Fédération des cégeps

Registre de risques liés à l'utilisation de l'intelligence artificielle

Risque	Description et/ou références
Dépendance technologique (IRIAES)	Le manque de vigilance dans l'utilisation de l'IA expose les organisations à des risques d'erreurs, de biais et de désinformation, pouvant compromettre la qualité des décisions. Il peut également entraîner une dépendance excessive aux outils automatisés. Une utilisation responsable exige donc un esprit critique constant et une compréhension claire des limites de ces technologies. Le manque de transparence dans les décisions générées par l'IA crée un risque important d'opacité, rendant difficile la compréhension des critères utilisés et la justification des résultats. Cette absence de visibilité peut compromettre la confiance, limiter la capacité à détecter les biais et compliquer la responsabilité organisationnelle. Pour garantir une utilisation fiable, il est essentiel d'adopter des systèmes explicables et des processus de gouvernance clairs.
Perte d'expertise humaine (IRIAES)	
Décisions automatisées sans garde-fou (IRIAES)	
Manque de transparence concernant la prise de décision ou le mode de fonctionnement (IRIAES)	
Manque de vigilance (IRIAES)	Malgré leurs performances, certains modèles d'IA générative, dont ChatGPT, produisent souvent des affirmations fausses ou approximatives (mésinformation). L'information fournie par les outils d'IA générative est aussi parfois biaisée. Elle peut être porteuse de valeurs, de croyances ou d'interprétations susceptibles de reproduire ou d'amplifier certains biais sociaux. Ces risques qui touchent à la qualité de cette information, semblent particulièrement problématiques dans un contexte où les personnes enseignantes ou étudiantes pourraient avoir tendance à accorder une confiance exagérée à l'IA, au détriment de leur propre jugement. (*Intelligence artificielle générative en enseignement supérieur : enjeux pédagogiques et éthiques, p. XIII*) Les systèmes d'IA se basent sur des données existantes. Ces données peuvent avoir des biais comme des préjugés ou des erreurs. Les contenus produits par l'IA peuvent donc contenir des biais et causer des dommages. (Gouvernement du Québec) L'IA peut parfois se tromper et inventer des informations qui ne sont pas vraies. Ce phénomène est connu sous le nom « D'hallucination ». Cela signifie que l'IA peut communiquer de fausses informations, ce qui présente un risque de mésinformation. Il est donc impératif de vérifier que les informations sont correctes et que les sources sont fiables. (Gouvernement du Québec)
Déresponsabilisation humaine (IRIAES)	
Manque de rigueur (IRIAES)	
Biais algorithmiques (IRIAES) et Biais dans les données (IRIAES)	
Risque de discrimination (IAGEPE)	
Risque de mésinformation (IAGEPE)	
Non-respect du principe de la fiabilité et de la robustesse (EPUIA)	
Objectifs mal définis (IRIAES)	Un mauvais encadrement de l'utilisation de l'intelligence artificielle pourrait entraîner des conséquences majeures sur les résultats obtenus et la confiance dans les résultats de recherche.
Absence de consentement éclairé (IRIAES)	L'absence de consentement éclairé dans l'usage de l'IA crée un risque majeur pour les individus, qui peuvent voir leurs données utilisées sans comprendre les finalités, les implications ou les limites des systèmes automatisés. Assurer un consentement véritablement informé exige transparence, pédagogie et contrôle effectif sur l'usage des données.

Risque	Description et/ou références
Imposition de l'outil sans consultation ou formation (IRIAES)	Il est important de se doter de structures d'évaluation des outils (tels une EFVP ou une ARP) et de mettre en place les processus nécessaires à l'adoption et l'intégration responsables des outils d'IA. L'utilisation non-structurée de l'IA introduit plus de risque pour un organisme que de retombées bénéfiques.
Absence de mécanisme de révision (IRIAES)	L'absence de mécanismes de révision pour les systèmes d'IA limite la capacité à détecter, corriger ou contester les erreurs, les biais ou les décisions injustifiées qu'ils peuvent produire. Cette absence de contrôle affaiblit la responsabilité organisationnelle et augmente les risques opérationnels, juridiques et éthiques. Mettre en place des processus de révision rigoureux est essentiel pour garantir la fiabilité, la transparence et la conformité des usages de l'IA.
Accumulation des données sensibles (IRIAES)	La multiplication des outils IA peut générer une accumulation de données sensibles et une décentralisation du contrôle exercé sur les données d'un organisme public. Il est essentiel d'assurer un encadrement structuré.
Utilisation des failles d'un modèle d'IA (Gouvernement du Québec) Utilisation détournée (usage malveillant, manipulation, censure) (IRIAES) Non-respect des principes de confidentialité (IRIAES) Non-conformité aux considérations éthiques et obligations légales (EPUIA) Risques pour la vie privée associés à l'utilisation de l'IA générative (IAGEPE) Risques liés à la protection des données et à la sécurité de l'information	L'utilisation de l'IA dans les établissements d'enseignement supérieur doit se faire en tout respect des politiques officielles du gouvernement québécois, notamment la *Loi sur la gouvernance et la gestion des ressources informationnelles* et la *Loi sur l'accès aux documents des organismes publics et sur la protection des renseignement personnels*. Les systèmes d'IA collectent fréquemment des données et les gardent en mémoire. Quand vous utilisez ces outils, vous risquez de partager des informations confidentielles ou des renseignements personnels sans le vouloir. Ces données peuvent être gardées en mémoire par l'IA et réutilisées plus tard pour répondre à la demande d'une autre personne. (Gouvernement du Québec) Les modèles d'IA sont vulnérables aux cyberattaques comme : **L'attaque par empoisonnement** Ce type d'attaque consiste à insérer des données incorrectes ou malveillantes dans celles utilisées pour entraîner l'AI. Cela peut amener l'IA à fournir des résultats erronés, biaisés ou dangereux. (Gouvernement du Québec) **L'extraction de données** Des individus malveillants peuvent essayer d'accéder à des informations confidentielles enregistrées dans les systèmes d'IA. Des renseignements personnels, des informations confidentielles ou même des secrets d'État pourraient être volés si les pratiques de sécurité ne sont pas respectées. (Gouvernement du Québec) **La violation par débridage** Le débridage est une technique où des personnes malveillantes manipulent un système d'IA pour produire du contenu normalement interdit. Cela vise à contourner les mesures de protection éthiques intégrées dans ces systèmes. (Gouvernement du Québec)
Non-respect du principe de la souveraineté numérique (EPUIA)	Les organismes publics doivent viser à diminuer leur dépendance technologique envers les fournisseurs étrangers et à conserver le plus possible un contrôle sur les modèles d'intelligence artificielle, leurs données et les infrastructures qui les hébergent. Un organisme doit donc favoriser, lorsque possible, des fournisseurs québécois ou autrement canadien offrant des solutions hébergées au Québec (EPUIA). La production d'une EFVP permet à un organisme de consigner dans un rapport écrit ses preuves de conformité des principes d'utilisation responsable de l'IA et de justifier son respect des principes directeurs.

Il est à noter que ce registre de risques liés à l'utilisation de l'intelligence artificielle est à titre indicatif seulement et ne constitue pas une liste exhaustive. Chaque organisme est responsable d'évaluer les risques spécifiques liés à l'utilisation d'un outil s'appuyant sur l'intelligence artificielle en effectuant les évaluations nécessaires (EFVP, EFVP-R, ARP) telles que requises par la règlementation.

Sources :

IRIAES - Intégration Responsable de l'Intelligence Artificielle en Enseignement Supérieur

EPUIA - Énoncé de Principe pour une Utilisation Responsable de l'IA

IAGEPE – Intelligence artificielle générative en enseignement supérieur : enjeux pédagogiques et éthiques

Gouvernement du Québec - Risques liés à l'intelligence artificielle | Gouvernement du Québec