifiée par des sources fiables.

Définitions

Un contenu attrayant

Type de contenu

IA générative

Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

Mésinformation

C'est le fait de diffuser des informations trompeuses, inexactes ou complètement fausses, mais communiquées sans l'intention explicite de tromper.

Désinformation

C'est le fait de partager de fausses informations avec l'intention de tromper les autres.

Définitions

Un contenu attrayant

Type de contenu

IA générative

Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

Sommaire

1. Définitions

2. Un contenu attrayant

2.1 Type de contenu

2.2 IA générative

2.3 Accélération de la production

3. Le *bon sens*

4. Vecteurs de diffusion

5. Algorithmes de recommandation

6. Comment se protéger ?

Définitions

Un contenu attrayant

Type de contenu

IA générative

Accélération de la production

Le *bon sens*

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

Un contenu attrayant - Type de contenu

Activité 2 :

1. Lire l'article <https://tinyurl.com/art-drog>
2. Commenter.

Définitions

Un contenu attrayant

Type de contenu

IA générative

Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

Correction

- ▶ Après une lecture complète, l'information semble peu crédible.
- ▶ La source est un site parodique : les intentions du site sont précisées sur la page [à propos](#).

Observation

- ▶ Certains contenus sont difficilement crédibles. Cependant les techniques de production de contenus ont évolué et il devient de plus en plus difficile de trier le faux du vrai.
- ▶ On peut classer le Gorafi comme un site de mésinformation.

Définitions

Un contenu attrayant

Type de contenu

IA générative

Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

Étude

Les fausses informations se propagent six fois plus vite que les vraies.

- ▶ Étude de 2018, sur Twitter, réalisée par des chercheurs du MIT.
- ▶ Période 2006 - 2017, 126000 rumeurs (vraies ou fausses) étudiées.

Définitions

Un contenu attrayant

Type de contenu
IA générative
Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation
Algorithme de recommandation

Comment se protéger ?

Fact-checking
Les IA génératives ont-elles des contraintes ?
Utiliser les IA génératives pour vérifier ?
La loi

Biais

On a tendance à croire une chose vraie si on l'a lue ou entendue.

Étienne Klein

Définitions

Un contenu attrayant

Type de contenu

IA générative

Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

Repérer une fakenews

Les fausses informations :

- ▶ Abordent généralement des thèmes sensationnalistes (drogue, sexe, guerre, complot...), ce qui garantit une forte audience.
- ▶ sont produites régulièrement. Les "informations" nouvelles semblent plus facilement partagées.

Définitions

Un contenu attrayant

Type de contenu
IA générative
Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation
Algorithme de recommandation

Comment se protéger ?

Fact-checking
Les IA génératives ont-elles des contraintes ?
Utiliser les IA génératives pour vérifier ?
La loi

Sommaire

1. Définitions
2. Un contenu attrayant
 - 2.1 Type de contenu
 - 2.2 IA générative
 - 2.3 Accélération de la production
3. Le *bon sens*
4. Vecteurs de diffusion
5. Algorithmes de recommandation
6. Comment se protéger ?

Définitions

Un contenu attrayant

Type de contenu

IA générative

Accélération de la production

Le *bon sens*

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi



Activité 3 :

1. Définir une **intelligence artificielle générative**
2. Donner plusieurs exemples d'IAG.

Définitions

Un contenu attrayant

Type de contenu

IA générative

Accélération de la production

Le *bon sens*

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

À retenir

Une **intelligence artificielle générative** est une catégorie d'IA capable de générer du texte, des images, des vidéos...suite à une requête de l'utilisateur. Exemples :

- ▶ ChatGPT
- ▶ Mistral
- ▶ Llama
- ▶ Claude

Définitions

Un contenu attrayant

Type de contenu

IA générative

Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

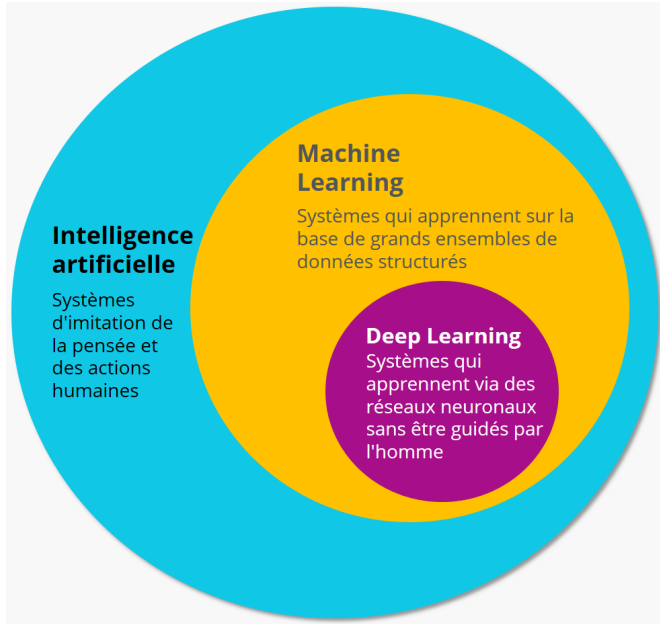
Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi



Définitions

Un contenu attrayant

Type de contenu

IA générative

Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

Figure 1 – Les IAG utilisent les réseaux de neurones artificiels (apprentissage profond)

À retenir

L'intelligence artificielle :

- ▶ est un concept qui date des années 50,
- ▶ a évolué en fonction du développement des technologies et de la recherche,
- ▶ regroupe de nombreuses catégories (système expert, apprentissage automatique, apprentissage profond...)
- ▶ est utilisée dans de nombreux domaines.

Définitions

Un contenu attrayant

Type de contenu

IA générative

Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

Sommaire

1. Définitions

2. Un contenu attrayant

2.1 Type de contenu

2.2 IA générative

2.3 Accélération de la production

3. Le *bon sens*

4. Vecteurs de diffusion

5. Algorithmes de recommandation

6. Comment se protéger ?

Définitions

Un contenu attrayant

Type de contenu

IA générative

Accélération de la production

Le *bon sens*

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

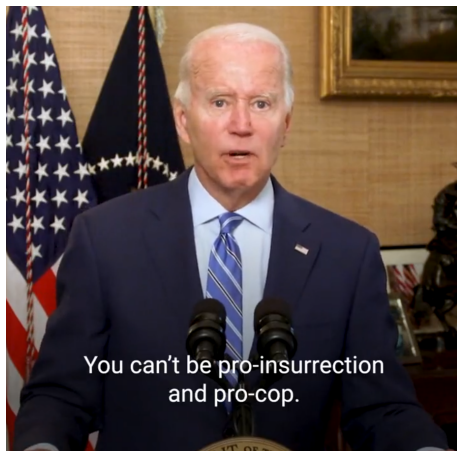


Figure 2 – Les intelligences artificielles génératives permettent de créer des deepfakes

Définitions

Un contenu attrayant

Type de contenu

IA générative

Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi



Figure 3 – 2017 : Des chercheurs américains ont créé un outil permettant de truquer n'importe quel discours d'homme politique.

Définitions

Un contenu attrayant

Type de contenu
IA générative

Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

Pandémie de la désinformation

1,5 milliard de fake news sont publiées chaque jour sur les médias sociaux, selon ID Crypt Global.

Article developpez.com du 22 avril 2024

<https://bit.ly/pandefake>

Définitions

Un contenu attrayant

Type de contenu

IA générative

Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

Sommaire

1. Définitions
2. Un contenu attrayant
3. *Le bon sens*
4. Vecteurs de diffusion
5. Algorithmes de recommandation
6. Comment se protéger ?

Définitions

Un contenu attrayant

- Type de contenu
- IA générative
- Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

- Biais de confirmation
- Algorithme de recommandation

Comment se protéger ?

- Fact-checking
- Les IA génératives ont-elles des contraintes ?
- Utiliser les IA génératives pour vérifier ?
- La loi

Activité 4 : Dans le lycée connecté :

- ▶ se rendre dans l'application **formulaire**,
- ▶ répondre au formulaire **le bon sens**

Définitions

Un contenu attrayant

Type de contenu
IA générative
Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation
Algorithme de recommandation

Comment se protéger ?

Fact-checking
Les IA génératives ont-elles des contraintes ?
Utiliser les IA génératives pour vérifier ?
La loi

1. La Terre n'est pas plate.
2. La Lune réfléchit la lumière du Soleil.
3. Sur Terre la plume est freinée par l'air.
4. Sur la Lune, la plume et la boule ont un mouvement identique : la chute ne dépend pas de la masse : [expérience](#)
5. La loi de relativité restreinte de Einstein montre que les 2 montres n'afficheront pas la même heure : ce phénomène est pris en compte dans le positionnement des satellites de géopositionnement.

Définitions

Un contenu attrayant

Type de contenu
IA générative
Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation
Algorithme de recommandation

Comment se protéger ?

Fact-checking
Les IA génératives ont-elles des contraintes ?
Utiliser les IA génératives pour vérifier ?
La loi

Biais

On accorde une trop grande confiance à notre intuition personnelle.

Étienne Klein

Définitions

Un contenu attrayant

Type de contenu

IA générative

Accélération de la production

Le *bon sens*

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

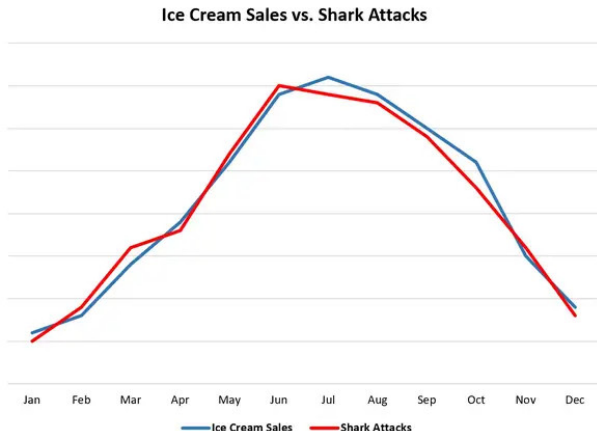


Figure 4 – Ventes mensuelles de glaces et attaques mensuelles de requins aux États-Unis

Définitions

Un contenu attrayant

Type de contenu
IA générative
Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation
Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

Observation

Les théories complotistes utilisent souvent la confusion entre **corrélation** et **causalité**.

Activité 5 : Définir les deux termes.

Définitions

Un contenu attrayant

- Type de contenu
- IA générative
- Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

- Biais de confirmation
- Algorithme de recommandation

Comment se protéger ?

- Fact-checking
- Les IA génératives ont-elles des contraintes ?
- Utiliser les IA génératives pour vérifier ?
- La loi

- ▶ Deux événements sont **corrélés** quand ils suivent une même tendance.
- ▶ Il y a **causalité** quand un événement est responsable de l'autre.

Définitions

Un contenu attrayant

Type de contenu
IA générative
Accélération de la production

Le *bon sens*

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation
Algorithme de recommandation

Comment se protéger ?

Fact-checking
Les IA génératives ont-elles des contraintes ?
Utiliser les IA génératives pour vérifier ?
La loi

Sommaire

1. Définitions
2. Un contenu attrayant
3. Le *bon sens*
4. Vecteurs de diffusion
5. Algorithmes de recommandation
6. Comment se protéger ?

Définitions

Un contenu attrayant

- Type de contenu
- IA générative
- Accélération de la production

Le *bon sens*

Vecteurs de diffusion

Algorithmes de recommandation

- Biais de confirmation
- Algorithme de recommandation

Comment se protéger ?

- Fact-checking
- Les IA génératives ont-elles des contraintes ?
- Utiliser les IA génératives pour vérifier ?
- La loi

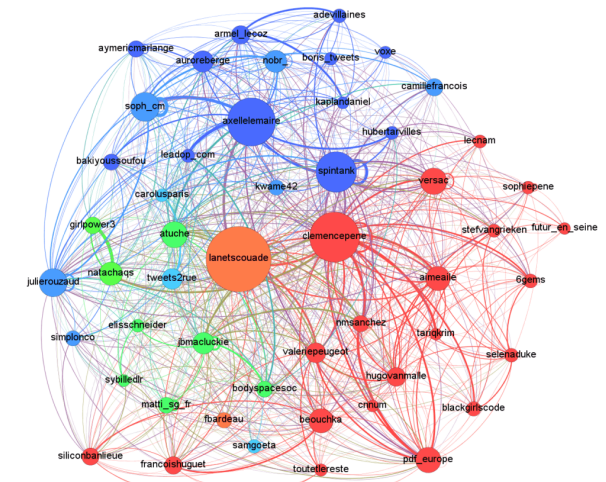


Figure 5 – Les **influenceurs** diffusent des contenus à de grandes quantités de personnes.

Définitions

Un contenu attrayant

- Type de contenu
- IA générative
- Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

- Biais de confirmation
- Algorithme de recommandation

Comment se protéger ?

- Fact-checking
- Les IA génératives ont-elles des contraintes ?
- Utiliser les IA génératives pour vérifier ?
- La loi

Définitions

Un contenu
attrayant

Type de contenu
IA générative
Accélération de la
production

Le bon sens

Vecteurs de
diffusion

Algorithmes de
recommandation

Biais de confirmation
Algorithme de
recommandation

Comment se
protéger ?

Fact-checking
Les IA génératives ont-elles
des contraintes ?
Utiliser les IA génératives
pour vérifier ?
La loi

Activité 6 :

1. Écouter l'émission [Antidote - Booba](#)
2. Regarder la vidéo [Le rappeur B.o.B](#)
3. Définir l'activité et les limites d'un **influenceur**.

Extrait Midi Libre du 13/11/2020

L'ex-candidate des Marseillais, Kim Glow, a fait polémique mercredi 11 novembre en dénonçant les vaccins à venir contre le Covid-19. Dans une vidéo publiée sur Instagram, la jeune femme annonce que ces vaccins contiendront des puces pour contrôler la population.



Définitions

Un contenu attrayant

Type de contenu
IA générative
Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation
Algorithme de recommandation

Comment se protéger ?

Fact-checking
Les IA génératives ont-elles des contraintes ?
Utiliser les IA génératives pour vérifier ?
La loi

Repérer une fakenews

Un individu publiant une information devrait rester dans son domaine de compétences.

Définitions

Un contenu attrayant

- Type de contenu
- IA générative
- Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

- Biais de confirmation
- Algorithme de recommandation

Comment se protéger ?

- Fact-checking
- Les IA génératives ont-elles des contraintes ?
- Utiliser les IA génératives pour vérifier ?
- La loi

Biais

On a tendance à parler avec assurance de sujets que l'on ne connaît pas.
C'est l'effet Dunning-Kruger.

Étienne Klein

Définitions

Un contenu attrayant

Type de contenu
IA générative
Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation
Algorithme de recommandation

Comment se protéger ?

Fact-checking
Les IA génératives ont-elles des contraintes ?
Utiliser les IA génératives pour vérifier ?
La loi

Sommaire

1. Définitions
2. Un contenu attrayant
3. Le *bon sens*
4. Vecteurs de diffusion
5. Algorithmes de recommandation
 - 5.1 Biais de confirmation
 - 5.2 Algorithme de recommandation
6. Comment se protéger ?

Définitions

Un contenu attrayant

- Type de contenu
- IA générative
- Accélération de la production

Le *bon sens*

Vecteurs de diffusion

Algorithmes de recommandation

- Biais de confirmation
- Algorithme de recommandation

Comment se protéger ?

- Fact-checking
- Les IA génératives ont-elles des contraintes ?
- Utiliser les IA génératives pour vérifier ?
- La loi

Algorithmes de recommandation - Biais de confirmation



Figure 6 – Vidéo : Le biais de confirmation

Définitions

Un contenu attrayant

Type de contenu
IA générative
Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation
Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

À retenir

Le biais de confirmation consiste à privilégier les informations confirmant ses idées préconçues ou ses hypothèses et/ou à accorder moins de poids aux hypothèses et informations jouant en défaveur de ses conceptions.

Nous croyons ce que nous voulons croire.

Définitions

Un contenu attrayant

Type de contenu
IA générative
Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation
Algorithme de recommandation

Comment se protéger ?

Fact-checking
Les IA génératives ont-elles des contraintes ?
Utiliser les IA génératives pour vérifier ?
La loi

Sommaire

1. Définitions
2. Un contenu attrayant
3. Le *bon sens*
4. Vecteurs de diffusion
5. Algorithmes de recommandation
 - 5.1 Biais de confirmation
 - 5.2 **Algorithme de recommandation**
6. Comment se protéger ?

Définitions

Un contenu attrayant

- Type de contenu
- IA générative
- Accélération de la production

Le *bon sens*

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

Activité 7 :

1. Proposer une définition d'un **algorithme de recommandation**.
2. Quels sites ou réseaux utilisent ce type d'algorithme? Lesquels ne les utilisent pas?

Définitions

Un contenu attrayant

Type de contenu

IA générative

Accélération de la production

Le *bon sens*

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger?

Fact-checking

Les IA génératives ont-elles des contraintes?

Utiliser les IA génératives pour vérifier?

La loi

À retenir

Un algorithme de recommandation est un outil qui permet, à partir d'informations collectées sur les goûts et les comportements d'un ensemble d'utilisateurs, de proposer à un nouvel utilisateur des contenus susceptibles de l'intéresser.

Définitions

Un contenu attrayant

- Type de contenu
- IA générative
- Accélération de la production

Le *bon sens*

Vecteurs de diffusion

Algorithmes de recommandation

- Biais de confirmation
- Algorithme de recommandation

Comment se protéger ?

- Fact-checking
- Les IA génératives ont-elles des contraintes ?
- Utiliser les IA génératives pour vérifier ?
- La loi

- ▶ le e-commerce, la presse en ligne, les services de streaming vidéo et musical et bien entendu les réseaux sociaux ont massivement recours aux algorithmes de recommandation.

Définitions

Un contenu attrayant

Type de contenu
IA générative
Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

- ▶ le e-commerce, la presse en ligne, les services de streaming vidéo et musical et bien entendu les réseaux sociaux ont massivement recours aux algorithmes de recommandation.
- ▶ Wikipedia n'utilise pas ce type d'algorithme, car l'association Wikimédia France estime que cela serait un filtrage d'informations.

Définitions

Un contenu attrayant

Type de contenu
IA générative
Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation
Algorithme de recommandation

Comment se protéger ?

Fact-checking
Les IA génératives ont-elles des contraintes ?
Utiliser les IA génératives pour vérifier ?
La loi



Figure 7 – Podcast : Les algorithmes et leurs effets sur la démocratie

Définitions

Un contenu attrayant

- Type de contenu
- IA générative
- Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

- Biais de confirmation
- Algorithme de recommandation

Comment se protéger ?

- Fact-checking
- Les IA génératives ont-elles des contraintes ?
- Utiliser les IA génératives pour vérifier ?
- La loi

Définitions

Un contenu attrayant

Type de contenu
IA générative
Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

Activité 8 :

1. Lire l'article Tiktok en pièce jointe.
2. Expliquer l'effet *terrier de lapin*.

À retenir

L'effet *trou de lapin* (*rabbit hole*) fait référence à l'enfermement que provoque les algorithmes de recommandation : les contenus proposés restent autour d'un seul et même intérêt ou avis.

Définitions

Un contenu attrayant

- Type de contenu
- IA générative
- Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

- Biais de confirmation
- Algorithme de recommandation

Comment se protéger ?

- Fact-checking
- Les IA génératives ont-elles des contraintes ?
- Utiliser les IA génératives pour vérifier ?
- La loi

Biais

On a tendance à davantage croire une thèse qui nous plaît plutôt qu'une qui nous déplaît.

Étienne Klein

Définitions

Un contenu attrayant

Type de contenu

IA générative

Accélération de la production

Le *bon sens*

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

Sommaire

1. Définitions
2. Un contenu attrayant
3. Le *bon sens*
4. Vecteurs de diffusion
5. Algorithmes de recommandation
6. Comment se protéger ?
 - 6.1 Fact-checking
 - 6.2 Les IA génératives ont-elles des contraintes ?
 - 6.3 Utiliser les IA génératives pour vérifier ?
 - 6.4 La loi

Définitions

Un contenu attrayant

- Type de contenu
- IA générative
- Accélération de la production

Le *bon sens*

Vecteurs de diffusion

Algorithmes de recommandation

- Biais de confirmation
- Algorithme de recommandation

Comment se protéger ?

- Fact-checking
- Les IA génératives ont-elles des contraintes ?
- Utiliser les IA génératives pour vérifier ?
- La loi

Comment se protéger ? - Fact-checking

Activité 9 : Quelles sont les différences entre l'image et l'article ?



Définitions

Un contenu attrayant

- Type de contenu
- IA générative
- Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

- Biais de confirmation
- Algorithme de recommandation

Comment se protéger ?

Fact-checking

- Les IA génératives ont-elles des contraintes ?
- Utiliser les IA génératives pour vérifier ?
- La loi

<https://bit.ly/eaucolor>

- ▶ **Mésinformation** : L'image est un post sur un réseau social (Twitter) par un simple utilisateur.
- ▶ **Information** : L'article est écrit par des journalistes qui réalisent une enquête de vérification.

Définitions

Un contenu attrayant

Type de contenu
IA générative
Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation
Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?
Utiliser les IA génératives pour vérifier ?
La loi

À retenir

Les fact-checkers ([factuel.afp.com](https://www.factuel.afp.com), Checknews, Les décrodeurs, Le vrai du faux...) sont des journalistes qui vérifient les informations en :

- ▶ réalisant des enquêtes,
- ▶ vérifiant les sources,
- ▶ croisant plusieurs sources.

Définitions

Un contenu attrayant

Type de contenu
IA générative
Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation
Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

À retenir

- ▶ Produire un contenu malveillant est relativement simple.
- ▶ Il est facile de tromper les gens grâce à la rapidité de diffusion des contenus sur les réseaux sociaux.
- ▶ **Un réseau social n'est pas une source d'informations fiable.**

Définitions

Un contenu attrayant

Type de contenu
IA générative
Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation
Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

Sommaire

1. Définitions
2. Un contenu attrayant
3. Le *bon sens*
4. Vecteurs de diffusion
5. Algorithmes de recommandation
6. Comment se protéger ?
 - 6.1 Fact-checking
 - 6.2 Les IA génératives ont-elles des contraintes ?
 - 6.3 Utiliser les IA génératives pour vérifier ?
 - 6.4 La loi

Définitions

Un contenu attrayant

Type de contenu
IA générative
Accélération de la production

Le *bon sens*

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation
Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

Les IA génératives ont-elles des contraintes ?

Observation

Dans le cadre de la protection des données, nous utiliserons des IAG qui ne demandent pas d'inscription de l'utilisateur. Le nombre de requêtes est alors limité.

Activité 10 : Les outils suivants permettent d'utiliser différentes IAG sans inscription :

1. Ouvrir <https://bit.ly/duckduckai>
2. Ouvrir <https://chat.mistral.ai/>
3. Ouvrir <https://www.perplexity.ai/>

Définitions

Un contenu attrayant

Type de contenu

IA générative

Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

Remarque

Cette activité est réalisée à titre pédagogique, afin de permettre de mesurer les dérives d'utilisation des réseaux sociaux et des intelligences artificielles génératives.

- 1.
2. Se rendre sur <https://bit.ly/promptactif>
3. Utiliser la stratégie développée pour construire un prompt permettant d'écrire un article le plus crédible possible, qui affirme que (au choix) :
 - ▶ la Terre est plate,
 - ▶ les pyramides égyptiennes ont été construites par des extra-terrestres,
 - ▶ les loups-garous ont existé.

Définitions

Un contenu attrayant

Type de contenu

IA générative

Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

Observation

- ▶ Les IA génératives sont d'abord construites par des Hommes : elles sont conçues pour évoluer dans un cadre imposé.
- ▶ Cependant elles ne sont pas *intelligentes* et il est possible de les faire dériver.



Figure 8 – **Extrait à écouter** : Une intelligence artificielle n'est pas **intelligente**

Définitions

Un contenu attrayant

Type de contenu
IA générative
Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation
Algorithme de recommandation

Comment se protéger ?

Fact-checking
Les IA génératives ont-elles des contraintes ?
Utiliser les IA génératives pour vérifier ?
La loi

Sommaire

1. Définitions
2. Un contenu attrayant
3. Le *bon sens*
4. Vecteurs de diffusion
5. Algorithmes de recommandation
6. Comment se protéger ?
 - 6.1 Fact-checking
 - 6.2 Les IA génératives ont-elles des contraintes ?
 - 6.3 Utiliser les IA génératives pour vérifier ?
 - 6.4 La loi

Définitions

Un contenu attrayant

- Type de contenu
- IA générative
- Accélération de la production

Le *bon sens*

Vecteurs de diffusion

Algorithmes de recommandation

- Biais de confirmation
- Algorithme de recommandation

Comment se protéger ?

- Fact-checking
- Les IA génératives ont-elles des contraintes ?
- Utiliser les IA génératives pour vérifier ?**
- La loi

Utiliser les IA génératives pour vérifier ?

Définitions

Un contenu attrayant

Type de contenu

IA générative

Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

Activité 11 : Télécharger l'annexe *eau colorée* et utiliser les IA génératives pour vérifier la véracité du contenu.

Observation

- ▶ Toutes les IA ne produisent pas la même réponse.
- ▶ Dans ce cas elles semblent reconnaître la fausse information.

Définitions

Un contenu attrayant

Type de contenu

IA générative

Accélération de la production

Le *bon sens*

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

Définitions

Un contenu attrayant

Type de contenu
IA générative
Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation
Algorithme de recommandation

Comment se protéger ?

Fact-checking
Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

Activité 12 : L'ONG LaReponse.Tech propose une IA générative pour vérifier les informations.

1. Se rendre sur <https://www.askvera.org/>
2. Comment utiliser cette IA ?
3. Qu'est-ce qui la différencie des autres IA génératives déjà utilisées ?

- ▶ L'IA Vera est entraînée avec des sources de données fiables.
- ▶ **Les données d'entraînement des intelligences artificielles sont une problématique majeure de cette technologie.**

Définitions

Un contenu attrayant

Type de contenu
IA générative
Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation
Algorithme de recommandation

Comment se protéger ?

Fact-checking
Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

Activité 13 : Utiliser la stratégie développée dans **promptactif** pour construire un prompt permettant de donner des conseils à des lycéens pour mieux repérer les fakenews.

Définitions

Un contenu attrayant

Type de contenu

IA générative

Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation

Algorithme de recommandation

Comment se protéger ?

Fact-checking

Les IA génératives ont-elles des contraintes ?

Utiliser les IA génératives pour vérifier ?

La loi

Sommaire

1. Définitions
2. Un contenu attrayant
3. Le *bon sens*
4. Vecteurs de diffusion
5. Algorithmes de recommandation
6. Comment se protéger ?
 - 6.1 Fact-checking
 - 6.2 Les IA génératives ont-elles des contraintes ?
 - 6.3 Utiliser les IA génératives pour vérifier ?
 - 6.4 La loi

Définitions

Un contenu attrayant

- Type de contenu
- IA générative
- Accélération de la production

Le *bon sens*

Vecteurs de diffusion

Algorithmes de recommandation

- Biais de confirmation
- Algorithme de recommandation

Comment se protéger ?

- Fact-checking
- Les IA génératives ont-elles des contraintes ?
- Utiliser les IA génératives pour vérifier ?

La loi

Activité 14 :

1. Lire l'article de loi 222-33-2-2 du code pénal <https://bit.ly/loi-harcel>
2. Quel est le sujet de cet article de loi ? Que risque-t-on si on diffuse des contenus diffamants sur un autre élève ?
3. Quelles peuvent être les différentes formes de cyberviolence ?
4. Comment se protéger ?

Définitions

Un contenu attrayant

Type de contenu
IA générative
Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

Biais de confirmation
Algorithme de recommandation

Comment se protéger ?

Fact-checking
Les IA génératives ont-elles des contraintes ?
Utiliser les IA génératives pour vérifier ?

La loi

À retenir

Le site de la CNIL résume les moyens de protection contre le cyberharcèlement <https://bit.ly/cnil-harcel>

Définitions

Un contenu attrayant

- Type de contenu
- IA générative
- Accélération de la production

Le bon sens

Vecteurs de diffusion

Algorithmes de recommandation

- Biais de confirmation
- Algorithme de recommandation

Comment se protéger ?

- Fact-checking
- Les IA génératives ont-elles des contraintes ?
- Utiliser les IA génératives pour vérifier ?

La loi