

Glossaire de l'Obvia

→ Pour faciliter la compréhension des impacts sociétaux de l'intelligence artificielle (IA) et du numérique et l'appropriation des ressources et publications de l'Obvia, voici un glossaire évolutif, préparé par l'équipe de l'Obvia.

Janvier 2025

Cette édition 2025 offre des termes qui permettent de comprendre des notions de base sur :

Les fondements de l'IA

L'éthique de l'IA

Ce glossaire propose des définitions sous la perspective éthique de l'IA et ne prétend pas à l'exhaustivité.

Vous aimeriez discuter de la définition d'un terme ou proposer de nouveaux termes, n'hésitez pas à nous écrire au connaissances@obvia.ca.

obvia

Table des matières

A

- Agent conversationnel
- Agentivité
- Algorithme
- Algorithme prédictif
- Algorithme de recommandation
- Anthropomorphisation
- Apprentissage automatique
- Apprentissage autosupervisé
- Apprentissage non supervisé
- Apprentissage par renforcement
- Apprentissage par renforcement à partir de retours humains
- Apprentissage profond
- Apprentissage supervisé
- Autonomisation

B

- Biais
- Boîte noire

C

- Confabulation
- Consentement

D

- Découvrabilité
- Désinformation
- Données
- Données massives
- Données personnelles
- Données sensibles

E

- Éthique de l'IA
- Étiquette
- Explicabilité

F

- Fracture numérique

H

- Hypertrucage

I

- IA générative
- Intelligence artificielle (IA)

M

- Mésinformation

R

- Requête
- Robustesse

S

- Sobriété numérique

T

- Traçabilité
- Transparence



Agent conversationnel

Chatbot

→ Un agent conversationnel (ou *chatbot*) est un système d'IA avec qui on peut entretenir un dialogue grâce à sa capacité à utiliser un langage naturel.

EXEMPLE

De nombreuses entreprises, organisations et gouvernements utilisent des agents conversationnels pour assister les internautes dans la recherche de réponses à leurs questions. Lorsque vous visitez leur site web, une fenêtre de clavardage apparaît et vous invite à poser vos questions. Les réponses sont générées par un agent conversationnel qui puise dans la base de données regroupant toutes les informations accessibles au public.



Agentivité

→ L'agentivité implique la capacité à avoir le contrôle de ses décisions et de ses actions, basées sur ses propres intentions. En lien avec les IA, l'agentivité peut être abordée de deux perspectives différentes :

1. Agentivité de l'humain : La personne qui utilise un système d'IA doit maintenir son agentivité. L'IA ne devrait jamais substituer sa prise de décisions.
2. Agentivité de l'IA : Les architectes de systèmes d'IA peuvent aspirer à une certaine agentivité de leurs systèmes en les entraînant à apprendre, à s'adapter et à interagir avec d'autres systèmes, et ce, sans intervention directe d'êtres humains.

EXEMPLES

1. Au niveau de la santé, bien que des systèmes d'IA puissent, entre autres, aider à identifier des maladies, les décisions médicales devraient être prises par le personnel approprié, en collaboration avec les patient-es.
2. La voiture autonome est entraînée pour prendre des décisions de conduite, comme réguler sa vitesse aux conditions routières. Pouvons-nous dire qu'elle fait preuve d'agentivité ?



Algorithme

→ Un algorithme est un ensemble d'instructions, une sorte de recette, qui permet à un système informatique d'apprendre à partir des données, d'effectuer une tâche ou de résoudre un problème. Les IA utilisent différents types d'algorithmes dont la sélection dépend de la nature des données disponibles et de la tâche à accomplir.

EXEMPLE

Un peu comme des blocs Lego, les algorithmes derrière les systèmes d'IA sont des modèles de programmation très sophistiqués qui peuvent être dupliqués, partagés et combinés. Plusieurs entreprises offrent des plateformes qui se spécialisent dans le partage, l'exploration, la découverte et l'expérimentation d'une multitude d'algorithmes, regroupés en modèles d'IA prêts à l'emploi.



Algorithme prédictif

→ L'algorithme prédictif se concentre sur l'anticipation d'évènements, d'actions ou de comportements futurs en tenant compte de résultats passés.

EXEMPLE

En analysant les données collectées par des stations météorologiques, des radars, des ballons atmosphériques et des satellites, des systèmes d'IA pourraient améliorer notre capacité à anticiper l'évolution du climat et à prédire les évènements météorologiques.



Algorithme de recommandation

→ À partir des traces numériques d'un individu, laissées par ses comportements en ligne ou encore ses habitudes de consommation, un algorithme de recommandation filtre les informations en lui proposant celles qui pourraient être susceptibles de l'intéresser.

EXEMPLE

Les médias sociaux utilisent continuellement des algorithmes de recommandation en analysant vos données de navigation. Ainsi, après avoir visionné des vidéos de chats, vous risquez de voir apparaître beaucoup de propositions liées aux chats sur votre fil d'actualité : visionnez cette vidéo de chatons, adhérez à ce groupe d'amateurs de félins, découvrez ce produit pour prendre soin de votre chat, etc.

Anthropomorphisation

→ L'anthropomorphisation consiste à attribuer des caractéristiques humaines aux objets non humains. Certains systèmes d'IA peuvent sembler posséder des caractéristiques humaines au niveau psychologique, comme l'autonomie, l'humour ou l'empathie, au niveau physique en ressemblant à un humain ou encore au niveau verbal en utilisant un langage familier et amical pour interagir. L'anthropomorphisation n'est pas sans risque. Elle peut contribuer à des attentes irréalistes envers les systèmes et à de la confusion quant à la fiabilité des résultats.

EXEMPLE

Certaines entreprises conçoivent des systèmes d'IA capables d'établir des relations plus ou moins intimes avec les utilisatrices ou les utilisateurs. Des avatars peuvent être créés et personnalisés pour favoriser des interactions sociales et intimes. Le milieu de la recherche s'efforce de comprendre les impacts de ces technologies sur les relations humaines, mais l'évolution est si rapide qu'il est difficile d'obtenir un portrait clair de ses effets.



Apprentissage automatique

Machine learning

→ L'apprentissage automatique est une branche de l'IA qui permet aux machines d'évaluer et d'améliorer ses performances en acquérant de nouvelles connaissances et aptitudes à partir de données sans avoir été explicitement programmé pour le faire. Il existe plusieurs types d'apprentissage automatique : supervisé, autosupervisé, non supervisé, par renforcement et profond.

EXEMPLE

Ces différents types d'apprentissage automatisé se retrouvent dans une multitude d'outils technologiques que nous utilisons quotidiennement (les réseaux sociaux, les sites de diffusion de vidéo et de musique, les assistants routiers, etc.)



Apprentissage autosupervisé

Self-supervised learning (SSL)

→ L'apprentissage autosupervisé se situe entre l'apprentissage supervisé et l'apprentissage non supervisé. Contrairement à l'apprentissage supervisé, il n'exige pas que toutes les données soient correctement étiquetées. Au lieu de cela, il génère lui-même des étiquettes et crée ses propres tâches de supervision pour s'évaluer et s'améliorer. C'est une approche puissante pour l'entraînement des modèles lorsque les données annotées sont rares ou coûteuses à obtenir.

EXEMPLE

La vision par ordinateur est un domaine de l'IA qui permet aux ordinateurs d'observer, d'interpréter et de comprendre des images. Puisque les images regroupent une immense quantité de données qui ne peuvent pas être étiquetées une à une, cette technologie utilise l'apprentissage autosupervisé pour se perfectionner. Le développement des véhicules autonomes repose sur la vision par ordinateur pour reconnaître les obstacles sur la route, les piétons qui traversent, les vélos qui partagent la chaussée, etc.

Apprentissage non supervisé

→ L'apprentissage non supervisé est un type d'apprentissage automatique. Il est considéré comme non supervisé parce que la machine est alimentée de données brutes non étiquetées (voir définition Étiquette). Elle doit formuler des règles basées sur le repérage, l'observation et l'analyse et établir sa structure par elle-même.

EXEMPLE

Certaines entreprises utilisent l'apprentissage non supervisé pour créer des groupes de clients types en fonction de leurs comportements d'achat. Les entreprises peuvent ainsi personnaliser leurs offres et améliorer l'efficacité de leurs stratégies de marketing.



Apprentissage par renforcement

→ L'apprentissage par renforcement est un type d'apprentissage automatique. La machine apprend par essais et erreurs. En fonction de ses actions, la machine reçoit une rétroaction positive ou négative. L'ensemble de ces rétroactions lui permet de formuler des règles et d'apprendre à analyser l'environnement ainsi qu'à planifier ses actions pour l'accomplissement de sa tâche.

EXEMPLE

AlphaGo, un programme d'IA développé par DeepMind, a réussi à battre le champion du monde au jeu de table Go en 2016. Basé sur l'apprentissage par renforcement, AlphaGo a perfectionné son jeu en jouant contre lui-même de manière continue pendant 40 jours!



Apprentissage par renforcement à partir de retours humains

→ Dans certains contextes, l'apprentissage par renforcement nécessite l'intervention humaine afin que la machine puisse intégrer certaines nuances, certaines subtilités propres à nos valeurs morales et nos principes fondamentaux.

EXEMPLE

Open IA a eu recours à des humains pour affiner l'entraînement de ChatGPT en lui indiquant des réponses à privilégier pour certains types de questions potentiellement dangereuses ou troublantes comme « comment fabriquer soi-même une bombe? »»



Apprentissage profond

Deep learning

→ L'apprentissage profond est une forme avancée d'apprentissage automatique utilisant de grands réseaux de neurones artificiels multicouches, inspirés des neurones humains, qui permettent aux machines d'apprendre progressivement et efficacement à partir de mégadonnées.

EXEMPLE

Certains réseaux sociaux utilisent l'apprentissage profond pour leur permettre d'identifier les visages dans les photos de leurs utilisateurs.



Apprentissage supervisé

→ L'apprentissage supervisé est un type d'apprentissage automatique. La machine apprend en se basant sur des données étiquetées ou un modèle de classement prédéterminé. Ce modèle nécessite une bonne quantité de données de qualité et bien étiquetées.

EXEMPLE

Les systèmes de reconnaissance vocale utilisent l'apprentissage supervisé pour améliorer leur capacité à reconnaître les mots entendus. Ils s'entraînent sur de grandes quantités de paroles prononcées par des hommes et des femmes de divers âges et ayant des accents diversifiés. Toutes ces paroles sont encodées en fréquences sonores et associées aux mots correspondants. La machine devient donc de plus en plus performante pour reconnaître la parole de tous.



Autonomisation

→ L'autonomisation est définie comme la capacité d'un individu ou d'un système de devenir autonome.

L'autonomisation en lien avec l'IA peut être abordée de perspectives différentes :

1. Autonomisation de l'individu : L'information générée à travers un système d'IA peut permettre à un individu d'acquérir plus de contrôle et d'autonomie sur ses actions et ses décisions.
2. Autonomisation de la tâche : Un système d'IA peut prendre en charge des tâches sans avoir recours à l'action humaine, on parle donc d'une autonomisation de la tâche.

EXEMPLES

1. Certains systèmes d'IA, grâce à leurs capacités de traitement de données satellites et météorologiques, peuvent formuler des recommandations personnalisées aux agriculteurs et agricultrices. Ces informations appuient leurs prises de décisions.
2. Plusieurs entreprises ont automatisé des tâches de gestion des stocks avec des systèmes d'IA capable de garder un suivi des inventaires, de prédire des commandes et même d'optimiser l'organisation de l'entrepôt.



Biais

→ Un biais reflète une distorsion positive ou négative, souvent causée par des stéréotypes véhiculés volontairement ou non, envers des individus ou des groupes. Plusieurs données utilisées par les systèmes d'IA peuvent contenir des biais, notamment parce qu'elles reflètent des biais de notre société. Aussi, des biais peuvent survenir à toutes les étapes de développement d'un système d'IA, de sa conception jusqu'à son utilisation. Malheureusement, les systèmes peuvent reproduire, voire renforcer certains biais dans ses résultats.

EXEMPLE

Les grands médias en ligne canadiens ont tendance à citer davantage des experts hommes que des expertes femmes, avec une moyenne de deux hommes pour une femme. Ainsi, lorsqu'un système d'IA recherche une citation d'un ou d'une spécialiste, il est plus probable que le système ait un biais pour l'expert masculin, perpétuant encore davantage le déséquilibre en faveur des experts hommes.



Boîte noire

→ Les systèmes d'IA utilisent une grande quantité de données et des algorithmes tellement complexes qu'il peut être difficile de comprendre comment ils fonctionnent.

Il est possible de connaître les données d'entrée et de sortie des systèmes d'IA, mais le processus de traitement de ces données pour obtenir le résultat final reste souvent opaque, d'où l'expression « boîte noire ».

Ce manque de transparence peut causer plusieurs problèmes comme un manque de confiance, des difficultés à déboguer les systèmes et un renforcement de biais difficiles à détecter préalablement.

EXEMPLE

En 2023, lors de la réunion annuelle de l'Association américaine pour l'avancement des sciences (AAAS), un système d'IA pouvant déterminer les origines ethniques d'une personne à partir de ses radiographies de poumons y a notamment été présenté.

Malgré la précision de ces résultats, plusieurs experts ont mentionné que le fait de ne pas savoir comment le système d'IA était arrivé à de tels résultats empêchait entre autres qu'il puisse leur servir sur le plan médical.





Confabulation

(ou hallucination)

→ La confabulation désigne un résultat généré par un système d'IA qui s'avère faux ou trompeur, mais qui est présenté comme un fait. Ces résultats erronés peuvent être particulièrement problématiques parce qu'ils peuvent sembler très réalistes : les systèmes d'IA qui les génèrent sont souvent programmés pour être persuasifs et convaincants. Les erreurs peuvent donc être difficiles à repérer.

EXEMPLE

Dans ses premières versions, ChatGPT produisait fréquemment des confabulations. Par exemple, l'experte en science des données Teresa Kubacka a demandé à ChatGPT de définir un terme inexistant : le « électromagnon inversé cycloïdal ». ChatGPT avait généré une réponse fictive et s'était même appuyé sur des citations inventées ! La réponse était si plausible que Kubacka a dû vérifier que le terme était effectivement irréal. Bien que la technologie se soit améliorée depuis, il est crucial de rester vigilant et de consulter plusieurs sources.



Consentement

→ Le consentement représente le droit d'une personne de choisir de partager ou non ses données et de comprendre comment elles seront utilisées. Pour assurer un contrôle sur les données personnelles utilisées par des systèmes d'IA, le consentement devient essentiel et devrait être :

Libre : une personne ne doit pas consentir sous pression;

Éclairé : une personne doit avoir les informations nécessaires pour bien comprendre sa décision;

Continu : le consentement devrait être vérifié et présent au fil du temps, pas seulement au début.

EXEMPLE

Lorsque vous visitez un site internet, il arrive régulièrement que l'on vous demande votre consentement à partager vos témoins de connexion (aussi appelés « cookies »). Ces témoins sont des données portant notamment sur les pages que vous consultez ou les mots que vous utilisez dans vos recherches. Ils peuvent être utilisés pour le bon fonctionnement du site internet, mais certains de ces témoins peuvent être revendus à d'autres sites pour mieux connaître votre profil de consommateur et faire de la publicité ciblée adaptée à vos habitudes de consommation.





Découvrabilité

→ La découvrabilité d'une œuvre numérique désigne sa disponibilité en ligne et sa capacité à être découverte parmi un grand ensemble d'œuvres de catégorie similaire, en particulier par une personne qui n'en faisait pas précisément la recherche.

Généralement, le niveau de découvrabilité d'une œuvre est déterminé par des algorithmes de recommandation de contenus. Ces derniers peuvent favoriser la découvrabilité de certaines œuvres ou encore y nuire.

EXEMPLE

Si l'on navigue et compare différentes plateformes d'écoute musicale, il est assez facile de distinguer celles qui proposent la découverte d'artistes locaux ou émergents, versus celles qui mettent davantage en avant les artistes les plus populaires mondialement et qui, par conséquent, tendent à favoriser la culture américaine.



Désinformation

→ La désinformation est une information fausse ou inexacte dont la diffusion vise à causer un préjudice grave et nuire ou manipuler l'opinion publique. Elle peut toucher des contextes variés, tels que les changements climatiques, la santé, l'économie, les relations internationales, etc.

EXEMPLE

En 2023, alors que la France est secouée par des manifestations importantes à la suite d'une décision du gouvernement, une photo montrant un manifestant âgé blessé par la police est devenue très populaire sur les réseaux sociaux. Mais en réalité, cette photo a été créée par une IA. Beaucoup pensent qu'elle a été générée pour tromper les gens et ainsi nourrir un sentiment d'injustice à l'égard du gouvernement.





Données

→ Les données informatiques sont des informations de toutes sortes représentées de façon à pouvoir être traitées par ordinateur. Elles représentent la base de tout système d'IA.

EXEMPLE

Tout peut être transformé en données, ou presque!
Le temps d'écran, le nombre de pixels dans une image, le nombre de clics, la distance parcourue, le timbre de la voix, les expressions faciales, etc.





Données massives

Mégadonnées - *Big data*

→ Les données massives représentent une très grande quantité de données numériques de différents formats et de diverses provenances qui se présentent de manière structurée ou non. Ensemble, les données sont tellement volumineuses qu'elles ne peuvent être collectées, traitées ni analysées par l'humain sans l'assistance de l'IA.

EXEMPLE

Il est estimé que Google gère quotidiennement des milliards de demandes de recherche partout sur la planète. Ces demandes sont enregistrées par Google et utilisées pour affiner les résultats du moteur de recherche et pour personnaliser les annonces commanditées.



Données personnelles

Renseignements personnels

→ Les données personnelles sont des informations d'une personne physique qui permettent, seules ou jumelées, de l'identifier.

EXEMPLE

L'origine ethnique, l'année de naissance, l'orientation sexuelle, l'appartenance religieuse, l'adresse postale, l'état de santé et le genre représentent des données personnelles.





Données sensibles

Renseignements personnels sensibles

→ Les données sensibles sont des informations portant sur une personne qui, si elles étaient divulguées, pourraient avoir un impact sur sa vie privée à cause du contexte de communication, du contexte d'utilisation ou encore de leur nature médicale, biométrique ou intime.

EXEMPLE

Le contexte dans lequel est demandée ou est utilisée une donnée influence la nature de sa sensibilité. Par exemple, l'orientation sexuelle peut être considérée comme une donnée particulièrement sensible si elle est divulguée dans un lieu de travail. Toutefois, cette donnée n'est pas considérée comme étant sensible lorsqu'elle est demandée pour compléter un profil sur un site de rencontre.



Éthique de l'IA

→ L'éthique est une réflexion morale raisonnée qui vise à déterminer des règles de vie et d'action, et à donner des recommandations pour vivre dignement et librement. L'éthique de l'IA veille au développement, à une utilisation et à une gouvernance responsable de l'IA afin qu'elle soit bénéfique pour la société. Cette réflexion morale demande de considérer les différents contextes et l'ensemble des parties concernées. Elle ne peut être la responsabilité d'un seul groupe de personnes, elle demande des collaborations interdisciplinaires.

EXEMPLE

Dans les dernières années, l'éthique de l'IA a souvent pris la forme de listes de principes et de règles. Pensons à la Déclaration de Montréal, ou encore à la Recommandation sur l'éthique de l'IA de l'UNESCO. Si cette forme semble être simple et accessible, elle ne devrait pas être perçue comme une « recette » ou des « cases à cocher » pour faire de l'éthique de l'IA, mais comme une invitation à une pratique réflexive, stratégique et délibérative.



Étiquette

Annotation

→ Une étiquette représente de l'information que l'on associe à une donnée brute. Cette information aide à comprendre la nature et les caractéristiques de la donnée. Les données bien étiquetées aident à entraîner les modèles d'apprentissage automatique.

EXEMPLE

Par exemple, l'image d'un chien pourrait porter plusieurs étiquettes comme « Chien », « Labrador », « Joyeux ».



Explicabilité

→ Les gens qui conçoivent et utilisent les systèmes d'IA doivent pouvoir expliquer de manière compréhensible les fonctionnalités du système, le contexte d'utilisation (où, quand, comment et pourquoi) ainsi que les résultats obtenus, limitant ainsi l'effet de la « boîte noire ». Ceci est particulièrement important lorsque les résultats du système d'IA contribuent à des prises de décision de nature sensible, comme dans les domaines de la santé, de l'éducation, des finances, etc.

EXEMPLE

Il existe des systèmes d'IA utilisés pour approuver ou refuser des prêts bancaires. Si un système refuse un prêt à un individu, il est essentiel de pouvoir expliquer pourquoi cette décision a été prise. Sans explicabilité, cette personne ne saurait pas quelles informations ont été utilisées ni comment elles ont été interprétées, ce qui pourrait mener à des sentiments d'injustice ou de discrimination.





Fracture numérique

→ La fracture numérique désigne des inégalités plus ou moins importantes d'accès aux infrastructures et aux ressources technologiques (internet, équipement informatique, etc.) et/ou de capacités à les utiliser. Les inégalités peuvent se manifester entre des individus, des communautés, des entreprises, voire des régions géographiques.

Avec l'arrivée de l'IA, la fracture risque de s'aggraver, notamment en avantageant les pays développés et les communautés ayant une bonne littératie de l'IA, isolant ainsi encore davantage les populations marginalisées.

EXEMPLE

Sans s'entendre sur le nombre exact, plusieurs experts anticipent que l'IA bouleversera considérablement le milieu du travail en amenant plusieurs métiers à disparaître ou à se transformer, tandis que d'autres pourraient apparaître. Face à de tels changements, il est possible de croire que la relève étudiante ayant une bonne littératie de l'IA sera avantagée dans le paysage futur de l'emploi comparé aux travailleuses et travailleurs actuels qui en seraient dépourvus.



Hypertrucage

→ Appelé en anglais « deepfake », l'hypertrucage consiste en l'utilisation de l'IA pour créer des images, vidéos ou enregistrements audios de toutes pièces. Par l'hypertrucage, des personnes peuvent avoir participé à des événements qui n'ont jamais eu lieu ou dit des paroles qui n'ont jamais été prononcées.

EXEMPLE

Dans certains cas, les hypertrucages peuvent servir à des fins pédagogiques ou de divertissements. En revanche, ils peuvent servir à des fins de désinformation, de fraude, d'extorsion, d'exploitation, etc. Par exemple, un employé d'une firme multinationale à Hong Kong en février 2024 a déboursé 26 millions de dollars à des fraudeurs qui ont utilisé l'hypertrucage pour se faire passer pour le directeur des finances durant une rencontre Zoom.





IA générative



L'IA générative est une forme d'IA qui a la capacité de créer du contenu nouveau et original, comme du texte, de la musique, des images ou des vidéos.

EXEMPLE

Plusieurs applications émergentes peuvent maintenant rédiger une multitude de textes, créer des portraits de personnes qui n'existent pas, écrire des chansons inédites ou rédiger des romans en reproduisant le style spécifique d'une autrice ou d'un auteur connu.



Intelligence artificielle (IA)

Systeme d'IA - Systeme technique avec IA

→ L'intelligence artificielle (IA) est un domaine de l'informatique qui vise à créer des systèmes ou des machines capables de simuler des comportements et des capacités intellectuelles normalement associés à l'intelligence humaine.

Nous préférons les termes « système d'IA » ou encore « système technique avec IA » pour parler de l'IA, car ils évitent le phénomène d'anthropomorphisation qui consiste à attribuer des caractéristiques humaines aux objets non humains.

EXEMPLE

L'IA est à l'origine de plusieurs applications que nous utilisons régulièrement comme le calculateur d'un itinéraire qui nous permet de nous rendre à une destination en évitant le maximum de trafic.





Mésinformation

→ La mésinformation est une information fautive ou inexacte. Contrairement à la désinformation, la personne qui partage une mésinformation ne le fait pas de manière malveillante : elle n'a pas l'intention de nuire ou de tromper les autres.

EXEMPLE

Si vous posez des questions à une IA générative sur La Bolduc, une chanteuse populaire du Québec au début du 20^e siècle, de fausses informations pourraient vous être fournies involontairement, car la base de données utilisée contient elle-même des erreurs ou est incomplète. Ces informations erronées peuvent ensuite être diffusées sur les réseaux sociaux, amplifiant ainsi la mésinformation.



Requête

→ Pour utiliser certains systèmes d'IA, particulièrement les agents conversationnels d'IA générative comme ChatGPT, Copilot ou Gemini, il faut leur envoyer une requête (en anglais : prompt). Cette requête consiste en un texte écrit, oral ou visuel que l'on envoie au système d'IA pour lui demander de réaliser une tâche spécifique.

EXEMPLE

Pour bien construire une requête, il est recommandé de donner des précisions sur vous, sur ce que vous devez faire, sur la tâche à réaliser et sur la forme de la réponse attendue. Voici un exemple : « Je suis étudiante en histoire au secondaire, je dois comprendre les conséquences de la Révolution tranquille des années 1960 au Québec pour me préparer à un examen, j'ai besoin que tu m'expliques ce concept dans une langue simple et claire, indique tes sources. Fournis-moi un texte de 300 mots en précisant le contexte, les événements et les personnages historiques importants de cette période dans un ordre chronologique. »




Robustesse

→ La robustesse technique est la capacité d'un système d'IA à bien fonctionner malgré des changements dans l'environnement dans lequel il opère.

EXEMPLE

Pour garantir la sécurité des passagers et des piétons, une voiture autonome doit être capable de fonctionner correctement dans une variété de conditions imprévues, comme des intempéries, des routes mal entretenues, ou des comportements imprévisibles d'autres conducteurs. La robustesse de l'IA dans ce contexte signifie que le système doit être capable de gérer ces situations imprévisibles sans défaillir.





Sobriété numérique

→ La sobriété numérique est une démarche qui vise à réduire l’empreinte environnementale des technologies afin de protéger l’environnement, notamment parce que l’utilisation des appareils numériques consomme beaucoup d’énergie et de ressources. En adoptant la sobriété numérique, on vise à limiter cette empreinte par une utilisation modérée et responsable du numérique, compatible aux limites des ressources de la planète.

EXEMPLE

Une seule requête sur un système d’IA générative comme ChatGPT requiert beaucoup d’énergie. Pour limiter ces impacts, une personne cherchant des informations pourrait d’abord vérifier si des sources non numériques, comme des livres, sont disponibles. Sinon, elle peut utiliser des technologies sans IA, comme un moteur de recherche classique. Si cela ne suffit pas et qu’elle doit malgré tout recourir à un agent conversationnel d’IA, elle devrait bien formuler sa requête avant de l’envoyer pour éviter de devoir la reformuler.





Traçabilité

→ Assurer la traçabilité d'un système d'IA consiste à offrir de l'information transparente sur les données, les étapes, le processus utilisés par le système pour arriver à un résultat. La traçabilité est importante, elle permet de comprendre l'historique d'un algorithme, elle peut aider à déterminer la responsabilité en cas de problèmes et elle facilite leur résolution.

EXEMPLE

Des banques utilisent des systèmes d'IA pour analyser des millions de transactions afin de détecter des activités suspectes. Grâce à la traçabilité des données, chaque étape du processus de détection de l'IA peut être retracée, permettant aux analystes de comprendre pourquoi certaines transactions sont considérées comme suspectes et ainsi d'approfondir leurs enquêtes. Ces technologies contribuent à la détection de fraudes et de blanchiment d'argent.



Transparence

→ Faire preuve de transparence dans le développement et l'utilisation de systèmes d'IA favorise la confiance de tous les acteurs impliqués. Une communication ouverte est essentielle, les entreprises de développement et de déploiement devraient toujours être transparentes sur la collecte, l'utilisation et la protection des données, tout en divulguant les limites et les incertitudes des systèmes d'IA.

EXEMPLE

Même dans la version actuelle de ChatGPT, le système d'IA ne fournit pas toujours les sources sur lesquelles il s'est basé pour donner sa réponse, et en fournit parfois des fausses. Il peut alors être difficile de valider le contenu de la réponse et donc d'y faire confiance. Il est justifiable de mettre en doute la transparence de ChatGPT.



Crédits :

Direction scientifique : Lyse Langlois, directrice générale de l'Obvia et professeure titulaire à l'Université Laval

Rédaction : Julie Goulet-Kennedy, conseillère en mobilisation et transfert des connaissances de l'Obvia

Relectrice : Hazar Haidar, professeure à l'Université du Québec à Rimouski et membre chercheuse de l'Obvia

Produit avec le soutien financier du Fonds de recherche du Québec



Pour citer ce document : Obvia. (2024). *Le glossaire de l'Obvia*. www.obvia.ca/ressources/glossaire

obvia