Policy Horizons Canada | Horizons de politiques Canada

RAPPORT

Prospective sur l'IA

Considérations politiques

L'intelligence artificielle (IA) est une technologie qui change la donne. Le rythme des progrès s'accélère. Au fur et à mesure que de nouvelles technologies d'IA sont mises sur le marché, l'incertitude grandit quant à leurs répercussions possibles, qu'elles soient positives ou négatives. La vitesse du développement technologique pourrait dépasser la capacité des décideurs à suivre le rythme.

Ce rapport présente dix idées sur les facteurs qui pourraient façonner l'évolution de l'IA, en termes de capacités techniques, d'adoption et d'utilisation. Il s'agit du premier rapport produit par le projet interministériel d'Horizons de politiques Canada sur l'avenir de l'IA, qui complète les travaux relatifs à l'IA menés par le gouvernement du Canada.

En aidant les lecteurs à comprendre les répercussions que l'IA pourrait avoir sur la gouvernance, la société et l'économie, en mettant l'accent sur les éléments qui se situent au-delà de l'horizon, le rapport vise à aider les décideurs à réfléchir à l'avenir de l'IA.

Canada

Prospective sur l’IA

Considérations politiques

PDF : PH4-210/2025F-PDF
ISBN : 978-0-660-74946-4

AVERTISSEMENT :
Horizons de politiques Canada (Horizons de politiques) est le centre d’excellence en prospective du gouvernement du Canada. Notre mandat est de doter le gouvernement du Canada d’une perspective et d’un état d’esprit tournés vers l’avenir afin de renforcer la prise de décisions. Le contenu de ce document ne représente pas nécessairement le point de vue du gouvernement du Canada ou des agences et ministères participants.

Avant-propos

L'intelligence artificielle (IA) évolue rapidement et présente à la fois des possibilités et des défis pour le Canada. Alors que l'IA continue de progresser, il est essentiel de comprendre ses répercussions potentielles sur la gouvernance, la société et l'économie.

Horizons de politiques Canada (Horizons de politiques) se consacre à l'étude de la manière dont l'IA pourrait façonner notre avenir. En nous engageant auprès d'un large éventail de partenaires et de parties prenantes, nous visons à déterminer les principaux domaines de changement et à soutenir les responsables politiques et les décideurs dans leur navigation au sein de ce paysage dynamique.

Au nom d'Horizons de politiques, je tiens à remercier ceux qui nous ont fait don de leur temps et fait part de leurs connaissances et de leurs réflexions.

Nous espérons que ce rapport vous donnera à réfléchir et vous sera utile.

Kristel Van der Elst
Directrice générale
Horizons de politiques Canada

Introduction

Ce rapport de prospective sur l'IA complète de nombreuses réflexions sur l'avenir de l'IA (voir encadré 1) au sein du gouvernement du Canada. Il vise à aider les décideurs, impliqués dans la mise en œuvre de l'IA ou dans l'élaboration de politiques liées à l'IA, en explorant les facteurs qui pourraient façonner l'évolution de l'IA, en termes de capacités techniques, d'adoption et d'utilisation, et qui pourraient être « *au-delà de l'horizon* ». Le rapport ne fournit pas d'orientations politiques particulières et n'est pas censé prédire l'avenir. Son objectif est de soutenir la réflexion prospective et d'éclairer la prise de décision.

Dans le cadre de ce travail, Horizons a procédé à une analyse documentaire, a recherché les avancées en cours dans ce domaine, s'est entretenu avec des analystes politiques et des décideurs au sein du gouvernement, et a eu des conversations approfondies avec des experts clés en matière d'IA.

Les dix idées présentées dans ce rapport explorent les capacités futures possibles de l'IA, les risques et les possibilités à long terme, ainsi que les incertitudes liées aux hypothèses pertinentes pour les politiques. Les lecteurs peuvent chercher à comprendre les répercussions que l'IA pourrait avoir sur la gouvernance, la société et l'économie. En lisant ce rapport, les lecteurs sont invités à se poser les questions suivantes :

- Comment les progrès futurs en matière de matériel, de logiciels et d'interfaces créeront-ils de nouvelles possibilités et de nouveaux risques pour le Canada et ses alliés
- Où l'IA pourrait-elle apporter les perturbations les plus importantes et les plus inattendues à la gouvernance, à la société et aux marchés
- Quelles hypothèses concernant le développement et le déploiement de l'IA à l'avenir pourraient devoir être remises en question ou approfondies avant de servir de base à la prise de décision

Les dix idées sont synthétisées dans le tableau 1 et développées plus loin dans le document.

Définir l'IA

Il existe de nombreuses façons de définir l'intelligence artificielle, ainsi que de nombreux débats sur l'opportunité de continuer à utiliser ce terme[1][2][3]. Pour les besoins de ce travail, Horizons de politiques utilise la définition d'un système d'IA de l'Organisation de coopération et de développement économiques (OCDE) : « un système basé sur une machine qui, pour des objectifs explicites ou implicites, déduit, à partir des données qu'il reçoit, comment générer des résultats tels que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer des environnements physiques ou virtuels. Les différents systèmes d'IA varient dans leurs niveaux d'autonomie et d'adaptabilité après le déploiement[4]. »

L'IA est un système basé sur une machine qui, pour des objectifs explicites ou implicites, déduit, à partir des données qu'il reçoit, comment générer des résultats tels que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer des environnements physiques ou virtuels. Les différents systèmes d'IA varient dans leurs niveaux d'autonomie et d'adaptabilité après le déploiement.

Tableau 1 : 10 idées sur les facteurs qui façonnent l'avenir de l'IA et avec l'IA

1	**L'IA pourrait briser Internet tel que nous le connaissons actuellement** Les nouveaux outils d'intelligence artificielle pourraient ébranler le modèle économique de la publicité qui a servi de base à Internet pendant la majeure partie des 20 dernières années. Internet à l'ère de l'IA pourrait être très différent, avec plus d'autonomie et de contrôle, mais aussi moins utile et moins sûr.
2	**L'IA pourrait renforcer les acteurs non étatiques et submerger les organisations de sécurité** À l'avenir, l'IA plus accessible et plus polyvalente aura des répercussions sur la sécurité. Les acteurs non étatiques, bons ou mauvais, auront accès à des capacités traditionnellement détenues uniquement par les États. Ils pourraient être en mesure de les déployer plus rapidement que les États, ce qui maintiendrait les organisations de sécurité dans une course constante pour les rattraper.
3	**Le manque de confiance dans l'IA pourrait entraver son adoption** On ne sait pas comment évoluera la confiance dans l'IA. Les défaillances fréquentes ou non corrigées des systèmes d'IA, ou une seule défaillance importante, pourraient éroder la confiance et entraver l'adoption, mettant en péril des secteurs entiers. Les nouvelles formes de certification, de vérification et d'efforts pour réparer les préjudices pourraient encourager la confiance des utilisatrices et utilisateurs et l'adoption du système.
4	**Les biais des systèmes d'IA pourraient perdurer** La partialité est une caractéristique de la prise de décision humaine et de l'IA. Comme les données utilisées pour former l'IA sont souvent entachées de biais difficiles à corriger, le recours croissant à l'IA dans les systèmes de prise de décision pourrait répandre les biais et entraîner des dommages importants. Les préjugés ne peuvent jamais être éliminés, en partie à cause des perspectives contradictoires sur l'équité.

5	**L’utilisation de l’IA pour prédire le comportement humain pourrait s’avérer inefficace** Si l’IA fait parfois des prédictions impressionnantes sur le comportement humain, nombre d’entre elles sont inexactes. Prendre des décisions sur la base de ces prédictions peut avoir des conséquences désastreuses pour les personnes. Il pourrait être impossible d’améliorer la technologie à un niveau tel que ses avantages l’emportent sur les coûts.
6	**L’IA pourrait devenir plus « légère » et fonctionner sur des appareils courants** Plutôt que quelques grands modèles d’IA fonctionnant sur des superordinateurs en nuage, les futurs modèles d’IA pourraient être diversifiés et personnalisés, certains d’entre eux fonctionnant sur de petits appareils locaux tels que les téléphones intelligents. Cela pourrait compliquer la réglementation et le contrôle, et multiplier les risques de cybersécurité.
7	**Des environnements intelligents pilotés par l’IA partout** De nombreux produits pourraient être vendus avec l’IA par défaut, créant ainsi des environnements « intelligents » capables d’apprendre et d’évoluer pour s’adapter aux besoins des propriétaires et des utilisateurs. Il peut être difficile pour les gens de comprendre les capacités des environnements intelligents ou de s’en passer.
8	**L’IA compromet encore plus la vie privée** À mesure que les appareils dotés d’IA collectent davantage de données en ligne et dans la vie réelle, les efforts visant à transformer les données en nouvelles sources de revenus pourraient se heurter à des attitudes et des appareils plus soucieux de la protection de la vie privée. Un nouveau statu quo pourrait voir le jour, très différent de la manière opaque dont les utilisateurs échangent aujourd’hui leurs données contre des services gratuits.

9	**Les données collectées sur les enfants pourraient remodeler leur vie au présent et à l'avenir** Les juridictions s'inquiètent de la protection de la vie privée des enfants à mesure que les technologies de l'IA deviennent plus omniprésentes. La collecte omniprésente de données pendant l'enfance pourrait offrir de nouvelles possibilités en matière d'accessibilité et d'éducation, mais aussi aggraver les vulnérabilités existantes, éroder la vie privée et remodeler la vie des adultes à l'avenir.
10	**L'IA pourrait remodeler nos relations avec les autres** Les outils d'IA pourraient médiatiser davantage les interactions sociales- dans des contextes publics ou professionnels, ou en privé avec des amis, des membres de la famille ou des partenaires romantiques. Ces outils pourraient être utilisés pour signaler des comportements suspects ou nuisibles et aider à éviter les maladresses sociales-- mais ils pourraient aussi aider à manipuler les autres et à en faire des proies.

Idée 1 : L'IA pourrait briser Internet tel que nous le connaissons actuellement

Les nouveaux outils d'intelligence artificielle pourraient ébranler le modèle économique de la publicité qui a servi de base à Internet pendant la majeure partie des 20 dernières années. Internet à l'ère de l'IA pourrait être très différent, avec plus d'autonomie et de contrôle, mais aussi moins utile et moins sûr.

Aujourd'hui

Internet fait partie intégrante de la vie quotidienne des Canadiennes et Canadiens. Les jeunes, en particulier, y trouvent une source d'amitiés[5] et d'informations[6]. Dans l'ensemble de la population, 95 % de la population canadienne âgée de plus de 15 ans utilisent Internet[7] et 75 % effectuent des opérations bancaires en ligne[8] et des achats en ligne[9]. Près de la moitié des ménages disposent d'appareils intelligents connectés à Internet[10].

Bien qu'il n'y ait jamais eu autant de sites Web, l'expérience de la plupart des gens sur Internet est dominée par une petite poignée d'entreprises massives. Comme le dit une blague en ligne, « Internet, c'est cinq sites géants montrant des captures d'écran et des textes des quatre autres »[11]. Aujourd'hui, 65 % de l'ensemble du trafic Internet se fait vers des domaines appartenant à Alphabet, Meta, Netflix, Microsoft, Tik Tok, Apple, Amazon ou Disney[12]. Google représente 91 % de toutes les recherches sur Internet[13].

La publicité finance la fourniture de services en ligne gratuits et l'économie des créateurs en ligne.[14]Alphabet, Meta, Apple, Microsoft et Amazon gagnent chacun des milliards grâce à la publicité en ligne[15]. Les entreprises ont investi 46,7 milliards de dollars américains en 2021 dans l'optimisation de la conception de leur site Web afin d'obtenir un meilleur classement dans les moteurs de recherche, d'augmenter le trafic et de générer davantage de recettes publicitaires[16].

Le contenu généré par l'IA devient rapidement plus réaliste et plus humain. Jusqu'à récemment, la plupart des contenus en ligne étaient générés par l'humain, car les contenus générés par ordinateur étaient généralement de mauvaise qualité. Cela a commencé à changer en 2022 avec la sortie de Dall-E 2, Midjourney et ChatGPT. Les grands modèles de langage (GML) peuvent produire des textes de haute qualité ressemblant à des textes humains. Les générateurs d'images d'IA peuvent produire des images photoréalistes. Les générateurs de vidéos d'IA sont suffisamment avancés pour intéresser Hollywood[17]. Les générateurs de voix ont rendu populaires les reprises de chansons en IA[18]. Si la plupart des contenus générés par l'IA peuvent encore être identifiables grâce à des signes révélateurs subtils, il est de plus en plus difficile de les distinguer des contenus créés par l'humain.

L'IA ne joue pas encore un rôle important dans la cybersécurité, mais les incidents se multiplient. Soixante-dix pour cent des Canadiennes et Canadiens ont signalé un incident de cybersécurité en 2022, par rapport à 58 % en 2020. Bien qu'il s'agisse principalement de pourriels et de tentatives d'hameçonnage peu sophistiquées[19], les cas de fraude impliquant des « hypertrucages » ont augmenté de 477 % en 2022[20]. Les escrocs ont commencé à faire de faux appels de rançon en utilisant des voix générées par l'IA des proches de la cible[21]. Des cas de fraude institutionnelle liés à l'hypertrucage apparaissent également, entraînant des millions de dollars de pertes potentielles pour les entreprises et les gouvernements[22].

Avenirs

Les agents et les moteurs de recherche alimentés par l'IA pourraient transformer la façon dont les gens interagissent avec Internet. Au lieu que les utilisatrices et utilisateurs se rendent sur des sites Web précis, les outils d'IA pourraient créer des interfaces personnalisées et sur mesure, alimentées par des contenus provenant de l'ensemble d'Internet. Ils pourraient également les aider à trouver des contenus de niche et des communautés au-delà des principales plateformes de médias sociaux.

Ces outils pourraient bouleverser les modèles économiques basés sur la publicité sur Internet. Si une part importante du trafic Web est constituée de robots d'intelligence artificielle recueillant des informations pour leurs utilisatrices et utilisateurs, les sites Web et les moteurs de recherche risquent de tirer moins de revenus de l'affichage de publicités. Ils devront peut-être trouver d'autres moyens de générer de l'argent, tels que l'introduction d'abonnements, de péages informatiques ou la monétisation directe des données des utilisatrices et utilisateurs.

Internet pourrait être dominé par des contenus générés par l'IA, qu'il serait impossible de distinguer des contenus générés par l'humain. Les plateformes en ligne pourraient créer des contenus multimédias adaptés à chaque utilisatrice et utilisateur. Internet pourrait être inondé de sites Web générés par l'IA et remplis de pourriels, de fausses informations, de robots et de fausses critiques de produits. Il pourrait devenir difficile pour les utilisatrices et utilisateurs de différencier les contenus de qualité des contenus indésirables. Si la vérification des faits par l'IA ne s'améliore pas, cela pourrait devenir encore plus difficile.

Le sentiment général de confiance et de sécurité que les Canadiennes et Canadiens ressentent en ligne pourrait être fortement diminué. Lorsque les appels vidéo peuvent être simulés de manière convaincante, il peut être difficile pour une personne de savoir si un nouvel ami en ligne est une personne réelle ou une escroquerie par hameçonnage de l'IA. Les campagnes de désinformation alimentées par l'IA pourraient devenir plus sophistiquées, sapant davantage la confiance dans les institutions. Les outils d'IA devenant de plus en plus accessibles et puissants, toute personne ayant une présence en ligne, même minime, pourrait être exposée à un risque croissant de préjudice.

Implications

- **Les moteurs de recherche alimentés par l'IA peuvent être tenus responsables des résultats** affichés aux utilisatrices et utilisateurs. Cela pourrait avoir des répercussions juridiques et porter atteinte à la confiance, en particulier si les résultats sont erronés, voire dangereux

- Le contenu et les services qui étaient **autrefois gratuits risquent d'être placés derrière des péages informatiques** à mesure que les outils d'intelligence artificielle sapent le modèle économique de la publicité en ligne
    - Les sites Web peuvent tenter de **monétiser directement les données et le contenu des utilisatrices et utilisateurs**, par exemple en les cédant sous licence à des entreprises spécialisées dans l'IA en tant que matériel de formation
    - Le contenu commandité, le placement de produits et d'autres formes de publicité sur le site **pourraient devenir plus courants**
- S'il est peu probable que le contenu généré par l'humain disparaisse, les **créateurs de contenu pourraient avoir du mal à rivaliser** avec le contenu généré par l'IA, qui est bon marché et adapté
    - Les créateurs de contenu peuvent ressentir une plus grande **pression pour monétiser** leurs publics
    - Le goût humain et la conservation pourraient être valorisés. **Les créateurs de contenu peuvent céder la place aux conservateurs de contenu**, qui accumulent des adeptes sur la base de leur sélection de contenu en ligne
    - Un contenu unique et personnalisé pourrait amener les gens à **se sentir isolés avec moins de points de contact culturels**
- Les outils d'IA peuvent **transférer le contrôle de la conception, de la mise en page et de l'expérience d'un site Web** des concepteurs de sites Web aux utilisatrices et utilisateurs**.** Cela pourrait permettre à ces derniers d'éviter plus facilement les motifs de dépendance ou de manipulation connus sous le nom de « motifs sombres »
    - **Il pourrait devenir très difficile** de naviguer sur Internet sans les outils d'IA
    - Les sites Web pourraient **cesser d'exister tels qu'on les connaît actuellement**, devenant plutôt des réservoirs de données à gratter par l'IA. Les entreprises pourraient ne plus avoir besoin de concepteurs de sites Web

- **La méfiance pourrait être l’attitude dominante en ligne** à mesure que les risques de cybersécurité augmentent et que les contenus générés par l’IA dominent Internet
    - **Les systèmes d’hameçonnage par IA** pourraient devenir plus sophistiqués
    - Les gens peuvent devenir plus **sélectifs quant aux informations qu’ils diffusent en ligne**
    - **De nouvelles mesures d’authentification** pourraient voir le jour pour tenter de rétablir la confiance en ligne
    - Si la confiance dans l’IA continue de diminuer (voir l’idée 3), les gens pourraient **craindre d’être manipulés** par des flux de contenus adaptés à l’IA
- Si les moteurs de recherche ne peuvent pas trier efficacement le contenu de qualité du pourriel de l’IA, ils **risquent de ne plus être des sources de référence efficaces** pour toutes les requêtes
    - Les gens pourraient se fier à **quelques sources fiables** pour obtenir des informations en ligne
- Les modèles actuels de commerce électronique pourraient être perturbés de manière imprévue

Idée 2 : L’IA pourrait renforcer les acteurs non étatiques et submerger les organisations de sécurité

À l’avenir, l’IA plus accessible et plus polyvalente aura des répercussions sur la sécurité. Les acteurs non étatiques, bons ou mauvais, auront accès à des capacités traditionnellement détenues uniquement par les États. Ils pourraient être en mesure de les déployer plus rapidement que les États, ce qui maintiendrait les organisations de sécurité dans une course constante pour les rattraper.

Aujourd’hui

L’IA abaisse les barrières à l’accès et réduit le coût des attaques.[23]Par exemple, l’IA peut aider une personne ayant des compétences limitées en programmation à concevoir des logiciels malveillants[24]. Les principaux modèles d’IA à code source ouvert ont des capacités légèrement inférieures à celles de l’IA polyvalente la plus puissante, GTP-4 Turbo[25]. Leur caractère polyvalent les rend également utiles pour tous les types de problèmes, y compris les activités nuisibles. On ne sait pas qui utilisera l’IA le plus efficacement et le plus rapidement : les organes gouvernementaux des nations rivales ou les acteurs non étatiques[26]. Toutefois, l’IA pourrait habiliter des acteurs non étatiques, des entreprises ou des nations qui ne sont pas soumis à des contraintes juridiques ou éthiques et qui sont prêts à appliquer la technologie d’une manière que d’autres États ne peuvent pas adopter.

Avenirs

De nombreux nouveaux acteurs pourraient avoir accès à la surveillance à grande échelle à l’avenir. La capacité de l’IA à analyser de grandes quantités de données de sources ouvertes pourrait permettre à de nouveaux acteurs de suivre et de prévoir les mouvements des forces policières et militaires[27]. Les outils d’IA peuvent aider à écrire des codes informatiques malveillants, ce qui rend la cyberdéfense plus difficile. De même, ChatGPT a été utilisé pour créer des logiciels

malveillants évolutifs, c'est-à-dire des logiciels malveillants capables de modifier leur code d'origine pour échapper aux cyberdéfenses[28].

L'IA peut également être utilisée dans le cadre d'attaques non conventionnelles, afin de réduire le coût des dommages physiques ou des attaques contre les infrastructures. Par exemple, l'IA peut faciliter le processus d'impression 3D de pièces dangereuses, comme celles nécessaires à la fabrication d'armes nucléaires[29]. L'IA pourrait également être utilisée pour automatiser des essaims de drones bon marché afin d'écraser les défenses aériennes[30], ce qui donnerait un avantage aux acteurs plus petits qui souhaitent cibler les zones urbaines ou affronter les armées modernes. Si l'IA augmente considérablement l'accès et automatise les dommages, cela pourrait accroître la pression sur le secteur de la sécurité et modifier la manière dont il assure la sécurité des citoyens.

Implications

- L'IA à code source ouvert pourrait donner aux acteurs non étatiques de nouveaux outils et **éroder les avantages traditionnellement détenus par les États**, tels que la surveillance et le contrôle[31]
- **La capacité des organismes d'application de la loi à recueillir des renseignements pourrait être limitée** par rapport à celle des acteurs non étatiques
- **Davantage de communautés pourraient contester l'utilisation de l'IA** par les **organismes d'application de la loi**
- L'utilisation innovante de l'IA pourrait **dépasser la capacité d'adaptation des organisations de défense et de sécurité**. Les échecs en matière de sécurité publique pourraient affaiblir la confiance des institutions ou modifier l'attitude du public quant à l'utilisation appropriée de l'IA par les pouvoirs publics[32]
- Les entreprises privées spécialisées dans l'IA pourraient devenir les principaux acteurs des secteurs de la cybersécurité et du renseignement, y

compris dans des espaces traditionnellement considérés comme relevant du domaine public

Idée 3 : Le manque de confiance dans l'IA pourrait entraver son adoption

On ne sait pas comment évoluera la confiance dans l'IA. Les défaillances fréquentes ou non corrigées des systèmes d'IA, ou une seule défaillance importante, pourraient éroder la confiance et entraver l'adoption, mettant en péril des secteurs entiers. Les nouvelles formes de certification, de vérification et d'efforts pour réparer les préjudices pourraient encourager la confiance des utilisatrices et utilisateurs et l'adoption du système.

Aujourd'hui

La confiance est essentielle à l'acceptation de l'IA et, au Canada, la confiance dans l'IA est en déclin[33]. L'indice CanTrust montre que la confiance des Canadiennes et Canadiens dans l'IA a diminué de 6 % entre 2018 et 2024.[34] L'étude mondiale IPSOS AI Monitor montre que l'anglosphère, y compris le Canada, fait moins confiance à l'IA que d'autres régions : par exemple, 63 % des Canadiennes et Canadiens sont nerveux à l'égard des produits et services qui utilisent l'IA, contre seulement 25 % des Japonaises et Japonais[35].

La confiance dans l'IA dépend du contexte dans lequel elle est utilisée. Par exemple, la confiance est plus élevée pour des tâches simples telles que le réglage d'un thermostat, et plus faible pour des tâches liées à la sécurité personnelle telles que les voitures autopilotées[36]. La confiance du public dans les voitures autonomes est faible et en baisse. En 2023, seuls 22 % des Canadiennes et Canadiens déclarent faire confiance aux voitures autonomes et autres moyens de transport sans conducteur basés sur l'IA[37]; contre 37 % des Américaines et Américains, soit une baisse par rapport aux 39 % de 2022 et aux 41 % de 2021[38].

Malgré la baisse de confiance, l'utilisation des outils d'IA au Canada augmente. Selon un sondage de 2024 de Legar, 30 % des Canadiennes et Canadiens utilisent désormais l'IA, contre 25 % il y a un an. Les jeunes utilisent davantage l'IA que les personnes plus âgées; 50 % des 18-35 ans déclarent utiliser l'IA, contre seulement 13 % des 55 ans et plus[39].

Les risques et les échecs liés aux technologies de l'IA ont fréquemment attiré l'attention du public au cours de l'année écoulée. Dans certains cas, la mise au point et les tests de nombreux outils d'IA ont été effectués après le lancement public, ce qui contraste fortement avec les essais de médicaments cliniques qui nécessitent de longues périodes de test avant d'être rendus publics. De nouvelles initiatives ont vu le jour pour recenser les incidents liés à l'IA et en rendre compte, comme la base de données des incidents liés à l'IA et le moniteur des incidents liés à l'IA de l'OCDE[4041].

L'adoption de l'IA peut sembler forcée, plutôt que choisie par la personne. La volonté actuelle d'intégrer l'IA partout peut signifier que les préoccupations valables concernant la sécurité des données, l'équité, les conséquences environnementales et la sécurité de l'emploi sont minimisées[42]. Forcer les gens à adopter l'IA dans leur vie quotidienne sans faire d'efforts pour rendre la technologie plus digne de confiance peut limiter les répercussions transformationnelles potentielles de la technologie.[43] La réaction actuelle contre l'utilisation croissante de la technologie de reconnaissance faciale dans les aéroports est un exemple de l'interaction entre l'adaptation forcée et l'absence de confiance[44].

Avenirs

L'amélioration des technologies, des pratiques et des systèmes pourrait contribuer à renforcer la confiance dans l'IA. Par exemple, de nouvelles capacités telles que l'IA neuro-symbolique, qui combine des réseaux neuronaux avec un traitement symbolique basé sur des règles, promettent d'améliorer la transparence et l'explicabilité des modèles d'IA. L'adoption par les entreprises de nouveaux modèles d'étiquetage, de certification ou d'assurance pourrait compenser une partie de la méfiance à l'égard de l'IA[45,46]. Certains fournisseurs mettent actuellement au point des moyens d'évaluer la sécurité et la fiabilité des modèles d'IA, en offrant des garanties pour vérifier leurs performances[47,48]. À l'avenir, les systèmes d'IA pourraient fournir un intervalle de confiance pour tout, des résultats de recherche aux véhicules autopilotés, aidant les utilisateurs à évaluer les risques et les incertitudes[49].

Un déploiement plus stratégique et réfléchi de l'IA pourrait renforcer la confiance. À l'avenir, l'IA deviendra probablement la bonne solution à certains problèmes, mais pas à d'autres. La confiance dans l'IA pourrait être renforcée si les gens perçoivent qu'elle leur facilite la vie[50], plutôt que de remplacer des tâches qu'ils apprécient ou de sembler être une solution à la recherche d'un problème. La familiarité individuelle avec l'IA peut renforcer la confiance dans un domaine du travail ou de la vie, sans nécessairement se traduire par des niveaux de confiance accrus dans l'ensemble de l'écosystème de l'IA[51].

Les échecs très médiatisés et l'appréciation croissante des risques pourraient éroder la confiance. Le scepticisme et la méfiance pourraient s'accroître à mesure que les risques de l'IA sont mieux connus et documentés et que davantage de tâches à fortes répercussions sont déléguées à l'IA. Les groupes qui subissent les effets négatifs de l'IA s'opposent activement à son utilisation dans certains domaines, comme les écrivaines, écrivains et artistes qui s'organisent collectivement pour limiter ce qu'ils considèrent comme le pouvoir destructeur de la technologie[52]. La méfiance pourrait être alimentée non seulement par des récits décrivant l'IA comme une menace d'extinction, mais aussi par son association avec l'accroissement des inégalités[53]. De même,les échecs technologiques très médiatisés pourraient faire planer l'ombre d'une perte de confiance en l'avenir. Par exemple, la confiance du public et le soutien à l'énergie nucléaire au Canada ont considérablement diminué à la suite de l'accident nucléaire de Fukushima Daiichi en 2011, et les préoccupations du public concernant la sûreté nucléaire ont entravé la croissance du secteur pendant des années[54]. Une perte de confiance semblable dans les technologies de l'IA telles que les voitures autonomes pourrait mettre en péril non seulement une entreprise, mais aussi des secteurs entiers.

Implications

- Le manque de confiance pourrait être un **obstacle majeur** à l'intégration de l'IA dans certains secteurs
- Un seul **incident aberrant et très médiatisé** impliquant un système d'IA bien établi pourrait nuire de manière disproportionnée à la confiance dans l'IA et à

son adoption; par exemple, une crise financière déclenchée par un contenu généré par l'IA et le commerce algorithmique à haute fréquence

- Les gens pourraient faire confiance à l'IA pour effectuer certaines tâches **plus qu'ils ne font confiance à d'autres humains**
- **Les différents niveaux de confiance dans l'IA** selon les groupes ou les cas d'utilisation pourraient **unir les gens au-delà des divisions sociétales habituelles ou les polariser d'une nouvelle manière**
- **Une confiance excessive** dans certains résultats de l'IA **pourrait accroître la désinformation**, avec des conséquences pour la démocratie et la cohésion sociétale
- **Une mauvaise expérience avec un système d'IA pourrait conduire à une perte de confiance dans d'autres systèmes d'IA**, alors qu'une expérience positive avec un outil d'IA pourrait conduire à une confiance accrue dans d'autres applications d'IA
- **La jurisprudence et la législation** qui déterminent la responsabilité des décisions prises par ou avec l'IA **pourraient influencer la confiance et l'adoption**
- L'émergence de **nouveaux labels et certifications pourrait nuire à la confiance des consommateurs** dans l'IA, tels que les labels d'avertissement, ou les labels analogues à ceux du commerce équitable ou des produits biologiques[55]
- **Les régimes d'imputabilité et de responsabilité seront clarifiés**, et de nombreux systèmes devront déterminer qui est responsable des défaillances de l'IA

Idée 4 : Les biais dans les systèmes d'IA pourraient perdurer

La partialité est une caractéristique de la prise de décision humaine et de l'IA. Comme les données utilisées pour former l'IA sont souvent entachées de biais difficiles à corriger, le recours croissant à l'IA dans les systèmes de prise de décision pourrait répandre les biais et entraîner des dommages importants. Les préjugés ne peuvent jamais être éliminés, en partie à cause des perspectives contradictoires sur l'équité.

Aujourd'hui

Les préjugés dans l'IA sont considérés comme un problème majeur capable d'automatiser la discrimination à grande échelle d'une manière qui peut être difficile à déterminer. Si les décisions humaines sont également biaisées, l'un des principaux risques de l'automatisation des décisions à fort enjeu est qu'elles se généralisent, augmentant ainsi la possibilité d'erreurs et de préjudices systémiques. Alors qu'un gestionnaire partial pourrait décider d'accorder des notes d'entretien plus élevées aux quelques candidats qui lui ressemblent et parlent comme lui, un modèle d'IA biaisé pourrait avoir un effet semblable sur des milliers de personnes au sein d'organisations, de secteurs ou de pays.

De nombreux produits d'IA prétendent être moins biaisés que les décideurs humains, mais les enquêtes indépendantes ont révélé des échecs et des rejets systématiques[56]**.** Par exemple, un audit de deux outils d'embauche par IA a révélé que les types de personnalité prédits variaient selon que le candidat soumettait son CV en format Word ou en texte brut[57]. Des outils semblables ont été discriminatoires à l'égard des femmes[58] ou des personnes handicapées[59]. Les biais sont intégrés à l'IA à de nombreux stades de son cycle de vie : données de formation, développement algorithmique, interaction avec l'utilisateur et retour d'information[60].

Les biais peuvent être impossibles à éliminer parce que les données utilisées pour former les modèles d'IA sont elles-mêmes souvent biaisées d'une manière qui ne peut pas être facilement corrigée. Le contrôle des résultats peut également poser des problèmes. Par exemple, un modèle d'IA qui apprend à écarter

les mots à connotation raciste pourrait omettre des informations importantes sur l'Holocauste ou l'esclavage[61]. En outre, les algorithmes ne peuvent souvent pas calculer différentes notions d'équité en même temps, ce qui conduit à des résultats constamment différents pour certains groupes[626364].

Avenirs

Dans un avenir où les préjugés ne pourront jamais être éliminés, qu'ils soient humains ou algorithmiques, les sociétés devront peut-être repenser les idées actuelles sur l'équité et la meilleure façon d'y parvenir. Les gens ne sont pas nécessairement d'accord sur le sens du mot « équitable ». Par exemple, certains considèrent que la discrimination positive est juste, d'autres non. Les institutions pourraient adopter des normes visant à distribuer les ressources, emplois, subventions, prix ou autres biens, d'une manière qui tente explicitement de réparer les injustices historiques. Les organisations qui cherchent à éviter les préjugés systémiques peuvent utiliser une approche de « pluralisme algorithmique », qui implique divers éléments dans le processus de prise de décision et garantit qu'aucun algorithme ne limite sévèrement les possibilités[65].

Des efforts pourraient être faits pour réduire les biais dans les systèmes d'IA à un niveau acceptable, bien qu'il soit impossible de les éliminer complètement. L'utilisation des technologies de l'IA dans certains domaines sensibles, tels que la police ou l'embauche, pourrait continuer à susciter des réactions négatives. Par ailleurs, ces technologies pourraient continuer à s'améliorer et devenir moins biaisées à l'avenir. Quoi qu'il en soit, il est probable que l'on continuera à s'efforcer de réduire les biais dans les technologies d'IA.

Implications

- **Les préjudices ou les échecs systémiques pourraient être institutionnalisés** dans des contextes où des algorithmes uniques sont autorisés à prendre des décisions globales concernant l'accès des personnes à certaines ressources (p. ex., des emplois, des prêts, des visas)

- **Les préjugés humains pourraient devenir plus importants** parmi ceux qui utilisent les systèmes d'IA, car les gens apprennent et reproduisent les perspectives biaisées de l'IA, portant les préjugés avec eux au-delà de leurs interactions
- **Les désaccords sur la meilleure façon de coder l'équité algorithmique** peuvent résulter de définitions différentes de ce que signifie réellement l'équité. Cela peut conduire à des résultats complètement différents pour des technologies ou des systèmes semblables
- L'incapacité à éliminer les préjugés des algorithmes pourrait **conduire à des divisions politiques, sociales ou économiques**
- Si les décisions sont de plus en plus distribuées, notamment divers algorithmes et humains à différents stades d'un processus, **il pourrait être difficile de faire des réclamations pour discrimination** ou de déterminer une partie responsable de la discrimination
- Des cas très médiatisés de discrimination algorithmique pourraient entraîner **une perte de confiance dans les systèmes décisionnels de l'IA**, en particulier dans les domaines de la police et des soins de santé, ainsi qu'une augmentation des litiges

Idée 5 : L'utilisation de l'IA pour prédire le comportement humain pourrait s'avérer inefficace

Si l'IA fait parfois des prédictions impressionnantes sur le comportement humain, nombre d'entre elles sont inexactes. Prendre des décisions sur la base de ces prédictions peut avoir des conséquences désastreuses pour les personnes. Il pourrait être impossible d'améliorer la technologie à un niveau tel que ses avantages l'emportent sur les coûts.

Aujourd'hui

De plus en plus de gouvernements et d'institutions utilisent l'IA pour prédire le comportement humain et prendre des décisions concernant les personnes. Par exemple, plus de 500 écoles aux États-Unis utilisent un modèle d'IA appelé Navigate pour prédire la réussite des élèves[66]. Aux États-Unis, les travailleurs sociaux ont utilisé l'IA pour prédire quels appels de protection de l'enfance doivent faire l'objet d'une enquête plus approfondie[67]. Tous deux sont des exemples d'« **optimisation prédictive** »[68]. D'éminents ingénieurs en IA ont affirmé que les algorithmes d'optimisation prédictive reposaient sur une science erronée, les prédictions de l'IA n'étant que légèrement plus précises que le hasard d'un tirage à pile ou face[69]. Malgré cela, ils continuent d'être utilisés parce qu'ils externalisent des tâches complexes telles que l'élaboration de règles de prise de décision (p. ex., quels sont les critères à prendre en compte pour rechercher des comportements frauduleux ou comment décider si un enfant risque d'être victime d'abus). Les règles de prise de décision générées par l'humain peuvent sembler subjectives et inexactes par rapport à celles des modèles prédictifs d'IA, qui prétendent représenter des modèles objectifs dans le monde réel.

Optimisation prédictive

L'utilisation de l'IA pour prédire des résultats futurs sur la base de données historiques, afin de prendre des décisions concernant des personnes.

Les modèles prédictifs ne sont pas toujours justes. Les modèles d'IA prédictifs sont confrontés à de nombreux problèmes, notamment des erreurs dues à une inadéquation entre les données d'entraînement et les données de déploiement. Comme l'IA prédictive doit être formée sur des données antérieures, elle ne peut

pas prendre en compte les variables émergentes et complexes du monde, et des comportements humains individuels. Les modèles peuvent être incapables de prendre en compte des facteurs nouveaux et inattendus. En outre, l'IA ne peut pas filtrer les effets des pratiques racistes dans le monde réel, telles que le maintien de l'ordre disproportionné dans les quartiers ou les communautés noirs, qui entraîne une augmentation des fausses arrestations[70]. Cela a conduit à des prévisions inexactes pour les personnes vulnérables[71].

Les modèles d'IA prédictifs ne peuvent pas comprendre pourquoi le comportement dans le monde réel diffère de leurs prédictions. Les modèles peuvent supposer que les personnes agiront de manière rationnelle ou suivront les mêmes règles et modèles humains dans l'ensemble. Les modèles peuvent ne pas tenir compte des facteurs structurels qui expliquent les différences entre les comportements prévus et les comportements réels. L'accent mis sur la prédiction peut empêcher la découverte de processus susceptibles d'engendrer de nouveaux comportements, comme lorsque la simplification du langage utilisé dans les convocations au tribunal a diminué le taux de personnes ne se présentant pas au tribunal[72].

Bien que cela soit parfois justifié par des économies de coûts, certains gouvernements ont ressenti des répercussions importantes après avoir utilisé des modèles d'optimisation prédictive. Par exemple, en 2021, le gouvernement néerlandais a démissionné à la suite d'un scandale impliquant l'adoption par l'administration fiscale d'une IA autoapprenante pour prédire les fraudes aux allocations de garde d'enfants[73]. L'IA a identifié à tort des dizaines de milliers de familles comme ayant des dettes excessives envers l'administration fiscale. Plus d'un millier d'enfants ont été placés dans des familles d'accueil et certaines victimes se sont suicidées. Les amendes infligées aux agences gouvernementales et les indemnités versées aux victimes pourraient atteindre plusieurs centaines de millions d'euros.

Avenirs

À l'avenir, l'optimisation prédictive pourrait être utilisée dans certaines juridictions, mais pas dans d'autres. Elle pourrait être interdite dans certaines

juridictions, en particulier là où les gouvernements ont dû faire face à des coûts élevés et à un examen minutieux en raison d'échecs. Cela pourrait encore permettre au secteur privé d'étendre ses utilisations actuellement opaques de l'optimisation prédictive[74]. D'autres juridictions peuvent continuer à utiliser des algorithmes d'optimisation prédictive malgré les risques. Cela peut s'expliquer par le fait que les personnes concernées sont moins à même de saisir la justice ou que leurs gouvernements ne sont pas liés par des normes démocratiques. D'autres peuvent considérer l'optimisation prédictive comme un outil inévitablement imparfait, mais dont l'utilisation peut être justifiée par des économies de coûts. Les institutions, y compris les gouvernements, qui adoptent l'IA pour l'optimisation prédictive et constatent que les coûts dépassent les avantages pourraient maintenir les systèmes en service bien plus longtemps qu'elles ne le devraient ou ne le voudraient, en raison des montants élevés déjà investis ou des difficultés liées à l'annulation d'un déploiement. Certains pourraient considérer l'IA prédictive comme inacceptable d'un point de vue éthique pour la prise de décision, et préférer travailler sur des interventions visant à réduire autant que possible les résultats négatifs prévus.

Implication

- Les gouvernements et les entreprises qui utilisent l'optimisation prédictive sans être transparents sur les règles de prise de décision de l'IA **pourraient être considérés comme indignes de confiance**
- Si les institutions utilisent l'IA pour l'optimisation prédictive alors que la charge de la preuve pour contester des prédictions inexactes incombe aux personnes concernées, **les populations déjà vulnérables risquent de voir leur situation s'aggraver**. Cela pourrait créer de nouveaux goulets d'étranglement bureaucratiques et engorger les tribunaux avec des litiges portant sur des préjudices algorithmiques, y compris des affaires liées aux droits de l'humain ou à des violations de la Charte

- Les tentatives de sacrifier les droits individuels au profit de gains collectifs peuvent **bénéficier aux populations privilégiées au détriment des plus vulnérables,** créant ainsi de plus grandes divisions socioéconomiques
- L'adoption de modèles d'optimisation prédictive pourrait créer **des économies initiales qui cèdent rapidement la place à de nouveaux coûts** : pour lutter contre les litiges liés à des prédictions inexactes, pour resigner des contrats avec des fournisseurs afin de recycler et de réajuster les modèles, et pour créer de nouvelles voies pour les plaintes et l'indemnisation des dommages
- Si le processus décisionnel de l'IA punit préventivement les personnes sur la base d'hypothèses biaisées, il pourrait **réduire l'autonomie des populations vulnérables** et dresser de nouveaux obstacles dans leur parcours de vie

Idée 6 : L'IA pourrait devenir plus « légère » et fonctionner sur des appareils courants

Plutôt que quelques grands modèles d'IA fonctionnant sur des superordinateurs en nuage, les futurs modèles d'IA pourraient être diversifiés et personnalisés, certains d'entre eux fonctionnant sur de petits appareils demainlocaux tels que les téléphones intelligents. Cela pourrait compliquer la réglementation et le contrôle, et multiplier les risques de cybersécurité.

Aujourd'hui

Les améliorations apportées aux techniques d'entraînement et de compression de l'IA permettent aux modèles plus petits et moins gourmands en ressources de devenir plus performants. La taille d'un modèle d'IA est souvent utilisée pour désigner sa puissance, sa capacité et sa qualité. Alors que les plus grands modèles sont souvent les plus puissants et les plus performants, les développeurs d'IA mettent sur le marché des versions plus petites et comprimées dérivées de modèles plus grands. Cela permet au modèle plus petit de conserver la plupart des performances du modèle plus grand tout en étant moins gourmand en énergie et fonctionnant sur du matériel moins puissant[75]. C'est pourquoi les modèles plus petits et plus récents sont plus performants que les modèles plus anciens et plus grands. Par exemple, Phi-3, qui a été lancé début 2024 et ne compte que 3,8 milliards de paramètres, a des performances comparables à GPT-3.5, qui a été lancé fin 2022 et compte 175 milliards de paramètres[76]. Des entreprises telles que Meta[77] et Mistral[78] ont publié des modèles d'IA en code source libre qui rivalisent avec les performances de ChatGPT, mais qui peuvent être exécutés sur un ordinateur portable. Les chercheurs dans le domaine du TinyML développent une IA plus petite et capable de fonctionner sur des appareils moins puissants pour permettre un Internet

Taille du modèle

La taille d'un modèle d'IA est déterminée par le nombre de ses paramètres. Les paramètres sont des variables d'un système d'IA dont les valeurs sont ajustées au cours de l'entraînement. Les petits modèles peuvent avoir des millions de paramètres, voire moins, tandis que les grands modèles peuvent en avoir plus de 400 milliards.

des objets (IdO) « intelligent ». Par exemple, le Raspberry Pi, un ordinateur de la taille d'une carte de crédit, très prisé des amateurs de programmation et d'ingénierie informatique, peut désormais exécuter une série de modèles d'IA, y compris la reconnaissance faciale[79].

Avenirs

Nous pourrions voir des milliers de modèles d'IA différents capables de fonctionner localement sur tous les types d'appareils numériques, des téléphones intelligents aux petits ordinateurs[80].

Ces modèles pourraient être développés par des amateurs, des jeunes poussespousses ou des criminels. Ils pourraient être basés sur des modèles à source ouverte et adaptés à différents objectifs par l'intermédiaire d'un entraînement sur des ensembles de données largement accessibles. Par exemple, Venice AI est un service d'IA basé sur le Web, construit à partir d'une poignée de modèles d'IA à source ouverte, qui permet aux utilisatrices et utilisateurs de générer du texte, du code ou des images avec peu ou pas de garde-fous et qui est vendu comme « privé et sans permission »[81]. Le déploiement de modèles d'IA de différentes tailles pourrait donner naissance à un écosystème de modèles d'IA présentant divers degrés d'interopérabilité. Les petits modèles pourraient interagir avec les grands modèles accessibles au public et basés sur l'informatique en nuage, en tirant parti de leur puissance pour effectuer des tâches ou apprendre (voir la figure 1). Ces petits modèles localisés peuvent manquer de mesures de sécurité et être déployés à grande échelle à l'insu de toute autorité.

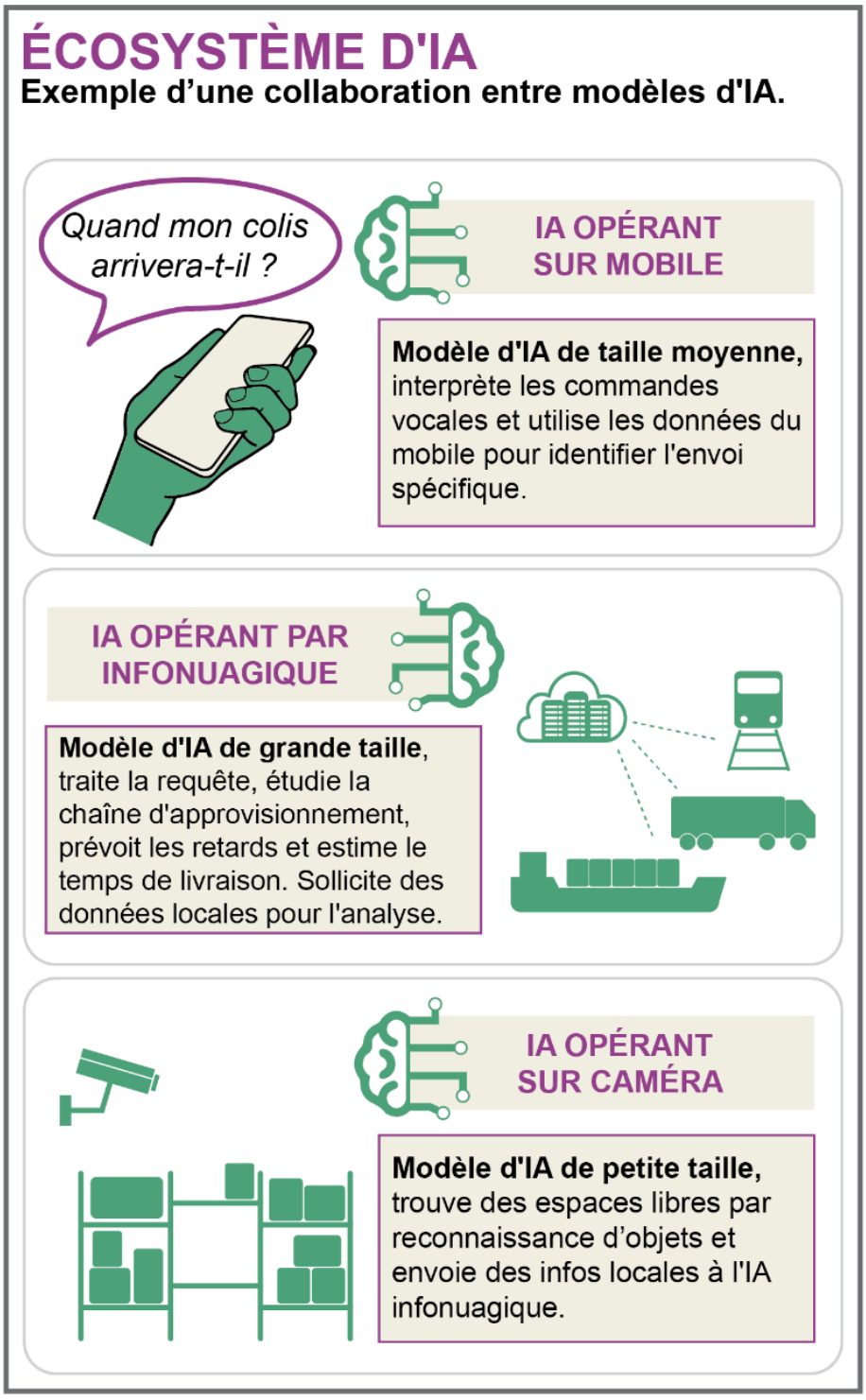

ÉCOSYSTÈME D'IA
Exemple d'une collaboration entre modèles d'IA.
Quand mon colis arrivera-t-il ?
IA OPÉRANT SUR MOBILE
Modèle d'IA de taille moyenne, interprète les commandes vocales et utilise les données du mobile pour identifier l'envoi spécifique.
IA OPÉRANT PAR INFONUAGIQUE
Modèle d'IA de grande taille, traite la requête, étudie la chaîne d'approvisionnement, prévoit les retards et estime le temps de livraison. Sollicite des données locales pour l'analyse.
IA OPÉRANT SUR CAMÉRA
Modèle d'IA de petite taille, trouve des espaces libres par reconnaissance d'objets et envoie des infos locales à l'IA infonuagique.

Figure 1

Implications

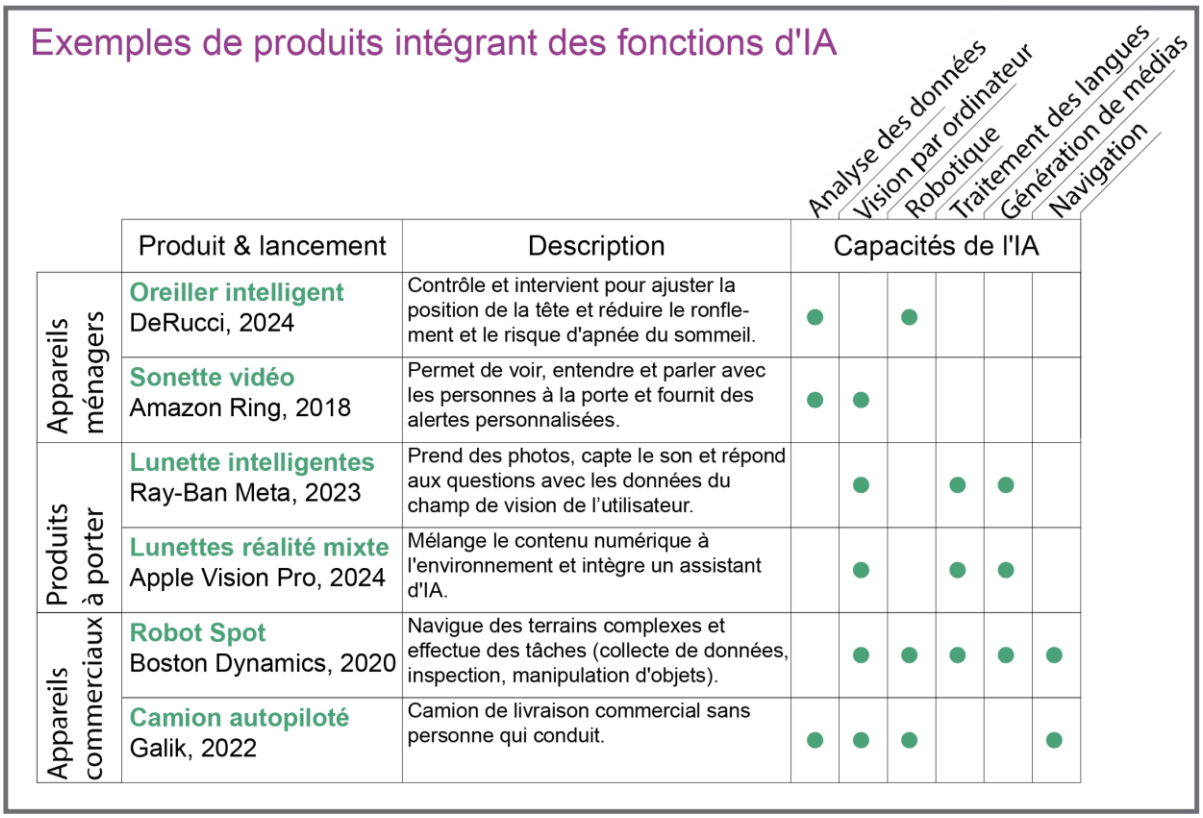
- Les réglementations axées uniquement sur les grands modèles d'IA **risquent d'être inefficaces**[82]
- **L'IA en libre accès pourrait permettre la circulation de modèles qui posent problème**, que ce soit parce qu'ils comportent des biais, manquent de mesures de sécurité ou facilitent des activités illégales[83]
- **Il pourrait être difficile de repérer les mauvais acteurs** qui forment ou utilisent des modèles d'IA petits, mais puissants
- En analysant les données localement, les modèles d'IA qui opèrent directement sur appareil pourraient aider les personnes **à protéger leurs données et leur vie privée**
- **Les petites entreprises pourraient personnaliser leurs propres outils d'IA** pour mieux répondre à leurs besoins[84]
- La compatibilité entre les appareils dotés d'IA pourrait **offrir plus d'options aux utilisatrices et utilisateurs**, mais aussi créer des vulnérabilités en matière de cybersécurité[85]

Idée 7 : Des environnements intelligents pilotés par l'IA partout

De nombreux produits pourraient être vendus avec l'IA par défaut, créant ainsi des environnements « intelligents » capables d'apprendre et d'évoluer pour s'adapter aux besoins des propriétaires et des utilisateurs. Il peut être difficile pour les gens de comprendre les capacités des environnements intelligents ou de s'en passer.

Aujourd'hui

Les dispositifs autonomes et les robots sont de plus en plus présents dans notre vie quotidienne. Par exemple, les restaurants utilisent des robots pour livrer les repas[86]. Les robots nettoyeurs sont couramment utilisés dans les espaces commerciaux[87]. Dans le secteur agricole, les machines autonomes et semi-autonomes sont de plus en plus utilisées pour les cultures. Dans les foyers, l'IA est ajoutée aux appareils de tous les jours. La figure 2 présente d'autres exemples. Ces



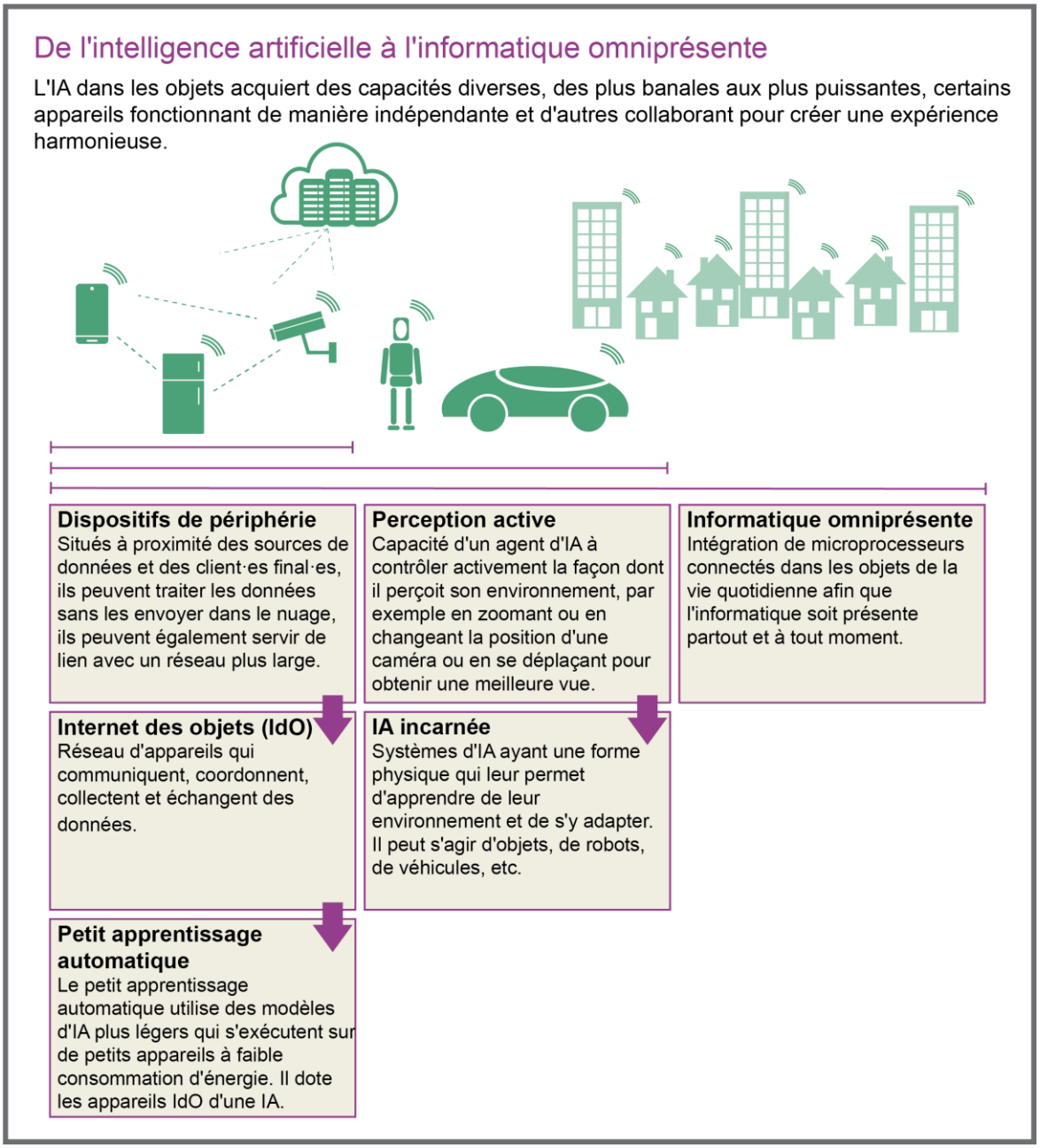
	Produit & lancement	Description	Capacités de l'IA					
Appareils ménagers	**Oreiller intelligent** DeRucci, 2024	Contrôle et intervient pour ajuster la position de la tête et réduire le ronflement et le risque d'apnée du sommeil.	●		●			
	Sonette vidéo Amazon Ring, 2018	Permet de voir, entendre et parler avec les personnes à la porte et fournit des alertes personnalisées.	●	●				
Produits à porter	**Lunette intelligentes** Ray-Ban Meta, 2023	Prend des photos, capte le son et répond aux questions avec les données du champ de vision de l'utilisateur.		●		●	●	
	Lunettes réalité mixte Apple Vision Pro, 2024	Mélange le contenu numérique à l'environnement et intègre un assistant d'IA.		●		●	●	
Appareils commerciaux	**Robot Spot** Boston Dynamics, 2020	Navigue des terrains complexes et effectue des tâches (collecte de données, inspection, manipulation d'objets).		●	●	●	●	●
	Camion autopiloté Galik, 2022	Camion de livraison commercial sans personne qui conduit.	●	●	●			●

Figure 2

appareils pourraient continuer à bénéficier de nouvelles fonctionnalités au fur et à mesure de la sortie de modèles d'IA plus performants[88].

Les chercheurs et l'industrie pourraient avoir besoin de plus de données sur le monde physique pour entraîner des IA plus avancées. L'IA qui recueille des informations en temps réel sur son environnement physique est appelée « IA incarnée » (voir la figure 3)[89]. L'IA peut s'incarner dans n'importe quoi, des

Figure 3



téléphones intelligents aux appareils ménagers, en passant par des robots à l'apparence humaine. Connectée à des capteurs et dotée d'une certaine mobilité, l'IA peut interagir avec les personnes et les espaces physiques, par exemple en ouvrant des portes ou en appelant des ascenseurs[90]. Le fait de doter l'IA d'un corps peut lui permettre d'apprendre en interagissant avec le monde, comme le font les humains, ce qui pourrait constituer une voie vers le développement d'une IA plus avancée[91].

Il devient de plus en plus difficile de comprendre les capacités des appareils qui nous entourent. Certains appareils sont qualifiés de « robots » alors qu'ils ne sont pas dotés de capacités d'intelligence artificielle[92]. D'autres appareils peuvent avoir plusieurs fonctions d'IA. Par exemple, les touristes peuvent louer des vélos électriques alimentés par l'IA qui peuvent offrir une visite guidée de la ville[93]. Les ornithologues peuvent acheter des jumelles dotées d'IA qui identifient la faune et la flore[94].

Les anciens appareils peuvent souvent être dotés de nouvelles fonctionnalités qui ne sont pas visibles de l'extérieur. Par exemple, une trousse d'intelligence artificielle peut rendre un tracteur existant entièrement autonome[95]. Les caméras de sécurité qui fonctionnent depuis longtemps peuvent être connectées à un logiciel de reconnaissance faciale[96].

Avenirs

À l'avenir, il est possible que l'on trouve de plus en plus d'appareils alimentés par l'IA dans un plus grand nombre d'environnements, des lieux de travail aux espaces de loisirs en passant par les habitations. Il peut devenir impossible d'éviter d'interagir avec ces dispositifs. Le nombre d'appareils IdO (Internet des objets) pourrait atteindre 75 milliards d'ici 2025, soit plus que doubler en quatre ans[97] et le marché mondial des logiciels d'IA pourrait être multiplié par cinq environ[98] entre 2022 et 2027.

Les fabricants d'appareils pourraient être incités à ajouter des capacités d'IA à un plus grand nombre d'appareils, soit pour les vendre, soit pour collecter des données. Les données peuvent être utiles non seulement pour générer de

nouvelles sources de revenus, mais aussi pour entraîner de nouveaux modèles. Cela pourrait être particulièrement pertinent si l'IA incarnée s'avère utile pour construire la prochaine génération de modèles d'IA de frontière, ou si les entreprises atteignent les limites des données d'entraînement de qualité existantes[99]. Par exemple, en déployant une flotte de voitures intelligentes, une entreprise pourrait utiliser des données sur le paysage urbain, la circulation et le comportement des piétons pour former des modèles d'IA encore plus puissants.

Les appareils de tous les jours pourraient finir par être dotés de capacités d'IA plus puissantes que nécessaire. Il peut être plus facile d'équiper un appareil d'une IA standard à usage général, telle que ChatGPT ou Copilot, que de personnaliser un modèle avec des fonctionnalités plus ciblées. Les appareils intelligents pourraient devenir la norme dans les nouveaux logements, prêts à s'adapter aux nouveaux propriétaires ou locataires. Les appareils pourraient être vendus avec certaines fonctions verrouillées par un modèle d'accès payant, comme cela a été le cas pour l'Amazon Ring[100], et pour les voitures Tesla[101] et Mercedes[102].

L'IA à usage général pourrait devenir une norme, ce qui brouillerait de plus en plus les frontières entre les catégories de produits de consommation. Par exemple, les montres intelligentes et les moniteurs d'activité physique ont suscité des inquiétudes quant au fait qu'ils pourraient occuper une zone grise réglementaire entre les dispositifs médicaux et les produits de consommation à faible enjeu[103]. Le capteur domestique Aqara peut être utilisé pour tout, du contrôle des lumières à la surveillance de la sécurité ou à la détection des chutes[104]. L'apparence de ces objets peut ne pas signaler clairement leurs capacités. Les robots de type humain pourraient avoir des yeux capables de voir à travers les murs, par exemple, ou les mêmes capteurs pourraient être entièrement cachés.

Implications

- Les gens pourraient avoir besoin de **nouvelles compétences pour naviguer dans les espaces alimentés par l'IA**. Les fabricants devront peut-être recourir à de nouveaux types d'étiquetage ou d'instructions pour divulguer les capacités de leurs dispositifs d'IA afin que les consommatrices et consommateurs prennent des décisions en connaissance de cause

- Les personnes qui ne veulent pas ou ne peuvent pas s'engager dans des espaces dotés d'IA peuvent se retrouver **dans l'impossibilité d'accéder à certains services**
- Les compagnies d'assurance pourraient **encourager certains types de surveillance de l'IAou l'exiger comme condition de couverture**[105]. Par exemple, la reconnaissance faciale pour confirmer l'identité d'un conducteur afin de réduire les vols de voitures
- **Les droits et les intérêts des personnes pourraient entrer en conflit de nouvelles façons.** Par exemple, le port de lunettes intelligentes dans les espaces publics ou l'envoi d'un robot pour faire les courses pourraient remettre en cause le droit à la vie privée. La confiance est nécessaire pour garantir que les appareils ne recueillent pas l'image des personnes sans leur consentement[106]. Les propriétaires de biens immobiliers pourraient installer des dispositifs alimentés par l'IA pour protéger leur investissement ou aider à l'entretien. Les locataires peuvent se retrouver dans une maison intelligente avec des services qu'ils ne souhaitent pas ou des paramètres qu'ils ne peuvent pas modifier
- Les environnements intelligents pourraient modifier les stratégies publicitaires. Les **appareils dotés d'une IA pourraient devenir habituels et proposer aux utilisateurs des publicités personnalisées** en temps réel. Par exemple, les voitures intelligentes peuvent rediriger les conducteurs vers certains commerces et les inciter à s'arrêter pour effectuer un achat

Idée 8 : L’IA compromet encore plus la vie privée

À mesure que les appareils dotés d’IA collectent davantage de données en ligne et dans la vie réelle, les efforts visant à transformer les données en nouvelles sources de revenus pourraient se heurter à des attitudes et des appareils plus soucieux de la protection de la vie privée. Un nouveau statu quo pourrait voir le jour, très différent de la manière opaque dont les utilisateurs échangent aujourd’hui leurs données contre des services gratuits.

Aujourd’hui

Les progrès de l’IA exacerbent les problèmes de protection de la vie privée liés à la technologie. La plupart des Canadiennes et Canadiens ont pris l’habitude d’accéder à des services en ligne gratuits, tels que les sites de médias sociaux, les plateformes d’IA générative ou les applications mobiles. Dans de nombreux cas, ils consentent sans le savoir à ce que des entreprises collectent leurs données, les vendent à des tiers et utilisent l’IA pour les exploiter et en tirer des conclusions à leur sujet. Par exemple, Facebook utilise l’IA pour déduire le risque de suicide des utilisateurs sur la base de leurs publications sur les médias sociaux[107]. Les progrès de l’IA permettent aux entreprises d’analyser une plus grande variété et une plus grande quantité de données, et de les transformer en flux de revenus d’une nouvelle manière.

Non seulement les environnements en ligne, mais aussi les espaces physiques deviennent moins privés. Comme il est indiqué dans l’idée 7, les objets ménagers de tous les jours sont équipés de capteurs pour collecter des données; des toilettes aux brosses à dents en passant par les jouets. La réalité virtuelle (RV) et les jeux vidéo peuvent collecter des données sur le comportement des utilisateurs à la maison et utiliser l’IA pour déduire les émotions et les traits de personnalité[108]. En dehors de la maison, des dispositifs tels que les lunettes intelligentes[109] et les épingles d’intelligence artificielle[110] soulèvent de nouvelles questions sur la protection de la vie privée en public. Les fragments d’ADN humain, connus sous le nom d’ADN environnemental, collectés dans les espaces publics à des fins telles que la surveillance des maladies, peuvent potentiellement être utilisés

pour suivre les personnes, prélever illégalement des génomes et s'engager dans des formes cachées de surveillance et d'analyse génétiques[111,112].

Les voitures intelligentes posent des problèmes particuliers en matière de respect de la vie privée. En 2023, la Fondation Mozilla a enquêté sur 25 marques de voitures et a constaté que chacune d'entre elles collectait des données personnelles qui n'étaient pas nécessaires au fonctionnement du véhicule[113]. Généralement obtenues à partir d'appareils mobiles connectés aux voitures par l'intermédiaire d'applications, ces données peuvent inclure le revenu annuel d'une personne, son statut d'immigré, sa race, ses informations génétiques, son activité sexuelle, ses photos, son calendrier et sa liste de choses à faire. Sur les 25 marques, 22 utilisent ces données pour faire des déductions; par exemple, à partir de la localisation et des contacts téléphoniques, et 21 communiquent ou vendent des données. Treize recueillent des informations sur les conditions météorologiques, l'état de la chaussée, les panneaux de signalisation et les « autres environnements », qui peuvent inclure les passants[114]. Quatre-vingt-quinze pour cent des nouveaux véhicules seront des véhicules connectés d'ici à 2030[115].

Avenirs

L'Internet des objets devenant l'« IA des objets », les données pourraient prendre encore plus de valeur, ce qui inciterait à en extraire toujours plus. Il pourrait devenir possible de tirer des conclusions plus sophistiquées pour prédire le comportement humain, les mouvements ou identifier les personnes, comme il est indiqué dans l'idée 5.

Toutefois, le refus de la réglementation internationale pourrait remodeler le paysage de la protection de la vie privée. De plus en plus de juridictions adoptent et appliquent de nouvelles lois sur la confidentialité des données[116], telles que l'American Privacy Rights Act[117] aux États-Unis. Cela pourrait modifier certains, voire de nombreux aspects du « capitalisme de surveillance »[118] en donnant aux utilisatrices et utilisateurs un plus grand contrôle sur leurs données. De futures réformes juridiques pourraient redéfinir les déductions en tant qu'informations personnelles[119], ce qui rendrait plus difficile leur vente à des tiers.

Les nouvelles technologies pourraient également modifier l'équilibre en matière de protection de la vie privée. L'informatique périphérique, qui fait référence aux réseaux ou aux appareils physiquement proches de l'utilisateur, pourrait améliorer la confidentialité et la sécurité des données[120]. Lorsque les données des utilisateurs sont stockées et traitées sur un appareil appartenant à l'utilisateur, il peut être plus difficile pour les entreprises de les collecter et de les vendre[121]. Cependant, l'informatique en périphérie peut également introduire de nouveaux risques, comme la reconnaissance faciale sur les appareils locaux et un accès potentiellement plus facile pour les acteurs malveillants[122].

Implications

- Les **distinctions** telles que public par rapport au privé et en ligne par rapport à hors ligne pourraient devenir **de plus en plus floues**. Les maisons et autres espaces peuvent être perçus comme plus ou moins privés en fonction de l'utilisation des appareils. Les visiteurs des maisons et les passagers des voitures peuvent exiger **de nouveaux protocoles de consentement pour protéger leur vie privée**.
- Les données transmises à des tiers pourraient conduire à ce que **les informations sensibles soient communiquées** aux compagnies d'assurance[123]
- Les écoles et les services de garde d'enfants pourraient utiliser les **protections de la vie privée comme un avantage concurrentiel** pour attirer les familles
- La surveillance pourrait modifier les **pratiques de la police et des criminels**
    - Certaines formes de **criminalité** pourraient s'enfoncer davantage dans la clandestinité et **devenir plus organisées** afin d'échapper à la détection
    - Les nouvelles capacités technologiques pourraient créer **de nouvelles possibilités de piratage, de fraude et de harcèlement**
    - La police de la circulation pourrait être moins nécessaire, car la surveillance des conducteurs par les gouvernements et les compagnies d'assurance permet de **remettre automatiquement des contraventions**

- Les militants[124] et les journalistes[125] pourraient de plus en plus utiliser l'informatique omniprésente pour « retourner le regard » en **recueillant des informations sur des organisations ou des personnes puissantes**. Cette pratique est connue sous le nom de « sousveillance » ou « équivalence ». Il peut s'agir du piratage d'informations sensibles sur la vie privée de représentants politiques ou d'autres personnalités publiques[126]
- Les régimes de protection des données pourraient devenir **plus complexes** et moins harmonisés au niveau mondial
- Les juridictions pourraient avoir du mal à trouver un équilibre entre la protection de la vie privée et le besoin des chercheurs de disposer d'ensembles de données représentatifs dans des domaines tels que la médecine[127]
- Les juridictions dont les lois sur la protection de la vie privée sont moins strictes pourraient devenir des destinations de travail ou de voyage de plus en plus « risquées »
- **Les dispositifs de protection de la vie privée** pourraient faire pencher la balance du pouvoir en faveur des utilisateurs

Idée 9 : Les données collectées sur les enfants pourraient remodeler leur vie au présent et à l'avenir

Les juridictions s'inquiètent de la protection de la vie privée des enfants à mesure que les technologies de l'IA deviennent plus omniprésentes. La collecte omniprésente de données pendant l'enfance pourrait offrir de nouvelles possibilités en matière d'accessibilité et d'éducation, mais aussi aggraver les vulnérabilités existantes, éroder la vie privée et remodeler la vie des adultes à l'avenir.

Aujourd'hui

Les jeunes sont un groupe particulièrement vulnérable en ce qui concerne la confidentialité des données[128]**.** Leur sens de l'autonomie et leur capacité à prendre des décisions sont encore en cours de développement. Un développement sain de l'enfant implique la capacité d'expérimenter et de faire des erreurs sans conséquences graves et durables. De plus en plus de juridictions étudient les moyens de protéger les droits des enfants en matière de données et de vie privée, et de relever les défis liés à l'obtention d'un consentement valable[129].

Les cas d'acteurs malveillants utilisant les données des enfants de manière à nuire à leur santé mentale et à leur bien-être sont de plus en plus nombreux. Les critiques soulignent la manière dont les grandes entreprises technologiques captent l'attention et les revenus par l'intermédiaire d'expériences d'utilisation addictives[130] et de motifs sombres[131]. Les « motifs sombres » sont un type de conception de sites Web ou d'applications qui peut être utilisé pour influencer votre prise de décision lorsque vous utilisez une application ou naviguez sur un site Web; par exemple, en rendant intentionnellement difficile l'annulation d'un service[132]. Les algorithmes des médias sociaux exacerbent les problèmes liés à une image corporelle négative[133]. L'IA générative a été impliquée dans la circulation[134] et le développement[135] de matériel pédopornographique, tant réel que généré par l'IA, par des adultes et des enfants.

Les parents disposent d'une marge de manœuvre extraordinaire pour obtenir et offrir une visibilité sur la vie privée de leurs enfants. Par exemple, les applications d'enregistrement de frappe peuvent permettre aux parents de voir non seulement les messages envoyés par un enfant, mais aussi les messages qu'il a tapés, mais qu'il a décidé de ne pas envoyer[136]. Les parents peuvent également diffuser les informations privées de leurs enfants à d'autres personnes. Certains gèrent des comptes d'enfants influents qui génèrent des revenus et diffusent régulièrement des informations personnelles et des images de leurs enfants. Ces comptes sont parfois ouvertement suivis par des pédophiles, qui bénéficient des politiques de la plateforme qui récompensent l'engagement[137].

Les ombres de « données » non désirées pourraient suivre les enfants jusqu'à l'âge adulte. Que les données soient diffusées par les parents ou collectées par des plateformes ou des appareils[138], elles peuvent créer une « ombre de données » qui suit les enfants tout au long de leur vie[139]. Cette ombre de données peut commencer avant la naissance; par exemple, lorsque les parents utilisent des services de tests ADN pour connaître la susceptibilité génétique de leurs enfants à certaines maladies[140]. Puisque ces vastes ensembles de données pouvant être stockés indéfiniment, les futurs systèmes d'IA pourraient s'en servir pour faire de nouvelles déductions sur les personnes au fur et à mesure qu'elles deviennent adultes.

Les écoles collectent de plus en plus d'informations sur les élèves, à l'aide de logiciels tiers et d'outils d'analyse à intelligence artificielle. Depuis la pandémie, l'utilisation des applications de gestion des élèves a augmenté de façon exponentielle dans les crèches, les écoles primaires et les écoles secondaires du Canada. Par exemple, on estime que 70 % des écoles primaires utilisent Class Dojo. Sa politique de confidentialité indique qu'elle peut communiquer des données à des fournisseurs de services tiers, notamment Facebook et Google[141].

Les violations de données ont nui aux enfants et aux jeunes au Canada et ailleurs. En 2024, les photos scolaires de 160 élèves de l'Alberta ont été volées lorsque des pirates ont accédé au fournisseur de stockage en nuage d'une société d'annuaires scolaires[142]. Des gangs de rançongiciels ont pris pour cible des écoles publiques américaines, divulguant des données sensibles sur la santé mentale des élèves, les agressions sexuelles et les plaintes pour discrimination[143]. En 2023, des attaques de rançongiciels ont touché des institutions telles que le Sick Kids Hospital

de Toronto[144], Family and Children's Services of Lanark, Leeds and Grenville[145], et le registre et réseau des bons résultats dès la naissance de l'Ontario, où 3,4 millions de dossiers médicaux ont fait l'objet d'une violation[146].

Les entreprises et les ONG détiennent un grand nombre de données sensibles sur les jeunes, qui peuvent faire l'objet d'une violation. Des applications privées de contrôle parental ont été piratées, exposant les données des enfants surveillés[147]. En 2023, TikTok[148], Microsoft[149], et Amazon[150] ont été condamnés à des amendes pour violation de la vie privée des enfants dans diverses juridictions. En tant qu'organisation non gouvernementale, Jeunesse, J'écoute, qui détient la plus grande base de données sur la santé mentale des jeunes au Canada[151], déclare avoir mené une évaluation des incidences sur la vie privée[152] et avoir regroupé et anonymisé ses données[153].

Avenirs

La vente, la circulation et l'analyse opaques des données relatives aux enfants deviendront plus courantes, commenceront beaucoup plus tôt dans la vie et seront utilisées à l'avenir à des fins imprévues. Le nombre de dispositifs de collecte de données avec lesquels les enfants interagissent, à la maison, à l'école et ailleurs, augmentera (voir les idées 7 et 8). Certains de ces appareils pourraient être plus vulnérables aux violations d'informations sensibles[154].

Les technologies de surveillance alimentées par l'IA pourraient devenir plus importantes, mais elles pourraient aussi être détournées. Les parents pourraient se tourner vers des technologies de surveillance alimentées par l'IA pour les aider à contrôler les activités en ligne de leurs enfants et à garder le contrôle sur des environnements informationnels et médiatiques de plus en plus complexes[155]. Cependant, les jeunes pourraient aussi développer des moyens de plus en plus sophistiqués pour échapper au contrôle parental.

Les enfants et les jeunes pourraient vivre dans des environnements médiatiques plus personnalisés. Les contenus de divertissement et la publicité pourraient de plus en plus être générés ou conçus par des compagnons d'IA personnalisés. Les sous-cultures et les cultures partisanes pourraient devenir de plus en plus personnalisées et politisées. Les sentiments d'isolement social pourraient devenir plus fréquents, ainsi qu'une cohésion sociale réduite. Certains

jeunes pourraient être désillusionnés par les technologies invasives alimentées par l'IA et choisir de passer plus de temps hors ligne. Toutefois, compte tenu de l'omniprésence de l'IA, cette option pourrait ne pas être envisageable à l'avenir.

Le marché des données sur les jeunes pourrait devenir plus compétitif à mesure que les préoccupations concernant la confidentialité des données sur les jeunes augmentent. Cela pourrait conduire les entreprises technologiques à développer des moyens plus insidieux d'extraire et de commercialiser les données des jeunes. L'âge limite pour être considéré comme un « enfant » peut varier selon les contextes. Les données pourraient devoir être divulguées lorsque l'enfant atteint l'âge de la majorité[156]. Les technologies de vérification de l'âge[157], comme celles qui sont actuellement utilisées dans certains États américains pour les sites Web pornographiques[158], pourraient être utilisées plus largement pour protéger les jeunes des adultes prédateurs et des espaces réservés aux adultes.

Malgré les nombreuses préoccupations qu'elles soulèvent, les nouvelles technologies basées sur l'IA pourraient également collecter des données de manière à favoriser l'accessibilité[159]. Ils pourraient être utilisés pour développer des outils d'apprentissage individualisés qui aideraient les élèves à progresser à leur propre rythme. Ils pourraient également améliorer la qualité des soins de santé pédiatriques en facilitant le diagnostic, le suivi des patients et la médecine de précision[160].

Implications

- Les enfants d'aujourd'hui pourraient être confrontés à des violations de données **plus fréquentes et plus dévastatrices** tout au long de leur vie
- Ces violations peuvent donner lieu à des formes d'**usurpation d'identité** qui entraînent des pertes financières ou la divulgation d'informations personnelles sensibles
- **La réidentification des données personnelles anonymes** pourrait devenir plus facile à mesure que les violations de données deviennent plus courantes et que les technologies progressent; les données qui semblent **privées aujourd'hui pourraient ne plus l'être demain**

- Des restrictions laxistes pourraient conduire à l'utilisation de données pour faire **des déductions formulées par l'IA** sur les jeunes qui **nuisent à leurs relations et à leur accès à l'emploi, au crédit ou à l'assurance** dans l'enfance et à l'âge adulte
- **L'utilisation accrue des technologies de contrôle parental** pourrait entraîner une surveillance indue et **une perte de vie privée et d'autonomie pour les enfants**
- L'IA pourrait **compliquer l'identification par les parents des contenus problématiques ou préjudiciables**, ou faciliter la dissimulation par les jeunes de leur participation à ces contenus
- Si la prise de conscience des problèmes liés à la confidentialité des données des enfants s'accroît, davantage de développeurs pourraient être amenés à lancer **des applications et des plateformes propres aux enfants, soumises à des normes de confidentialité plus strictes**[161] ou prenant en compte des questions telles que la santé mentale et la dépendance

Idée 10 : L'IA pourrait remodeler nos relations avec les autres

Les outils d'IA pourraient médiatiser davantage les interactions sociales-- dans des contextes publics ou professionnels, ou en privé avec des amis, des membres de la famille ou des partenaires romantiques. Ces outils pourraient être utilisés pour signaler des comportements suspects ou nuisibles et aider à éviter les maladresses sociales-- mais ils pourraient aussi aider à manipuler les autres et à en faire des proies.

Aujourd'hui

L'IA joue déjà un rôle important dans la médiation de nos relations avec les étrangers, les amis et la famille dans les espaces en ligne. Les algorithmes de recommandation agissent comme un filtre social, déterminant quel contenu un utilisateur voit, de quelle personne et dans quel ordre[162]. Ces algorithmes peuvent encourager les utilisateurs à s'engager avec des influenceurs et des créateurs de contenu qui offrent un accès apparent à leur vie[163].Pour certains utilisateurs, ces « intimités » peuvent se transformer en relations parasociales, où les personnes se sentent émotionnellement liées ou attachées à de parfaits inconnus[164].

Les dispositifs d'IA servent de médiateur à un nombre croissant d'interactions professionnelles et personnelles. Par exemple, les médecins utilisent déjà l'IA pour aider à diagnostiquer ou à surveiller les patients[165]. Les gens utilisent l'IA pour rédiger des profils[166] ou des messages[167] sur les applications de rencontre. L'IA peut même analyser et signaler le ton utilisé par les personnes dans les messages qu'elles s'adressent les unes aux autres, par exemple dans les applications utilisées pour faciliter la communication dans les situations de coparentalité difficiles[168].

Les dispositifs portables qui introduisent l'IA dans de nouveaux aspects de notre vie peuvent brouiller les frontières entre les espaces réels et numériques. Ces dispositifs peuvent utiliser la réalité virtuelle (RV), la réalité augmentée (RA) et une combinaison de RA et de RV connue sous le nom de réalité mixte (RM)[169]. La recherche laisse supposer que les environnements immersifs peuvent avoir une incidence émotionnelle plus importante que les espaces en ligne

traditionnels[170]. Les expériences collectives dans la RV peuvent constituer un nouveau type de rassemblement social enrichissant pour des groupes géographiquement éloignés. Les préjudices subis dans la RV, tels que les agressions, peuvent avoir des effets psychologiques semblables à ceux de l'équivalent hors ligne[171].

Les personnes peuvent développer des liens émotionnels avec les compagnons de l'IA. Des millions de personnes se tournent vers des compagnons d'IA pour soulager leur solitude, accéder à une thérapie, obtenir des conseils et nouer des relations amoureuses[172][173][174]. Lorsqu'un modèle d'IA produit des textes, des discours et des images qui ne se distinguent pas de ceux des humains, il est facile d'anthropomorphiser le modèle en attribuant un motif et une intention à ses réponses[175].

Les utilisateurs de plateformes populaires vivent dans des univers de plus en plus personnalisés et privés, l'IA se chargeant de sélectionner les contenus qu'ils voient. Les algorithmes des médias sociaux proposent souvent aux utilisateurs des contenus qui laissent entendre qu'ils les « connaissent » mieux que ne le feraient des amis proches. Toutefois, au fil du temps, la consommation de contenus créés par l'IA, par opposition aux contenus diffusés par des amis, peut déformer les représentations de soi[176]. En faisant défiler le contenu seul, les utilisateurs peuvent entrer dans ce que les chercheurs appellent un état de transe[177].

L'IA modifie les relations des parents avec leurs enfants. Les outils d'IA peuvent offrir aux parents un niveau de visibilité et de contrôle sans précédent sur les applications utilisées par leurs enfants, le contenu qu'ils consomment et les messages qu'ils écrivent, comme l'explique l'idée 7. Les téléphones intelligents ou les traceurs peuvent donner aux parents des informations en temps réel, 24 heures sur 24 et 7 jours sur 7, sur l'endroit où se trouvent leurs enfants[178]. Ces outils peuvent éroder l'autonomie, la vie privée et l'indépendance des enfants au fur et à mesure qu'ils grandissent et mûrissent. Des outils semblables utilisés dans les relations amoureuses peuvent faciliter les comportements abusifs et le harcèlement[179].

Avenirs

À l'avenir, l'IA pourrait jouer un rôle plus important dans la médiation des interactions professionnelles, limitant ainsi les possibilités de nouer de nouvelles amitiés. L'IA pourrait améliorer l'efficacité de la communication entre les clients et les employés d'une entreprise et modifier les flux de travail entre les personnes et les équipes. La culture du lieu de travail pourrait devenir plus impersonnelle, avec moins de possibilités de socialisation.

Les outils d'IA pourraient également servir de médiateurs pour des interactions sociales plus personnelles, même à la maison entre les membres d'une même famille. Ces outils peuvent être des agents d'IA, des algorithmes de plateforme ou des dispositifs portables, tels que des lunettes de réalité augmentée. Plus d'informations et de visibilité sur la vie intérieure des personnes, qu'elle soit physiologique ou psychologique, pourraient devenir normales. Cela pourrait améliorer la communication dans les relations. Elle pourrait également modifier la dynamique des relations, entraînant une diminution de la confiance et de l'autonomie, et une augmentation des problèmes de santé mentale[180].

Les personnes pourraient de plus en plus se tourner vers l'IA pour trouver de la compagnie ou des réponses à leurs problèmes personnels. L'IA pourrait aider les personnes socialement isolées à entrer en contact avec les autres[181]. Les thérapeutes de l'IA pourraient fournir des soins de santé mentale sur mesure aux populations qui n'y ont pas accès : des applications telles que Black Female Therapist, par exemple, utilisent l'IA entraînée pour souligner l'importance du racisme systémique[182]. D'autre part, les compagnons de l'IA pourraient isoler davantage les personnes s'ils remplacent les relations avec les humains. Les personnes qui en viennent à préférer les relations synthétiques aux relations réelles pourraient se retrouver déconnectées de la communauté, même si elles ne sont pas nécessairement seules.

Certaines personnes pourraient chercher à établir des liens humains en échangeant et en comparant leurs flux médiatiques. Les expériences médiatiques devenant de plus en plus personnalisées, il pourrait y avoir un intérêt accru pour la compréhension des mondes distincts que les gens habitent. Il peut s'agir d'une « analyse de flux » dans un cadre thérapeutique, d'un échange de flux en présence d'amis ou même d'événements publics d'échange de flux[183].

À l’avenir, il pourrait devenir impossible de faire la distinction entre les humains et les agents IA hyperréalistes lors des interactions dans les espaces en ligne. La technologie de l’IA pourrait être utilisée pour créer des répliques numériques de proches décédés ou éloignés, ou de célébrités et d’influenceurs. Les agents d’IA pourraient être perçus comme manifestant des émotions humaines telles que l’empathie et l’amour. Les personnes pourraient avoir ce qui ressemble à une relation intime avec une personne, mais qui est en fait une interaction parasociale avec un robot conversationnel. Cela pourrait remplacer entièrement les relations sociales humaines pour certaines personnes vulnérables ou solitaires.

Implications

- L’IA pourrait **contribuer à réduire les inégalités pour les personnes confrontées à des barrières linguistiques** ou à des difficultés à s’orienter dans des interactions sociales complexes
- Les relations avec les compagnons de l’IA pourraient être identiques aux relations humaines, voire plus faciles ou meilleures, pour certaines personnes
- **Les compagnons ou thérapeutes de l’IA pourraient avoir plus d’influence sur les comportements d’une personne** que sa famille ou ses amis proches
- **Les compétences sociales pourraient s’atrophier**. Des compétences telles que l’écoute et l’empathie pourraient être érodées si les utilisatrices et utilisateurs s’appuient trop sur l’aide de l’IA pour les interactions sociales ou personnalisent les agents de l’IA pour qu’ils représentent leurs besoins et leurs préférences
- Les taux de mariage pourraient diminuer et la solitude augmenter
- L’expérience de l’identité pourrait changer. Une autosurveillance plus précoce et plus fréquente, ainsi que l’application d’analyses prédictives aux processus biologiques et mentaux, pourraient déboucher sur **de nouvelles façons de comprendre et d’optimiser le soi**

- **De nouvelles formes d’abus et de criminalité virtuelle pourraient voir le jour** et remettre en question les définitions de l’agression et du harcèlement
- **Les prédateurs pourraient plus facilement gagner la confiance des enfants et des adultes**, ce qui augmenterait le risque de fraude, de harcèlement ou d’autres abus
- **L’utilisation d’outils d’IA pour communiquer avec les gens pourrait modifier la langue** au fil du temps, potentiellement dans le sens d’une plus grande homogénéisation et d’une stérilisation
- **Les outils d’IA pourraient signaler les comportements suspects**, signaler les abus au moment où ils se produisent et aider les personnes à s’orienter dans des relations toxiques ou dangereuses
- **Les brimades et le harcèlement pourraient devenir plus omniprésents** et nuire à la santé mentale s’ils se produisent dans des environnements immersifs réalistes ou avec l’utilisation de l’IA générative

Publication des travaux d’Horizons de politiques Canada relatifs à l’IA :

Perturbations à l’horizon

Les avenirs de la création de sens : une crise de la certitude

Vies futures : Incertitudes

Commercialisation des données biologiques

Métavers

L’avenir de l’IA générative

La géotechnomie

Trois façons dont ChatGPT pourrait soutenir la prospective stratégique

Remerciements

Ce rapport résume les réflexions, les idées et les analyses de nombreux contributeurs et contributrices dans le cadre de recherches, d'ateliers, et de conversations.

Horizons de politiques Canada souhaite remercier les membres du comité directeur des sous-ministres et le sous-ministre adjoint principal, Elisha Ram, pour leurs conseils, leur soutien et leur expertise, ainsi que tous les collègues qui ont contribué à l'élaboration de ce travail.

Horizons de politiques Canada souhaite également remercier les nombreux expert.e.s qui ont généreusement donné de leur temps pour soutenir ce travail, y compris celles et ceux qui ont choisi de demeurer anonymes :

Blair Attard-Frost
Chargé de cours, Université de Toronto

Stephanie Baker
Chercheuse, Systèmes électroniques et ingénierie IdO, Université James Cook

Michael Beauvais
Candidat au doctorat en droit, Faculté de droit de l’Université de Toronto

Olivier Blais
Cofondateur et vice-président de Decision Science, Moov AI

Ana Brandesescu
Candidate au doctorat, Université McGill

Francesca Campolongo
Directrice de la transformation numérique et des données, Commission européenne.

Ashley Chisholm
Conseillère en politique stratégique, Culture médicale et bien-être des médecins, Association médicale canadienne

Sherif Elsayed-Ali
Co-fondateur, Nexus Climate

Kay Firth-Butterfield
Directeur général, Good Tech Advisory LLC

Michael Geist
Professeur titulaire, Section de common law, Faculté de droit, Chaire de recherche du Canada en droit de l'Internet et du commerce électronique, Université d'Ottawa

N. Katherine Hayles
Professeure de recherche distinguée à l'Université de Californie, Los Angeles, et professeure émérite James B. Duke à l'Université de Duke.

Matissa Hollister
Professeure adjointe (enseignement), comportement organisationnel, École de gestion Desautels, Université McGill

Sun-Ha Hong
Professeur adjoint, École de communication, Université Simon Fraser

Kai-Hsin Hung
Candidat au doctorat, HEC Montréal

Ian Scott Kalman
Professeur associé, Université Fulbright du Vietnam

Andrew J. Kao
Chercheur, Université de Harvard

Sayash Kapoor
Candidat au doctorat, Université de Princeton

Kristin Kozar
Directeur général, Centre d'histoire et de dialogue sur les pensionnats autochtones, Université de la Colombie-Britannique

Nicholas Lane
Professeur d'informatique et de technologie, Université de Cambridge

Sasha Luccioni
Responsable climatique, Hugging Face

Arvind Narayanan
Directeur/professeur, Centre pour la politique en matière de technologies de l'information, Université de Princeton

David Nielson
Directeur du laboratoire de réalité mixte, Institut des technologies créatives de l'USC

Deval Pandya
Vice-président de l'ingénierie de l'IA, Vector Institute

Manish Raghavan
Professeur de développement de carrière Drew Houston (2005) et professeur adjoint de technologie de l'information à la Sloan School of Management du MIT

Mark Riedl
Professeur/directeur associé, Georgia Tech, School of Interactive Computing/Machine Learning Center

Julie Robillard
Professeure associée de neurologie, Université de la Colombie-Britannique

Stephen Sanford
Directeur général, U.S. Government Accountability Office

Teresa Scassa
Membre du corps professoral et Chaire de recherche du Canada en droit et politique de l'information, professeur titulaire, Section de common law, Faculté de droit

Mona Sloane
Professeure adjointe de science des données et d'études des médias, Université de Virginie

Nick Srnicek
Maître de conférences en économie numérique au département des sciences humaines numériques du Kings College de Londres

Luke Stark
Professeur adjoint, Université de Western Ontario

Yuan Stevens
Associé académique – Recherche en santé et gouvernance de l'IA, Centre pour la génomique et la politique

Catherine Stinson
Boursière nationale du Queen's en implications philosophiques de l'intelligence artificielle et professeure adjointe au département de philosophie et à l'école d'informatique de l'Université Queen's.

Mark Surman
Président et directeur général de la Fondation Mozilla

Liana Tang
Deuxième directeur, Smart Nation Strategy Office, ministère des communications et de l'information, Singapour

Agnes Venema
Chercheuse à l'Académie Nationale de Renseignements 'Mihai Viteazul', Ministère de la Défense, Roumanie

Wendy Wong
Professeure et chaire de recherche principale, Université de la Colombie-Britannique

Agnieszka Wykowska
Chercheure principale titularisée et chercheure principale, cognition sociale dans l’interaction humain-robot, Institut italien de technologie

Un remerciement spécial est adressé à l'équipe du projet :

John Beasy, analyste
Martin Berry, analyste principal
Leah Desjardins, analyste
Miriam Havelin, analyste
Nicole Rigillo, analyste principale
Kristel Van der Elst, directrice générale
Claire Woodside, gestionnaire

Et aux collègues actuel.le.s et ancien.ne.s d'Horizons de politiques suivant.e.s : Katherine Antal, Imran Arshad, Marcus Ballinger, Fannie Bigras-Lafrance, Mélissa Chiasson, Steffen Christensen, Suesan Danesh, Pierre-Olivier Desmarchais, Nicole Fournier-Sylvester, Chris Hagerman, Laura Gauvreau, Pascale Louis-Miron, Leona Nikolic, Megan Pickup, Simon Robertson, Julie-Anne Turner, Alexa Van Every, et Andrew Wright (externe) pour leur soutien dans ce projet.

Notes de fin de texte

[1] Pretz, Kathy. « Stop Calling Everything AI, Machine-Learning Pioneer Says ». *IEEE Spectrum*. Consulté le 8 août 2024. https://spectrum.ieee.org/stop-calling-everything-ai-machinelearning-pioneer-says.
[2] Lanier, Jaron. « There Is No A.I. » *The New Yorker*, 20 avril 2023. https://www.newyorker.com/science/annals-of-artificial-intelligence/there-is-no-ai.
[3] Drage, Eleanor, et Kerry Mackereth. « The Good Robot Podcast: Featuring Emily M. Bender and Alex Hanna ». Consulté le 8 août 2024. https://aihub.org/2024/02/09/the-good-robot-podcast-featuring-emily-m-bender-and-alex-hanna/.
[4] Instruments juridiques de l'OCDE. « Recommandation du Conseil sur l'intelligence artificielle », 2 mai 2024. https://legalinstruments.oecd.org/fr/instruments/OECD-LEGAL-0449.
[5] Lenhart, Amanda. « Teens, Technology and Friendships ». *Pew Research Center* (blogue), 6 août 2015. https://www.pewresearch.org/internet/2015/08/06/teens-technology-and-friendships/.
[6] Kaiser & associés. « Les jeunes Canadiens développent une confiance croissante à l'égard des informations largement partagées sur les médias sociaux », 15 novembre 2023. https://kaiserpartners.com/fr/les-jeunes-canadiens-developpent-une-confiance-croissante-a-legard-des-informations-largement-partagees-sur-les-medias-sociaux/.
[7] Statistique Canada. « Le quotidien - Enquête canadienne sur l'utilisation d'Internet, 2022», 20 juillet 2023. https://www150.statcan.gc.ca/n1/daily-quotidien/230720/dq230720b-fra.htm.
[8] Association des banquiers canadiens. « Fiche info - Les Canadiens et leurs activités bancaires », 31 mars 2022. https://cba.ca/technology-and-banking?l=fr.
[9] International Trade Administration. « Canada - Country Commercial Guide - eCommerce », 4 novembre 2023. https://www.trade.gov/country-commercial-guides/canada-ecommerce.
[10] Statistique Canada. « Le quotidien - Enquête canadienne sur l'utilisation d'Internet, 2022», 20 juillet 2023. https://www150.statcan.gc.ca/n1/daily-quotidien/230720/dq230720b-fra.htm.
[11] Doctorow, Cory. « As Platforms Decay, Let's Put Users First ». Electronic Frontier Foundation, 9 mai 2023. https://www.eff.org/deeplinks/2023/04/platforms-decay-lets-put-users-first.
[12] « Global Internet Phenomena ». Sandvine, mars 2024. https://www.sandvine.com/phenomena, p. 8 et 9.
[13] StatCounter Global Stats. « Search Engine Market Share Worldwide ». Consulté le 23 mai 2024. https://gs.statcounter.com/search-engine-market-share.
[14] Alexander, Julia. « Creators Finally Know How Much Money YouTube Makes, and They Want More of It ». The Verge, 4 février 2020. https://www.theverge.com/2020/2/4/21121370/youtube-advertising-revenue-creators-demonetization-earnings-google.

[15] Ball, James. « Big Tech Can't Escape the Ad Business ». The Atlantic, 1er juin 2023. https://www.theatlantic.com/technology/archive/2023/06/advertising-revenue-google-meta-amazon-apple-microsoft/674258/.
[16] Acumen Research and Consulting. « Search Engine Optimization Services - Global Market and Forecast Till 2030 », février 2023. https://www.acumenresearchandconsulting.com/search-engine-optimization-services-market.
[17] Buckley, Thomas, Shaw, Lucas et Ghaffary, Shirin. « OpenAI Courts Hollywood in Meetings With Film Studios, Directors ». *Bloomberg.Com*, 22 mars 2024. https://www.bloomberg.com/news/articles/2024-03-22/openai-courts-hollywood-in-meetings-with-film-studios-directors.
[18] Pasion, Lorenz. « Artists Fear Lack of Job Security, Regulations as AI-Made Song Covers Go Viral in TikTok ». *RAPPLER* (blogue), 11 septembre 2023. https://www.rappler.com/technology/features/artists-fear-lack-job-security-regulations-ai-generated-song-covers-viral-tiktok/.
[19] Statistique Canada. « Le quotidien - Enquête canadienne sur l'utilisation d'Internet, 2022», 20 juillet 2023. https://www150.statcan.gc.ca/n1/daily-quotidien/230720/dq230720b-fra.htm.
[20] Zandt, Florian. « Infographic: How Dangerous Are Deepfakes and Other AI-Powered Fraud? » Statista Daily Data, 13 mars 2024. https://www.statista.com/chart/31901/countries-per-region-with-biggest-increases-in-deepfake-specific-fraud-cases.
[21] Al-Sibai, Noor. « Bone-Chilling AI Scam Fakes Your Loved Ones'Voices to Demand Hostage Ransom ». Futurism, 9 mars 2024. https://futurism.com/the-byte/ai-voice-hostage-scam.
[22] Magramo, Kathleen. « British Engineering Giant Arup Revealed as $25 Million Deepfake Scam Victim | CNN Business ». CNN, 17 mai 2024. https://www.cnn.com/2024/05/16/tech/arup-deepfake-scam-loss-hong-kong-intl-hnk/index.html.
[23] Kreps, Sarah et Li, Richard. « Cascading Chaos: Nonstate Actors and AI on the Battlefield ». Brookings, 1er février 2022. https://www.brookings.edu/articles/cascading-chaos-nonstate-actors-and-ai-on-the-battlefield/.
[24] Scroxton, Alex. « Research Team Tricks AI Chatbots into Writing Usable Malicious Code ». Computer Weekly, 24 octobre 2023. https://www.computerweekly.com/news/366556692/Research-team-tricks-AI-chatbots-into-writing-usable-malicious-code.
[25] Huggingface. « LMSys Chatbot Arena Leaderboard - a Hugging Face Space by Lmsys ». Consulté le 19 février 2024. https://huggingface.co/spaces/lmsys/chatbot-arena-leaderboard.
[26] Hirsh, Michael. « How AI Will Revolutionize Warfare ». Foreign Policy (blogue), 11 avril 2023. https://foreignpolicy.com/2023/04/11/ai-arms-race-artificial-intelligence-chatgpt-military-technology/.
[27] Homeland Security. « Addressing Risks From Non-State Actors' Use of Commercially Available Technologies », 2022. https://www.dhs.gov/sites/default/files/2022-09/Addressing%20Risks%20from%20Non-State%20Actors.pdf.

[28] Mascellino, Alessandro. « ChatGPT Creates Polymorphic Malware ». Infosecurity Magazine, 18 janvier 2023. https://www.infosecurity-magazine.com/news/chatgpt-creates-polymorphic-malware/.
[29] Volpe, Tristan. « Dual-Use Distinguishability: How 3D-Printing Shapes the Security Dilemma for Nuclear Programs ». Fondation Carnegie pour la paix internationale. Consulté le 19 février 2024. https://carnegieendowment.org/2019/08/22/dual-use-distinguishability-how-3d-printing-shapes-security-dilemma-for-nuclear-programs-pub-79910.
[30] Ware, Jacob. « Terrorist Groups, Artificial Intelligence, and Killer Drones ». War on the Rocks, 24 septembre 2019. https://warontherocks.com/2019/09/terrorist-groups-artificial-intelligence-and-killer-drones/.
[31] « Algorithms and Terrorism: The Malicious Use of Artificial Intelligence for Terrorist Purposes ». Centre des Nations Unies pour la lutte contre le terrorisme et Institut interrégional de recherche des Nations unies sur la criminalité et la justice, 2021. https://unicri.it/sites/default/files/2021-06/Malicious%20Use%20of%20AI%20-%20UNCCT-UNICRI%20Report_Web.pdf.
[32] Kreps, Sarah. « Democratizing Harm: Artificial Intelligence in The Hands of Nonstate Actors ». Brookings, novembre 2021. https://www.brookings.edu/wp-content/uploads/2021/11/FP_20211122_ai_nonstate_actors_kreps.pdf.
[33] Gillespie, Nicole; Steven, Lockey; Caitlin Curtis; Javad Pool et Ali Akbari. « Trust in Artificial Intelligence: A Global Study ». The University of Queensland and KPMG Australia, 2023. https://www.aiunplugged.io/wp-content/uploads/2023/10/Trust-in-Artificial-Intelligence.pdf.
[34] Proof Strategies. « L'Indice CanTrust 2024 ». Proof Strategies, 13 février 2024. https://getproof.com/trust/cantrust/.
[35] Carmichael, Matt et Jamie Stinson. "The Ipsos AI Monitor 2024: Changing Attitudes and Feelings about AI and the Future It Will Bring" Ipsos AI Monitor. Paris : Ipsos, 6 juin 2024. https://www.ipsos.com/en/ipsos-ai-monitor-2024-changing-attitudes-and-feelings-about-ai-and-future-it-will-bring.
[36] Karadeglija, Anja. « Poll Finds More Canadians Using AI despite 'deep-Rooted' Fears ». Le National Post, 9 février 2024. https://nationalpost.com/news/more-canadians-using-ai-tools-despite-deep-rooted-fears-about-the-tech-poll.
[37] Chhim, Chris et Sanyam Sethi. « Canadians Among Least Likely to Believe Artificial Intelligence Will Make Their Lives Better ». Paris : Ipsos, 14 janvier 2022. https://www.ipsos.com/en-ca/news-polls/Canadians-Least-Likely-AI-Make-Lives-Better.
[38] Stilgoe, Jack. « What Does It Mean to Trust a Technology? » *Science*, vol.382, nº 6676 (s.d.) : 9782.
[39] « Utilisation d'outils d'IA ». Léger, 5 février 2024. https://leger360.com/use-of-ai-tools/.
[40] L'Observatoire OCDE des politiques de l'IA. « AIM: The OECD AI Incidents Monitor, an Evidence Base for Trustworthy AI ». Consulté le 7 août 2024. https://oecd.ai/en/incidents.

[41] « Bienvenue dans la base de données des incidents d'IA ». Consulté le 7 août 2024. https://incidentdatabase.ai/fr/.
[42] O'Gorman, Marcel. « Opinion: When It Comes to AI, It Feels like We're Doing More Adaptation than Adoption These Days. This Is Not a Good Feeling ». *The Globe and Mail*, 21 juin 2024. https://www.theglobeandmail.com/opinion/article-when-it-comes-to-ai-it-feels-like-were-doing-more-adaptation-than/.
[43] Marr, Bernard. « Will AI Really Revolutionize Every Industry? A Critical Analysis ». Forbes, 23 juillet 2024. https://www.forbes.com/sites/bernardmarr/2024/07/23/will-ai-really-revolutionize-every-industry-a-critical-analysis/.
[44] Vermes, Jason. « Airports want to scan your face to make travelling easier. Privacy experts caution it's not ready for takeoff ». *CBC News.* 3 mars 2024. Airports want to scan your face to make travelling easier. Privacy experts caution it's not ready for takeoff | CBC Radio
[45] Schuman, Leslie. « Leading Corporations Introduce Data Provenance Standards ». *Businesswire*, 30 novembre 2023. https://www.businesswire.com/news/home/20231130851266/en/Leading-Corporations-Introduce-Data-Provenance-Standards.
[46] The Data & Trust Alliance. « The Data & Trust Alliance ». Consulté le 7 août 2024. https://dataandtrustalliance.org/.
[47] « LLM Safety Leaderboard - a Hugging Face Space by AI-Secure ». Consulté le 7 août 2024. https://huggingface.co/spaces/AI-Secure/llm-trustworthy-leaderboard.
[48] Masse, Bryson. « Armilla Offers Verification and Warranties for Enterprises Using AI Models ». *VentureBeat* (blogue), 3 octobre 2023. https://venturebeat.com/ai/armilla-offers-verification-and-warranties-for-enterprises-using-ai-models/.
[49] Sanz Sáiz, Beatriz. « How your organization can have confidence in the opportunities AI brings ». *Ernst & Young* (blogue), 15 janvier 2024. https://www.ey.com/en_gl/insights/ai/how-your-organization-can-have-confidence-in-the-opportunities-ai-brings
[50] Gillespie, Nicole; Lockey, Steven; Curtis, Caitlin; Pool, Javad et Ali Akbari. 'Trust in Artificial Intelligence: A global study'. *The University of Queensland* et *KPMG*. (2023). https://doi.org/10.14264/00d3c94.
[51] Horowitz, Michael C., Lauren Kahn, Julia Macdonald et Jacquelyn Schneider. « Adopting AI: How Familiarity Breeds Both Trust and Contempt ». *AI & SOCIETY*, 12 mai 2023. https://doi.org/10.1007/s00146-023-01666-5.
[52] Schneider, Michael. « SXSW Audiences Loudly Boo Festival Videos Touting the Virtues of AI ». Variety, 13 mars 2024. https://variety.com/2024/tv/news/sxsw-audiences-boo-videos-artificial-intelligence-ai-1235940454/ .
[53] Perrigo, Billy. « AI Poses Extinction-Level Risk, State-Funded Report Says ». *Time Magazine*, 11 mars 2024. https://time.com/6898967/ai-extinction-national-security-risks-report/.

[54] Personnel de la Presse canadienne. « Parti Quebecois Government to Close Gentilly-2 Nuclear Power Plant | Globalnews.Ca ». *Global News*, 12 septembre 2012. https://globalnews.ca/news/285716/parti-quebecois-government-to-close-gentilly-2-nuclear-power-plant/.
[55] Huet, Ellen. « AI Certification Program Verifies Systems Are 'Fairly Trained' - Bloomberg ». *Bloomberg*. Consulté le 13 février 2024. https://www.bloomberg.com/news/articles/2024-01-17/ai-certification-program-verifies-systems-are-fairly-trained.
[56] Helhoski , Anna. « AI Could Prevent Hiring Bias - Unless It Makes It Worse ». NerdWallet, 12 juin 2023. https://www.nerdwallet.com/article/finance/ai-hiring-decisions.
[57] Rhea, Alene K.; Markey, Kelsey; D'Arinzo, Lauren; Schellmann, Hilke; Sloane, Mona; Squires, Paul; Arif Khan, Falaah et Stoyanovich, Julia. « An external stability audit framework to test the validity of personality prediction in AI hiring ». *Data Mining and Knowledge Discovery*, vol. 36, nº 6 (2022) : p. 2153-2193. https://link.springer.com/article/10.1007/s10618-022-00861-0.
[58] Andrew, Lori et Hannah Bucher. « Automating Discrimination: AI Hiring Practices and Gender Inequality ». Cardozo Law Review. Consulté le 15 août 2024. https://cardozolawreview.com/automating-discrimination-ai-hiring-practices-and-gender-inequality/.
[59] Brown, Lydia; Shetty, Ridhi et Richardson, Michelle. « Algorithm-Driven Hiring Tools: Innovative Recruitment or Expedited Disability Discrimination? », 3 décembre 2020. https://cdt.org/insights/report-algorithm-driven-hiring-tools-innovative-recruitment-or-expedited-disability-discrimination/.
[60] Samuel, Sigal. « Why It's so Damn Hard to Make AI Fair and Unbiased ». Vox, 19 avril 2022. https://www.vox.com/future-perfect/22916602/ai-bias-fairness-tradeoffs-artificial-intelligence.
[61] Ferrara, Emilio. « Éliminer les biais dans l'IA peut être impossible – un informaticien explique comment l'apprivoiser à la place ». La Conversation, 19 juillet 2023. http://theconversation.com/eliminating-bias-in-ai-may-be-impossible-a-computer-scientist-explains-how-to-tame-it-instead-208611.
[62] Kleinberg, Jon. « Inherent Trade-Offs in Algorithmic Fairness ». 10 avril 2018. https://www.youtube.com/watch?v=p5yY2MyTJXA&list=TLPQMjkxMjIwMjOfavJbOg-0WQ&index=2.
[63] Dwork, Cynthia. « The Emerging Theory of Algorithmic Fairness ». 6 septembre 2018. https://www.youtube.com/watch?v=g-z84_nRQhw.
[64] Raghavan, Manish. « What Should We Do When Our Ideas of Fairness Conflict? » *Communications of the ACM*, vol. 67, nº 1 (janvier 2024) : p. 88-97. https://doi.org/10.1145/3587930.
[65] Shomik, Jain; Suriyakumar, Vinith; Creel, Kathleen et Wilson, Ashia. « Algorithmic Pluralism: A Structural Approach To Equal Opportunity ». arXiv, 21 septembre 2023. https://doi.org/10.48550/arXiv.2305.08157.

[66] Feathers, Todd. « Texas A&M Drops “Race” from Student Risk Algorithm Following Markup Investigation ». The Markup, 30 mars 2021. https://themarkup.org/machine-learning/2021/03/30/texas-am-drops-race-from-student-risk-algorithm-following-markup-investigation.
[67] PBS News. « AP Report: DOJ Examining AI Screening Tool Used by Pa. Child Welfare Agency ». PBS NewsHour, 31 janvier 2023. https://www.pbs.org/newshour/nation/ap-report-doj-examining-ai-screening-tool-used-by-pa-child-welfare-agency.
[68] Wang, Angelina; Kapoor, Sayash; Barocas, Solon et Narayanan, Arvind. « Against Predictive Optimization: On the Legitimacy of Decision-Making Algorithms That Optimize Predictive Accuracy ». SSRN Scholarly Paper. Rochester (New York) 4 octobre 2022. https://papers.ssrn.com/abstract=4238015.
[69] Wang, Angelina; Kapoor, Sayash; Barocas, Solon et Narayanan, Arvind. « Against Predictive Optimization: On the Legitimacy of Decision-Making Algorithms That Optimize Predictive Accuracy ». SSRN Scholarly Paper. Rochester (New York) 4 octobre 2022. https://papers.ssrn.com/abstract=4238015.
[70] Knox, Dean; Lowe, Will et Mummolo, Jonathan. « Administrative Records Mask Racially Biased Policing ». *American Political Science Review*, vol. 114, nº 3 (août 2020) : p. 619-637. https://doi.org/10.1017/S0003055420000039.
[71] Wang, Angelina; Kapoor, Sayash; Barocas, Solon et Narayanan, Arvind. « Against Predictive Optimization: On the Legitimacy of Decision-Making Algorithms That Optimize Predictive Accuracy ». SSRN Scholarly Paper. Rochester (New York) 4 octobre 2022. https://papers.ssrn.com/abstract=4238015.
[72] Fishbane, Alissa; Ouss, Aurelie et Shah, Anuj K. ‘Behavioral Nudges Reduce Failure to Appear for Court’. *Science*, vol. 370, nº 6517 (6 novembre 2020) : eabb6591. https://doi.org/10.1126/science.abb6591.
[73] Brenninkmeijer, Alex et Seldam, Björn ten. « The Dutch Benefits Scandal: A Cautionary Tale for Algorithmic Enforcement’ ». Application du droit de l’UE (blogue), 30 avril 2021. https://eulawenforcement.com/?p=7941.
[74] Rhea, Alene K.; Markey, Kelsey; D’Arinzo, Lauren; Schellmann, Hilke; Sloane, Mona; Squires, Paul; Arif Khan, Falaah et Stoyanovich, Julia. « An external stability audit framework to test the validity of personality prediction in AI hiring ». *Data Mining and Knowledge Discovery*, vol. 36, nº 6 (2022) : p. 2153-2193. https://link.springer.com/article/10.1007/s10618-022-00861-0.
[75] Hou, Jilei. « Quantization: What It Is & How It Impacts AI ». *Qualcomm* (blogue), 11 mars 2019. https://www.qualcomm.com/news/onq/2019/03/heres-why-quantization-matters-ai.
[76] Edwards, Benj. « Microsoft’s Phi-3 Shows the Surprising Power of Small, Locally Run AI Language Models ». Ars Technica, 23 avril 2024. https://arstechnica.com/information-technology/2024/04/microsofts-phi-3-shows-the-surprising-power-of-small-locally-run-ai-language-models/.

[77] Ramlochan, Sunil. « How Does Llama-2 Compare to GPT-4/3.5 and Other AI Language Models ». Prompt Engineering Institute, 1er septembre 2023. https://promptengineering.org/how-does-llama-2-compare-to-gpt-and-other-ai-language-models/.
[78] Ali Awan, Abid. « Running Mixtral 8x7b On Google Colab For Free ». KDnuggets, 12 janvier 2024. https://www.kdnuggets.com/running-mixtral-8x7b-on-google-colab-for-free.
[79] Dunn, Caroline. « How to Train Your Raspberry Pi for Facial Recognition ». Tom's Hardware, 17 septembre 2022. https://www.tomshardware.com/how-to/raspberry-pi-facial-recognition.
[80] Mearian, Lucas. « GenAI Is Moving to Your Smartphone, PC and Car - Here's Why ». Computerworld, 30 janvier 2024. https://www.computerworld.com/article/3712601/genai-is-moving-to-your-smartphone-pc-and-car-heres-why.html.
[81] Irwin, Kate. « Venice's Privacy-Focused AI Chatbot Won't Store Your Data, Judge Your Questions ». PC Magazine, 9 août 2024. https://www.pcmag.com/news/venices-privacy-focused-ai-chatbot-wont-store-your-data-judge-your-questions.
[82] La Maison-Blanche. « Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence », 30 octobre 2023. http://web.archive.org/web/20250115010337/https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/.
[83] Thiel, David. « Identifying and Eliminating CSAM in Generative ML Training Data and Models », 2023. https://doi.org/10.25740/KH752SM9123.
[84] Pham, Nguyen. « Open Source Tools as an Opportunity for SMEs to Use AI? » foojay, 2 juin 2021. https://foojay.io/today/open-source-tools-as-an-opportunity-for-smes-to-use-ai/.
[85] Apple. « Bâtir un écosystème fiable pour des millions d'apps », juin 2021. https://www.apple.com/ca/fr/privacy/docs/Building_a_Trusted_Ecosystem_for_Millions_of_Apps.pdf.
[86] Cotnam, Hallie. « Meet the Robotic Cat Serving Diners at a Gatineau Restaurant ». *CBC News*, 27 octobre 2021. https://www.cbc.ca/news/canada/ottawa/robot-cat-gatineau-restaurant-1.6224125.
[87] Shibu, Sherin. « Sam's Club achète des centaines de robots laveurs de sols autonomes pour ses magasins ». *PC Magazine*, 22 octobre 2020. https://www.pcmag.com/news/sams-club-buys-hundreds-of-autonomous-floor-scrubbing-robots-for-its-stores.
[88] Kucharski, Kyle. « Meta's Ray-Ban Smart Glasses Just Got Another Useful Feature for Free (and a New Style) ». *ZDNET*, 23 avril 2024. https://www.zdnet.com/article/metas-ray-ban-smart-glasses-just-got-another-useful-feature-for-free-and-a-new-style/.
[89] Zia, Dr. Tehseen. « Embodied AI: How It Bridges the Gap Between Mind and Matter ». *Techopedia* (blogue), 12 septembre 2023. https://www.techopedia.com/embodied-ai-bridging-the-gap-between-mind-and-matter.
[90] Diligent Robotics. « Moxi ». Consulté le 23 mai 2024. https://www.diligentrobots.com/moxi.

[91] Berruti, Federico. « An Executive Primer on Artificial General Intelligence ». *McKinsey*, 29 avril 2020. https://www.mckinsey.com/capabilities/operations/our-insights/an-executive-primer-on-artificial-general-intelligence.
[92] Teisceira-Lessard, Philippe. « Ambulances : Des robots pour ramener des patients à la vie ». *La Presse*, 15 mars 2024. https://www.lapresse.ca/actualites/grand-montreal/2024-03-15/ambulances/des-robots-pour-ramener-des-patients-a-la-vie.php.
[93] Dent, Steve. 'Urtopia's Fusion e-Bike Has Fully Integrated ChatGPT'. *Engadget*, 10 janvier 2024. https://www.engadget.com/urtopias-fusion-e-bike-has-fully-integrated-chatgpt-144429572.html.
[94] Geschwindt, Siôn. « These AI Binoculars Just Made Birdwatching a Whole Lot Easier ». *TNW | Deep-Tech*, 12 janvier 2024. https://thenextweb.com/news/smart-ai-binoculars-birdwatching.
[95] Martin, Diana. « Autonomous Tractor Retrofit Arrives in Canada ». *Alberta Farmer Express* (blog), 27 octobre 2023. https://www.albertafarmexpress.ca/news/autonomous-tractor-retrofit-arrives-in-canada/.
[96] « I Already Have Security Cameras Installed. How Can I Add Face Recognition? » *LinkSprite*, 6 mars 2019. http://www.linksprite.com/i-already-have-security-cameras-installed-how-can-i-add-face-recognition/.
[97] « The How's and Why's of IoT ». *Sogeti*, Consulté le 23 mai 2024. http://web.archive.org/web/20230322222427/https://www.sogeti.com/globalassets/global/downloads/the-hows-and-whys-of-iot-adoption.pdf.
[98] Shirer, Michael. « IDC Forecasts Revenue for Artificial Intelligence Software Will Reach $307 Billion Worldwide in 2027 ». *IDC*, 31 octobre 2023. https://www.idc.com/getdoc.jsp?containerId=prUS51345023.
[99] Villalobos, Pablo; Sevilla, Jaime; Heim, Lennart; Besiroglu, Tamay; Hobbhahn, Marius et Ho, Anson. « Allons-nous manquer de données? Une analyse des limites de la mise à l'échelle des ensembles de données dans l'apprentissage automatique ». *arXiv*, 25 octobre 2022. https://doi.org/10.48550/arXiv.2211.04325.
[100] Tuohy, Jennifer Pattison. « More Ring Camera and Alarm Features Will Soon Require Subscriptions ». *The Verge*, 3 mars 2023. https://www.theverge.com/2023/3/3/23623523/ring-alarm-camera-features-subscription.
[101] Rivers, Stephen. « $4,500 Bill To Unlock Extra Battery Capacity Has People Taking Sides Between Tesla And Customer ». *Carscoops*, 9 juillet 2023. https://www.carscoops.com/2023/07/4500-bill-to-unlock-extra-battery-capacity-has-people-taking-sides-between-tesla-and-customer/.
[102] Charlton, Alistair. « Mercedes Wants To Charge $1,200 Subscription To Unlock Quicker EV Performance ». *Forbes*. Consulté le 23 mai 2024. https://www.forbes.com/sites/alistaircharlton/2022/11/24/mercedes-wants-to-charge-1200-subscription-to-unlock-quicker-ev-performance/.

[103] Dharamshi, Alannah et Lipsey, Adrienne. « Exercising Privacy: Policy Options for Privacy and Wellness Wearables ». *Groupe CSA*, février 2002. https://www.csagroup.org/wp-content/uploads/CSA-Group-Research-Exercising-Privacy-Policy-Options-Privacy-for-Wellness-Wearables.pdf.
[104] Diaz, Maria. « Aqara Just Launched a Smart Home Presence Sensor with Fall Detection ». *ZDNET*, 21 avril 2024. https://www.zdnet.com/home-and-office/smart-home/aqara-just-launched-a-smart-home-presence-sensor-with-fall-detection/.
[105] Alini, Erica. « Insurance Apps That Track Your Driving Could Now Yield Premium Increases - National ». *Global News*, 21 mars 2021. https://globalnews.ca/news/7704732/auto-insurance-app-usage-based-insurance-surcharges-canada/.
[106] Chowdhary, Krishi. « Meta's New AI-Enabled Ray-Ban Raises Privacy Concerns ». *Tom's Guide*, 5 janvier 2024. https://www.tomsguide.com/news/metas-new-ai-enabled-ray-ban-raises-privacy-concerns.
[107] Singer, Natasha. « In Screening for Suicide Risk, Facebook Takes On Tricky Public Health Role ». *The New York Times*, 31 décembre 2018, sec. Technology. https://www.nytimes.com/2018/12/31/technology/facebook-suicide-screening-algorithm.html
[108] Kröger, Jacob Leon; Raschke, Philip; Percy Campbell, Jessica et Ullrich, Stefan. « Surveilling the Gamers: Privacy Impacts of the Video Game Industry ». *Entertainment Computing*, vol. 44 (1er janvier 2023) : 100537. https://doi.org/10.1016/j.entcom.2022.100537.
[109] Chen, Brian X. « Comment la nouvelle caméra faciale de Meta annonce une nouvelle ère de surveillance ». *The New York Times*, 13 décembre 2023, sec. Technology. https://www.nytimes.com/2023/12/13/technology/personaltech/meta-ray-ban-glasses.html.
[110] Brianna R. « Humane AI: Privacy Implications of This New AI-Powered Lapel ». *Medium* (blogue), 22 décembre 2023. https://medium.com/@cyber-news/humane-ai-privacy-implications-of-this-new-ai-powered-lapel-c1ac377fe630.
[111] While, Jenny; Farrell, Jessica Alice et La Conversation. « The DNA You Shed Could Identify You ». Scientific American, 15 mai 2023. https://www.scientificamerican.com/article/the-dna-you-shed-could-identify-you/
[112] Whitmore, Liam; McCauley, Mark; Farrell, Jessica A.; Stammnitz, Maximilian R.; Koda, Samantha A.; Mashkour, Narges; Summers, Victoria; Osborne, Todd; Whilde, Jenny et Duffy, David J. "Inadvertent Human Genomic Bycatch and Intentional Capture Raise Beneficial Applications and Ethical Concerns with Environmental DNA." *Nature Ecology & Evolution*, vol. 7, nº 6 (juin 2023) : p. 873-888. https://doi.org/10.1038/s41559-023-02056-2.
[113] Caltrider, Jen; Rykov, Misha et MacDonald, Zoe. « Les voitures sont la pire catégorie officielle de produits en matière de confidentialité que nous ayons jamais examinée » *Confidentialité non incluse de Mozilla* (blogue), 6 septembre 2023.

https://foundation.mozilla.org/en/privacynotincluded/articles/its-official-cars-are-the-worst-product-category-we-have-ever-reviewed-for-privacy/.

[114] Caltrider, Jen; Rykov, Misha et MacDonald, Zoe. « Les voitures sont la pire catégorie officielle de produits en matière de confidentialité que nous ayons jamais examinée » *Confidentialité non incluse de Mozilla* (blogue), 6 septembre 2023. https://foundation.mozilla.org/en/privacynotincluded/articles/its-official-cars-are-the-worst-product-category-we-have-ever-reviewed-for-privacy/.

[115] McKinsey. « Unlocking Connected Cars with Corporate Business Building », 31 août 2023. https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/corporate-business-building-to-unlock-value-in-automotive-connectivity.

[116] Rimol, Meghan. « Gartner Identifies Top Five Trends in Privacy Through 2024 ». *Gartner Press Release* (blogue), 22 mai 2022. https://www.gartner.com/en/newsroom/press-releases/2022-05-31-gartner-identifies-top-five-trends-in-privacy-through-2024.

[117] Hunton Andrews Kurth. « New Bipartisan Federal Privacy Proposal Unveiled: American Privacy Rights Act », 23 avril 2024. https://www.hunton.com/privacy-and-information-security-law/new-bipartisan-federal-privacy-proposal-unveiled-american-privacy-rights-act.

[118] Zuboff, Shoshana. *« L'ère du capitalisme de surveillance : la lutte pour l'avenir à la nouvelle frontière du pouvoir* ». Affaires publiques, 2019.

[119] Commissariat à la protection de la vie privée du Canada. « Principes pour des technologies de l'intelligence artificielle (IA) générative responsables, dignes de confiance et respectueuses de la vie privée », 7 décembre 2023. https://www.priv.gc.ca/fr/sujets-lies-a-la-protection-de-la-vie-privee/technologie/intelligence-artificielle/gd_principes_ia/.

[120] Schulze, Kris. « Cloud Computing vs. Edge Computing | Blog. » Scale Computing, 2 mai 2024. https://www.scalecomputing.com/blog/decoding-the-differences-between-cloud-computing-vs-edge-computing.

[121] Earney, Shandra. « Is the Edge or Cloud Better for Security and Privacy? » Xalient, 17 octobre 2022. https://xailient.com/blog/is-the-edge-or-cloud-better-for-security-and-privacy/.

[122] Swabey, Pete. « Why Edge Computing Is a Double-Edged Sword for Privacy ». *Tech Monitor* (blogue), 31 mars 2023. https://techmonitor.ai/focus/privacy-on-the-edge-why-edge-computing-is-a-double-edged-sword-for-privacy.

[123] Beasy, John; Hagerman, Chris; Joy, Amanda; Rigillo, Nicole; Robertson, Simon; Thomas, Tieja; Van der Elst, Kristel et Wester, Meaghan. *Vies futures : Incertitudes*. Ottawa : Horizons de politiques Canada, 2024. https://horizons.service.canada.ca/fr/2024/vies-futures-incertitudes/index.shtml.

[124] Miller, Lloyd. « RECON VILLAGE - Applied OSINT For Politics: Turning Open Data Into News - TIB AV-Portal ». Présenté lors de la DEF CON (Las Vegas), 2018. https://av.tib.eu/media/39947.

[125] Vincent, Subramaniam. « How Open Source Intelligence Can Help Journalists Cover Conflicts." *Markkula Center for Applied Ethics »* (blogue), 3 octobre 2023. https://www.scu.edu/ethics/all-about-ethics/how-open-source-intelligence-can-help-journalists-cover-conflicts/.
[126] Pierson, Chris. « Celebrities Are a Big Target for Hackers - Cyber Threats ». *BlackCloak* | *Protect Your Digital Life™* (blogue), 15 octobre 2020. https://blackcloak.io/online-threats-put-celebrities-digital-lives-in-crosshairs/.
[127] Lobel, Orly. « The Problem With Too Much Data Privacy ». *Time Magazine*, 7 octobre 2022. https://time.com/6224484/data-privacy-problem/.
[128] Irwin, Jasmine; Dharamshi, Alannah et Zon, Noah. « La sécurité et la confidentialité des enfants à l'ère numérique ». Groupe CSA, mars 2021. https://www.csagroup.org/wp-content/uploads/Groupe-CSA-Recherche-Securite-et-confidentialite-des-enfants-ere-numerique.pdf.
[129] Irwin, Jasmine; Dharamshi, Alannah et Zon, Noah. « La sécurité et la confidentialité des enfants à l'ère numérique ». Groupe CSA, mars 2021. https://www.csagroup.org/wp-content/uploads/Groupe-CSA-Recherche-Securite-et-confidentialite-des-enfants-ere-numerique.pdf.
[130] Murphy, Chris. « Opinion | Algorithms Are Making Kids Desperately Unhappy ». *The New York Times*, 18 juillet 2023, sec. Opinion. https://www.nytimes.com/2023/07/18/opinion/big-tech-algorithms-kids-discovery.html.
[131] Leiser, M. R. « Protecting Children from Dark Patterns and Deceptive Design ». *SSRN*, 11 décembre 2023. http://web.archive.org/web/20240616110633/https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4660222.
[132] Bureau de la consommation. « Les interfaces truquées »'. Gouvernement du Canada, 19 July 2023. https://ised-isde.canada.ca/site/bureau-consommation/fr/interfaces-truquees.
[133] Hill, Amelia. « Social Media Triggers Children to Dislike Their Own Bodies, Says Study ». *The Guardian*, 1er janvier 2023, sec. Society. https://www.theguardian.com/society/2023/jan/01/social-media-triggers-children-to-dislike-their-own-bodies-says-study.
[134] David, Emilia. « AI Image Training Dataset Found to Include Child Sexual Abuse Imagery - The Verge ». *The Verge*, 20 décembre 2023. https://www.theverge.com/2023/12/20/24009418/generative-ai-image-laion-csam-google-stability-stanford.
[135] Duboust, Oceane. « Society Needs to Be Alert": Most People Are Unaware AI Is Being Used to Create Child Abuse Content | Euronews ». *Euronews Next*, 19 février 2024. https://www.euronews.com/next/2024/02/19/society-needs-to-be-alert-most-people-are-unaware-ai-is-being-used-to-create-child-abuse-c.
[136] SecureKin. « Keylogger App To Record Your Child's Keystrokes ». Consulté le 23 mai 2024. https://securekin.com.

[137] Valentino-DeVries, Jennifer et Keller, Michael H. « A Marketplace of Girl Influencers Managed by Moms and Stalked by Men ». *The New York Times*, 23 février 2024, sec. U.S. https://www.nytimes.com/2024/02/22/us/instagram-child-influencers.html.
[138] Owen. « New AI Toys Spark Privacy Concerns for Kids ». *GZERO Media*, 12 décembre 2023. https://www.gzeromedia.com/gzero-ai/gzero-ai-video/new-ai-toys-spark-privacy-concerns-for-kids.
[139] Irwin, Jasmine; Dharamshi, Alannah et Zon, Noah. « La sécurité et la confidentialité des enfants à l'ère numérique ». Groupe CSA, mars 2021. https://www.csagroup.org/wp-content/uploads/Groupe-CSA-Recherche-Securite-et-confidentialite-des-enfants-ere-numerique.pdf.
[140] Bala, Nila. « Opinion | Why Are You Publicly Sharing Your Child's DNA Information? » *The New York Times*, 2 janvier 2020. https://www.nytimes.com/2020/01/02/opinion/dna-test-privacy-children.html.
[141] ClassDojo. « Third Party Service Providers ». Consulté le 23 mai 2024. https://www.classdojo.com/en-gb/third-party-service-providers/?redirect=true.
[142] DaSilva, Tomasia. 'Hackers Steal Children's School Photos Following a Privacy Breach'. *Global News*, 14 février 2024. https://globalnews.ca/news/10294971/hackers-childrens-school-photos-edge-imaging/.
[143] Bajak, Frank; Hollingsworth, Heather et Fenn, Larry. « Ransomware Criminals Are Dumping Kids' Private Files Online after School Hacks ». *Canadian Security Magazine*, 5 juillet 2023. https://www.canadiansecuritymag.com/ransomware-criminals-are-dumping-kids-private-files-online-after-school-hacks/.
[144] Omstead, Jordan. « SickKids Cyberattack: Ransomware Group LockBit Apologizes Saying "partner" Was behind Attack ». *CTV News*, 3 janvier 2023. https://toronto.ctvnews.ca/ransomware-group-lockbit-apologizes-saying-partner-was-behind-sickkids-attack-1.6214906.
[145] Lee, Austin. « I Am Deeply Troubled": Data Breach Impacts Clients at Lanark County Family Services Organization ». *CTV News*, 14 février 2024, sec. Ottawa. https://ottawa.ctvnews.ca/i-am-deeply-troubled-data-breach-impacts-clients-at-lanark-county-family-services-organization-1.6769384.
[146] Tsekouras, Phil. « Did You Give Birth between 2010 and 2023 in Ontario? Your Personal Health Information Was "likely" Impacted by a Data Breach ». *CP24*, 25 septembre 2023. https://www.cp24.com/news/did-you-give-birth-between-2010-and-2023-in-ontario-your-personal-health-information-was-likely-impacted-by-a-data-breach-1.6576525?cache=yes%3F.
[147] Ali, Suzan; Elgharabawy, Mounir; Duchaussoy, Quentin; Mannan, Mohammad et Youssef, Amr. « Betrayed by the Guardian: Security and Privacy Risks of Parental Control Solutions ». In *Proceedings of the 36th Annual Computer Security Applications Conference*, p. 69-83. ACSAC 2020. New York (N.Y., États-Unis) : Association for Computing Machinery, 2020. https://doi.org/10.1145/3427228.3427287.

[148] Woollacott, Emma. « TikTok condamné à une amende de 345 millions d'euros pour violation de la vie privée des enfants ». *Forbes*, 18 septembre 2023. https://www.forbes.com/sites/emmawoollacott/2023/09/18/tiktok-hit-with-345m-fine-for-failing-to-protect-children/.

[149] Matza, Max. « Microsoft va payer 20 millions de dollars pour violations de la vie privée des enfants ». *BBC*, 6 juin 2023. https://www.bbc.com/news/world-us-canada-65817558.

[150] Singer, Natasha. « Amazon to Pay $25 Million to Settle Children's Privacy Charges ». *The New York Times*, 31 mai 2023, sec. Technology. https://www.nytimes.com/2023/05/31/technology/amazon-25-million-childrens-privacy.html.

[151].Jeunesse, j'écoute. « Jeunesse, J'écoute et l'Institut Vector annoncent un partenariat important pour des innovations améliorées axées sur l'humain pour les services de santé mentale en ligne pour les jeunes d'un océan à l'autre ». Consulté le 23 mai 2024. https://jeunessejecoute.ca/publications/jeunesse-jecoute-et-linstitut-vector-annoncent-un-partenariat-important-pour-des-innovations-ameliorees-axees-sur-lhumain-pour-les-services-de-sante-mentale-en-ligne-pour-les-jeunes-dun-ocean-a-lautre/?_ga=2.3904055.629371948.1726846218-976181761.1726846218&_gl=1*184n6jl*_ga*OTc2MTgxNzYxLjE3MjY4NDYyMTg.*_ga_TQY135CFXS*MTcyNjg0NjIxOC4xLjAuMTcyNjg0NjIxOC42MC4wLjExMDU1OTQ2MTA.

[152] Jeunesse, j'écoute. « Réponses aux questions à propos de Données de Jeunesse, J'écoute! » Accessed 23 May 2024. https://jeunessejecoute.ca/obtenir-des-donnees/donnees-faq/?_ga=2.66936105.629371948.1726846218-976181761.1726846218&_gl=1*5deznc*_ga*OTc2MTgxNzYxLjE3MjY4NDYyMTg.*_ga_TQY135CFXS*MTcyNjg0NjIxOC4xLjEuMTcyNjg0NjMwNy42MC4wLjExMDU1OTQ2MTA.

[153].Jeunesse, j'écoute. « Jeunesse, J'écoute et l'Institut Vector annoncent un partenariat important pour des innovations améliorées axées sur l'humain pour les services de santé mentale en ligne pour les jeunes d'un océan à l'autre ». Consulté le 23 mai 2024. https://jeunessejecoute.ca/publications/jeunesse-jecoute-et-linstitut-vector-annoncent-un-partenariat-important-pour-des-innovations-ameliorees-axees-sur-lhumain-pour-les-services-de-sante-mentale-en-ligne-pour-les-jeunes-dun-ocean-a-lautre/?_ga=2.3904055.629371948.1726846218-976181761.1726846218&_gl=1*184n6jl*_ga*OTc2MTgxNzYxLjE3MjY4NDYyMTg.*_ga_TQY135CFXS*MTcyNjg0NjIxOC4xLjAuMTcyNjg0NjIxOC42MC4wLjExMDU1OTQ2MTA.

[154] Bennett, Drake. « How a Massive Hack of Psychotherapy Records Revealed a Nation's Secrets ». *Bloomberg*, 22 avril 2024. https://www.bloomberg.com/news/features/2024-04-22/a-massive-therapy-hack-shows-just-how-unsafe-patients-files-can-be.

[155] Wood, Stuart. « Explorer la sensibilisation et l'utilisation des contrôles parentaux pour soutenir la sécurité numérique ». *Internet Matters* (blogue), 21 juillet 2023. https://www.internetmatters.org/hub/research/research-tracker-awareness-usage-parental-controls/.

[156] Beauvais, Michael, and Leslie Regan Shade. « How Will Bill C-27 Impact Youth Privacy? » *Schwartz Reisman Institute* (blogue), 8 octobre 2022. https://srinstitute.utoronto.ca/news/how-will-bill-c-27-impact-youth-privacy.
[157] McConvey, Joel R. « Wizz Dials up Biometrics from Yoti to Prevent Sextortion, Achieve EU Compliance ». *Biometric Update* (blogue), 15 février 2024. https://www.biometricupdate.com/202402/wizz-dials-up-biometrics-from-yoti-to-prevent-sextortion-achieve-eu-compliance.
[158] Mithani, Jasmine. « The 19th Explains: Why Some States Are Requiring ID to Watch Porn Online ». The 19th, 29 janvier 2024. https://19thnews.org/2024/01/states-age-verification-adult-content-online/.
[159] Costabel, Milagros. « I'm Totally Blind. Artificial Intelligence Is Helping Me Rediscover the World ». *Slate*, 11 octobre 2023. https://slate.com/technology/2023/10/ai-image-tools-blind-low-vision.html.
[160] Shu, Li-Qi; Sun, Yi-Kan; Tan, Lin-Hua; Shu, Qiang et Chang, Anthony C. "Application of Artificial Intelligence in Pediatrics: Past, Present and Future." *World Journal of Pediatrics*, vol. 15, nº 2 (1er avril 2019) : p. 105-108. https://doi.org/10.1007/s12519-019-00255-1.
[161] UNICEF. « Children and AI: Opportunities and Risks ». Genève : UNICEF, 2018. https://www.unicef.org/innovation/sites/unicef.org.innovation/files/2018-11/Children%20and%20AI_Short%20Verson%20%283%29.pdf.
[162] Madrigal, Alexis C. « How the Facebook News Feed Algorithm Shapes Your Friendships ». *The Atlantic*, 20 octobre 2010. https://www.theatlantic.com/technology/archive/2010/10/how-the-facebook-news-feed-algorithm-shapes-your-friendships/64996/.
[163] Narayanan, Arvind. « Understanding Social Media Recommendation Algorithms ». *The Knight First Amendment Institute*, mars 2023. http://knightcolumbia.org/content/understanding-social-media-recommendation-algorithms.
[164] Daniels, Nicole. « Do You Feel You're Friends With Celebrities or Influencers You Follow Online? » *The New York Times*, 13 mai 2021. https://www.nytimes.com/2021/05/13/learning/do-you-feel-youre-friends-with-celebrities-or-influencers-you-follow-online.html.
[165] HealthSnap. « AI in Remote Patient Monitoring: The Top 4 Use Cases in 2024 », 6 septembre 2023. https://healthsnap.io/ai-in-remote-patient-monitoring-the-top-4-use-cases-in-2024/.
[166] LoveGenius. « LoveGenius - Magic AI That Generates Your Tinder & Bumble Bio ». Consulté le 23 mai 2024. https://www.lovegenius.io/.
[167] Elliott, Jaleesa. « Singles in America" Study: Daters Breaking the Ice with AI », 20 février 2024. https://phys.org/news/2024-02-singles-america-daters-ice-ai.html.
[168] Jade, Isobella. « How the OurFamilyWizard Co-Parenting App Saved My Divorce ». *The Daily Beast*, 6 février 2024. https://www.thedailybeast.com/how-the-ourfamilywizard-co-parenting-app-saved-my-divorce.

[169] Bookker. « How Do Augmented Reality and Artificial Intelligence Interact? », 10 mars 2023. https://www.bookkercorp.com/en/how-do-augmented-reality-and-artificial-intelligence-interact/.
[170] Sanchez, Bailey et Spivack, Jameson. « Youth Privacy in Immersive Technologies: Regulatory Enforcement, Self-Regulatory Guidance, and Remaining Uncertainties ». Future of Privacy Forum, mars 2024. https://fpf.org/wp-content/uploads/2024/03/Final-Youth_immersive-tech-regulatory-guidance.pdf
[171] Parshall, Allison. « Why an Assault on Your VR Body Can Feel so Real ». *Scienceline*, 29 juin 2022. https://scienceline.org/2022/06/virtual-reality-assault-psychology/.
[172] De Freitas, Julian; Uguralp, Ahmet K.; Uguralp, Zeliha O. et Puntoni, Stefano. « AI Companions Reduce Loneliness ». arXiv, 9 juillet 2024. https://doi.org/10.48550/arXiv.2407.19096.
[173] Robb, Alice. « He Checks in on Me More than My Friends and Family': Can AI Therapists Do Better than the Real Thing? » The Guardian, 2 mars 2024, sec. Life and style. https://www.theguardian.com/lifeandstyle/2024/mar/02/can-ai-chatbot-therapists-do-better-than-the-real-thing.
[174] Sahota, Neil. « AI Cupid: Enhancing Romance And Deepening Connections In Relationships ». *Forbes*, février 2024. https://www.forbes.com/sites/neilsahota/2024/02/12/ai-cupid-enhancing-romance-and-deepening-connections-in-relationships/.
[175] Marchesi, Serena; Ghiglino, Davide; Ciardo, Francesca; Perez-Osorio, Jairo; Baykara, Ebru et Wykowska, Agnieszka. « Do We Adopt the Intentional Stance Toward Humanoid Robots? » *Frontiers in Psychology*, vol. 10 (15 mars 2019). https://doi.org/10.3389/fpsyg.2019.00450.
[176] Rodgers, Harry et Lloyd-Evans, Emily Christine. « Intimate Snapshots: TikTok, Algorithm, and the Recreation of Identity ». *Anthways*, 18 septembre 2021. https://doi.org/10.5281/ZENODO.5515620.
[177] Collu, Samuele. « # Zoombies: Cybernetic Trance in Pandemic Times ». Dans *Planetary Health Humanities and Pandemics*, p. 199-217. Routledge. Consulté le 8 août 2024. https://www.taylorfrancis.com/chapters/edit/10.4324/9781003367581-13/zoombies-samuele-collu.
[178] Elsa. « 4 Ways to Track My Child's Phone without Them Knowing [2024] ». AirDroid, 3 janvier 2024. https://www.airdroid.com/parent-control/track-childs-phone-without-knowing-free/.
[179] Valentino-DeVries. « Hundreds of Apps Can Empower Stalkers to Track Their Victims - The New York Times ». The New York Times, 19 mai 2018. https://www.nytimes.com/2018/05/19/technology/phone-apps-stalking.html.
[180] Emch, Lilly. « Parents Should Trust, Not Track, Their Kids [colonne] ». *LancasterOnline*, 29 octobre 2023. https://lancasteronline.com/opinion/columnists/parents-should-trust-not-track-their-kids-column/article_6d14039c-7434-11ee-bc71-effc2c075d90.html.
[181] New York State Office for the Aging. « NYSOFA's Rollout of AI Companion Robot ElliQ Shows 95% Reduction in Loneliness », 1er août 2023. https://aging.ny.gov/news/nysofas-rollout-ai-companion-robot-elliq-shows-95-reduction-loneliness.
[182] Blackett, L'Oréal. « I Found A New Black Therapist & It's An AI Chatbot ». *Refinery 29*. Consulté le 8 août 2024. https://www.refinery29.com/en-us/artificial-intelligence-chat-gpt-black-mental-health.
[183] Collu, Samuele. Into the Loop. Affect, Therapy, Screens. Durham, Caroline du Nord : Duke University Press, sous presse.