

Depuis un peu plus d'un an, des modèles génératifs d'images, basés sur l'intelligence artificielle, permettent de créer des images de plus en plus esthétiques à partir d'une simple commande textuelle. Cela a pour effet de bouleverser le marché de l'art. De plus, ces modèles sont entraînés en utilisant les travaux d'artistes, ce qui soulève des questions de droits d'auteur, en plus de révolter les auteurs qui attaquent en justice les entreprises de l'IA.

Les artistes révoltés

Le travail des artistes volés ?

Ceux-ci sont les premiers concernés par cette controverse. Leur art est scrapé (collecté, souvent illégalement) sur de nombreuses plateformes de diffusion, puis utilisé pour entraîner la majorité des modèles génératifs que nous connaissons, comme l'ensemble de données LAION (5,85 milliards de paires image-texte), ce qui pose des problèmes de droits d'auteur. Certains artistes considèrent qu'entraîner des modèles n'est ni une utilisation légitime ni éthique de leur travail. Une petite coalition d'artistes attaquent actuellement en justice MidJourney, Stability AI ainsi que Deviant Art (les principales plateformes de génération), et appellent à une redéfinition de la notion de droit d'auteur.

Les modèles dans le collimateur

Les plateformes de génération sont accusées d'avoir stocké des millions d'images sur des serveurs pour une utilisation professionnelle sans autorisation. Les artistes accusent également les images générées d'être des travaux dérivés des originaux.

LAION, mentionné en tant que tierce partie dans le procès de Getty Image contre Stable Diffusion, a déjà été attaqué en justice en 2023 pour faire retirer certaines images du jeu de données. L'entreprise admet ouvertement ne pas se porter garante du respect des droits d'auteurs, affirmant ne faire que référencer les images déjà présentes sur Internet.

Qu'en dit la loi ?

Les licences

Celles-ci déterminent ce qu'on peut faire ou non avec une image, la plupart autorisent la reproduction et le partage d'images, à condition que cela ne soit pas à des fins commerciales.

Images libres

A condition de citer l'auteur, toutes les utilisations des images sous licence CC-BY sont autorisées, comme l'entraînement des modèles, théoriquement. Les artistes appellent à une modification pour refuser cette utilisation sans autorisation.

Productions de l'IA

Les images produites par l'IA peuvent parfois s'apparenter à des copies modifiées d'images existantes. Cependant, certains modèles vont plus loin et apprennent des concepts plus généraux à partir de vastes ensembles de données. Considère-t-on alors l'image générée comme une copie ?

Le fine tuning pour copier le style

Le fine tuning consiste à spécialiser un modèle déjà entraîné, appelé modèle de fondation, souvent pour imiter fidèlement un artiste spécifique. C'est un point d'ombre légalement. Peut-on empêcher quelqu'un d'imiter le style d'un autre ?

Il existe des cas de procès similaires entre artistes, comme Deborah Roberts en 2023 contre Lynthia Edwards. Néanmoins, empêcher quelqu'un d'utiliser le travail d'autrui pour devenir un artiste, même en imitant son style, peut sembler hors-propos.



Les travaux de Lynthia Edwards (gauche) et de Deborah Roberts (droite) mis côte à côte.

Les enjeux sociaux

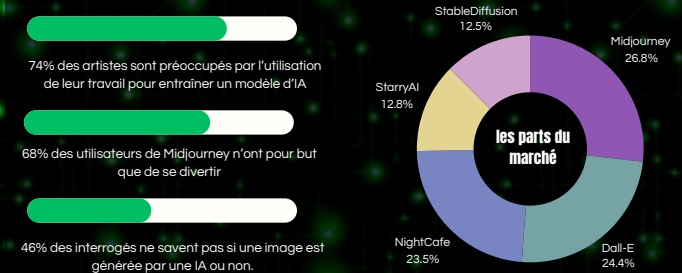
Une accessibilité accrue

D'après les plateformes de génération dont certaines sont gratuites ou peu chères, l'IA générative est très utile pour rendre l'art plus accessible. Peu importe la formation, les compétences, ou les moyens financiers (ceux n'ayant pas les moyens de payer un artiste), tout le monde peut prétendre à la création d'œuvres artistiques. Cela permet à de nombreuses personnes d'exprimer leur créativité, en utilisant les bon prompts pour faire générer le résultat imaginé.

Des biais, parfois volontaires

La totalité des modèles, entraînés sur un ensemble de données arbitraire, disposent de nombreux biais. Certains modèles génèrent des images avec un style très occidental, par exemple. Les entreprises derrière les modèles peuvent également choisir d'inclure volontairement un biais, ce qui pourrait poser problème. Ces biais ont déjà conduit à plusieurs scandales (comme une IA proposée par Twitter pour centrer une image, qui ne centrerait que des personnes blanches).

Voir le QR code pour les sources des études !



En résumé ?

L'utilisation abusif d'images sous licence pose un problème légal de droit d'auteur. La tension monte entre les artistes qui jugent que l'entraînement des modèles n'est pas un usage légitime de leur travail et les plateformes de génération d'images. Celles-ci mettent en avant la démocratisation de l'art comme un bénéfice pour la société, et ont des raisons économiques de vouloir se développer. Plusieurs plateformes commencent cependant à proposer des IA génératives entraînées sur leur propre licence d'image, comme Adobe Firefly, dont le rôle n'est pas de remplacer les artistes, mais de devenir le « copilote de la conception graphique » (The Guardian).

Les conséquences économiques

Dans l'industrie de l'art ?

Les artistes voient leurs services peu à peu remplacés par des IA. Par exemple, Disney a récemment utilisé une IA pour produire un générique, celui de la série Secret Invasion.

Avec l'ère de l'art digital, les réseaux sociaux sont devenus le nouveau lieu d'exposition pour les artistes. Les artistes digitaux voient les IA génératives voler une part de leur audience... avec des images générées sur des modèles entraînés sur leurs propres créations (comme Midjourney, plus de 16 millions d'utilisateurs).

La controverse des NFTs

Les NFTs, des actifs numériques uniques basés sur la technologie de la blockchain permettent de certifier la propriété et l'authenticité d'un fichier, souvent des œuvres d'art numériques. Ainsi, bien qu'une image numérique puisse être copiée, le possesseur du NFT détient "l'original", qu'il peut choisir de revendre comme on vendrait un tableau. De très nombreux internautes se servent d'IAs pour créer des collections de NFTs imitant un artiste. La rémunération va à la personne ayant généré l'image et non à l'artiste qui a dessiné les images utilisées pour l'entraînement du modèle, comme c'est arrivé à Tome Hanuka.



Trois images de la collection de NFTs « Punks by Hanuka » générée par IA.

https://midjourney.com

Qu'est-ce qu'un modèle génératif ?

Ce sont des modèles généralement basés sur les réseaux de neurones. Ceux-ci sont entraînés sur d'énormes ensembles de données, tels que le dataset LAION qui contient des milliards d'images provenant de diverses sources, souvent sans autorisation explicite des créateurs.

La méthode de diffusion

Celle-ci est utilisée dans les modèles les plus connus comme MidJourney, Stable Diffusion ou encore Dall-E. Les grands modèles de langage interprètent le texte d'un prompt pour en extraire la signification et le convertir en vecteur numérique (word embedding). Ce vecteur devient le guide pour le processus de diffusion, où le modèle apprend à générer et affiner des images en introduisant et en éliminant progressivement du bruit dans un ensemble de données.

Ces vecteurs sont ensuite traités par des couches supplémentaires (couches de convolution) dans le modèle de langage, qui peuvent inclure des réseaux de neurones récurrents (RNN), des réseaux de neurones convolutifs (CNN) pour traiter le texte comme une séquence, ou des transformateurs qui utilisent des mécanismes d'attention pour mieux capter les relations complexes entre les mots dans des phrases plus longues. Le résultat est un ensemble de features qui encapsule le contenu du prompt textuel, plus le prompt est clair plus l'image est mieux réalisée.

Ce processus permet à Midjourney de transformer un amas de pixels en une image qui correspond à la description de l'utilisateur, le tout en à peu près une minute.

