

L'appel d'une centaine d'ONG : « L'IA telle qu'elle est développée alimente un système d'exploitation global »

Tribune

Un collectif, parmi lesquelles Amnesty International et la Ligue des droits de l'homme, recommande, dans une tribune au « Monde », de placer les droits humains et la justice environnementale au cœur de la régulation de l'intelligence artificielle.

L'intelligence artificielle (IA) connaît un développement foudroyant, et nos dirigeants ne semblent pas pressés de réfléchir aux enjeux humains, sociaux et environnementaux de ces nouvelles technologies, uniquement vues sous le prisme de la croissance, des gains de productivité et des profits. L'IA telle qu'elle est développée perpétue cependant les discriminations, aggrave les inégalités, détruit la planète et alimente un système d'exploitation global. Parce que ces constats ne figureront pas au programme officiel du Sommet mondial sur l'IA [qui se tient à Paris les 10 et 11 février], nous, organisations de la société civile, vous les rappelons ici.

Lire aussi l'Article réservé à nos abonnés Jusqu'où ira l'intelligence artificielle ?

Se concentrer uniquement sur d'éventuels futurs risques existentiels à venir de l'IA est un leurre : ces technologies ont déjà des effets très concrets pour les populations les plus vulnérables et les plus discriminées et portent largement atteinte aux droits humains. En s'appuyant sur des bases de données biaisées et en intégrant les préjugés de ses concepteurs, l'IA perpétue les stéréotypes, renforce les inégalités sociales et limite l'accès aux ressources et opportunités. A cela s'ajoute le fait que le déploiement de ces systèmes d'IA s'inscrit dans le contexte des structures discriminatoires et inégalitaires qui existent dans les sociétés du monde entier. Le recours à ces technologies, souvent sur fond de politiques d'austérité, amplifie les discriminations dans l'accès à la santé, à l'emploi, aux services publics ou aux prestations sociales. En témoignent les scandales ayant éclaté ces dernières années : biais sexistes et racistes des algorithmes de santé, algorithme des services de l'emploi autrichien qui refuse d'orienter les femmes vers le secteur informatique, profilage et discrimination des usagers de la Caisse nationale des allocations familiales en France, au Danemark ou aux Pays-Bas.

Or les technologies sont rarement la solution à des problèmes en réalité systémiques. Il est préférable de s'attaquer à la racine de ces problèmes plutôt que de prendre le risque d'aggraver les violations des droits humains avec des systèmes d'IA. Tandis que l'on confie de plus en plus de décisions aux algorithmes, leurs biais peuvent avoir des conséquences dramatiques sur nos vies. Les IA prédictives se substituent à la justice et à la police, risquant d'amplifier le racisme systémique. Par exemple, aux Etats-Unis, une IA calculant les risques de récidive désignait deux fois plus les accusés noirs comme étant « à haut risque » que les accusés blancs. Et quand bien même on réduirait ces biais, se concentrer sur les outils prédictifs nous empêche de penser à des réformes plus globales du système carcéral.

Menaces pour l'Etat de droit

Ces systèmes sont aussi utilisés à des fins de surveillance et d'identification dans le cadre du contrôle des frontières ou de conflits, comme Lavender, cette IA qui, en désignant des cibles terroristes, a provoqué la mort de milliers de civils gazaouis. Et bien souvent, ces technologies sont développées par les pays occidentaux, comme les outils créés par des pays européens utilisés pour surveiller la population ouïghoure en Chine.

Les systèmes d'IA générative sont également instrumentalisés à des fins de désinformation et de déstabilisation par des régimes répressifs et des acteurs privés. « Bots » utilisés pour manipuler l'information sur des questions liées à la santé, désinformation à caractère raciste durant les dernières élections européennes, deepfakes audios et vidéo mettant en scène des candidats aux

élections : ces technologies sont autant de menaces pour l'Etat de droit. Les montages crédibles générés par IA sont aussi un danger pour les femmes et les enfants : 96 % de ces deepfakes sont des contenus non consentis à caractère sexuel [selon le rapport 2019 du cabinet de conseil en gestion de risques DeepTrace], massivement utilisés dans le but de nuire aux femmes et de générer des contenus pédocriminels.

Par ailleurs, ces effets s'inscrivent dans un système d'exploitation global. L'IA, et notamment l'IA générative, constitue un véritable désastre pour l'environnement. D'ici à 2027, l'IA générative nécessitera une alimentation en électricité équivalente à celle de pays comme l'Argentine ou les Pays-Bas [comme le rapporte un article du New York Times d'octobre 2023]. Les émissions de CO2 des « géants de la tech » ont augmenté de 30 à 50 % en 2024 en raison du développement fulgurant de ces technologies. Et ce sont les pays du Sud global qui sont les premiers touchés : les data centers y pullulent, et l'extraction de minerais, comme le cobalt, utilisé entre autres dans les batteries, met en péril la santé des populations, entraîne la pollution des eaux et des terres et alimente violences et conflits armés.

L'affaire de toutes et tous

Les inégalités entre les pays du Nord et du Sud sont également aggravées par les technologies déployées pour la modération de contenus en ligne. Les géants du numérique qui allouent plus de moyens aux pays du Nord privilégient ainsi certaines langues et récits culturels, déjà dominants, au détriment des autres. Enfin, n'oublions pas que ces systèmes d'IA sont majoritairement entraînés par des travailleurs et travailleuses du Sud global, exploités et sous-payés. Selon les informations du magazine Time, la société OpenAI a ainsi rémunéré des Kényans moins de deux dollars (1,95 euro) de l'heure pour labelliser des contenus toxiques, un travail particulièrement violent et éprouvant.

Face à ces constats alarmants, le règlement européen sur l'IA, présenté comme un instrument de protection des droits et libertés, reste très imparfait, notamment sur les questions de surveillance et de police prédictive. Par ailleurs ce règlement ne s'appliquera pas hors des frontières de l'Union européenne, alors même que la menace pour les droits humains et l'environnement est globale et que l'exportation des IA de surveillance génère du profit pour les entreprises européennes.

Nos gouvernements ne cessent de parler de souveraineté de l'IA, mais les défis posés par ces systèmes transcendent les frontières. Loin d'être un sujet technologique, l'IA est l'affaire de toutes et tous. Tout le monde doit pouvoir choisir la direction de ses développements, quitte à les refuser s'ils ne correspondent pas à notre projet de société. Un cadre contraignant élaboré démocratiquement, dans une perspective de solidarité internationale et avec les communautés les plus touchées, qui place les droits humains et la justice environnementale au cœur de la régulation de l'IA, voilà le véritable progrès.