

L'IA appliquée à la sphère électorale et considérations éthiques #3

En 2024, des rapports d'enquête ont révélé l'existence de [réseaux sur Telegram](#), dans la région des Balkans, au travers desquels des centaines de milliers d'utilisateurs faisaient circuler un logiciel manipulé par l'IA, permettant de « dénuder » des femmes photographiées. Ces photos modifiées numériquement se sont transformées en outils de chantage, de coercition et d'humiliation publique, forçant ainsi au silence les femmes ciblées et les mettant au ban de la vie publique. Si l'IA générative est devenue une fonctionnalité incontournable sur les médias sociaux, [plus de la moitié des contenus en ligne étant produits par l'IA](#), une dure réalité semble absente des débats : il est estimé en effet que [96 % des « deepfakes »](#) représentent des images intimes de femmes diffusées à leur insu. Ces violences sexistes facilitées par la technologie sont particulièrement alarmantes dans le contexte électoral, car ces attaques personnelles non seulement accentuent les stéréotypes néfastes fondés sur le genre, mais maintiennent également les femmes à l'écart de la sphère publique et politique, ou les dissuadent d'y entrer.

L'affaire Telegram débouche sur un constat préoccupant : sans garanties suffisantes pour prévenir les abus des systèmes d'IA, les femmes et d'autres groupes marginalisés peuvent subir de graves préjudices. La violence fondée sur le genre en ligne existait bien avant que l'IA ne devienne un outil aussi accessible ; toutefois, cette technologie peut accélérer et amplifier la propagation des discours de haine et des images à caractère sexuel non consenties, entraînant des atteintes graves aux droits humains fondamentaux. Les risques posés par l'IA ne sont pas uniquement liés à son utilisation délibérément abusive ; ils sont également intrinsèques à la structure même de ces systèmes et à la manière dont ils sont construits et déployés. Au-delà des cas comme celui évoqué plus haut concernant la plateforme Telegram, où l'IA a été utilisée comme une arme pour cibler des personnes en particulier, il faut prendre conscience des atteintes et dommages, parfois involontaires, que l'IA peut provoquer.

La maîtrise de l'IA doit donc aboutir à une meilleure compréhension de la manière dont les biais systématiques, intégrés aux ensembles de données utilisés pour l'entraîner, contribuent à perpétuer la discrimination et l'exclusion, souvent dans l'ignorance des utilisateurs. Ces biais peuvent accentuer certaines normes de genre, les stéréotypes

raciaux et d'autres préjugés, et découler précisément des décisions prises au moment de la conception du système, de la collecte et de l'étiquetage des données, ou de la manière dont les modèles sont développés, appliqués et mis à disposition. Ces tendances peuvent conduire à accroître le risque de préjudice, fondé sur la discrimination, l'exclusion, sans même que l'utilisateur en ait conscience. Qu'il s'agisse d'abus délibérés ou de failles involontaires, l'IA exige une approche éthique urgente et des mécanismes clairs de redevabilité à chaque niveau.

Les grands ensembles de données utilisés pour former les modèles d'IA nécessitent également une attention particulière quant à la manière dont elles sont collectées et stockées, en particulier celles sensibles relatives à l'identité des électeurs, utilisées notamment pour l'identification biométrique, la reconnaissance faciale et la surveillance électorale. La protection de ces données est essentielle pour garantir de manière éthique la vie privée des électeurs et préserver la confiance du public dans l'administration électorale, composante clé de toute démocratie saine. Cela dit, ces protections sont souvent coûteuses et requièrent le déploiement de moyens importants de la part des organismes de gestion électorale (OGE). Ces derniers doivent donc mettre en place une infrastructure de cybersécurité adéquate pour protéger les données sensibles des électeurs et garantir la transparence sur la nature et le mode de traitement des données.

Les considérations éthiques sont essentielles à chaque phase de développement d'une technologie fondée sur l'IA, depuis la conception initiale jusqu'au déploiement, en passant par le suivi et l'évaluation continus du système. Les décisions prises à chacun de ces stades ont des conséquences itératives, car les négligences ou la non prise en compte des valeurs démocratiques peuvent affecter les communautés minoritaires et conduire à la violation de leurs droits civiques et politiques fondamentaux. Dans la sphère électorale, la manière dont l'IA est déployée a des effets directs sur sa dimension éthique. L'IA appliquée à la sphère électorale doit tenir compte des réalités culturelles, politiques et sociales spécifiques afin de ne pas accentuer les inégalités existantes. Pour atténuer ces risques, les systèmes d'IA doivent être transparents et mettre à la disposition des électeurs des mécanismes clairs de mise en cause de la responsabilité afin de garantir leur équité. Une surveillance continue est également essentielle pour détecter et corriger les déviations éthiques, ce qui suppose une solide alphabétisation numérique et d'importantes capacités techniques au niveau des OGE.

L'IA pouvant agir comme un amplificateur des problèmes systémiques préexistants, le poids de ces préjudices pèse souvent sur les personnes déjà marginalisées, notamment les minorités ethniques et religieuses, les femmes et les personnes de statut socio-économique inférieur.

Sans une surveillance humaine stricte, de tels dommages, souvent involontaires, passent inaperçus ; pourtant la menace pour les droits humains est indéniable. Lorsque l'IA est employée dans n'importe quel aspect relatif aux processus électoraux, notamment l'[organisation des élections](#), il est fondamental de prêter attention à ces risques.

Le projet « [IA pour les acteurs électoraux](#) » d'International IDEA conduit à reconnaître que les considérations éthiques et relatives aux droits humains doivent être au cœur de toute initiative impliquant le recours à l'IA dans le cadre électoral. Conformément à ce principe, l'un des piliers clés de l'atelier – et le deuxième présenté dans cette série d'articles – est consacré à l'IA, la dimension éthique et les droits de la personne. (Découvrez les [cinq piliers](#) qui constituent le fondement des principes entourant l'IA dans les élections et permettant d'aborder ses implications sur les plans démocratique, technique, juridique et éthique.)

Cet article s'appuie sur les discussions échangées au cours du deuxième atelier, organisé à Tirana, en Albanie, et qui a réuni des représentants d'organismes de gestion des élections et d'organisations de la société civile de la région des Balkans et d'Europe de l'Est afin d'examiner ensemble les fondements éthiques nécessaires pour garantir un alignement de l'utilisation de l'IA dans le cadre électoral sur les valeurs démocratiques.

Ces échanges ont permis de démêler la question complexe de la protection des membres des communautés marginalisées pour éviter qu'ils soient persécutés ou privés de leurs droits civiques, en introduisant des garanties éthiques.

Pilier n° 2, IA, éthique et droits de la personne

Les OGE considèrent de plus en plus l'adoption des outils d'IA pour améliorer l'intégrité des élections et l'accomplissement de fonctions telles que la mise à jour des listes électorales (en évitant la duplication des électeurs), l'optimisation de l'emplacement des bureaux de vote, l'authentification des électeurs ou la détection des fraudes. Toutefois, ces outils peuvent également être utilisés par d'autres acteurs dans le cadre électoral, par exemple lors des campagnes menées par les partis politiques. Chacune de ces applications a des implications importantes sur le plan éthique et pour les droits humains, notamment au regard de la confidentialité, de la sécurité et de la transparence, ainsi que de la durabilité – l'IA consommant d'énormes quantités d'énergie. De plus, les régions qui ne bénéficient pas d'un accès égal aux infrastructures publiques numériques ou d'une culture numérique suffisante sont confrontées à des défis supplémentaires pour garantir un déploiement juste et équitable de l'IA. Le renforcement des connaissances en la matière, notamment parmi les acteurs électoraux, est donc aujourd'hui une priorité essentielle.

Lors de l'atelier organisé à Tirana, les représentants des OGE ont jugé leurs connaissances dans le domaine de l'IA relativement faibles, ce qui a suscité des inquiétudes quant à leur capacité à éviter les atteintes éventuelles aux droits des personnes. Mais en dépit de ce constat, certains représentants ont indiqué qu'ils utilisaient déjà l'IA pour surveiller les réseaux sociaux et fournir des informations sur l'administration électorale par le biais de chatbots.

Les chatbots alimentés par l'IA ont parfois tendance à générer des informations inexactes, également appelées hallucinations, car leurs réponses comportent une part aléatoire. Des hallucinations ont déjà été constatées lors des [élections européennes de 2024](#), lorsque les plus grands chatbots du monde ont diffusé des informations erronées sur les dates des élections, les modalités de vote et les personnes habilitées à voter. Sans limitations intégrées au système, sans réglementation et sans contrôle humain, les systèmes basés sur l'IA risquent non seulement de saper la confiance dans les institutions électorales, mais aussi de priver les électeurs de leur droit.

Les hallucinations mettent en évidence un problème éthique plus général lié à l'utilisation de l'IA dans la sphère publique : le manque de transparence. Les décisions prises par les algorithmes sont souvent inexplicables, même pour les développeurs qui en sont à l'origine, un phénomène connu sous le nom de « [problème de la boîte noire](#) ». Étant donné qu'il existe peu d'informations sur la manière dont un système d'IA produit certains résultats, il peut être difficile d'expliquer exactement les raisons pour lesquelles le système parvient à certaines conclusions et de quelle façon. Dans le cas d'un chatbot électoral, le système peut par exemple inventer de faux bureaux de vote dans des zones pour lesquelles il n'existe pas suffisamment de données pertinentes, créant ainsi un obstacle systématique pour l'électorat vivant dans ces zones. Ces problèmes sont particulièrement susceptibles d'affecter les communautés au demeurant déjà défavorisées, où le manque de données est assez fréquent, aggravant ainsi les injustices sociales préexistantes. Les inexactitudes ou fausses informations qui filtrent dans les réponses apportées par l'IA sont difficiles à prévoir, mais elles sont une atteinte au droit inaliénable à l'information, qui est primordial pour prendre des décisions de vote libres et éclairées.

Le [manque de confiance](#) dans l'IA a été identifié comme un obstacle majeur à la volonté de certains de mettre l'IA au service de l'ensemble des secteurs. Cette méfiance généralisée ne permet pas de tirer parti de tous les avantages offerts par l'IA pour améliorer l'efficacité et l'équité des processus électoraux. Pour que l'IA permette de bâtir une base démocratique solide, la conscience éthique et la reddition de comptes doivent être intégrées à tous les niveaux – les entreprises qui les développent, celles qui les déploient et utilisateurs finaux.

Cela est particulièrement important dans le domaine électoral, où les questions relatives aux droits de l'homme et les normes éthiques doivent rester au premier rang des priorités. Les responsables des organes électoraux doivent donc faire preuve de prudence lorsqu'ils sont contactés par les fournisseurs d'IA, et assurer un contrôle humain réel à chaque étape du déploiement de la technologie, en prenant des mesures adéquates pour identifier les risques possibles et y remédier.

Il est certain que le recours aux outils de l'IA a jusqu'à présent été marginal dans la région des Balkans occidentaux et en Europe de l'Est et il n'existe donc pas de cas documenté de biais induits par l'IA dans les processus électoraux. Néanmoins, l'histoire de la région, marquée par des tensions ethniques, une instabilité politique et la manipulation des médias, laisse supposer des risques liés à l'introduction de l'IA sans une gestion prudente, au nombre desquels, par exemple, l'amplification des biais ethniques, raciaux, géographiques ou médiatiques si la technologie n'est pas utilisée de manière réfléchie.

Des préoccupations en ce sens sont apparues en Bosnie-Herzégovine, où l'affaire [Pilav c. Bosnie-Herzégovine \(2016\)](#) a servi à montrer que le système électoral du pays était discriminatoire en fonction de l'origine ethnique et du lieu de résidence. En limitant l'accès à des fonctions politiques à trois « peuples constitutifs » – les Bosniaques, les Croates et les Serbes –, le système avait exclu les personnes n'appartenant pas à ces ethnies ou résidant dans certaines régions. Ces biais fondés sur l'origine ethnique ou géographique illustrent le fait que des règles a priori neutres peuvent entraîner des résultats faussés dans la pratique.

Des discriminations liées à l'âge ont également été signalées durant le processus électoral en Albanie. Nos experts et les participants ont fait observer que les électeurs âgés avaient du mal à utiliser le système de vote électronique, nécessitant souvent une assistance pouvant compromettre la confidentialité du votant. Un tel exemple montre que les opérations de scrutin basées sur l'IA ou la technologie peuvent présenter des obstacles à la participation si elles ne sont pas pensées et mises en œuvre selon une approche inclusive.

Pris ensemble, ces exemples – bien entendu non spécifiques à cette région – soulignent à quel point les biais de l'IA peuvent exacerber les clivages sociaux et les problèmes qui nuisent à l'équité des élections. Ils mettent en évidence la nécessité d'une approche prudente et réfléchie pour intégrer l'IA aux processus électoraux. Même si cette nouvelle technologie offre des avantages évidents, en contribuant par exemple à la simplification administrative ou à la réduction des erreurs humaines, il demeure indispensable de prévoir des garanties solides et un contrôle transparent.

Les développeurs d'IA et les autorités électorales ne doivent pas perdre de vue les considérations éthiques, doivent consulter diverses parties prenantes et surveiller rigoureusement la mise en œuvre afin de garantir que l'IA renforce plutôt qu'elle ne fausse l'équité en matière électorale et les règles démocratiques.

Le quatrième volet de cette série d'articles permet d'explorer le troisième pilier fondamental pour une IA au service de la démocratie : à quoi sert concrètement l'IA dans la sphère électorale et quelles considérations doivent être prises en compte pour sa mise en œuvre responsable et démocratique ? La discussion s'appuiera sur les enseignements tirés de notre troisième atelier « IA pour les acteurs électoraux » organisé à Johannesburg, la première semaine d'avril 2025.

Ne perdez aucun article de la série et lisez aussi le [premier article](#) publié pour avoir une vue d'ensemble des fondements nécessaires pour une IA démocratique et le [deuxième article](#) sur la maîtrise de l'IA et le premier atelier qui s'est déroulé à Kuala Lumpur.