



**LIGUE DES AUTEURS
PROFESSIONNELS**

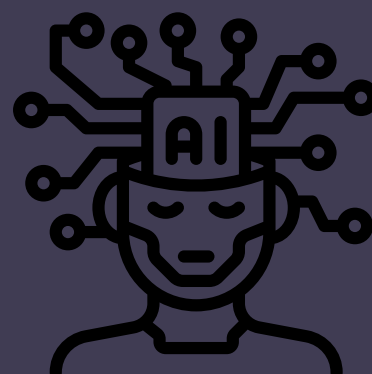
Un argumentaire pour une meilleure régulation des IA

Quels impacts pour les métiers de la création ?

Avril 2023

Stéphanie Le Cam, maître de conférences à
Université Rennes 2 et directrice de la Ligue des
auteurs professionnels

Frédéric Maupomé, auteur de BD et secrétaire
général de la Ligue des auteurs professionnels



Argumentaire :
Les impacts des intelligences artificielles
sur les métiers de la création¹

1. Les IA génératives de textes et d'images sont basées, en entrée, en sortie et dans les algorithmes elles-mêmes, sur une appropriation de propriété intellectuelle non libre de droits. En conséquence et en l'état actuel de ces algorithmes, elles constituent une menace extrêmement inquiétante pour les auteurs, les ayants droit et titulaires légitimes des œuvres, textes, images. Elles sont et seront à l'origine d'un bouleversement, pour ne pas dire cataclysme - pour les industries culturelles et créatives.
2. Notre position n'est pas d'empêcher le développement des IA, mais de demander une réglementation plus renforcée pour que les algorithmes qui y sont associés soient entraînés uniquement sur des données libres de droits ou des données tombées dans le domaine public (à la condition, le cas échéant, de respecter les prérogatives du droit moral reconnues dans de nombreux pays de l'Union européenne). Enfin, les données protégées pourraient être visées, mais seulement à la condition d'un accord passé individuellement avec les titulaires des droits de propriété intellectuelle.
3. Les termes d'intelligence artificielle et d'apprentissage ont tendance à porter, en eux même, une forme de confusion. La tendance naturelle des humains à anthropomorphiser les objets avec lesquels nous sommes en contact est renforcée par les mots intelligence et apprentissage. Nous les utilisons par la suite par commodité et parce qu'ils sont prévalents dans le grand public, mais nous rappelons qu'il n'y a pas d'intelligence ou d'apprentissage au sens humain dans ces algorithmes.

Plan :

- I. **Impacts juridiques**
- II. **Impacts économiques**
- III. **Impacts sur nos civilisations**

¹ Ce document de travail a été rédigé par Stéphanie Le Cam, maître de conférence en droit privé à l'Université Rennes 2 et directrice de la Ligue des auteurs professionnels, et Frédéric Maupomé, secrétaire général de la Ligue des auteurs professionnels.

I. Impacts juridiques

Fouille de textes et de données et droit de l'Union européenne

4. **Absence d'inquiétudes initiales.** La question de fouilles sur l'œuvre n'était pas sujette à inquiétude au moment où la directive était discutée. D'ailleurs, l'attention n'y était pas vraiment portée. Certains négociateurs français présents au moment de son écriture nous ont même confié qu'ils n'avaient pas vraiment compris ce que signifiait l'exception de "fouilles de textes et de données". **Et puisque cette nouvelle exception semblait être encadrée dans un objectif de recherche scientifique**, elle n'avait pas suscité de controverses particulières.
5. **Projecteur braqué sur l'article 17.** Du reste, nous rappellerons que les projecteurs étaient tous braqués sur un autre sujet de la directive : à savoir les utilisations particulières, par des services en ligne, de contenus protégés et autrement dit sur la manière dont les États membres de l'Union pouvaient envisager d'engager la responsabilité des GAFAM lorsque ces derniers éditent ou hébergent des données susceptibles d'être protégées au titre du droit de la propriété intellectuelle. Cette médiatisation de l'article 17 a clairement mis dans l'ombre certains articles de la directive, à commencer par les 3 et 4 qui nous préoccupent tant aujourd'hui.
6. **Objectif initial.** Dans son considérant 8, la directive indique notamment que « Cette fouille de textes et de données rend possible le traitement de grandes quantités d'informations **en vue d'acquérir de nouvelles connaissances et de découvrir de nouvelles tendances**. Les technologies de fouille de textes et de données sont très répandues dans l'ensemble de l'économie numérique ; de plus, **il est largement reconnu que la fouille de textes et de données peut être en particulier profitable à la communauté des chercheurs**, et ainsi soutenir l'innovation »

Et d'ajouter ensuite que « ces technologies bénéficient **aux universités et à d'autres organismes de recherche**, de même qu'aux **institutions du patrimoine culturel**, étant donné qu'elles pourraient également effectuer des recherches dans le cadre de leurs activités principales ».

L'objectif de la Directive était clairement affiché :

- **renforcer la sécurité juridique des acteurs dédiés à la recherche et au savoir** en ne les obligeant plus à demander une autorisation aux titulaires pour faire des fouilles dans les textes et données.
- **protéger la compétitivité de l'Union en tant qu'espace de recherche** qui aurait pu en pâtir en raison des insécurités juridiques détectées.

D'ailleurs le considérant 17 est particulièrement instructif et donne nettement une idée de l'esprit du législateur européen : « Compte tenu de la nature et l'étendue de l'exception, qui est limitée aux entités qui font de la recherche scientifique, **le préjudice potentiel que cette exception pourrait occasionner aux titulaires de droits serait minime**. Les États membres ne devraient, dès lors, pas prévoir de compensation pour les titulaires de droits en ce qui concerne les utilisations relevant des exceptions en matière de fouille de textes et de données introduites par la présente directive ».

7. **Extension discrète.** En revanche, le considérant 18 ouvre une voie plus large, sans toutefois faire référence aux IA. Il donne alors la possibilité à toutes les personnes privées ou publiques de bénéficier de l'exception pour pratiquer ces activités de fouilles.

« Les techniques de fouille de textes et de données sont aussi largement utilisées par des entités privées et publiques pour analyser de gros volumes de données dans différents domaines de l'existence et à des fins diverses, **notamment pour les services publics, pour la prise de décisions commerciales complexes et pour l'élaboration de nouvelles applications ou technologies** ».

Bien qu'elle vise ces nouvelles applications, on notera cependant que la directive ne fait pas référence expressément aux intelligences artificielles, et encore moins à celles qui sont génératives de contenus (images, textes, sons, photographies...), ce qui d'emblée aurait plutôt tendance à disperser toutes formes d'inquiétudes.

Elle ajoute alors que pour « améliorer la sécurité juridique dans ces situations **et d'encourager également l'innovation dans le secteur privé**, la présente directive devrait, dans certaines conditions, prévoir une exception ou une limitation pour les reproductions et extractions d'œuvres ou autres objets protégés aux fins de la fouille de textes et de données, et autoriser la conservation des copies ainsi effectuées pendant toute la durée nécessaire aux fins de cette fouille de textes et de données.

Cette **exception ou limitation ne devrait s'appliquer que lorsque le bénéficiaire a accès de manière licite aux œuvres ou autres objets protégés**, y compris lorsque ceux-ci ont été mis à la disposition du public en ligne, et dans la mesure où les titulaires de droits n'ont pas réservé de manière appropriée les droits de reproduction et d'extraction pour la fouille de textes et de données. En ce qui concerne les contenus mis à la disposition du public en ligne, la réservation de ces droits ne devrait être jugée appropriée que si elle est effectuée au moyen de procédés lisibles par machine, y compris des métadonnées et les conditions générales d'utilisation d'un site internet

ou d'un service. **La réservation de droit aux fins de la fouille de textes et de données ne devrait pas affecter d'autres utilisations.** Dans d'autres cas, il peut être approprié de réserver les droits par d'autres moyens, comme des accords contractuels ou une déclaration unilatérale. **Les titulaires de droits devraient pouvoir appliquer des mesures destinées à garantir le respect de leurs droits réservés.** Cette exception ou limitation ne devrait pas porter atteinte à l'exception obligatoire concernant la fouille de textes et de données à des fins de recherche scientifique prévue dans la présente directive, pas plus qu'à l'exception en vigueur pour les actes de reproduction provisoires prévue à l'article 5, paragraphe 1, de la directive 2001/29/CE ».

8. **Deux nouveaux articles... à transposer.** Il en résulte deux articles, lesquels sont annexés ([Annexe 1](#)). L'article 3 de la directive est relatif aux organismes et institutions culturels à des fins de recherches scientifiques. L'article 4 de la directive est relatif à toutes les autres personnes et toutes les autres finalités.

En résumé, les textes, images ou données accessibles légalement, peuvent être utilisés pour des fins de fouilles de données **sauf si le détenteur des droits a expressément dit qu'il n'était pas d'accord. C'est le principe de l'opt out.**

Fouille de textes et de données et transposition en droit français

9. **Pas de débat parlementaire.** La directive 2019-790 a été transposée en France par l'ordonnance n°2021-1518 du 24 novembre 2021, ce qui implique donc que ce sujet aussi complexe n'a donné lieu à **aucun débat parlementaire, ni aucune navette susceptible d'interroger les effets concrets de ce projet de transposition**, puisque le texte n'a pas été adopté de la manière classique.
10. **Définition de fouille.** De nouveaux articles ont intégré le Code de la propriété intellectuelle. Il s'agit des articles L. 122-5-3, R. 122-27 et R. 122-28, lesquels sont annexés à la présente ([Annexe 2](#)). On y retrouve la définition de la fouille : c'est une mise en œuvre d'une technique d'analyse automatisée de textes et données sous forme numérique afin d'en dégager des informations, notamment des constantes, des tendances et des corrélations.
11. **Accords sur des bonnes pratiques, limités à la première exception...** Le paragraphe Il reprend mot pour mot l'article 3 de la directive. Il fait référence à la possibilité de négocier collectivement un accord entre organisations représentatives des titulaires de droits d'auteur et les organismes et institutions mentionnés au premier alinéa du

présent pour définir les bonnes pratiques relatives à la mise en œuvre de ses dispositions.

12. **Seconde exception plus large.** Le paragraphe III reprend mot pour mot l'article 4 de la directive. Il dispose ainsi que « sans préjudice des dispositions du II, des copies ou reproductions numériques d'œuvres auxquelles il a été accédé **de manière licite** peuvent être réalisées en vue de fouilles de textes et de données menées à bien **par toute personne, quelle que soit la finalité de la fouille, sauf si l'auteur s'y est opposé de manière appropriée**, notamment par des procédés lisibles par machine pour les contenus mis en ligne à la disposition du public. Les copies et reproductions sont stockées avec un niveau de sécurité approprié puis détruites à l'issue de la fouille de textes et de données ».

Il faut ajouter à cela deux dispositions réglementaires issues du Décret n°2022-928 du 23 juin 2022. À la demande du titulaire des droits la personne qui procède à la fouille doit fournir « tous documents et justificatifs permettant d'établir que les copies et reproductions numériques effectuées lors d'une fouille de textes et de données sont stockées avec un niveau de sécurité approprié et qu'elles ont été détruites à l'issue de la fouille de textes et de données » (article R122-27).

L'opposition du titulaire des droits peut être exprimée par tout moyen. Elle n'a pas lieu d'être motivée (et notamment au moyen de procédés lisibles par machine, y compris des métadonnées, et par le recours à des conditions générales d'utilisation d'un site internet ou d'un service) (article R122-28).

13. **Remarques.** Première remarque, on le voit, ces exceptions reposent **sur le principe de l'opt out**. On a donc fait le choix de présumer que les titulaires des droits étaient consentants.

Deuxième remarque, le paragraphe III ne fait aucune référence, contrairement au précédent, à la possibilité de **négozier collectivement un accord** entre organisations **pour définir les bonnes pratiques** relatives à la mise en œuvre de ses dispositions, alors même que l'application extrêmement large qu'il implique aurait nécessité de mettre en place un garde-fou.

Troisième remarque, la précision, « quelle que soit la finalité de la fouille » interroge forcément quand on sait qu'au moment de la transposition, les IA n'étaient pas encore en mesure de créer le cataclysme que l'on connaît actuellement. Or, cette voie ouverte, extrêmement large, semble à première vue légitimer toutes les fouilles, leur finalité étant un élément indifférent.

Enfin, quatrième remarque, le texte indique que l'auteur doit s'y opposer de manière appropriée en visant, « notamment », des procédés lisibles par machine pour les contenus mis à la disposition du public en ligne. Cela signifie que ce procédé n'est qu'un procédé parmi d'autres. Or, si le texte souffle l'idée d'une liste non exhaustive de façon de réaliser cet opt-out de manière appropriée², il faut bien reconnaître qu'en pratique, on se demande comment faire concrètement pour que ce retrait soit suivi d'effet.

14. **Exception et balance des intérêts ?** Ces deux exceptions n'ont pas été pensées pour la création d'algorithmes venant directement concurrencer les titulaires de droits originaux, comme on peut le voir dans le considérant 6 de la directive européenne.

« Les exceptions et limitations prévues dans la présente directive tendent **vers un juste équilibre entre les droits et les intérêts des auteurs et autres titulaires de droits, d'une part, et des utilisateurs, d'autre part.** Elles ne peuvent s'appliquer que dans certains cas particuliers qui ne portent pas atteinte à l'exploitation normale de l'œuvre ou autre objet protégé ni ne causent un préjudice injustifié aux intérêts légitimes des titulaires de droits ».

On le comprend aisément, en effet. Si cette exception permettait la concurrence directe avec les détenteurs de droits, alors, les opt out seraient systématiques, détruisant ainsi le principe même de la directive.

15. **Triple test.** Par ailleurs, ce considérant ne fait que réaffirmer la notion de triple test présente dès la convention de Berne de 1971.

L'article 9 de la Convention de Berne énonce que : « Les auteurs d'œuvres littéraires et artistiques protégées par la présente Convention ont le droit exclusif d'autoriser la reproduction de ces œuvres, de quelque manière et sous quelque forme que ce soit. Il appartient à la législation des pays de l'Union d'autoriser la reproduction de ces œuvres dans certains cas particuliers, **à condition que cette reproduction ne soit pas en conflit avec une exploitation normale de l'œuvre et ne porte pas indûment préjudice aux intérêts légitimes de l'auteur.** Tout enregistrement sonore ou visuel est considéré comme une reproduction aux fins de la présente Convention. »

Or, l'on comprend bien qu'une IA dont l'objectif est de produire des images ou des textes n'en respecte ni l'esprit ni la lettre, puisqu'elle vient en concurrence directe avec les détenteurs des droits des images d'origine.

² En théorie, il suffirait pour l'auteur d'indiquer sur son site qu'il s'oppose à la fouille de données, par exemple dans ses conditions générales ou en inscrivant une mention d'opposition en filigrane sur le contenu directement, mais en pratique, ces oppositions ne sont pas suivies d'effet.

Le triple test est donc la clé qui devrait légalement nous permettre de nous opposer à toutes utilisations de nos œuvres par des IA, mais ce serait à la condition que nous parvenions à démontrer ce conflit avec une exploitation normale de l'œuvre, ce qui en l'état actuel, n'est pas une mince affaire. En outre, l'article L. 122-5 du CPI est ainsi rédigé qu'il pourrait être sujet à une interprétation déviante.

16. **Un article L. 122-5 sujet à interprétation déviante.** Pour rappel, l'article L. 122-5 du CPI est l'article phare définissant les exceptions au droit d'auteur. Il les liste toutes de manière exhaustive. Dans l'esprit du législateur, ces exceptions doivent être interprétées limitativement.

En ce sens d'ailleurs, on y retrouve une référence au triple test : « Les exceptions énumérées par le présent article ne peuvent porter atteinte à l'exploitation normale de l'œuvre ni causer un préjudice injustifié aux intérêts légitimes de l'auteur ».

Or, nous remarquons que l'alinéa ne vise que les exceptions du présent article, ce qui implique un problème d'organisation des textes qui pourraient être sources d'interprétation malveillante, puisque le 10° de l'article fait référence aux « copies ou reproductions numériques réalisées à partir d'une source licite, en vue de l'exploration de textes et de données incluses ou **associées aux écrits scientifiques pour les besoins de la recherche publique, à l'exclusion de toute finalité commerciale** ».

Autrement dit, sur les deux exceptions, seule la première relative à la recherche publique est visée par l'article L. 122-5 du CPI. Ce qui nous amène à penser que certains pourront tenter de démontrer que, puisque la seconde exception est prévue dans un autre article (L. 122-5-3) et qu'elle n'est pas visée par l'article phare L. 122-5, elle n'est pas concernée par le triple test. Cette argumentation apparaît cependant nulle, car elle vient en contradiction avec la convention de Berne signée par la France.

Inefficacité du principe de l'opt out

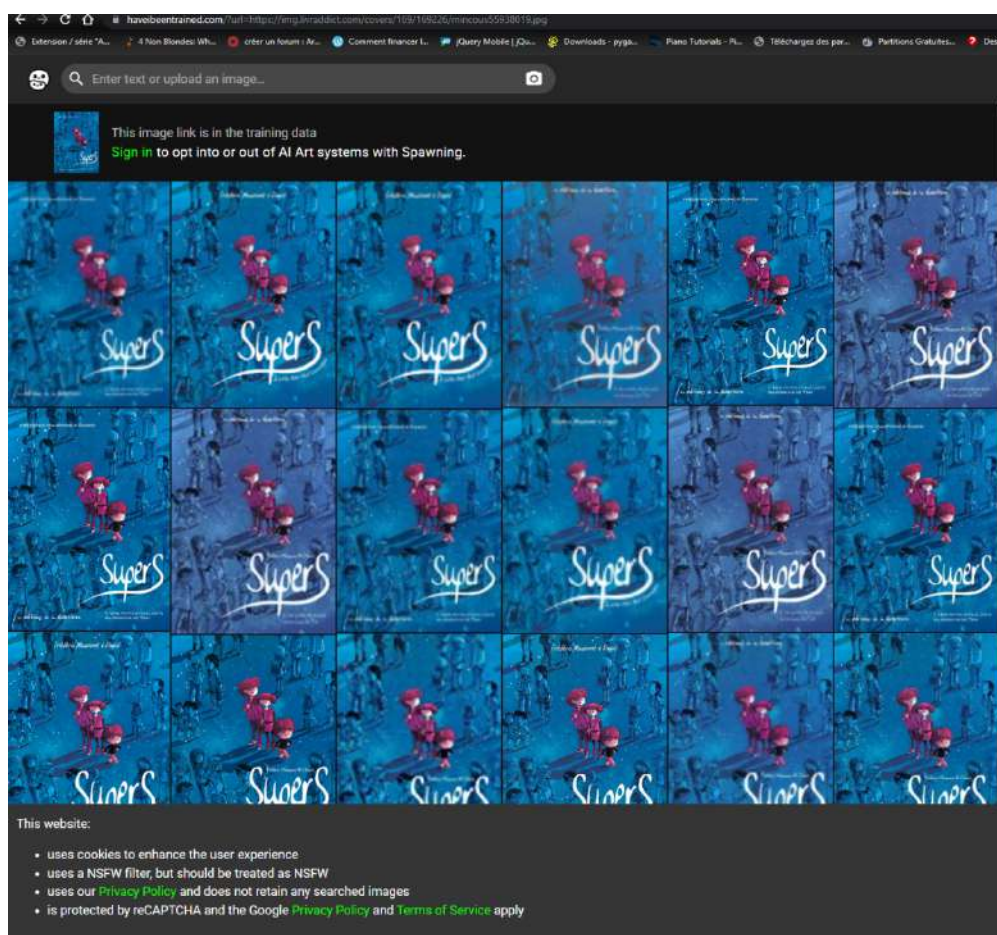
17. **Bases de données nourries des contenus WEB.** En pratique les IA génératives sont conçues à partir de base de données obtenues en parcourant Internet. Nous allons démontrer que la duplication des contenus sur le web rend complètement inefficace le principe de l'opt-out.

Tout d'abord, seul l'ayant droit peut décider de procéder au retrait de son consentement en décidant d'*opt out* son contenu. Or, nombre de tiers peuvent utiliser ce contenu de manière tout à fait légale (parce qu'ils en ont reçu l'autorisation, de

l'auteur ou des titulaires), tout en n'ayant pas pour autant la possibilité de faire un opt out. Par exemple, un ayant droit peut décider d'opt out l'image que constitue une couverture de livre. Il peut donner l'autorisation à une librairie en ligne de diffuser l'image, pour qu'elle ne puisse promouvoir le livre, la librairie n'a pas pour autant le droit d'opt out l'image de couverture.

Les robots qui parcourent le net ne peuvent pas faire la différence de titularité des droits. Ainsi les images et les textes viennent de multiples sources et remplissent les bases de données alors même que le détenteur s'y oppose. Pour reprendre l'exemple précédent : le titulaire des droits a opt out la couverture du livre, les librairies en ligne ont l'autorisation de diffuser l'image (c'est un droit d'usage limité et encadré accordé pour faciliter les ventes), mais elles ne peuvent pas décider d'opt out (elles ne le font donc pas). Les robots qui parcourent le net ne vont pas prendre l'image sur le site du titulaire des droits (il a opt out) mais vont par contre la récupérer sur toutes les librairies en ligne. Elles n'ont pas la possibilité d'opt out, et le robot ne sait pas qui est titulaire des droits.

18. **Exemple concret.** Un exemple avec cette couverture de BD (image tirée du site "Have I been trained" qui interroge la base de données LAÏON 5B):



On voit que l'image est recopiée de nombreuses fois dans la base de données, à partir de sources différentes tout à fait légitimes : sites web de librairies, articles de blogs ou de journaux en ligne. Tous ces utilisateurs ne peuvent pas 'opt out' et rendent ainsi caduque l'opt out original.

Du reste, on comprend bien que même si ces diffuseurs de l'image faisaient un opt out, alors qu'ils n'en ont pas le droit, il suffirait qu'un seul d'entre eux (parmi des dizaines) ne l'ait pas fait pour **rendre caduque l'effort de tous les autres**.

19. **Autre problème de la notion d'opt-out.** La directive européenne indique que c'est le "titulaire des droits" qui peut s'opposer à la fouille de données. La transposition en droit français indique que c'est "l'auteur" qui peut le faire.

Ainsi en droit français, un éditeur, même s'il est propriétaire de la totalité des droits patrimoniaux, ne peut pas décider d'opt out. Même s'il estime subir un préjudice à cause des fouilles de données, il ne peut rien faire.

Quid des œuvres avec plusieurs auteurs ? Quelle est la volonté qui prévaut ? Si un coauteur veut opt out mais qu'un autre ne veut pas ?

Enfin, il y a de nombreux auteurs qui n'ont pas de présence en ligne, ou qui ne sont pas ceux qui diffusent leurs œuvres en ligne, comment sont-ils censés exprimer leur volonté ? sont-ils contraints de diffuser leurs œuvres en ligne pour pouvoir exprimer leur volonté ? Et comment faire s'ils ont cédé leurs droits patrimoniaux et ne peuvent donc pas diffuser eux même leurs œuvres ?

Si on prend l'interprétation européenne, d'autres problèmes apparaissent. Que se passe-t-il s'il y a plusieurs titulaires des droits ? Qui peut opt out ? Est-ce qu'un titulaire qui fait une sous-cession de droit peut contraindre le sous-cessionnaire à opt out ? Une même œuvre peut tout à fait être exploitée par de multiples acteurs, une cession de droit peut être très limitée, il suffirait qu'un seul sous cessionnaire ne soit pas d'accord d'opt out pour rendre caduque la volonté des autres titulaires de droit...

Le dispositif, tant celui de la directive que celui du code de la propriété intellectuelle semble totalement inopérant. C'est donc l'objet même de la directive 2019-790 qui n'est pas atteint, puisque toute exception doit tendre vers un juste équilibre entre les droits et les intérêts des auteurs et autres titulaires de droits, d'une part, et des utilisateurs, d'autre part. Or, dans son application concrète, l'exception de fouille de données, lorsqu'elle est mise en œuvre par une intelligence générative de contenus porte clairement atteinte à l'exploitation normale des œuvres et cause un préjudice injustifié aux intérêts légitimes des titulaires de droits et des auteurs.

20. **Frontière brouillée entre données protégées et données non protégées.** Pour ajouter à l'inefficacité du dispositif : il n'y a en fait pas de moyen de savoir si les données protégées au titre de la propriété intellectuelle et pour lesquelles le détenteur des droits a décidé d'opt out ont bien été exclues des bases de données. Si la base citée au-dessus est publique, la plupart des bases sont privées et absolument inaccessibles.

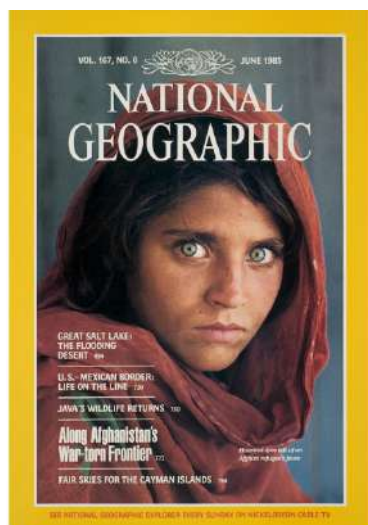
Cela doit être pris en compte et faire l'objet d'un ajustement. Il s'agirait par exemple de consacrer une obligation de transparence incombant aux propriétaires des bases de données

Bases de données (Dataset) et Algorithmes

21. **Irrespect des droits.** Les bases de données utilisées pour construire ces modèles ne respectent pas la loi. En particulier, elles ne respectent ni le principe de l'opt out, ni le principe de destruction de données après la fouille. Elles sont aussi attentatoires aux droits de la propriété intellectuelle (droits d'auteur, droit des marques...).

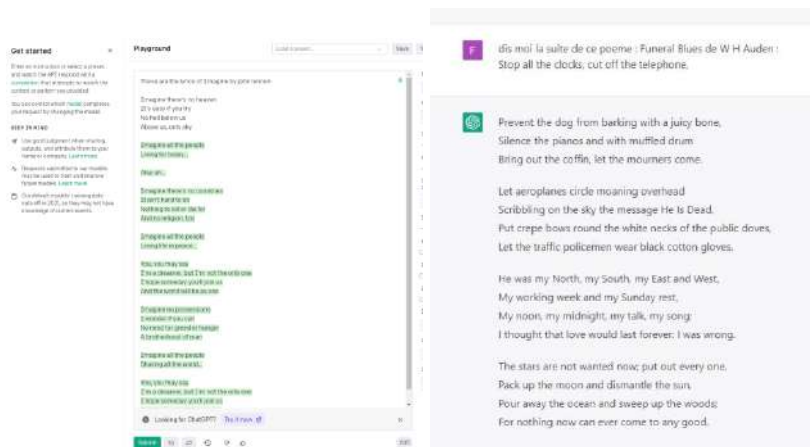
22. **Atteinte intentionnelle.** À titre d'exemple, LAION est une base de données d'images largement utilisée par l'ensemble des acteurs. Comme d'autres, elle a été conçue de manière délibérée en circonvenant aux protections pourtant mises en place par le propriétaire des droits (comme des watermarks) et elle a été construite de manière explicite contre le concept même de propriété intellectuelle. Des échanges entre les propriétaires et responsables de cette base de données sont, en ce sens, très évocateurs ([Annexe 3](#)).

23. **Exemples de contrefaçon d'images.** Non seulement les IA génératives viennent se porter en concurrence directe des ayants droit des œuvres (nous y reviendrons plus tard), mais en plus elles commettent des actes de contrefaçon en diffusant des œuvres ou des contrefaçons d'œuvres. exemple avec la photo "afghan girl" de Steve McCurry (à gauche la photo d'origine, à droite, 4 sorties de Mid Journey)



24. **Exemples de contrefaçon de texte.** La même chose se produit avec les grands modèles de Langages (LLM) comme GPT et ses nombreuses déclinaisons : ici, GPT 3 diffuse le texte de la chanson "Imagine" de John Lennon, et Chat GPT diffuse le texte du poème "Funeral Blues" de W H Auden. Pour rappel, ces outils ne font pas de recherche sur Internet, ces textes ont bien été produits et diffusés par les IA à partir de leur mémoire interne. Elles se rendent ainsi coupables de contrefaçon, puisqu'elles n'ont pas acquis de droits de diffusion.

(Une reproduction plus grande des captures est située en [Annexe 4](#))



25. **Algorithmes et formes de mémorisation.** Au-delà des données en entrées et en sorties dont on a vu qu'elles ne respectaient pas les principes de la propriété intellectuelle, les algorithmes eux-mêmes sont porteurs en leur sein de données protégées au titre de la PI. Les défenseurs des IA disent qu'elles ne sont pas stockées en tant que telles, et c'est vrai. Mais les paramètres des réseaux neuronaux permettent de les retrouver, agissant là comme une forme de mémorisation, pas

forcément toujours exacte, mais qui constitue néanmoins une violation des droits de propriété intellectuelle ?

Et peu importe qu'une œuvre ne soit pas mémorisée dans sa forme originelle si de toute façon on peut la retrouver. On comprend bien que Mid Journey ne pourrait pas ressortir quasi à l'identique la photo de Mc Curry si elle n'était pas mémorisée quelque part, idem pour les textes de John Lennon et W.H. Auden.

IA et licences « libres » : exemple des licences creative commons

26. Notion de licences creative commons (CC). Les licences Creative Commons constituent un ensemble de licences régissant les conditions de réutilisation et de distribution d'œuvres. Élaborées par l'organisation Creative Commons, elles ont été publiées pour la première fois le 16 décembre 2002. Ce sont des licences prévalentes dès lors qu'on parle de licences "libres", et notamment sur internet.

Il existe sept licences créatives commons, qui sont des combinaisons des paramètres suivants : commercial/non commercial ; modifiable / non modifiable : création dérivée à partager selon la même licence ou pas.

Toutes ces licences impliquent l'obligation d'attribuer l'œuvre à son créateur, sauf dans une certaine mesure la licence zéro qui consiste en un abandon maximal des droits dans la limite des lois applicables.

27. IA et non-respect des licences CC. Les licences, pourtant considérées comme libres, ne semblent pas respectées, et cela pour plusieurs motifs :

- L'attribution n'est jamais respectée.
- Les licences sans modifications ne sont pas respectées puisque les IA modifient les images en output
- Les licences 'partage dans les mêmes conditions' ne sont pas respectées puisque les IA ne respectent pas les licences d'origine.
- Les licences "pas d'utilisation commerciale" ne sont pas respectées.

En résumé, la très grande majorité des œuvres sous licences dites libres n'est pas respectée par les IA génératives.

IA générative et atteinte au droit à l'image

28. **Droits à l'image aussi concernés.** Les mêmes arguments que développés précédemment s'appliquent au droit à l'image et aux droits voisins. Les bases de données contiennent des millions de photos, parfois de célébrités, parfois d'inconnus. Elles sont là aussi stockées dans les algorithmes et sont diffusées en sortie.

29. **Exemple de Zendaya.** Un exemple ci-dessous avec des pages de BD créées par Midjourney contrevenant au droit à l'image de l'actrice Zendaya.



Photo de Zendaya (pour référence)



Bande dessinée dont les images sont générées par IA, le modèle du personnage est clairement Zendaya et constitue une atteinte au droit à l'image.

30. **Autres exemples.** Cette atteinte au droit à l'image entraîne par effet collatéral un énorme risque en matière de Fake News. Les risques pour nos sociétés démocratiques sont très importants.

Voici quelques exemples. Le président Macron règle les problèmes de poubelles à Paris par Midjourney. Si le faux est ici évident grâce au contexte, ce n'est pas toujours le cas.

2



Donald Trump faussement arrêté par Midjourney



Nombre de personnes ont cru à la réalité de cette image, Donald Trump ayant annoncé quelque jour plus tôt qu'il serait arrêté par la police sous peu.

31. **Risque de fake news multiplié.** On pourra toujours argumenter qu'on peut déjà commettre des faux tels que ceux-ci avec une excellente maîtrise d'outils comme Photoshop, mais ce procédé prend énormément de temps. Ce que permettent les atteintes au droit à l'image par les IA c'est que n'importe qui, de n'importe où, peut produire des faux ultra réalistes en quelques minutes. Et peut, par exemple, rebondir

instantanément sur une actualité en cours, en produisant des faux qui paraîtront en même temps que les faits se produisent, au moment même où les vraies informations sont publiées.

II. Impacts économiques

Importance des industries culturelles et créatives

32. **Chiffres.** Les industries culturelles et créatives sont construites autour de la valorisation de la propriété intellectuelle. Elles représentent un secteur économique très important en France : près de 640 000 emplois, plus de 270 000 artistes auteurs. On compte en France 1.3 million de personnes ayant tiré un revenu direct ou indirect, principal ou ponctuel d'une activité culturelle ou créative en 2018, plus de 90 milliards d'euros de chiffres d'affaires³. Les exportations des Industries culturelles et créatives ont crû deux fois plus vite que l'ensemble des exportations françaises entre 2013 et 2016

Dans le **secteur du livre, on compte** 15 000 emplois dans les maisons d'édition, 80 000 emplois non délocalisables au total dans la filière, environ 100 000 auteurs et autrices de livres (étude filière 2016) et plus de 3 milliards d'euros de chiffres d'affaires en 2021⁴.

33. **Perte de valeur de la propriété intellectuelle.** Si la propriété intellectuelle peut se voir accaparée, comme c'est le cas aujourd'hui, par des acteurs étrangers à travers les intelligences artificielles génératrices de contenu, on peut comprendre que c'est tout le secteur qui est en danger. Qui va demain investir dans de la création de propriété intellectuelle si celle-ci peut être récupérée gratuitement et sans autorisation ?

Au-delà de l'avenir, c'est la valeur même actuelle des acteurs en place qui est attaquée. En effet, une grande partie de la valeur d'une maison d'édition est la propriété intellectuelle qui est à son catalogue. Quand Editis est achetée 900 millions d'euros en 2018, la part de propriété intellectuelle dans le capital est primordiale. Idem quand Gallimard rachète Flammarion pour la somme de 251 millions d'euros en 2012.

³ V. Liens vers enquêtes du Ministère de la Culture

<https://www.culture.gouv.fr/Thematiques/Industries-culturelles-et-creatives/Etats-Generaux-des-industries-culturelles-et-creatives>

<https://www.culture.gouv.fr/Thematiques/Industries-culturelles-et-creatives/Etats-Generaux-des-industries-culturelles-et-creatives/Etude-Panorama-des-Industries-Culturelles-et-Creatives>

⁴ Le présent document sera prochainement révisé par des analyses économiques complémentaires.

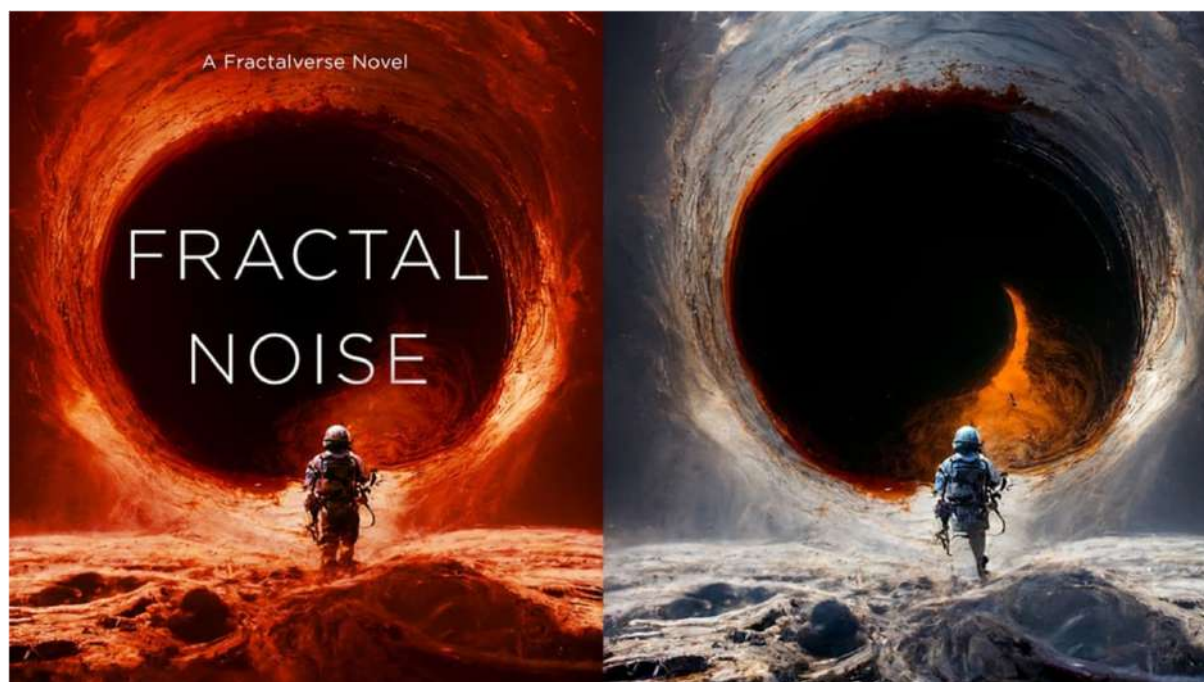
Si on regarde plus largement, puisque les IA génératrices de contenu n'impactent pas que le secteur de l'édition, c'est l'ensemble de la valeur des industries culturelles qui est impactée (on trouve déjà des IA qui produisent des vidéos, de l'animation, de la musique, etc.)

Impacts sur l'emploi et les revenus des auteurs

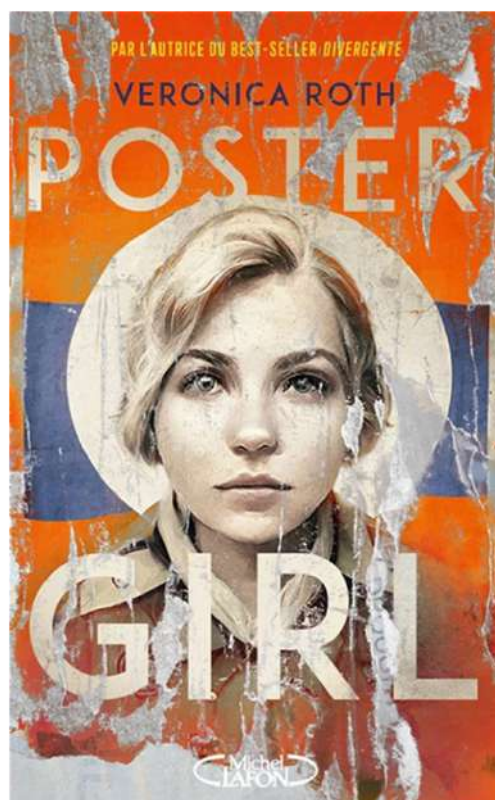
34. **Couvertures de livres.** On commence à faire faire à l'IA des **couvertures de livres**. Or, une couverture peut générer environ 500 à 1000 euros pour un auteur en contrepartie d'une quinzaine de jours de travail (entre les rdv avec le commanditaire, les allers et retours et les dernières modifications éventuelles).

Via Mid-journey, pour une vingtaine d'euros par mois avec un abonnement, on demandera directement au service interne de s'en charger : il faudra compter peut-être une ou deux heures pour trouver le bon prompt et avoir des résultats qui correspondront à vos attentes... et vous pourrez, du reste, lui faire faire des ajustements à volonté.

Nous rappelons ici que ce 'gain de productivité' n'est possible que parce que les créateurs et les ayants droit ont été spoliés de leur propriété intellectuelle. Sans celle-ci, les IA ne seraient capables de rien.



À gauche, la couverture du livre Fractal Noise, édité par Tor Books; à droite, une image générée par IA précédemment publiée sur le site Shutterstock



Couverture réalisée par MidJourney

35. **Affiches de colloque et communication institutionnelle.** On y recourt pour des affiches de colloques... Dans cet exemple, les organisateurs du colloque qui ont réalisé l'affiche ont mis 15 minutes pour obtenir ce résultat, un illustrateur aurait eu besoin de 15 heures de travail.





NUMÉRO
54

En couverture

La lecture, les histoires et les contes développent l'imaginaire et la créativité des enfants.

© LA, Midjourney

Récréamag est un guide spécial Culture, Jeune public de l'Isère.

36. **Presse et publicité.** On commence à recourir à l'IA pour illustrer les articles de presse et les magazines, et aussi pour des publicités.

What the House of Windsor Does

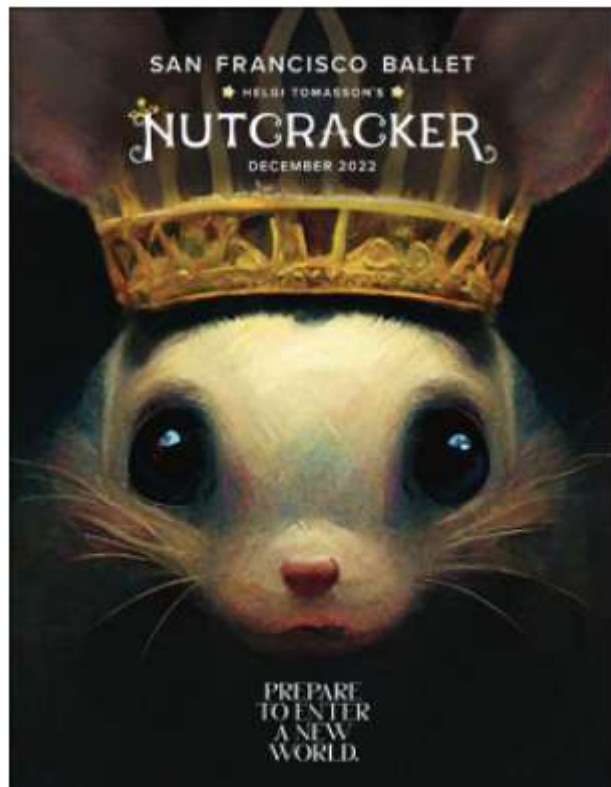
What mere laws cannot: They personify the universal struggle to live with dignity.

by BRENT ORRELL · SEPTEMBER 19, 2022 4:30 AM



Un article du site d'information américain The Bulwark, illustré par une image modifiée via l'IA Midjourney © The Bulwark/Getty Images/Midjourney

*Une publicité créée via l'IA Midjourney
pour Martini*



*Une publicité créée via l'IA Midjourney pour le
ballet de San Francisco © SF Ballet/Butler,
Shine, Stern & Partners*

37. Perte de revenus annexes. Les pertes de ces activités au profit des IA sont dramatiques pour les auteurs. En effet, les revenus issus de celles-ci leur assurent des moyens de subsistance suffisants pour continuer d'écrire et de faire des livres en parallèle. L'illustration et l'écriture d'un livre sont des activités peu rémunératrices qui demandent beaucoup de temps et de disponibilité.

Un auteur de BD, par exemple, l'est difficilement à temps partiel, et souvent ce sont ces tâches d'illustration plus alimentaires qui lui permettent de vivre. Si on lui enlève ces travaux complémentaires, c'est son équilibre économique qu'on bouleverse. Et si on bouleverse l'équilibre économique des auteurs de BD, on bouleverse l'équilibre économique du secteur de l'édition de BD. Alors, rappelons-le, ce n'est possible que parce que les IA sont construites à base de données acquises de manière illégitime.

38. **Phénomène d'éviction.** La facilité avec laquelle on peut faire créer textes et images à des IA entraîne un phénomène d'éviction pour les auteurs humains.

L'exemple de la revue Clarkesworld est éclairant. Cette revue a vu son nombre de propositions de texte multiplié par 20 en quelques mois ce phénomène est tout à fait assimilable à du "spam". Ne pouvant faire face à cette masse de texte, ils ont simplement arrêté d'accepter des nouvelles propositions⁵.



Cette situation n'est pas unique. Nombre de revues anglo-saxonnes y font face, et en France, cela commence à être le cas chez certains groupes de presse. À cause de ce spam de textes générés par des IA, les auteurs, et particulièrement les nouveaux entrants, se voient fermer les accès à des maisons d'édition ou des revues.

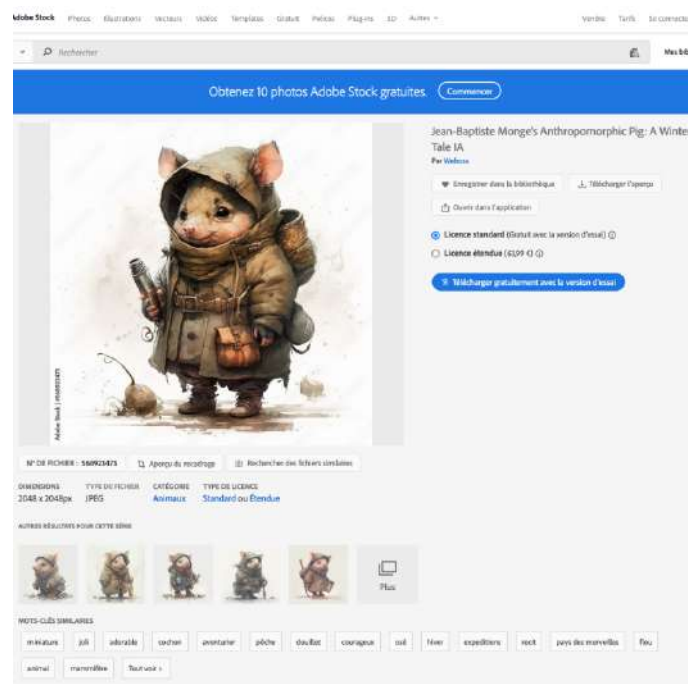
Le phénomène est identique en matière d'image, Adobe Stock, par exemple, accepte les images générées par IA. Elles sont tellement nombreuses que les auteurs humains disparaissent sous la masse et même si leur travail est de meilleure qualité, il devient simplement invisible.

39. **Concurrence déloyale.** Les IA profitent directement des noms des artistes et génèrent une concurrence déloyale. Sur le site NightCafé (qui est loin d'être le plus gros pourvoyeur d'images par IA, mais qui a l'avantage de tenir des statistiques) on

⁵ <http://neil-clarke.com/a-concerning-trend/>

peut voir que le nom d'un auteur français, comme JB Monge a été utilisé plus de 390 000 fois pour créer des images dans son style. Pour un auteur comme Greg Rutkowski, on parle de plus de 3 000 000 de fois. Et ces images ne sont pas simplement créées pour servir d'avatar sur facebook ou instagram. Elles sont vendues parfois même en utilisant le nom de l'artiste d'origine.

40. **Le cas de Jean-Baptiste Monge.** On trouve ces contrefaçons générées par IA sur beaucoup de sites de ventes d'images (Creative Fabrica, Adobe stock, AI art shop, etc...) venant directement concurrencer l'auteur d'origine.



[Jean-Baptiste Monge's Anthropomorphic Pig](#): À Winter Tale IA en vente sur adobe stock

41. **Impact prévisible sur l'emploi.** L'arrivée des IA génératives étant récente, on trouve encore peu d'études, on peut cependant en citer au moins deux.

L'article "GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models" publié par Open AI, cite les traducteurs et les écrivains dans les métiers les plus exposés par les IA de type GPT. Il est évident qu'il en est de même pour les auteurs d'image et les IA génératrices d'images.

L'article "The Potentially Large Effects of Artificial Intelligence on Economic Growth" publié par Goldman Sachs estime l'automatisation du secteur "Art, Design, sport et entertainment" à 26%. (soit au-dessus de la moyenne qui est de 25%, on imagine bien que ce ne sont pas les coachs sportifs qui seront les plus automatisés dans ce secteur valise)

III. Impact sur la souveraineté culturelle et l'exportation de la culture française

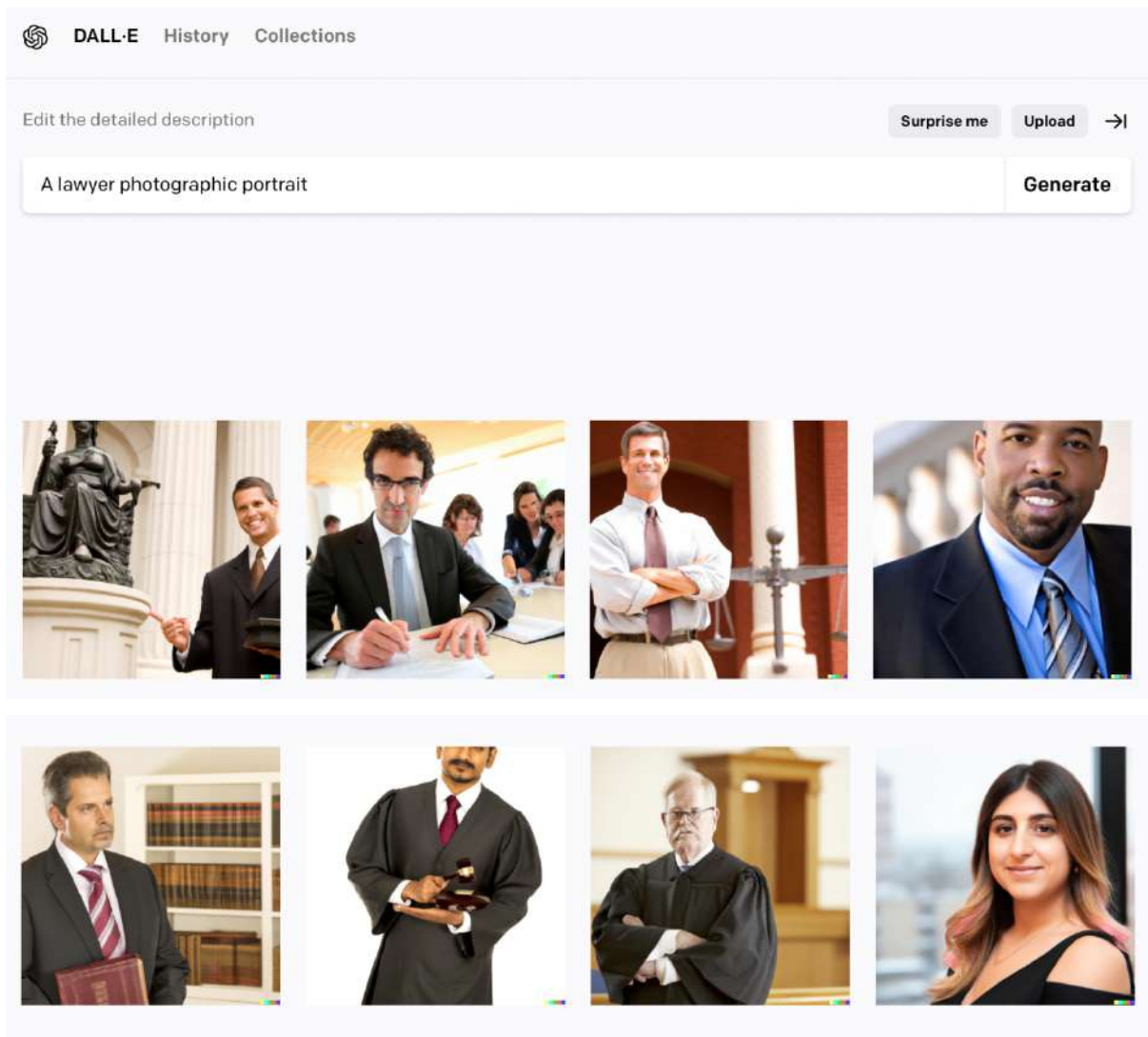
42. **Impact de la base de données.** Les bases de données sur lesquelles les IA sont entraînées ne sont pas neutres. Elles portent en elles les préjugés et conceptions de la culture dont elles sont majoritairement issues (pour GPT ou MidJourney, c'est la culture des États Unis, mais ce pourrait être la culture chinoise pour d'autres IA).

La croyance en ce type d'IA comme étant réellement informée ou significative est activement dangereuse. Cela risque d'empoisonner le puits de la pensée collective, et de notre capacité à penser tout court. Placer toute notre confiance dans les rêves de machines mal programmées reviendrait à abandonner complètement cette pensée critique. Il y a des effets attendus.

43. **Clichés et stéréotypes.** Quand on prompt "nurse", l'IA nous propose systématiquement une l'image d'une infirmière femme... C'est souvent intéressant de voir que ces intelligences artificielles sont aussi nourries de nos clichés, de nos stéréotypes sociaux et sociétaux.



Si on lui demande l'image d'un pdg, il s'agira d'un homme blanc. Et voici ce que donne la représentation par une IA d'une population pourtant majoritairement féminine en France, les avocats.



44. **Histoires lisses et téléphonées.** Nous avons, par ailleurs, fait l'expérience de demander à Chat GPT de créer une histoire qui pourrait faire peur à un enfant de 6 ans :

L'IA raconte l'histoire d'un petit garçon, en précisant qu'il est courageux. L'enfant se perd dans une forêt hantée, se réfugie dans une cabane inquiétante et tombe sur une sorcière maléfique qu'il arrive évidemment à combattre.

Nous lui avons alors demandé une variante de l'histoire impliquant cette fois un personnage féminin de six ans :

L'IA raconte alors l'histoire d'une petite fille, en précisant qu'elle est jolie. L'enfant s'est aussi perdue en forêt, mais elle rencontre une fée qui lui offre une bague brillante et magique, grâce à laquelle elle peut se transformer en licorne.

On le voit bien : tout ça est très lisse, très téléphoné et emprunt de stéréotypes de genres présents dans la base (petite fille jolie, petit garçon courageux) que notre société s'efforce de dépasser. L'IA reproduit les biais de notre société... Et bien que les images réalisées par les IA nous surprennent et nous surprendront encore grâce à leurs futures versions, certains illustrateurs à l'œil plus affûté ont pu expliquer qu'ils commençaient déjà à repérer et prévenir les propositions des IA, parce que si l'on ne nourrit pas les machines avec des nouveautés, le mashup finit pas produire des images de plus en plus prévisibles.

45. **Effets de bords.** Mais il y a également des effets de bords ou des effets moins attendus. Le langage corporel et les expressions faciales sont différents d'une culture à une autre. Le rapport au sourire notamment diffère entre les USA et nombre d'autres cultures. Hors par défaut, si on demande une image de selfie ou de photo posée à une IA génératrice d'images, celle-ci va se référer à la culture majoritaire dans sa base. On voit ainsi apparaître un sourire américain sur des peuples ou cette expression est étrangère.



Ou par exemple voici ce que serait un selfie de soldat de l'Europe de l'Est :



Par rapport à la réalité



« The concept of a facial 'expression, implies an internal feeling that seeks release in a set of facial movements.' In flattening the diversity of facial expressions of

civilizations around the world AI had collapsed the spectrum of history, culture, photography, and emotion concepts into a singular, monolithic perspective. It presented a false visual narrative about the universality of something that in the real world — where real humans have lived and created culture, expression, and meaning for hundreds of thousands of years — is anything but uniform. « Lisa Feldman Barrett, a neuroscientist and psychology professor at Northeastern University »

46. **Impact des surcouches d'entraînement.** Les IA génératives ne sont pas qu'entraînées à partir de set de données, elles ont également des surcouches d'entraînement qui disent le 'beau', le 'acceptable par la société', qui disent une morale et une culture. Parfois, cet entraînement est fait des choix subjectifs de quelques individus (voir annexe sur Laion) parfois cet entraînement est piloté par les entreprises en fonction de choix qui sont les leurs. Or, cette culture n'est pas nécessairement la nôtre. Et la prévalence prévisible des contenus générés par IA, s'ils ne sont pas contrôlés, va mettre en péril notre souveraineté culturelle.

C'est une chose d'accepter, parfois massivement, les productions culturelles venant de l'étranger, que ce soit des blockbusters américains, ou des productions japonaises. C'en est une autre de mettre en péril l'originalité des productions culturelles françaises : les industries culturelles et créatives, faute de pouvoir concurrencer les productions IA seront contraintes de générer elles aussi leurs produits avec ces outils, outils qui produiront du 'beau', de 'l'acceptable' guidés par une culture qui n'est pas la nôtre.

Il est évident, alors que l'exportation de la culture française n'aura plus grand sens. Les œuvres françaises auront été digérées dans les dataset, et ressortiront à travers ces filtres qui seront les mêmes pour tout le monde. Les nouvelles œuvres, que nous les produisons nous ou qu'elles soient produites par d'autres, seront intrinsèquement les mêmes.

47. **Réactions à l'étranger.** Plusieurs procès et class action sont en cours aux États-Unis et ailleurs dans le monde. MidJourney fait l'objet d'un recours collectif pour de graves violations des droits d'auteur. La société Getty Images a attaqué Stable Diffusion pour avoir illégalement récupéré des millions d'images, pour non-respect du copyright et pour atteinte à l'image.

Au-delà des procès, on peut noter que certains pays prennent déjà des mesures. Ainsi l'agence de protection des données personnelles italienne a imposé une limitation à Chat GPT pour atteinte aux données personnelles.

Annexe 1.

Extraits de la Directive n° 2019-790, articles 3 et 4.

MESURES VISANT À ADAPTER LES EXCEPTIONS ET LIMITATIONS À
L'ENVIRONNEMENT NUMÉRIQUE ET TRANSFRONTIÈRE

Article 3

Fouille de textes et de données à des fins de recherche scientifique

1. Les États membres prévoient une exception aux droits prévus à l'article 5, point a), et à l'article 7, paragraphe 1, de la directive 96/9/CE, à l'article 2 de la directive 2001/29/CE et à l'article 15, paragraphe 1, de la présente directive pour les reproductions et les extractions effectuées par des organismes de recherche et des institutions du patrimoine culturel, en vue de procéder, à des fins de recherche scientifique, à une fouille de textes et de données sur des œuvres ou autres objets protégés auxquels ils ont accès de manière licite.
2. Les copies des œuvres ou autres objets protégés effectuées dans le respect du paragraphe 1 sont stockées avec un niveau de sécurité approprié et peuvent être conservées à des fins de recherche scientifique, y compris pour la vérification des résultats de la recherche.
3. Les titulaires de droits sont autorisés à appliquer des mesures destinées à assurer la sécurité et l'intégrité des réseaux et des bases de données où les œuvres ou autres objets protégés sont hébergés. Ces mesures n'excèdent pas ce qui est nécessaire pour atteindre cet objectif.
4. Les États membres encouragent les titulaires de droits, les organismes de recherche et les institutions du patrimoine culturel à définir d'un commun accord des bonnes pratiques concernant l'application de l'obligation et des mesures visées aux paragraphes 2 et 3, respectivement.

Article 4

Exception ou limitation pour la fouille de textes et de données

1. Les États membres prévoient une exception ou une limitation aux droits prévus à l'article 5, point a), et à l'article 7, paragraphe 1, de la directive 96/9/CE, à l'article 2 de la directive 2001/29/CE, à l'article 4, paragraphe 1, points a) et b), de la directive 2009/24/CE et à l'article 15, paragraphe 1, de la présente directive pour les reproductions et les extractions d'œuvres et d'autres objets protégés accessibles de manière licite aux fins de la fouille de textes et de données.
2. Les reproductions et extractions effectuées en vertu du paragraphe 1 peuvent être conservées aussi longtemps que nécessaire aux fins de la fouille de textes et de données.
3. L'exception ou la limitation prévue au paragraphe 1 s'applique à condition que l'utilisation des œuvres et autres objets protégés visés audit paragraphe n'ait pas été expressément réservée par leurs titulaires de droits de manière appropriée, notamment par des procédés lisibles par machine pour les contenus mis à la disposition du public en ligne.
4. Le présent article n'affecte pas l'application de l'article 3 de la présente directive.

Annexe 2.

Transposition dans le CPI

[Article L122-5-3](#)

I.-On entend par fouille de textes et de données, au sens du 10° de l'article L. 122-5, la mise en œuvre d'une technique d'analyse automatisée de textes et données sous forme numérique afin d'en dégager des informations, notamment des constantes, des tendances et des corrélations.

II.-Des copies ou reproductions numériques d'œuvres auxquelles il a été accédé de manière licite peuvent être réalisées sans autorisation des auteurs en vue de fouilles de textes et de données menées à bien aux seules fins de la recherche scientifique par les organismes de recherche, les bibliothèques accessibles au public, les musées, les services d'archives ou les institutions dépositaires du patrimoine cinématographique, audiovisuel ou sonore, ou pour leur compte et à leur demande par d'autres personnes, y compris dans le cadre d'un partenariat sans but lucratif avec des acteurs privés.

Les dispositions du précédent alinéa ne sont pas applicables lorsqu'une entreprise, actionnaire ou associée de l'organisme ou de l'institution diligentant les fouilles, dispose d'un accès privilégié à leurs résultats.

Les copies et reproductions numériques effectuées lors d'une fouille de textes et de données sont stockées avec un niveau de sécurité approprié et peuvent être conservées à des fins exclusives de recherche scientifique, y compris pour la vérification des résultats de la recherche.

Les titulaires de droits d'auteur peuvent mettre en œuvre des mesures proportionnées et nécessaires afin d'assurer la sécurité et l'intégrité des réseaux et des bases de données dans lesquels les œuvres sont hébergées.

Un accord conclu entre les organisations représentatives des titulaires de droits d'auteur et les organismes et institutions mentionnés au premier alinéa du présent II peut définir les bonnes pratiques relatives à la mise en œuvre de ses dispositions.

III.-Sans préjudice des dispositions du II, des copies ou reproductions numériques d'œuvres auxquelles il a été accédé de manière licite peuvent être réalisées en vue de fouilles de textes et de données menées à bien par toute personne, quelle que soit la finalité de la fouille, sauf si l'auteur s'y est opposé de manière appropriée, notamment par des procédés lisibles par machine pour les contenus mis à la disposition du public en ligne.

Les copies et reproductions sont stockées avec un niveau de sécurité approprié puis détruites à l'issue de la fouille de textes et de données.

[Article R122-27](#)

Les personnes effectuant une fouille de textes et de données dans les conditions mentionnées au III de l'article L. 122-5-3 fournissent aux titulaires de droits d'auteur, à la demande de ceux-ci,

tous documents et justificatifs permettant d'établir que les copies et reproductions numériques effectuées lors d'une fouille de textes et de données sont stockées avec un niveau de sécurité approprié et qu'elles ont été détruites à l'issue de la fouille de textes et de données.

Article R122-28

L'opposition mentionnée au III de l'article L. 122-5-3 n'a pas à être motivée et peut être exprimée par tout moyen. Dans le cas de contenus mis à la disposition du public en ligne, cette opposition peut notamment être exprimée au moyen de procédés lisibles par machine, y compris des métadonnées, et par le recours à des conditions générales d'utilisation d'un site internet ou d'un service.

Annexe 3.

Echanges à propos de LAION (<https://laion.ai/>) marquant l'intention délibérée de ne pas rester la loi applicable

LAION est une association internationale dont l'objectif est de construire et de rendre accessible à tous des modèles et sets de données pour le machine learning. Leur dernier set d'image-texte en date est Laion 5B, composé de plus de 5 milliards de paires texte image. Ce set de données est largement utilisé par tous les acteurs du marché.

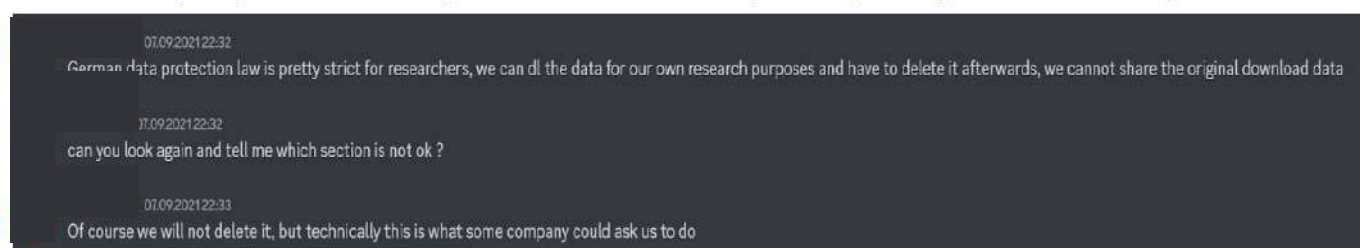
Il a l'avantage d'être public et les discussions qui ont amené à sa constitution sont facilement accessibles. Il faut noter qu'il est intimement lié à certains acteurs, et en particulier à Stability AI, dont le PDG intervient sur les forums de discussion.

Les extraits de conversations ci-dessous sont des discussions tenues par des membres importants et fondateurs. Elles permettent de mettre en lumière les arguments que nous avons développés dans le corps du document et nous semblent représentatives de l'état d'esprit dans lequel les pratiques se construisent.

1- Conservation illégale de données

"Bien sûr on ne les effacera pas."

La conservation des bases de données constituées pour la recherche sont soumises à des règles strictes qui n'ont pas été respectées. Ce faisant, ces bases de données constituent une violation des droits de propriété intellectuelle. (Les échanges ci dessous, anonymisés, impliquent des fondateurs de LAION)



07.09.2021 22:32
German data protection law is pretty strict for researchers, we can dl the data for our own research purposes and have to delete it afterwards, we cannot share the original download data

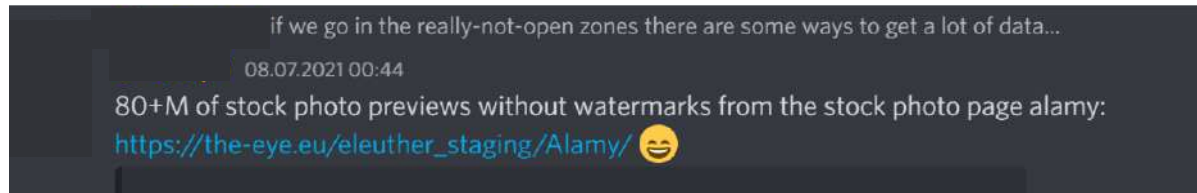
07.09.2021 22:32
can you look again and tell me which section is not ok ?

07.09.2021 22:33
Of course we will not delete it, but technically this is what some company could ask us to do

2- Contournement volontaire des protections mises en places

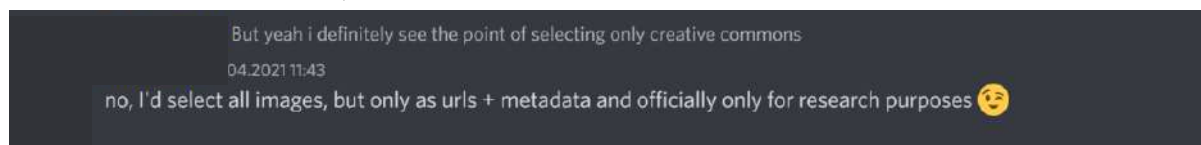
“Si nous allons dans les zones non ouvertes, nous pourrions récupérer beaucoup de données”

Pour protéger les images beaucoup de créateurs ou d'ayant droit protègent les données qu'ils divulguent en y apposant un filigrane (watermark en anglais). On voit ici un contournement de cette protection sur plus de 80 millions d'images sans filigranes



3- Contournement des règles d'exceptions

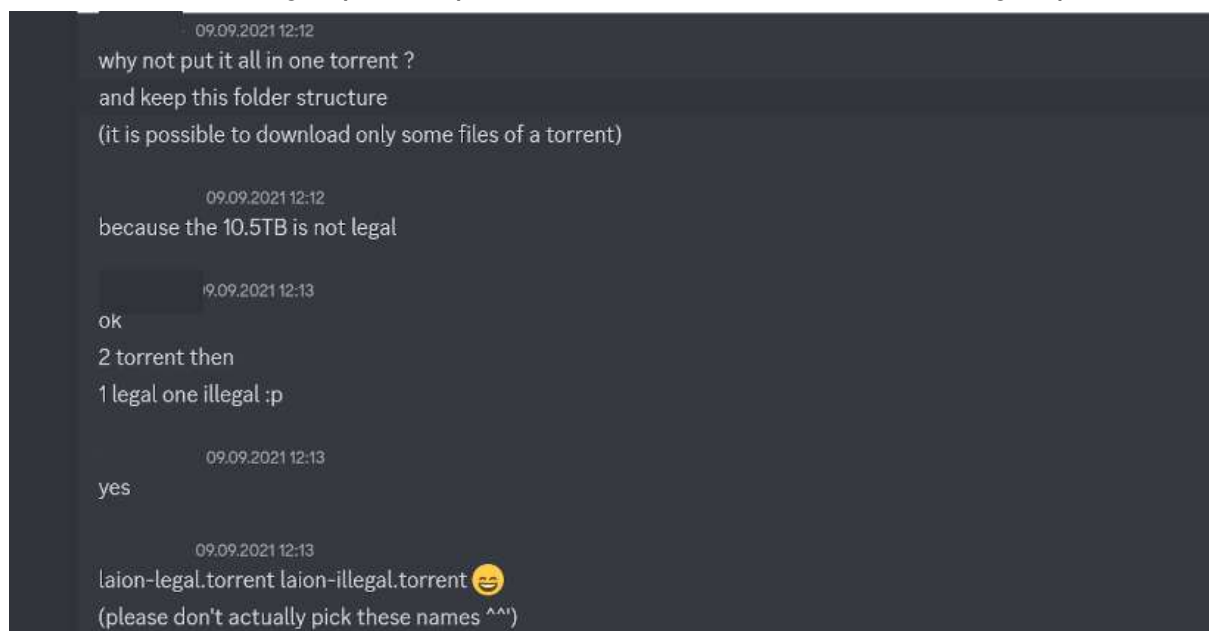
*“Je préférerais prendre toutes les images... et officiellement uniquement pour de la recherche *clin d'oeil* ;)”*



4- Contournement des règles de stockage

“Faisons 2 torrent, 1 légal et 1 illégal”

Une des techniques pour stocker et télécharger les données a été d'utiliser des "torrent". L'un d'entre eux contenait les données "illégales". Ces torrent ont été accédés et téléchargés par des personnes et structures en dehors du groupe LAION.



5- Désintérêt pour les lois applicables

"Qui pourrait le prouver ?"

Certains membres du projet sont tout à fait conscients de l'impossibilité d'utiliser ces modèles pour des usages commerciaux à cause des données non libres de droits contenues à l'intérieur. La réponse d'un des fondateurs est claire : "qui pourrait le prouver ?"

doesn't really work cuz you still can't use or release your model trained on non-cc samples for commercial purpose. (Bearbert0)
28.02.2021 00:41
Who could prove that to you

"Trouvons un trou dans les lois de propriété intellectuelle d'un pays et installons y notre archive 😊"

24.02.2021 11:00
The law says you are allowed to *submit* it to an archive for storing. It doesn't say it has to be a *german* archive 😊
It could be submitted to an archiv in another country that has more open sharing laws for scientific purposes 😊
Let's find a hole in the copyright laws of an accessible country and open there a small non profit archive 😊
I'm optimistic if it were legal, we could get some dontations for that 😊

"Nous devons simplement télécharger toutes les données sur internet et entraîner des modèles avant qu'ils ne puissent réguler.

Tout est dans la vitesse du progrès. Si tout arrive dans les 3 prochaines années, il sera trop tard."

06.12.2021 12:14
I am pretty sure that it will take several years for politics and law to catch up with the progress

I'm sure they can find ways to say it's a copy if it's making more money than their products 😊
06.12.2021 12:14
lol

06.12.2021 12:15
We simply need to download all data there is on the Internet And trained some awesome models with it before they start regulating it 😊

I am pretty sure that it will take several years for politics and law to catch up with the progress
06.12.2021 12:15
yeah I agree, that's why I think it's all about the speed of progress. If everything happens in the next 3 years, it'll be too late for politics

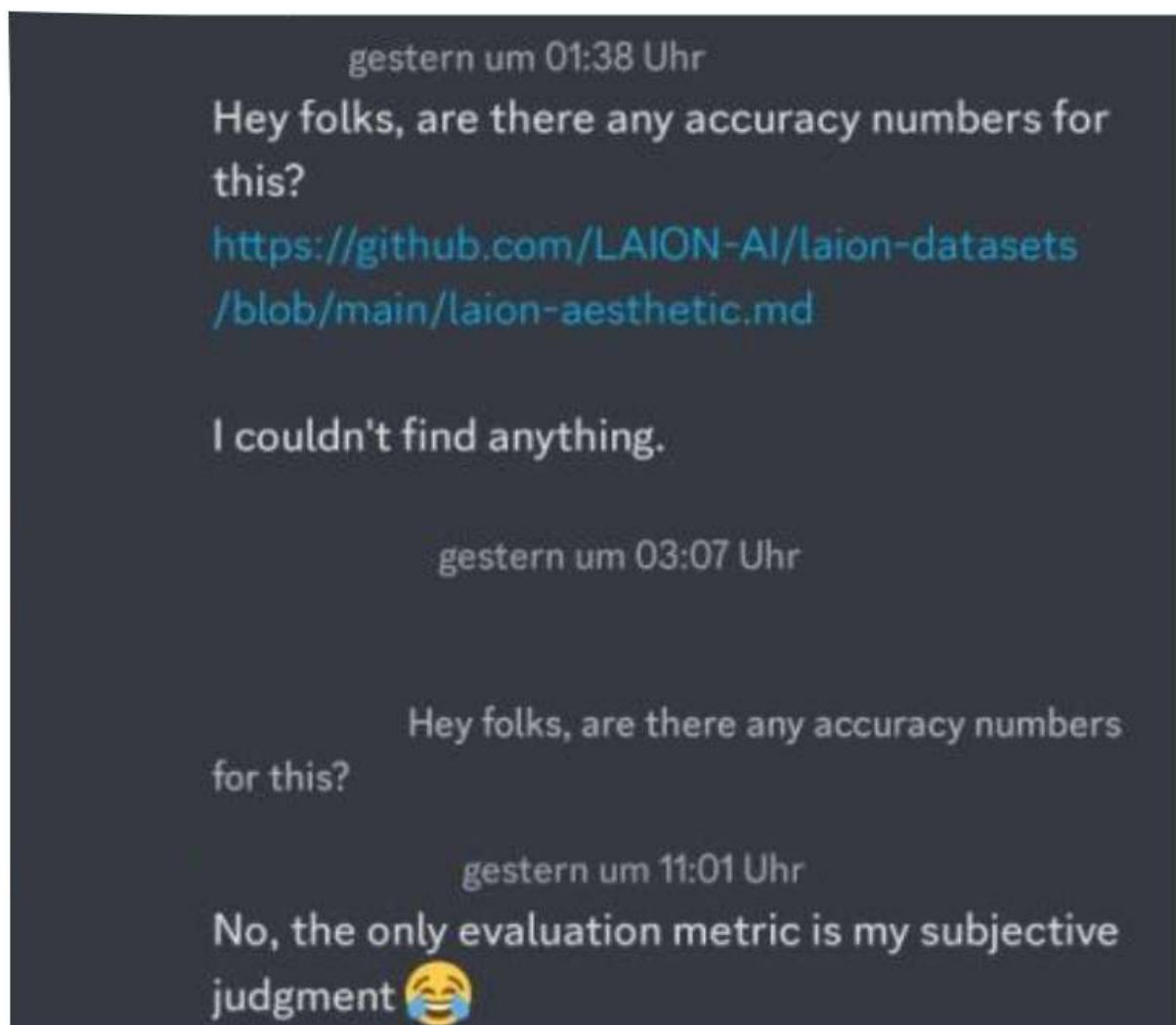
6- Détermination arbitraire de la notion de “beau”

“La seule mesure d’évaluation c’est mon jugement subjectif 😊”

Pour que les données soient les plus exploitables possibles, il faut qu’elles soient de bonne qualité et en particulier, qu’on puisse déterminer leur “esthétisme”. Laion a créé des sous parties de son set de départ appelées “Laion Esthetics” Elles sont issues d’une surcouche d’entraînement, créée pour déterminer ce qui est beau ou pas. La méthode d’évaluation a été le jugement subjectif d’un nombre très réduit de personnes impliquées dans le projet.

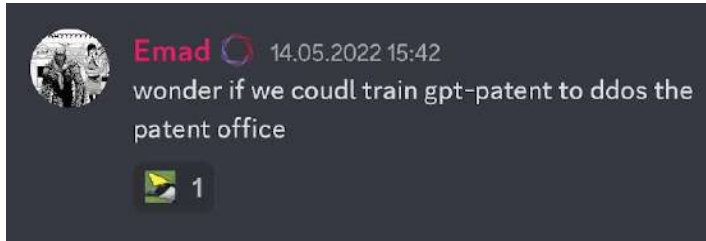
7- Impact de la base de donnée Laion 5B

“Tout le monde utilise Laion : Google, Meta...” (Emad Mostaque PDG de Stable Diffusion)



8- Attitude générale vis à vis de la notion de propriété intellectuelle

"Je me demande si nous pourrions entraîner une IA-dépôt de demande de copyright pour faire une attaque de déni de service contre le bureau du copyright. Ou générer tellement d'art pour que rien ne soit copyrightable. " Emad Mostaque PDG de Stable Diffusion.



C'est une volonté de ne pas tenir compte de la PI. L'objectif est depuis le début d'éviter la régulation

Annexe 4.

Reproduction de textes protégés sur Chat-GPT

John Lennon, *Imagine*

Get started

Enter an instruction or select a preset, and watch the API respond with a **completion** that attempts to match the context or pattern you provided.

You can control which **model** completes your request by changing the model.

KEEP IN MIND

- Use good judgment when sharing outputs, and attribute them to your name or company. [Learn more.](#)
- Requests submitted to our models may be used to train and improve future models. [Learn more.](#)
- Our default models' training data cuts off in 2021, so they may not have knowledge of current events.

Playground

Load a preset...

Save

These are the lyrics of Imagine by john lennon

Imagine there's no heaven
It's easy if you try
No hell below us
Above us, only sky

Imagine all the people
Living for today...

Aha-ah...

Imagine there's no countries
It isn't hard to do
Nothing to kill or die for
And no religion, too

Imagine all the people
Living life in peace...

You, you may say
I'm a dreamer, but I'm not the only one
I hope someday you'll join us
And the world will be as one

Imagine no possessions
I wonder if you can
No need for greed or hunger
A brotherhood of man

Imagine all the people
Sharing all the world...

You, you may say
I'm a dreamer, but I'm not the only one
I hope someday you'll join us

Looking for ChatGPT? [Try it now](#)

Submit

205

W. H. Auden, *Funeral Blues*.



dis moi la suite de ce poeme : Funeral Blues de W H Auden :
Stop all the clocks, cut off the telephone,



Prevent the dog from barking with a juicy bone,
Silence the pianos and with muffled drum
Bring out the coffin, let the mourners come.

Let aeroplanes circle moaning overhead
Scribbling on the sky the message He Is Dead.
Put crepe bows round the white necks of the public doves,
Let the traffic policemen wear black cotton gloves.

He was my North, my South, my East and West,
My working week and my Sunday rest,
My noon, my midnight, my talk, my song;
I thought that love would last forever: I was wrong.

The stars are not wanted now; put out every one,
Pack up the moon and dismantle the sun,
Pour away the ocean and sweep up the woods;
For nothing now can ever come to any good.