# 7. L'IA est-elle SEXISTE ?

LA REPRÉSENTATION STÉRÉOTYPÉE DES FEMMES GÉNÉRÉE PAR L'INTELLIGENCE ARTIFICIELLE INTERROGE. LES ALGORITHMES AMPLIFIENT-ILS DES CLICHÉS QUI PERDURENT ? DÉCRYPTAGE.

**LE PAPE EN DOUDOUNE IMMACULÉE,** Emmanuel Macron dans une manif… Ces derniers temps, pas une semaine ne passe sans qu'une nouvelle photo totalement *fake* ne fasse le buzz sur les réseaux sociaux. Derrière ces fausses images, bluffantes mais fabriquées, on trouve des intelligences artificielles (IA) dites génératives. Les plus connues ? Dall-E (qui se lit « Dali », jeu de mots), Stable Diffusion, ou encore Midjourney. À partir d'un simple texte, appelé « prompt », et avec une simplicité quasi enfantine, ces nouveaux outils permettent de créer des images de synthèse. Si jusqu'à peu les visuels réalisés manquaient de précision et restaient facilement traçables (notamment au niveau du rendu des mains), la nouvelle mise à jour de Midjourney, en mars dernier, a marqué un saut technologique : les images sont quasi parfaites !
Mais ces derniers temps, en plus des questions éthiques, de droits d'auteur et de manipulation de l'information, les IA créatrices d'images font polémique : et si elles ne produisaient que des représentations de femmes ultrasexualisées ? Et si ces robots étaient… sexistes ? De nombreux internautes reprochent aux algorithmes, ces boîtes noires au cœur des systèmes, d'avoir des « biais ». On le sait, les IA ingèrent des gigaoctets d'informations (on appelle ça le *machine learning*). Le problème ? Si les machines se basent sur des données d'apprentissage subjectives, le processus d'entraînement dévie – et les IA recrachent alors des réponses orientées. En clair, si vous demandez à Midjourney (et aux autres) de vous « dessiner un médecin », elle vous « inventera » majoritairement une photo… d'homme. Idem si vous préférez qu'elle crée un visuel de PDG.

**LES APPLIS NE SONT PAS EN RESTE.** Fin 2022, l'application Lensa, qui utilise l'IA de Stable Diffusion, faisait un carton (plus de 4 millions de téléchargement dès son

PHOTOS S. P.

PAR **SÉVERINE PIERRON**

apparition). Le concept ? Après avoir nourri Lensa d'une dizaine de selfies, on recevait une version « améliorée » de soi. Problème, les femmes se retrouvaient quasiment toutes transformées en avatars ultra-sexualisés, avec seins gonflés et lèvres pulpeuses. Sur TikTok, le filtre Bold Glamour (plus de 36 millions de vues) fait lui aussi appel à l'intelligence artificielle. Contrairement aux anciens filtres simples, qui ressemblaient à des masques plaqués sur le visage (en gros, la machine repérait par calcul les yeux et la bouche), avec ce filtre nouvelle génération, l'image est travaillée au niveau du pixel – et le résultat est bluffant, même en vidéo. Bold Glamour vous offre une version de vous-même qui pousse à fond tous les curseurs de la beauté à la Kardashian : nez ultrafin, pommettes liftées, yeux de chat et bouche façon Russian lips. Mais alors, pourquoi les IA ont-elles tendance à proposer des représentations biaisées des femmes ? La réponse se trouve au cœur de la « boîte noire » des IA, leurs algorithmes de calcul. Dès 2018, Benoît Raphaël (1), entrepreneur et fondateur de la start-up Flint, a conçu Ella, une agente intelligente de curation de contenus pour #JamaisSansElles, une association qui promeut l'égalité et la mixité (2). Afin d'obtenir des réponses plus vertueuses, l'IA d'Ella a été nourrie par des données pointées comme « non sexistes » par des experts humains. Benoît Raphaël précise : « Il est tout à fait possible d'entraîner une IA pour qu'elle soit féministe. L'IA n'est pas sexiste, car elle n'a pas de pensée sexiste. Les biais culturels nous impriment tous. »

**POUR LUI, CES BIAIS DE REPRÉSENTATION DES FEMMES** s'expliquent en partie à cause de la sur-représentation des hommes dans le numérique depuis des décennies (à peine 30 % de femmes dans la tech à l'échelle mondiale, et seulement 17 % de femmes diplômées en France, selon une enquête publiée en mars 2022 par Global Contact). Bref, le sexisme des IA reflète le sexisme de notre société tout entière. C'est aussi l'avis de Grégory Chatonsky, un artiste pionnier du Net art, régulièrement exposé à Paris, au palais de Tokyo ou au Centre Pompidou. Lui travaille sur les intelligences artificielles depuis 2009 : « Les IA sont un miroir déformant que nous nous tendons à nous-mêmes. En représentant nos préjugés, elles nous permettent de les explorer et d'en prendre conscience. » Mais pour l'artiste, dire que les IA sont sexistes est un « faux procès ». Il précise : « Les biais de genre ou raciaux deviennent des paramètres que l'on peut ajuster dans ce qu'on appelle l'espace latent. Une IA générative d'images fonctionne grâce à un grand nombre de données que l'on fournit à un logiciel qui va apprendre la probabilité d'apparition des pixels. » Et c'est là que l'on peut faire varier les paramètres : en rééquilibrant les données que l'on soumet au logiciel. Pour Grégory Chatonsky, « l'IA est un outil d'enquête de nos représentations imaginaires, et c'est comme cela que je m'en saisis en tant qu'artiste. On a tendance à doter les IA d'une espèce de conscience, ce qui est une projection anthropomorphique. Malheureusement, il est plus facile de corriger les inégalités de représentation dans l'espace numérique que dans l'espace réel ». En 2020, le Laboratoire de l'égalité signait un pacte pour une « intelligence artificielle égalitaire ». Le but ? Agir sur les données et les algorithmes « pour une IA aux résultats non sexistes ». Pour l'instant, un vœu pieux. ●

> "Il est tout à fait *possible d'entraîner* une *IA* pour qu'elle soit FÉMINISTE"
>
> Benoît Raphaël, entrepreneur et fondateur de la start-up Flint

*(1) Auteur d'« Information : l'indigestion », Éditions Eyrolles, 316 p., 19 €. (2) jamaissanselles.fr*

Les filtres de nouvelle génération, comme Bold Glamour sur TikTok, n'appliquent pas les effets des filtres classiques mais utilisent l'IA générative… aux résultats troublants.