

Le questionnement éthique des internautes face à l'intelligence artificielle. Le cas ChatGpt sur X

Article inédit, mis en ligne le 23 novembre 2024.

Karen Nuvoli

Maîtresse de conférences en Sciences de l'information et de la communication, Université de Lorraine, laboratoire CREM (Centre de recherche sur les médiations)
karen.nuvoli@univ-lorraine.fr

Plan de l'article

Introduction
IA et éthique
Recueillir et analyser des tweets
Le questionnement éthique des internautes face à ChatGPT
Conclusion
Notes
Références bibliographiques

RÉSUMÉ

De nombreux modèles génératifs d'intelligence artificielle ont récemment vu le jour, tels que l'agent conversationnel ChatGPT. Face aux bouleversements qu'annonce l'arrivée de tels outils, cet article examine la manière dont le questionnement et sa mise en œuvre sont tributaires de contextes situés, qui sont culturellement et socialement diversifiés. Cet article propose donc d'analyser et de comparer les principales stratégies discursives des usagers, afin de mieux comprendre les notions abordées par ces derniers ainsi que les préoccupations éthiques associées. Pour cela, nous proposons d'étudier deux corpus constitués à partir de tweets français et italiens.

Mots clés

Intelligence artificielle, éthique, ChatGPT, discours, techniques, X.

THE ETHICAL QUESTIONING OF INTERNET USERS REGARDING ARTIFICIAL INTELLIGENCE. THE CHATGPT CASE ON X

Abstract

Many generative AI models have emerged during the last months. One of them is ChatGPT, a conversational agent which can answer an infinite number of questions. In light of the disruptions heralded by the arrival of such tools, this article examines how ethical questioning manifests in the practices of internet users on X. This inquiry and

its implementation are contingent on situated contexts, which are culturally and socially diversified. The purpose of this study is to analyze and compare the main discursive strategies taken by the users, to better understand the concepts and the related concerns. To this end, we propose to examine two corpora consisting of French and Italian tweets.

Keywords

Artificial intelligence, Ethics, ChatGPT, discourse, technique, X

EL CUESTIONAMIENTO ÉTICO DE LOS INTERNAUTAS ANTE LA INTELIGENCIA ARTIFICIAL. EL CASO DE CHATGPT EN X

Resumen

En los últimos meses han surgido varios modelos de inteligencia artificial generativa, como el agente conversacional ChatGPT. Frente a las transformaciones que anuncia la llegada de tales herramientas, este artículo examina cómo se manifiestan las cuestiones éticas en las prácticas de los usuarios de internet en X. Este cuestionamiento y su implementación dependen de contextos situados, que son cultural y socialmente diversificados. Por ello, este artículo propone analizar y comparar las principales estrategias discursivas utilizadas por los usuarios, con el fin de comprender mejor los conceptos abordados y las preocupaciones éticas asociadas. Para ello, proponemos estudiar dos corpus formados por tweets en francés e italiano.

Palabras clave

Inteligencia artificial, ética, ChatGPT, discurso, técnico, X.

INTRODUCTION

« Nous pouvons espérer que les machines pourront égaler les hommes dans tous les domaines purement intellectuels ». C'est ainsi que le grand logicien Alan Turing (1912-1954) conclut un célèbre article publié en 1950 intitulé « *Computing Machinery and Intelligence* ». Dans cet article qui prend pour point de départ la question suivante : « Les machines peuvent-elles penser ? », il précise que les candidates auxquelles il songe sont à l'image des ordinateurs qui venaient tout juste d'être inventés, en grande partie à partir d'idées qu'il avait lui-même élaborées. Les ingrédients du rêve de l'intelligence artificielle, une expression qui n'avait pas encore été inventée, étaient alors réunis : la pensée, l'intelligence et l'ordinateur (Andler, 2023).

L'intelligence artificielle (IA), est définie comme la capacité des machines à effectuer des tâches qui nécessitent normalement une intelligence humaine, telles que la reconnaissance visuelle ou la compréhension du langage naturel. Son champ d'application s'est considérablement élargi au cours des dernières décennies. La prolifération de l'IA dans la société soulève cependant des questions complexes qui concernent essentiellement la sécurité, la vie privée, l'éthique, ou encore les conséquences socio-économiques. Bien que cette technique soit prometteuse, les nombreuses recherches consacrées à ce domaine (Tsamados *et. al.*, 2021 ; Bruneault, Sabourin Laflamme, 2021) soulignent différents enjeux éthiques tels que les risques de biais, la discrimination algorithmique ou encore l'utilisation non autorisée des données personnelles.

Quand on aborde la question éthique, il est important de rappeler les trois positionnements avancés par le groupe d'études et de recherche sur l'éthique et le numérique en information-communication (GER GENIC), abordant les spécificités liées au numérique (Domenget *et alii.*, 2022) :

L'éthique sous le prisme info-communicationnel : l'éthique numérique doit interroger les changements sociotechniques qui accompagnent les usages des dispositifs d'information-communication, lesquels s'inscrivent dans un ensemble plus large de pratiques sociales.

L'éthique comme pratique située : « l'éthique n'est ni de nature juridique, ni une expertise, mais une forme de questionnement sur des situations problématiques concrètes qui s'exprime dans l'action. Selon l'approche de l'éthique située, l'éthique est avant tout un questionnement, une préoccupation, un souci de vigilance qui s'exerce en référence à un système de valeurs, comme la justice sociale, la responsabilité, la compensation de vulnérabilités individuelles, etc. (Zacklad et Rouvroy, 2022) » (Domenget, Roelens, 2024, p. 3).

L'éthique dans sa dimension internationale : « l'éthique doit être abordé sous une focale internationale, ce qui rappelle la nécessaire prise en compte des contextes. En effet, bien que l'éthique repose sur des principes universels, sa mise en œuvre dépend de contextes situés et localisés, culturellement et socialement divers (AoIR, 2019) » (Domenget, Roelens, 2024, p. 3).

Ainsi, l'éthique, bien qu'elle soit souvent perçue comme un principe universel, se révèle profondément influencée par les contextes culturels, sociaux et géographiques dans lesquels elle est appliquée.

Dans ce contexte, cet article propose de s'intéresser au questionnement éthique des internautes autour des IA génératives, en se concentrant sur les modèles les plus récents tels que ChatGPT. L'objectif n'est pas tant d'aborder les questionnements éthiques dans leur ensemble, mais plutôt d'explorer - en partant d'un terrain de recherche défini - le questionnement éthique des internautes en s'appuyant ici sur une analyse comparative entre la France et l'Italie, en tenant compte de la diversité des perspectives et des enjeux locaux.

Dans ce contexte, la décision d'entreprendre une analyse comparative¹ entre l'Italie et la France découle de l'évaluation des particularités et des traits communs aux deux pays, qui se prêtent valablement à une étude de ce type. Les deux pays sont caractérisés par des avancées techniques similaires et une adoption stricte du Règlement européen sur la protection des données. Ils manifestent tous deux également une attention particulière à l'égard de ChatGPT bien que depuis le 31 mars 2023, l'Autorité de régulation italienne² ait opté pour la suspension temporaire du robot conversationnel dans le pays pour non-respect du RGPD européen. En France, la CNIL a, de son côté, préféré encadrer plutôt qu'interdire, en proposant un plan d'action³ pour appréhender le fonctionnement des systèmes d'IA et leurs importances sur les personnes.

Nous avons donc constitué deux corpus, le premier faisant référence à des tweets français et le second à des tweets italiens. Les limites méthodologiques liées à l'usage des données issues de X⁴ sont multiples. Parmi celles-ci, citons la non-représentativité des profils sociaux

...

1. Pour d'autres analyses comparatives France/Italie voir : Volpe, V., Zaza, O. (2024) ; Garzonio, E., Nuvoli, K. (2022).

2. Provvedimento del 30 marzo 2023 [9870832] : <https://www.garantepriacy.it/web/guest/home/docweb/-/docweb-display/docweb/9870832>

3. Intelligence artificielle : le plan d'action de la CNIL : <https://www.cnil.fr/fr/intelligence-artificielle-le-plan-daction-de-la-cnil>

4. Depuis 2023 et son rachat par Elon Musk, le réseau social Twitter a été rebaptisé X.

qui s'expriment sur les réseaux et sur les médias numériques. Dans le cas de notre étude, X n'est considéré ni comme un échantillon représentatif de la population ni comme un espace prédictif des opinions de la population. Les travaux de Julien Boyadjian (2016) sur les publics qui y publient des messages à caractère politique rendent compte de l'existence d'une « très forte sélection sociale » qui se matérialise par une surreprésentation de publics jeunes « issus pour la plupart de milieux sociaux relativement aisés, masculins, diplômés, cadres et politisés » (p.114). Cette absence de représentativité peut devenir un atout pour des recherches ciblées sur certains aspects. Dans notre étude, les discussions sur X reflèteraient « *avant tout les préoccupations et thématiques mises à l'agenda par les grands médias (Neuman et al., 2014) mais sont aussi l'expression des citoyens lorsqu'ils commentent et partagent l'actualité* (Mercier, 2021, p.52) ».

IA ET ÉTHIQUE

Le développement de l'intelligence artificielle, tributaire de l'avancement du numérique et des techniques de l'information et de la communication (TIC), occupe aujourd'hui une place importante dans le débat public. Ces dernières années, l'IA a régulièrement fait la une des médias, en particulier depuis le lancement de ChatGPT fin 2022.

« *Si les termes d'IA et d'algorithme dominent aujourd'hui le débat, ils restent cependant très imprécis tant ils recouvrent des ensembles de dispositifs plus ou moins définis et complexes (robotique, algorithme, dialogue naturel, machine learning, apprentissage profond, système autonome, etc.)* » (Crepel, Cardon, 2022, p. 133).

Définir l'IA n'est donc pas chose simple. En effet, elle se présente comme une boîte opaque et comme un domaine complexe « fondé autour d'un objectif ambitieux : comprendre comment fonctionne la cognition humaine et la reproduire » (Villani, 2018, p. 9). Pour Julien Nocetti, chercheur à l'Institut français des relations internationales (IFRI), spécialisé dans la gouvernance du web : « L'IA consiste avant tout en des applications concrètes – reconnaissance faciale, traitement automatisé du langage, vision par ordinateur, voiture autonome, etc. » (Nocetti, 2019, p. 10). Il s'agit donc de méthodes qui donneront la possibilité aux ordinateurs de se comporter « intelligemment ». Aujourd'hui, un consensus semble émerger sur le fait qu'un système d'IA repose sur une machine qui a la possibilité, pour un ensemble donné d'objectifs définis par l'homme, d'émettre des prédictions ou des recommandations influençant des environnements réels ou virtuels (Meneceur, 2022). L'IA désigne donc tout d'abord « un objet qu'on cherche à créer : un système doté d'une certaine propriété, mais l'expression désigne aussi une discipline, qui se donne pour but de concevoir, c'est-à-dire de caractériser et de construire, cet objet » (Andler, 2023, p. 22). Le terme est apparu pour la première fois dans les années 1950, après la publication de l'article d'Alan Mathison Turing intitulé « *Computer machinery and intelligence* » (Turing, 1950). Cet article est considéré comme le premier ouvrage évoquant l'IA. Le « Test de Turing », ou *Imitation Game*, qui y est décrit, avait pour objectif de démontrer la capacité cognitive d'une machine.

Selon plusieurs auteurs (Andler, 2023 ; Ménessier, 2022 ; Villani, 2018), la naissance de l'IA est souvent associée à une conférence organisée à Dartmouth en 1956, qui réunissait une vingtaine de chercheurs en cybernétique. Parmi eux, John McCarthy travaillait à la meilleure manière de doter les ordinateurs de comportements intelligents, et Marvin Minsky, avait conçu une machine neuronale imitant le cerveau d'un rat. C'est lors de ce colloque qu'a été présenté le programme informatique *Logic Theorist*, démontrant qu'une machine avait la possibilité de résoudre un problème non chiffré en développant un raisonnement humain. A ce sujet, le projet de recherche des « pères » de l'IA stipule leur

hypothèse selon laquelle « *chaque aspect de l'apprentissage ou n'importe quelles caractéristiques de l'intelligence peuvent, en principe, être décrites de manière suffisamment précise afin qu'une machine puisse être construite pour la simuler* » (McCarthy *et al.* 1955, p. 1). Les différentes définitions de l'IA de cette époque laissent entendre que l'on fabriquerait une nouvelle intelligence concurrente de l'intelligence humaine. Depuis la conférence de Dartmouth, « *l'intelligence artificielle s'est développée, au gré des périodes d'enthousiasme et de désillusion qui se sont succédées, repoussant toujours un peu plus les limites de ce qu'on croyait pouvoir n'être fait que par des humains* » (Villani, 2018, p. 9).

Les évolutions techniques des années 1990, notamment dans le domaine informatique, ainsi que l'amélioration des capacités de calcul des ordinateurs permettant de développer le *deep learning*, ou « apprentissage profond » – à différencier du *machine learning*, « apprentissage machine » qui s'appuie sur une quantité bien plus réduite de données –, entraînent un renouveau dans la recherche et font de l'IA un domaine de compétence et de recherche à part entière.

Ces dernières années, les progrès considérables dans les techniques de *deep learning* ont favorisé l'émergence de modèles capables de générer, à partir d'instructions simples, des contenus de manière autonome, avec un niveau de qualité et de complexité sans précédent. On parle alors d'IA générative, c'est-à-dire une IA entraînée sur des données massives et générant de nouveaux contenus (texte, image, audio, vidéo) à partir d'une information d'entrée (appelée « prompt ») saisie par l'utilisateur (Gozalo-Brizuela, Garrido-Merchan, 2023).

Pour entraîner les modèles d'IA génératives, les architectures antagonistes génératifs (GANs) et les *transformers*, tels que GPT et BERT⁵, apparaissent comme les architectures les plus utilisées. Celles-ci décomposent les textes par des algorithmes en *tokens* (suite de caractères), en utilisant le mécanisme d'auto-attention permettant d'établir des relations entre les mots et de déterminer le sens d'un mot en fonction du contexte (Vaswani *et alii*, 2017). Parmi ces modèles massifs de langage, en anglais *large language models* (LLM), nous trouvons celui qui génère à l'heure actuelle le plus de débats au sein du public, à savoir ChatGPT (à sa quatrième version début 2024). Lancé officiellement le 30 novembre 2022 par Open AI, ChatGPT est un modèle de langage capable de produire du contenu écrit. Le terme « ChatGPT » est la contraction entre le mot « chat » pour conversation et GPT pour « *Generative Pre-trained Transformer* » (transformateur génératif pré-entraîné). Il s'agit d'un modèle de traitement de langage naturel de grande taille (LLM : *large language Model*), reposant sur des réseaux de neurones. Ces modèles sont pré-entraînés sur de vastes ensembles de données textuelles afin de permettre à ChatGPT de produire du langage et de réaliser des tâches.

Ce système conversationnel a suscité à la fois un grand enthousiasme et des controverses significatives, notamment sur ses aspects éthiques, amplifiés par une médiatisation spectaculaire : deux mois après son lancement, ChatGPT dépassait les 100 millions de comptes enregistrés, faisant de cet outil l'application ayant connu la croissance la plus rapide à ce jour dans le monde (Hu, 2023).

L'éthique, en particulier en philosophie, est souvent définie comme « l'ensemble des principes et valeurs guidant des comportements sociaux et professionnels, et inspirant des règles déontologiques » (Farjat, 2004, p. 157). En matière d'IA, l'éthique interroge donc l'ensemble de la chaîne de conception, création, réalisation, utilisation et régulation d'outils techniques. Les publications académiques dans le domaine sont nombreuses :

...

5. BERT (Bidirectional Encoder Representations from Transformers) est un modèle de traitement de langage naturel de type Transformers conçu par Google.

« Des recherches tentent de proposer des modes de programmation qui assurent que les IA soient «éthiques par leur conception» (*ethical by design*). Des premiers résultats, prometteurs, visent à assurer des qualités telles que l'exécution correcte des algorithmes, la confidentialité des données, l'identification et l'élimination de biais dans les données... Mais les outils de formalisation nécessaires à la traduction de concepts éthiques de haut niveau contraignant le code semblent encore inaccessibles » (Linden, 2020, p. 24).

Outre les publications académiques, un grand nombre de déclarations et chartes présentant des principes éthiques, des recommandations ou des lignes directrices a été publié au cours des dernières années. En avril 2019, le groupe d'experts de haut niveau en IA (HLEG AI) nommé par la Commission européenne avait publié ses règles éthiques pour une IA de confiance. Leur rapport final indique quatre principes éthiques pour le développement et le déploiement de l'IA : le respect de l'autonomie humaine ; la prévention de toute atteinte ; l'équité (*fairness*) et l'explicabilité. Plus récemment, en 2020, l'Académie Pontificale pour la Vie, Microsoft, IBM, la FAO et le gouvernement italien, ont signé l'« Appel pour une éthique de l'IA », un texte visant à promouvoir une approche éthique de l'IA, c'est-à-dire l'adoption d'une intelligence artificielle centrée sur l'homme et digne de confiance. Dans le document « *les promoteurs de l'appel expriment leur désir de travailler ensemble, pour promouvoir une «algor-éthique», à savoir l'utilisation éthique de l'IA telle que définie par les principes suivants : transparence ; inclusion ; responsabilité ; impartialité ; fiabilité ; sécurité et respect de la vie privée* » (Appel de Rome pour une éthique de l'IA, 2020, p. 111). On remarque le caractère hétérogène de la liste fournie, « *hétérogénéité qui repose sur le mélange entre des critères purement techniques validant un usage fiable de la technologie basée sur la robustesse de cette dernière, et des critères plutôt éthiques basés sur les valeurs de la démocratie (telle que la non-discrimination)* » (Ménissier, 2022, p. 82).

Bien que principe universel, la compréhension de l'éthique et sa mise en œuvre sont tributaires de contextes situés et localisés, qui sont également culturellement et socialement divers (Franzke *et al.*, 2020). Contrairement à certains points de vue, l'éthique ne se réduit pas à une conception prédéterminée ou à un comportement prédéfini qu'il faut adopter. Elle se présente plutôt comme « *une réflexion sur nos pratiques individuelles et sociales et sur les valeurs qu'elles actualisent, menant à leur évaluation à travers l'exercice d'un jugement pratique et à des actions conséquentes* » (Lacroix *et alii*, 2017, p. 37-46). Comme le souligne Boudreau (2019), il est question d'« IA éthique » associée à une démarche, une réflexion, un principe, une valeur, etc. Cela signifie que les problèmes éthiques sont pensés et traités de manière concrète et contextualisée.

RECUEILLIR ET ANALYSER DES TWEETS

Concernant la phase de collecte des données⁶ et de construction du corpus, nous avons choisi d'utiliser Twitter Archiver, une extension de X qui permet d'extraire et de sauvegarder les tweets sur des feuilles de calcul Google *via* des requêtes de recherche, notamment la recherche par hashtags, par géotags ou par mots-clés. La constitution des deux corpus repose sur la définition de mots clés (#ChatGPT #etica #ethique) faisant référence à notre problématique.

...

6. Twitter fournit des API pour la collecte des tweets. Ces APIs ont de nombreuses limites sur la taille, la durée et le type de corpus qui peut être collecté. Cependant, comme ils sont disponibles ils ont attiré l'attention de nombre de chercheurs. Pour en saisir le paysage, son étendu, ses évolutions et ses limites, voir : (Boyadjian 2014 ; Severo, Lamarche-Perrin, 2018 ; Longhi, 2020).

Les données sont constituées initialement de 33 195 tweets pour le corpus français et 9 517 pour le corpus italien collectés⁷ entre le 8 décembre 2022 et le 8 avril 2023 : cette période englobe à la fois le lancement de l'application en Europe et la mise en place d'actions d'interdiction dans différents pays, dont l'Italie le 1 avril 2023. Ils ont ensuite été testés dans une pré-captation (non intégrée dans notre corpus) pour vérifier qu'ils étaient bien en référence avec le sujet. Notons que le corpus italien est bien plus petit en nombre de tweets collectés, ce qui s'explique en partie par un nombre d'utilisateurs X environ deux fois plus faible qu'en France⁸. Le processus de collecte de données a permis de relever des mots-dièse (#) utilisés pour structurer les préoccupations liées aux questionnements éthiques de ChatGPT (ex : #biais, #plagiat, #privacy etc.). Ces derniers ont été définis par l'étude des échanges observés sur X pendant les premiers débats relatifs à la mise à disposition de ChatGPT en libre accès pour les citoyens en France et en Italie. Après suppression des doubles tweets et ceux exploitant les trends hashtags pour promouvoir un contenu de nature différente, nous avons obtenu deux corpus à partir des mots clés précédemment cités. Les deux corpus ont été analysés avec la méthode Reinert (Reinert, 1997) implémentée dans le logiciel libre IRaMuTeQ :

« Cette méthode permet de déterminer les différentes thématiques qui structurent un corpus textuel. Elle repose sur une classification hiérarchique descendante qui peut être décrite comme une succession de bipartitions reposant sur une analyse factorielle des correspondances » (Ratinaud, Smyrnaiois, et alii., 2019, p.190).

Cette méthode donne la possibilité de constituer des « classes terminales, décrites à partir du lexique qui les caractérise. Ce lexique est constitué des mots qui sont significativement surreprésentés dans la classe si on la compare à l'ensemble des autres classes (sur la base d'un Chi²). Les « mondes lexicaux » (Reinert, 1997) qui se dégagent présentent alors les différentes thématiques abordées dans le corpus » (Ratinaud, Smyrnaiois, et alii 2019, p.190).

LE QUESTIONNEMENT ÉTHIQUE DES INTERNAUTES FACE À CHATGPT

Les premières observations concernent la structuration des deux corpus. En analysant les deux corpus, nous avons isolé une série d'arguments que nous avons réunis en deux macro-catégories (données personnelles et biais ; données personnelles et *privacy*), chacune ayant des caractéristiques syntaxiques et discursives propres.

Données personnelles et biais

Pour le corpus français, nous avons identifié quatre clusters principaux (Figure 1) ; au-delà, les clusters devenant répétitifs et trop hétérogènes. D'un point de vue hiérarchique, le premier cluster qui se dégage est le 4, suivi du cluster 3, puis de deux classes jumelles constituées des clusters 1 et 2.

Le cluster 3 (31,3%) est le plus important en termes de volume, il porte sur la thématique des données. Les mots clés associés à ce cluster comprennent les termes « donnée » (chi² 143,46), « partager » (36,13), « entraînement » (33,34) et « confidentiel » (28,84). Les tweets analysés à l'intérieur de ce cluster concernaient essentiellement le traitement des données personnelles en l'absence du consentement préalable des personnes concernées et la collecte des données dans le but d'entraîner les IA. En effet, certaines IA reposent

...

7. La collecte automatique des données à travers les API de X était possible jusqu'au 30 mars 2023. Pour les données du 31 mars 2023 au 8 avril, nous avons utilisé la collecte manuelle à travers la fonction recherche avancée disponible sur X.

8. Statistiques publiés par X, dans son 1er « Rapport de transparence », voir : DSA Transparency Report - April 2024 - <https://transparency.x.com/dsa-transparency-report.html#/>

sur l'utilisation d'une quantité considérable de données pour former et améliorer les algorithmes. Cette dépendance accrue envers les données a transformé celles-ci en des ressources de grande valeur, suscitant un intérêt marqué de la part de divers acteurs, qu'ils soient publics ou privés. Cela a engendré une série de défis éthiques étroitement liés à la nature fondamentale de ces données, ainsi qu'aux modalités de leur collecte, de stockage et d'utilisation.

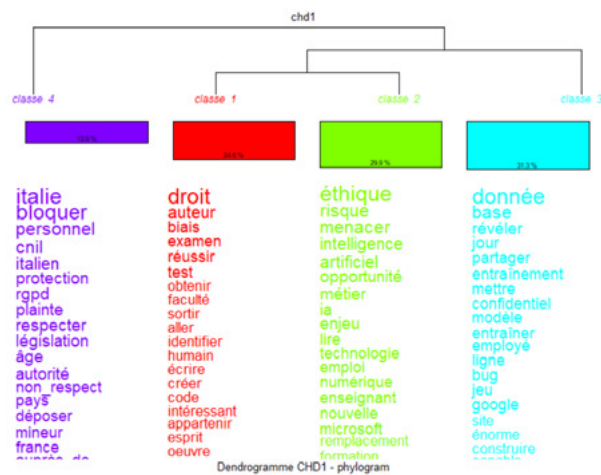


Figure 1. Classification par dendrogramme du corpus français par Iramuteq

Le deuxième cluster le plus important en volume est le cluster 2 (29,9%), il fait directement référence aux risques liées à l'utilisation de l'IA. Les mots clés associés à ce cluster comprennent les termes « éthique » (chi2 169,27), « risqué » (81,81), et « menacer » (80,67). Ce cluster renvoie plus précisément aux chatbots utilisés dans le domaine professionnel. En effet, d'autres items notables, tels que « métier » (chi2 52,95), « emploi » (23,95) et « enseignant » (22,2), font également écho à l'inquiétude que les IA puissent un jour remplacer l'humain dans le domaine scolaire. Ce deuxième cluster est lié au cluster 1 (24,8%) qui se distingue néanmoins par l'emploi d'un vocabulaire évoquant les risques liés aux biais et au plagiat. En effet, ce cluster contient des termes tels que « droit » (chi2 289,51), « auteur » (110,21), et « biais » (52,51). D'autres items tels que « écrire » (20,22), « chercheur » (16,47) et « étudiant » (13,91) rappellent également le cluster 2 et les inquiétudes liées à l'usage de certaines IA en contexte scolaire. Ces items font également écho à l'inquiétude de ne plus pouvoir distinguer les écrits réalisés par ChatGPT de ceux réalisés par un humain. Certes, avec l'avènement de cet outil, le plagiat a pris une forme nouvelle. Des systèmes automatisés peuvent désormais générer du contenu à la place de véritables auteurs, mettant en question la notion même de plagiat. Dans ce sens, un nouveau mot vient d'être créé, *algjarism* ou *AI-assisted plagiarism*, pour définir le plagiat par chatbot. Ainsi, certains acteurs éducatifs ont déjà interdit à leurs élèves le recours à ChatGPT⁹.

Enfin, le quatrième cluster (13,9 %) qui est le plus isolé renvoie à la situation italienne et plus précisément à la décision du gouvernement italien de bloquer l'usage de ChatGPT sur le territoire. Les mots clés correspondant respectivement aux deux valeurs maximales d'association au sein de ce cluster sont « Italie » (chi2 547,95), et « bloquer » (441,83). De fait, le 31 mars 2023, l'autorité de protection des données italienne (*Garante per la*

...

9. ChatGPT va-t-il faire les devoirs de nos enfants ? - <https://numeriqueethique.fr/ressources/articles/chat-gpt-devoirs-enfants>

protezione dei dati personali) a suspendu l'utilisation de ChatGPT en Italie, et cela jusqu'à ce que l'outil soit conforme à la réglementation en matière de protection de la vie privée. Les items « RGPD »¹⁰ (chi2 159,5), « plainte » (154,82), « respecter » (148,06) et « européen » (75,03) suggèrent que ce cluster évoque non seulement les dispositions prises par le gouvernement italien mais également par l'Europe de manière plus générale. D'autres items notables, tels que « CNIL » (chi2 285,4), et « France » (84,59) font également écho aux thèmes de l'autorité et des décisions relatives aux grands principes qui devraient être appliqués à l'IA en France.

Données personnelles et privacy

La procédure de clustering du corpus italien a également conduit à l'identification de quatre clusters (Figure 2) regroupés en deux classes jumelles constituées respectivement des clusters 2 et 4 et des clusters 1 et 3.

Le cluster 1 est le plus important en termes de volume (40,4%). Il identifie une zone sémantique dont les mots-clés sont « privacy » (chi2 194,99), « Italia » (185,83), et « garante privacy » (125,51). L'item « garante privacy », autrement dit le garant de la protection des données personnelles, souligne une attention particulière à l'éthique. Rappelons à nouveau que « *d'importants enjeux éthiques liés au développement de l'IA ont trait à la gouvernance des données* » (Martineau, 2023, p. 61), ainsi qu'à la façon dont elles sont collectées, utilisées et partagées. En complément, les items « bloccare » (chi2 108,94), « blocco » (37,82) et « chiudere » (37,61) font référence à la décision de suspendre l'utilisation de ChatGPT.

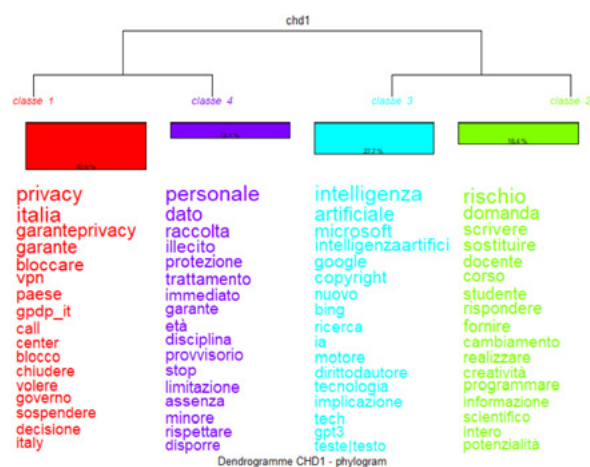


Figure 2. Classification par dendrogramme du corpus italien par Iramuteq

Le deuxième cluster le plus important en taille est le cluster 3 (27,7%) faisant référence au champ lexical des nouvelles techniques et des moteurs de recherche - thème qui semble plus négligé dans le corpus français. Les items les plus caractéristiques sont : « Microsoft » (chi2 86,57), suivi de « Google » (74,69), et « Bing » (53,11), tous faisant référence à la technologie et l'innovation.

Le cluster 2 (18,4%) identifie un monde lexical qui fait référence aux usages spécifiques de ChatGPT en contexte scolaire et au possible risque que cet outil puisse remplacer les enseignants. En particulier, les items tels que « rischio » (chi2 62,3), « scrivere » (37,7),

...

10. Le Règlement général sur la protection des données (RGPD) entré en vigueur dans l'Union européenne (UE) en 2018, est un texte réglementaire qui encadre le traitement des données de manière égalitaire sur tout le territoire de l'Union européenne.

« sostituire » (36,79) et « docente » (35,68) rappellent cette thématique. Ce cluster est similaire, tant par sa taille que par son objet au cluster 1 français qui faisait également mention des risques liées à l'usage de certaines IA en contexte scolaire, notamment celles spécialisées dans le traitement de texte.

Le cluster 4 (13,1%) correspond au champ lexical de la protection des données personnelles. En particulier, les items tels que « dato » (χ^2 526,79), « raccolta » (444,99) et « trattamento » (243,61) rappellent cette thématique. Les tweets analysés à l'intérieur de ce cluster concernaient essentiellement le traitement de données personnelles en l'absence de consentement préalable des personnes concernées ou encore l'absence de vérification de l'âge des utilisateurs.

Au terme de cette analyse et contrairement au corpus français, nous identifions ici deux macro-dimensions argumentatives bien distinctes :

- Une première dimension éthique concerne les aspects liés à la protection des données personnelles (cluster 1 et 4)
- Une seconde dimension concerne les usages des IA, notamment en contexte scolaire (cluster 2 et 3)

En définitive, le corpus italien montre donc davantage les enjeux et risques en matière de données personnelles que les risques de discrimination algorithmique, de plagiat ou de biais retrouvés dans le corpus français. Il s'agit ici d'un élément de différenciation, qui suggère que malgré la présence d'un certain questionnement éthique dans les deux pays étudiés, ceux-ci se focalisent sur des problèmes très différents. Les imaginaires et les enjeux éthiques sont donc propres à chaque contexte et varient selon la culture et les pays.

CONCLUSION

Nous avons exploré les questionnements éthiques liés à ChatGPT. Notre objectif n'était pas de définir ce qu'est l'éthique de l'IA en général, ni de juger le caractère éthique ou non de certains systèmes d'IA. Nous avons voulu analyser les diverses manières, souvent variées, d'appréhender le questionnement éthique. Notre objectif était de mettre en évidence que la façon dont les internautes répondent aux questions d'éthique est conditionnée par le contexte, notamment politique, culturel ou social.

Bien que nous ne puissions guère interpréter les *hashtags* comme des identifiants directs des sujets ou des thématiques, l'analyse démontre une importante diversité du côté de la production de messages. Comme le souligne Longhi (2020), l'analyse thématique « *permet de voir dans quelles configurations les mots sont utilisés, et ainsi percevoir leur orientation discursive, en les rattachant aux discours dans lesquels ils s'intègrent* » (p.20).

Notre analyse confirme que l'éthique est « *étroitement imbriquée dans le monde social. Elle entre en relation avec les groupes sociaux qui le composent, produisant une co-construction entre éthique et rapports sociaux* » (Marques, 2022 p.21). Le discours sur l'éthique de ChatGPT mobilise des répertoires ancrés tant sur le plan national (références à des normes et décisions propres à chaque pays) que sur le plan européen (référence à la législation européenne). Certains thèmes génèrent plus de réactions que d'autres et les questions dominantes sont sujettes à des réinterprétations propres au contexte national.

Un premier élément de diversification important entre les deux corpus étudiés dans cet article est représenté par le nombre de tweets collectées (33 195 tweets pour le corpus français et 9 517 pour le corpus italien). Cette grande différence entre les deux corpus s'explique en partie par le nombre d'utilisateur qui diffère dans les deux pays. Mais elle

démontre aussi qu'en Italie, les questions éthiques liées à ChatGPT ne représentent pas un sujet très débattu malgré la décision de l'autorité de protection des données italienne (la *Garante per la protezione dei dati personali*) d'en suspendre l'utilisation. Une attention qui semble bien plus présente pour les internautes français, qui d'ailleurs font référence à la situation italienne (voir cluster 4).

Un autre élément de diversification est représenté par les thèmes émergeant des deux principaux clusters : en France, nous constatons une attention particulière aux biais et aux plagiat. Le corpus italien, quant à lui, reflète davantage une attention aux questions liées à la protection des données personnelles. Sur le plan comparatif, il est également intéressant de souligner que si deux des trois clusters italiens renvoient directement à des controverses actuelles, notamment la violation de la réglementation en matière de données personnelles, le corpus français reflète plutôt une inquiétude générale et met prioritairement l'accent sur les conséquences à long terme du développement de l'IA, en soulignant notamment les enjeux et les risques associés à l'usage de ChatGPT en milieu scolaire.

L'IA, en particulier les outils comme ChatGPT, suscite des débats complexes et variés, orientés par les contextes nationaux. Alors que certains mettent en avant des préoccupations immédiates comme la protection des données, d'autres se concentrent sur les impacts à long terme, notamment dans des domaines sensibles comme l'éducation. Dans ce cadre, Marques (2022) met en garde contre une approche trop globalisée de l'éthique de l'IA, soulignant que « celle-ci devrait être appréhendée de manière à travailler avec les cultures, tout en s'interrogeant sur la façon dont localement, les standards peuvent être appliqués en accord avec les valeurs du pays » (p. 19).

Le défi aujourd'hui consiste donc à établir un équilibre entre la promotion du déploiement de ces innovations techniques, tout en veillant à ce qu'une réglementation adaptée soit mise en place pour encadrer leur utilisation, en prenant en compte les spécificités nationales et les enjeux éthiques internationaux.

RÉFÉRENCES BIBLIOGRAPHIQUES

Andler, Daniel (2023), *Intelligence artificielle, intelligence humaine : la double énigme*, Paris : Gallimard (Collection « NRF Essais »).

Boydjjan, Julien (2016), *Analyser les opinions politiques sur internet. Enjeux théoriques et défis méthodologiques*, Paris : Dalloz.

Boudreau, Marie Claude (2019), *La compétence éthique en milieu de travail : Une perspective pragmatiste pour sa conceptualisation et son opérationnalisation*, thèse de doctorat en philosophie - Université de Sherbrooke.

Bruneault, Frederick ; Sabourin Laflamme, Andreane (2021), « Éthique de l'intelligence artificielle et ubiquité sociale des technologies de l'information et de la communication : comment penser les enjeux éthiques de l'IA dans nos sociétés de l'information ? » *TIC & Société*, vol.15, p. 159-189.

Crepel, Maxime ; Cardon, Dominique (2022), « Robots vs algorithmes. Prophétie et critique dans la représentation médiatique des controverses de l'IA », *Réseaux*, vol. 2, n°232-233, p. 129-167.

Domengot, Jean-Claude ; Wilhelm Carsten ; Arruabarrena, Béatrice ; Alloing Camille, (2022), « Introduction », *Revue française des sciences de l'information et de la communication*, [en ligne], consulté le 01 juin 2024, <http://journals.openedition.org/rfsic/13164>.

- Domenget, Jean-Claude ; Roelens, Camille (2024), « Éthique et numérique au XXIème siècle. Approches compréhensives, normatives et critiques. Regards interdisciplinaires », *Interfaces numériques*, vol. 13-1, p. 1-11.
- Farjat, Gérard (2004), *Pour un droit économique*, Paris : PUF (collection « Les voies du droit »).
- Franzke Shakti, Aline ; Bechmann, Anja ; Zimmer, Michael ; M. Ess, Charles, (2020), *Internet Research : Ethical Guidelines 3.0*, [en ligne], consulté le 07 juin 2024, <https://aoir.org/reports/ethics3.pdf>.
- Garzonio, Emma ; Nuvoli, Karen (2022), « Covid-19, la «dittatura sanitaria». Uno studio comparativo della narrazione antivaccinista in Italia e Francia », *Problemi dell'informazione, Rivista quadrimestrale*, n°3, p. 383-405.
- Gozalo-Brizuela, Roberto ; Garrido-Merchan, Eduardo C. (2023), *ChatGPT is not all you need. A State of the Art, Review of large Generative AI models* [en ligne], consulté le 25 juillet 2024, <https://arxiv.org/abs/2301.04655#>.
- Hu, Krystal (2023), *ChatGPT sets record for fastest-growing user base – analyst note*, [en ligne], consulté le 05 juin 2024, <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>.
- Lacroix André ; Marchildon, Allison ; Bégin, Luc (2017), *Former à l'éthique en organisation, Québec*, Québec : Presses de l'Université du Québec.
- Linden, Isabelle (2020), « Entre rêves et illusions...L'intelligence artificielle en question », *Revue d'éthique et de théologie morale*, vol. 307, n. 3, p. 11-27.
- Longhi, Julien (2020), « Explorer des corpus de tweets : du traitement informatique à l'analyse discursive complexe », *Corpus*, vol. 20 [en ligne], consulté le 20 juin 2024 <http://journals.openedition.org/corpus/4567>.
- Marques, Julie (2022), « L'Intelligence Artificielle, une approche intersectionnelle », *Interfaces numériques*, vol. 11, [en ligne], consulté le 20 juin 2024 <https://doi.org/10.25965/interfaces-numeriques.4796>
- Martineau, Joé T. (2023), « Transition numérique et intelligence artificielle : d'importants enjeux éthiques à surveiller », *Gestion*, vol. 48, p. 60-64.
- Meneceur, Yannick (2022), Analyse des principaux cadres supranationaux de régulation des applications de l'intelligence artificielle. Des éthiques de l'intelligence artificielle à la conformité ? In Mendoza-Caminade, Alexandra (dir.), *L'entreprise et l'intelligence artificielle - Les réponses du droit*, Toulouse : Presses de l'Université Toulouse Capitole.
- Ménissier, Thierry (2022), « Jusqu'où l'institution peut-elle être augmentée ? Pour une éthique publique de l'IA », *Quaderni*, [en ligne], consulté le 05 juin 2024, <http://journals.openedition.org/quaderni/2234>.
- Mercier, Arnaud (2021), « Les modalités de la colère citoyenne sur Twitter », *Quaderni* [en ligne], consulté le 05 juin 2024, <http://journals.openedition.org/quaderni/2134>.
- McCarthy, John ; Minsky, Marvin L. ; Rochester, Nathaniel ; Shannon, Claude E. (2006), «A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence», *AI Magazine*, vol. 27, p. 12-14.
- Neuman, Russel W. ; Guggenheim, Lauren, *et al.* (2014), « The Dynamics of Public Attention: Agenda-Setting Theory meets Big Data », *Journal of Communication*, vol. 64, p. 193-214.
- Nocetti, Julien (2019), *Rapport Intelligence artificielle et politique internationale*, IFRI, [en ligne], consulté le 05 juin 2024, <https://www.ifri.org/fr/etudes/intelligence-artificielle-et-politique-internationale-les-impacts-dune-rupture-technologique>

Ratinaud, Pierre, *et al.* (2019), « Structuration des discours au sein de Twitter durant l'élection présidentielle française de 2017. Entre agenda politique et représentations sociales », *Réseaux*, vol. 2, n. 214-215, p. 171-208.

Reinert, Max (1997), *Les « mondes lexicaux » et leur « logique » à travers l'analyse statistique de divers corpus*, [en ligne], consulté le 10 juin 2024 <http://lexicometrica.univ-paris3.fr/jadt/jadt1998/reinert.htm>.

Severo, Marta ; Lamarche-Perrin Robin (2018), « L'analyse des opinions politiques sur Twitter. Défis et opportunités d'une approche multi-échelle », *Revue française de sociologie*, vol. 3, n. 59, p. 507-532.

Turing, Alan Mathison (1950), « Computing machinery and intelligence », *Mind*, vol. 59, p. 433-460.

Tsamados, Andreas ; Aggarwal, Nikita *et al.* (2021), « The ethics of algorithms : key problems and solutions », *Ethics, Governance, and Policies in Artificial Intelligence*, Springer International Publishing, p. 97-123.

Vaswani Ashish, *et al.* (2017), « Attention is all you need », *Advances in Neural Information Processing Systems 30 (NIPS 2017)*, [en ligne], consulté le 10 juin 2024, https://papers.nips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html.

Villani, Cédric (2018), Donner un sens à l'intelligence artificielle. Pour une stratégie nationale et européenne, Mission parlementaire du 8 septembre 2017 au 8 mars 2018, [en ligne], consulté le 15 juin 2024, <https://www.enseignementsup-recherche.gouv.fr/fr/rapport-de-cedric-villani-donner-un-sens-l-intelligence-artificielle-ia-49194>.

Volpe, Valeria ; Zaza, Ornella (2024), « Les arrangements de l'innovation numérique dans les territoires ruraux. Une comparaison France-Italie », *Espaces et sociétés*, vol. 191, p. 31-47.