



F A É C U M

L'UTILISATION DES TECHNOLOGIES DE L'INFORMATION DANS UN CONTEXTE DE FRAUDE ET PLAGIAT

Le 30 novembre 2022, la mise en ligne de ChatGPT est venue bouleverser le monde universitaire, notamment parce qu'il est possible de générer des travaux universitaires en entier grâce à son aide. Puisque ChatGPT est un programme informatique utilisant d'énormes bases de données et de l'intelligence artificielle (IA), il est capable de fournir des réponses plausibles à des questions parfois très complexes. Pour être utilisé, ce programme ne demande que la création d'un compte gratuit ce qui le rend facilement accessible.

Ce faisant, plusieurs universités se sont penchées sur la modernisation de leurs règlements encadrant la fraude et le plagiat ainsi que sur de meilleurs outils informatiques afin de détecter ces cas. L'Université de Montréal (UdeM) ne fait pas exception à cette tendance. À cet effet, le 14 février 2023, une modification a été apportée au *Règlement disciplinaire sur le plagiat et la fraude concernant les étudiants [et les étudiantes] de premier cycle* et à son équivalent aux cycles supérieurs afin d'interdire l'utilisation d'outils basés sur de l'intelligence artificielle dans le cadre d'évaluations, à moins que la personne enseignant le cours n'ait autorisé explicitement leur utilisation.

En parallèle, l'Université de Montréal est récemment entrée en appel d'offres afin de se pourvoir d'un logiciel de détection des similitudes, notamment dans une optique de faciliter la détection de plagiat. À cet effet, l'Université de Montréal est présentement en discussion avec Turnitin. Afin de baliser l'utilisation de ce genre de logiciel, la Fédération s'est dotée en 2011 de l'*Avis sur la situation du plagiat à l'Université de Montréal*¹.

¹ FAÉCUM. 2011. *Avis sur la situation du plagiat à l'Université de Montréal*. <https://www.faecum.qc.ca/ressources/documentation/avis-memoires-recherches-et-positions-1/avis-sur-la-situation-du-plagiat-a-l-universite-de-montreal>

De plus, la FAÉCUM s'est dotée de nouvelles positions sur la détection de la fraude et du plagiat récemment, notamment en 2020 par l'entremise de la note *La télésurveillance des évaluations à distance*². Cela étant dit, force est de constater que l'évolution rapide des technologies utilisées tant pour produire des documents qui pourraient être frauduleux que celles utilisées pour leur détection nécessite que la FAÉCUM s'assure que son discours sur les outils technologiques utilisés en contexte de fraude et plagiat soit à jour.

LOGICIELS D'INTELLIGENCE ARTIFICIELLE

Selon l'Office québécois de la langue française (OQLF), l'intelligence artificielle est le domaine d'étude tentant de reproduire de manière artificielle les facultés cognitives humaines et l'exécution de tâches qui requièrent en temps normal l'intelligence humaine. L'OQLF indique également que, par extension, l'intelligence artificielle désigne les systèmes informatiques produits dans le but de simuler les capacités cognitives humaines ou accomplir les tâches qui demandent normalement ce genre de capacité³. Plusieurs programmes informatiques entrent dans cette catégorie et permettent de réaliser diverses tâches. Par exemple, DeepAI permet de transformer du texte en une image et DeepL permet de traduire du texte dans plus de 30 langues.

Plus précisément, ChatGPT est un agent conversationnel développé par OpenAI, c'est-à-dire un logiciel créé dans le but d'entretenir une conversation avec une personne utilisatrice donnée. Pour réaliser cette tâche, des millions de documents qui sont utilisés de plusieurs manières différentes ont été soumis au programme. De cette façon, ChatGPT a pu apprendre à analyser une requête et à repérer les informations les plus pertinentes pour répondre à la requête de façon structurée, c'est-à-dire sous forme d'un texte continu. Par conséquent, en analysant une énorme quantité de données, ChatGPT est capable de trouver la réponse qui est statistiquement la plus sensée à une requête donnée^{4,5}. Ce faisant, il est capable de rédiger des textes longs, mais aussi donner des réponses courtes pour répondre à une requête. Il peut également modifier un texte soumis afin d'en retirer certaines erreurs factuelles ou grammaticales. De plus, l'outil est capable de corriger ou créer un code informatique puisqu'il s'agit également d'une forme de texte.

Malgré l'intérêt qu'ils soulèvent, ces logiciels sont présentement accompagnés de plusieurs limites. Par exemple, ChatGPT invente fréquemment de fausses citations ou sources. De plus, puisque ce genre de logiciel est entraîné sur d'immenses bases de données, il a tendance à se tromper lorsque le sujet est général ainsi que lorsqu'il est trop niché. De même, ChatGPT reproduit les biais existants dans les données.

² FAÉCUM. 2020. *La télésurveillance des évaluations à distance*. <https://www.faecum.qc.ca/ressources/documentation/avis-memoires-recherches-et-positions-1/note-sur-la-telesurveillance-des-evaluations-a-distance>

³ Office québécoise de la langue française. 2017. *Intelligence artificielle*. <https://vitrinelinguistique.oqlf.gouv.qc.ca/fiche-gdt/fiche/8385376/intelligence-artificielle>.

⁴ Brown, Tom ; Mann, Benjamin ; Subbiah, Melanie ; Kaplan, Jared *et al.* 2020. *Language Models are Few-Shot Learners*. DOI : <https://doi.org/10.48550/arXiv.2005.14165>.

⁵ Radford, Alec ; Wu, Jeffrey ; Child, Rewon ; Luan, David *et al.* 2019. *Language Models are Unsupervised Multitask Learners*. *OpenAI blog* 1 (8) : 9. <https://lile-extension.github.io/2020/05/27/GPT-技术初探/language-models.pdf>.

Puisqu'une très grande quantité de documents est compilée et que cette base de données inclut des articles scientifiques de grande qualité, mais aussi des articles provenant de sites web véhiculant des informations douteuses, le logiciel peut avoir intégré des informations incorrectes⁶. De même, comme l'outil fait usage de statistiques afin de générer des réponses, des erreurs peuvent être produites. Cependant, il est probable que plusieurs de ces lacunes soient réglées dans les prochaines années.

De plus, plusieurs autres entreprises ont annoncé travailler sur leur propre agent conversationnel comparable à ChatGPT. Par exemple, Google veut lancer son logiciel nommé Bard⁷ et l'entreprise chinoise Baidu développe son propre agent conversationnel⁸. De même, Microsoft travaille à l'intégration de ChatGPT à son moteur de recherche⁹. En somme, ce genre d'outils se développe rapidement. Ce rythme est si rapide que les outils permettant de détecter les textes générés par des logiciels d'intelligence artificielle n'arrivent pas à suivre le rythme. Ce faisant, il est déjà très difficile de détecter les textes générés par des IA. OpenAI a développé un outil de détection de textes produits par leur IA, ChatGPT. Or, seulement 26% des textes rédigés par ChatGPT ont été identifiés¹⁰. Bien que les outils de détection des textes générés par IA ne soient pas très efficaces, d'autres logiciels permettant de réduire le taux de détection existent déjà.

Cette réalité est similaire pour l'ensemble des autres outils basés sur l'intelligence artificielle disponible présentement. Les programmes permettant de détecter leur utilisation ne sont pas au point et peuvent, de toute manière, être fréquemment contournés. Ce faisant, il ne serait pas utile d'interdire strictement toute utilisation d'IA dans les évaluations produites par la communauté étudiante, puisqu'il serait assez difficile de veiller à l'application de cette règle.

Également, comme ces logiciels sont disponibles librement en général, il est essentiel que la communauté étudiante puisse apprendre à en faire un bon usage. L'apprentissage à l'aide de ces logiciels pourrait avoir une grande valeur pédagogique ou considérablement faciliter certaines étapes d'une évaluation. Par exemple, ces outils peuvent être utilisés comme assistant de recherche pour trouver des informations sur un sujet donné qu'une personne étudiante pourrait utiliser dans le cadre d'une rédaction¹¹. De même, dans le cadre d'une évaluation, il pourrait être demandé de traduire un texte à l'aide de DeepL et de corriger les erreurs générées par cette traduction. Ce faisant, les personnes évaluées apprendraient à se familiariser

⁶ Lajoie, Guillaume. 2023. *Qu'est-ce que ChatGPT : Une brève introduction*. (Conférence).

⁷ Urbain, Thomas. 2023. « Après ChatGPT, Google lance son propre robot conversationnel, nommé Bard ». *La Presse*, 6 février. <https://www.lapresse.ca/affaires/entreprises/2023-02-06/intelligence-artificielle/apres-chatgpt-google-lance-son-propre-robot-conversationnel-nomme-bard.php>

⁸ TVA Nouvelles. 2023. « Le chinois Baidu va lancer son propre robot conversationnel face à ChatGPT ». *TVA Nouvelle*, 7 février. <https://www.tvanouvelles.ca/2023/02/07/le-chinois-baidu-va-lancer-son-propre-robot-conversationnel-face-a-chatgpt>

⁹ Boily, André. 2023. « C'est parti pour la recherche sur Bing de Microsoft qui annonce l'intégration de l'IA ChatGPT ». *Le Journal de Montréal*, 10 février. <https://www.journaldemontreal.com/2023/02/10/cest-parti-pour-la-recherche-sur-bing-de-microsoft-qui-annonce-lintegration-de-lia-chatgpt>

¹⁰ Agence France-Presse. 2023. « L'entreprise derrière ChatGPT peine à trouver un antidote ». *Radio-Canada*, 31 janvier. <https://ici.radio-canada.ca/nouvelle/1952532/chatgpt-intelligence-artificielle-open-ai-chatbot>

¹¹ Dassonneville, Ruth. 2023. *ChatGPT pour les étudiant.es et pour la recherche - Utilité et limites de ChatGPT*. (Conférence).

avec les limites de l'outil. Par conséquent, il vaut mieux laisser le corps enseignant baliser l'utilisation de l'IA dans leurs cours plutôt que l'interdire complètement.

Recommandation 1

Que l'Université de Montréal n'interdise pas complètement l'utilisation de logiciels d'intelligence artificielle dans le cadre d'évaluations et qu'elle laisse son corps enseignant établir les limites de son utilisation dans leurs cours.

Cela étant dit, comme ces technologies se développent très rapidement, il est essentiel que le corps enseignant de l'Université de Montréal soit outillé afin qu'il puisse bien baliser l'utilisation des logiciels d'IA existants ou qui seront développés. Afin de faciliter l'intégration de ces technologies dans les salles de classe et favoriser les bonnes pratiques pédagogiques entourant l'IA, il est primordial que l'Université de Montréal développe des formations et des guides qui pourront être mis à la disposition du corps enseignant. Ces formations et guides devront être mis à jour périodiquement afin qu'ils suivent les développements des nouvelles technologies d'IA. À cet effet, le Centre de pédagogie universitaire (CPU) offre déjà plusieurs guides et outils à destination du corps enseignant et pourrait donc développer ces nouveaux outils.

Recommandation 2

Que le Centre de pédagogie universitaire développe des formations et des guides destinés au corps enseignant afin qu'il soit outillé pour faire une utilisation pédagogique des logiciels d'intelligence artificielle. Que ces formations et guides soient mis à jour périodiquement.

Finalement, puisque les règlements sur la fraude et le plagiat ont été modifiés afin d'interdire l'utilisation de logiciels d'intelligence artificielle à moins que la personne responsable du cours l'autorise, il est essentiel que ces modifications soient rapidement communiquées à la communauté étudiante. À cet égard, il est déjà obligatoire d'inclure une rubrique sur la fraude et le plagiat dans les plans de cours. Il serait donc simple d'y ajouter une mention de l'interdiction d'utiliser l'IA dans les évaluations. Également, il est nécessaire que la communauté étudiante soit au courant des modalités selon lesquelles ChatGPT ou toute autre IA peuvent être utilisées dans chacun de leurs cours lorsque permis par le corps enseignant. Afin qu'il ne règne aucun doute sur la nature des modalités d'utilisation, il est primordial que les instructions soient fournies à l'écrit avec les consignes entourant les évaluations. Ces méthodes permettraient à toutes les personnes étudiantes de prendre connaissance à la fois de l'interdiction généralisée et des modalités d'utilisation qui s'appliquent dans son cours lorsqu'elle est permise.

Recommandation 3

Que la rubrique sur la fraude et le plagiat des plans de cours inclue systématiquement une mention que l'utilisation de logiciels d'intelligence artificielle est interdite à moins d'indication contraire de la personne responsable du cours. Lorsque l'utilisation de logiciels d'intelligence artificielle est permise, que les

modalités d'utilisation soient communiquées à l'écrit à la classe et qu'elles se trouvent dans les directives des évaluations.

LOGICIELS DE DÉTECTION DES SIMILITUDES

Les logiciels de détection des similitudes sont des outils qui permettent de repérer les éléments identiques entre deux textes, de même que les paraphrases qui ont pu être formulées à partir d'un premier texte vers un second. Comme ces logiciels ont accès à d'énormes bases de données composées de millions de textes, ils peuvent aisément détecter le plagiat et identifier la source originale du passage plagié.

Au cours de la pandémie de COVID-19, plusieurs universités ont vu les cas de fraude et de plagiat augmenter. À l'Université de Montréal, le nombre de cas de fraude et de plagiat analysés pour l'année 2020-2021 représenterait près du double des cas reçus lors d'une année moyenne avant la pandémie de COVID-19¹². Puisque l'utilisation de tels logiciels vient faciliter le travail des personnes qui corrigent des évaluations, plusieurs universités ont entamé des démarches pour se doter de tels logiciels. C'est le cas de l'Université de Montréal où Turnitin pourrait être implanté aussi tôt que durant l'année 2023-2024.

Cela étant dit, plusieurs logiciels du genre offrent l'option pour les personnes étudiantes d'y déposer leurs travaux avant leur remise finale afin de faire valider les aspects méthodologiques concernant le plagiat. Dans cette optique, il est essentiel que l'implantation d'un tel logiciel à l'Université de Montréal offre cette possibilité. D'autant plus que d'autres universités faisant affaire avec Turnitin offrent cette option, comme l'Université McGill qui utilise le logiciel depuis l'année 2007-2008¹³. Ainsi, il est fondamental que l'Université de Montréal utilise ce logiciel avant tout dans un but de prévention et de formation. Il ne doit pas remplacer d'autres outils, comme les campagnes de sensibilisation, et il ne doit pas être utilisé comme un simple outil de détection de fraude et de plagiat.

De plus, il est important que l'accès au logiciel ne soit pas limité pour les personnes étudiantes. Ces dernières peuvent avoir des travaux à remettre dans plusieurs cours ou elles pourraient désirer resoumettre un travail pour un même cours. Dans le deuxième cas, l'évaluation pourrait avoir subi des modifications à la suite d'une première soumission et il serait nécessaire de le resoumettre afin de valider son intégrité. Limiter le nombre de soumissions pourrait entraver la possibilité qu'aurait une personne de faire valider ses travaux et de les remettre exempts de plagiat. De même, imposer un délai entre deux soumissions pourrait nuire à la communauté étudiante qui ne pourrait pas tester ses évaluations aussi fréquemment qu'elle le désire. Cette limite serait particulièrement négative lors de périodes où plusieurs évaluations sont à remettre dans de courts laps de temps comme les périodes de finaux ou la mi-session.

¹² Morasse, Marie-Ève. 2021. « La pandémie fait bondir les cas de plagiat ». *La Presse*, 15 septembre. <https://www.lapresse.ca/actualites/education/2021-09-15/universites/la-pandemie-fait-bondir-les-cas-de-plagiat.php>

¹³ Conférence des recteurs et des principaux des universités du Québec. 2011. *Les logiciels de détection de similitudes : Une solution au plagiat électronique ?*. https://www.uqac.ca/plagiat/documents/crepuq_plagiat.pdf.

Amendement à la position 1217

Que tout système de détection des similitudes ~~automatique de plagiat~~ serve d'abord à aider les personnes étudiantes lors de leur processus d'apprentissage des règles de méthodologie.

Recommandation 4

Si un logiciel de détection des similitudes est rendu disponible à l'Université de Montréal, que la communauté étudiante puisse soumettre sans pénalité ses évaluations à ce logiciel avant leur remise finale afin d'obtenir un rapport de similitudes. Que ces soumissions puissent être réalisées sans aucune restriction.

Bien qu'ils facilitent la détection des similitudes et du plagiat, ces logiciels peuvent produire des rapports contenant des erreurs. Une partie non négligeable des éléments rapportés dans ces rapports ne constituent pas du plagiat, comme certaines expressions communes qui ne sont pas soumises à un droit d'auteur. De même, une similitude pourrait avoir été détectée en raison d'un oubli d'un guillemet pour fermer une citation alors que le guillemet initial, la note de bas de page ou même la référence à la personne autrice originale pourrait être présente. Il s'agirait alors d'une erreur méthodologique. Puisque d'autres éléments présents attesteraient de l'intention de citer correctement le passage, il est essentiel de ne pas traiter l'erreur comme un cas de plagiat. Par conséquent, le rapport de similitudes d'un tel logiciel ne doit pas être soumis tel quel dans le cadre d'une mise en cause pour fraude ou plagiat.

De même, il est essentiel que le rapport de similitudes ne puisse pas être déposé tel quel comme preuve pour mettre en cause une personne étudiante dans un cas de fraude et de plagiat. Il est nécessaire que le rapport de similitudes soit traité par la personne responsable du cours afin qu'elle cible les éléments qui sont réellement répréhensibles et sanctionnables en vue des règlements sur le plagiat et la fraude. De plus, comme ces rapports peuvent contenir plusieurs éléments superflus, une personne devant se défendre d'avoir plagié ne pourrait pas savoir quels éléments de preuves sont retenus contre elle. Cela rendrait la préparation d'une audience en comité ou d'une lettre de plaidoyer extrêmement difficile. De plus, déposer directement le rapport rendrait beaucoup plus compliquée l'analyse d'un dossier par un comité disciplinaire. Ce faisant, la charge de travail serait complètement déplacée vers la personne étudiante et le comité. En outre, autant le comité que la personne étudiante doivent avoir accès au texte original complet susceptible d'avoir été plagié et pas seulement à un extrait de texte afin de pouvoir bien juger de la situation ou se défendre.

Amendement à la position 1204

Que ~~lors de l'envoi de~~ la lettre de mise en cause envoyée aux personnes étudiantes soupçonnées d'avoir plagié ~~comporte à l'étudiant soupçonné d'avoir plagié,~~ un rapport écrit délimitant clairement les éléments de l'infraction reprochée. lui soit envoyé. Que ~~copie~~ la mise en cause comporte la preuve amassée et les sources originales qui auraient été plagiées. lui soit aussi rendue disponible.

Amendement à la position 1218

Que les rapports résultats provenant d'un outil de détection ~~des similitudes automatique de plagiat~~ soient jugés inadmissibles comme élément de preuve dans un rapport de mise en cause, toujours considérés comme étant seulement un commencement de preuve.

Plusieurs logiciels de détection des similitudes, tels que Turnitin, offrent l'option aux établissements d'enseignement d'archiver les évaluations soumises afin de pouvoir les comparer à de nouvelles soumissions plus tard¹⁴. Or, il est fondamental que les personnes étudiantes conservent le plein contrôle de l'usage qui est fait de leurs travaux, de leurs textes ou de tout autre document soumis à un logiciel de détection des similitudes. Ainsi, le choix doit être donné aux personnes étudiantes quant à l'archivage de leurs travaux par l'entreprise responsable du logiciel. De plus, pour ne pas influencer ce choix, l'option par défaut du logiciel doit être de ne pas archiver une copie afin que la personne étudiante puisse délibérément et en pleine connaissance de cause choisir l'option d'archiver ses travaux. Les personnes étudiantes qui choisissent d'archiver leurs travaux doivent le faire de manière libre et éclairée. Ainsi, l'usage des travaux et la durée de l'archivage doivent être clairement indiqués lorsque les personnes étudiantes font ce choix. Également, les personnes étudiantes ne désirent pas partager leurs travaux ne doivent pas voir leur expérience utilisatrice réduite à cause de ce choix.

Amendement à la position 1219

Que les personnes étudiantes puissent choisir de manière libre et éclairée si ~~choisisse que son~~ leur évaluation est archivée ~~travail soit ajouté à une~~ dans une banque de données liée à un outil de détection ~~des similitudes automatique de plagiat~~. et que l'option par défaut soit de ne pas archiver l'évaluation.

CONCLUSION

L'apparition de nouveaux logiciels, comme ChatGPT, et l'évolution rapide des technologies de l'information demandent une veille constante. Il est également important de réfléchir aux meilleurs moyens pour mettre les nouvelles technologies au service de la pédagogie et de la communauté universitaire sans affecter l'intégrité des évaluations et des apprentissages réalisés par la communauté étudiante. De plus, comme ces technologies existent à l'extérieur des murs de l'Université de Montréal, il vaut mieux former les prochaines cohortes à leur utilisation de manière éthique plutôt que de les interdire complètement. Il vaut donc mieux travailler à utiliser ces technologies de manière pédagogiques afin de créer des environnements propices à la réussite et au développement de compétences pour toute la communauté universitaire.

¹⁴ Turnitin. 2022. "Vie privée et sécurité ». <https://help.turnitin.com/fr/vie-privee-et-securite.html>.