



EUROPEAN GUILD
FOR ARTIFICIAL
INTELLIGENCE
REGULATION

Nous sommes un groupe d'artistes, de créatifs, d'éditeurs et d'associations de toute l'Europe, unis pour porter à l'attention du public la façon dont nos données et nos propriétés intellectuelles sont exploitées sans notre consentement, à une échelle jamais vue auparavant. Une telle situation aussi inédite nous a poussés à unir nos forces pour atteindre les institutions européennes et faire entendre nos voix. Si tu crois que tes données et ton travail créatif ne devraient pas être exploités impunément à des fins de profit par une poignée d'entreprises, rejoins-nous pour soutenir cette bataille.

L'été 2022 a vu l'émergence d'une nouvelle technologie incroyable : L'IA texte-image (TTI) ou image-image. Chaque fois qu'une technologie révolutionnaire voit le jour, notre société doit superviser son déploiement afin d'éviter tout préjudice ou toute atteinte aux droits de l'homme. Cela ne s'est pas encore produit avec la technologie de l'intelligence artificielle. Ces formes d'intelligence artificielle peuvent générer des images ou d'autres contenus à partir d'invites textuelles ou d'autres supports, visuels, audios, donnés par l'utilisateur. Pour ce faire, une IA doit être entraînée sur un ensemble de données de médias. La qualité d'une IA générative est définie par la qualité de son ensemble de données - par exemple, en ce qui concerne les images, plus une IA s'entraîne sur des images et des illustrations, plus l'IA est capable de reproduire des styles et plus elle peut faire de choses. Par conséquent, les produits vendus par les entreprises d'IA sont le résultat d'opérations sur des ensembles de données, qui contiennent toutes sortes de données, y compris des millions d'images protégées par des droits d'auteur, de photos privées et d'autres documents sensibles. Ces fichiers ont été collectés en grattant indistinctement l'Internet sans le consentement des propriétaires et des personnes qui y sont représentés et sont actuellement utilisés par les entreprises d'IA à des fins lucratives. De plus, certaines de ces entreprises utilisent les noms de certains des artistes dont elles ont gratté les œuvres pour faire la publicité de leur IA et des styles qu'elles peuvent reproduire. Cette exploitation de notre travail et de nos données non seulement ne répond pas aux exigences minimales imposées par les droits de l'homme qui sont à la base de notre société : elle porte aussi gravement atteinte au marché de l'art, le marquant potentiellement à jamais. Nous considérons que ce n'est que le début d'une crise qui touchera toutes sortes d'emplois et de professions, qu'il s'agisse d'emplois créatifs ou non. Le marché de l'art est le premier à être touché uniquement en raison de ses vulnérabilités structurelles, qui en font une proie facile. Il est temps de changer cela.

Ce sont les points clés que nous voulons faire respecter :

- 1) Toute donnée relative aux personnes ou aux œuvres, sous quelque forme que ce soit, qu'il s'agisse de données numériques - telles que des fichiers texte, des audios, des vidéos ou des images - ou capturées de la réalité par une caméra, des microphones ou tout autre moyen d'enregistrement, ne doit pas être utilisée pour former un modèle d'IA sans le consentement explicite et informé de son propriétaire. Nous demandons une extension aux IA des principes de protection des données personnelles précédemment introduits par le GDPR et l'introduction d'une nouvelle forme de protection spécifique à ce type d'exploitation : le "droit de formation".
- 2) L'utilisation de noms de personnes, de noms de scène ou de titres d'œuvres non couvertes par une licence d'exploitation pour l'apprentissage de l'IA est interdite pour les logiciels qui permettent d'utiliser une invite textuelle ou vocale pour générer des images, des vidéos, des textes ou du son.
- 3) L'utilisation de vidéos, d'images, d'audios et de textes non couverts par une licence d'exploitation pour la formation à l'IA est interdite pour les logiciels qui permettent de télécharger des contenus multimédia pour générer une image, une vidéo, un texte ou un audio.
- 4) Un système d'indexation et de certification "lisibles par l'homme et la machine" doit être mis en place, signalant toutes les activités des IA et le contenu complet de leurs ensembles de données d'images, de textes, de vidéos et de sons, qu'ils soient entièrement ou partiellement reproduits. Des légendes telles que "entièrement réalisé par l'IA", "réalisé à l'aide de matériel généré par l'IA" devraient devenir la norme.
- 5) La distinction entre "matériel protégé par des droits d'auteur" et "domaine public" ne permet plus d'identifier ce qui peut et ne peut pas être utilisé pour les ensembles de données. Les ensembles de données d'apprentissage contiennent des données personnelles sensibles, protégées par les lois sur la confidentialité, mais pas par le droit d'auteur. Nous pouvons trouver des exemples de matériel publié alors qu'il n'aurait pas été possible de prévoir son utilisation dans un ensemble de données pour former un modèle d'IA. Toute donnée utilisée pour former un modèle doit être conservée et autorisée par son propriétaire légitime et insérée volontairement dans l'ensemble de données par son auteur en toute connaissance de cause. Les entreprises d'IA doivent produire en interne du matériel original pour la formation ou accorder une licence pour du matériel externe en suivant les conditions et les contrats préalablement établis avec les auteurs ou les détenteurs légitimes dudit matériel.

Pour en savoir plus, consulter le site www.egair.eu.