

## Comprendre les outils d'intelligence artificielle.



Créé en partenariat avec



Années scolaires : de la 6<sup>e</sup> année à la 5<sup>e</sup> année secondaire

Durée : de 60 à 90 minutes, plus les activités après le cours

### Objectifs d'apprentissage :

#### Concepts/grandes idées clés :

- **Les médias sont des concepts** : Les algorithmes ne « pensent » pas comme les humains; ils font des liens en trouvant des tendances dans les données.
- **Les médias ont des implications sociales et politiques** : Certaines étapes doivent être respectées pour s'assurer que les algorithmes sont utilisés de manière juste et éthique.
- **Les médias numériques ont un public inattendu** : Les algorithmes se servent de nos données personnelles pour prendre des décisions nous concernant.
- **Les expériences avec les médias numériques sont façonnées par les outils que nous utilisons** : Ce que les algorithmes savent (ou pensent savoir) sur nous peut changer la manière dont nous interagissons avec eux.
- **Les interactions sur les médias numériques ont de véritables conséquences** : Nous devons garder un esprit critique et agir de manière éthique lorsque nous utilisons ou interagissons avec des algorithmes.

#### Questions clés :

- Comment fonctionnent les algorithmes?
- Comment savoir si quelque chose a été créé par une intelligence artificielle?

- Comment savoir si les algorithmes sont utilisés de manière juste?
- Comment contrôler la manière dont les algorithmes utilisent nos renseignements personnels?
- Quelles questions essentielles devrions-nous poser lorsque nous utilisons des algorithmes?
- Comment se servir d'un algorithme de manière juste?

#### Les idées fausses à bannir :

- L'intelligence artificielle fonctionne comme l'intelligence humaine, délibérément et intentionnellement.
- Les algorithmes sont toujours justes, car ils prennent des décisions selon des données uniquement.

#### Préparation:

##### Préparez-vous à distribuer ce qui suit :

- Document : Comprendre les outils d'IA (page 10)
- Document : Comprendre les outils de l'IA – Notes de présentation (page 8)
- Document : Comment utiliser les outils de clavardage intelligent de manière sûre et éthique? (page 14, imprimez ce document recto-verso)

**Devoirs** : Comprendre les outils d'IA – Règles pour des outils de clavardage intelligent sûrs et éthiques (page 16)

## Procédure

**Commencez par demander aux élèves :** Combien d'entre vous ont déjà utilisé un outil de clavardage intelligent ou un assistant virtuel? Lesquels avez-vous utilisés? (Vous pouvez citer quelques exemples tels que : ChatGPT, Character AI, My AI, Alexa, Siri.)

Distribuez le document **Comprendre les outils d'IA** et demandez aux élèves de remplir ce sondage

**Pensez-vous que les outils de clavardage intelligent...**

- Ont des émotions?
- Peuvent ressentir la douleur ou le rejet?
- Peuvent choisir de faire ou ne pas faire certaines choses?
- Sont intelligents?
- Vous donnent toujours la bonne information?
- Savent s'ils vous transmettent la bonne information ou non?
- Peuvent penser au bien ou au mal de certaines actions?

Passez en revue les points du sondage et demandez aux élèves s'ils ont répondu « oui », « non » ou « peut-être ». Pour chaque point, demandez-leur d'approfondir leur réponse.

Faites savoir que de nombreux enfants, même à leur âge, ne savent pas ce genre de choses : selon une étude, un quart des enfants de 11 ans pensent que « Alexa » (une enceinte intelligente qui utilise l'IA) ressent des émotions et peut penser par elle-même, et un autre tiers pensent que ces suppositions étaient « peut-être » vraies.

Dites aux élèves que les outils d'IA ont bel et bien l'air intelligent (d'où l'emploi du terme « intelligence artificielle ») et pourraient sembler capables de ressentir des émotions. Ils prennent même des décisions – p. ex. ils pourraient refuser un ordre que vous leur donnez, et si vous leur posez une question, ils peuvent « choisir » ce qu'ils vont dire ou omettre dans leur réponse.

## Mais alors, qu'est-ce qu'un « outil d'IA », vraiment?

L'intelligence artificielle se définit comme toute technologie de l'information capable de résoudre des problèmes complexes qui se poseraient normalement à des humains ou à des animaux.

L'IA comprend de nombreuses technologies. **L'IA comprend de nombreuses technologies. Ce plan de cours porte spécifiquement sur les algorithmes d'apprentissage-machine.**

Demandez aux élèves ce que signifie selon eux « IA ». Laissez-leur quelques minutes pour écrire leurs réponses sur le document, demandez s'il y a un ou une volontaire pour partager ses réponses.

Expliquez que la plupart des choses que l'on appelle « IA » sont des algorithmes d'apprentissage-machine, et demandez aux élèves s'ils peuvent citer d'autres exemples d'algorithmes. (Exemples : moteurs de recherche, le contenu recommandé sur YouTube ou Netflix, les publicités diffusées sur les médias sociaux, les algorithmes de création de contenu/image comme MidJourney ou Dall-E, des algorithmes de saisie semi-automatique.)

Dites aux élèves qu'un algorithme est une série d'instructions qui dirigent le processus décisionnel. L'utilisation d'algorithmes rend les décisions plus cohérentes et plus équitables, car les mêmes normes sont appliquées pour tous – en revanche, si vous ne faites pas attention, vous pourriez ne pas réaliser lorsqu'un algorithme n'est **pas** équitable.

**Point no1 à considérer : L'IA est-elle toujours juste?**

Lisez la section **Point no1 à considérer** du document. Expliquez aux élèves qu'il y a toujours plus de personnes en attente d'une greffe de rein que de reins disponibles, et qu'il peut être très difficile de choisir qui recevra un rein.

Parlez de l'équation  $PG = EV$  du document et expliquez aux élèves que ce simple algorithme a été créé pour décider qui passera en priorité : PG signifie « priorité pour une greffe » et EV signifie « espérance de vie ». En d'autres termes, une personne a plus de chances de recevoir une greffe qu'une autre si elle est susceptible de vivre plus longtemps que l'autre personne. Laissez aux élèves quelques minutes pour répondre à ces questions :

- Est-ce que cela vous paraît raisonnable?
- Est-ce que cela paraît juste? Pourquoi?

Si les élèves ne posent pas les questions suivantes, demandez :

- Comment peut-on vraiment calculer une espérance de vie? Comment peut-on savoir, ou deviner, de combien d'années dispose une personne si elle reçoit un rein?
- Plus une personne est vieille, moins elle aura d'années devant elle, même avec une greffe de rein. Est-il juste de donner plus de chances à une personne jeune de recevoir un rein?
- L'espérance de vie varie en fonction des groupes démographiques : les groupes de personnes qui subissent des discriminations ou ceux qui ont moins accès aux soins de santé ont des espérances de vie plus courtes. Est-ce juste que les personnes qui ont déjà des espérances de vie plus longues aient la priorité pour des greffes de rein?

À la fin de la conversation, écrivez sur le tableau le principe selon lequel **les algorithmes peuvent renforcer les préjugés du monde réel.**

Expliquez que les personnes qui conçoivent les algorithmes pour les greffes de rein se sont rendu compte du problème et ont décidé d'y remédier. (Le véritable algorithme prend en compte d'autres variables. Par exemple, il prend en compte l'état de santé du donneur de rein – les patients plus jeunes ne sont pas forcément plus susceptibles de recevoir un rein, mais ils recevront un rein en meilleure santé.)

**Point no2 à considérer : en quoi l'IA est différente des simples algorithmes?**

Lisez à présent le **Point no2 à considérer** et expliquez aux élèves que presque tous les algorithmes qui font partie de leur vie – les moteurs de recherche, les algorithmes de recommandation que YouTube et TikTok utilisent, les algorithmes de tri comme le fil Instagram (qui décide quelles publications apparaîtront sur votre fil d'actualité et dans quel ordre) – ne sont pas que de simples algorithmes comme ceux que nous venons d'aborder, mais des algorithmes d'apprentissage-machine ou d'IA. Laissez-leur quelques minutes pour écrire une réponse à la question « Comment cela peut rendre plus difficile de s'assurer que le IA prend des décisions justes? », puis reprenez cette question en classe.

Assurez-vous d'aborder les éléments suivants :

- L'IA est entraînée sur des données du monde réel, donc tout comme l'algorithme des dons de reins, l'algorithme va refléter les préjugés, à moins d'y faire attention.
- Si nous ne savons pas pourquoi elles ont pris telle ou telle décision, nous pourrions ne pas réaliser que cette décision est fondée sur des préjugés.

À la fin de la discussion, écrivez sur le tableau le principe selon lequel **nous ne savons pas toujours comment l'IA prend des décisions.**

**Point no3 à considérer : Les IA prennent-elles toujours les bonnes décisions?**

Lisez à présent le **Point no3 à considérer**. Laissez quelques minutes aux élèves pour répondre à cette question et reprenez-la en classe.

Assurez-vous d'aborder les éléments suivants :

- À mesure que le temps passe, plus vous aurez de crème glacée et moins vous aurez de brocoli. Au bout du compte, vous ne mangerez plus que des crèmes glacées et plus du tout de brocoli. Si votre dîner est composé de plus de deux éléments, à un moment donné vous ne garderez que votre nourriture favorite et rien d'autre.
- C'est peut-être super de manger moins de brocoli et plus de crème glacée, mais ce n'est pas forcément bon pour votre santé!

Expliquez aux élèves que la plupart des algorithmes présents dans nos vies fonctionnent de la même manière : ils sont optimisés pour vous présenter des choses qui vous feront rester sur l'application ou la plateforme que vous utilisez.

Demandez à présent aux élèves : Quels types de « crème glacée » (c'est-à-dire les choses mauvaises pour vous, que ce soit en petite ou grande quantités) l'IA est-elle susceptible de vous présenter? Quels exemples de « brocoli » l'AI devrait-elle vous offrir même si vous n'en avez pas toujours envie? Demandez aux élèves de réfléchir à cette question dans le contexte de leurs fils d'actualité sur les médias sociaux.

Exemples de « crème glacée » (les choses que vous voyez souvent sur votre fil d'actualité, mais que vous regrettez avoir lues ou vues) :

- Les « pièges à clics » ou les « fausses nouvelles » (les choses qui paraissent géniales, mais qui sont fausses ou trompeuses)
- Les choses qui vous mettent en colère ou qui vous contrarient
- Les choses qui vous font vous sentir mal dans votre peau
- Les choses qui vous incitent à continuer de regarder
- Les choses dangereuses (comme le contenu qui encourage les défis viraux dangereux)

Des exemples de « brocoli » (les choses que vous devez chercher sur votre fil et qui vous font vous sentir bien après les avoir lues ou vues) :

- Les nouvelles
- Les alertes d'urgence
- L'information fiable
- Les choses qui vous font vous sentir bien dans votre peau
- Les choses qui vous encouragent à entreprendre quelque chose de sain ou porteur de sens

À la fin de la discussion, écrivez sur le tableau le principe selon lequel **l'IA vous donne ce qu'elle pense que vous voulez, pas ce qui est bon pour vous.**

**Point à considérer no4 : D'où proviennent les données utilisées par l'IA pour prendre des décisions?**

Lisez à présent le **Point no4 à considérer**. Laissez quelques minutes aux élèves pour répondre à cette question et reprenez-la en classe.

Assurez-vous d'aborder les éléments suivants :

- Assurez-vous d'aborder les éléments suivants :
- L'IA peut refléter l'historique de recherche, de visionnement ou d'achat des gens avec des comportements plus ou moins risqués. Par exemple, acheter acheter un skateboard, ou regarder des vidéos de skateboard.
- Cela pourrait être avantageux pour les personnes considérées comme « moins à risque », car elles paieraient probablement moins, mais les personnes considérées comme « à risque » paieraient sans doute plus.

- Les données pourraient faire que l'IA prend des décisions plus précises que celles qui se basent sur des catégories très larges comme le genre. Les connexions faites peuvent être inexactes dans deux cas :
  - Des données peuvent être mal interprétées (p. ex. : si vous avez acheté un skateboard comme cadeau pour quelqu'un d'autre).
  - La connexion établie entre un type de risque et un autre peut être inexacte (p. ex. vous pourriez aimer les sports extrêmes, mais quand même conduire très prudemment).

Expliquez à présent aux élèves que quasiment tout l'IA fonctionne comme cela. Elle utilise soit :

- Des données collectées vous concernant (p. ex. les questions posées à un clavardage intelligent, qui repose sur l'IA, ou de l'information que votre ordinateur a automatiquement envoyée comme votre localisation ou votre adresse IP)
- L'information achetée par l'entreprise (soit une autre entreprise ou de courtiers en données qui achètent les données depuis de nombreux endroits et les rassemblent pour en faire des profils)
- Ou bien de l'information partagée entre différents secteurs d'une même entreprise (p. ex. Google peut utiliser vos vues YouTube pour décider quels résultats de recherche vous proposer, et vice versa, car Google et YouTube font partie de la même entreprise)

Écrivez maintenant sur le tableau le principe selon lequel **l'IA prend des décisions selon ce qu'elle pense sur vous.**

**Point no5 à considérer : En quoi les clavardages intelligents sont-ils similaires à d'autres algorithmes d'apprentissage-machine?**

Lisez à présent le **Point no5 à considérer** avec la classe.

Ce n'est pas grave si vous ne vous sentez pas à l'aise de répondre à plus de questions détaillées sur ce contenu : un expert a un jour dit « personne sur terre ne comprend parfaitement le fonctionnement des grands modèles de langage ». Si vous souhaitez une explication plus détaillée, lisez l'article [Les grands modèles de langage vulgarisés](#)

## Mais alors, qu'est-qu'un outil d'IA, vraiment?

Demandez aux élèves de retourner à la section « Mais alors, qu'est-ce qu'un outil d'IA, vraiment? » du document de compréhension des outils d'IA et revoyez la définition qu'ils ont écrite. Demandez-leur si leur définition a changé après la discussion avec la classe; si c'est le cas, demandez-leur d'écrire une nouvelle définition.

Rappelez à présent aux élèves les exemples de clavardages intelligents mentionnés au début de la leçon, puis les quatre principes qui ont émergé au cours de la discussion :

1. Les algorithmes peuvent renforcer les préjugés du monde réel;
2. Nous ne savons pas toujours comment l'IA prend des décisions;
3. L'IA vous donne ce qu'elle pense que vous voulez, pas ce qui est bon pour vous;
4. L'IA prend des décisions selon ce qu'elle pense savoir sur vous.

Divisez à présent les élèves en groupes de 3 à 4 personnes et attribuez l'un de ces principes à chaque groupe. (chaque principe peut être attribué à plus d'un groupe selon le nombre d'élèves.)

Demandez à chaque groupe de discuter de la manière dont chaque principe s'applique aux clavardages intelligents. À mesure que la conversation entre les élèves avance, passez entre les groupes et posez-leur les questions suivantes :

**Groupe 1:** Quels genres de préjugés peut-on trouver dans les messages sur lesquels les clavardages intelligents s'entraînent? Parlez des publications des médias sociaux, ou des vieux livres.) Quelle incidence cela pourrait-il avoir sur leurs réponses?

**Groupe 2:** Les personnes qui créent les clavardages intelligents devraient-elles être capables d'expliquer comment ils fonctionnent? Comment le fait que même certains concepteurs ne savent souvent pas comment l'IA prend leurs décisions rend-il plus difficile de créer des mécanismes de protection ou des « protecteurs » qui pourront empêcher certains autres problèmes?

**Groupe 3:** Pouvez-vous penser à d'autres sujets ou renseignements que vous aimeriez aborder ou recevoir d'un clavardage intelligent, mais que vous savez ne pas être bons pour vous? Que pourrait-il mal se passer si vous demandez conseil au clavardage intelligent, ou que vous lui parliez d'un problème personnel?

**Groupe 4:** Comment pensez-vous que le clavardage intelligent récolte de l'information sur vous? Pensez-vous que c'est juste? Pensez-vous que c'est exact?

Demandez à présent aux groupes qui travaillaient sur le même principe de se rassembler et de partager leurs discussions. Demandez-leur d'identifier les mêmes points discutés dans les groupes et discutez de ceux qui n'ont été discutés que dans un groupe. Les groupes n'ont pas à être d'accord sur tous les points, mais ils doivent identifier les points sur lesquels ils ne sont pas d'accord.

Ensuite, demandez aux groupes rassemblés de faire une courte présentation à la classe dans laquelle ils expliqueront la manière dont leur principe s'applique aux clavardages intelligents. Ces groupes doivent également émettre une suggestion sur la manière dont les personnes qui utilisent des clavardages intelligents peuvent résoudre les problèmes identifiés ou améliorer les points faibles.

- Si c'est possible, demandez aux groupes d'écrire leurs points sur un bout de papier qu'ils afficheront au mur ou au tableau après leur présentation.
- Avant le passage d'un groupe, distribuez le document « Comprendre les outils d'IA : notes de présentation (page 10) ». Assurez-vous que les élèves comprennent qu'ils devront se servir des présentations des autres groupes dans la prochaine partie de la leçon et encouragez-les à utiliser le document pour prendre des notes.

## Comprendre les outils d'IA : Les règles pour des clavardages intelligents sûrs et éthiques

Distribuez la fiche de travail « Comprendre les outils d'IA : Les règles pour des clavardages intelligents sûrs et éthiques (page 16) » et étudiez-la avec les élèves.

Expliquez-leur qu'ils doivent inscrire au moins deux règles, codes ou directives pour les quatre principes, et pour un total d'au moins huit.

- Ces règles, codes ou directives seront destinés aux personnes qui utilisent ou conçoivent des clavardages intelligents ou les utilisent dans leurs services (p. ex. un réseau social ou un moteur de recherche qui intègre un clavardage intelligent dans sa plateforme).
- Les élèves doivent également écrire quelques phrases sous chaque règle, code ou directive pour expliquer pourquoi c'est important pour eux.

Ces règles et explications doivent également démontrer le savoir et la compréhension des élèves concernant les algorithmes et les clavardages intelligents acquis grâce au document Comprendre l'IA, ainsi que les discussions de classe, et les présentations des autres groupes sur la manière dont les quatre principes de l'IA s'appliquent aux clavardages intelligents

## Question de réflexion : Comment utiliser les outils de clavardage intelligents de manière sûre et éthique?

Distribuez le document de réflexion : Comment utiliser les outils de clavardage intelligents de manière sûre et éthique? (page 14) Demandez aux élèves de lire la question et de prendre quelques minutes pour y répondre.

Une fois la réponse écrite, demandez-leur de tourner la feuille et de lire les conseils notés. Ces conseils répondent-ils aux interrogations mentionnées dans leurs réponses?

## Extension: Utiliser les clavardages intelligents en classe

Les activités précédentes ont été créées sans besoin d'accès à un service de clavardage intelligent, à cause des préoccupations relatives à la vie privée qui leur sont associées. Si vous ou l'un des élèves avez accès à un clavardage intelligent et que vous souhaitez l'intégrer dans la présentation, vous pouvez faire certaines ou toutes les activités suivantes.

Les modèles de clavardages intelligents sont continuellement en développement. Il vous est donc recommandé de toujours mettre à l'essai ces exercices avant de les faire en classe.

Assurez-vous de bien connaître les politiques de votre conseil ou de votre division concernant l'utilisation de plateformes commerciales en ligne dans votre classe.

**1. Utilisez l'IA pour votre propre bénéfice :** Faites remarquer aux élèves que pendant longtemps, les étudiants n'avaient pas le droit d'utiliser de calculatrices en classe de maths, peu importe leur niveau d'études. À présent, les élèves plus âgés sont autorisés à les utiliser afin de pouvoir consacrer plus de temps à des problèmes de maths complexes plutôt qu'à l'arithmétique.

- Demandez : Quelles sont les choses qui, dans notre classe, ressemblent à de l'arithmétique? (Les choses que vous devez répéter alors que vous les avez déjà apprises). Comment pouvez-vous utiliser les clavardages intelligents comme une calculatrice, pour libérer plus de temps à la pensée critique et d'autres tâches de plus haut niveau?
- Demandez aux élèves d'expérimenter en utilisant un clavardage intelligent pour les tâches que vous avez identifiées.
- Demandez aux élèves de créer un processus et un code d'éthique pour l'utilisation des clavardages intelligents à l'école.

**2. L'IA de vérification des faits :** Posez une question à un clavardage intelligent sur un sujet récemment abordé en classe, et demandez aux élèves de lire et de commenter le résultat. (Vous pouvez également demander à un élève de poser une question sur un sujet qu'il connaît ou maîtrise.)

- Les résultats sont-ils corrects? (Les grands modèles de langage ont tendance à « halluciner » des détails.)

- Le résultat a-t-il démontré une bonne compréhension du sujet? (Les grands modèles de langage accordent souvent trop d'importance ou trop peu d'importance aux différents détails en comparaison à ce qu'un expert humain considérerait comme important. En raison de la manière dont ils s'entraînent, ils considèrent aussi souvent des idées fausses largement répandues comme des faits.)
- La formulation de la question change-t-elle l'exactitude des résultats? (Par exemple, utiliser la phrase « écrire un essai sur » conduira souvent à des résultats moins précis que le fait de poser une question.)

**3. Résumé critique :** Demandez aux élèves de lire une histoire courte très connue, puis demandez à la classe de faire un résumé des points importants de l'intrigue et de l'interprétation de sa signification. Puis, demandez au clavardage intelligent de faire la même chose, et demandez aux élèves de comparer les résultats. À quel point sont-ils différents?

**4. Vos références? :** Posez une question à un clavardage intelligent concernant un sujet académique et demandez-lui de donner des références de sources académiques. Demandez aux élèves de vérifier les sources. Combien d'entre elles sont réelles? Combien d'entre elles avaient le bon auteur et la bonne date de publication?

**5. Augmenter le niveau :** Choisissez une idée ou un concept étudié en classe, et demandez au clavardage intelligent de l'expliquer à trois niveaux différents : aux enfants de l'école primaire, aux élèves du secondaire, et aux professeurs qui s'apprentent à enseigner le sujet. Demandez-leur de comparer les résultats. En quoi sont-ils différents? Chaque explication est-elle adaptée au public visé?

## Activité : Explorer les algorithmes

Ces activités vont plus loin que la leçon principale, afin de voir d'autres aspects de nos vies touchés par l'IA. Assurez-vous de bien connaître les politiques de votre conseil ou de votre division concernant l'utilisation de plateformes commerciales en ligne dans votre classe.

**1. Comment entraîner l'algorithme de TikTok :** TikTok, deuxième plateforme en ligne la plus populaire chez les jeunes Canadiens selon des études de HabiloMédias, se repose beaucoup sur son fil d'actualité « Pour Toi », trié par un algorithme.

### Demandez aux élèves :

- Si vous utilisez TikTok, aimez-vous ce que votre fil d'actualité Pour Toi vous montre?
- Savez-vous comment est « entraîné » le fil d'actualité Pour Toi?
- Voici comment l'entraîner de manière intentionnelle :
  - Si vous aimez une vidéo et aimeriez voir plus de vidéos similaires, regardez-la jusqu'à la fin.
  - Si vous voulez voir plus de vidéos sur un sujet particulier, recherchez ce sujet et regardez les vidéos correspondantes.
  - Si vous n'aimez pas une vidéo et que vous aimeriez ne plus en voir de similaire, appuyez sur la vidéo pendant quelques secondes avec votre doigt et cliquez sur « Pas intéressé. »
  - Si vous voulez que l'algorithme oublie tout ce qu'il sait sur vous pour recommencer, rendez-vous sur Profil -> Menu -> Paramètres et confidentialité -> Préférences de contenu et cliquez sur Actualiser ton fil d'actualité.
- Pensez-vous que l'algorithme fonctionnait ainsi?
- Allez-vous intentionnellement encourager ou décourager l'algorithme de vous montrer certaines vidéos?
- Si vous avez décidé d'actualiser votre fil d'actualité, que ferez-vous de différent maintenant que vous savez comment entraîner l'algorithme?

**2. Voir les préjugés :** La manière dont les algorithmes reproduisent (et parfois amplifient) les préjugés sera peut-être plus facile à comprendre si elle présentée au format visuel. L'outil [Stable Diffusion Bias Explorer](#) vous permet de comparer les images créées par différents générateurs d'images utilisant l'IA et constater la manière dont certaines caractéristiques, comme les adjectifs et les professions, reflètent les préjugés présents dans leur entraînement.

# Comprendre les outils d'IA.

## Notes de présentation

Utilisez cette fiche pour prendre des notes sur les présentations des groupes (y compris la présentation de votre propre groupe). Vous aurez besoin de ces dernières pour la suite de la leçon, posez donc des questions si vous avez manqué quelque chose ou n'avez pas compris autre chose.

## Les algorithmes peuvent renforcer les préjugés du monde réel :

Notes

---

---

---

---

---

---

---

---

---

---

## Nous ne savons pas toujours comment les IA prennent des décisions :

Notes

---

---

---

---

---

---

---

---

---

---

**L'IA vous donne ce qu'elle pense que vous voulez, pas ce qui est bon pour vous**

Notes

---

---

---

---

---

---

---

---

---

---

**L'IA prend des décisions selon ce qu'elle pense savoir sur vous :**

Notes

---

---

---

---

---

---

---

---

---

---

# Comprendre les outils d'IA..

## Sondage :

Répondez aux questions suivantes par « oui », « non » ou « peut-être ».

Pensez-vous que les outils de clavardage intelligent...

- Ont des émotions? \_\_\_\_\_
- Peuvent ressentir la douleur ou le rejet? \_\_\_\_\_
- Peuvent choisir de faire ou ne pas faire certaines choses? \_\_\_\_\_
- Sont intelligents? \_\_\_\_\_
- Vous donne toujours la bonne information? \_\_\_\_\_
- Savent s'ils vous donnent la bonne information? \_\_\_\_\_
- Peuvent penser au bien ou au mal de certaines actions? \_\_\_\_\_

**Mais alors, qu'est-qu'un outil d'IA, vraiment?** Inscrivez la réponse qui va paraître être la bonne :

---

---

---

---

**Maintenant que nous avons abordé le sujet en classe, changeriez-vous votre définition?** Si oui, écrivez votre nouvelle réponse ici :

---

---

---

---

## Approfondir

**Point no1 à considérer : Les algorithmes sont-ils toujours justes?**

Les greffes de reins peuvent offrir quelques années de vie en plus. Il y a toujours plus de personnes en attente de greffe que de reins disponibles, et quelqu'un doit décider à qui attribuer un rein. Il est important d'aborder la décision de manière juste, alors des médecins ont, en 2007, inventé un algorithme pour décider :

**PG = EV, c'est-à-dire :**

- « PG » signifie « priorité pour une greffe » : lorsqu'une personne est en attente de greffe. Plus la PG sera importante, plus la personne sera susceptible d'obtenir un rein.
- « EV » signifie « espérance de vie » : le nombre d'années qu'une personne est susceptible de vivre après la greffe de rein.

Donc PG = EV signifie que vous avez plus de chance d'obtenir un rein si ce rein vous offre 20 ans de vie en plus plutôt que 10 ans de vie en plus.

**Est-ce que cela vous paraît raisonnable?**

---

**Est-ce que cela paraît juste? Pourquoi?**

---

---

**Point no2 à considérer : En quoi l'IA est différente des simples algorithmes?**

Cela signifie qu'on lui donne un objectif (comme décider de qui obtiendra un rein), puis un ensemble de données (comme une liste de personnes qui ont reçu un rein par le passé et le nombre d'années vécues après), et qu'on lui demande de chercher des tendances qu'elle utilisera pour prendre des décisions. Nous ne savons pas toujours pourquoi un outil d'IA prend une décision particulière. Même les personnes qui créent ces outils peuvent ne pas savoir pourquoi!

**Comment cela rend-il plus difficile de s'assurer que les IA prennent des décisions justes?**

---

---

**Point no3 à considérer : L'IA prend-elle toujours les bonnes décisions?**

L'IA se différencie également des simples algorithmes, car elle peut se changer elle-même.. Imaginez qu'un outil d'IA décide du lieu où vous allez manger, et qu'il a été optimisé pour vous proposer le repas que vous préférez. Cela lui serait possible en observant la quantité de nourriture que vous avez mangé des différents aliments qu'il vous a proposés, et en s'ajustant par rapport à cela : si vous n'avez pas mangé tous vos brocolis sur une journée, il vous en proposera moins le lendemain, et si vous avez mangé toute la crème glacée en une journée, il vous en proposera plus le lendemain.

**Pensez-vous que cela serait bon pour vous? Pourquoi?**

---

---

**Point no4 à considérer : D'où proviennent les données utilisées par l'IA pour prendre des décisions?**

Imaginez-vous à la conception d'un algorithme pour une compagnie d'assurance destiné à déterminer la démographie la plus susceptible d'avoir un accident. (Plus vous êtes susceptible d'avoir un accident, plus vous paierez cher pour votre assurance.)

Tout comme l'algorithme des greffes de reins, les compagnies d'assurance ont déjà des prix fixés selon les données des différentes démographies : les hommes paient en moyenne plus que les femmes pour une assurance auto, car ils ont plus d'accidents de la route.

Qu'en serait-il si votre IA savait quelles vidéos étaient visionnées en ligne par la population, et quels achats elle avait faits en ligne, ou ce qu'elle avait recherché sur Google, et pouvait lier toutes ces données avec son historique d'accidents?

**Quels types de tendances pensez-vous que l'IA trouverait? En quoi cela pourrait-il être bénéfique ou défavorable pour les gens?**

---

---

**Quelle serait l'exactitude de ces données? Qu'est-ce qui pourrait les rendre moins précises?**

---

---

**Pensez-vous qu'il serait juste pour une IA d'utiliser ces données pour prendre des décisions sur la probabilité d'avoir un accident? Pourquoi?**

---

---

## Point no5 à considérer : En quoi les clavardages intelligents sont-ils similaires à d'autres algorithmes d'apprentissage-machine?

Les clavardages intelligents sont des sortes d'IA appelées « grand modèle de langage ».

### Qu'est-ce que ça veut dire? Prenons les mots un par un :

- **Modèle** – Comme pour les autres algorithmes d'apprentissage-machine, la plupart des capacités des clavardages intelligents ne sont pas programmées, mais viennent de leur entraînement à partir de nombreux textes. Ils y trouvent des tendances pour créer un modèle de fonctionnement de langage.
- **Langage** – Contrairement aux algorithmes plus anciens, comme les outils de saisie semi-automatique, ils peuvent lire et écrire des phrases, des paragraphes et même des articles complets. Ils y parviennent en observant à quel point certains mots sont similaires ou différents dans des contextes (ou « dimensions ») variés.
  - Par exemple, si nous prenons seulement deux dimensions, la rondeur et la rougeur, une pomme et un camion de pompier pourraient ne pas se ressembler au niveau de la rondeur, mais se retrouver au niveau de la rougeur, alors qu'une balle de baseball se rapprocherait d'une pomme au niveau de la rondeur, mais s'en éloignerait au niveau de la rougeur.
  - Les clavardages intelligents peuvent faire des suppositions par « l'observation » des différentes dimensions : s'ils commençaient par « roi » et qu'ils prenaient la dimension de genre, ils verraient « reine », alors que s'ils prenaient la dimension de l'âge, ils verraient « prince », et s'ils prenaient les deux dimensions, ils verraient « princesse ».

– Cela permet à l'IA de faire de meilleures suppositions concernant les mots qui devraient se suivre selon les différentes parties d'une phrase ou d'un paragraphe. Par exemple, si vous écrivez « Frida a bu du chocolat », un simple algorithme, comme les algorithmes de saisie semi-automatique, pourrait toujours suggérer que le mot précédant « chocolat », devrait être « pépite » car c'est ce qui précède le plus ce mot en moyenne.

Mais si vous demandez à un clavardage intelligent « quel genre de boisson a bu Frida? », il pourrait repérer le mot « boire » et calculer que le mot le plus proche de « chocolat » de la dimension liquide était « lait ».

- **Grand** – ChatGPT a été entraîné à partir d'un ensemble de données comprenant environ 500 milliards de mots. Chaque mot possède en moyenne 96 dimensions, et l'IA entreprend plus de 900 opérations à chaque fois qu'elle devine un nouveau mot. Les différents clavardages intelligents ont été entraînés selon différentes données, mais la majorité a été entraînée à l'aide de ce que l'on appelle le « web ouvert » : des sites comme Wikipédia ou Reddit, des publications publiques sur les médias sociaux, des sites Internet ouverts au public comme les médias d'information et les blogues, ainsi que les textes numériques comme les livres accessibles au public sur le site du Projet Gutenberg. La plupart de ces textes ont été « moissonnés » – c'est-à-dire lus sans permission expresse du créateur et utilisés dans l'ensemble de données.



## Conseils pour utiliser les clavardages intelligents de manière sûre et éthique

Lisez la liste ci-dessous des conseils à retenir lorsque vous interagissez avec un clavardage intelligent.

À combien de conseils avez-vous pensé?

Avez-vous pensé à des conseils qui ne figurent pas dans la liste?

### Éthique

- Ne jamais s'attribuer le travail d'une IA.
- Avant d'utiliser un clavardage intelligent pour vos devoirs, parlez-en à votre professeur.
- Si vous utilisez une IA pour vos devoirs, vous devez montrer clairement où et comment vous l'avez utilisée.
- Ne vous appuyez pas sur l'information fournie par un clavardage intelligent sans avoir fait vos propres recherches; sans votre validation, vous ne pouvez pas savoir à quel point elle est exacte.

### Sécurité et vie privée

- Faites attention à ce que vous dites ou écrivez lorsque vous utilisez des clavardages intelligents. Rappelez-vous que l'entreprise peut enregistrer et lire tous vos messages et l'information que vous fournissez peut ultérieurement modifier la manière d'agir de l'IA. De plus, les choses que vous dites maintenant peuvent tout influencer, des résultats de recherche aux offres d'emploi que vous trouverez dans dix ans.
- Découvrez combien de données l'IA collecte sur vous et ce que vous pouvez y faire. Rendez-vous sur <https://tosdr.org/> pour lire un résumé de sa politique en matière de protection de la vie privée. Si elle n'apparaît pas, consultez la Politique en matière de protection de la vie privée et cherchez un titre comme « Comment nous utilisons vos renseignements personnels ». Cherchez ensuite un titre comme « Vos choix » ou « Vos droits » pour consulter vos choix (si vous en avez).

- Limitez le nombre de données que l'IA a sur vous. Sur un iPhone ou iPad, paramétrez à « Demander à l'app de ne pas suivre mes activités ». Vous pouvez également utiliser des applications comme Blokada ou DuckDuckGo et des extensions de navigateurs comme Privacy Badger ou Ghostery pour empêcher les applications et sites Internet de collecter des données.
- Souvenez-vous que lorsque vous parlez à un clavardage intelligent, vous pouvez avoir l'impression de parler à un humain, mais les clavardages intelligents ne sont que des simulateurs qui devinent ce que vous voulez entendre. Traitez-les comme des jouets, pas comme des amis.
- Les clavardages intelligents peuvent être très amusants, mais comme tout ce que vous faites en ligne, vous devez vous assurer que le temps que vous leur consacrez ne vous empêche pas d'en passer avec vos amis, votre famille, à l'école, au sport, et n'empiète pas sur votre sommeil.

# Comprendre les outils d'IA - Règles pour des clavardages intelligents sûrs et éthiques.

Pour ce devoir, vous définirez des règles et directives qui pourraient être utilisées par les personnes qui créent ou utilisent les clavardages intelligents. Ces règles devraient s'attaquer aux problèmes mentionnés dans les présentations de groupe.

Passez en revue les notes que vous avez prises pendant les présentations et pensez à **deux** règles destinées à répondre aux principes discutés :

1. Les clavardages intelligents peuvent renforcer les préjugés du monde réel;
2. Nous ne savons pas toujours comment les clavardages intelligents prennent des décisions;
3. Les clavardages intelligents vous donnent ce qu'ils pensent que vous voulez, pas ce qui est bon pour vous; et
4. Les clavardages intelligents prennent des décisions selon ce qu'ils pensent savoir sur vous.

Cela fait un total de **huit** règles et directives.

Expliquez l'importance de chacune de ces règles.

Ces règles et explications doivent montrer vos connaissances et votre compréhension acquises grâce au document Comprendre les outils d'IA et la discussion de classe, ainsi que les présentations de groupe.

Principe	Règle	En quoi cette règle est-elle importante?
1. Les clavardages intelligents peuvent renforcer les préjugés du monde réel.	1.	
	2.	
2. Nous ne savons pas toujours comment les clavardages intelligents prennent des décisions.	3.	
	4.	
3. Les clavardages intelligents vous donnent ce qu'ils pensent que vous voulez, pas ce qui est bon pour vous.	5.	
	6.	
4. Les clavardages intelligents prennent des décisions selon ce qu'ils pensent savoir de vous.	7.	
	8.	