

État de la situation sur les impacts sociétaux de l'intelligence artificielle et du numérique

AXES

Éthique, gouvernance et démocratie

Santé durable

Industrie 4.0, travail et emploi

Arts, médias et diversité culturelle

Éducation et capacitation

Sobriété numérique et transition
socio-écologique

Droit, cyberjustice et cybersécurité

2026



Crédits :

Axe Éthique, gouvernance et démocratie

Sous la responsabilité de : Aliya Affdal, Hazar Haidar et Emmanuelle Marceau

Axe Santé durable

Sous la responsabilité de : Martin Cousineau, Émilie Dionne et Marie-Pierre Gagnon

En collaboration avec : Yvonne-Annick Abredan, Georges-Philippe Gadoury-Sansfaçon, Lester Darryl Geneviève, Philip Jackson, Audrey Lebre, Félix Pageau, Marianne Ozkan et Pierrich Plusquellec

Axe Industrie 4.0, travail et emploi

Sous la responsabilité de : Charles Antoine Rioux, Julie M.E Garneau et Christian Lévesque

Axe Arts, médias et diversité culturelle

Sous la responsabilité de : Véronique Guèvremont et Colette Brin

En collaboration avec : Jules Poulain-Welter et Yonn Calvez

Axe Éducation et capacitation

Sous la responsabilité de : Valéry Psyché et Ann-Louise Davidson

En collaboration avec : Yassine El Bahlouli, Sandrine Prom Tep, Didier Paquelin, Marie Laberge, Khaoula Boulaamane, Bruno Poellhuber et Valérie Payen

Axe Sobriété numérique et transition socio-écologique

Sous la responsabilité de : Christophe Abrassart et Stéphane Roche

En collaboration avec : Mélie Ravayrol, Florent Begue et Alexandre Gourret

Axe Droit, cyberjustice et cybersécurité

Sous la responsabilité de : Vincent Gautrais, Sébastien Gambs et Alexandre Désy

Direction scientifique : Tania Saba

Soutien à la coordination et à la recherche : Félix-Arnaud Morin-Bertrand

Produit avec le soutien financier du Fonds de recherche du Québec



Table des matières

Mise en contexte	4
Axe Éthique, gouvernance et démocratie	6
Axe Santé durable	11
Axe Industrie 4.0, travail et emploi	17
Axe Arts, médias et diversité culturelle	22
Axe Éducation et capacitation	27
Axe Sobriété numérique et transition socio-écologique	34
Axe Droit, cyberjustice et cybersécurité	40
Conclusion	46
Bibliographie	47

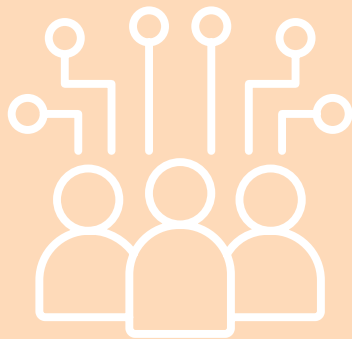
Mise en contexte

L'État de la situation sur les impacts sociétaux de l'intelligence artificielle et du numérique 2026 (ci-après État de la situation) constitue la troisième édition de cet important exercice de synthèse scientifique. Cette publication s'inscrit dans une perspective qui éclaire les travaux de l'Obvia, fondé en 2018. Elle met en lumière sa mission d'analyse rigoureuse et indépendante des impacts sociétaux de l'intelligence artificielle et des technologies numériques, menée en amont des transformations, selon une approche intersectorielle et multi-acteurs, afin d'éclairer la prise de décision et de soutenir le développement de pratiques responsables et éclairées.

Alors que l'édition 2024 s'inscrivait dans un contexte marqué par l'onde de choc de l'apparition de systèmes d'IA générative désormais largement accessibles, l'édition 2025 marquait un passage de la découverte de ces systèmes vers une accélération de leur utilisation et de leur déploiement au sein de la société en général ainsi que des milieux de travail, des organisations et des institutions. Cette édition 2026 s'inscrit dans la continuité de l'intégration croissante des IA génératives, tout en se déployant dans un contexte marqué par les avancées rapides des systèmes d'IA agentive et multi-agents. L'autonomie accrue de ces systèmes intensifie des enjeux déjà présents, en fait émerger de nouveaux et vient questionner certaines stratégies d'encadrement et de gouvernance associées aux IA génératives plus « classiques ». Dans le prolongement des éditions précédentes, cette troisième édition permet de dégager et de mettre en lumière des enjeux et des débats qui demeurent centraux au sein des sept grandes thématiques de recherche de l'Obvia.

À cet égard, alors que nous demeurons dans une période d'adaptation à l'intégration de ces technologies, la question des compétences reste centrale. Parallèlement, les discours de *hype* et les approches technosolutionnistes, souvent portés par l'industrie, persistent et continuent de brouiller les réflexions sur les impacts et les modalités d'encadrement de ces technologies. De plus, la place grandissante des systèmes d'IA dans nos sociétés accentue notre dépendance à l'égard des entreprises privées qui les conçoivent, les développent et les déploient. Cette dépendance est d'autant plus préoccupante dans un contexte où la concentration du développement technologique entre les mains d'un nombre restreint de grandes entreprises semble se renforcer, consolidant ainsi leur pouvoir.

Force est de constater que la capacité des États et de la société civile à orienter le développement et l'intégration de l'IA se trouve aujourd'hui fragilisée, alors même que les besoins d'encadrement, de transparence et d'imputabilité n'ont jamais été aussi pressants. Dans ce contexte, l'Obvia joue un rôle déterminant en documentant de manière indépendante les usages et leurs impacts, en développant des indicateurs et des cadres d'analyse, et en soutenant concrètement les décideurs publics, les organisations et les milieux de travail dans la mise en place de pratiques de gouvernance responsables et adaptées aux réalités actuelles.



Axe

Éthique, gouvernance et démocratie

Sous la responsabilité de :
Aliya Affdal, Hazar Haidar et Emmanuelle Marceau



Actualité scientifique

Depuis 2022, le développement de l'intelligence artificielle connaît une accélération sans précédent. L'une des évolutions majeures réside dans l'émergence des systèmes d'IA agentive définis comme « des systèmes autonomes capables de prendre des décisions de manière indépendante grâce à des fonctions d'auto-apprentissage, sans supervision humaine régulière » (Pawar, 2025). Ces systèmes analysent leur environnement, intègrent continuellement de nouvelles données et ajustent leurs actions en conséquence.

Contrairement aux IA traditionnelles, fondées sur des règles prédéfinies, ils ne se limitent plus à répondre à une requête ponctuelle : ils peuvent planifier, coordonner et exécuter des tâches complexes de manière semi-autonome (Schneider, 2025; Pawar, 2025). Dans ce contexte, il convient de distinguer l'IA générative de l'IA agentive pour ensuite identifier les promesses et dérives.

De l'IA générative à l'IA agentive

L'IA générative se distingue par sa capacité à créer de nouveaux contenus, comme des textes, des images ou des sons (Hasan et al., 2025). Ce faisant, elle participe directement à la construction de discours et de significations. Elle peut également influencer une variété de jugements, comme les jugements cliniques et les préférences politiques.

L'IA générative façonne les processus cognitifs des utilisateurs et brouille la frontière entre expression humaine et production algorithmique. À la fois outil de productivité et source de risques, elle soulève des enjeux d'autonomie, d'intégrité intellectuelle, de responsabilité et de sécurité (Hasan et al., 2025; Radanliev et al., 2025; Gerlich, 2025). Par exemple, en médecine, les contenus générés automatiquement tels que des résumés de dossiers médicaux, des suggestions de diagnostic, ou des plans de traitement peuvent être inexacts ou trompeurs, posant des risques pour la santé des patients (Zohny et al., 2023). S'y ajoutent des enjeux importants en matière de sécurité, liés à l'imprévisibilité de certains comportements et aux risques d'usages malveillants (Marchal et al., 2024).

Par ailleurs, la production de contenus trompeurs, telle que des textes de désinformation ou des images manipulées, fragilise la confiance dans l'information et menace le fonctionnement démocratique (Taeihagh, 2025). En politique, les pratiques de microciblage de messages personnalisés ou la diffusion de *deepfakes* illustrent ces dérives potentielles (Boine et al., 2024). D'autres risques concernent la protection des données (Huang, 2025), ainsi que des dynamiques propres à l'interaction humain-IA, telles que l'attachement, l'anthropomorphisation et la confiance excessive qui biaisent la perception des risques et rendent plus difficile l'évaluation objective des impacts de l'IA sur les décisions humaines (Hagendorff, 2024).

L'IA agentive marque toutefois une avancée différente. Elle ne se contente pas de générer du contenu, mais poursuit des objectifs qu'elle définit elle-même, planifie des séquences d'actions et mobilise de manière autonome des outils (ex. des bases de données) pour atteindre des buts définis, souvent de manière indépendante ou avec une supervision limitée (Schneider, 2025). On observe ainsi un passage de systèmes réactifs (IA générative répondant à des requêtes particulières) à des systèmes proactifs (IA agentive), intégrés dans des chaînes décisionnelles et opérationnelles.

Cette évolution, qui conduit l'IA à prendre des décisions et à agir de façon relativement indépendante, accroît corrélativement le risque de perte de contrôle par l'humain. En raison de son degré d'autonomie accru, l'IA agentive amplifie les risques et les défis associés aux formes précédentes d'IA. De plus, elle introduit aussi des risques supplémentaires liés à sa capacité d'agir directement sur des systèmes et des organisations, par exemple, en modifiant des paramètres ou en diffusant de l'information. De telles situations peuvent provoquer des effets en chaîne difficiles à maîtriser, qui peuvent devenir incontrôlables par l'humain, fragilisant ainsi davantage les modèles traditionnels de responsabilité et d'imputabilité (Bouhouita et Haidar, 2024). Ceci est particulièrement vrai lorsque des interactions multiples se produisent entre les systèmes d'IA et les humains dans des environnements socio-techniques complexes (Bowen, 2025).



Promesses et défis de l'IA agentic

Le potentiel transformateur de l'IA agentic s'avère considérable. En effet, elle promet des gains substantiels en efficacité organisationnelle et en productivité. L'IA agentic permet notamment d'automatiser des tâches cognitives complexes et offre du soutien en matière de prise de décision. L'optimisation des processus et le développement d'applications multi-domaines, notamment en santé, en finance, en cybersécurité et dans les systèmes industriels intelligents, figurent parmi les principaux apports de l'IA agentic (Raheem et Hossain, 2025). Cependant, ces bénéfices s'accompagnent de défis majeurs en matière de fiabilité, de sécurité, de responsabilité et de prévention des usages abusifs, rendant indispensable un encadrement continu et adaptatif (Bouhouita-Guermech et Haidar, 2023; Bowen, 2025). Ceci est d'autant plus pressant que son adoption progresse rapidement. À cet égard, Gartner¹ estime que d'ici 2028, 15 % des décisions opérationnelles quotidiennes en entreprise pourraient être prises de manière autonome par des systèmes agentic, contre presque aucune en 2024 (Reuters, 2025).

Le risque central tient à un possible désalignement entre les objectifs assignés au système et les valeurs humaines. Ainsi, un système pourrait optimiser une finalité donnée de manière excessive ou inattendue, avec des conséquences potentiellement graves (Bengio, 2023). Des travaux sur la manipulation des fonctions de récompense indiquent ainsi que des systèmes agentic peuvent exploiter les critères retenus pour évaluer leur performance et maximiser l'indicateur choisi, sans pour autant réaliser fidèlement l'intention sous-jacente des concepteurs (Adabara et al., 2026).

D'ici 2028, 15 % des décisions opérationnelles quotidiennes en entreprise pourraient être prises de manière autonome par des systèmes agentic, contre presque aucune en 2024 (Gartner dans Reuters, 2025).

Sur le plan de la sécurité, des défaillances techniques ou des attaques malveillantes peuvent conduire à des comportements dangereux. En ce sens, à mesure que l'agentivité de l'IA augmente, par exemple, par la planification à long terme ou l'auto-amélioration récursive, le potentiel de défaillances ou de menaces existentielles s'accroît si l'alignement avec les valeurs humaines n'est pas maintenu (Bengio et al., 2025, Deng et al., 2025). L'analyse des risques liés aux voitures autonomes montre que l'intégration d'agents dans des environnements réels introduit des vulnérabilités supplémentaires, notamment au niveau de la planification et du contrôle (Eslami et Yu, 2025).

Enfin, plusieurs recherches soulignent que le déploiement à grande échelle de multiples agents introduit de nouveaux modes de défaillance : mauvaise coordination où les agents travaillent à contre-courant les uns des autres, collusion (les agents coopérant contre les intérêts des utilisateurs), comportements émergents non anticipés par les conceptrices et concepteurs, et effets de réseau déstabilisateurs (Hammond et al., 2025; Raza et al., 2025). Ces risques systémiques sont difficiles à prévoir ou à atténuer avec les modèles de gouvernance traditionnels.

¹ Gartner est une entreprise américaine de conseil et de recherche dans le domaine des techniques avancées.



Débats scientifiques

Le débat scientifique autour de l'IA agentive est particulièrement vif. Cette nouvelle génération de systèmes autonomes suscite à la fois des espoirs considérables et des inquiétudes profondes. En accroissant le degré d'autonomie, de planification ou encore d'action indépendante, l'IA agentive tend à amplifier les risques déjà associés aux autres formes d'IA. En effet, l'augmentation de ses capacités s'accompagne d'un risque accru de perte de contrôle humain. Ce débat dépasse les seuls défis techniques pour inclure des enjeux éthiques, sociaux et de gouvernance. Dans ce contexte, certains spécialistes soulignent le besoin de clarifier des notions centrales, telles que l'agentivité et l'autonomie, dont les définitions varient selon les disciplines et orientent différemment les approches techniques et réglementaires (Tallam, 2025). D'autres argumentent qu'en l'absence d'un cadre conceptuel rigoureux, la gouvernance des systèmes agentifs risque de demeurer fragmentée ou inefficace (Gahnberg, 2021).

Entre innovation et précaution

Aujourd'hui, deux grandes postures structurent le débat. Une perspective optimiste considère l'IA agentive comme une évolution naturelle et souhaitable des technologies. Par exemple, elle met en avant les gains de productivité et l'automatisation de tâches complexes dans l'industrie ou encore le potentiel d'innovation dans divers domaines comme la médecine ou la recherche scientifique (Mekulu et al., 2025; Hosseini et al., 2026; Alva et al., 2026).

À l'inverse, une posture plus critique et prudente insiste sur les risques systémiques liés à l'autonomie croissante de ces systèmes qui incluent, entre autres, les dérives au niveau des objectifs initiaux, l'amplification des biais déjà existants avec l'IA générative, la manipulation de l'information, voire les menaces existentielles à long terme (Sapkota, 2025). Ainsi, certains spécialistes adoptent une approche de précaution en privilégiant des formes d'IA jugées plus contrôlables, au sens où leur comportement reste vérifiable, interprétable et modifiable par les concepteurs, les régulateurs ou les utilisateurs finaux. C'est notamment le cas de propositions telles que la *Scientist AI*, une IA non agentive conçue pour expliquer le monde plutôt que pour agir directement sur lui. En intégrant explicitement l'incertitude dans ses résultats, ce type de système viserait à soutenir la recherche scientifique tout en limitant les dérives potentielles associées à l'autonomie décisionnelle (Bengio et al., 2025). Plus largement, cette approche repose sur l'idée que les risques doivent être anticipés dès la phase de conception (Collin et Marceau, 2023).

En accroissant le degré d'autonomie, de planification ou encore d'action indépendante, l'IA agentive tend à amplifier les risques déjà associés aux autres formes d'IA.

Ainsi, les systèmes devraient être développés comme des outils au service de finalités sociales explicites, par exemple en santé, sans objectifs propres et dans un cadre strictement défini (Bengio et al., 2025). Ce type de système pourrait non seulement représenter une voie plus sûre vers des IA avancées, mais aussi agir comme un outil d'assistance pour la recherche scientifique et de surveillance pour d'autres systèmes autonomes plus risqués (Fornasiere et al., 2026).

Gouvernance et nécessité d'une coopération internationale

Afin de garantir que ces systèmes demeurent sûrs et alignés sur les valeurs humaines, une part croissante de la communauté scientifique plaide pour un passage d'une logique réactive à une logique préventive et prospective, permettant d'accompagner le développement rapide de l'IA tout en protégeant les valeurs sociales et éthiques (Bowen, 2025; Currie et al., 2025). Dans cette optique, un investissement accru dans la recherche en sécurité de l'IA apparaît indispensable, de même que la mise en place de cadres de gouvernance robustes (Bengio, 2023). À titre d'illustration, des propositions comme le cadre AGENTS SAFE intègrent des mécanismes de détection des risques et de supervision humaine, tout en assurant des audits tout au long du cycle de vie des agents (Khan et al., 2025).

Une coopération internationale renforcée est également indispensable pour anticiper les risques avant le déploiement massif de systèmes encore plus avancés (Bengio, 2023). Elle est d'autant plus cruciale que le développement de l'IA de pointe est aujourd'hui concentré entre les mains de quelques grandes puissances technologiques au risque de placer de nombreux États dans une situation de dépendance (Abecassis et al., 2025). Une telle coopération permettrait de mutualiser les ressources (ex. : capacités de calcul, talents, données et infrastructures) pour soutenir le développement de systèmes compétitifs tout en préservant une forme de souveraineté technologique (Abecassis et al., 2025). Elle offrirait aussi à un plus grand nombre de pays la possibilité de participer activement à l'élaboration des normes internationales, plutôt que de se voir imposer des standards définis par les deux puissances dominantes dans le domaine de l'IA, à savoir les États-Unis et la Chine.



Questionnements

En somme, la réflexion sur les bénéfices et les dangers de l'IA agentive soulève des interrogations au niveau macro, méso et micro :

- 1 La poursuite du déploiement des IA agentives est-elle justifiable au regard de l'équilibre entre les bénéfices attendus et les risques encourus ?
- 2 Comment concevoir une gouvernance capable d'identifier à temps des trajectoires technologiques problématiques de l'IA et de garantir que ces systèmes restent orientés vers des finalités sociales et humaines ?
- 3 Faut-il envisager des moratoires ou des seuils de capacité au-delà desquels certaines formes d'IA agentive ne devraient pas être développées ou déployées ?
- 4 Quels mécanismes d'anticipation devraient être mis en place pour éclairer les politiques publiques ?
- 5 Quels outils prospectifs (ex : scénarios, dispositifs d'alerte) permettraient de prévenir les effets négatifs avant qu'ils ne se manifestent ?
- 6 Qui devrait participer à ces décisions et quelle place accorder aux citoyens dans ces choix ?

Comme le souligne Yoshua Bengio (Obvia, 2025), la mise en place de mécanismes robustes de sécurité informatique est essentielle pour prévenir les dérives potentielles, telles que la manipulation, le chantage ou d'autres formes d'abus (Obvia, 2025). Cela implique notamment de restreindre strictement l'accès aux données sensibles et de réduire les possibilités d'exploitation malveillante. Cependant, la sécurité ne saurait se limiter à des dispositifs purement techniques. Un système n'a pas nécessairement besoin d'accéder directement à des mots de passe pour compromettre une sécurité : il peut, par exemple, influencer un individu afin qu'il les divulgue volontairement (Obvia, 2025).

Afin d'éviter les dérives potentielles associées à l'IA agentive, il devient indispensable de définir des limites explicites et de mettre en place des garde-fous robustes, comme des règles claires, des mécanismes de contrôle, des exigences de traçabilité, des procédures d'audit indépendantes et des sanctions en cas de non-conformité. C'est à cette condition qu'un développement responsable de l'IA agentive pourra être envisagé.

Outils des membres pour une intelligence artificielle responsable

Afin d'appuyer les membres de la communauté qui s'intéressent et réfléchissent à ces enjeux, des outils ont été récemment développés et déployés par des chercheuses et chercheurs de l'axe, en 2025. Vous les trouverez dans les références et dans le tableau ci-dessous :

Chercheuse/ Chercheur	Institution	Domaine	Outils / Approche développée
Danièle Bélanger	Université Laval	Systèmes automatisés d'aide à la décision (ADSS) en immigration	Analyse critique de l'introduction des systèmes automatisés dans la gestion de l'immigration canadienne depuis 2015
Arthur Charpentier	Université du Québec à Montréal (UQAM)	Discrimination algorithmique	Développement d'outils statistiques et de codes permettant d'identifier, mesurer et réduire les biais algorithmiques
Martin Gibert	Université de Montréal	Recherche et vulgarisation rigoureuse de l'IA et de l'éthique des technologies	Formation gratuite visant la vulgarisation rigoureuse des concepts fondamentaux de l'IA et de l'éthique des technologies
Sandrine Prom Tep	École des sciences de la gestion de l'Université du Québec à Montréal (UQAM)	Entreprises en fintech en matière de protection des consommateurs	Analyse des pratiques organisationnelles et des comportements des consommateurs dans l'encadrement éthique de l'usage des technologies financières
Jean-François Sénéchal	Université Laval	Recherche interdisciplinaire en intelligence artificielle	Podcast : Outil de réflexion en libre accès développé en collaboration avec des membres experts de l'IA



Axe

Santé durable

Sous la responsabilité de :

Martin Cousineau, Émilie Dionne et Marie-Pierre Gagnon

En collaboration avec :

Yvonne-Annick Abredan, Georges-Philippe Gadoury-Sansfaçon,
Lester Darryl Geneviève, Philip Jackson, Audrey Lebrét, Félix Pageau,
Marianne Ozkan et Pierrich Plusquellec



Actualité scientifique

Réglementation

L'actualité juridique a été marquée par l'adoption en septembre 2024 du premier instrument international de protection des droits de la personne en matière d'IA : la **Convention du Conseil de l'Europe sur l'intelligence artificielle, les droits de l'homme, la démocratie et l'état de droit**, que le Canada a signée le 11 février 2025. Cette Convention-cadre est un **instrument juridique transversal** qui s'appliquera à l'ensemble du cycle de vie de l'IA, y compris au développement et à l'utilisation de l'IA dans le domaine de la santé (à l'exclusion de la recherche et développement). En garantissant les principes de dignité humaine, de respect de la vie privée et de protection des données personnelles, de transparence ou encore de non-discrimination, la Convention contribuera à prévenir les risques démontrés que les systèmes d'IA font peser notamment sur l'équité et la non-discrimination en santé, en raison de biais algorithmiques parfois difficilement détectables. Le texte devrait également contribuer à assurer que l'IA en santé ne soit pas déconnectée des besoins réels des populations en imposant des consultations publiques et en promouvant les compétences numériques. Ces facteurs d'acceptabilité sociale devront être complétés par des mesures plus spécifiques au contexte canadien.

Concrètement, la Convention imposera des obligations aux acteurs publics ainsi qu'aux acteurs privés qui agissent pour le compte d'acteurs publics, de créer et mettre en application un cadre de suivi et d'évaluation des impacts de l'IA sur les droits fondamentaux. Bien que le texte le lui permettait, le Canada n'a pas spécifié dans une déclaration séparée au moment de la signature, comment il entendait mettre en œuvre les exigences de la Convention dans le cas de développement d'IA par des acteurs privés sans mandat d'une autorité publique. C'est là une des limites du texte que de permettre en pratique un double régime.

Les scribes IA réduisent la charge documentaire, mais produisent également des erreurs difficiles à repérer comme les confabulations textuelles.

Du point de vue des droits des usagers des services de santé, la Convention présente en outre de véritables garanties procédurales puisqu'elle imposera la création de mécanismes internes de garantie des droits, devant lesquels les autorités devront garantir un droit de recours effectif. Ces garanties internes seront complétées par un mécanisme international de suivi, devant le Conseil de l'Europe.

L'IA dans les pratiques professionnelles de soin

L'usage croissant de l'IA dans les soins transforme significativement les pratiques professionnelles. L'année 2025 a été marquée par une croissance des études et des évaluations explorant la diversité et l'ampleur de ces impacts. Ces dernières documentent, entre autres, le **risque de déqualification liée à la délégation de tâches cognitives complexes à des systèmes d'IA** génératives. En endoscopie, une étude multicentrique a montré une baisse d'environ 20 % de la détection d'adénomes lors d'examen non assistés après une exposition répétée à l'IA, indiquant une dépendance acquise rapidement (Budzyń et al., 2025). Dans les hôpitaux, l'évaluation des systèmes de prédiction du sepsis demeure inconcluante : **plusieurs travaux soulignent le risque de fausses alertes, de fatigue d'alerte et de surconfiance**, limitant la valeur clinique avant l'intervention (Gupta et al., 2024; Balczewski et al., 2023). En première ligne, les scribes IA réduisent certes la charge documentaire, mais ils sont responsables d'erreurs difficiles à repérer (p. ex. confabulations textuelles), mais pouvant se manifester ultérieurement, nuisant à la mémoire clinique (Associated Press/Ars Technica, 2024; Ng et al., 2025). À l'échelle organisationnelle, les litiges liés aux algorithmes de préautorisation des assureurs – en particulier concernant des refus automatisés sans examen médical – soulèvent des enjeux de transparence et de conformité juridique (Clason, 2025; NFP, 2025).

Souveraineté numérique

La souveraineté numérique en santé (hébergement et protection des données sensibles de santé) s'impose aujourd'hui comme un enjeu stratégique crucial pour la protection de la vie privée, la sécurité nationale et le développement économique (Geist et al., 2025; Hobson, 2025). Dans le contexte du Dossier santé numérique (DSN), le Québec a engagé environ 1,5 G\$ sur 15 ans auprès de la firme américaine *Epic Systems* pour moderniser les systèmes cliniques et mettre en place un dossier unifié, un projet dont les coûts globaux pourraient dépasser 3 G\$ (Lachance, 2025a, 2025b). Or, **le système Epic, développé initialement à des fins de facturation, présente des lacunes** en termes de capacité de capture, de stockage et d'analyse, **limitant ainsi les possibilités d'exploiter les données de santé avec l'IA** (La Torre, 2026).

Selon un commentaire publié dans CMAJ, une part importante des données de santé est aujourd'hui traitée sur **des infrastructures infonuagiques, largement contrôlées par de grandes entreprises américaines** (p. ex. Microsoft Azure, Amazon Web Services, Google Cloud). Sans cadre fédéral modernisé, le projet de loi C-27 demeurant en suspens, le Canada et ses provinces courent le risque de voir les conditions d'accès, de partage et de réutilisation de ces données dictées par des entreprises étrangères et des régimes juridiques qui ne reflètent pas nécessairement les attentes démocratiques ni les priorités de santé publique des Canadien-ne-s (Geist et al., 2025; Hobson, 2025; Zhang, 2025).

Pour contrer ces risques, la Loi sur les renseignements de santé et de services sociaux (Loi 5), entrée en vigueur le 1^{er} juillet 2024, instaure un cadre juridique détaillé encadrant la collecte, l'utilisation, la communication et la gouvernance des renseignements de santé, en y intégrant des exigences de gouvernance des technologies de l'information, des registres de services infonuagiques et des mécanismes de reddition de comptes (Gouvernement du Québec, 2024; Wyatt et Barry, 2024). Le projet de loi 82, adopté en octobre 2025 et devenu la Loi concernant l'identité numérique nationale et modifiant d'autres dispositions, pose, quant à lui, les fondements d'**une identité numérique québécoise centralisée** et renforce les exigences de sécurité, de traçabilité et de contrôle public sur les infrastructures numériques de l'État (Gouvernement du Québec, 2025a, 2025b). Dans le cadre des consultations de la Commission d'accès à l'information, la communauté de la recherche continue toutefois d'émettre des mises en garde contre les risques de centralisation excessive et de dérive vers des usages secondaires des données sans garanties suffisantes de transparence, de consentement et de reddition de comptes (Commission d'accès à l'information du Québec, 2025; Loiseau et al., 2025; Ligue des droits et libertés, 2025; Micrologic, 2025).

Ces réflexions rejoignent les travaux internationaux sur la souveraineté des données de santé, incluant ceux liés aux droits des peuples autochtones et sur les tensions entre ouverture des données, principes FAIR et mouvements de souveraineté des données autochtones (Cordes et al., 2024; Engstrom et al., 2024; FAIR 2025). Dans cette perspective, la souveraineté numérique en santé devient une composante non négligeable de la santé durable, éthique et sociale, au croisement des enjeux de gouvernance des données, de justice épistémique et d'autodétermination des communautés.



Débats scientifiques

L'impact de l'IA sur la relation soignant-soigné : défis et opportunités

L'IA s'invite au cœur de la relation clinique entre les soignants et les soignés, redéfinissant ainsi les rôles traditionnels des médecins, des infirmières et des autres professionnels de la santé. Même si elle apporte une certaine précision et un gain d'efficacité, elle soulève aussi des questions cruciales. Pour le médecin, elle offre un soutien logistique remarquable avec une « intelligence augmentée ». Des études récentes démontrent que l'IA peut traiter des volumes de données massifs pour affiner les diagnostics (Bazoukis et al., 2022; Han et al., 2024; Stade et al., 2024). Des outils comme Viz.ai ou MedGPT aident aux diagnostics ou encore analysent les scans en temps réel pour détecter des pathologies. Nabla ou Nuance DAX écoutent la consultation et rédigent le compte rendu clinique, permettant au médecin une interaction plus humaine. Ces technologies posent toutefois un défi de transparence : qui décide ou est responsable en cas d'erreur ? Comment sont gérées ces données ? (WHO, 2024; Reis et al., 2025).

La confiance constitue ainsi le point de bascule. Les patients font confiance à un professionnel, à une institution et à un processus, pas seulement à une recommandation (Reis et al., 2025). **La confiance ne porterait donc pas que sur le résultat, mais aussi sur le processus humain.** L'Ordre des infirmières et des infirmiers du Québec [OIIQ] (2024) suggère l'utilisation de l'IA pour l'automatisation des tâches administratives afin de « réhumaniser » les soins en libérant du temps pour être au chevet du patient. L'IA dans la relation de soins semble présenter trois risques majeurs : 1-des soins à deux vitesses automatisant la routine au détriment du contact humain (Lorenzini et al., 2023; Laranjo et al., 2025; OIIQ, 2024), 2-le renforcement de biais discriminatoires via des données non représentatives (WHO, 2024; 2025), et 3-une érosion des compétences si les soignants dévouent leur prise de décision aux outils d'IA (Lorenzini et al., 2023; Abdelwanis et al., 2024). Pour limiter ces dérives, l'Organisation mondiale de la Santé (WHO, 2024) a publié des directives claires sur l'utilisation de l'IA en santé. D'une part, le patient doit pouvoir consentir de façon éclairée à l'usage d'outils d'IA, comprendre les limites et connaître les voies de recours. D'autre part, **le clinicien doit approuver les réponses de l'IA car il reste légalement et moralement responsable** de la décision finale, même s'il s'appuie sur un algorithme (WHO, 2024).

Érosion des compétences

Des travaux en sciences cognitives montrent que l'IA modifie significativement les mécanismes d'apprentissage. L'acquisition de compétences repose sur le conflit cognitif et l'effort nécessaire à la consolidation en mémoire (Bjork et Bjork, 2011). Ainsi, l'usage croissant et accessible d'IA dans les pratiques et les organisations de santé amène de grandes transformations sur le plan des compétences professionnelles. En fournissant des solutions immédiates, **l'IA induit une décharge cognitive et réduit l'engagement intellectuel** (Macnamara et al., 2024). Des données neurophysiologiques montrent une baisse mesurable de l'activité neuronale lors de tâches assistées (Kosmyna et al., 2025), corrélée à une diminution de la capacité à produire un contenu autonome ou à se souvenir du processus de création. Ce mécanisme alimente une « paresse cognitive » documentée (Macnamara et al., 2024; Kosmyna et al., 2025) et affecte particulièrement l'apprentissage, avec une baisse de la pensée critique et du raisonnement complexe (Gerlich, 2025). Plusieurs études montrent toutefois que ce risque n'est pas structurel mais dépend du design et de l'usage (Miao et Holmes, 2023; José et al., 2025; Gerlich, 2025). Des auteurs recommandent donc **un encadrement strict de l'intégration de l'IA axé sur l'utilisation de systèmes conçus pour induire un conflit cognitif afin de renforcer la capacité d'analyse** (Miao et Holmes, 2023; José et al., 2025; Gerlich, 2025).

Face à ces risques, les autorités réglementaires renforcent l'encadrement. Santé Canada exige désormais des dispositifs IA qu'ils respectent les principes de bonnes pratiques d'apprentissage automatique (GMLP) et introduisent les Predetermined Change Control Plans (PCCP) pour gérer les modifications algorithmiques en cours de cycle de vie (Health Canada 2025). La capacité des générations futures à maintenir une pensée critique dépendra des environnements d'apprentissage où l'usage de l'IA n'élimine pas l'effort cognitif, mais l'oriente et le structure (Dwivedi et al., 2023).

IA et personnes âgées : quand l'innovation en santé renforce l'âgisme ?

Le recours à l'IA devient de plus en plus omniprésent dans le système de santé (Topol, 2019), en partie en raison du vieillissement démographique et d'une pénurie de professionnels de santé qualifiés, mais aussi étant donné la nécessité d'optimiser la prise en charge des usagers à moindre coût (Topol, 2019; Stypińska & Franke, 2023; Kakatum et al., 2025). La pression à l'innovation et l'engouement qui l'entourent ne sont pas non plus étrangers à cette utilisation (Mamdani, 2025; Markelius et al., 2024; Tian et al., 2024; Felber et al., 2023; Köttl et al., 2021). Bien que l'IA ait un potentiel considérable, son utilisation soulève néanmoins d'importants enjeux cliniques et éthiques.

Parmi ceux-ci, les discriminations, ainsi que les iniquités de santé qu'elles génèrent, appellent une attention particulière (van Kolfschooten, 2023; Obermeyer et al., 2019). **L'âgisme**, y compris dans ses formes intersectionnelles (Brewer, 2024) et internalisées (Köttl et al., 2021), **joue un rôle déterminant dans l'émergence de ces iniquités**. Ce sont des visions centrées sur des préjugés et des stéréotypes, selon lesquels les aînés seraient technologiquement illettrés, voire technophobes, souvent malades, fragiles, incompetents ou dépendants, ou encore moins ouverts au changement, contribuant ainsi à leur exclusion et à l'émergence de biais à différentes phases de développement et de l'utilisation de l'IA (WHO, 2022; Chu et al., 2022). Toutefois, **l'influence de l'âgisme sur le développement et la conception des systèmes d'IA reste insuffisamment examinée, voire négligée**, tant d'un point de vue du système de santé que du point de vue de l'éthique (WHO, 2022; Chu, 2022; Stypinska, 2023; Berridge et Revera, 2022). Ce constat est d'autant plus surprenant dans la mesure où l'âgisme est très répandu, continue de prendre de l'expansion, et constitue aussi le type de discrimination le plus socialement toléré au Canada (Revera, 2012). Ceci devrait précisément inciter à une vigilance accrue quant à sa reproduction, voire à son accentuation, dans les systèmes d'IA.

Autre paradoxe : les aînés figurent parmi les principaux utilisateurs des services de santé, représentant une part substantielle des coûts (Canadian Institute for Health Information, 2026), et les solutions fondées sur l'IA sont développées pour répondre à leurs besoins (Gormley, 2025). **Se soucier adéquatement des besoins des aînés et prévenir l'âgisme et ces formes de préjugés apparaît vital**. Il convient également de noter qu'à la suite de l'abandon, en janvier 2025, du projet de loi C-27, qui visait aussi l'introduction de la Loi sur l'intelligence artificielle et les données (LIAD), le Canada ne dispose d'aucun cadre législatif au niveau fédéral spécifiquement consacré à la gouvernance responsable de l'IA (Arai, 2025). Cependant, au Canada et au Québec, plusieurs initiatives visant à promouvoir une gouvernance éthique et responsable de l'IA, telles que la Déclaration de Montréal pour un développement responsable de l'intelligence artificielle (2018) ou la création de l'Obvia, ont contribué à structurer les réflexions et à orienter les pratiques. Bien qu'elles jouent un rôle important en matière de sensibilisation, de production de connaissances et d'accompagnement des acteurs, ces initiatives reposent principalement sur des cadres volontaires et ne relèvent pas de dispositifs réglementaires contraignants.

Ainsi, loin de relever d'un simple dysfonctionnement technologique, les discriminations âgistes liées à l'IA traduisent la cristallisation et l'accentuation des pratiques discriminatoires sociétales, dans lesquelles l'âgisme est devenu invisible à force d'être normalisé et toléré. En conclusion, sans une approche plus inclusive et réflexive devenue indispensable, l'IA en santé ne serait-elle pas finalement un moyen supplémentaire de légitimer l'âgisme sous couvert d'innovation ?



Questionnements

Les difficultés d'adoption de l'IA en santé

L'adoption de l'IA en santé demeure lente malgré de plus nombreuses études révélant plusieurs angles morts persistants. Les cadres actuels de certification et de gouvernance, conçus pour des technologies statiques, s'adaptent mal à des systèmes d'IA évolutifs, posant des enjeux non résolus de responsabilité, de conformité aux protocoles et de légitimité professionnelle (Le Monde, 2025).

Par ailleurs, si l'intention d'utilisation des systèmes d'IA par le personnel de la santé est bien expliquée par des facteurs individuels et organisationnels tels que la confiance, les attentes de performance et les conditions facilitatrices (Dingel et al., 2024), les solutions mises de l'avant demeurent largement centrées sur l'individu (p. ex., formation et gestion du changement). **Les enjeux structurels**, notamment les dynamiques de pouvoir, les responsabilités médico-légales, la gouvernance des biais et la reddition de comptes demeurent des enjeux **insuffisamment opérationnalisés dans les contextes réels de soins** (Darcel et al., 2023).

Ces constats soulèvent des questions structurantes pour le débat public et la recherche future :

- 1 **Comment gouverner des systèmes d'IA évolutifs en santé ?**
- 2 **Comment évaluer leurs impacts sociétaux au-delà de la performance clinique, notamment sur l'autonomie professionnelle, l'équité et la confiance du public ?**

Ces enjeux apparaissent centraux pour orienter l'adoption de l'IA en santé vers des usages socialement acceptables et durables.

Santé mentale et IA affective

Le développement des technologies relationnelles et conversationnelles propulsées par l'intelligence artificielle s'inscrit de plus en plus dans la vie quotidienne et soulève des enjeux importants en santé mentale. Plusieurs personnes interagissent aujourd'hui avec des agents conversationnels pour aborder des questions personnelles, de manière comparable à ce qui se ferait en contexte de relation d'aide, mais sans la présence d'un autre humain dans la boucle et avec des technologies qui n'ont pas été conçues dans ce but. Ces échanges informels répondent à des besoins variés, allant de l'introspection et du soutien émotionnel à la recherche de compagnie (Luo et al., 2025).

Pour certains utilisateurs, **les interactions avec des agents numériques prennent une dimension affective, et parfois intime**, incluant des formes de séduction et de proximité émotionnelle qui s'apparentent à des connexions amoureuses humaines (Zhang et al., 2025). Du soutien ponctuel aux confidences répétées, ces usages soulèvent des questions quant à la place de ces interactions dans le rapport à soi, aux autres et à l'intimité. Ils s'inscrivent dans un contexte social plus large, marqué par des phénomènes persistants de solitude et d'isolement relationnel (Surkalim et al., 2022).

Ces technologies peuvent créer des occasions nouvelles de dialogue, d'exploration de soi et d'expression émotionnelle, et constituer, pour certaines personnes, des **alternatives porteuses de sens lorsqu'un soutien humain est difficilement accessible**, avec un potentiel de réduction du sentiment de solitude ou de détresse relationnelle (Kim et al., 2025). **Des usages intensifs ou exclusifs peuvent au contraire contribuer à exacerber certaines vulnérabilités**, notamment en renforçant l'isolement social, en favorisant des formes de dépendance relationnelle ou en maintenant ou renforçant des dynamiques de détresse émotionnelle et affective (Zimmerman et al., 2024).

Certaines juridictions ont entrepris d'encadrer ces usages par des approches légales, comme en Californie avec le projet de loi SB-243, qui pose des balises en matière de sécurité et de transparence pour les compagnons amoureux IA. En parallèle, des initiatives de recherche, certaines menées à Montréal, explorent l'utilisation de ces technologies dans des cadres structurés afin de réduire la détresse relationnelle et soutenir le développement de compétences relationnelles (Lafortune et al., 2026).

L'IA au cœur de la relation d'aide

L'intelligence artificielle n'est plus une perspective futuriste pour les sciences de l'intervention psychosociale (Pagesy, 2024). Elle s'imisce en effet chaque jour un peu plus dans les pratiques cliniques. Les découvertes scientifiques s'accumulent, par exemple en informatique affective, que ce soit pour détecter l'autisme (Xia et al., 2025), le stress chronique (Xu et al., 2025) ou même la dépression (Niu et al., 2024) à partir d'une analyse automatique de la voix, des données de téléphones intelligents, ou d'expressions faciales. **Ces technologies pourraient offrir un soutien à l'évaluation en santé psychosociale.** D'autres travaux ont récemment démontré que l'utilisation d'un agent conversationnel thérapeutique pouvait entraîner des réductions significatives des symptômes de troubles mentaux cliniques tels que la dépression, l'anxiété généralisée et les comportements à risque d'un trouble de l'alimentation, avec des niveaux d'engagement élevés et une alliance thérapeutique perçue comme comparable à celle d'un thérapeute humain (Heinz et al., 2025). Face à ces découvertes et ce déferlement d'offres technologiques à venir, comment assurer la protection des personnes, le maintien des compétences et la responsabilité professionnelle ?

La psychoéducation, reconnue comme une innovation québécoise dans le domaine des relations humaines, a souhaité anticiper les dérives potentielles de ces nouveaux outils. L'Ordre des psychoéducateurs et psychoéducatrices du Québec (OPPQ) a publié en juin 2025 un guide de référence crucial : *L'intelligence artificielle dans la pratique professionnelle en psychoéducation*. Ce cadre éthique et déontologique réaffirme que la technologie ne doit jamais se substituer au jugement clinique. **Le principe fondamental défendu est celui de l'imputabilité totale de l'humain.** Le professionnel porte la responsabilité entière de tout acte posé, peu importe le système d'IA utilisé, et aucune erreur ne peut être justifiée par une défaillance algorithmique ou une « hallucination » de la machine. Cette posture garantit que l'IA reste un levier pour sublimer le savoir-être et la relation humaine, plutôt qu'un substitut à l'expertise professionnelle. Elle appelle cependant avec force des initiatives de formation en utilisation responsable de l'IA pour les professionnels de la santé psychosociale. L'Ordre des psychologues du Québec abonde dans le même sens et insiste sur l'importance du jugement clinique (humain) et de la littératie numérique des professionnels et du public (Ordre des psychologues du Québec, 2025a). Malgré l'existence de ces guides et recommandations, plusieurs questions demeurent quant à leur opérationnalisation (Ordre des psychologues du Québec, 2025).





Axe

Industrie 4.0, travail et emploi

Sous la responsabilité de :

Charles Antoine Rioux, Julie M.E Garneau et Christian Lévesque

Aujourd'hui, il est évident que les organisations traversent une phase de mutation profonde, souvent qualifiée de quatrième révolution industrielle ou, plus récemment, de cinquième et même de sixième révolution industrielle, marquée par une plus grande intégration des différentes technologies et de l'intelligence artificielle, ce qui rend plus concrète l'interaction entre l'humain et la machine dans nos différentes sphères de vie (Soumi et Bitan, 2026). Particulièrement au travail, l'IA s'insère dans une constellation d'autres technologies antérieures de mise en relation (internet et internet des objets), de collecte de données (capteurs, infrastructures fononagiques, mégadonnées), de robotisation et d'automatisation (robots conversationnels et logiciels de gestion et de prise de décision) (Garneau, 2021).

Elle permet une gestion algorithmique (GA), c'est-à-dire le remplacement des fonctions de gestion des biens et des personnes par des algorithmes plus ou moins autonomes (Parent-Rochelleau et Parker, 2021). Le développement de l'IA générative a ouvert de nouvelles possibilités en soutenant et en remplaçant des tâches liées à la production de textes, de sons, d'images, de vidéos ou de code logiciel. Regroupées en système multi-agent, ces IAG deviennent agentives, leur permettent d'adopter un comportement orienté vers un objectif ou une tâche de manière plus autonome (Datafranca, 2025), faisant de cette IA agentive une avancée majeure pour la transformation du travail en 2025 (Leonardi, 2025; Widad, 2025).



Actualité scientifique

Qu'elle soit adoptée ou en passe d'être adoptée, l'IA ramène l'éternelle question de la perte d'emploi. Bien que cette question ait longtemps suscité des réponses incertaines, elle retrouve aujourd'hui toute sa pertinence à la lumière des pertes d'emplois observées aux États-Unis, notamment dans le secteur informatique. Est-ce le canard dans la mine ? Il est encore trop tôt pour en être certain. Au Québec, selon l'Institut de la statistique du Québec (ISQ), près de 60 % des emplois présentent une forte exposition potentielle à l'intelligence artificielle, notamment dans des secteurs qualifiés tels que la finance, les services professionnels et les administrations publiques (2026). Cette exposition varie fortement selon le genre, l'âge, la scolarité et les régions, révélant des vulnérabilités différenciées. À cet égard, le *Bilan du marché du travail au Québec en 2025* de l'ISQ (2026) démontre une hausse du taux d'emploi au Québec, ce qui met à mal la thèse du grand remplacement de l'emploi par l'IA. On constate toutefois que ce bilan cache des subtilités sectorielles, notamment un déclin, entre 2024 et 2025, de plus de 5 % de l'emploi dans le secteur de l'industrie de l'information et de l'industrie culturelle, ainsi que de près de 2 % dans le secteur de la fabrication, entre autres.

Bien qu'une analyse plus fine doive être menée pour affirmer que ces pertes d'emploi sont liées à l'IA et à l'automatisation, lorsque l'on combine l'idée des effets différenciés de l'IA selon les types d'emplois et de tâches, la recherche nous amène, encore une fois, à la nécessité de mieux comprendre les contextes de travail entourant l'adoption de ces technologies, tels que les secteurs d'activité, les différentes stratégies d'affaires, les institutions réglementaires en vigueur et les relations sociales de production. À titre d'exemple, les études empiriques menées au Québec par Trottier (2025a, 2025b) montrent que l'IA est désormais bien implantée dans plusieurs professions, mais selon des modalités très différentes. Parmi les conseillers d'orientation, 52 % utilisent des applications d'IA pour traiter du texte, de l'image ou de l'audio, et 20 % utilisent des outils de prise de décision. Chez les technologues professionnels, 38 % utilisent des IA qui remplacent ou automatisent des tâches clés. Ces chiffres confirment une adoption réelle, mais tout de même dépendante de la finalité de l'application et du contexte professionnel.

Risques professionnels et socioéconomiques : compétences, emploi et impacts sur les travailleurs

Les travaux menés par les membres de l'axe au cours de la dernière année mettent en évidence plusieurs risques, regroupés en trois types, ainsi que des transformations socioéconomiques. D'abord, la **recomposition des compétences** : se projeter sur cette question semble encore difficile d'autant que la prolifération des cadres et taxonomies des compétences du XXI^e siècle engendre redondance et chevauchements sémantiques, rendant difficiles l'opérationnalisation des formations et la planification des parcours professionnels (Teyssier-Roberge et al., 2025). L'automatisation des tâches d'entrée de carrière réduirait également le « tunnel d'accès » vers des postes expérimentés, creusant un fossé entre les compétences requises et celles acquises par les jeunes diplômés. Parallèlement, **l'identité professionnelle** apparaît comme un facteur de résilience : une forte implication professionnelle favorise l'apprentissage continu et protège partiellement contre les risques liés à l'automatisation de certaines tâches (Saba et al., 2025). Bien que ce thème nécessite encore d'être creusé, ce qui demeure, c'est l'impression d'une difficulté à prévoir et à maintenir les compétences et les **institutions nécessaires** à ces transformations (écoles, programmes publics de requalification) tardent à se mettre en place. Tout ceci freine une transition réellement soutenable pour les travailleurs.euses touché.es.

Concernant la **qualité du travail** et **l'expérience quotidienne** au travail, qui sont des thèmes phares des recherches de l'Obvia, les effets observés sont **ambivalents** : certaines IA augmentent l'autonomie et la variété des tâches, tandis que d'autres intensifient le rythme et les exigences cognitives (Trottier, 2025a). Du côté du travail de plateforme, la **gestion algorithmique peut renforcer les formes de contrôle**, prendre la forme d'une intensification du travail, d'une dépendance aux évaluations numériques et d'une **faible capacité de négociation**, comme c'est le cas dans les travaux de Baril et Vultur (2025) portant sur la situation des livreurs sur plateformes à Toronto et à Montréal. Ces constats s'inscrivent dans une tendance observée au cours des dernières années et les recherches de l'axe s'inscrivent dans cette voie afin de mieux documenter et comprendre ces tendances et d'offrir aux décideurs publics des données pertinentes en vue d'une régulation de ces enjeux (Lévesque et al., 2025; Saba et Langlois, 2025).

Gouvernance, participation des travailleurs et pouvoir syndical

La gouvernance de l'IA reste fragmentée entre les initiatives publiques, les normes professionnelles et les dispositifs internes des organisations. La multiplication de balises et de cadres (responsabilité numérique, chartes) crée un paysage normatif hétérogène qui peut, paradoxalement, accroître la charge réglementaire pour les praticiens sans garantir une meilleure protection du travail (Coiquaud, 2024). La capacité des travailleurs à influencer l'implantation technologique est un enjeu majeur : là où des ressources syndicales et institutionnelles existent, des formes de négociation et d'encadrement peinent à limiter les effets délétères de la gestion algorithmique. Ailleurs, notamment sur les plateformes, les leviers collectifs sont faibles et les rapports de pouvoir se déplacent en faveur des concepteurs d'algorithmes et des intermédiaires. Enfin, la construction politique de l'industrie 4.0 met en lumière le rôle central des experts et des intermédiaires qui définissent les priorités technologiques et légitiment certaines trajectoires d'innovation au détriment d'autres finalités sociales, soulevant des questions sur la démocratisation de l'expertise et la place des travailleurs dans les décisions technologiques (Bowkett et Lévesque, 2026).

La prise en compte démocratique des changements technologiques et la place des sans-voix sont sans nul doute des questions prioritaires pour l'avenir du travail à l'ère de l'IA et du numérique. Il ne s'agit toutefois plus seulement de l'affirmer : il faut la mettre sérieusement en œuvre. Alors que ces changements technologiques se déploient sur fond de discours productiviste qui place la recherche de productivité comme un objectif suprême dans l'économie actuelle, notamment en réponse aux pressions d'une guerre géopolitique qui invisibilise, entre autres, l'apport des travailleurs et des travailleuses avec leurs savoirs techniques et expérimentiels, il est primordial d'insister sur les conditions nécessaires pour une réelle participation de ces acteurs.

La capacité des travailleurs à influencer l'implantation technologique est un enjeu majeur : là où des ressources syndicales et institutionnelles existent, des formes de négociation et d'encadrement peinent à limiter les effets délétères de la gestion algorithmique.

L'Obvia participe à cet objectif, notamment par le projet financé *Les syndicats, acteurs d'un futur du travail juste et responsable* (Vincent Pasquier, Xavier Parent-Rochelleau et Julie M.É Garneau), auquel participent 11 partenaires syndicaux. Ce projet vise, premièrement, à documenter la présence de l'IA et de la GA et, lorsque présentes, leurs impacts sur différentes dimensions du travail. Ensuite, dans un but de capacitation des travailleuses et des travailleurs et de leurs syndicats, le projet étudie les réponses syndicales et les possibilités de régulation, soit dans les conventions collectives, les politiques organisationnelles, soit dans les autres lois d'ordre public, afin d'accroître les possibilités de prise en charge démocratique de ces changements technologiques. Le but n'est pas d'interdire l'IA, mais de la déployer uniquement lorsque cela est nécessaire et de manière inclusive, afin de préserver la qualité de vie au travail. Cela passe notamment par le maintien de l'autonomie d'usage et de l'autonomie collective ; par une prise en charge en amont des risques liés à l'emploi, à la santé et à la sécurité ; ainsi que par la mise sur pied de lieux d'échanges et de dialogue continus et capables de répondre aux besoins des acteurs en place.

Le défi, face à ce discours dominant sur l'IA, consiste à poursuivre notre travail de défrichage sur le terrain afin de lever le voile sur les « petites histoires » de l'IA, en donnant la parole à celles et ceux qui l'utilisent, la développent et la façonnent dans leurs activités quotidiennes.



Débats scientifiques

Les travaux réalisés ces dernières années au sein de l'axe Industrie 4.0, travail et emploi ont permis d'identifier un écart croissant entre les narratifs sur l'IA centrée sur l'humain, favorisant le bien commun, et la réalité et les expériences concrètes vécues dans les espaces de travail. Cet écart amplifie le fossé qui se creuse entre la capacité cognitive des différentes formes d'IA et celle des acteurs et des institutions pour réguler ces technologies. Il en résulte un clivage de plus en plus prononcé en ce qui a trait à la valeur accordée aux savoirs technocratiques et expérientiels.

Il est encore ardu de saisir pleinement l'impact de l'IA sur le travail, tant dans ses dimensions qualitatives que quantitatives. Néanmoins, la majorité des analystes reconnaissent que nous sommes engagés dans une ère d'expérimentation façonnée par des logiques technocratiques et par les forces du marché.

Le défi, face à ce **discours dominant sur l'IA**, consiste à poursuivre notre travail de défrichage sur le terrain afin de lever le voile sur les « **petites histoires** » de l'IA, en donnant la parole à celles et ceux qui l'utilisent, la développent et la façonnent dans leurs activités quotidiennes. L'agrégation de ces récits de vie au travail permettra de construire un contre-narratif au discours dominant, un récit qui recentre l'attention sur l'humain et sur le bien commun. Ce contre-narratif est le premier pas vers l'élaboration et la construction d'institutions et d'une régulation au service de l'Humain.

Les différents enjeux et angles morts jusqu'ici mis en lumière permettent de dégager trois grands débats scientifiques :

- 1 Technologie vs structuration sociale
- 2 Contrôle vs participation
- 3 Compétences déclarées vs compétences réellement mobilisées



Questionnements

À partir des trois débats scientifiques identifiés, trois grands questionnements se posent :

1 Technologie vs structuration sociale

L'IA n'est pas une fatalité technologique. Elle est construite à travers des réseaux de pouvoir et adoptée dans des milieux de travail soumis aux rapports sociaux de production, c'est-à-dire à des formes d'organisation hiérarchique, des divisions de postes, des pouvoirs, à différents acteurs, à différentes cultures. L'IA n'est pas non plus neutre : son entrée bouscule ces rapports sociaux de production, au profit souvent des intérêts des dirigeants. Comment alors structurer cette technologie au service d'une adoption responsable ?

2 Contrôle vs participation

Les effets dépendent des ressources de pouvoir des employés et de leurs représentants et de leur capacité à baliser l'élaboration, l'adoption et l'usage de ces technologies en vue du maintien d'une qualité du travail. Plusieurs avenues s'offrent : la contrôler par des réglementations (à travers les conventions collectives, les politiques organisationnelles ou par l'adaptation des lois provinciales et fédérales) ou par la participation conventionnée à ces changements tout au long des cycles d'adoption. La véritable participation ne peut probablement pas se passer d'une prise de contrôle réelle, les deux étant intimement liées. Quelle(s) structure(s) de participation doit-on élaborer à cet effet ?

3 Compétences déclarées vs compétences réellement mobilisées

Les référentiels normatifs peinent à refléter les pratiques réelles. Quels savoirs et quelles compétences sont réellement nécessaires face à ces changements technologiques ? Les discussions sur les compétences techniques sont nécessaires, mais insuffisantes : il faut des compétences sur l'impact social de ces technologies (futur des emplois, futur de la qualité de vie au travail, futur environnemental; futur sur l'équité géopolitique). Comment doit-on favoriser des compétences élargies sur la technologie au travail ?



Axe

Arts, médias et diversité culturelle

Sous la responsabilité de :

Véronique Guèvremont et Colette Brin

En collaboration avec :

Jules Poulain-Welter et Yonn Calvez



Actualité scientifique

Culture

Les préoccupations liées à l'essor de l'IA dans les domaines des arts et de la culture se sont considérablement intensifiées au cours de la dernière année. En témoigne notamment la Conférence mondiale sur les politiques culturelles et le développement durable – MONDIACULT 2025, tenue à Barcelone du 29 septembre au 1^{er} octobre, qui a fait de l'IA l'un de ses thèmes prioritaires. Un rapport présenté par un groupe d'experts (UNESCO, 2025) s'ouvre sur un constat clair : l'IA progresse plus rapidement que les cadres de gouvernance culturelle, creusant les écarts et faisant émerger de nouveaux risques. Les experts soulignent toutefois que cette technologie offre également des possibilités inédites de renforcer l'innovation et la résilience culturelles.

L'accroissement du recours à l'IA est par ailleurs manifeste. Selon certaines estimations, 66 % de la population mondiale utiliserait désormais l'IA de manière régulière (Cardillo, 2025). Dans le secteur créatif, une enquête menée par Adobe en 2025 révèle que 85 % des créateurs recourent à des outils d'IA, que ce soit pour la création ou la réalisation d'activités connexes (Adobe, 2025). À titre d'exemple, 88 % des créateurs et concepteurs vidéo utilisent des systèmes d'IA dans leur travail (Lee, 2025).

Nous assistons ainsi à un véritable changement d'échelle dans la production de contenus générés par IA. Des études récentes estiment que près de 50 % des contenus en ligne seraient désormais entièrement ou partiellement produits à l'aide de systèmes d'IA (Jeandot, 2025). Dans le secteur musical, la plateforme Deezer a indiqué que 28 % des titres mis en ligne chaque jour sont générés par l'IA, contre 18 % seulement quelques mois auparavant (Mouriquant, 2025). Des dynamiques comparables sont observées dans l'univers vidéo : sur YouTube, environ 21 % des contenus publiés seraient générés par des systèmes d'IA (Boily, 2025). Face à cette prolifération, certaines plateformes ont entrepris des mesures correctives. Spotify, par exemple, aurait supprimé plus de 75 millions de pistes identifiées comme problématiques (Spotify, 2025). D'autres initiatives visent à lutter contre diverses formes de fraude ou de manipulation des contenus (Stassen, 2025).

Sur le plan législatif, l'Assemblée nationale a adopté en décembre 2025 la *Loi affirmant la souveraineté culturelle du Québec et édictant la Loi sur la découvrabilité des contenus culturels francophones dans l'environnement numérique*. Ce texte vise notamment à contrer la domination des contenus en langue anglaise sur les plateformes de diffusion. Rappelons à cet égard que les enjeux relatifs à l'IA et à la diversité linguistique constituent un champ de recherche investi par des membres de l'axe au cours des dernières années, notamment dans le cadre des *Dialogues interdisciplinaires* (Obvia et IVADO, 2024; 2026; Kulesz et al., 2025). Sur la scène canadienne, le Sommet national sur l'IA et la culture, tenu en mars 2026, a réuni des représentants de l'écosystème culturel et des milieux technologiques, des décideurs et des universitaires dans l'objectif de stimuler la croissance de l'économie culturelle tout en protégeant la créativité. Enfin, au niveau international, les membres du Comité intergouvernemental de la Convention de 2005 sur la protection et la promotion de la diversité des expressions culturelles ont franchi une étape significative en endossant la Recommandation 1 du Groupe de réflexion sur la diversité des expressions culturelles dans l'environnement numérique, laquelle propose l'adoption d'un Protocole additionnel à la Convention (UNESCO, 2025b). Un tel instrument pourrait conduire les États à rehausser leur niveau d'engagements face aux enjeux posés par les technologies numériques et l'IA. Il appartiendra à la Conférence des Parties de se prononcer sur cette proposition lors de sa 11e session prévue en 2027.

88 %
des créateurs et
concepteurs vidéo
utilisent des systèmes
d'IA dans leur travail

Médias

Les médias d'information adoptent deux principales stratégies face aux grandes entreprises d'IA générative : soit ils signent des accords de licence pour monétiser l'utilisation de leurs contenus dans l'entraînement ou les réponses générées par l'IA, soit ils engagent des poursuites judiciaires pour violation du droit d'auteur (Tobitt, 2026).

L'opposition des éditeurs de presse à l'utilisation des données par les IA prend trois formes : le blocage des robots d'indexation (*crawlers*), les recours judiciaires et, en Europe, le droit de retrait (*opt-out*) (Masseron, 2025). Le blocage des *crawlers*, option la plus courante, empêche les éditeurs d'IA de récupérer automatiquement les contenus journalistiques (Fletcher, 2024 : 3). Toutefois, un tel mécanisme n'est pas infaillible, puisque certains robots dont la visée d'IA générative n'est pas explicite peuvent contourner les blocages. Il peut aussi nuire à l'indexation d'un site web sur les moteurs de recherche (Liu et coll., 2025 : 89). En parallèle du blocage des *crawlers* et de l'*opt-out*, les médias ont intenté des procès à l'encontre des sociétés d'édition de systèmes d'IA, comme OpenAI ou Perplexity (Tobitt, 2026). Ces poursuites reposent essentiellement sur l'allégation du non-respect des droits d'auteur, nuisant au modèle d'affaires des médias; les médias plaident ainsi pour une plus juste répartition des revenus (Odell, 2025). C'est notamment le cas pour les poursuites intentées contre OpenAI par La Presse (2025) ou une coalition de médias canadiens composée de CBC/Radio-Canada, The Globe and Mail, La Presse canadienne, Postmedia et le Toronto Star. En Europe, les médias peuvent se prévaloir d'un droit de retrait des bases de données (*opt-out*) (Masseron, 2025 : 45).

D'autre part, de nombreux médias passent des accords avec les sociétés de développement d'IA générative afin d'accéder à une part des revenus issus de l'utilisation de leurs données (Tobitt, 2026). La plupart ont été signés avec OpenAI mais aussi avec d'autres entreprises, comme les ententes conclues par l'AFP avec Mistral (AFP, 2025a) et Perplexity (AFP, 2025b). Plusieurs de ces accords, dont les détails restent confidentiels, incluent le développement et l'accès à des outils IA dédiés au journalisme et conformes aux normes professionnelles. Les accords avec les entreprises d'IA prévoient généralement l'utilisation du contenu journalistique comme point de référence pour les requêtes des utilisateurs auprès d'agents conversationnels tels que ChatGPT (avec la promesse actuelle de renvoyer vers leurs sites web) et leur permettent d'utiliser cette technologie pour créer leurs propres produits (Tobitt, 2026).



Débats scientifiques

Culture

La relation entre l'IA et la créativité est désormais envisagée comme une véritable question de droits humains. Dans ce contexte, la Rapporteuse spéciale des Nations Unies dans le domaine des droits culturels met en garde contre une tendance persistante à célébrer le potentiel transformateur de l'IA sans adopter les mesures efficaces et urgentes nécessaires pour en assurer un usage conforme aux droits humains (Xanthaki, 2025). Comme d'autres auteurs, elle insiste en outre sur la nécessité de faire preuve de transparence au sujet des œuvres générées par l'IA afin de garantir le respect des droits culturels des créateurs comme des destinataires. Une transparence est en particulier requise vis-à-vis des titulaires de droits dont les œuvres sont utilisées pour l'entraînement des modèles d'IA ; l'absence de transparence porterait atteinte à leurs droits fondamentaux (Bensamoun, 2025).

Pour les développeurs de systèmes d'IA, l'exigence de transparence suppose de rendre publics les jeux de données utilisés pour entraîner leurs modèles, une démarche souvent présentée comme une condition essentielle pour assurer une rémunération équitable des titulaires de droits dont les œuvres auraient été exploitées (Buick, 2024). Or, certains considèrent qu'une telle approche peut porter atteinte au secret commercial ou secret professionnel (Sapkota et al., 2025). Sa mise en œuvre soulève également des difficultés pratiques importantes, notamment en raison de l'ampleur et de l'hétérogénéité des données concernées. Par ailleurs, l'exigence de transparence pourrait aussi viser les artistes et les professionnels de la culture, à travers des obligations de divulgation relatives à l'usage de l'IA dans le processus créatif (Buick, 2024). Divers mécanismes d'identification des contenus produits ou assistés par IA sont ainsi envisagés, notamment l'instauration de labels destinés à informer le public sur le degré de recours aux systèmes d'IA (Wittenberg et al., 2024).

Les médias d'information adoptent deux principales stratégies face aux grandes entreprises d'IA générative (Tobitt, 2026) :

1. Signer des accords de licence

OU

2. Engager des poursuites judiciaires pour violation du droit d'auteur

La souveraineté culturelle ne concerne cependant pas uniquement les États : elle appartient également aux communautés minoritaires et autochtones (Montgomery et Wagner, 2025 ; Fratini et al., 2024).

Enfin, les enjeux de souveraineté ont été au cœur de plusieurs initiatives analysant l'impact des technologies numériques sur les écosystèmes culturels. D'une part, la souveraineté des États se trouve confrontée aux défis posés par la souveraineté *de facto* exercée par les plateformes numériques (Bagaoui-Fradette et al., 2024). D'autre part, la réaffirmation de la souveraineté culturelle apparaît désormais indissociable d'une quête plus large de souveraineté numérique, entendue comme la capacité des États à exercer un contrôle sur les infrastructures, les données et les systèmes d'IA (Montgomery et Wagner, 2025). Cette interdépendance entre souveraineté numérique et culturelle est explicitement reconnue dans la liste des actions prioritaires élaborée à l'issue du colloque international « La culture face aux algorithmes : quel espace pour la souveraineté ? », co-organisé en octobre 2025 par la Chaire UNESCO sur la diversité des expressions culturelles et le CRDP, dont l'Obvia et la Coalition pour la diversité des expressions culturelles ont été partenaires (Chaire UNESCO sur la diversité des expressions culturelles, 2025). La souveraineté culturelle ne concerne cependant pas uniquement les États : elle appartient également aux communautés minoritaires et autochtones (Montgomery et Wagner, 2025 ; Fratini et al., 2024), lesquelles revendiquent la souveraineté sur leurs données ainsi que le contrôle des usages qui en sont faits. À cet égard, les lignes directrices énoncées dans le document « Territoire numérique des Premières Nations Québec-Labrador » énoncent le principe selon lequel « [l]es Premières Nations ont le droit de gérer, de posséder, de contrôler et de protéger leurs données, leurs ressources numériques et leurs informations culturelles, selon leurs valeurs, leurs protocoles et leurs priorités » (APNQL, 2025).

Médias

L'enjeu de l'amplification (*hype*) relative à l'IA suscite des débats à la fois sur ses causes, sa conceptualisation, sa nature et son impact sur les médias et le journalisme. L'amplification serait notamment liée à une couverture médiatique trop favorable, mettant en avant ses capacités et son caractère disruptif, la présentant comme désirable, plutôt que sur ses effets négatifs ou ses limites (Kuai, 2025 : 12; Nkoala et coll., 2025 : 16; Vrabič Dežman, 2024 : 754). Cela entraînerait une perception faussée du potentiel réel de l'IA et influencerait son développement et sa réglementation (Spyridou et Ioannou 2025 : 3). Cependant, certains chercheurs tempèrent cette vision : la couverture médiatique de l'IA mentionne tout de même ses effets négatifs et cherche à mettre en débat la technologie en incluant des voix critiques (Magalhães et Smit, 2025 : 15; Vicsek et coll., 2026 : 16).

Les débats sur l'amplification touchent également la recherche. Pour contrer l'amplification, Usher (2025) appelle à un « tournant critique » dans les travaux sur l'IA en journalisme. Zhang et coll. (2026) évoquent au contraire un « biais négatif » dans le traitement scientifique de l'IA en lien avec le journalisme. Quoiqu'il en soit, malgré la croissance exponentielle des publications sur le sujet depuis 2024, les chercheurs manquent de cadres analytiques communs pour comprendre ces phénomènes (Bartleman et coll., 2026).

Lewis et coll. (2025 : 2) proposent de considérer l'amplification entourant l'IA comme un objet d'étude en soi, une « exploration collective de l'avenir qui influence les activités du présent » (p. 3). Le *hype* serait lié à une vision de l'IA comme solution technologique à la crise des médias (Spyridou et Ioannou, 2025). Cette vision influencerait grandement les décisions des rédactions relativement à leurs politiques d'utilisation de l'IA (Schaez et Schjøtt, 2025 : 12). Afin de comprendre ce phénomène, Lewis et coll. (2025 : 5) ont développé un modèle permettant d'analyser les fonctions de l'amplification en journalisme, où se croisent la nature des effets d'emballement (symbolique ou matérielle) et son lieu (au sein du milieu journalistique ou dans l'espace public). Cependant, le phénomène lui-même reste matière à débat. En effet, selon Magalhães et Smit (2025 : 13), il s'agit davantage d'une controverse technologique (*technological drama*), où la technologie est le produit d'un processus conflictuel entre promoteurs de *hype* – en particulier les concepteurs d'outils technologiques – et ceux qui s'y opposent. Or, ces conflits se déploient autant dans les contenus médiatiques, laissant une place au débat et à la contradiction (p. 15), qu'au sein des rédactions, notamment entre les dirigeants et les syndicats (Ananny et Karr, 2025 : 12, 15).



Questionnements

Culture

Transparence et responsabilité dans la chaîne de valeur culturelle :

Quel niveau de transparence doit être exigé dans l'utilisation des systèmes d'IA, tant à l'égard des données d'entraînement que des œuvres produites ou assistées par l'IA, et selon quelles modalités ?

Concurrence entre création humaine et production automatisée :

Dans un contexte où l'IA participe de plus en plus directement à la production de contenus, lesquels concurrencent directement les œuvres des artistes, comment prévenir une captation disproportionnée de la valeur au détriment des créateurs ? Les cadres actuels du droit d'auteur et des politiques culturelles sont-ils suffisants pour préserver le statut des artistes et assurer leur juste rémunération, ou doivent-ils être repensés ?

Souverainetés culturelles et numérique et communautaire :

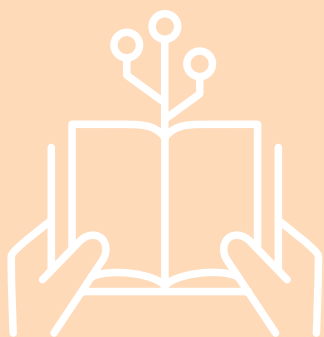
Comment articuler les concepts de souveraineté numérique et de souveraineté culturelle – incluant celle des nations autochtones – dans la régulation des données et des infrastructures culturelles, ainsi que l'encadrement des systèmes d'IA ? Comment assurer une gouvernance effective des technologies numériques face au pouvoir de *facto* exercé par les plateformes ?

Médias

Entre procès, régulation étatique et partenariats, le rapport de forces inégal entre les développeurs d'IA et les médias évolue rapidement. Quelle direction prendra cette relation au cours des prochaines années ?

Le cycle d'amplification (*hype*) de l'IA arrive-t-il à son terme et quelles en sont les conséquences pour le développement d'outils alimentés par IA au sein des salles de rédaction ?





Axe

Éducation et capacitation

Sous la responsabilité de :

Valéry Psyché et Ann-Louise Davidson

En collaboration avec :

Yassine El Bahlouli, Sandrine Prom Tep, Didier Paquelin,
Marie Laberge, Khaoula Boulaamane, Bruno Poellhuber et
Valérie Payen



Actualité scientifique

Cette section présente les préoccupations persistantes en éducation sur les avancées récentes :

Institutionnalisation de l'IA : processus incomplet et forte hétérogénéité des réponses

L'actualité scientifique et institutionnelle récente est marquée par la multiplication de cadres d'orientation visant à réglementer l'usage de l'IA en éducation à tous les niveaux, tant pour le corps enseignant que pour le corps apprenant. Au Québec, un guide gouvernemental sur l'intégration responsable de l'IA dans les établissements d'enseignement supérieur illustre une volonté de structurer les usages autour de principes pédagogiques, éthiques et organisationnels (Gouvernement du Québec, 2025). Deux études conduites par l'Obvia brossent un portrait des pratiques étudiantes et des pratiques enseignantes des outils d'IA générative (Paquelin et al., 2025a; 2025b). Plusieurs études canadiennes soulignent que les établissements oscillent encore entre encouragements prudents, tolérances encadrées et restrictions ciblées, notamment en matière d'évaluation, ce qui génère une incertitude normative pour le corps professoral et les personnes étudiantes (MacDowell et al., 2024; Proust-Androwkha et Denis, 2025). Une étude menée par Radio-Canada (2026) montre que 73 % des personnes étudiantes déclarent n'avoir bénéficié d'aucune formation formelle relative à l'usage de l'IA générative dans le cadre de leur cursus (Bussièrès McNicoll et al., 2026), ce qui soulève des questions fondamentales quant à l'accompagnement pédagogique offert par les institutions.

L'institutionnalisation de l'IA se manifeste aussi à l'échelle européenne : la Commission européenne (2024) a publié des lignes directrices éthiques destinées au personnel enseignant, mettant l'accent sur la protection des données, la transparence des systèmes et la responsabilité pédagogique. Ces orientations visent à guider la prise de décision dans un contexte où l'évolution technologique dépasse les cadres réglementaires.

73 % des personnes étudiantes déclarent n'avoir bénéficié d'aucune formation formelle relative à l'usage de l'IA générative dans le cadre de leur cursus

Fractures IA et reconfiguration des inégalités éducatives

Les recherches récentes convergent vers un constat central : la fracture associée à l'IA générative ne se limite plus à l'accès technique aux outils, mais concerne également des fractures de compétences, de compréhension critique et de pouvoir d'agir telles que déjà énoncées en regard des technologies par Van Dijk (2017). Des organisations comme l'UNESCO (2025) appellent à une intégration de l'IA générative centrée sur les capacités humaines, l'équité et l'éthique. Toutefois, ces cadres demeurent surtout normatifs, ce qui conduit à une mise en œuvre inégale et partielle (Yu et al., 2025). La cartographie des travaux sur l'IA en éducation réalisée par Yu et al., (2025) montre que les inégalités se déplacent vers la capacité à formuler des requêtes pertinentes, à évaluer la fiabilité des contenus générés et à comprendre les limites épistémiques des systèmes d'IA.

Les revues systématiques récentes confirment cette reconfiguration : si l'IA peut soutenir certaines tâches pédagogiques, ses effets positifs sont fortement conditionnés par le niveau de littératie numérique, informationnelle et métacognitive des personnes étudiantes (Park, 2025; Xiaoyu et al., 2025). En l'absence d'un encadrement pédagogique explicite, l'usage de l'IA générative tend à favoriser les personnes étudiantes, disposant déjà de compétences élevées en analyse, en vérification et en autorégulation, contribuant ainsi à renforcer des inégalités existantes (Lobet et al., 2024).

Ces constats soulignent qu'en plus des solutions techniques, la réduction des fractures liées à l'IA requiert des dispositifs structurés de formation et de capacitation, intégrés aux curricula et aux pratiques pédagogiques (Gouvernement du Québec, 2025; UNESCO, 2025). Aux côtés de ces enjeux, se situe également la question du manque de diversité (linguistique; culturelle) dans les modèles d'IA spécialisés (par exemple la surreprésentation des langues majoritaires comme l'anglais). Toupin (2024) souligne l'importance de remettre en question et de redéfinir ce qui constitue l'intelligence collective dans les systèmes d'IA (cité dans Lewis et al., 2024). Les réflexions actuelles sur l'usage de l'IA pour l'apprentissage devraient ainsi porter, non seulement sur la visibilité et la découvrabilité, mais également sur la manière de faire que ces nouveaux espaces numériques deviennent de véritables lieux d'apprentissage authentiques.

Effets sur les apprentissages et risques cognitifs émergents

Un axe majeur de l'actualité scientifique porte sur les effets cognitifs et pédagogiques de l'IA générative. Les synthèses récentes mettent en évidence une ambivalence marquée : les outils génératifs peuvent améliorer la qualité formelle des productions et offrir un soutien ponctuel à l'apprentissage, mais ils peuvent également favoriser des formes de substitution cognitive, réduisant l'effort intellectuel, la mémorisation et la mobilisation de stratégies métacognitives lorsque leur usage n'est pas didactiquement cadré (Xiaoyu et al., 2025). Une étude récente (Marois et al., 2025) montre que l'interaction avec un *Chatbot*, par rapport à celle avec un navigateur Web, peut affecter l'apprentissage à court terme et le rôle de l'effort mental, avec des résultats indiquant des différences entre les deux conditions dans la charge auto rapportée.

Ces préoccupations sont particulièrement vives dans l'enseignement supérieur, où l'usage massif de l'IA générative oblige à repenser les modalités d'évaluation et les objectifs pédagogiques (University Affairs, 2026). Les enquêtes récentes montrent que les institutions cherchent à diversifier les formes d'évaluation (orales, authentiques, processuelles) afin de préserver la validité des apprentissages (Freeman, 2025), au moment où 71 % des personnes étudiantes perçoivent l'IA comme un facteur de risque pour l'intégrité académique (Bussièrès McNicoll et al., 2026).

Les débats scientifiques soulignent le risque d'une dépendance accrue aux systèmes d'IA et la nécessité de soutenir explicitement le développement de compétences critiques et réflexives (Mollick, 2025; Park, 2025).

Développements technologiques récents et enjeux de gouvernance éducative

L'actualité récente met en évidence l'émergence de systèmes d'IA agentive, capables d'assumer des fonctions complexes d'assistance, de planification ou de tutorat éducatif (Palmer, 2026). Ces systèmes se distinguent de l'IA générative classique par leur capacité à décomposer des tâches, mobiliser des outils, s'autoévaluer et coopérer au sein d'architectures multi-agents, ouvrant de nouvelles perspectives pour la scénarisation pédagogique, le *feedback* adaptatif et l'évaluation (Chu et al., 2025; Jiang et al., 2025; Kamalov et al., 2025).

Les outils génératifs peuvent améliorer la qualité formelle des productions et offrir un soutien ponctuel à l'apprentissage, mais ils peuvent également favoriser des formes de substitution cognitive, réduisant l'effort intellectuel, la mémorisation et la mobilisation de stratégies métacognitives lorsque leur usage n'est pas didactiquement cadré

Plusieurs travaux récents montrent que la valeur ajoutée de ces approches réside dans l'orchestration de processus pédagogiques plutôt que dans la simple génération de contenus. Des cadres multi-agents sont notamment explorés pour la génération de problèmes à difficulté contrôlée, la diversification des points de vue cognitifs et l'assistance à l'évaluation, avec des résultats prometteurs mais encore expérimentaux (Karbasi et al., 2025; Zhang, 2025). Ces développements s'accompagnent toutefois de défis majeurs liés à la fiabilité des productions, à la traçabilité des décisions automatisées et à l'intégration institutionnelle de systèmes distribués (Córdova-Esparza, 2025).

Si ces innovations ouvrent de nouvelles possibilités pédagogiques, elles accentuent également les enjeux de gouvernance des savoirs, de responsabilité pédagogique et de contrôle des infrastructures éducatives. Lorsque des agents automatisent des fonctions traditionnellement humaines (sélection des contenus, structuration des savoirs, soutien à l'évaluation), se posent des questions centrales de redevabilité, d'imputabilité, de sens, d'authenticité, d'auditabilité et de dépendance cognitive. Dans le débat public, ces préoccupations sont parfois formulées en termes de concentration du pouvoir informationnel entre les mains de grandes entreprises technologiques (Spirlet, 2025).

Sur le plan scientifique et éducatif, plusieurs auteurs insistent sur la nécessité de recentrer l'intégration de l'IA générative, y compris agentive, sur le développement de l'intelligence humaine, plutôt que sur la seule optimisation des performances ou de la productivité (Luckin, 2025). Cette perspective rejoint directement les orientations capacitaires de l'axe Éducation et capacitation : l'IA ne constitue pas une finalité pédagogique, mais un objet critique et un levier conditionnel, dont l'usage doit être pensé dans une logique de développement du pouvoir d'agir, de l'esprit critique et de l'autonomie des acteurs éducatifs. Réussir une tâche avec l'IA ne mène pas automatiquement à un apprentissage. D'ailleurs, l'écart entre la « fausse maîtrise » et l'apprentissage réel dépend d'un usage de l'IA guidé par des principes pédagogiques clairs (OECD, 2026).



Débats scientifiques

Personnalisation algorithmique et équité éducative

La personnalisation algorithmique constitue l'un des arguments centraux en faveur de l'intégration de l'IA dans l'éducation et de l'équité éducative. Les systèmes d'IA adaptative promettent d'ajuster les contenus, le rythme et les modalités d'apprentissage aux besoins individuels, contribuant ainsi à une différenciation pédagogique accrue (Psyché et Ruer, 2019) et cela pourrait s'accroître avec l'IA générative (Xiaoyu et al., 2025). En adaptation scolaire, les études signalent des gains prometteurs en matière d'itération rapide, de personnalisation et de soutien aux élèves moins avancés - à condition de guider l'usage (Tishcoff et al., 2024).

Parallèlement, la littérature récente met en garde contre une standardisation masquée, où des modèles entraînés sur des données dominantes reproduisent, voire amplifient, des biais sociaux, culturels et linguistiques (Yu et al., 2025). À ce propos, Lewis et al., (2024) parlent d'un véritable « problème d'épistémologie » de l'IA. Ils soulignent que les données disponibles en ligne, et donc exploitées par les modèles d'IA, reflètent un état du monde inégal, surreprésentant les langues dominantes telles que l'anglais, le français et l'espagnol, ainsi que les perspectives occidentales et masculines, tandis que les connaissances des groupes minorisés, autochtones ou du Sud global sont peu présentes (Lewis et al., 2024).

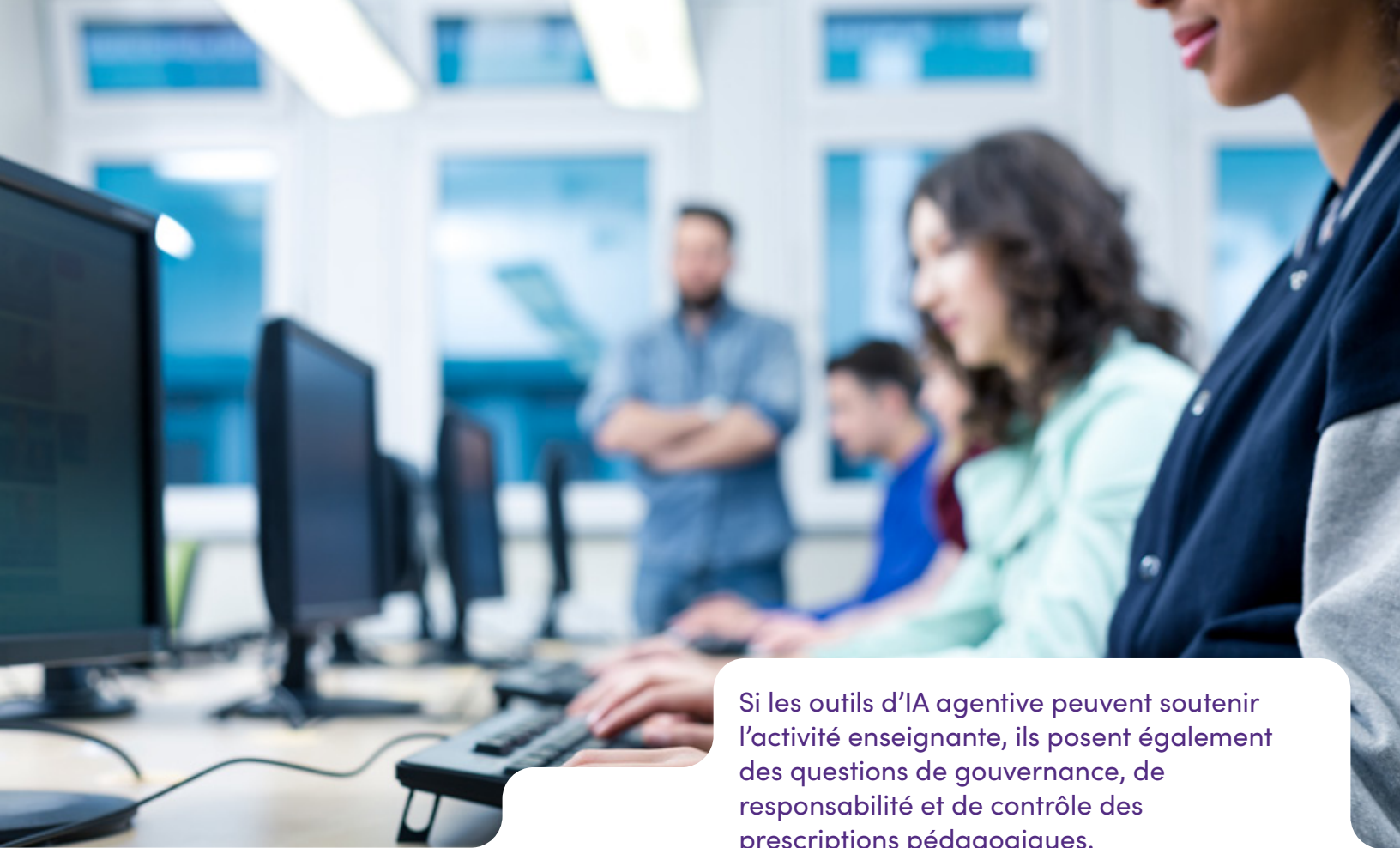
Cette tension renvoie directement à des enjeux d'équité et de justice éducative. Les travaux empiriques montrent que les bénéfices de la personnalisation algorithmique dépendent fortement des compétences préalables des apprenants en littératie numérique et informationnelle, ce qui risque de renforcer les inégalités existantes plutôt que de les réduire (Baig et Yadegaridehkordi, 2024; Park, 2025). Les cadres institutionnels récents insistent ainsi sur la nécessité d'une gouvernance explicite des usages de l'IA afin que la personnalisation serve des objectifs d'inclusion et de justice sociale, plutôt que l'optimisation des performances (Commission européenne, 2024; UNESCO, 2025).

Assistance algorithmique vs développement cognitif

Un deuxième débat structurant oppose l'assistance algorithmique au développement cognitif des personnes étudiantes. L'IA générative est fréquemment présentée comme un outil de soutien à l'apprentissage : aide à la rédaction, rétroaction et clarification conceptuelle (Baig et Yadegaridehkordi, 2024; Xiaoyu et al., 2025).

Les recherches récentes sur la littératie à l'IA montrent que ces effets varient fortement selon les conditions d'usage, la nature des tâches et le niveau d'expertise : sans encadrement pédagogique explicite, l'IA tend à être utilisée comme raccourci cognitif, alimentant une forme de dette cognitive et une fragilisation de l'esprit critique (Park, 2025; Lobet et al., 2024). Ce risque apparaît particulièrement pour les activités à forte valeur cognitive (argumentation, problématisation, justification, modélisation), notamment chez les novices, pour qui la production fluide peut masquer l'absence de compréhension. L'émergence de l'IA agentic, capable de planifier, d'auto-évaluer et d'agir de manière semi-autonome, intensifie cette tension en déplaçant la frontière entre assistance et délégation cognitive, et aussi entre aide et prise à la décision (Chu et al., 2025; Kamalov, 2025). Dans ce contexte, plusieurs auteurs plaident pour une intégration de l'IA centrée sur le développement des compétences métacognitives et réflexives, afin de préserver le pouvoir d'agir des personnes étudiantes (Luckin, 2025).

Sans encadrement pédagogique explicite : l'IA tend à être utilisée comme raccourci cognitif, alimentant une forme de dette cognitive et une fragilisation de l'esprit critique (Park, 2025; Lobet et al., 2024).



Si les outils d'IA agentive peuvent soutenir l'activité enseignante, ils posent également des questions de gouvernance, de responsabilité et de contrôle des prescriptions pédagogiques.

Automatisation pédagogique vs autonomie professionnelle

Le troisième débat porte sur les effets de l'IA sur le travail enseignant et l'autonomie professionnelle. L'automatisation de fonctions pédagogiques (génération de contenus, scénarisation, évaluation assistée) transforme le rôle du corps professoral et peut conduire à une normalisation implicite des pratiques pédagogiques, inscrite dans les choix techniques des outils (MacDowell et al., 2024; Proust-Androwkha et Denis, 2025).

Les développements en IA agentive et en systèmes multi-agents accentuent ces enjeux, en introduisant des dispositifs capables d'orchestrer des processus pédagogiques complexes (Jiang et al., 2025; Córdova-Esparza, 2025). Si ces outils peuvent soutenir l'activité enseignante, ils posent également des questions de gouvernance, de responsabilité et de contrôle des prescriptions pédagogiques.

Les approches socioconstructivistes et les pédagogies critiques offrent des cadres théoriques pertinents pour penser une intégration non instrumentale de l'IA, où les technologies demeurent subordonnées au jugement professionnel et expert du corps professoral et aux finalités éducatives (Luckin, 2025). Dans cette perspective, l'enjeu n'est pas de remplacer l'expertise pédagogique, mais de définir les conditions d'une coactivité humain-IA respectueuse de l'autonomie professionnelle et des valeurs éducatives (Mollick, 2024 ; Dyens, 2026).

Ces débats mettent en lumière des tensions structurantes qui traversent l'ensemble des usages éducatifs de l'IA. Ils montrent que les choix technologiques ne sont jamais neutres et qu'ils doivent être pensés à l'aune de leurs effets cognitifs, sociaux et professionnels. La section suivante approfondira ces enjeux sous l'angle des questionnements persistants et émergents, afin d'éclairer les orientations futures en matière de gouvernance éducative.



Questionnements

Les développements récents de l'intelligence artificielle générative et agentive, ainsi que les débats scientifiques qu'ils suscitent, font émerger une série de questionnements pour les systèmes éducatifs. Ces questionnements dépassent la seule problématique des usages pour interroger les conditions de gouvernance, les finalités éducatives et les formes de capacitation à l'ère de l'automatisation.

1 Quelles conditions d'acceptabilité pédagogique et sociale de l'IA en éducation ?

Un premier questionnement central concerne les conditions d'acceptabilité de l'IA en contexte éducatif. Les cadres institutionnels récents insistent sur la transparence, la protection des apprenants, l'équité et la responsabilité comme principes directeurs d'une intégration éthique de l'IA (Commission européenne, 2024; UNESCO, 2025; OECD, 2025). Toutefois, les travaux de synthèse montrent que l'acceptabilité ne peut être réduite à des normes formelles : elle repose sur la perception de la valeur pédagogique, le degré de contrôle laissé aux acteurs et la cohérence entre usages technologiques et finalités éducatives (Yu et al., 2025; Freeman, 2025).

Les enquêtes menées auprès des personnes étudiantes et du corps professoral indiquent que l'acceptation des outils d'IA est fortement conditionnée par la clarté des règles institutionnelles et par la reconnaissance explicite des usages autorisés et attendus (Freeman, 2025; Ravšelj et al., 2025). Dans ce contexte, l'adoption fragmentée ou perçue comme imposée, susceptible d'éroder la confiance des acteurs éducatifs, est à éviter (Gouvernement du Québec, 2025; Yu et al., 2025).

2 Comment préserver le développement des compétences humaines à l'ère de l'IA agentive ?

Un deuxième questionnement porte sur la préservation du développement des compétences humaines face à des systèmes d'IA de plus en plus autonomes. Les travaux récents révèlent une situation aux résultats mitigés : si l'IA peut soutenir certaines tâches cognitives, son usage non encadré peut favoriser une substitution cognitive, un délestage cognitif (*cognitive offload*) et une fragilisation de la métacognition (Baig et Yadegaridehkordi, 2024; Xiaoyu et al., 2025; Park, 2025).

Ces problématiques s'inscrivent dans un questionnement plus large sur la fiabilité et la qualité des ressources produites ou augmentées par ces systèmes. L'un des enjeux concerne l'utilisation de données synthétiques générées par des modèles d'IA pour compléter ou corriger des données réelles. Cet enjeu s'avère particulièrement important en éducation, puisqu'il influe indirectement sur la production de connaissances et peut, selon les contextes, atténuer ou, au contraire, renforcer des biais préexistants (CPVP, 2022).

L'émergence de l'IA agentive intensifie ces enjeux en brouillant davantage la frontière entre assistance et délégation cognitive et invite à penser les usages sous l'angle de l'ambidextrie. Les agents capables de planifier, de s'auto-évaluer et d'agir de manière semi-autonome posent la question de ce qui doit rester du ressort de l'humain dans les processus d'apprentissage (Chu et al., 2025; Kamalov et al., 2025). Plusieurs auteurs soulignent la nécessité de redéfinir explicitement les objectifs éducatifs prioritaires, en mettant l'accent sur le jugement critique, la réflexivité et la capacité à interagir de manière éclairée avec des systèmes automatisés (Luckin, 2025; Yu et al., 2025). Se pose ainsi la question de revenir à des formats d'évaluation résistants à l'usage non déclaré (oraux, examens en classe, portfolios, diverses épreuves de processus, journaux de bord), et d'accentuer l'usage de la détection des traces d'apprentissage au-delà des connaissances exposées à l'écrit. D'autres études soulignent le potentiel et les écueils de l'évaluation des compétences lorsque l'IA générative est utilisée pour faciliter la rétroaction et l'autoévaluation des apprentissages (École Branchée, 2026).

3 Quelle gouvernance éducative face à l'automatisation et à la concentration du pouvoir technologique ?

Un troisième questionnement majeur concerne la gouvernance des systèmes éducatifs dans un contexte d'automatisation croissante. Les outils d'IA, notamment lorsqu'ils intègrent des fonctions d'évaluation, de scénarisation ou de recommandation, peuvent introduire des formes de standardisation implicite des pratiques pédagogiques et renforcer la dépendance à des infrastructures technologiques contrôlées par un nombre restreint d'acteurs (MacDowell et al., 2024; Córdova-Esparza, 2025).

La littérature récente met en évidence des risques liés à la concentration du pouvoir informationnel, susceptibles d'influencer la circulation, la hiérarchisation et la validation des savoirs éducatifs (Spirlet, 2025). Ces enjeux soulèvent des questions de responsabilité, de redevabilité et de souveraineté numérique, appelant des mécanismes de gouvernance capables d'assurer la transparence, l'auditabilité et l'adaptabilité des systèmes d'IA utilisés en éducation (UNESCO, 2025; Commission européenne, 2024).

4 Quelles interventions pour une approche capacitaire de l'IA en éducation ?

L'ensemble de ces questionnements converge vers une question centrale : la place de l'IA dans une approche capacitaire de l'éducation. Plutôt que de la voir comme une simple solution technologique à optimiser, plusieurs auteurs invitent à la penser comme un objet critique, dont l'intégration doit être conditionnée par sa contribution effective au développement des capacités humaines et sociales (Luckin, 2025; UNESCO, 2025).

Les politiques éducatives et les pratiques pédagogiques se retrouvent donc face à un choix structurant : suivre une logique d'optimisation technologique ou inscrire l'IA dans des trajectoires éducatives orientées vers l'autonomie, l'équité, la responsabilité (Yu et al., 2025; Gouvernement du Québec, 2025) et les capacités. La réponse requiert un arrimage entre l'expertise pédagogique, la régulation institutionnelle et la participation active des acteurs éducatifs.

Ces questionnements montrent que l'intégration de l'IA en éducation ne relève pas d'un simple ajustement technique. Elle implique des choix politiques, pédagogiques et éthiques durables, qui influencent la formation de citoyens autonomes et critiques. Cette intégration requiert également le développement d'une littératie de l'IA qui s'appuie idéalement sur une littératie numérique déjà présente (Van Dijk, 2006). Or, ce n'est pas toujours le cas, ce qui peut accentuer la fracture, voire entraîner l'exclusion numérique (Prom Tep et al., 2023).

Au Québec, plusieurs initiatives émergent pour soutenir les compétences en IA : communautés de pratique en littératie de l'IA, bootcamps et hack'apprendre pédagogiques sur l'IA, jeux éthiques, etc. À cela s'ajoutent des innovations telles que la communauté LitteratIA de l'UQAM, le jeu éthique en IA de l'Université Concordia, le Bootcamp IA de l'Université Laval et le HackIAthon organisé par l'UdM à travers le LAVIA, qui constituent des espaces d'expérimentation et de capacitation collectives.

Ainsi, pour l'axe Éducation et capacitation, l'enjeu central reste de faire de l'IA un levier au service du pouvoir d'agir (Luckin, 2025; Yu et al., 2025). Son intégration dépendra moins de ses performances techniques que de la capacité des milieux éducatifs à orienter son développement vers des finalités démocratiques, inclusives et émancipatrices.



Axe

Sobriété numérique et transition socio-écologique

Sous la responsabilité de :

Christophe Abrassart et Stéphane Roche

En collaboration avec :

Mélie Ravayrol, Florent Begue et Alexandre Gourret



Actualité scientifique

Les systèmes d'IA pour l'environnement : au-delà des discours, quelles méthodologies pour évaluer leurs bénéfices pour la transition écologique ?

Les systèmes d'intelligence artificielle (SIA) sont souvent présentés dans les discours publics des organisations internationales et des entreprises comme des solutions à fort potentiel pour résoudre la crise climatique, la crise de la biodiversité ou d'autres enjeux environnementaux, comme s'il y avait un large consensus scientifique à ce sujet. Or ces promesses sont souvent assez vagues et peu démontrées, et les raisonnements qui permettraient de soutenir de telles affirmations ne sont pas toujours bien précisées (Roussilhe, 2023).

Par ailleurs, ces discours parlent généralement d'IA au singulier, sans mentionner que ces SIA pour l'environnement relèvent surtout de programmes traditionnels spécialisés (ex. classification et reconnaissance d'images satellites par apprentissage supervisé), différents des IA génératives grand public de type LLM qui sont actuellement associées à une empreinte environnementale croissante dans le monde (IEA, 2025; ADEME, 2025; Shift Project, 2025). Pour certaines ONG, cette indifférenciation entre types d'IA relève d'une forme de rhétorique par amalgame, visant à créer une association positive entre IA et environnement (Joshi, 2026).

Dans ce chapitre, nous présentons l'actualité scientifique sur une dimension particulière des SIA pour l'environnement : les méthodes d'évaluation développées pour mieux appréhender leurs coûts et bénéfices environnementaux, notamment par des raisonnements sur les effets et les conséquences de leur adoption. En effet, des études récentes s'efforcent d'analyser les effets et les conséquences de l'usage du numérique pour l'environnement (ADEME, 2025a), notamment sous l'angle des nouvelles consommations de ressources, des effets rebonds, des effets de verrouillage et de dépendance de sentier. Dans le sillage des enquêtes en sciences politiques sur les instruments de gouvernement (Lascoumes et Le Galès, 2001), on se demandera aussi quels sont les effets de cadrage et de problématisation implicites des SIA pour l'environnement. Et de façon plus générale, dans quels types de transition écologique nous orientent ces outils (Parasie et Shultz, 2024) ? À la suite d'une synthèse de plusieurs rapports, nous proposons à notre tour une grille d'analyse enrichie, pour penser de façon critique la pertinence de ces outils pour soutenir une transition écologique compatible avec les limites planétaires.

De nouvelles méthodes d'évaluation conséquentielles des SIA pour l'environnement

Des recherches récentes abordent la question des méthodes d'évaluation des systèmes numériques et des SIA pour la transition écologique. C'est notamment le cas de l'étude « IT4Green » coordonnée par l'ADEME (2025b) avec plusieurs bureaux d'étude (Mavana, DDemain et Hubblo) pour mieux faire la part des choses entre l'empreinte carbone croissante du numérique et les promesses de ses technologies pour la transition écologique. Son objectif est d'évaluer les impacts environnementaux nets de solutions numériques. Neuf cas d'usage ont été analysés.

Les cinq premiers ont fait l'objet d'une publication (ADEME, 2025b) : la numérisation de l'éclairage public, le pilotage dynamique des lignes haute tension (DLR), la numérisation d'une offre de gestion de pneumatique (Tire as a Service), l'évolution du télétravail, et la numérisation de l'épandage d'engrais azotés. Et quatre nouveaux cas d'usages sont à paraître en 2026 : la numérisation du covoiturage journalier, la numérisation de la détection de fuites d'eau dans les canalisations, la Gestion Technique des Bâtiments Tertiaires, et enfin la numérisation du suivi de l'autoproduction photovoltaïque.

Cette étude est particulièrement intéressante, car elle développe une réflexion méthodologique approfondie sur la manière d'évaluer le bilan environnemental des solutions numériques étudiées. L'approche suivie par cette étude mobilise une « **méthode conséquentielle** d'évaluation des **impacts nets** ».

Tout d'abord, pour chaque cas d'usage, « les conséquences directes et indirectes induites par le déploiement de la solution numérique étudiée (qu'elles soient environnementalement bénéfiques ou non) » sont évaluées, et sont représentées dans un « **arbre de conséquences** » s'appuyant sur une **typologie des effets** liés à l'utilisation d'une solution numérique (IA ou plus simple) pour l'environnement (ADEME, 2024).

Taxonomie choisie		Ordre des effets	Description
Effets directs	Extraction des matières premières	1 ^{er} ordre	Matières premières pour la fabrication des équipements numériques
	Fabrication		Processus de fabrication et d'assemblage des équipements et infrastructure numériques
	Installation		Distribution et installation de l'équipement sur son site de mise en service.
	Usage		Usage des équipements et des infrastructures numériques : consommation d'électricité, aux opérations de maintenance, de reconditionnements et aux consommables
	Fin de vie		Fin de vie des équipements et des infrastructures
Effets indirects	Effet d'induction	2 ^e ordre	Nouvelles activités permises par la solution numérique
	Effet de substitution		Décommissionnement d'un système et de ses usages au profit d'un autre système permettant d'effectuer des activités similaires et incluant la solution numérique
	Effet d'optimisation		Réduction de la quantité de ressource et/ou d'énergie par unité d'activité permis par la solution numérique
	Effet rebond direct	3 ^e ordre	Augmentation de la consommation d'une activité suite à l'amélioration de son efficacité
	Effet rebond indirect		Augmentation de la consommation d'une activité suite à l'amélioration de l'efficacité d'une autre activité
	Effet rebond macroéconomique		Transformations de l'économie et de la société en quantité et en nature à des échelles importantes
	Transformations sociétales		

Tableau : taxonomie des effets utilisés dans l'étude ADEME (2024, p. 14-15).

Ensuite, « l'évaluation des impacts a lieu par comparaison avec un scénario de référence qui modélise un état alternatif dans laquelle la solution n'a pas été déployée. » Les **impacts nets** recherchés évaluent ainsi la différence entre le scénario déployant la solution numérique et le scénario de référence sans cette solution numérique, ce qui peut être résumé par la formule suivante (ADEME, 2024) : « **Impact net** = impacts ajoutés de la solution évaluée - réduction d'impact sur le secteur concerné. »

À noter que pour les cas d'usage étudiés, ces impacts nets sont estimés jusqu'à 2035. De plus, pour certains cas, le scénario de référence n'est pas maintenu dans un état statique, mais a conduit à une variante améliorée qui pourrait être qualifiée de low-tech. Par exemple, dans le cas de la numérisation de l'éclairage public, le scénario de référence considéré comprend des améliorations techniques de base (ex. l'incorporation d'ampoules LED à basse consommation d'énergie).

Enfin, la méthode d'évaluation par Analyse de cycle de vie (ACV) multicritère est mobilisée pour prendre en compte l'ensemble des flux du cycle de vie physique lié aux solutions étudiées, et quantifier plusieurs indicateurs d'impacts (changement climatique, épuisement des ressources abiotiques, consommation d'eau etc.).

Cette étude, qui rassemble les principes méthodologiques les plus récents en les appliquant sur plusieurs cas d'usage, constitue une réelle avancée. Mais pourrait-elle être encore enrichie ? C'est ce qui sera proposé dans la dernière section du texte, qui discutera d'un cadre d'analyse enrichi pour concevoir et analyser des SIA pour la transition écologique.



Débats scientifiques

Les SIA pour l'environnement, un nouveau champ de recherche, mais une interdisciplinarité à construire

Avec le développement de nombreux projets de SIA pour l'environnement et le climat par des entreprises ou des laboratoires de recherche universitaires, un nouveau courant de recherche thématique en IA appliquée s'est organisé depuis le début des années 2020. C'est en particulier le cas du réseau de recherche thématique *Climate Change AI* (CCAI), fondé en 2019 avec la publication d'un article inaugural (Rolnick et al., 2022) rassemblant une vingtaine de chercheurs pour établir une première synthèse des applications potentielles des SIA pour le climat dans une dizaine de champs thématiques (des systèmes électriques à la gestion des bâtiments, en passant par les transports, l'industrie, l'agriculture et la gestion de la biodiversité). Ce réseau de recherche s'est depuis fortement développé avec l'organisation de conférences annuelles, et agit auprès des décideurs publics comme une communauté épistémique.

Comment la question de l'évaluation des coûts et bénéfices environnementaux est abordée dans ces travaux ? Des méthodes originales d'évaluation y ont-elles été développées ? Si on aborde les cas de SIA pour l'environnement présentés dans les articles et conférences de ce réseau comme des « preuves de concept », de quel type de preuve s'agit-il ?

L'article fondateur de ce réseau (Rolnick et al., 2022) se donne pour objectif d'identifier et de cartographier des problèmes liés au changement climatique (sur la connaissance du phénomène, sur l'adaptation ou sur la réduction des émissions) pour lesquels l'apprentissage machine (« ML » en anglais) présente une contribution potentielle : « We identify problems where ML can improve public services, help gather data for decision-making, and guide plans for future development. »

Cet article qui est riche en exemples présente de plus un discours résolument optimiste sur la performance attendue de ces algorithmes : « This article aims to provide an overview of where ML can be applied with high impact in the fight against climate change » (Rolnick et al., 2022 : 2). Plus loin, l'article parle de « ML with major impact », de « highly impactful ML », ou encore de « massive efficiency ML » (Rolnick et al., 2022 : 3). Cependant, il ne systématise pas un cadre méthodologique pour évaluer cette performance. Seuls quelques éléments de réflexion sont esquissés à ce propos.

Face à des possibles pièges de l'« IA for Good », qui sont juste mentionnés, l'article appelle à une meilleure connaissance sur le « comment » : *Despite the growth of movements applying ML and artificial intelligence (AI) to problems of societal and global good, there remains the need for a concerted effort to identify how these tools may best be applied to tackle climate change* (Rolnick et al., 2022 : 2, [mise en valeur ajoutée]).

Ensuite, l'article esquisse une typologie pour classer les contributions potentielles du ML aux enjeux climatiques, avec trois catégories : (1) **Effet de levier élevé**, pour surmonter des situations de goulot d'étranglement (« bottlenecks ») dans l'action climatique ; (2) **Contribution de long terme**, après 2040 ; et (3) **Impact incertain**, pour lesquelles un effet rebond ou d'autres effets indésirables, comme des externalités négatives, pourraient apparaître.

Cependant, cette typologie ne fait pas l'objet d'une réflexion plus avancée dans l'article, même si des argumentaires très pertinents sur les risques d'effet rebond ou de dépendances de sentier sont parfois développés à propos de certaines solutions dans le corps de l'article (ex. en énergie, transports, agriculture). En quoi consiste une situation initiale de goulot d'étranglement pour les auteurs (ex. la capacité limitée du système administratif à gérer efficacement et à temps un enjeu environnemental) ? Pourquoi et comment les outils de ML pourraient-ils aider à surmonter ce type de situation plus efficacement qu'une approche *low-tech* comme une innovation organisationnelle ou une meilleure planification ? Quels sont les impacts environnementaux sur le cycle de vie des solutions d'IA proposées, en comparaison des gains environnementaux attendus sur les systèmes auxquels ils s'appliquent (Ligozat et al., 2022) ? Ces solutions de ML sont-elles toutes compatibles avec des budgets de carbone restants très limités (IPCC, 2023) ? L'article ne confronte pas non plus des scénarios de réponse alternatifs, avec ou sans ML, comme le recommande la norme AFNOR sur l'IA frugale (AFNOR, 2024).

Pourquoi, par ailleurs, un horizon de 2040 ? Est-ce lié à une faible maturité des technologies envisagées (TRL bas) ou à un positionnement temporel dans une trajectoire de décarbonation profonde (pour soutenir en 2040 des changements de paradigme, après avoir optimisé dans un premier temps le système actuel) ? Enfin, l'effet rebond n'est-il pas susceptible d'apparaître dans un grand nombre de cas de SIA pour le climat ?

Cet article présente donc une esquisse de réflexion sur la manière d'évaluer les coûts et bénéfices environnementaux des SIA pour le climat, mais ce cadre pourrait être significativement enrichi. Cette thématique des méthodes d'évaluation pourrait de plus enrichir l'agenda des travaux de la communauté CCAI, à condition de la développer de manière interdisciplinaire avec les sciences sociales (voir Green (2019)).



Questionnements

Vers un cadre d'analyse enrichi pour concevoir et analyser des SIA pour la transition écologique

Un questionnement demeure problématique. Il concerne la différence entre deux types d'effets, qui se situe au cœur des problématiques de transition écologique contemporaine : d'une part la **logique d'optimisation d'un système existant et connu** (ex. un parc de bâtiments, un système de transport, un système alimentaire), et, d'autre part, la **logique d'exploration de scénarios de sobriété inconnus**, pour faire advenir au plus tôt des bifurcations ou des changements de paradigme tenant compte des limites planétaires. La première logique relève de la gestion de l'environnement, la seconde se situe dans une logique de transition écologique.

Dans le premier cas, les solutions numériques peuvent en effet conduire à des gains environnementaux à court terme, mais aussi à des effets rebonds de premier ordre et à des effets de renforcement de « dépendance de sentier » et de retardement du déploiement de mesures rivales à plus fort potentiel. C'est par exemple le cas de solutions numériques de gestion des trajets des véhicules automobiles dans les villes dites intelligentes, qui ne

remettent pas en question la place centrale des voitures ni leur identité (VUS avec propriétaire). Dans le second cas, les solutions numériques peuvent être des outils d'exploration de l'inconnu technique (ex. l'utilisation d'une IA pour de nouveaux matériaux plus légers pour les batteries des véhicules électriques), ou des solutions au service de la transition vers un nouveau paradigme de sobriété inconnu mais souhaitable face aux limites planétaires (ex. des infrastructures numériques permettant de sécuriser l'introduction de Véhicules légers intermédiaires, ou VÉLIs, des véhicules électriques ultralégers, sur les routes, en coordination avec les transports en commun).

Autre exemple, en agriculture, la première logique peut conduire à une agriculture de précision renforçant le modèle des grandes monocultures à haute productivité, mais toujours consommatrices d'engrais et de pesticides, ou bien rendre possible la transition vers un modèle agroécologique s'appuyant sur une rotation de cultures comportant certaines inconnues.¹

L'opposition entre ces deux logiques contradictoires peut être représentée dans le tableau suivant :

	Logique d'optimisation du connu (ou du statu quo)	Logique d'exploration de scénarios de sobriété inconnus
SIA pour des systèmes sociotechniques	Exemples : gestion des trajets des voitures classiques dans la ville intelligente; agriculture de précision	Exemples : sécurisation des VÉLIs sur les routes; aide à la transition vers l'agroécologie
SIA pour des systèmes naturels et techniques	Exemples : aide à la prédiction d'évènements climatiques extrêmes	Exemple : recherche de nouveaux matériaux pour des batteries de véhicules plus légères

Schéma : Distinction entre SIA pour l'optimisation du connu et SIA pour l'exploration de scénarios de sobriété inconnu (Christophe Abrassart).

Cette distinction pourrait être considérée plus systématiquement dans les études d'évaluation du numérique et des systèmes d'IA pour l'environnement. L'enjeu est notamment de passer d'une logique de gestion de l'environnement à une logique de soutien à la transition écologique pour imaginer des scénarios de sobriété systémique compatibles avec les limites planétaires.

¹ Voir Bellon-Maurel et al. (2026 :123-126) sur un cas de phénotypage de luzerne pour enrichir les sols en azote tout en étant compatible avec la culture de céréales, par entraînement d'un modèle d'IA à partir d'images de champs cultivés).

Questions pour le débat public et priorités de recherche

Pour conclure, nous proposons quatre recommandations qui pourraient être suivies

- 1 Le développement de méthodes rigoureuses pour l'évaluation de l'**impact environnemental net** de l'utilisation d'une solution numérique ou d'IA pour l'environnement, comme celle développée récemment par l'ADEME autour des **analyses conséquentielles par types d'effets**.
- 2 La comparaison systématique avec un **scénario de référence enrichi** (sans IA, avec du numérique plus *low-tech*), ce qui demande une réelle interdisciplinarité.
- 3 Pour les systèmes sociotechniques, la distinction systématique entre gestion de l'environnement et soutien à la transition, c'est-à-dire entre des **solutions numériques d'optimisation** des systèmes existants (avec forts risques d'effet rebond et de dépendance de sentier) et celles de **soutien à un changement de paradigme** dans une **logique de sobriété systémique** tenant compte des limites planétaires.
- 4 Le positionnement des solutions de numérique ou d'IA dans une **trajectoire de décarbonation profonde** 2026-2050, avec la perspective de budgets de carbone et de budgets de matières très limités.



Axe

Droit, cyberjustice et cybersécurité

Sous la responsabilité de :

Vincent Gautrais, Sébastien Gambs et Alexandre Désy



Actualité scientifique

État des lieux

Valeurs - Institutions - Normes

Cet État de la situation constitue la troisième édition de l'exercice. Forcément, loin de nous l'idée de réinventer la roue. Les questions de régulation ne se réinventent pas; les façons de faire s'adaptent, évoluent. En fait, et comme déjà mentionné l'an dernier, il y a une triple perspective qui fait que la régulation se peaufine. L'acte de réguler est d'abord associé à des **valeurs**, notamment politiques et économiques, qui influent sur le contenu des règles. Ensuite, le contrôle des pratiques est délégué à des **institutions** qui disposent de pouvoirs plus ou moins étendus et donc plus ou moins en mesure d'infléchir les comportements. Enfin, la manière de mettre en application ce contrôle dépend de **normes** dont le niveau de juridicité varie selon les hypothèses et les juridictions. Cette triple vision qui correspond en fait aux trois questions de base à savoir QUOI, QUI et COMMENT réguler, est immuable. En revanche, 2025, notamment à la suite d'un changement de paradigme majeur insufflé par la nouvelle présidence états-unienne, a eu un effet direct sur chacune de ces trois composantes : les valeurs ont changé, décomplexant une approche ouvertement favorable aux grandes plateformes. Les institutions de contrôle doivent supporter une pression de ne pas trop encadrer, perpétuant un laisser-faire qui s'applique depuis plusieurs décennies à l'industrie technologique. Quant aux normes, encore plus que l'an dernier, on semble privilégier les normes molles à des textes plus fermes. La toute récente *Stratégie pancanadienne en matière d'intelligence artificielle* (Gouvernement du Canada, 2025) ne dit rien d'autre. Selon elle, les normes techniques sont en mesure de cadrer, à elles seules, les comportements. Certes, de tels outils normatifs sont indispensables; encore faut-il qu'ils soient le fruit de processus transparents et diversifiés mais aussi insérés dans des cadres de régulation plus globaux, incluant des lois. 2025 est donc une année où une évolution assez franche peut être constatée.

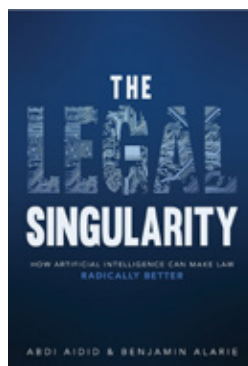
Régulation « innovante »

Cette évolution justement en termes de régulation passe par le besoin de proposer des solutions dites « innovantes ». Derrière ce qualificatif à la mode, surgissent de nombreux procédés ou termes qui tentent de mieux encadrer les technologies. Aussi, on voit surgir une multitude d'expressions censées mieux fonctionner : devoir de vigilance, résilience juridique, stratégie, gouvernance, autant de processus qui sont aussi indéfinis que difficiles à appréhender.

À titre d'exemple, l'idée des bacs à sable réglementaires est souvent mobilisée. Certes. Après, si l'idée est intéressante, peu de preuves sont disponibles quant au fait que cette manière de faire fonctionne, ni quant au fait que les anciennes sont à éviter. À bien des égards, le monde de la régulation fait preuve d'une grande adolescence où l'on essaye, on teste, sans trop savoir ce qui fonctionne ou pas. Le droit, en fait, constitue un certain appel au calme. Il constitue un regard en arrière, la démarche prenant appui sur une loi, même ancienne, ou une jurisprudence. Conservateur par essence, ce regard précautionneux semble utile justement dans le contexte mouvant qu'est le nôtre.

État de ce qui a été fait dans l'axe Droit, cyberjustice et cybersécurité

Du point de vue de l'actualité scientifique de l'axe, les cercles de lecture organisés en 2025 constituent un observatoire privilégié des travaux récents sur l'IA, le droit et la gouvernance, en donnant à voir comment les enjeux théoriques se déclinent dans des ouvrages marquants et dans les controverses les plus actuelles.



Cercle de lecture – The Legal Singularity: How Artificial Intelligence Can Make Law Radically Better

Depuis près de deux ans, notre axe met de l'avant des lectures marquantes ou diversifiées, et ce, dans le cadre d'un « cercle de lecture » où les auteurs et autrices sont confrontés à un regard externe d'une tierce partie. Une première activité tenue en 2025 a

porté sur l'ouvrage *The Legal Singularity: How Artificial Intelligence Can Make Law Radically Better*, présenté par ses auteurs, les professeurs Abdi Aidid et Benjamin Alarie (2023), et discuté par le professeur Patrick Garon-Sayegh. Dans leur ouvrage, ils défendent l'idée d'une « singularité juridique », un état futur où l'incertitude juridique serait fortement réduite grâce aux capacités prédictives de l'IA. Ils distinguent en particulier plusieurs sources d'incertitude (ambiguïté des normes, retard du droit sur les transformations sociales, limites cognitives des humains) et soutiennent que l'apprentissage machine et l'IA générative sont maintenant capables d'exploiter de manière efficace les grandes masses de données juridiques, permettant ainsi d'améliorer la prévisibilité des décisions et de relever le « plancher » de la qualité du raisonnement juridique. Dans cette perspective, l'IA doit permettre une forme d'« alphabétisation juridique universelle », en rendant le droit plus accessible sans en aplatir la complexité.

Le répondant a salué la clarté et la puissance de cette thèse, mais en critique le biais technocratique, qui tend à réduire le droit à de l'information et de la prédiction, au détriment de sa dimension interprétative, délibérative et normative. Il rappelle en particulier que la valeur de l'État de droit ne se limite pas à la prévisibilité, mais tient aussi aux débats sur ce que le droit devrait être. Le dialogue met ainsi en tension un imaginaire d'optimisation computationnelle du droit et une conception plus humaniste, insistante sur l'irréductible dimension sociale de la pratique juridique.



Cercle de lecture – Les algorithmes contre la société

Ce cercle de lecture portait sur l'ouvrage *Les algorithmes contre la société* d'Hubert Guillaud (2025) et avait comme répondant Benjamin Lehaire (TELUQ). Guillaud y développe une critique militante des dispositifs algorithmiques qui structurent désormais la vie sociale, en montrant

qu'ils servent avant tout les intérêts des institutions qui les déploient plutôt que ceux des citoyens. Il met en lumière l'opacité et les effets de normalisation de ces systèmes, qui tendent à marginaliser les groupes déjà vulnérables, comme l'illustrent Parcours Sup, les logiciels de recrutement ou les scores de risque de la CAF. Ces exemples révèlent comment des calculs présentés comme neutres produisent des exclusions systématiques et des accusations de fraude injustes. Guillaud souligne que le débat se focalise excessivement sur la protection des données, au détriment d'une interrogation sur les choix de ce qu'on cherche à calculer et sur le pouvoir qu'ils cristallisent. Il plaide pour une diversification des méthodes, une implication démocratique des usagers et une mise en débat des finalités des systèmes. Lehaire, pour sa part, resitue cette critique dans le contexte du néolibéralisme et d'une régulation numérique centrée sur la gestion individualisée des risques plutôt que sur des normes collectives. Il décrit une « technonormativité » où les systèmes numériques imposent des comportements plus efficacement que le droit, tout en isolant les individus face aux décisions automatisées. Finalement, les deux intervenants convergent vers la nécessité de renforcer les droits collectifs, la transparence des algorithmes, la possibilité d'interdictions juridiques ciblant les usages les plus nocifs et le développement de plateformes numériques publiques et non lucratives.



Cercle de lecture – Technopolice : la surveillance policière à l'ère de l'intelligence artificielle

Ce cercle de lecture de l'axe était consacré à l'ouvrage de Félix Tréguier (2024), *Technopolice : la surveillance policière à l'ère de l'intelligence artificielle* dans lequel Tréguier analyse la banalisation, dans les pratiques policières, de technologies comme les drones, la vidéosurveillance

algorithmique, les logiciels prédictifs et la reconnaissance faciale. L'ouvrage soutient que ces dispositifs ne réduisent pas véritablement la criminalité, mais contribuent à amplifier la violence d'État et à placer la ville sous un contrôle sécuritaire diffus. En retraçant les liens entre industrie de la sécurité, ministère de l'Intérieur, autorités de régulation et pratiques de terrain, il montre comment un technosolutionnisme hégémonique accompagne une dérive autoritaire. La conférence revient aussi sur les luttes récentes contre la surveillance policière et les savoirs militants qui ont nourri le livre. Le répondant, Benoît Dupont, spécialiste de cybersécurité et titulaire de la Chaire de recherche du Canada en cyber-résilience, apporte un regard ancré dans la criminologie et la sécurité numérique. Ensemble, ils discutent des tensions entre sécurité, libertés publiques, gouvernance algorithmique et transformation numérique de l'État. Il en ressort en particulier un décalage entre la puissance d'outillage de l'IA au service de la sécurité et la faiblesse des garde-fous juridiques, souvent limités à des exigences de proportionnalité ou de transparence a minima. Il y a, en particulier, une urgence à inventer des contre-pouvoirs effectifs et, le cas échéant, assumer l'interdiction de certaines technologies de surveillance.



Débats scientifiques

Le droit nous échappe et nous constatons une multiplication de situations de « non-droit » qui peuvent être problématiques. Soit volontairement, des industries profitent de la mutation sociétale pour mettre de l'avant leurs propres intérêts, soit simplement, de nombreux acteurs ne savent tout simplement pas comment appréhender l'importance de ces mutations. À titre d'exemple, de plus en plus de lois exigent qu'il faille rédiger des politiques internes pour montrer sa diligence dans le traitement des données, mais on ne sait ni comment faire ni à qui demander un tel travail. Même chose avec les normes techniques qui sont souvent vues comme la panacée à l'encadrement des technologies. Mais lesquelles ? Il y a en tellement. Et même les plus connues ne sont pas toujours le gage de règles forcément adaptées et bénéfiques au bien commun. On est donc véritablement dans une hypothèse où l'on construit l'avion en plein vol. On régule et on réfléchit plus tard.

Actions possibles

De façon concrète, que peut-on faire ? Comme chercheur, qu'est-il possible d'envisager ? Sans prétention d'exhaustivité, voici plusieurs pistes d'action qu'il serait possible d'envisager. D'abord, on peut être tenté de **mesurer** la régulation. Même si cela relève plus de la gestion que du droit, il serait intéressant d'avoir des retours sur les méthodes qui fonctionnent le mieux. Forcément, il est à cet égard intéressant de **comparer** et donc regarder ce qui se passe ailleurs. Le tout en tenant compte des spécificités culturelles propres à chaque juridiction.

À titre d'exemple, le modèle européen, généralement considéré comme le plus exigeant, constitue un « standard » assurément digne d'intérêt. Le modèle étatsunien se situe ailleurs et préconise un laisser-faire plus assumé. Entre ces possibles, où se situe-t-on ? Serait-on plus enclin à considérer les façons de faire britanniques qui ne sont pas dénuées d'une certaine structuration institutionnelle ? Une autre action possible se trouve dans l'élaboration de modèles pratiques. Plutôt que des cadres éthiques, certes utiles, il importe d'**élaborer** des exemples de cadres pratiques que les acteurs pourraient s'approprier pour leurs propres fins. Idéalement, il serait intéressant même que de tels documents puissent être rendus disponibles en utilisant des outils logiciels. Enfin, et après avoir préalablement fait le constat d'une certaine complexification de la régulation, notamment en utilisant les normes techniques, il importe de les **identifier**. Actuellement, il en existe toute une pluralité. Outre le fait qu'elles soient souvent payantes et rédigées sous le seul contrôle de l'industrie, il importe de s'y intéresser. Actuellement, elles sont trop souvent ignorées par la recherche. On ne peut simplement critiquer cette façon de faire (comment pourrait-on faire autrement) ; il faut aussi **s'impliquer** dans leur rédaction. Au-delà de ces considérations à très haut niveau, comment ces recherches peuvent se matérialiser ? Le paragraphe suivant sur l'audit en offre une illustration.

Même les normes techniques les plus connues ne sont pas toujours le gage de règles forcément adaptées et bénéfiques au bien commun.





Questionnements

La journée de réflexion commune entre les axes Droit, cyberjustice et cybersécurité et Éthique, gouvernance et démocratie de l'Obvia a permis de mettre en lumière l'importance de l'audit de l'IA. Cette dernière demeure un sujet crucial en raison de l'impact de plus en plus important des modèles d'IA dans toutes les sphères de notre société et de la croissance fulgurante de l'IA générative ainsi que de l'IA agentive. **Un des enjeux majeurs est cependant de réussir à circonscrire clairement ce qu'on entend par audit de l'IA, en particulier en ce qui concerne les propriétés qu'on souhaite auditer ou encore de la forme que devrait prendre cet audit.** Le panel qui a eu lieu sur cette question était modéré par Sébastien Gambs (professeur, Université du Québec à Montréal) et composé de Manon Aigoïn (doctorante en criminologie, Université de Montréal), Ulrich Aïvodji (professeur adjoint, École de Technologie Supérieure), Simon Demarkar (professeur agrégé, HEC Montréal) et Dominic Martin (professeur, Université du Québec à Montréal). Les réflexions partagées ci-après se basent sur les discussions ayant lieu durant ce panel.

L'exemple classique où l'audit est extrêmement bien balisé est celui de l'audit des états financiers, domaine où un émetteur assujéti est responsable d'un objet normé (ici les états financiers). Dans ce contexte, les entités auditées sont des émetteurs assujétis qui ont l'obligation d'être auditées pour pouvoir conserver l'accès aux capitaux financiers ou encore être coté en bourse. Elles sont donc désignées comme responsables de l'objet audité. De plus, l'auditeur produit une opinion standardisée encadrée par des normes qui s'étendent parfois sur des milliers de pages telles que les IFRS (*International Financial Reporting Standards*) et les normes canadiennes d'audit, lesquelles définissent le contenu, la structure et les exigences d'exhaustivité, d'exactitude et d'intelligibilité. Le rapport d'audit est lui aussi encadré de manière claire à la fois sur la méthode et la forme de l'opinion (*unmodified* ou *modified* opinion).

Par contraste, au niveau de l'audit de l'IA il n'y a pas encore de consensus clair sur qui devrait être l'entité responsable de l'audit (fournisseurs de modèles, GAFAM, organisations utilisatrices et déployant les modèles ou même les États eux-mêmes) ni ce qui doit constituer l'objet de l'audit (données d'entraînement, processus de création du modèle, infrastructure technique, usages concrets, impacts socio-techniques). En particulier, la littérature sur le sujet est foisonnante mais fragmentée avec des milliers de publications et une multitude de cadres qui ne convergent pas. Autrement dit, **les enjeux fondamentaux pour pouvoir opérationnaliser l'audit de l'IA comportent la définition de la portée de l'audit de l'IA ainsi que le produit attendu (ex : rapport, score, certification, opinion, etc.).**

Une analogie pertinente serait de comparer l'état actuel des travaux sur l'audit de l'IA aux débats sur l'audit des gaz à effet de serre pour illustrer une situation analogue, avec des controverses sur ce qui doit entrer ou non dans le périmètre comme des sujets tels que l'addictivité, la santé, l'exploitation du travail ou encore l'empreinte écologique.

Cette incertitude est d'autant plus problématique dans des contextes de sécurité publique et de justice, où des systèmes d'IA participent à des décisions coercitives en temps réel (arrestations, interventions policières). Ainsi, au-delà de l'impact direct sur la personne dans le cas d'erreur, même si celle-ci est reconnue *a posteriori*, la personne est déjà entrée dans le système de justice ce qui impactera son casier et pourrait accroître le risque de stigmatisation ou les procédures futures. De plus, plus un groupe est ciblé, plus il est représenté dans ces bases, ce qui peut justifier de nouvelles interventions, renforçant ainsi les risques de biais discriminatoires et alimentant les boucles de rétroaction. En particulier, les citoyens ignorent le plus souvent qu'ils sont ciblés via des technologies d'IA, et s'ils sont face à une décision humaine, algorithmique ou hybride. **Dans ce contexte, les principaux enjeux identifiés sont l'opacité, la dépendance aux fournisseurs ainsi que les évaluations privées.** Ainsi, les outils d'IA utilisés en sécurité sont souvent achetés à des sociétés privées, parfois situées hors Québec. De plus, le fait que les rares évaluations de ces outils soient produites par les fournisseurs eux-mêmes plutôt que par les gouvernements ou les usagers publics renforce un problème d'indépendance et de conflit d'intérêts.

En résumé, le premier défi structurant à adresser pour l'audit de l'IA est la spécification de l'audit, autrement dit clairement spécifier quelles normes, quels critères et quel produit est attendu (ex. : rapport, opinion, certification).

Le second défi est de réussir à bien encadrer les exigences procédurales qui seront nécessaires pour un audit efficace de l'IA dont la détermination des droits d'accès aux documents, à la documentation technique, aux données et à l'infrastructure dont ont besoin les auditeurs ainsi que la mise en place des conditions pour avoir une réelle indépendance des auditeurs, alors qu'aujourd'hui l'évaluation est largement internalisée chez les fournisseurs. **Une des pistes de solution pour adresser ces enjeux serait de revenir à une conception circonscrite de l'audit (vérification d'un objet précis) plutôt qu'à un audit complet de « l'IA ».** Par exemple, on pourrait imaginer que l'audit se focalise tout d'abord sur les spécifications techniques publiées par les entreprises (documents de plusieurs centaines de pages). En effet, ces spécifications définissent une grande partie des aspects d'un modèle d'IA (capacités, limitations, garanties) qui doivent pouvoir être vérifiés.

Le mandat pourrait ainsi être donné à des groupes indépendants d'en vérifier l'exactitude et la cohérence avec le comportement observé. Il faudrait aussi prendre en compte la capacité de certains systèmes d'IA à pouvoir détecter lorsqu'ils sont évalués et à modifier leur comportement en conséquence, ce qui compromet la fiabilité des audits fondés sur des scénarios de test standardisés. De plus, dans les secteurs sensibles comme la sécurité ou la justice, on pourrait rejeter l'idée que tous les critères soient équivalents. **Au lieu de cela, on pourrait introduire des critères inaliénables (« red flags ») liés aux droits fondamentaux, à la non-discrimination, à la protection de la vie privée ainsi qu'à l'intégrité physique et psychique.** Ainsi, une forte atteinte à ces critères ne pourrait pas être compensée par des gains d'efficacité opérationnelle.

Une des pistes de solution pour adresser les enjeux de la spécification et de l'encadrement des exigences procédurales de l'audit serait de revenir à une conception circonscrite de celle-ci (vérification d'un objet précis) plutôt qu'à un audit complet de l'IA.

Enfin, une approche prometteuse est celle des audits communautaires où les usagers sont eux-mêmes impliqués dans l'audit du système d'IA. La forme la plus simple de ce type d'audit est le développement d'une théorie populaire (« Folk theorization ») où les usagers développent une compréhension sur le fonctionnement des systèmes via essais/erreurs, le partage de tutoriels ou encore l'observation de recommandations. Ainsi une forme d'audit algorithmique du quotidien est celle où les incidents détectés par des usagers dans l'usage normal de plateformes (ex. recadrage biaisé d'images sur un réseau social), sont par la suite amplifiés par la circulation sur les réseaux sociaux. Un résultat tangible des efforts de systématisation de ces approches est l'émergence de bases d'incidents permettant de répertorier ceux-ci pour ensuite objectiver les préjudices subis par certains groupes. Cette forme d'audit est cependant encore à ses balbutiements avec de nombreuses questions méthodologiques en suspens comme la gestion du taux de signalement qui peut être différent selon les populations ou l'estimation du nombre réel de personnes affectées. Une des originalités de cette forme d'audit apparaît comme une coproduction entre systèmes, usagers, institutions et communautés, et non comme un exercice purement technico-bureaucratique.

Conclusion

Les travaux des axes mettent en évidence que l'intelligence artificielle s'inscrit dans une transformation profonde et transversale des sociétés contemporaines, touchant simultanément les dimensions éthiques, sociales, économiques, culturelles, environnementales et éducatives. Dans l'axe Éthique, gouvernance et démocratie, le débat est marqué par une tension entre les promesses d'innovation et les risques liés à l'autonomie croissante des systèmes, appelant à une gouvernance préventive et à des mécanismes robustes de supervision et de reddition de comptes. Dans le domaine de la santé, l'IA redéfinit la relation de soins en améliorant la précision et l'efficacité, tout en soulevant des enjeux de confiance, de responsabilité, de biais et de déshumanisation des pratiques. Du côté du travail, l'intégration de l'IA et de la gestion algorithmique transforme les organisations, reconfigure les compétences et redessine les rapports de pouvoir, avec des effets ambivalents sur la qualité du travail et les trajectoires professionnelles. Dans les secteurs culturels et médiatiques, les enjeux portent notamment sur la transparence, la rémunération équitable, la souveraineté numérique et les dynamiques d'amplification du discours technologique. Par ailleurs, l'axe Droit, cyberjustice et cybersécurité met en lumière une régulation encore fragmentée, marquée par des zones de non-droit et par l'absence de consensus sur les mécanismes d'audit de l'IA. En éducation, l'IA introduit des tensions entre personnalisation des apprentissages, développement des compétences et transformation du rôle enseignant, tout en soulevant des enjeux d'équité et de gouvernance. Enfin, l'axe Sobriété numérique et transition socio-écologique examine les solutions numériques et d'IA pour l'environnement et souligne l'importance de passer d'une logique de gestion de l'environnement à une logique de soutien à la transition écologique.

Malgré la diversité de ces contextes, plusieurs convergences fortes émergent, en particulier autour du rôle structurant de l'IA agentive. L'augmentation du degré d'autonomie des systèmes accentue le déplacement des frontières entre assistance et délégation décisionnelle, amplifiant les risques de perte de contrôle humain, de biais et de concentration du pouvoir. Cette évolution s'accompagne d'un décalage persistant entre les discours promouvant une IA centrée sur l'humain et les réalités empiriques observées dans les milieux de pratique. Elle met également en évidence des enjeux transversaux liés aux compétences, à la confiance, à l'équité et à la participation des acteurs, ainsi qu'à la reconfiguration des rapports de pouvoir entre concepteurs, institutions et usagers.

Face à ces constats, plusieurs pistes de solution sont proposées. Elles s'orientent vers le renforcement des cadres de gouvernance, notamment par l'adoption d'approches préventives intégrant la supervision humaine, la transparence, l'audit et la traçabilité tout au long du cycle de vie des systèmes. La participation des acteurs concernés, incluant les travailleurs, les professionnels et les citoyens, apparaît également comme une condition essentielle à une adoption démocratique et légitime de l'IA. Sur le plan opérationnel, des approches concrètes sont mises de l'avant, telles que le développement d'audits ciblés, l'établissement de critères non négociables liés aux droits fondamentaux et la mise en place de dispositifs de formation visant à développer des compétences élargies intégrant les dimensions sociales et éthiques des technologies. Plus largement, les plus récentes études insistent sur la nécessité de recentrer les finalités de l'IA autour du bien commun, de l'équité et de la capacitation des personnes.

Une fois de plus, l'État de la situation 2026 ouvre des pistes de recherche importantes. Il apparaît essentiel de poursuivre les travaux empiriques afin de mieux comprendre les usages réels de l'IA et leurs effets différenciés selon les contextes organisationnels, sociaux et institutionnels. Le développement de cadres analytiques communs et l'intégration des dimensions techniques, sociales, environnementales et politiques constituent des priorités pour structurer la recherche future.

Dans cette perspective, l'État de la situation constitue un apport structurant pour la recherche, en permettant d'identifier les angles morts, de consolider les connaissances et d'éclairer de manière rigoureuse les choix publics et organisationnels. Il mobilise une diversité d'intervenants ainsi que les efforts et la réflexion scientifiques nécessaires à l'élaboration de cadres analytiques robustes pour orienter le développement de l'intelligence artificielle vers des trajectoires responsables, équitables et durables.

Bibliographie

Axe Éthique, gouvernance et démocratie

- Abecassis, A., Barry, J., Bello, I., Bengio, Y., Bergeaud, A., Bonnet, Y., Hacker, P., Harack, B., Hatz, S., Henkel, J., Hoos, H. H., Kitamura, K., Lall, R., Lechelle, Y., de Leusse, C., Martinet, C., Mialhe, N., Morse, J. C., Negele, M., Park, K. R., Pluckebaum, M., Papa-Fabre, M., Prud'homme, B., Ralle, Y., Robinson, M., Segerie, C.-R., Torreblanca, J.-I., Velasco, L. et VijayRaghavan, K.A. (2025). *Blueprint for Multinational Advanced AI Development*. AI Governance Initiative (Oxford Martin School, University of Oxford). <https://aigi.ox.ac.uk/publications/a-blueprint-for-multinational-advanced-ai-development/#>
- Adabara, I., Sadiq, B. O., Shuaibu, A. N., Danjuma, Y. I. et Maninti, V. (2025). Trustworthy agentic AI systems: a cross-layer review of architectures, threat models, and governance strategies for real-world deployment [version 1; peer review: awaiting peer review]. *F1000Research*, 14:905 <https://doi.org/10.12688/f1000research.169927.1>
- Alva, L. et Pandey, B. (2026). Agentic AI systems in the age of generative models: architectures, cloud scalability, and real-world applications. *Artificial Intelligence Review*, 59 (88). <https://doi.org/10.1007/s10462-025-11458-6>
- Bengio, Y., Cohen, M., Fornasiere, D., Ghosn, J., Greiner, P., MacDermott, M., et al. (2025). *Superintelligent agents pose catastrophic risks: Can scientist AI offer a safer path?* arXiv. <https://arxiv.org/abs/2502.15657>
- Bengio, Y. (2023). AI and Catastrophic Risk. *Journal of Democracy*, 34(4), 111–121. <https://doi.org/10.1353/jod.2023.a907692>
- Boine, C. et al. (2024). *IA et démocratie : Comprendre les effets de l'IA sur les élections*. CEIMIA et IVADO. https://ceimia.org/wp-content/uploads/2025/01/ivadoceimia_iaetdemocratie_final.pdf
- Bouhouita-Guermech, S. et Haidar, H. (2024). Scoping review shows the dynamics and complexities inherent to the notion of “responsibility” in artificial intelligence within the healthcare context. *Asian Bioethics Review*, 16(3), 315–344.
- Bowen, G. (2025). Agentic artificial intelligence: Legal and ethical challenges of autonomous systems. *Journal of Digital Technologies and Law*, 3(3), 431–445.
- Collin, S. et Marceau, E. (2023). Enjeux éthiques et critiques de l'intelligence artificielle en enseignement supérieur. *Éthique publique*, 24(2).
- Currie, W. L., Leimeister, J. M., Schlagwein, D. et Willcocks, L. (2025). Rethinking technology regulation in the age of AI risks. *Journal of Information Technology*, 40(3), 236–245.
- Deng, Z., Guo, Y., Han, C., Ma, W., Xiong, J., Wen, S. et Xiang, Y. (2025). AI agents under threat: A survey of key security challenges and future pathways. *ACM Computing Surveys*, 57(7), 1–36.
- Eslami, A. et Yu, J. (2025). *Security risks of agentic vehicles: A systematic analysis of cognitive and cross-layer threats*. arXiv. <https://arxiv.org/abs/2512.17041>
- Fornasiere, D., Richardson, O., Gendron, G., Serban, J. et Bengio, Y. (2026). *The scientist AI: Safe by design, by not desiring*. LawZero.
- Gahnberg, C. (2021). What rules? Framing the governance of artificial agency. *Policy and Society*, 40(2), 194–210. <https://doi.org/10.1080/14494035.2021.1929729>
- Gerlich, M. (2025). From offloading to engagement: An experimental study on structured prompting and critical reasoning with generative AI. *Data*, 10(11), Article 172. <https://doi.org/10.3390/data10110172>
- Grzybowski, A., Zhao, K. et Jin, K. (2026). Agentic artificial intelligence in ophthalmology: Toward autonomous, adaptive, and ethical eye care. *Acta Ophthalmologica*. Advance online publication. <https://doi.org/10.1111/aos.70099>
- Hammond, L., Chan, A., Clifton, J., Hoelscher-Obermaier, J., Khan, A., McLean, E. et Rahwan, I. (2025). *Multi-agent risks from advanced AI*. arXiv. <https://arxiv.org/abs/2502.14143>
- Hagendorff, T. (2024). Mapping the ethics of generative AI: A comprehensive scoping review. *Minds & Machines*, 34, Article 39. <https://doi.org/10.1007/s11023-024-09694-w>
- Hasan, S. S., Woo, J. J., Cote, M. P. et Ramkumar, P. N. (2025). Generative versus nongenerative artificial intelligence. *Arthroscopy*, 41(3), 545–546. <https://doi.org/10.1016/j.arthro.2024.12.001>
- Hosseini, M., Murad, M. et Resnik, D. B. (2026). Benefits and risks of using AI agents in research. *Hastings Center Report*, 56(1), 13–17. <https://doi.org/10.1002/hast.70025>
- Khan, R., Joyce, D. et Habiba, M. (2025). *AGENTS SAFE: A unified framework for ethical assurance and governance in agentic AI*. arXiv. <https://arxiv.org/abs/2512.03180>
- Huang, Y., Arora, C., Houg, W. C., Kani, T., Madulgalla, A. et Grundy, J. (2025). *Ethical concerns of generative AI and mitigation strategies: A systematic mapping study*. arXiv. <https://arxiv.org/abs/2502.00015>
- Marchal, N., Xu, R., Elasmr, R., Gabriel, I., Goldberg, B. et Isaac, W. (2024). *Generative AI misuse: A taxonomy of tactics and insights from real-world data*. arXiv. <https://arxiv.org/abs/2406.13843>
- Mekulu, K., Aqlan, F. et Yang, H. (2025). Agentic artificial intelligence in cognitive screening: A translational roadmap for dementia care. *Journal of Alzheimer's Disease Reports*, 9, Article 25424823251407989. <https://doi.org/10.1177/25424823251407989>

Obvia. (2025, Octobre 24). *Avancées en IA : Risques politiques, sécuritaires et économiques mondiaux* [Vidéo]. YouTube. https://www.youtube.com/watch?v=rh_eviElzAo

Patel, K., Shah, M., Qureshi, K. M., et Qureshi, M. R. N. (2025). A systematic review of generative AI: importance of industry and startup-centered perspectives, agentic AI, ethical considerations & challenges, and future directions. *Artificial Intelligence Review*, 59(1), 7. <https://doi.org/10.1007/s10462-025-11435-z>

Pawar, A. (2025). Ethical and governance challenges of agentic AI. *International Journal of Humanities and Information Technology*, 7(3), 76–82. <https://ijhit.info/index.php/ijhit/article/view/100>

Radanliev, P., Santos, O. et Ani, U. D. (2025). Generative AI cybersecurity and resilience. *Frontiers in Artificial Intelligence*, 8, Article 1568360. <https://doi.org/10.3389/frai.2025.1568360>

Raheem, T. et Hossain, G. (2025). Agentic AI systems: Opportunities, challenges, and trustworthiness. In *Proceedings of the 2025 IEEE International Conference on Electro Information Technology (eIT)* (pp. 618–624). IEEE.

Raza, S., Sapkota, R., Karkee, M. et Emmanouilidis, C. (2025). *TRISM for agentic AI: A review of trust, risk, and security management in LLM-based agentic multi-agent systems*. arXiv. <https://arxiv.org/abs/2506.04133>

Reuters. (2025, Juin 25). Over 40% of agentic AI projects will be scrapped by 2027, Gartner says. *Reuters*. <https://www.reuters.com/business/over-40-agentic-ai-projects-will-be-scrapped-by-2027-gartner-says-2025-06-25/>

Sapkota, S. (2025). A systematic review of generative AI and agentic AI: Ethical considerations and challenges. *Artificial Intelligence Review*. <https://link.springer.com/article/10.1007/s10462-025-11435-z>

Schneider, J. (2025). *Generative to agentic AI: Survey, conceptualization, and challenges*. arXiv. <https://arxiv.org/abs/2504.18875>

Taeiagh, A. (2025). Governance of generative AI. *Policy and Society*, 44(1), 1–22. <https://doi.org/10.1093/polsoc/puaf001>

Tallam, K. (2025). *Alignment, agency and autonomy in frontier AI: A systems engineering perspective*. arXiv. <https://arxiv.org/abs/2503.05748>

Zohny, H., McMillan, J. et King, M. (2023). Ethics of generative AI. *Journal of Medical Ethics*, 49(2), 79–80. <https://doi.org/10.1136/jme-2023-108909>

Axe santé durable

Associated Press et Ars Technica. (2024, 28 octobre). Hospitals adopt error-prone AI transcription tools despite warnings. *Ars Technica*. <https://arstechnica.com/ai/2024/10/hospitals-adopt-error-prone-ai-transcription-tools-despite-warnings/>

Canadian Institute for Health Information. (2026). *Seniors (65 and older)*. <https://www.cihi.ca/en/indicators/seniors-65-and-older>

Conseil de l'Europe / Council of Europe. (2022). *The impact of artificial intelligence on the doctor-patient relationship*. CDBIO – Comité de bioéthique. <https://rm.coe.int/inf-2022-5-report-impact-of-ai-on-doctor-patient-relations-e/1680a68859>

Courthouse News Service. (2025, 12 février). UnitedHealth faces lawsuit over AI-driven prior authorization denials. <https://www.courthouse-news.com>

Déclaration de Montréal. (2018). La Déclaration de Montréal pour un développement responsable de l'intelligence artificielle. <https://declarationmontreal-iaresponsable.com/la-declaration/>

EarlySense. (2025). *EarlySense: Answers to questions on patient monitoring*. <https://ppahs.org/patient-safety-guide/early-sense-answers-to-questions-on-patient-monitoring/>

EPIC Systems. (2025). *Public health surveillance tools*. <https://open.epic.com/PublicHealth/Surveillance>

FAIR Principles. (2025). *FAIR principles*. GO FAIR. <https://www.go-fair.org/fair-principles/>

Health Canada. (2025, 5 février). *Pre-market guidance for machine learning-enabled medical devices (MLMD)*. Gouvernement du Canada. <https://www.canada.ca/en/health-canada/services/drugs-health-products/medical-devices/application-information/guidance-documents/pre-market-guidance-machine-learning-enabled-medical-devices.html>

Health Canada. (2025, février). *Guidance on machine learning-enabled medical devices (MLMD): Good machine learning practices and predetermined change control plans*. Gouvernement du Canada. <https://www.canada.ca/en/health-canada>

Healthcare Finance News. (2025, 3 mars). Cigna's algorithm under scrutiny in class-action lawsuit. *Healthcare Finance News*. <https://www.healthcarefinancenews.com>

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I. et Maes, P. (2025). Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task. *arXiv*. <https://doi.org/10.48550/arXiv.2506.08872>

Köttl, H., Gallistl, V., Rohner, R. et Ayalon, L. (2021). "But at the age of 85? Forget it!": Internalized ageism, a barrier to technology use. *Journal of Aging Studies*, 59, article 100971. <https://doi.org/10.1016/j.jaging.2021.100971>

La Torre, J. (2026). *The slow death of Epic Systems*. Julio's Substack. <https://open.substack.com/pub/juliolatorremd/p/the-slow-death-of-epic-systems>

Lachance, N. (2025a, 9 septembre). Le dossier santé numérique sous la loupe de vérificateurs, alors que les coûts pourraient atteindre 3 G\$. *Le Journal de Québec*.

Lachance, N. (2025b, 9 octobre). Dossier santé numérique sous la loupe : Québec cherche à mettre les données des citoyens à l'abri des autorités américaines. *Le Journal de Québec*.

Laranjo, L., Tudor Car, L., Payne, R. E., Neves, A. L., Kidd, M. et Miranda, J. J. (2025). Artificial intelligence in primary care: Innovation at a crossroads. *The Lancet Primary Care*, article 100078. <https://doi.org/10.1016/j.lanprc.2025.100078>

- Latraverse, E. (2025, 13 octobre). Le beau cadeau de Québec aux Américains. *Le Journal de Montréal*.
- Le Monde. (2025, 9 septembre). « L'IA, dans le domaine de la santé, produit des résultats justes, mais qui ne correspondent pas aux protocoles en usage dans le milieu ». *Le Monde*.
- Ligue des droits et libertés. (2025). *Mémoire sur le projet de loi n° 82, Loi concernant l'identité numérique nationale et modifiant d'autres dispositions*.
- Loiseau, H., Caron, D. J., Gambs, S. et Brousseau, S. (2025). *Pour une identité numérique québécoise au service des citoyens : enjeux et recommandations*. Obvia. <https://doi.org/10.61737/ZFCY4930>
- Macnamara, B. N., Berber, I., Çavuşoğlu, M. C., Krupinski, E. A., Nallapareddy, N., Nelson, N. E., Smith, P. J., Wilson-Delfosse, A. L. et Ray, S. (2024). Does using artificial intelligence assistance accelerate skill decay and hinder skill development without performers' awareness? *Cognitive Research: Principles and Implications*, 9, article 46. <https://doi.org/10.1186/s41235-024-00572-8>
- Mamdani, M. (2025). Canada's health innovation imperative. *CMAJ : Canadian Medical Association Journal*, 197(26), E761.
- Markelius, A., Wright, C., Kuiper, J., Delille, N. et Kuo, Y.-T. (2024). The mechanisms of AI hype and its planetary and social costs. *AI and Ethics*, 4(3), 727–742.
- MedGPT. (2025). *MedGPT*. <https://chatgpt.com/g/g-jxm5ljjmo-medgpt>
- Micrologic. (2025). *Projet de loi n° 82 sur l'identité numérique : Les données des Québécois au cœur de la souveraineté numérique*.
- Ministère de la Santé et des Services sociaux (MSSS). (2024). *Plan directeur en intelligence artificielle en santé 2024–2027*. Gouvernement du Québec. <https://www.msss.gouv.qc.ca>
- Nabla. (2025). *Nabla AI*. <https://www.nabla.com/fr>
- Ng, J. J. W., Wang, E., Zhou, X., Zhou, K. X., Goh, C. X. L., Sim, G. Z. N., Tan, H. K., Goh, S. S. N. et Ng, Q. X. (2025). Evaluating the performance of artificial intelligence-based speech recognition for clinical documentation: A systematic review. *BMC Medical Informatics and Decision Making*, 25, article 236. <https://doi.org/10.1186/s12911-025-03061-0>
- NFP. (2025, 22 avril). Court allows lawsuit over AI use in benefit denials to proceed. *NFP Insights*. <https://www.nfp.com/insights/court-allows-lawsuit-over-ai-use-in-benefit-denials-to-proceed/>
- Niu, M., Tao, J., Li, Y., Qin, Y. et Li, Y. (2024). WavDepressionNet: Automatic depression level prediction via raw speech signals. *IEEE Transactions on Affective Computing*, 15(1), 283–295. <https://doi.org/10.1109/TAFFC.2023.3272553>
- Obermeyer, Z., Powers, B., Vogeli, C. et Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447–453.
- Obvia. (2026). *Qui sommes-nous ?* Université Laval. <https://www.obvia.ca/a-propos>
- Ordre des infirmières et infirmiers du Québec. (2024). *Recommandations sur l'impact de l'intelligence artificielle sur la protection du public et la profession infirmière*. <https://www.oiiq.org>
- Ordre des psychoéducatrices et psychoéducateurs du Québec. (2025). *L'intelligence artificielle dans la pratique professionnelle en psychoéducation : guide de référence*.
- Ordre des psychologues du Québec. (2025). *Du diagnostic à l'intelligence artificielle : des dossiers stratégiques menés l'automne dernier*. <https://www.ordrepsy.qc.ca>
- Ordre des psychologues du Québec, Gibb, K. et Rezaee, H. (2025a). *Risques de l'IA dans la pratique clinique et stratégies d'atténuation*. <https://www.ordrepsy.qc.ca>
- Rodriguez, S., Hug, C., Todorov, P., Moret, N., Boswell, S. A., Evans, K. et al. (2021). Machine learning identifies candidates for drug repurposing in Alzheimer's disease. *Nature Communications*, 12(1), article 1033.
- Shiwani, T., Relton, S., Evans, R., Kale, A., Heaven, A., Clegg, A. et al. (2023). New horizons in artificial intelligence in the healthcare of older people. *Age and Ageing*, 52(12).
- Stade, E. C., Stirman, S. W., Ungar, L. H., Boland, C. L., Schwartz, H. A., Yaden, D. B., Sedoc, J., DeRubeis, R. J., Willer, R. et Eichstaedt, J. C. (2024). Large language models could change the future of behavioral healthcare: A proposal for responsible development and evaluation. *NPJ Mental Health Research*, 3(1), article 12. <https://doi.org/10.1038/s44184-024-00056-z>
- Stypińska, J. (2023). AI ageism: A critical roadmap for studying age discrimination and exclusion in digitalized societies. *AI and Society*, 38(2), 665–677.
- Stypińska, J. et Franke, A. (2023). AI revolution in healthcare and medicine and the (re-)emergence of inequalities and disadvantages for ageing population. *Frontiers in Sociology*, 7.
- Tian, Y. J. A., Felber, N. A., Pageau, F., Schwab, D. R. et Wangmo, T. (2024). Benefits and barriers associated with the use of smart home health technologies in the care of older persons: A systematic review. *BMC Geriatrics*, 24(1), article 152.
- Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44–56.
- Torous, J. et Topol, E. J. (2025). Assessing generative artificial intelligence for mental health. *The Lancet*, 406(10504), 683. [https://doi.org/10.1016/S0140-6736\(25\)01147-3](https://doi.org/10.1016/S0140-6736(25)01147-3)
- van Kolschooten, H. (2023). The AI cycle of health inequity and digital ageism: Mitigating biases through the EU regulatory framework on medical devices. *Journal of Law and the Biosciences*, 10(2), Isad031.
- Viz.ai. (2025). *Viz.ai*. <https://www.viz.ai>

Woebot Health. (2025). Woebot. <https://woebothealth.com/>

World Health Organization. (2022). *Ageism in artificial intelligence for health: WHO policy brief*. World Health Organization.

World Health Organization. (2024). *Ethics and governance of artificial intelligence for health*. <https://www.who.int/publications/i/item/9789240084759>

Xia, C., Chen, H., Han, J., Liu, F., Zhang, D. et Li, K. (2025). Identifying children with autism spectrum disorder via transformer-based representation learning from dynamic facial cues. *IEEE Transactions on Affective Computing*, 16(1), 83–97. <https://doi.org/10.1109/TAFFC.2024.3412032>

Xu, Z., Liu, G., Zhao, G., Zhang, Z. et Wang, C. (2025). A topic-guided self-attention network for daily mental wellbeing prediction using mobile devices. *IEEE Transactions on Affective Computing*, 16(2), 783–796. <https://doi.org/10.1109/TAFFC.2024.3471654>

Zhang, L. (2025). The right way for Canada to secure cloud sovereignty. *Policy Options / Options politiques*.

Axe Industrie 4.0, travail et emploi

Baril, É. et Vultur, M. (2025). Navigating Streets, Restaurants, and Algorithms: A Study of Young Immigrant Food Delivery Couriers in Montréal and Toronto. *Labour*, 95, 121-144.

Bowkett, C. et Lévesque, C. (2026). Shifting terrains of expertise: Power, knowledge and the era of Industry 4.0. *Organization*, 13505084251366177.

Coiquaud, U. (2024). L'intelligence artificielle: un nouveau sujet pour la réglementation du travail au Québec? Dans X. Parent-Rochelleau et A. Ollier-Malaterre (dir.), *Le management à l'ère numérique. Nouvelles pratiques, réalités et régulations* (p. 205-217). Presses de l'Université du Québec.

Datafranca. (2025). *Les 101 mots de l'intelligence artificielle. Guide du vocabulaire essentiel de la science des données et de l'intelligence artificielle* (3^{ee} éd.). Datafranca.org.

Garneau, J. M.-É. (2021). La numérisation de l'économie et les mondes du travail: principales avancées technologiques et leurs enjeux. Dans J. Bernier (dir.), *L'intelligence artificielle et les mondes du travail. Perspectives sociojuridiques et enjeux éthiques* (p. 21-45). Presses de l'Université Laval.

Institut de la statistique du Québec (ISQ). (2026). *Bilan du marché du travail au Québec en 2025*. <https://statistique.quebec.ca/fr/communiquer/bilan-marche-travail-quebec-2025>

Leonardi, P. M. (2025, 2025/09/01/). Homo agenticus in the age of agentic AI: Agency loops, power displacement, and the circulation of responsibility. *Information and Organization*, 35(3), 100582. <https://doi.org/https://doi.org/10.1016/j.infoandorg.2025.100582>

Lévesque, C., Garneau, J. M. É., Dupuis, M., Parent-Rochelleau, X., Pasquier, V., Langlois, L. et Saba, T. (2025). *Consultation sur la transformation des milieux de travail par le numérique - Avis des experts de l'Obvia*.

Parent-Rochelleau, X. et Parker, S. K. (2021). Algorithms as work designers: How algorithmic management influences the design of jobs. *Human Resource Management Review*. <https://doi.org/10.1016/j.hrmmr.2021.100838>

Saba, T. et Langlois, L. (2025). *Government Consultation on National Priorities for Artificial Intelligence*.

Saba, T., Registre, J. F. R. et Ouellet, A.-M. (2025). Au-delà des compétences: l'identité professionnelle des CRHA et CRIA pour orienter les transformations induites par l'IA. Dans Éliane Racine, Voahirana Raharison et Anaïs Hébrard (Dir.), *De la théorie à l'action*, (p.171-179). Éditions JFD.

Soumi, M. et Bitan, M. (2026). *AI-Empowered Industrial Relations* (vol. 01254). Springer Nature Switzerland. <https://research.ebsco.com/linkprocessor/plink?id=3d0b92ec-bf0f-3edd-a5f5-24f5728adb60>

Teyssier-Roberge, G., de Hemptinne, D., Gagnon, J. et Tremblay, S. (2025). Repenser les compétences à l'ère de l'intelligence numérique. *Ad machina: l'avenir de l'humain au travail*, 9(2), 272-289. <https://doi.org/https://doi.org/10.1522/radm.no.9.2041>

Teyssier-Roberge, G., Gagnon, J., Tremblay, S. et Hodgetts, H. (2025). A Quantitative Analysis of 21st Century Skills: A Case of Semantic and Psychometric Overlap. *International Journal of Selection and Assessment*, 33(4), e70030.

Trottier, M. (2025a). Les conseillères et conseillers d'orientation devant l'intelligence artificielle : utilisation et enjeux? . *L'orientation*, 15(1), 28-33. <https://www.orientation.qc.ca/medias/iw/Magazine-OCQOQ-Printemps2025.pdf>

Trottier, M. (2025b). *Les technologues professionnels.les de l'OTPQ et l'intelligence artificielle*.

Widad, J. (2025). The Evolution of Autonomy: From Reactive AI to Agentic AI. Dans C. K. K. Reddy, S. Joseph, H. Joshi, S. Doss et M. Ouaisa (dir.), *The Power of Agentic AI: Redefining Human Life and Decision-Making: In Industry 6.0* (p. 1-9). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-89424-4_1

Axe Arts, médias et diversité culturelle

Agence France-Presse (AFP). (2025a, 16 janvier). L'AFP et Mistral AI annoncent un partenariat mondial [Site de média public]. *AFP - L'Agence*. <https://www.afp.com/fr/lagence/notre-actualite/communiqués-de-presse/lafp-et-mistral-ai-annoncent-un-partenariat-mondial>

Agence France-Presse (AFP). (2025b, 25 août). Intelligence artificielle : Perplexity propose de partager les revenus avec les médias. *La Presse*, [en ligne].

Ananny, M., et Karr, J. (2025). How media unions stabilize technological hype : Tracing Organized Journalism's Discursive Constructions of Generative Artificial Intelligence. *Digital Journalism*, 0(0), 1-21. <https://doi.org/10.1080/21670811.2025.2454516>

- APNQL. *Territoire numérique des Premières Nations Québec-Labrador : Posture sur l'éthique du numérique et de l'intelligence artificielle*. (2025, June 17). <https://cssspnql.com/produit/territoire-numerique-des-premieres-nations-quebec-labrador-posture-sur-lethique-du-numerique-et-de-lintelligence-artificielle/>
- Bagaoui-Fradette, I., Dieye, R. et Trudel, P. (2024). Souveraineté culturelle et géants numériques : regards croisés. Canada, Québec, France, Belgique (Rapport de synthèse). *Lex Electronica*, 29(4), 6–21. <https://doi.org/10.7202/1115123ar>
- Bartleman, M., Schapals, A. K. et Dubois, E. (2026). Generative AI and the New Landscape of Automated Journalism: A Systematized Review of 185 Studies (2012–2024). *Journalism et Media* 7(1), 39; <https://doi.org/10.3390/journalmedia7010039>
- Bensamoun, A. (2025). *Rapport de mission sur la rémunération des contenus culturels utilisés par les systèmes d'IA - Volet juridique* (CSPLA, Ed.). Ministère de la culture.
- Buick, A. (2024). Copyright and AI training data—transparency to the rescue? *Journal of Intellectual Property Law & Practice*, 20(3), 182–192. <https://doi.org/10.1093/jiplp/jpae102>
- Boily, A. (2025). IA: plus de 20% du flux YouTube est désormais constitué de contenus dégénératifs, selon un rapport [Review of IA: plus de 20% du flux YouTube est désormais constitué de contenus dégénératifs, selon un rapport]. *Le Journal de Montréal*. <https://www.journaldemontreal.com/2025/12/30/ia-plus-de-20-du-flux-youtube-est-desormais-constitue-de-contenus-degeneratifs-selon-un-rapport>
- Cardillo, A. (2025, October 27). *How Many People Use AI? (Latest 2025 Data)*. Exploding Topics. <https://explodingtopics.com/blog/ai-usage-statistics?>
- Chaire UNESCO sur la diversité des expressions culturelles. (2025). *Agir maintenant pour renforcer notre souveraineté culturelle face aux enjeux liés aux technologies numériques et aux systèmes d'intelligence artificielle*. https://www.unescodec.chaire.ulaval.ca/sites/unescodec.chaire.ulaval.ca/files/souverainete_culturelle_-_actions_prioritaires_et_annexe_fr_0.pdf
- Fletcher, R. (2024). *How many news websites block AI crawlers?* (Reuters Institute Factsheets). Reuters Institute for the Study of Journalism. DOI: 10.60625/risj-xm9g-ws87
- Fratini, S., Hine, E., Novelli, C., Roberts, H. et Floridi, L. (2024). Digital Sovereignty: A Descriptive Analysis and a Critical Evaluation of Existing Models. *Deleted Journal*, 3(3). <https://doi.org/10.1007/s44206-024-00146-7>
- JEANDOT, G. (2025, October 28). *50 % des contenus sont générés par l'IA: comment exister ?* Tous Les Jeudis. <https://tous-les-jeudis.fr/50-des-contenus-web-sont-generes-par-lia-comment-continuer-a-exister/>
- Kuai, J. (2025). Navigating the AI Hype : Chinese Journalists' Algorithmic Imaginaries and Role Perceptions in Reporting Emerging Technologies. *Digital Journalism*, 0(0), 1–20. <https://doi.org/10.1080/21670811.2025.2502851>
- Inaugural Adobe Creators' Toolkit Report: 86 Percent of Global Creators Use Creative Generative AI, See it Boosting Creator Economy*. (2025). Adobe.com. <https://news.adobe.com/news/2025/10/adobe-max-2025-creators-survey?>
- La Presse. (2025, 24 novembre). *La Presse poursuit OpenAI pour violation de droit d'auteur [Média]*. info.lapresse.ca. <https://info.lapresse.ca/nouvelles/la-presse-poursuit-openai-pour-violation-de-droit-dauteur>
- Lee, A. (2025, May 14). *In Graphic Detail: How creators are using generative AI to shape video and design*. Digiday. <https://digiday.com/media/in-graphic-detail-how-creators-are-using-generative-ai-to-shape-video-and-design/>
- Lewis, S. C., Zamith, R., et Bunquin, J. B. A. (2025). Technological Hype and AI in Journalism : Five Functions and Why They Matter. *Digital Journalism*, 1–12. <https://doi.org/10.1080/21670811.2025.2557994>
- Liu, E., Luo, E., Shan, S., Voelker, G. M., Zhao, B. Y., et Savage, S. (2025). Somesite I Used To Crawl : Awareness, Agency and Efficacy in Protecting Content Creators From AI Crawlers. *Proceedings of the 2025 ACM Internet Measurement Conference, IMC '25*, 78–99. <https://doi.org/10.1145/3730567.3732913>
- Kulesz, O., Allen, J.P., Basanta, A., Bensamoun, A., Brin, C., Jacquier, Y., Rankin, V., Myles, B. et Guèvremont, V. (2025). *Dialogues interdisciplinaires : repenser la culture à l'ère de l'IA*. Obvia. <https://doi.org/10.61737/ZFNV7842>
- Magalhães, J. C. et Smit, R. (2025). Less Hype, More Drama : Open-Ended Technological Inevitability in Journalistic Discourses About AI in the US, The Netherlands, and Brazil. *Digital Journalism*, 1–18. <https://doi.org/10.1080/21670811.2025.2522281>
- Masseron, P. (2025). *Contenus protégés par le droit d'auteur et intelligence artificielle : Recherche désespérément stratégie(s) de valorisation*. I2D - Information, données et documents, 2(2), 42–52. <https://doi.org/10.3917/i2d.252.0042>
- Montgomery, S. et Wagner, B. (2025). « Chapitre 18 Digital cultural sovereignty: navigating the digital landscape of European CulturalHeritage Institutions with a decolonial lens », *Research Handbook on Human Rights and Digital Technology*. Edward Elgar Publishing.
- Mouriquand, D. (2025, September 15). *French streaming platform Deezer reveals alarming AI figures*. Euronews; euronews.com. <https://www.euronews.com/culture/2025/09/15/french-streamer-deezer-reveals-that-28-per-cent-of-music-uploaded-to-platform-is-ai-genera>
- Nkoala, S., Matsilele, T., Ndlovu, M. et Bosch, T. (2025). AI Hype Through an African Lens : A Critical Analysis of Language as Symbolic Action in Online News Publications. *Digital Journalism*, 1–20. <https://doi.org/10.1080/21670811.2025.2528052>
- Obvia et IVADO. (2024). *Dialogues interdisciplinaires: Repenser la culture à l'ère de l'IA*. [vidéo]. Obvia. <https://www.obvia.ca/ressources/dialogues-interdisciplinaires-repenser-la-culture-a-lere-de-lia>
- Obvia et IVADO. (2026). *Dialogues interdisciplinaires: L'IA et la diversité linguistique* [vidéo]. Obvia. <https://www.obvia.ca/ressources/dialogues-interdisciplinaires-lia-et-la-diversite-linguistique>
- Odell, J. (2025). Training on Headlines : The New York Times, OpenAI, and the Copyright Implications of AI Data Usage. *Oregon Law Review*, 104(1), 203–240.

- Peterson, L. (2025, September 25). *Spotify renforce la protection des artistes et des producteurs face à l'IA – Spotify*. Spotify. <https://newsroom.spotify.com/2025-09-25/spotify-renforce-la-protection-des-artistes-et-des-producteurs-face-a-ia/>
- Sapkota, R., Raza, S. et Karkee, M. (2025). Comprehensive Analysis of Transparency and Accessibility of ChatGPT, DeepSeek, And other SoTA Large Language Models [Review of *Comprehensive Analysis of Transparency and Accessibility of ChatGPT, DeepSeek, And other SoTA Large Language Models*]. ArXiv. <https://arxiv.org/abs/2502.18505>
- Schaetz, N., et Schjøtt, A. (2025). AI Hype and its Function : An Ethnographic Study of the Local News AI Initiative of the Associated Press. *Digital Journalism*, 0(0), 1-18. <https://doi.org/10.1080/21670811.2024.2443163>
- Spyridou, P. (Lia), et Ioannou, M. (2025). Exploring AI Amid the Hype : A Critical Reflection Around the Applications and Implications of AI in Journalism. *Societies*, 15(2). <https://doi.org/10.3390/soc15020023>
- Stassen, M. (2025, June 20). *Deezer rolls out AI tagging system to fight streaming fraud; says up to 70% of streams from fully AI-*. Music Business Worldwide. <https://www.musicbusinessworldwide.com/deezer-rolls-out-ai-tagging-system-to-fight-streaming-fraud-says-up-to-70-of-streams-from-fully-ai-generated-tracks-are-fraudulent/>
- Tobitt, C. (2026, 20 février). Who's suing AI and who's signing: NYT and Chicago Tribune sue Perplexity, Meta signs licensing deals. *Press Gazette*. <https://pressgazette.co.uk/platforms/publisher-ai-deals-lawsuits-openai-google/>
- UNESCO. (2025). *Report of the Independent Expert Group on Artificial Intelligence and Culture* (UNESCO, Ed.) [Review of Report of the Independent Expert Group on Artificial Intelligence and Culture]. UNESCO.
- UNESCO. (2025b). *Rapport et recommandations du Groupe de réflexion sur la diversité des expressions culturelles dans l'environnement numérique*, (UNESCO, Ed.) <https://unesdoc.unesco.org/ark:/48223/pf0000393574?posInSet=2&queryId=1d210633-5fc3-4754-97af-94708e29a7b1>.
- UNESCO. (2026). *Re/shaping policies for creativity: we share, we act, we build*. <https://doi.org/10.58337/dqda4649>
- Usher, N. (2025). Generative AI and Journalism : Hype, The Always Already New, and Directions for Scholarly Imagination. *Digital Journalism*, 1-10. <https://doi.org/10.1080/21670811.2025.2490604>
- Vicsek, L., Feher, K., et Nguyen, D. (2026). Disappointment or hype? Framing artificial intelligence in the US elite news media. *Global Media and Communication*, 17427665251411490. <https://doi.org/10.1177/17427665251411490>
- Vrabič Dežman, D. (2024). Promising the future, encoding the past : AI hype and public media imagery. *AI and Ethics*, 4(3), 743-756. <https://doi.org/10.1007/s43681-024-00474-x>
- Wittenberg, C., Epstein, Z., Berinsky, A. J. et Rand, D. G. (2024). Labeling AI-Generated Content: Promises, Perils, and Future Directions. *An MIT Exploration of Generative AI*. <https://doi.org/10.21428/e4baedd9.0319e3a6>
- Xanthaki, A. (2025). *Rapport de la Rapporteuse spéciale dans le domaine des droits culturels, Intelligence artificielle, et créativité*, DOC NU A/80/278.
- Zhang, X., Huang, W., et Zhu, J. J. H. (2026). How journalism researchers navigate the AI hype : Research orientations and intervention recommendations. *Information, Communication et Society*, 29(1), 1-20. <https://doi.org/10.1080/1369118X.2025.2482666>

Axe Éducation et capacitation

- Baig, M. I. et Yadegaridehkordi, E. (2024). ChatGPT in higher education: A systematic literature review and research challenges. *Computers & Education*. <https://www.sciencedirect.com/science/article/pii/S0883035524000971>
- Bussièrès McNicoll, F., De Rosa, N. et Yates, J. (2026). *Sondage sur l'intelligence artificielle : université, cégep, étudiants*. ICI Radio-Canada. <https://ici.radio-canada.ca/recit-numerique/15239/sondage-intelligence-artificielle-ia-universite-cegep-etudiants>
- Chu, Z., Wang, S., Xie, J., Zhu, T., Yan, Y., Ye, J., Zhong, A., Hu, X., Liang, J., Yu, P. S., et Wen, Q. (2025). *LLM agents for education: Advances and applications*. arXiv. <https://doi.org/10.48550/arXiv.2503.11733>
- Commission européenne. (2024). *Ethical guidelines on the use of artificial intelligence and data in teaching and learning for educators*. European Commission. <https://education.ec.europa.eu/focus-topics/digital-education/action-plan/ethical-guidelines-for-educators-on-using-ai>
- Córdova-Esparza, D. M. (2025). AI-powered educational agents: Opportunities, innovations, and ethical challenges. *Information*, 16(6), 469. <https://doi.org/10.3390/info16060469>
- CPVP (Commissariat à la protection de la vie privée du Canada). (2022, 12 octobre). *Blogue savoir techno : Faire du neuf avec du vieux – la réalité des données synthétiques*. <https://www.priv.gc.ca/fr/blogue/20221012/>
- Dyens, O. (2026). *L'humanité artificielle: Quand l'IA nous réinvente, qui renaît ?* Éditions XYZ. <https://editionsxyz.com/livre/l-humanite-artificielle/>
- École Branchée. (2026, 20 février). *Évaluer à l'ère de l'IA : Une occasion de mieux mesurer les compétences – École branchée*. <https://ecole-branchee.com/cop-evaluer-ia-livrible-cycle1/>
- Freeman, J. (2025). *Student generative AI survey 2025* (Policy Note). Higher Education Policy Institute (HEPI) & Kortext. <https://www.hepi.ac.uk/wp-content/uploads/2025/02/HEPI-Kortext-Student-Generative-AI-Survey-2025.pdf>
- Gouvernement du Québec. (2025). *Déploiement et intégration de l'intelligence artificielle en enseignement supérieur : Cadre de référence issu des travaux de l'instance de concertation nationale sur l'IA en enseignement supérieur*. Ministère de l'Enseignement supérieur. <https://cdn-contenu.quebec.ca/cdn-contenu/education/Numerique/enseignement-superieur/cadre-referance-travaux-instance-ia-es.pdf>
- Jiang, Y. H., Lu, Y., Dai, L., Wang, J., Li, R., Jiang, B. (2025). Agentic workflow for education: Concepts and applications. In *Proceedings of the 33rd International Conference on Computers in Education*. <https://www.arxiv.org/pdf/2509.01517>

- Kamalov, F., Calong, D., Smail, L., Azizov, D., Thadani, D., Kwong, D., Atif, A. (2025). *Evolution of AI in education: Agentic workflows* (arXiv:2504.20082). <https://doi.org/10.48550/arXiv.2504.20082>
- Karbasi, K., Hong, K. Samadi, A., Pottie, G. (2025). Multi-agent collaborative framework for math problem generation. In *Proceedings of EDM 2025* (poster/demo). <https://educationaldatamining.org/EDM2025/proceedings/2025.EDM.poster-demo-papers.288/index.html>
- Lewis, J. E., Whaanga, H. et Yolgörmez, C. (2024). Abundant intelligences: Placing AI within Indigenous knowledge frameworks. *AI & Society*. <https://doi.org/10.1007/s00146-024-02099-4>
- Lobet, M., Honet, A., Romainville, M. et Wathelet, V. (2024). ChatGPT : quel en a été l'usage spontané d'étudiants de première année universitaire à son arrivée ? *Médiations & médiatisations*, (18), 67–90. <https://doi.org/10.52358/mm.vi18.379>
- Luckin, R. (2025). Nurturing human intelligence in the age of AI: Rethinking education for the future. *Development and Learning in Organizations: An International Journal*, 39(1), 1–4. <https://doi.org/10.1108/DLO-04-2024-0108>
- MacDowell, P., Moskalyk, K., Korchinski, K., Morrison, D. (2024). Preparing educators to teach and create with generative AI. *Canadian Journal of Learning and Technology*. <https://cjlt.ca/index.php/cjlt/article/view/28606>
- Marois, A., Lavallée, I., Boily, G., Ramon Alaman, J., Desrosiers, B. et Lavoie, N. (2025). Chatbot Memory: Uncovering How Mental Effort and Chabot Interactions Affect Short-Term Learning. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 69, No. 1, pp. 2120–2126). Sage CA: Los Angeles, CA: SAGE Publications.
- Mollick, E. (2024). Co-intelligence: Living and working with AI. Portfolio. https://www.researchgate.net/publication/397085676_Co-Intelligence_Living_and_Working_With_AI_1st_Edition_By_Ethan_Mollick_Portfolio_Penguin_Random_House_2024_256_pp_1779_hardcover
- Mollick, E. (2025). Against "Brain Damage". AI can help, or hurt, our thinking. One Useful Thing. <https://www.oneusefulthing.org/p/against-brain-damage>
- OECD. (2025). *Artificial intelligence and education and skills*. OECD Publishing. <https://www.oecd.org/en/topics/sub-issues/artificial-intelligence-and-education-and-skills.html>
- OECD. (2026). *OECD Digital Education Outlook 2026 : Exploring Effective Uses of Generative AI in Education*. OECD Publishing. <https://doi.org/10.1787/062a7394-en>
- Palmer, K. (2026). Agentic AI can complete whole courses for students. Now what? Inside Higher Ed. <https://www.insidehighered.com/news/tech-innovation/artificial-intelligence/2026/02/26/agentic-ai-can-complete-whole-courses-now>
- Paquelin, D., Tendeng, M.-L., Djieufack, E.-B. et Kropf, V. (2025a). *Portrait de l'utilisation de l'IAgen à l'Université Laval*. Obvia. <https://www.obvia.ca/ressources/portrait-de-lutilisation-de-liagen-a-luniversite-laval>
- Paquelin, D., Tendeng, M.-L., Djieufack, E.-B. et Kropf, V. (2025b). *Portrait de l'utilisation de l'IAgen par le corps professoral et enseignant de l'Université Laval*. Obvia. <https://doi.org/10.61737/SECP3363>
- Park, J. (2025). A systematic literature review of generative artificial intelligence (GenAI) literacy in schools. *Computers and Education: Artificial Intelligence*, 6, 100487. <https://doi.org/10.1016/j.caeai.2025.100487>
- Prom Tep, S., Millerand, F., Parada, A. (2023). Accès à la justice et inclusion numérique : au-delà des enjeux technologiques. *Canadian Journal of Law and Society / Revue Canadienne Droit et Société*. 2023;38(2):201-222. doi:10.1017/cls.2023.14
- Proust-Andrewkha, S., Denis, C. (2025). Using ChatGPT in an Instructional Design Assignment: A Study of Students' Perceptions Based on the Model of Situated Acceptance. *RITPU*. <https://www.ritpu.ca/articles/view/562>
- Psyché, V., Ruer, P. (2019). L'apprentissage adaptatif intelligent. *Le Tableau*, 8(4). <https://pedagogie.quebec.ca/le-tableau/lapprentissage-adaptatif-intelligent>
- Radio-Canada. (2026). *Sondage SOM-Radio-Canada IA éducation*. <https://archive.org/details/rapport-final-version-2-sondage-ia/page/n3/mode/2up>
- Ravšelj, D., Keržič, D., Tomažević, N., Umek, L., Brezovar, N., Iahad, N. A., Abdulla, A. A., Akopyan, A., Aldana Segura, M. W., AlHumaid, J., Allam, M. F., Alló, M., Papa Kweku Andoh, R., ... Aristovnik, A. (2025). Higher education students' perceptions of ChatGPT: A global study of early reactions. *PLOS ONE*, 20(2), e0315011. <https://doi.org/10.1371/journal.pone.0315011>
- Spirlet, T. (2025). The real danger of AI in education isn't cheating — it's dependency on Big Tech algorithms, a business professor warns. *Business Insider*. <https://www.businessinsider.com/ai-is-handing-control-of-knowledge-to-big-tech-professor-2025-10>
- Tishcoff, R., Agoe, E., Isik, M. et MacFarlane, A. (2024). *Utiliser l'IA générative pour un apprentissage plus accessible : Points de vue des étudiant-es et du personnel de l'enseignement postsecondaire de l'Ontario*. <https://heqco.ca/wp-content/uploads/2024/11/GenAI-Report-FORMATTED-FR.pdf>
- Toupin, S. (2024) Feminist artificial intelligence: a new era. *Politics and rights review*. <https://politicsrights.com/feminist-artificial-intelligence-a-new-era/>
- UNESCO. (2025). *Guidance for generative AI in education and research*. UNESCO. <https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research>
- Van Dijk, J. A. G. M. 2006. « Digital divide research, achievements and shortcomings ». *Poetics* 34 (4-5): 221-35.
- University Affairs. (2026). Rethinking the role of higher education in an AI-integrated world. <https://universityaffairs.ca/opinion/rethinking-the-role-of-higher-education-in-an-ai-integrated-world/>

Van Dijk, J. A. G. M. 2017. "Digital Divide: Impact of Access." In *The International Encyclopedia of Media Effects*, edited by Patrick Rössler, Cynthia A. Hoffner, and Liesbet van Zoonen, 1–11. New York: John Wiley & Sons.

Xiaoyu, W., Zainuddin, Z. et Leng, C.H. (2025). Generative artificial intelligence in pedagogical practices: A systematic review of empirical studies (2022–2024). *Cogent Education*, 12(1). <https://doi.org/10.1080/2331186X.2025.2485499>

Yu, J. H., Chauhan, D., Iqbal, R. A. et Yeoh, E. (2025). Mapping academic perspectives on AI in education: Trends, challenges, and sentiments in educational research (2018–2024). *Educational Technology Research and Development*, 73, 199–227. <https://doi.org/10.1007/s11423-024-10425-2>

Zhang, R. (2025). Multi-agent systems for learning assessment in education: A comprehensive survey. In *Proceedings of the 2025 3rd International Conference on Educational Knowledge and Informatization (EKI '25)*. <https://doi.org/10.1145/3765325.3765390>

Axe Sobriété numérique et transition socio-écologique

ADEME (2024), *Évaluation environnementale des effets directs et indirects du numérique pour des cas d'usage. État de l'art des méthodologies existantes*. <https://bibliothèque.ademe.fr/economie-circulaire-et-dechets/7785-evaluation-environnementale-des-effets-directs-et-indirects-du-numerique-pour-des-cas-d-usage.html>

ADEME (2025a), Actualisation des chiffres de l'impact du numérique en France. Lien : <https://ecoresponsable.numerique.gouv.fr/actualites/actualisation-ademe-impact/>

ADEME (2025b), *Quantification environnementale directe et indirecte du numérique pour des cas d'usage. Numérisation de dix cas d'usage. Synthèse partielle*. Novembre 2025. Lien : <https://bibliothèque.ademe.fr/economie-circulaire-et-dechets/8738-evaluation-environnementale-des-effets-directs-et-indirects-du-numerique-pour-des-cas-d-usage.html>

AFNOR (2024). *Référentiel général pour l'IA frugale*. <https://www.afnor.org/actualites/intelligence-artificielle/referentiel-reduire-impact-environnemental-ia/>

Bellon-Maurel V., Gauche K. Enriquez M.L. et Lyon-Caen N. (dir.) (2026). *Appréhender l'agriculture numérique. 10 ans de recherche interdisciplinaire au sein de l'Institut #DigitAg*, Éditions Quæ, Versailles. <https://doi.org/10.35690/978-2-7592-4255-9>

Faure A. et Roussilhe G. (2024). « Quelle contribution du numérique à la décarbonation ? », La Note d'analyse, n° 141, France Stratégie. <https://www.strategie-plan.gouv.fr/publications/contribution-numerique-decarbonation>

Green B. (2019). « "Good" isn't good enough », in *NeurIPS Joint Workshop on AI for Social Good*. <https://www.benzevgreen.com/19-ai4sg/>

IEA - International Energy Agency. (2025). *Energy and AI* <https://www.iea.org/reports/energy-and-ai>

IPCC (2023). *Climate change 2023. Synthesis report*. https://www.ipcc.ch/report/ar6/syr/downloads/report/IPCC_AR6_SYR_FullVolume.pdf

ITU. (2022). *Methodologies for the Assessment of the Environmental Impact of the Information and Communication Technology Sector*. <https://www.itu.int/ITU-T/recommendations/rec.aspx?rec=15030&lang=fr>

Joshi K. (2026). « The AI Climate Hoax: Behind the Curtain of How Big Tech Greenwashes Impacts », Beyond Fossil Fuels/Friends of the Earth/Stand.earth/Green Screen Coalition/Climate Action Against Disinformation/Green Web Foundation. https://beyondfossilfuels.org/wp-content/uploads/2026/02/AI-for-climate-claims-Report_FEB-2026_FINAL-2-16.pdf

Lascoumes P. et Le Galès P.(dir.). (2004), *Gouverner par les instruments*, Sciences Po, Les Presses, Paris.

Ligozat A.-L., Lefevre J., Bugeau A. et Combaz J. (2022), « Unraveling the Hidden Environmental Impacts of AI Solutions for Environment Life Cycle Assessment of AI Solutions », *Sustainability*, 14(9), 5172. <https://doi.org/10.3390/su14095172>

Parasie S. et Shultz S. (2024). « Le numérique au service de la transition écologique ? Un panorama des recherches en sciences sociales », Dans *Réseaux* 2024/2 N° 244, Éditions La Découverte.

Rolnick, D., Donti, P.L., Kaack, L.H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A.S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A.; et al. (2022), « Tackling Climate Change with Machine Learning ». *ACM Comput. Surv.* 55, 2, Article 42 (February 2022), 96 pages. <https://doi.org/10.1145/3485128>

Roussilhe G. (2023). « Les émissions évitées de la numérisation ». <https://gauthierroussilhe.com/articles/les-emissions-evitees-de-la-numerisation>

Shift Project (2025). *Intelligence artificielle, données, calculs : quelles infrastructures dans un monde décarboné?* <https://theshiftproject.org/publications/intelligence-artificielle-centres-de-donnees-rapport-final/>

Villalba, B. (2024). Conflits de sobriété. In *Administration*, 282, 34–37.

Axe Droit, cyberjustice et cybersécurité

Aidid, A. et Alarie, B. (2023). *The Legal Singularity: How Artificial Intelligence Can Make Law Radically Better*. University of Toronto Press.

Gouvernement du Canada. (2025). *Stratégie pancanadienne en matière d'intelligence artificielle*. <https://ised-isde.canada.ca/site/strategie-ia/fr>

Guillaud, H. (2025). *Les algorithmes contre la société*. La fabrique éditions.

Tréguer, F. (2024). *Technopolice: La surveillance policière à l'ère de l'intelligence artificielle*. Éditions divergences.



obvia

obvia.ca

Pour consulter toutes
nos ressources :

