

Les enjeux civiques et démocratiques posés par les IA génératives : possible court-circuit ?



Karine ROUSSEAU

Chercheur associé

Institut des Etudes de Géopolitique Appliquée

Dans son roman « Les Furtifs » paru en 2019, l'écrivain Alain Damasio imagine une société française en 2040 où les individus, hyper-contrôlés, aliénés à des objets connectés, vivent, en partie, dans des villes privatisées par de grandes entreprises multinationales, profitant d'Etats défaillants. Les impôts sont remplacés par des forfaits différenciés avec des zones urbaines privilégiées, créant des citoyennetés à géométrie variable et repliées dans leurs « technococons », avec leur consentement. Dans le scénario de la Red Team, équipe d'écrivains de SF au sein de l'Agence de l'Innovation de Défense, intitulé « Chronique d'une mort culturelle annoncée », les populations européennes sont enfermées dans des univers de réalité virtuelle (les « Safe sphères »), grâce à des implants oculaires neuronaux, déconnectées de la

réalité, influencées par de fausses informations et devenues méfiantes à l'égard des informations officielles du gouvernement et des messages de l'armée organisant l'évacuation lors d'une attaque bioterroriste. Ces deux récits d'anticipation soulèvent pourtant des questions bien réelles en tant que citoyens au sein de démocraties libérales. Quels sont les enjeux démocratiques et civiques posés par les technologies numériques, en particulier les IA génératives auxquelles elles sont de plus en plus adossées, récemment popularisées par des robots conversationnels tels que ChatGPT ? Il s'agira ici de se concentrer sur les risques générés par les IA pour nos démocraties, même si les technologies numériques peuvent être aussi un atout pour les citoyens et même pour les décideurs politiques. Mais c'est un autre sujet...

Des IA suscitant crainte et fantasmes

Les technologies numériques et les IA, font partie intégrante de nos vies quotidiennes depuis plusieurs décennies sans véritable interrogation citoyenne. Cependant, l'actualité des robots conversationnels a changé la donne. En 2020, le Forum pour l'Information et la Démocratie à New York a proposé la création d'un Observatoire de l'Information et de la Démocratie sur le modèle du GIEC, en réponse au danger représenté par la « marchandisation de l'information par les géants du numérique, affaiblissant ainsi les démocraties »¹. Cette préoccupation a été réaffirmée lors du forum de 2022, notamment par Reporter Sans Frontières.

¹ Marianne, « A New York, vers un GIEC de l'information et de la démocratie contre les Gafam ? », 24 septembre 2022

Selon Christophe Deloire, son secrétaire général, ces plateformes privatisées détenues par des acteurs tels que Google ou Facebook détiennent un pouvoir extraordinaire en décidant des informations pouvant circuler. « Ces géants se substituent aux parlements et affirment que les algorithmes ou le « code » font office de loi ». En mars 2023, l'ONG américaine « Future of Life Institute » a lancé une lettre ouverte appelant à une pause dans la recherche sur l'intelligence artificielle, en particulier concernant ChatGPT, en raison des risques de diffusion de fausses informations. Cette lettre était signée, de manière ironique, par des personnalités telles qu'Elon Musk et Steve Wozniak, eux-mêmes sur des projets d'IA. En mai 2023, l'administration Biden a rencontré les grandes entreprises de l'IA telles que Microsoft ou OpenAI. Le 17 mai 2023, à l'occasion de la Journée mondiale de la société de l'information (#WTISD2023), 100 chercheurs du monde entier ont lancé un appel en faveur de la création de cet Observatoire International sur l'Information et la Démocratie, soulignant l'importance d'adapter la technologie à la démocratie plutôt que l'inverse². Mais quelle est la nature de cette crainte ainsi exprimée? S'agit-il simplement d'une « hallucination », pour reprendre un terme « chatgptéiste », ou d'une peur fondée? L'IA est une création humaine, un outil technique qui reflète les valeurs qui lui sont insufflées. Ces IA, programmées et fabriquées par des ingénieurs grâce à des investissements souvent privés de grandes entreprises technologiques (Big Tech), génèrent du contenu en imitant remarquablement l'humain. La Cour suprême des Etats-Unis a récemment réagi en ce qui concerne la responsabilité des plateformes quant aux contenus qu'elles diffusent. Et un « AI Act » européen a également été adopté par le Parlement européen le 14 juin pour réguler les intelligences artificielles

avant la mise sur le marché « avec le triptyque sécurité, santé, droits fondamentaux ». ³ Et, tout récemment en France, des missions d'information ont été organisées par l'Assemblée Nationale sur le thème « Intelligence artificielle et protection des données personnelles ». Perçues longtemps comme le cadre de développement d'une « démocratie augmentée », les technologies numériques, avec les IA, suscitent donc davantage aujourd'hui de questionnements.

Des enjeux liés à l'exercice par le citoyen de sa faculté de penser et de décider par lui-même

L'un des piliers de la démocratie est la liberté, en particulier la liberté d'expression. Et les IA sont au cœur du débat à ce sujet. Tout d'abord, elles génèrent la production massive d'informations à une vitesse hypersonique et parmi elles de fausses informations, manipulées par des acteurs étatiques ou privés, avec des objectifs géopolitiques et politiques : influence en termes de puissance, valorisation de modèles politiques autoritaires. C'est le champ de la guerre informationnelle et de la guerre cognitive, où les récits sont déformés, les opinions manipulées. Ainsi, l'affaire récente des sites miroirs, diffusant des articles pro-russes sur de faux sites de grands médias, l'illustre clairement⁴. Les fausses informations peuvent être produites à grande échelle grâce aux IA génératives. Et la création de fausses images ou vidéos, telles que les « Deepfakes » ou les logiciels comme Midjourney, entrent aussi dans ce champ, avec « une volonté de destruction des modes de pensée des individus de manière durable et à grande échelle » ⁵. Les Deepfakes sont de plus en plus réalistes grâce aux IA. Par exemple, une voix peut être synthétisée à partir de seulement 30 secondes d'enregistrement public et utilisée en conjonction avec des techniques de

² Le Monde, « Un effort multilatéral est nécessaire pour préserver l'espace de l'information et de la communication », 17 mai 2023

³ Le Monde, « « AI Act » : comment l'UE investit déjà dans des intelligences artificielles à « haut risque » pour contrôler ses frontières », 22 juin 2023

⁴ France Info, « Guerre en Ukraine : ce que l'on sait de l'opération de désinformation russe « Doppelgänger » qui a visé la France », 14 juin 2023

⁵ La revue d'Histoire Militaire, « L'attaque des cerveaux : qu'est-ce que la guerre cognitive ? », 3 novembre 2022

phishing. Un exemple récent et sans précédent de l'utilisation des IA à des fins électoralistes est le clip des Républicains aux EU, intitulé « Beat Biden », mettant en scène un pays plongé dans le chaos si Biden était réélu⁶.

Dans tous les cas, les IA contribuent à renforcer les « bulles de filtre »⁷, confortant et nourrissant les opinions préexistantes de chacun. Ces bulles informationnelles représentent un risque de déstabilisation pour nos démocraties, qui peuvent se propager dans le monde réel. Les événements tels que l'invasion du Capitole le 6 janvier 2021 après l'élection de Biden aux Etats-Unis, ou l'invasion de la place des 3 pouvoirs le 8 janvier 2023 au Brésil, témoignent du rôle que peuvent jouer les réseaux sociaux, associés aux algorithmes de diffusion virale des messages de désinformation et d'appel au renversement du pouvoir. Des recherches sur des mouvements tels qu'« Occupy Wall Street », les « Indignados » ou les « Gilets Jaunes » ont montré l'importance des réseaux sociaux dans l'amplification de ces mouvements par des « actions connectives », ainsi que leur plus grande vulnérabilité aux « récits alternatifs »⁸. Et le développement potentiel du métavers suscite des inquiétudes quant à une amplification de ces risques. Les technologies immersives fournies par les IA pourront rendre les contre-récits, notamment dirigés contre les démocraties, plus persuasifs, en exploitant l'engagement corporel de l'individu et l'effet miroir, multipliant ainsi les possibilités de déstabilisation d'un pays par un autre, avec des armées d'avatars formés et engagés pour infiltrer les sphères du métavers et rejoindre des mouvements contestataires locaux, si besoin est. Cette menace de « subversion violente » est d'ailleurs prise en compte, en France, dans la

Stratégie Nationale de Renseignement (SNR). Une étude menée en partenariat avec l'OTAN en 2020 affirme que « le cerveau sera le champ de bataille du 21e siècle », menaçant les démocraties. Cela représente un risque pour l'intégrité des Etats, la cohésion nationale et une menace pour l'image des démocraties, au profit de régimes autocratiques et/ou d'intérêts privés. L'exemple de la campagne de désinformation entourant le faux « charnier » de Gossi au Mali, accusant les armées françaises, est révélateur des intentions de l'empire médiatique de Prigojine et du groupe Wagner, soutenus par la Russie à l'époque⁹. Enfin, le scandale de Cambridge Analytica, ayant manipulé les citoyens américains lors de la campagne présidentielle de 2016 au profit de l'équipe de Trump, met en évidence les enjeux électoraux pour une démocratie.

Par ailleurs, la question des jeunes et des futurs citoyens est un aspect central à prendre en compte. Le temps passé devant les écrans par les jeunes accroît les possibilités d'influence et de manipulation grâce à des IA et des algorithmes calibrés pour capter leur attention et leur recherche parfois frénétique du « like », générateur de dopamine¹⁰. Les travaux de Dominique Boullier démontrent le caractère addictif des applications telles que Facebook ou TikTok grâce à des IA extrêmement précises pour collecter les traces numériques de chacun, prédire les comportements et propager des données profilées et potentiellement orientées¹¹. D'où l'importance de l'éducation aux médias et aux IA, à l'esprit critique pour tous, et pas seulement pour les jeunes. Car même si des outils de lutte existent (en France, la L2I, Viginum, COMCYBER...), c'est un enjeu primordial pour les élus des sociétés démocratiques eux-mêmes, avec des scénarios déjà pensés d'IA remplaçant des politiques¹², dans des démocraties

⁶ Numérama, « Cette vidéo anti-Biden est faite par IA, une première dans une campagne », 25 avril 2023

⁷ Eli Pariser, *The Filter Bubble: What The Internet Is Hiding From You*, 2012

⁸ *The Conversation*, « « Gilets jaunes », médias et Internet : les premiers enseignements », 10 décembre 2018

⁹ RFI, *Les dessous de l'Infox*, « A Gossi, mise en scène d'un charnier

pour discréditer l'armée française », 22 juin 2022

¹⁰ Arte, « Dopamine, comment les applis piègent notre cerveau », juillet 2023

¹¹ Dominique Boullier, *Propagations, Un nouveau paradigme pour les sciences sociales*, Armand Colin, 2023

¹² Usbek&Rica, « Et si l'intelligence artificielle confisquait le pouvoir aux hommes politiques ? », 19 janvier 2023

où le poids des réseaux sociaux a réduit la place des corps intermédiaires, dans un mouvement de démocratie horizontale et ainsi ouvre la porte à des manipulations d'opinion plus importantes¹³

Des enjeux liés à la protection de la liberté et vie privée du citoyen

Quelle légitimité ont des acteurs privés mais aussi les gouvernements démocratiques eux-mêmes à utiliser les données des citoyens et dans quel but ? Avec quelle garantie de protection ?

Depuis février 2023, la Cour Suprême des Etats-Unis est saisie de deux affaires concernant la responsabilité des plates-formes numériques pour les contenus qu'elles diffusent, notamment la propagande de l'Etat islamique. La loi de 1996, connue sous le nom de « section 230 », les protégeait jusqu'à présent, mais elle pourrait être révisée. ChatGPT lui-même est également visé par cette loi. L'affaire Samsung, en avril dernier, a accru la prise de conscience des Etats quant aux failles en termes de sécurité des données, qui restent largement non contrôlées¹⁴. Les données ont été collectées via l'e-mail ou le numéro de téléphone demandés lors de l'inscription, en violation du règlement européen sur la protection des données (RGPD). En mars 2023, l'Italie est devenue le premier pays à interdire ce chatbot pour non-respect du RGPD. En mai 2023, l'Association française des correspondants à la protection des données (AFCDP) a, de même, souligné les risques liés à l'imprécision, voire à l'invention des données fournies par ces IA, et a insisté sur la nécessité d'imposer aux éditeurs concernés une plus grande transparence sur les sources exactes de données, en particulier celles contenant des données personnelles, ainsi que de respecter les droits

fondamentaux prévus par le RGPD.

Une autre problématique soulevée est celle de la conciliation entre la liberté d'expression, les règles à introduire sur les plateformes de réseaux sociaux et le risque de « techno-surveillance ». Etant donné que ces IA sont des propriétés privées, cela soulève la question de qui contrôle ces masses de données personnelles, dans quel but et avec quel cadre réglementaire. Par exemple, l'usage des caméras de surveillance « augmentées par l'IA » dans leur capacité de reconnaissance et d'identification, et demain éventuellement de nos émotions, a déjà suscité un débat animé en France en 2019, tandis que certaines villes aux Etats-Unis ont choisi de les interdire. L'exemple du « ranking social » mis en place en Chine offre un aperçu de ce que les technologies basées sur l'IA peuvent déjà réaliser dans un régime autoritaire. Bien que les démocraties disposent de garde-fous et d'un cadre juridique, elles ne sont pas à l'abri de risques réels, comme en témoigne le différend entre la CNIL (Commission Nationale de l'Informatique et des Libertés) et l'entreprise de reconnaissance faciale Clearview AI, dont le traitement des données est jugé « très intrusif »¹⁵. Ainsi, le smartphone pourrait devenir notre « ennemi intime ».

La juriste Mireille Delmas-Marty soulève, quant à elle, la nécessaire prise en compte des questions de nécessité et de proportionnalité de ces outils de surveillance dans une démocratie. Il est également crucial de faire une distinction claire entre le droit à la sûreté, constitutionnel, et le droit à la sécurité, qui englobe également la protection contre les abus de l'Etat¹⁶. Enfin, Asma Mhalla soulève le problème lié à l'association entre intérêts privés et intérêts publics, permettant ainsi la « neutralisation politique de la majorité passive » et le maintien sous contrôle de la

¹³ Institut Montaigne, « Corps intermédiaires : accords perdus de la démocratie ? », Blanche Leridon, avril 2023

¹⁴ RFI, « Des données sensibles de Samsung divulgués sur ChatGPT par des employés », 9 avril 2023

¹⁵ France Info, « Reconnaissance faciale : pour la CNIL, l'algorithme

de Clearview AI est un « traitement de données très intrusif » », 11 mai 2023

¹⁶ Interview sur TV5 Monde le 14 novembre 2019 « Reconnaissance faciale : une technologie compatible avec les valeurs de la République française ? »

« majorité silencieuse » sans avoir besoin de recourir à des actions directes sur le terrain, grâce à la collecte massive et au traitement algorithmique des données personnelles¹⁷. Le scandale des écoutes à l'échelle mondiale révélé par Snowden, qui a touché la démocratie étasunienne, est un cas emblématique des risques inhérents à ces technologies, y compris de la part de régimes démocratiques.

Les enjeux liés aux producteurs / concepteurs / utilisateurs d'IA : des Etats ou des acteurs privés porteurs de modèles, de valeurs mais aussi de biais

Effectivement, la question de la porosité entre les positionnements idéologiques étatiques ou personnels des acteurs des IA et les capacités de ces outils est centrale. Les orientations technologiques choisies en termes de stratégie économique peuvent influencer le développement et le cœur de « pensée » des IA. L'exemple d'Elon Musk, proche de l'alt-right et opposé au « wokisme », en est une illustration frappante. Il cherche à développer une IA anti-« woke », un « TruthGPT ». La guerre des IA, sur fond de post-vérité, a bien démarré selon Asma Mhalla¹⁸.

Concernant la présence de biais cognitifs dans les IA, la réalité rejoint parfois la fiction, comme aux Etats-Unis où des biais raciaux ont été identifiés dans des IA de prédiction utilisées par les services de police, les LegalTech, afin d'orienter leurs recherches ou interventions¹⁹. L'expérimentation a été ainsi abandonnée à Los Angeles. Le sociologue Gérald Bronner a soumis ChatGPT au test de réflexion cognitive (CRT) en février 2023 et le verdict est sans appel : les réponses étaient complètement fausses, avec des raisonnements intuitifs mais erronés²⁰. L'utilisation de telles IA dans un cadre démocratique

pose donc un problème de justice et de discrimination. Les travaux de la mathématicienne Cathy O'Neil sont également instructifs. Dans son livre *Algorithmes, la bombe à retardement* (Les Arènes, 2018), elle démontre comment l'intelligence artificielle accroît les inégalités entre les citoyens. Les biais algorithmiques sont en cause, reposant sur les valeurs portées, consciemment ou non, par les personnes qui collectent et utilisent les données. Cela se traduit par des algorithmes alimentés par des idéologies racistes, homophobes et sexistes. En juin 2023, la Fondation des femmes, Femmes Ingénieures, et Global Witness ont déposé deux plaintes liées à la discrimination sexiste des algorithmes de Facebook, après avoir testé des offres d'emploi dont les résultats ont montré la distribution genrée réalisée par les algorithmes²¹.

Conclusion : la régulation comme solution suprême ?

De nombreux Etats et organisations ont adopté ou prévoient d'adopter des cadres réglementaires juridiques pour réguler les IA. Cependant, les défis posés par les IA nécessitent des réponses politiques plus larges et ambitieuses, notamment de la part des Etats face aux grandes entreprises du numérique.

Les IA ne sont pas uniquement un marché économique, elles soulèvent des questions cruciales concernant la stabilité et la force de nos démocraties, l'intérêt général par rapport à l'intérêt privé, et la relation entre le corps social démocratique et l'individu isolé.

Il est sans doute nécessaire de réfléchir à de nouvelles lois pour les IA, comme le fit en son temps Asimov avec ses trois lois robotiques, ainsi que la création d'un GIEC de l'IA, comme cela a été

¹⁷ Le Grand Continent, « Techno-politique de la surveillance », Asmah Mhalla, 3 février 2023

¹⁸ Institut Montaigne, « Technopolitique de l'IA : luttes idéologiques, tensions géopolitiques, espoirs démocratiques », Asma Mhalla, 19 avril 2023

¹⁹ Le Meilleur des Mondes, France Culture, « Police prédictive,

LegalTech : les algorithmes font-ils la loi ? », 8 octobre 2021

²⁰ L'Express, « Quand ChatGPT a tout faux au célèbre test de réflexion cognitive », Gerald Bronner, 02 février 2023.

²¹ France Info, « Algorithmes : discriminations, sexisme et racisme, ce que vous devez savoir », Malika Boudiba, 21 juin 2023



récemment réclamé. De plus, il ne faut pas oublier que l'image d'un futur totalitaire et gouverné par la techno-surveillance, facilité par les IA, repose également sur le propre consentement des utilisateurs qui cèdent leurs données personnelles. Il s'agit d'une nouvelle forme de « servitude volontaire », 3.0.

La réglementation des IA soulève également des questions éthiques, notamment en ce qui concerne les droits humains, que ce soit dans le domaine militaire avec les armes autonomes létales ou dans le débat récent sur l'utilisation d'analyses prédictives pour les trajectoires des réfugiés, qui peut représenter un risque pour les droits fondamentaux et faciliter des pratiques illégales telles que le « pushback »²².

Pour conclure sur une note positive, il ne faut pas négliger les avantages des IA pour les citoyens, tels que la sécurité et la protection sur les réseaux sociaux grâce à des IA comme DarkBert, qui accélère la recherche sur le dark web pour lutter contre la cybercriminalité. Mais il demeure important de garder à l'esprit la loi d'Amara selon laquelle nous avons tendance à surestimer l'impact d'une nouvelle technologie à court terme et à le sous-estimer à long terme

²² Le Monde, « AI Act » : comment l'UE investit déjà dans des intelligences artificielles à « haut risque » pour contrôler ses frontières, 22 juin 2023